

# 基于 DRCoALTP 的印刷体文档图像多文种识别方法

吴正健<sup>1,2</sup>, 吾尔尼沙·买买提<sup>1</sup>, 杨耀威<sup>1</sup>, 阿力木江·艾沙<sup>1</sup>, 库尔班·吾布力<sup>1\*</sup>

(1.新疆大学计算机科学与技术学院, 新疆 乌鲁木齐 830017; 2.武汉大学计算机学院, 湖北 武汉 430072)

**摘要:**针对视觉结构类似导致的文种相似性问题,基于局部三值模式的相邻共生矩阵(co-occurrence of adjacent local ternary patterns, CoALTP)提出一种具有判别性和鲁棒性的局部三值模式的相邻共生矩阵(discriminant and robust co-occurrence of adjacent local ternary patterns, DRCoALTP)方法,用于获取图像纹理。计算文档图像的相邻稀疏局部三值模式(adjacent sparse local ternary patterns, ASLTP),将采样点数量设定为8,以便获得详细的局部纹理,设计出一种基于自适应中值滤波思想的半自适应阈值方法,用于提取灰度图像中心像素周边对角邻域像素的编码值。ASLTP在邻域像素位置存放稀疏局部三值模式(local ternary patterns, LTP)的值,提取灰度共生矩阵(gray-level co-occurrence matrix, GLCM),从4个方向统计使用ASLTP后灰度图像像素之间的频率关系。该算法在阿拉伯文、俄文、简体中文、哈萨克文、藏文、蒙古文、土耳其文、维吾尔文、英文、吉尔吉斯斯坦文和塔吉克斯坦文11个文种的自建印刷体文档图像数据集中验证。试验结果表明,相较于基线和先进的纹理方法,改进后的方法更具判别性,平均识别准确率为99.14%。为改善CoALTP方法可能产生低效分类特征的问题,提出半自适应阈值方法,有效提高识别率并抑制噪声。此外,针对算法产生的高维特征,采用基于均方差的特征选择方法,通过支持向量机(support vector machine, SVM)分类器特征选择后,识别速度提高284%,对11个文种的平均识别准确率达99.44%。

**关键词:**稀疏局部三值模式;灰度共生矩阵;文种识别;半自适应阈值;特征选择

**中图分类号:**TP391.4 **文献标志码:**A

**引用格式:**吴正健,吾尔尼沙·买买提,杨耀威,等.基于DRCoALTP的印刷体文档图像多文种识别方法[J].山东大学学报(工学版),2025,55(1):51-57.

WU Zhengjian, MAMAT Hornisa, YANG Yaowei, et al. A script identification method for printed document images based on DRCoALTP [J]. Journal of Shandong University (Engineering Science), 2025, 55(1):51-57.

## A script identification method for printed document images based on DRCoALTP

WU Zhengjian<sup>1,2</sup>, Hornisa Mamat<sup>1</sup>, YANG Yaowei<sup>1</sup>, Alimjan Aysa<sup>1</sup>, Kurban Ubul<sup>1\*</sup>

(1. School of Computer Science and Technology, Xinjiang University, Urumqi 830017, Xinjiang, China; 2. School of Computer Science, Wuhan University, Wuhan 430072, Hubei, China)

**Abstract:** Aiming at the script similarity due to small differences in visual structures, a discriminant and robust co-occurrence of adjacent local ternary patterns (DRCoALTP) was proposed to overcome the limitation of the co-occurrence of adjacent local ternary patterns (CoALTP) and obtain the image texture. The adjacent sparse local ternary patterns (ASLTP) of the document image were computed with eight sampling points in order to obtain the detailed texture, and a semi-adaptive threshold method based on the idea of adaptive median filtering was used to extract the coded values of the pixels in the diagonal neighborhood around the central pixel of the grayscale image. The ASLTP stored the neighborhood pixel positions with the values of sparse local ternary patterns (LTP) and extracted the gray-level co-occurrence matrix (GLCM) to count the frequency relationship between pixels of the grayscale image used ASLTP in four directions. The algorithm was verified in the self-built data sets of printed document images from eleven languages, including Arabic, Russian, Chinese, Kazakh, Tibetan, Mongolian, Turkish, Uygur, English, Kyrgyz and Tajik. Compared with the baseline and advanced texture methods, experimental results showed that the improved method was more discriminant, with an average recognition accuracy of 99.14%. To improve the inefficient classification features extracted by the

**收稿日期:**2023-11-22

**基金项目:**国家自然科学基金资助项目(62266044, 61862061, 61363064);2022年新疆维吾尔自治区重点研发专项-厅厅联动资助项目(2022B03035)

**第一作者简介:**吴正健(1995—),男,安徽芜湖人,博士研究生,主要研究方向为模式识别及计算机视觉。E-mail: wzj199538@gmail.com

**\*通信作者简介:**库尔班·吾布力(1974—),男,新疆巴楚人,教授,博士生导师,博士,主要研究方向为数字图像处理与模式识别。

E-mail: kurbanu@xju.edu.cn

CoALTP method, the proposed semi-adaptive threshold method effectively improved the accuracy rate and suppressed noise. Besides, a feature selection method based on mean square error was adopted for the high-dimensional features, improving the recognition speed by 284% through the support vector machine (SVM) classifier, and the average recognition accuracy of eleven languages reached 99.44%.

**Keywords:** sparse local ternary patterns; GLCM; script identification; semi-adaptive threshold; feature selection

## 0 引言

多文种识别是识别文档图像上文种类型的过程,通常在光学字符识别(optical character recognition, OCR)之前处理。对于书写而言,文种表示形式为图形,由语言和一系列的字符表达<sup>[1]</sup>。纹理是识别图像中感兴趣对象或区域的重要特征之一,文献[2]提出局部二值模式(local binary patterns, LBP),用于纹理分类,在更少的计算中表现出色。近几十年,局部纹理特征被广泛研究<sup>[3-9]</sup>。在速度和性能方面,LBP及其变种方法成功应用于很多领域,如纹理分类<sup>[10-12]</sup>、人脸识别<sup>[13-15]</sup>、文种识别<sup>[16-18]</sup>。文献[3]提出基于非统一模式改进的LBP,通过组合非统一模式提高LBP性能;文献[4]扩展表示金字塔变换域中的LBP描述符,与现有的多分辨率LBP描述符相比,该特征具有良好的性能和较低的计算成本;文献[5]提出一种局部二值模式方差(local binary patterns variance, LBPV)描述符,与现有的局部选择不变LBP比较,拥有更好的分类性能。

文档图像是由扫描仪或照相机等设备将纸质文档转换而成,转换过程中扫描设备、光照、纸质材料都可能引入图像噪声,要求选择的特征需要具有良好的鲁棒性。文献[8]通过建立一个局部三值模式的相邻共生矩阵(co-occurrence of adjacent local ternary patterns, CoALTP)和鲸鱼优化算法结合的分类框架,提高疾病严重性的预测水平;文献[9]提出一种基于梯度局部二值模式(gradient local binary patterns, GLBP)的表面缺陷检测方法,利用图像子块降低LBP数据矩阵维数,通过在附近8个方向上采用加权二进制输出值指示局部灰度变化,抑制光和噪声对检测结果的影响;文献[19]首次提出灰度共生矩阵(gray-level co-occurrence matrix, GLCM)方法,一幅图像的灰度共生矩阵能反映出图像灰度关于方向、相邻间隔、变化幅度的综合信息;文献[20]对4个文种的文档图像提取GLCM特征并在多个分类器中试验,验证GLCM在书写文种识别中的有效性。由于局部三值模式(local ternary patterns, LTP)和GLCM纹理特征方法具有不同的特点,本研究将二者

结合起来进行纹理分析。

LBP、LTP和GLCM纹理描述方法在文种识别领域也有研究。文献[8]将LTP的共生矩阵纹理描述算法应用到医学图像,对灰度化图像的LTP提取特征,实现对心血管疾病严重程度的预测,该算法集中LTP和GLCM,对光照和噪声敏感性降低;文献[18]提出一种有效基于块级别的LBP算法,对印度多语种手写文档图像进行识别;文献[21]提出一种结合LTP和GLCM的CoALTP算法,检索纹理图像和人脸图像;文献[22]提出一种结合梯度方向直方图的GLCM方法,识别场景中的多语言字符。

印刷体文档图像的特点是文字区域和背景区分度高,没有复杂的背景且文字结构规范。在灰度图像上,采用纹理特征进行多文种识别是一种合适的方法。相较于LTP,CoALTP描述了像素的空间关系,能在一定程度上反映文档图像的方向结构。因此,本研究提出一种基于CoALTP的改进方法识别多文种的印刷体文档图像。

## 1 方法

### 1.1 邻域稀疏局部三值模式的共生矩阵

LTP增强了LBP编码,在LBP描述符的基础上,通过估计相邻像素与中心像素之差确定阈值函数 $s(z)$ 。因此,整体LBP特征模式分布对噪声高度敏感,而LTP利用一个三值标号提高噪声容忍范围,阈值函数

$$s(z) = \begin{cases} 1, & z \geq t \\ 0, & t > z \geq -t, \quad t \in N, \\ -1, & -t > z \end{cases} \quad (1)$$

式中: $z$ 为像素; $t$ 为一个预定义的常量,属于某个非负整数 $N$ 。LTP特征函数

$$LTP_{q,w}(x,y) = \sum_{q=0}^{q-1} s(p(x,y) - p(x_e, y_e)) \cdot 3^q, \quad (2)$$

式中, $q$ 为周围邻域像素点个数, $w$ 为半径, $(x,y)$ 为中心像素, $(x_e, y_e)$ 为第 $e$ 个邻域像素, $p$ 为LTP阈值函数。当生成三值模式时,特征维度显著增加。为解决这一问题,将LTP描述符分为上模式和下模式,特征大小为 $2 \times 28$ 。

CoALTP方法先计算ASLTP,不同于典型的LTP,在ASLTP的计算中,不考虑8个相邻像素,仅

使用 4 个半径为 1 的对角像素,提取 4 个方向的共生矩阵作为最终特征集<sup>[22]</sup>。ASLTP 的优点是结合了 LTP 和 GLCM,使其特征拥有局部纹理、像素频率和相关性信息。相比垂直和水平的邻域像素,4 个对角的邻域像素提供了更多有关中心像素的信息。

结合文档图像的结构特点,本研究对灰度化后的文档图像计算距离为 1 的局部三值编码。在每个 3 像素×3 像素范围的相邻像素中,将 4 个对角邻域像素的 ASLTP 编码存放在 (3,3) 对角线像素位置而不是中心像素位置,保留更多邻域边缘信息,以便计算共生矩阵时获得更多的边缘频率。示例将距离设置为 1,通过给定像素的 4 个对角线邻域像素获得 ASLTP,如图 1 所示。

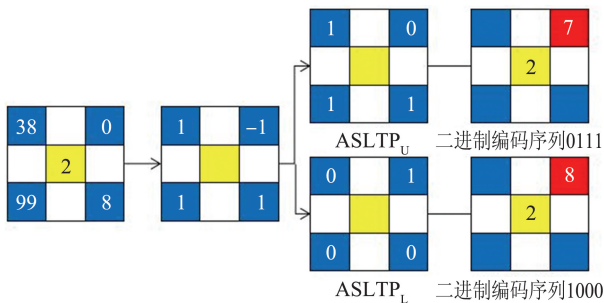


图 1 从文档图像计算给定像素的 ASLTP  
Fig.1 Calculate the ASLTP for a given pixel from the document image

在得到 ASLTP 编码后,通过距离 1 与 4 个不同位置方向进一步获得文档图像变换后的共生矩阵。GLCM 是一种考虑空间关系的统计方法<sup>[19]</sup>,通过计算特定方向的像素强度(灰度水平)  $i$  与像素频率  $j$  生成灰度水平的共生矩阵。统计一幅完全灰度图像(灰度范围  $[0, 255]$ ) 的 GLCM 是复杂的,一般将灰度图像的灰度水平  $i$  从 256 降低到 8 或 16。GLCM 计算示例如图 2 所示,GLCM 的灰度水平为 8,统计左边矩阵距离为 1、水平方向的结果,每个元素  $(i, j)$  由  $i$  与相邻(距离为 1)水平靠右方向的  $j$  的次数和得到。例如图 2 中右侧 GLCM(2,2) 位置的元素值为 1,是因为左侧灰度水平为 8 的输入图像在像素等于 2 和 2 的水平空间上只出现了一次。

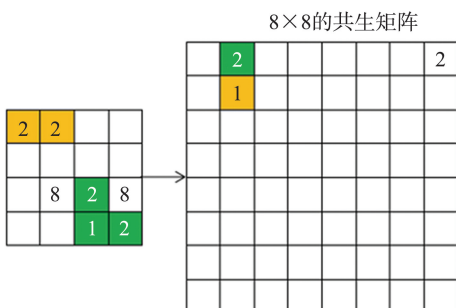
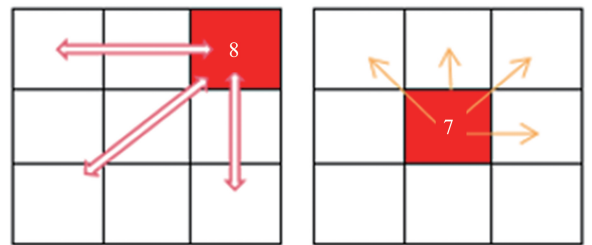


图 2 GLCM 计算示例  
Fig.2 GLCM calculation example

通过上述方法,将每个模式图(上 ASLTP 图和下 ASLTP 图)的灰度范围变为  $[0, 15]$ ,统计 4 个方向的 GLCM,最终特征和向量的长度为  $16 \times 16 \times 2 \times 4$ 。本研究方案的稀疏局部三值模式共生矩阵如图 3 所示,CoALTP 特征函数

$$\text{CoALTP}(\Delta x, \Delta y) = \sum_{u=1}^M \sum_{v=1}^N [\text{ASLTP}_U(u+\Delta x, v+\Delta y) \text{ASLTP}_L(u+\Delta x, v+\Delta y)], \quad (3)$$

式中:  $\text{ASLTP}_U$ 、 $\text{ASLTP}_L$  分别为图像的上、下局部三值函数映射;  $\Delta x$ 、 $\Delta y$  为原点坐标  $(u, v)$  在  $x$  轴和  $y$  轴方向上的偏移量,共有 4 个偏移方向,其中  $(0, 1)$  对应  $0^\circ$ ,  $(-1, 1)$  对应  $45^\circ$ ,  $(-1, 0)$  对应  $90^\circ$ ,  $(-1, -1)$  对应  $135^\circ$ 。最终特征向量是在 4 个方向上获得的 CoALTP 组合。



(a) 计算对角邻域 ASLTP (b) ASLTP 编码后统计 4 个方向的共生矩阵

图 3 计算 CoALTP 的示例  
Fig.3 An example of calculating CoALTP

## 1.2 改进的稀疏局部三值模式的共生矩阵

### 1.2.1 半适应 ASLTP 阈值方法

原始 ASLTP 方法的阈值是一个常数,对噪声比较敏感。同一图像中,不同的阈值可能产生不同的局部三值编码结果。此外,单个像素点容易受光照等自然噪声的影响,常量阈值仅对像素的部分单调变化具有抵抗性。因此,本研究提出一种半适应阈值函数  $s(z)$  提高特征的判别性。 $s(z)$  与诸多因素有关,如图像信噪比、灰度图像分布,本研究认为  $s(z)$  是关于图像邻域像素与中心像素灰度的函数。类似自适应中值滤波的思想,文档图像的每个邻域窗口在编码 ASLTP 时,以 ASLTP 同样的圆形邻域为局部窗口,通过中心像素  $C$  与对角邻域的灰度中值  $I_{med}$  搜索非噪声点,在邻域里若  $I_{med}$  在最大灰度  $I_{max}$  和最小灰度  $I_{min}$  之间,则当前邻域窗口的阈值  $T = I_{med}$ ,否则判断  $C$  是否为非噪声点。当前窗口搜索失败,则认为邻域外的窗口也存在噪声信号,停止搜索,此时阈值  $T = 7/255$ 。基于半适应阈值法的实现过程如图 4 所示,其中  $N_{max}$  和  $r$  分别为扩张窗口限制和扩张像素,  $\text{sort}(W)$  是对邻域像素灰度  $W$  排序的函数。

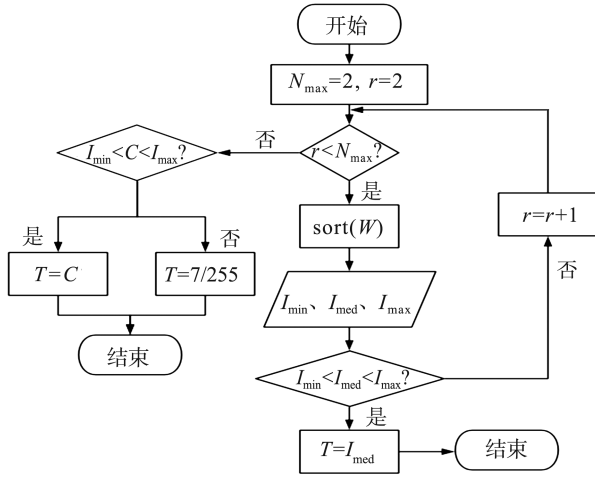


图4 ASLTP的半自适应阈值法流程图

Fig.4 Flow chart of semi-adaptive threshold method for ASLTP

$s(z)$ 是构建CoALTP的编码时,图像在第 $p$ 个邻域像素与中心像素比较时的空间响应编码阈值,决定编码时选择哪个三进制的标签值。相比继续通过扩大中心像素周围的分析窗口寻找局部中值的自适应中值滤波算法思想,半适应ASLTP阈值不仅减少了噪声干扰,也提升了特征提取速度。此外,在进一步提取共生矩阵时,不仅能获取图像丰富的频率信息,还能保持与常量阈值几乎近似的计算时间。

### 1.2.2 基于均方差的特征选择方法

CoALTP采用4个对角邻域计算稀疏的LTP,降低计算复杂度,但最终的特征数量由提取的共生矩阵向量决定。CoALTP特征数量

$$F_{\text{cas}} = 2dN_{\text{umLevels}}^2 \quad (4)$$

式中, $N_{\text{umLevels}}$ 为GLCM的灰度级别, $d$ 为提取GLCM的数量。当灰度级别为16时,提取4个方向的邻域局部三值模式共生矩阵特征数量为2048。由于本研究方法从稀疏LTP编码后的图像矩阵中提取共生矩阵,不可避免产生冗余的特征向量,尤其是稀疏纹理的图像。实际上,应用于多文种文档图像的CoALTP特征集中有许多低效的特征子集。此外,特征也受文档图像采集时的操作、纸质数据源的质量等因素影响。因此,特征选择或降维是必要的。文献[23]通过选择一组旋转不变的LBP和LTP增强分类器性能,基于此思路,本研究提出采用基于均方差(based on mean square error, BMS)的特征选择算法进一步筛选CoALTP特征子集,从2048个特征中筛选出499个特征子集。

#### 算法1 BMS的执行步骤

输入 特征集 $F$ 。

输出 特征子集 $S$ 。

(1)计算特征集 $F$ 中每列向量的均方差

$$S_i = \sqrt{|m_i - \bar{u}|^2},$$

式中: $m_i$ 为第 $i$ 列向量中的元素; $\bar{u}$ 为第 $i$ 列向量中元素的均值, $\bar{u} = \frac{1}{N} \sum_{i=1}^N m_i$ ,其中 $N$ 为特征集中列向量的数量。

(2)归一化处理。计算每列向量均方差占有所有向量均方差和的比例

$$\delta_i = \frac{S_i}{S_N},$$

式中, $S_N$ 为所有向量均方差的和, $S_N = \sum_{i=1}^N S_i$ 。针对 $\delta_i$ 进行降序排列。

(3)计算第1列 $\delta_1$ 到第 $i$ 列 $\delta_i$ 的累计和

$$K = \sum_{i=1}^N \delta_{1 \rightarrow i}.$$

(4)当 $K$ 在累计过程中超过阈值1,则保留达到阈值前的特征向量 $b_i$ ,组成新的特征子集

$$S = \{b_i, \dots, b_j\}, K > 1.$$

## 2 试验结果与分析

本研究将书籍、杂志等纸质资料通过扫描仪采集得到多文种的文档图像,建立包括阿拉伯文、俄文、简体中文、哈萨克文、藏文、蒙古文、土耳其文、维吾尔文、英文、吉尔吉斯斯坦文和塔吉克斯坦文共11个文种的文档图像数据库,其中每个文种包含2000张文档图像。本研究以1:1的比例随机从文档图像数据中选择训练集和测试集。试验1选择2048个CoALTP特征集进行试验,499个CoALTP特征子集进行训练和测试;试验2将原始CoALTP和本研究改进LTP编码方式后的CoALTP方法进行比较;试验3对提出的半自适应阈值函数的抗噪性进行测试。所有试验采用支持向量机(support vector machine, SVM)分类器,为避免过拟合,在训练时使用5折交叉验证并统计混淆矩阵。在每个试验中通过准确率评价试验结果。将所提方法与LBP及其统一模式(U\_LBP)、选择不变模式(RI\_LBP)、统一选择不变模式(RIU\_LBP)、LTP、GLCM和离散曲波变换方法<sup>[24]</sup>等先进纹理特征算法进行比较。

### 2.1 原方法与改进方法的对比试验

设置4组对比试验,试验结果如图5所示,其中

CoALTP 为原方法,CoALTP\_BMS 为 CoALTP 进行 BMS 特征选择的方法,DRCoALTP 为改进方法,DRCoALTP\_BMS 为 DRCoALTP 进行 BMS 特征选择的方法。由图 5 可以看出,改进的纹理描述方法 DRCoALTP 对印刷体文档图像数据集的 11 个文种具有比较好的区分效果,平均准确率达 99.14%,而 CoALTP 方法的特征集平均准确率为 88.11%,改进后的平均准确率提升约 11 个百分点;DRCoALTP\_BMS 特征集平均准确率为 99.44%,相比 CoALTP\_BMS 特征集平均准确率高出 11.33 百分点。另外,

采用 BMS 特征选择不仅减少超 3/4 的特征,还提高了识别准确率。CoALTP 算法对吉尔吉斯斯坦文的识别准确率为 1%,简体中文为 73.5%,塔吉克斯坦文为 78%,特征选择后 3 个文种的识别准确率分别为 93%、100%和 93.5%,但哈萨克文与维吾尔文的识别率明显降低,可能是 BMS 过滤了相应文种的判别性特征,导致特征子集识别特定文种的性能下降。类似地,分析 DRCoALTP\_BMS 与 DRCoALTP 的数据可以证明 BMS 特征选择算法的有效性。

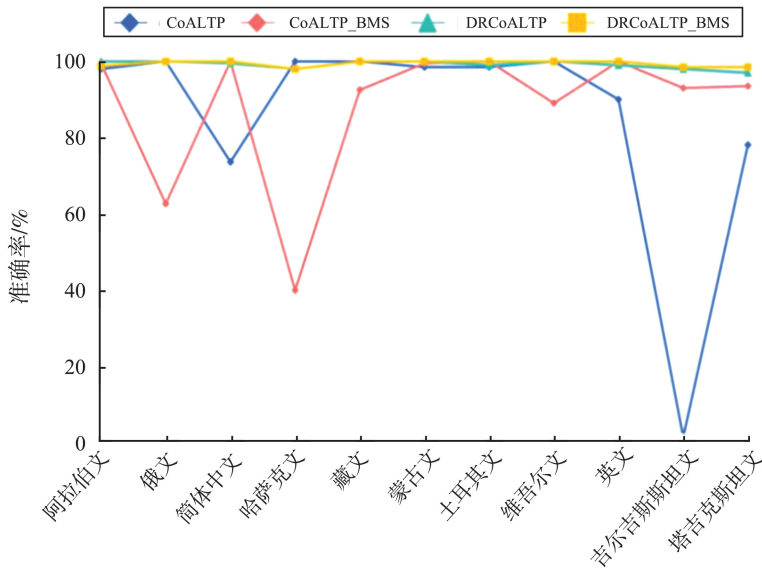


图 5 CoALTP、CoALTP\_BMS、DRCoALTP 和 DRCoALTP\_BMS 方法的试验结果比较

Fig.5 Comparison of experimental results between CoALTP, CoALTP\_BMS, DRCoALTP and DRCoALTP\_BMS methods

混淆矩阵是监督学习任务的一种评价方式,用于统计每个类别被正确识别与错误分类的数量,DRCoALTP\_BMS 方法试验结果的混淆矩阵如表 1 所示。混淆矩阵通过训练集获得,取 5 折交叉验证后的平均准确率作为结果。由表 1 可以看出:哈萨克文和吉尔吉斯斯坦文之间的误判相对较多,平均误差为 2.5%,这是由于 2 个文种的文字结构比较相似;维吾尔文、英文和哈萨克文中有近 1%的文档图像被错误认为是塔吉克斯坦文;土耳其文中也有近 1%的文档图像被错误地分类为英文;塔吉克斯坦文中有约 1%的文档图像被错误识别为哈萨克文和吉尔吉斯斯坦文。正预测类为每个文种的文档图像数量与预测为该类文种的数量的比,即平均准确率 99.0%以上。试验条件下哈萨克文、吉尔吉斯斯坦文与塔吉克斯坦文容易被机器混淆,与人类对这 3 个文种的视觉感受相似。

2.2 改进的 CoALTP 与其他方法的对比

不同纹理描述方法具有各自的特征维度,为了

验证本研究方法的先进性,在同一数据集上对本研究方法和其他方法进行平均查准率的比较,如表 2 所示。由表 2 可知:对图像进行离散曲波变换 (discrete curvelet transform, DCT) 获得系数矩阵后,提取其中 82 维的能量特征,在数据集上的平均查准率为 97.00%;与 DCT 相同,有限脊波变换 (finite ridgelet transform, FRIT) 也是多尺度多分辨率方法,提取其系数矩阵标准差的统计特征,在数据集上的平均查准率为 89.30%;梯度方向直方图 (histogram of oriented gradient, HOG) 的单元参数大小设置为 36,128 像素×128 像素的文档图像有 16 个单元,4 个单元为一个块,方向的区间数设定为 9 个,最终特征维度为 36,在数据集上的平均查准率达 90.13%;LBP 特征在数据集上的平均查准率为 85.65%;LTP 是基于 LBP 的改进方法,平均查准率达 90.80%;本研究对转换后的 GLCM 进一步提取 20 个统计特征,如对比度、相关性、能量和同质性等,平均查准率达 78.33%;本研究方法在数据集上的平均查准率达 99.44%。结果表明,相对于其他方

法,本研究方法具有更高的识别准确率。此外,使用局部三值模式的共生矩阵(co-occurrence of local ternary patterns, COLTP)对 LTP 编码值提取共生矩

阵的纹理方法,和本研究改进 LTP 编码方式后的 CoALTP 方法拥有相同的特征长度,但平均查准率却不高,仅有 81.53%。

表 1 DRCoALTP\_BMS 方法试验结果的混淆矩阵  
Table 1 Confusion matrix of experimental results of DRCoALTP\_BMS method

真实文种类别	预测文种类别的平均准确率/%										
	阿拉伯文	俄文	简体中文	哈萨克文	藏文	蒙古文	土耳其文	维吾尔文	英文	吉尔吉斯斯坦文	塔吉克斯坦文
阿拉伯文	>99							<1		<1	
俄文		100									
简体中文			100								
哈萨克文				97						2	1
藏文					100						
蒙古文						100					
土耳其文							>99		<1		
维吾尔文								>99			<1
英文									>99		<1
吉尔吉斯斯坦文	<1			3			<1			97	
塔吉克斯坦文				<1						<1	>98
正预测类	>99	100	100	97	100	100	>99	>99	>99	97	>98

表 2 不同描述符的平均查准率对比  
Table 2 Comparison of average precision in different descriptors

试验方法	特征维度	平均查准率/%
DCT	82	97.00
HOG	36	90.13
LBP	256	85.65
RIU_LBP	10	84.30
RI_LBP	36	86.40
U_LBP	59	87.70
LTP	512	90.80
GLCM	20	78.33
COLTP	2 048	81.53
FRIT	128	89.30
本研究方法	499	99.44

### 2.3 半自适应阈值方法的噪声测试

为了验证半自适应阈值函数的抗噪性能,本研究从数据集中随机划分训练集和测试集,维持训练集不变,在测试集中分别添加高斯噪声和椒盐噪声,加噪后 DRCoALTP 方法的试验结果如表 3 所示。文档图像未添加噪声的平均准确率为 99.44%,在增加高斯噪声和椒盐噪声后,11 个文种在测试集上的平均准确率分别为 97.81%和 98.12%,平均召回率分别为 99.42%和 99.17%,而 CoALTP 方法在增加噪声后识别准确率低于 50%,几乎失去了对文种的判别性。因此,半自适应阈值方法对椒盐噪声具有良好的鲁棒性。为进一步说明针对不同方法的敏感性,本试验在噪声测试中对随机 300 张文档图像统计其所有阈值类型的数量,如表 4 所示,3 像

素×3 像素的圆形邻域内,文档图像的阈值搜索窗口下高斯噪声信号较少,椒盐噪声信号较多,其中椒盐噪声常量占比为 54.21%,同时也反映出中心像素的灰度为阈值的情况很少,相反每个噪声的中值占比更高。

表 3 加噪后 DRCoALTP 方法的试验结果  
Table 3 Experimental results of DRCoALTP method after adding noise 单位:%

文种	准确率		召回率	
	高斯噪声	椒盐噪声	高斯噪声	椒盐噪声
阿拉伯文	100.00	100.00	100.00	99.33
俄文	100.00	99.01	100.00	99.33
简体中文	100.00	98.04	100.00	98.67
哈萨克文	79.00	100.00	96.00	100.00
藏文	100.00	100.00	100.00	100.00
蒙古文	100.00	100.00	100.00	100.00
土耳其文	99.33	100.00	99.53	100.00
维吾尔文	100.00	100.00	100.00	100.00
英文	100.00	100.00	100.00	100.00
吉尔吉斯斯坦文	100.00	84.27	100.00	95.00
塔吉克斯坦文	97.53	98.00	98.13	98.50

表 4 半自适应阈值方法各类型阈值数量  
Table 4 The number of threshold types in the semi-adaptive threshold method 单位:%

阈值类型	阈值数量占比	
	椒盐噪声	高斯噪声
中值	45.75	77.16
中心像素	0.04	3.71
常量	54.21	19.13

## 2.4 系统推理速度测试

在系统推理速度方面,研究对比 CoALTP、DRCoALTP 和 DRCoALTP\_BMS 方法在模型训练和测试上花费的时间,结果如表 5 所示。由表 5 可知:改进后的 DRCoALTP\_BMS 方法相对于原始 CoALTP 方法训练时间减少 184%,测试时间提高 137%,提高了系统的实用性。

表 5 不同方法训练时间和测试时间对比  
Table 5 Different methods comparison of training time and test time 单位:s

试验方法	训练时间	测试时间
CoALTP	75.804 8	0.013 9
DRCoALTP	77.444 8	0.021 4
DRCoALTP_BMS	26.655 3	0.005 8

## 3 结论

虽然 CoALTP 集合了 LTP 和 GLCM 的优点,但在实际应用时仍会受到图像内容、噪声等因素的限制。原始的 CoALTP 方法相对于 LTP 和 GLCM 等局部纹理方法,特征维度过高,且在提取共生矩阵过程中可能混入冗余信息。经过多组试验表明,相较于其他纹理特征方法,本研究提出的 DRCoALTP 方法更具判别性和鲁棒性,在自建数据集上的平均识别准确率达 99.14%。同时,本研究采用的半自适应阈值方法和基于均方差的特征选择方法对特征的抗噪性和高效性进一步改善,效果显著。在推理速度方面,DRCoALTP\_BMS 方法加入基于均方差的特征选择过程,通过特征降维使文种识别速度提升 284%,保证了系统的实时性。

在未来工作中,将继续扩充数据集的文种数量,进一步考虑图像中字体大小、形式等影响因素,研究更轻量级的多文种识别方法。

### 参考文献:

[1] WANG G, JIN Y, LIU L, et al. Identification of East Asian languages based on multi-feature fusion[J]. Computer Science, 2013, 40(1): 273-276.

[2] OJALA T, PIETIKAINEN M, HARWOOD D. A comparative study of texture measures with classification based on feature distributions[J]. Pattern Recognition, 1996, 29: 51-59.

[3] NANNI L, BRAHNAM S, ALESSANDRA L. A simple method for improving local binary patterns by considering non-uniform patterns[J]. Pattern Recognition, 2012, 45: 3844-3852.

[4] QIAN X, HUA X, CHEN P, et al. PLBP: an effective

local binary patterns texture descriptor with pyramid representation[J]. Pattern Recognition, 2011, 44: 2502-2515.

[5] GUO Z, ZHANG L, ZHANG D. Rotation invariant texture classification using LBP variance with global matching[J]. Pattern Recognition, 2010, 43: 706-716.

[6] GUO Z, ZHANG L, ZHANG D. A completed modeling of local binary pattern operator for texture classification [J]. IEEE Transactions on Image Processing, 2010, 19 (6): 1657-1663.

[7] GUO Y, ZHAO G, PIETIKAINEN M. Discriminative features for texture description[J]. Pattern Recognition, 2012, 45: 3834-3843.

[8] MUTHULAKSHMI M, KAVITHA G. An integrated multi-objective whale optimized support vector machine and local texture feature model for severity prediction in subjects with cardiovascular disorder [J]. International Journal of Computer Assisted Radiology and Surgery, 2020, 15: 601-615.

[9] LORIS N, ALESSANDRA L, SHERYL B. Survey on LBP based texture descriptors for image classification[J]. Expert Systems with Applications, 2012, 39 (3): 3634-3641.

[10] SHI M, HEALEY G. Hyperspectral texture recognition using a multiscale opponent representation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2003, 41(5): 1090-1095.

[11] REN J, JIANG X, YUAN J. Noise-resistant local binary pattern with an embedded error-correction mechanism [J]. IEEE Transactions on Image Processing, 2013, 22 (10): 4049-4060.

[12] SATPATHY A, JIANG X, ENG H. LBP-based edge-texture features for object recognition[J]. IEEE Transactions on Image Processing, 2014, 23 (5): 1953-1964.

[13] ZHANG B, GAO Y, ZHAO S, et al. Local derivative pattern versus local binary pattern: face recognition with higher-order local pattern descriptor[J]. IEEE Transactions on Image Processing, 2010, 19 (2): 533-544.

[14] TAN X, TRIGGS B. Enhanced local texture feature sets for face recognition under difficult lighting conditions [J]. IEEE Transactions on Image Processing, 2010, 19 (6): 1635-1650.

[15] PAPAKOSTAS G A, KOULOURIOTIS D E, KARAKASIS E G, et al. Moment based local binary patterns: a novel descriptor for invariant pattern recognition applications [J]. Neurocomputing, 2013, 99: 358-371.