

# 基于图理论聚类和二值纹理分析技术的 彩色文本图像二值化方法

李向丰 汪斌 刘峰 胡福乔

(上海交通大学图象处理与模式识别研究所, 上海 200030)

**摘要** 为了有效地对彩色文本图像进行分割, 提出了一种复杂背景下彩色图像中文本-背景分离的新方法。该方法首先应用颜色空间降维以及基于图理论的颜色聚类对彩色文本图像进行聚类, 并对应于聚类结果获得一系列二值图像, 这些二值图像以及它们之间的组合就构成了二值化的待选结果; 然后对与游程直方图以及空间-尺寸分布相关的两类纹理特征进行分析, 并结合线性判别分析分类器来从待选的二值图像中选取出具有最佳文本-背景分离效果的二值图像。实验结果显示, 该方法的二值化效果比现有方法有显著提高, 因而能更有效地对具有复杂背景的彩色文本图像进行分割。

**关键词** 彩色文本图像 二值化 二值纹理分析 游程分析 空间-尺寸分布特征 线性判别分析

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2004)03-0290-07

## The Binarization for Color Text Images Based on Graph-theoretical Clustering and Binary Texture Analysis

LI Xiang-feng, WANG Bin, LIU Feng, HU Fu-qiao

(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030)

**Abstract** Text is an important feature for computer vision, especially for information retrieval applications. In this paper, the authors have developed a novel algorithm for text-background separation, or binarization for color images of complicated backgrounds. In their algorithm, dimensionality reduction and graph-theoretical clustering are first performed. Corresponding to each cluster, a binary image can be obtained. Additional binary images are obtained through combination among these cluster-related binary images. Then, two kinds of features capable of effectively characterizing binary texture images, run-length histogram based and spatial-size distribution based features associated with each of these binary images are extracted out. Based on the analysis of these texture features, cooperating with an LDA classifier, the optimal binary image which gives the best text-background separation will be found out as the final binarization result. Experiments with images collected from Internet have been carried out, which show that their method can handle color text images with complex background effectively; comparison with existing techniques also presented a notable improvement brought by the proposed method.

**Keywords** color text image, binarization, binary texture analysis, run-length analysis, spatial-size distribution feature, linear discriminant analysis

### 1 引言

众所周知, 文本是计算机视觉中, 特别是信息检索应用中的一项重要特征。在与文本图像相关的应用中, 图像中的文本首先要被识别出来, 然后才能得

到它们对应的语义信息。由于目前的字符识别技术(OCR)大多只能处理处于简单图像背景情况下的文本<sup>[1]</sup>, 因而二值化过程是现有字符识别技术应用中一个不可或缺的预处理步骤。二值化处理的目的是将图像中的文本同周围背景分离开来, 以便得到 OCR 能够处理的文本图像。

现有的二值化方法大多属于阈值化方法,而阈值化方法可以分为全局阈值化方法和局部或自适应阈值化方法两大类。其中,全局阈值化方法期望通过单一阈值对图像进行二值化。在文献[2]中,Sahoo 等对一些有效的全局阈值化方法作了很好的综述,他们指出在常用的全局阈值化方法中,根据二值化以后图像的区域均一性以及边界的准确度为评价标准,基于判别分析的 Otsu 方法能获得良好的二值化结果。在文献[3]中,首先采用 Otsu 方法来得到一系列待选阈值,且每个阈值都对应一幅二值图像;然后提取各幅二值图像的纹理特征,用以指导选取对应最佳阈值的二值图像。文献[3]的方法适用于在灰度直方图上字符部分和背景部分明显可分,并且字符部分处于直方图的低灰度段的图像。

大家知道,在复杂图像背景情况下,全局阈值化方法往往不能取得满意的结果,而与全局阈值化方法不同,局部或自适应阈值化方法则是根据从局部获得的相关信息来动态地改变阈值。在文献[4]中,Wellner 等为 DigitalDesk 设计了一种自适应阈值化方法,该方法是通过估计处于移动窗口内的背景灰度来计算图像中各点对应的阈值。文献[5]针对 11 种自适应阈值化方法的性能作了分析后指出,在图像呈低对比度以及在背景和噪声变化的情况下,局部阈值化方法能够得到满意的二值化效果,但文献[5]中也指出,对于具有复杂背景的情况,仍然没有一种方法能在单一参数的情况下取得满意结果。

由于上述现有的二值化方法大多只适用于灰度图像,因此本文提出了一种彩色文本图像二值化的新方法,其不仅能有效地克服现有二值化方法应用于具有复杂背景的彩色图像时所表现的不足,而且不同于一般的阈值化方法,因为该方法有效地结合了颜色聚类、二值纹理特征分析和线性判别分析技术,并且通过游程直方图和空间-尺寸分布(SSD)综合分析来有效描述二值文本图像的纹理特征;同时,采用在不同聚类对应的二值图像之间进行合并的策略。这些技术使得该方法能有效处理具有复杂背景的彩色文本图像。

## 2 基于纹理特征的彩色文本图像二值化

如图 1 所示,本文提出的彩色文本图像二值化方法主要由颜色空间降维、图理论颜色聚类、纹理特征提取以及最佳二值图像的选取 4 个部分组成。

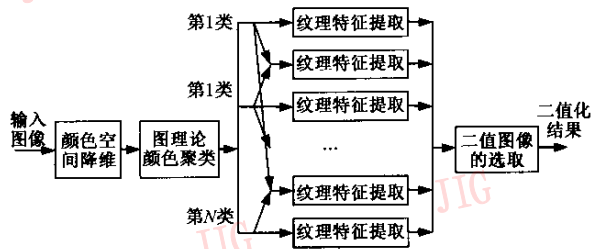


图 1 本文提出的二值化方法的框图

### 2.1 颜色空间降维

通常一幅彩色图像的每个像素用 24bit 来表示,这就意味着每个像素有超过 16M 的可能取值。如果直接在原始的颜色空间进行处理,则势必会降低算法的有效性。Jones 等经过对大量彩色图像的分析,发现在一幅通常的彩色图像中有 77% 的可能颜色不会出现<sup>[6]</sup>,于是,他们提出了一种基于直方图的颜色统计模型,同时提出将彩色图像的 R、G、B 3 个通道中的每个通道由原来的 8bit 表示降低为 5bit 表示,其实验结果显示,这种颜色空间降维使他们开发的皮肤检测系统获得了最好检测结果。

在文献[6]的基础上,通过试验发现,对一般彩色图像,RGB 各通道只以 8bit 中的高 4bit 来代替,图像视觉特性仍然没有显著的变化。由于此操作将图像颜色空间由原来的  $256 \times 256 \times 256$  变换到  $16 \times 16 \times 16$ ,因而为后面操作的有效执行奠定了基础,并且由于该颜色空间降维技术能很方便地实现,即只需将原始图像数据的每个字节右移 4bit 即可,这就大大降低了对计算的要求,从而提高了算法的效率。

### 2.2 基于图理论聚类的颜色聚类

经过颜色空间降维后的图像还不能直接进行文本提取,其颜色空间仍有  $16 \times 16 \times 16$  种颜色,因此需要进行进一步的信息聚合。本文使用基于图理论的颜色聚类来实现这个过程。

$G = (S, R)$  代表一个图,其中  $S$  是图  $G$  的顶点集合,  $R \subseteq S \times S$  是图  $G$  的边集合。若用  $(X, Y) \in R$  来定义顶点  $X$  和  $Y$  的相邻关系,则顶点  $X$  的邻域定义为

$$N(X) = \{Y | Y \in S, (X, Y) \in R\} \quad (1)$$

图  $G$  如果满足

$$(1) |G| > k;$$

(2)  $G$  的子图  $G - (W, R)$  是连通的,即  $G$  减去子图  $(W, R)$  所剩余的顶点和边构成的子图是连通的,其中,  $W$  是顶点子集,即  $W \subseteq S$ , 且  $|W| < k$ , 则图  $G$  为  $k$ -连通图。

由于图理论聚类的目标就是要将连通图  $G$  分解成一组单连通 ( $k=1$ ) 子图,  $G_i, i \in I, G_i$  满足

$$\bigcup_{i \in I} G_i = G \text{ 和 } G_i \cap G_j = \emptyset, \forall i, j \in I, i \neq j$$

因此聚类过程也是一个图的分解过程。

基于图理论的聚类算法需要如下两个参数: 直方图的长度以及用来搜索局部极大值分量的邻域宽度, 它们分别对应图  $G$  中的顶点个数和邻域范围。为了将图理论聚类算法应用于彩色图像, 首先需要计算彩色图像的 RGB-三维直方图, 然后在三维颜色空间中实现聚类过程。在本文提出的算法中, 控制聚类过程的两个参数分别是: ① RGB 每个通道的直方图长度都是 16; ② 搜索局部极大值分量的邻域宽度为 3。其中, 颜色空间中两个颜色分量之间的距离定义为它们之间的 City-block 距离  $d$  为

$$d(C_1(R_1, G_1, B_1), C_2(R_2, G_2, B_2))$$

$$= \max\{|R_1 - R_2|, |G_1 - G_2|, |B_1 - B_2|\} \quad (2)$$

基于图理论的聚类算法如下:

(1) 计算特征直方图, 使直方图的每一分量构成图  $G$  的一个顶点, 其邻域为特征空间中给定距离内所有点的集合, 即

$$N(X) = \{Y | Y \in R^n, d(X, Y) \leq D\}$$

其中,  $n$  为特征空间维数,  $D$  为邻域半径。计算结束后, 直方图每一分量拥有一个列表, 该列表包含了具有该特征值的所有像素。

(2) 顺序扫描直方图, 同时对直方图的每一分量, 在其给定邻域内, 找出对应于最大计数的直方图分量, 并在当前分量中, 保存一个指向该最大分量的链接, 即

$$N_{\text{new}}(X) = \{Y_m | H(Y_m) = \operatorname{argmax}\{H(Y) | Y \in N(X)\}\}$$

(3) 所有直方图分量扫描完成后, 原图  $G$  在新生成的邻接关系下就被分解成一组单连通子图, 而每一子图就对应聚类结果中的一类; 或者说, 这些链接就构成了一个森林结构, 其中每棵树的根节点对应直方图的一个局部极大值分量, 而对应每棵树中, 其包含的直方图分量的所有像素就形成了一个聚类。

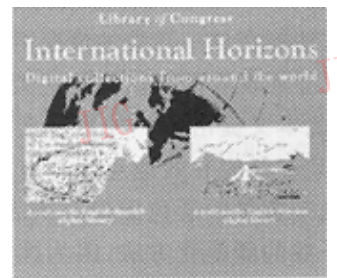
由于基于图理论的聚类算法不是直接作用于图像的各个像素, 而仅在彩色图像的三维直方图上执行, 因此其显著特点是聚类速度对图像大小的变化不敏感, 这就克服了基于像素聚类方法的缺陷, 同时, 颜色空间降维也降低了聚类计算要求。该聚类方法已经成功应用于文献[7]提出的彩色图像的文本定位方法中。图 2 显示了对彩色图像进行颜色空间降维和图理论聚类的结果。图 2(a) 是原图(原图为彩色, 现排成黑白), 其包含有 48 832 种不同颜色; 图 2(b) 为经过颜色空间降维后的图像, 仍然包含 778 种不同颜色; 图 2(c) 是经过聚类后, 聚类数为 3 的结果图像(分别呈现白色、灰色、偏黑色)。这样就为后面的处理提供了有利条件。



(a) 原图像



(b) 颜色空间降维结果



(c) 图理论颜色聚类结果

图 2 颜色空间降维和图理论颜色聚类

### 2.3 二值图像的生成

对应于颜色聚类后得到的原图像的每个聚类, 都能构造出一幅二值图像。原图像中属于该聚类的像素在对应的二值图像中取值 1, 即二值图像的前景部分; 而不属于该聚类的像素则在二值图像中取值为 0, 即二值图像的背景部分。将这些与聚类直接对应的二值图像进行合并后, 就可以得到附加的二值图像。在对二值图像进行合并之前, 需要对每一幅与聚类对应的二值图像进行连通分量分析, 以去除

二值图像中的一些主要干扰分量。

这样, 对应于不同的二值图像合并准则以及不同的输入图像, 就可以得到十几幅到几十幅二值图像。至此, 需要解决的主要问题就变成如何从这些待选二值图像中挑选出最佳的二值图像, 也就是具有最佳文本-背景分离效果的二值图像。为了有效评价各待选图像的二值化效果, 考察了如下两类纹理特征: 基于游程直方图的特征以及与空间-尺寸分布相关的特征。

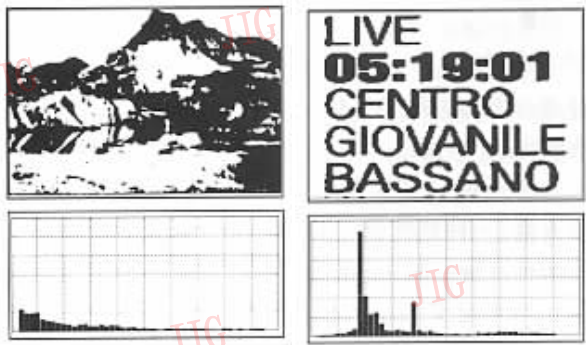
### 2.4 纹理特征提取

纹理是描述图像中某些区域内关联特性的术语,其主要反映区域内像素灰度的空域分布规律,而纹理特征则是图像处理和模式识别领域广泛应用的重要特征。纹理特征主要可以分成统计的、几何的、基于模型的以及基于信号处理的<sup>[8]</sup>等几类纹理特征。本文中提出的方法中,作者考察了两种纹理特征,即基于游程直方图的和基于空间-尺寸分布特性的两种纹理特征。这两种纹理特征都能有效地描述二值图像中包含的文本纹理。

#### 2.4.1 基于游程直方图的特征

游程是在一条扫描线上由具有相同灰度,并且在空间上连续的像素构成的像素集合,一个游程可以用灰度  $a$ ,起始位置  $b$ ,游程长度  $r$  和角度  $\theta$  来描述,即  $V(a, b, r, \theta)$ ,而描述纹理的特征则可以通过计算纹理区域中游程的长度和灰度的概率分布来获得<sup>[9]</sup>。文献<sup>[9]</sup>中给出了 5 种常用的游程统计特征,Tang 等又在他们提出的游程矩阵的基础上发展了几种新的纹理特征<sup>[10]</sup>。

现以  $M(a, r)$  表示所有考虑方向上长度为  $r$ ,灰度值为  $a$  的游程个数。由于本文的分析对象是二值图像,并且只考察前景部分水平方向的游程,因此,  $a \equiv 1$  以及  $r \in [1 \cdots L]$ ,其中  $L$  是允许的最大游程长度,而且  $M(a, r)$  可以进一步简化为  $M(r)$ 。令  $Z = [1 \cdots L]$ ,则  $M(Z)$  表示在游程容许长度范围内,由各游程频数组成一维数组,和图像灰度直方图相似,该数组称为对应二值图像的游程直方图。图 3 给出了普通二值图像和文本二值图像对应的游程直方图实例。可见,游程直方图包含能够描述对应二值图像特性的重要信息<sup>[3]</sup>。



(a) 一般二值图像及其游程直方图 (b) 文本二值图像及其游程直方图

图 3 不同类型二值图像的游程直方图对比

本文感兴趣的特征为最大频数特征和文本笔划宽度特征,其中最大频数特征定义为游程直方图中除去单位长度游程外的最大游程频数,因为单位长度游程往往由噪声构成。最大频数特征为

$$P_{\max} = \max_{r \in Z} M(r), r \neq 1 \quad (3)$$

对应于最大频数的游程长度为

$$l_{p_{\max}} = \arg \max_{r \in Z} M(r), r \neq 1 \quad (4)$$

由于文本二值图像的最大频数游程长度实际上反映了图像中主要文本的平均笔划宽度,因而在文献<sup>[3]</sup>中,称之为笔划宽度特征。

#### 2.4.2 空间-尺寸分布相关的特征

二值图像中,物体或者分量的空间特性以及它们的分布模式构成了二值纹理的主要内容,由于游程直方图与灰度直方图一样,是图像全局信息的一维表现,因而难以有效地反映图像中各物体真正的二维分布模式即图像中各物体的灰度、尺寸、颗粒及位置关系在内的综合分布模式。业已证明,数学形态学及其相关技术是进行二值图像分析的有效工具,其中颗粒及其尺寸分布分析是数学形态学应用于纹理分析的有效实例。这种颗粒尺寸分布也称为模式频谱<sup>[11]</sup>,它虽具有与信号频谱相似的特性,但由于模式频谱只能反映图像中各物体的尺寸分布,因而非常相异的纹理图像也可以具有相同的模式频谱。Ayala 等将图像中物体的空间分布分析与颗粒尺寸分布分析相结合,发展了一种新的空间-尺寸分布描述方法,从而有效地克服了模式频谱在描述二值图像纹理时表现出的不足<sup>[12]</sup>。

对应于给定图像  $G$ ,其在某一包含原点在内的紧致凸集合  $U$  上的  $(p, q)$  阶空间-尺寸分布函数(SSD)定义为颗粒面积  $A(\cdot)$  对位置变量  $h_1, \dots, h_q$  的多重积分<sup>[12]</sup>

$$F_{G,U}(\lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q) = \frac{1}{A(G)^{(q+1)}} \int_{\mu_q} \dots \int_{\mu_1} A(G + h_1 \cap \dots \cap G + h_q) A(\tilde{\Psi}(G) + h_1 \cap \dots \cap \tilde{\Psi}(G) + h_q) dh_1 \dots dh_q \quad (5)$$

其中,  $\lambda$  表示元素尺寸,  $\mu$  表示支撑集大小,以及

$$\tilde{\Psi}(G) = \Psi_{\lambda_p}^{(\rho)}(\dots(\Psi_{\lambda_1}^{(1)}(G))\dots) \quad (6)$$

表示不同颗粒的组合,并且

$$\Psi_{\lambda}(G) = G \circ \lambda B(0, 1) = G \circ B(0, \lambda) \quad (7)$$

表示以尺寸为  $\lambda$  的结构元素  $B$  对图像  $G$  做形态学开运算,而对应于  $F_{G,U}$  的联合密度函数为

$$f_{G,U}(\lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q) = \frac{\partial^{p+q} F_{G,U}(\lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q)}{\partial \lambda_1 \dots \partial \lambda_p \partial \mu_1 \dots \partial \mu_q} \quad (8)$$

对于离散集和离散的二值图像,上述联合密度函数可以表示成

$$f_{G,U}(\lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q) = \sum_{(u,v) \in T} (-1)^{\text{sgn}(u,v)} F_{G,U}(u,v) \quad (9)$$

其中

$$T = \{(u,v): u_i = \lambda_i \text{ 或 } u_i = \lambda_i - 1; v_j = \mu_j \text{ 或 } v_j = \mu_j - 1\} \quad (10)$$

以及

$$\text{sgn}(u,v) = \#\{u_i: u_i = \lambda_i - 1\} + \#\{v_j: v_j = \mu_j - 1\} \quad (11)$$

$\#$  为计数符号,即  $\#\{u_i: u_i = \lambda_i - 1\}$  表示满足条件  $u_i = \lambda_i - 1$  的  $u_i$  的个数。

由于本文考虑了二值图像前景分量的(1,1)阶和(2,1)阶 SSD 特征,且设其中元素尺寸  $\lambda$  取值与对应前景分量的笔划宽度特征值相同,支撑集大小  $\mu$  从 0 变化至 9,步长为 3,因此,对应于每幅二值图像,总共能够得到 8 个空间-尺寸分布特征值。

## 2.5 二值图像的选取

待选二值图像的合并与选取可通过以下几个步骤来实现:首先,根据游程直方图纹理特征,去除那些笔划宽度频数值较低的二值图像,而保留具有显著文本笔划特征的二值图像;然后,对初选后保留的

二值图像根据空间-尺寸分布纹理特征进行进一步验证;最后,将通过验证的二值图像进行合并来得到最终结果。

在对初选后保留的二值图像进行进一步验证的过程中,2.4 节介绍的对应每幅图像的 8 个空间-尺寸分布特征都被送入一个线性判别分析(LDA)分类器<sup>[13]</sup>进行分类判别。LDA 是一种利用给定特征来确定两类或者多类之间的分离边界的统计方法,LDA 的基本思想是寻找一种线性变换,使得在执行该变换后各个特征类别之间可分离性达到最大。本文应用的 LDA 分类器是通过 leave-one-out 进行训练,但通过 LDA 分类器验证的仍然可能是多幅二值图像,由于它们都包含有明显的文本特征,因而,最后应再将这此二值图像进行合并,才能得到本文方法的二值化结果。

图 4 给出了进行类间合并前后二值化结果的对比。从图 4 可以看出,类间合并规则的应用能明显改善二值化效果,特别是对于具有变化颜色的文本块,并且,现有方法的二值化结果中,得到的文本部分取值仍然不定,即文本块取值仍然可能为 1(前景部分取值)和 0(背景部分取值)。由于本文方法得到的文本取值都为前景部分(即取值 1),因而能减轻后续字符识别系统(OCR)的负担,即能提高系统的整体性能。



(a) 原彩色图像



(b) 合并前二值化结果



(c) 二值化后结果

图 4 合并前后二值化结果对比

## 3 实验结果

本文从 Internet 上收集了 500 多幅图像进行实验,所有这些图像中都包含有明显的文本块,但是图像背景的复杂程度、文本的字体、颜色和大小等则随图像不同而变化。其和通常文档图像处理中所考虑的对象不同,由于这些图像中大都只包含几个单词,因而应用一般的纹理分析技术难以有效对它们进行处理。为了考察本文提出方法的有效性,作者将它和两

种现有方法进行了比较,该两种方法分别是 Otsu 方法和文献[3]提出的基于游程直方图的阈值化方法。

为能进行多角度的对比,本文还设计了两种评测方案:一种是视觉上的评估,另一种是单词水平上的精确评价。其中,在视觉评估过程中,可接受的二值化结果定义为输入图像中至少有 80% 的文本被有效分离出来,并且结果中非文本背景部分的面积不大于文本面积;在单词水平的评价方案里,则首先对输入图像以及对应的二值化结果图像中可识别的单词个数进行统计;然后根据准确分离的单词的比

例,对不同方法进行对比。

图 5 给出了不同方法二值化结果的对比。图 5(a)是原彩色图像;图 5(b)是 Otsu 方法的二值化结果;图 5(c)是文献[3]方法的二值化结果;图 5(d)是本文提出方法的二值化结果。Otsu 方法对类似图 5(a)的具有复杂背景图像难以取得令人满意的二值化效果。与 Otsu 方法对比,文献[3]提出的方法二值化效果虽有所改善,但由于它仍然是基于阈值化的方法,因而仍然难以克服由复杂背景带来的影

响,并且在文献[3]提出的方法中,只能处理图像中单一笔划宽度的文本,而本文提出的方法则能够有效处理同一图像中不同笔划宽度的文本。如图 5(d)所示,和上述两种现有的方法相比较,本文方法的二值化效果有很大的改善。实验结果也显示,本文方法不仅能有效地将文本和图像背景分离开,而且倾向于获得更干净的二值化结果。表 1 中给出了实验统计数据。

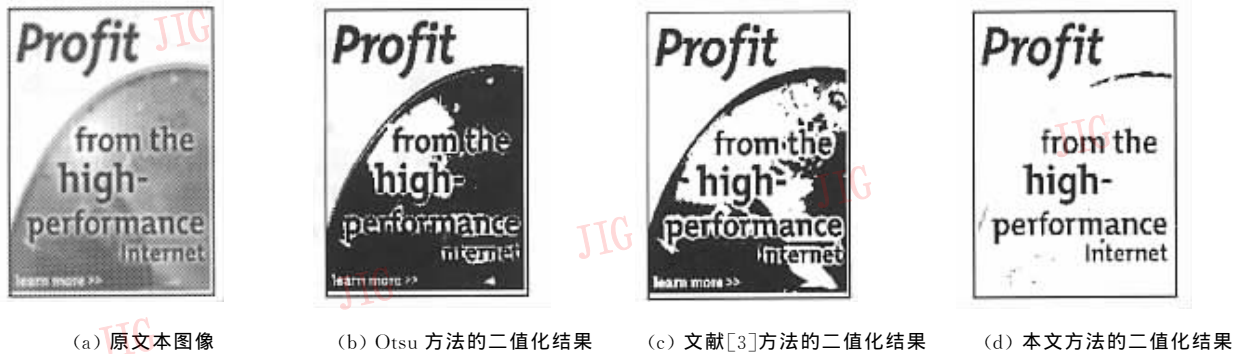


图 5 不同方法二值化结果对比

图 1 实验结果数据统计

	全部图像	Otsu 方法		文献[3]方法		本文提出的方法	
		正确数	正确率(%)	正确数	正确率(%)	正确数	正确率(%)
视觉评价	520 幅	353 幅	67.9	376 幅	72.3	446 幅	85.8
单词的正确分离率	3 519 个	2 397 个	68.1	2 441 个	69.4	2 850 个	81

## 4 结 论

二值化是计算机视觉,特别是 OCR 和文档或文本分析应用中一项重要的处理过程。由于大多现有的二值化方法都属于阈值化方法,因此在应用于具有复杂背景的图像时,它们往往不能获得满意的二值化结果。本文提出了一种新的结合颜色聚类和二值纹理分析技术的彩色文本图像二值化方法。实验结果显示,该方法能有效克服现有方法的不足,在图像具有复杂背景的情况下,仍能取得满意的结果。结合已经取得的图像文本定位的研究成果<sup>[7]</sup>,作为对现有工作的进一步深入和扩展,作者进一步的目标是构建一个基于图像文本的图像检索系统(image text based image retrieval system)。

## 参 考 文 献

1 Mori S, Suen C, Yamamoto K. Historical review of OCR research and development[J]. Proceedings of the IEEE, 1992,

80(7): 1029~1058.

- Sahoo P, Soltani S, Wong A. A survey of thresholding techniques [J]. Computer Vision, Graphics and Image Processing, 1988, 41(2): 233~260.
- Liu Y, Srihari S N. Document image binarization based on texture features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(5): 540~544.
- Wellner P D. Adaptive thresholding on the DigitalDesk [R]. Technical Report EPC-1993-110, Rank Xerox Research Center, Cambridge Laboratory, England, 1993.
- Trier O D, Taxt T. Evaluation of binarization methods for document images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(3): 312~315.
- Jones M, Rehg J. Statistical color models with application to skin detection [J]. International Journal of Computer Vision, 2002, 46(1): 81~96.
- 汪斌, 胡福乔. 基于图理论聚类的彩色图象文本提取[J]. 微电子学与计算机, 2003, 20(8): 89~93.
- Mihran T, Jain A K. Handbook of Pattern Recognition and Compute Vision [M]. River Edge, NJ, USA: World Scientific Publishing, 1993: 235~276.
- Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and

Machine Vision[M]. Pacific Grove, CA, USA: Brooks and Cole Publishing, 1999.

- 10 Tang X O. Texture information in run length matrices[J]. IEEE Transactions on Image Processing, 1998,7(11): 1602~1609.
- 11 Maragos P. Pattern spectrum and multiscale shape representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989,11(7):701~716.
- 12 Ayala G, Domingo J. Spatial size distributions: Applications to shape and texture analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001,23(12): 1430~1442.
- 13 Yang J, Yu H, Kunz W. An efficient LDA algorithm for face recognition[A]. In: Proceedings 6th International Conference on Control, Automation, Robotics and Vision[C], Singapore, Dec 2000.



李向丰 1979年生,2001年获西安交通大学学士学位,现为上海交通大学图像处理与模式识别专业硕士研究生。主要研究方向为图像文本分析和处理、图像检索。



汪斌 1979年生,2003年获上海交通大学图像处理与模式识别专业硕士学位。主要研究方向为图像文本分析和处理、医学图像分析。



刘峰 1975年生,1997年获得武汉大学学士学位,现为上海交通大学图像处理与模式识别专业硕士研究生。主要研究方向为无线通信、智能交通系统中交通数据的采集、传输及处理。

胡福乔 1957年生,上海交通大学图像处理与模式识别研究所副教授。主要研究方向为智能交通系统、模式识别及图像处理。