

基于旋转投影统计特征的手写数字识别方法

庄伟¹ 雷小锋¹ 宋丰泰¹ 戴斌¹ 谢昆青²

(中国矿业大学计算机学院 徐州 221116)¹

(北京大学信息科学技术学院智能科学系视觉与听觉国家重点实验室 北京 100871)²

摘 要 通过抽取数字的轮廓和骨架来提取几何特征,可以有效地反映手写数字的细节,但手写数字的不规范性导致其识别率并不高。运用统计分析理论可以克服这一缺点。首先提出了基于投影间隔比率和间隔变化的特征提取方法,通过数字投影计算间隔的像素数比率和变化趋势,并将其归一化作为特征向量。进一步通过旋转投影基准线,增加特征向量之间的正交性以减少信息冗余,基于这一思路提出旋转投影的识别方法。理论分析和实验证明了旋转投影可以在相同特征数量的情况下达到更高的识别率,并给出了推荐参数。此外,通过旋转投影,直接解决了倾斜数字的识别问题。

关键词 统计分析,旋转投影,朴素贝叶斯,最小近邻法,手写数字识别

中图法分类号 TP391.4 **文献标识码** A

Method of Statistical Property Extraction Based on Rotational Projection in Handwritten Digit Recognition

ZHUANG Wei¹ LEI Xiao-feng¹ SONG Feng-tai¹ DAI Bin¹ XIE Kun-qing²

(School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China)¹

(Department of Intelligence Science/National Laboratory on Machine Perception, Peking University, Beijing 100871, China)²

Abstract To extract geometrical property by abstracting the contour and skeleton of digits can obtain the particulars efficiently, but irregularity induces low recognition rate. The statistical analysis theory can overcome the weakness. First, Gap-Rate and Gap-Variation extraction methods were probed, by projecting and computing the rate and change of each gap as the eigenvector of the digit. Further, Rotational Projection was probed based on the idea of enhancing the orthogonality of eigenvectors and reducing the redundancy information in the way of rotating the datum line. Theoretical analysis and experiment proved the high recognition rate of rotational projection with the same statistical properties, also parameters were recommended in the end. Further more, it also recognizes the oblique digit.

Keywords Statistical analysis, Rotational projection, Naïve bayes, Minimum nearest neighbor method, Handwritten digit recognition

数字识别(Handwritten Numeral Recognition)是“光学字符识别技术”的一个分支,经过多年的研究与发展,针对印刷体的数字识别技术已经取得了一定成就,如车牌号码、身份证号码等。然而对于手写体数字,由于人们书写习惯的不同和书写的随意性,使其识别率比较低,离实用还有一定距离。

手写体数字识别的研究对象是如何使计算机自动辨认人书写的阿拉伯数字,其识别的准确率主要取决于数字特征的提取。经过数十年的发展,研究者们提出了许多方法,大体分为两类^[4]:基于几何结构和基于统计分析。前者通过抽取数字的轮廓和骨架来提取几何特征,包括端点、交叉点、弧线、环、直线及凹凸性等^[5],在识别过程中能有效结合几何结构知识,准确反映印刷体数字的特性。但是对于手写体数字,端点数、交叉点数、环数、直线性及凹凸性都不是数字的固有属性,

它们会随着书写的随意性而改变,因此需要进行许多细节处理,如基于结构特征分类的数字识别^[5]、基于加强贝叶斯分类的手写数字识别^[6]。而后者主要通过点密度测量、矩、特征区域等得到数字的统计特征^[5]。基于这些特征的分类器易于训练,在给定的训练集上能得到相对较高的识别率,并且基于统计的方法可以有效地减小噪声对识别结果的影响,但往往具有较大的计算复杂度,如基于射线的轮廓特征提取方法^[7]。

本文采用统计分析方法,基于大样本下数字基本特征和规律的不变性,利用实验分析基于投影和投影间隔变化两种特征提取方法的特点和优劣,提出了基于旋转投影的统计特征提取。其通过增加特征向量之间的正交性减少信息冗余,达到更高的识别率;将投影、基准线以一定角度间隔绕原点旋转,将倾斜数字在基准线上进行投影得到特征向量;对所有角

到稿日期:2010-12-02 返修日期:2011-03-24 本文受国家 863 高技术研究发展计划(2006AA12Z217)、江苏省基础研究计划(自然科学基金)(BK2009093)、中国矿业大学科技基金(OD080313)资助。

庄伟(1986—),男,硕士生,主要研究方向为图像处理、机器学习等,E-mail:zwvs@163.com;雷小锋(1975—),男,博士后,主要研究方向为时空数据库与时空数据挖掘、机器学习等;宋丰泰(1989—),男,硕士生,主要研究方向为数据挖掘与模式识别等;戴斌(1982—),男,硕士生,主要研究方向为机器学习等;谢昆青(1941—),男,博士生导师,主要研究方向为智能信息处理等。

度的特征向量进行曲面拟合,得到该数字对应的特征曲面,最后利用最小邻近法匹配。实验表明,基于旋转投影的特征提取相对前面两种方法有更高的识别率和更小的计算复杂度。

1 投影间隔比率和间隔变化特征

由于固有误差以及书写时的偶然性和习惯性,人们通常无法将脑海中的数字模式准确地描绘出来,一般会在端点及交叉点有毛刺和曲折现象,如图 1(a)中圆圈标识处,但数字的基本模式不会改变,而这些就是这个数字的统计特性,在识别的过程中,要尽可能提取出来。

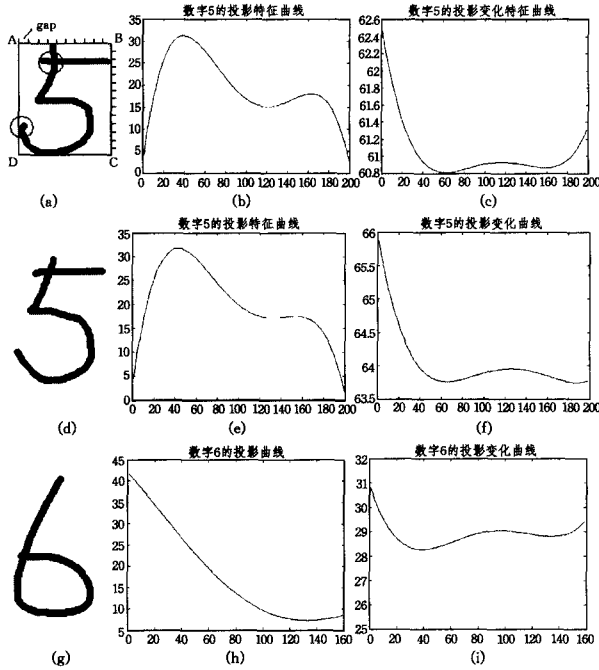


图 1 手写数字样本及其对应的特征曲线

在提取特征之前,通常需要利用彩色图像灰度化、二值化、去噪、平滑、细化、图像增强等方法进行预处理。本文基于统计分析的方法,绕开了复杂的细化和轮廓提取处理,只需对原始图像进行灰度化、二值化和去噪,降低了预处理的复杂度,减少了原始信息的损失。

1.1 基于投影间隔比率的特征提取

基于投影间隔比率的特征提取主要有以下步骤:

(1)提取数字外框。数字外框是指包含数字的最小矩形框,如图 1(a)所示的矩形 ABCD,在去除噪声的情况下,通过扫描一遍整幅图像,记录黑色像素出现的最早及最晚横纵坐标,即可得出数字的左上角和右下角坐标。数字外框为后续的特征提取和识别提供了方便,减少了计算量。

(2)选取间隔并投影。数字各个部分之间像素数的比例反映了数字的形状及内部结构,它是数字统计特性的重要表现,而利用数字在横向和纵向上间隔的像素数比率可以很好地反映这个特性。如图 1 所示,以一定间隔 gap 将数字外框在横向和纵向上切分成 n 等份,对外框中黑色像素进行投影,计算落在每个间隔 gap 内的累计像素数。

(3)计算间隔比率并归一化。以横向为例,将第(2)步得到的每个间隔 gap 的累计像素数除以数字总的像素数,就得到了每个间隔在整个数字中所占的比率 $X(x_0, x_1, \dots, x_{n-1})$:

$$x_k = \frac{\sum_{i=k \times gap}^{(k+1) \times gap - 1} \sum_{j=0}^{height-1} p(i, j)}{\sum_{i=0}^{width-1} \sum_{j=0}^{height-1} p(i, j)} \quad (1)$$

式中, $width$ 和 $height$ 表示数字外框的宽度和高度, gap 表示每个间隔的长度, $p(i, j)=1$ 表示在位置 (i, j) 处像素为黑色, 否则为 0。 x_k 表示一个随机事件,它反映了该数字在横向上第 k 个间隔中出现像素的可能性。对于同一类数字,在大样本下,该可能性在统计上趋向于某个数值——概率^[3]。由于其取值是 $(0, 1)$ 之间的任一小数,故间隔与间隔之间没有统一的可比性,需要将其归一化到一致的取值空间:

$$x_k = \left\lfloor \frac{x_k - \min(X)}{\max(X) - \min(X)} \times d + 0.5 \right\rfloor \quad (2)$$

式中, $\max(X)$ 和 $\min(X)$ 表示随机变量 $X(x_0, x_1, \dots, x_{n-1})$ 的最大值和最小值, d 表示新的取值空间中的最大值,加 0.5 表示四舍五入。经过归一化以后,随机变量 $X(x_0, x_1, \dots, x_{n-1})$ 的取值是一组 $[0, d]$ 之间的整数。对于纵向也进行类似的处理。

(4)设计朴素贝叶斯分类器。经过归一化后的间隔比率就可以作为设计分类器的特征向量 $X(x_0, x_1, \dots, x_{n-1})$ 。基于贝叶斯理论的朴素贝叶斯分类器是判断未知分类数据应该归属于哪一类的有效方法,其思想基础是对于给定的待分类项,求解在此项出现的条件下该项属于各个类别的概率,取归属概率最大的那一类作为待分类项最终所属的类别^[3]:

$$P(Y_i | X) = \frac{P(X | Y_i) P(Y_i)}{P(X)} \quad (3)$$

特征向量 $X(x_0, x_1, \dots, x_{n-1})$ 的各个分量分别代表各个间隔所占的比率,其取值是相互独立的。而对于一个给定的待分类项, $P(X)$ 是一个常数。因此计算给定待分类项属于类别 Y_i 的后验概率 $P(Y_i | X)$ 就变为计算在类别 Y_i 发生的情况下各个分量发生概率的乘积再乘上类别 Y_i 的先验概率:

$$P(Y_i | X) = C \times P(x_0 | Y_i) P(x_1 | Y_i) \dots P(x_{n-1} | Y_i) \times P(Y_i) \quad (4)$$

式中, C 是常数,先验概率 $P(Y_i)$ 可以通过计算训练集中属于每个类的训练样本所占的比例得到。最后,待分类项所属的类别是使得后验概率 $P(Y_i | X)$ 达最大的那个 Y_i 。

实验表明(见表 1),利用投影的方法提取数字的统计特征并结合朴素贝叶斯分类器,在减少复杂预处理工作的同时,对许多有微小倾斜和形状变形的数字也能正确识别,达到了较高的识别率。

1.2 基于投影间隔变化的特征提取

第 1.1 节对数字进行横向和纵向投影并计算比率可以较好地反映数字在各个间隔上的特性,而间隔之间像素数的变化趋势也在一定程度上反映了该数字的统计特性。例如对于手写数字 0、1、2、3、8,间隔与间隔之间的像素数变化比较缓慢,而数字 4、5、6 则呈逐渐上升的趋势,数字 7、9 有陡增的情况出现。因此,本节将间隔之间像素数的变化趋势作为数字的特征向量。特征提取和识别的步骤与第 1.1 节类似,不同之处在于第(3)步,在得到每个间隔的像素数以后,计算间隔与间隔之间的差分并归一化作为该数字的特征向量 $X(x_0, x_1, \dots, x_{n-1})$:

$$x_k = \frac{\sum_{i=(k+1) \times gap}^{(k+2) \times gap - 1} \sum_{j=0}^{height-1} p(i, j) - \sum_{i=k \times gap}^{(k+1) \times gap - 1} \sum_{j=0}^{height-1} p(i, j)}{gap} \quad (5)$$

由于后一间隔的像素数可能小于前一间隔,因此特征值 x_k 可能小于 0,同样需要利用式(2)进行归一化处理。

通过第 1.1 节介绍的投影方法对数字进行特征提取,可以得到一组特征向量,对其进行拟合所得到的特征曲线反映了数字在这种特征提取方法下的统计特性。实验表明,对同一类数字的不同样本,特征曲线在形状和变化趋势上表现相似。而对于不同类数字的不同样本,特征曲线有很大的差异,高内聚和低耦合性是提高数字识别率的重要指标。例如对于图 1(a)中的数字 5,图 1(b)和图 1(c)分别是该数字在投影间隔比率和间隔变化两种特征提取方法下所对应的特征曲线。同样,图 1(e)、图 1(f)、图 1(h)、图 1(i)分别是图 1(d)和图 1(g)在这两种方法下的特征曲线。由图可知,对于同一类数字 5 的两个不同样本图 1(a)和图 1(d),它们对应的特征曲线在形状和变化趋势上比较类似,而对于不同类的数字 6,特征曲线有较大的差异。通过比较可以发现,基于投影间隔变化所得到的特征曲线,其相似性不如基于投影间隔比率所得到的特征曲线,因此其识别的准确率也不如后者。后面的实验结果也证明了这一推断。

2 基于旋转投影的特征提取

第 1.1 节和 1.2 节介绍的方法只在横轴和纵轴两个方向上提取特征值。考虑到数字的统计信息在各个方向上都有均等的表达能力,我们设想:将投影基准线以一定角度间隔旋转,使得每条基准线上所取得的特征信息尽可能正交。这样减少了冗余,保证了特征信息的高效性和完整性,数字识别率也有进一步的提高,而且利用旋转的方法对于数字的倾斜也有很好的免疫作用。具体步骤如下:

(1)提取数字外框。方法与第 1.1 节相同。

(2)选取角度并投影。如图 2(a)所示,对于倾斜的数字,选取经过数字外框左下角与横轴夹角为 θ 的直线作为投影基准线,将黑色像素在这条基准线上投影。具体方法为假设数字外框左下角为坐标原点 $(0,0)$,则直线方程为 $Ax + By = 0$,像素 $p(i, j)$ 在直线上的垂足坐标 (x_{ij}, y_{ij}) 为: $(\frac{B \times B \times i - A \times B \times j}{A^2 + B^2}, \frac{A \times A \times j - A \times B \times i}{A^2 + B^2})$ 。计算垂足到原点的距离,当夹角 $\theta \in (\frac{\pi}{2}, \pi) \cup (\frac{3\pi}{2}, 2\pi)$ 时,垂足可能落在原点的两旁。因此,定义距离计算方法如下:

$$Dist(x_{ij}, y_{ij}) = \begin{cases} \sqrt{x_{ij}^2 + y_{ij}^2}, & y_{ij} \geq 0 \\ -\sqrt{x_{ij}^2 + y_{ij}^2}, & y_{ij} < 0 \end{cases} \quad (6)$$

(3)选取间隔并计算比率。首先将每个像素对应的距离 $Dist(x_{ij}, y_{ij})$ 归并为非负数:

$$Dist(x_{ij}, y_{ij}) = Dist(x_{ij}, y_{ij}) - \min(Dist) \quad (7)$$

$\min(Dist)$ 表示在这一角度下所有距离的最小值,其值可能小于零。然后选取一定间隔 $gap = \frac{\max(Dist)}{n}$,其中 $\max(Dist)$ 表示归并后的距离最大值, n 表示间隔数。计算落在每个间隔中的累计垂足数以及该间隔的比率:

$$Count(k) = Count(k) + 1, k = \left\lfloor \frac{Dist(x_{ij}, y_{ij})}{gap} + 0.5 \right\rfloor \in [0, n-1] \quad (8)$$

$$Rate(k) = \frac{Count(k)}{\sum_{i=0}^{n-1} Count(i)} \quad (9)$$

最后对 $Rate(k)$ 进行归一化,得到数字在角度为 θ 的直线上的投影特征向量 $X_\theta(x_0, x_1, \dots, x_{n-1})$:

$$x_k = \left\lfloor \frac{Rate(k) - \min(Rate)}{\max(Rate) - \min(Rate)} \times d + 0.5 \right\rfloor \quad (10)$$

式中, $\max(Rate)$ 和 $\min(Rate)$ 表示 $Rate(k)$, $k \in [0, n-1]$ 的最大值和最小值。

(4)利用最小邻近法匹配。对角度为 θ 时所产生的特征向量进行拟合,可以得到数字在该角度下的特征曲线。随着 θ 在 $[0, 2\pi)$ 的范围内逐渐变化,二维的特征曲线衍变成了三维的特征曲面。

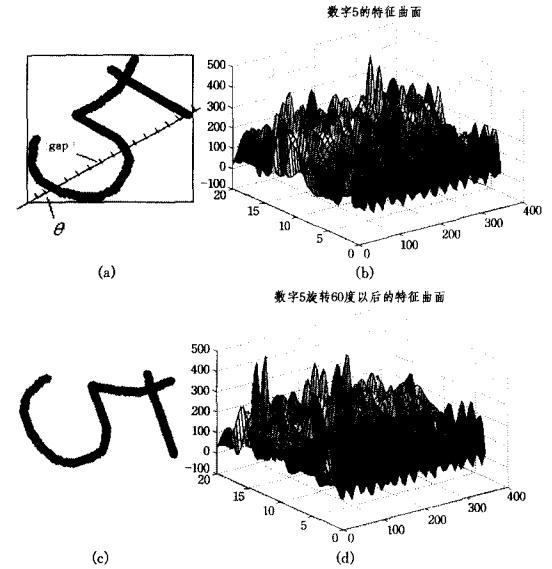


图 2 倾斜数字样本及其所对应的特征曲面

图 2(b)和图 2(d)分别是图 2(a)和图 2(c)的特征曲面。从图中可以看出,对同一个数字样本的不同倾斜角度 θ_1, θ_2 ,在特征曲面上的表现为 θ 轴上有一个位移 $\Delta\theta = \theta_2 - \theta_1$ 。将超出 2π 的部分曲面整个平移到最后端,以填补前面的移出空缺。为了减少计算量,在训练的时候,我们选取的所有样本都是正立的手写数字,对它们进行旋转投影得到每个数字样本的特征曲面。对属于同一类数字的所有特征曲面求平均,得到该类数字的平均特征曲面:

$$X_\theta = \frac{\sum_{i=0}^{s-1} X_{\theta_i}}{s} \quad (11)$$

式中, X_{θ_i} 表示第 i 个样本在角度为 θ 的直线上投影所得到的特征向量, s 表示总的样本数。平均特征曲面就是将得到的平均特征向量 X_θ 在 $\theta \in [0, 2\pi)$ 上拟合。在识别的时候,只需将待识别数字的特征曲面在 θ 轴上平移,并计算与样本平均特征曲面的差,取最小的差值作为待识别数字与该类数字的差异值:

$$Diffc = \min_{0 \leq \alpha < 2\pi} \left(\sum_{\theta=0}^{2\pi} |X_\theta - X_{\theta+\alpha}| \right) \quad (12)$$

式中, $X_{\theta+\alpha}$ 表示待识别数字的特征曲面平移 α 角以后在 θ 处的特征向量。最后差异值最小的那个类就是待识别数字所属的类 $C = \min_{0 \leq \alpha < 9} (Diffc)$ 。实验证明(见表 1),该方法无需对数字进行倾斜校正即可达到很高的识别率,具有较好的通用性。

3 实验内容及结果分析

对上述 3 种特征提取方法,我们选取了 1000 个正立的手

写数字作为训练样本,300个正立的和200个倾斜的作为测试样本,其中每个样本是以单个bmp文件形式保存,因此其读取速度比较慢,得到的实验结果如表1所列。

表1 3种特征提取方法对应的识别率及消耗时间

	0(%)	1(%)	2(%)	3(%)	4(%)	5(%)	6(%)
实验1	86.67	80.00	100	93.33	100	100	100
实验2	90.00	66.67	76.67	60.00	73.33	86.67	100
实验3	100	100	100	100	95.00	100	100

	7(%)	8(%)	9(%)	识别率 (%)	训练时间 (s)	测试时间 (s)
实验1	96.67	100	96.67	95.33	2.211	4.138
实验2	96.67	73.33	100	82.33	2.178	4.127
实验3	100	100	100	99.50	130.52	82.61

实验1、2、3依次代表第1.1节、1.2节、2节所介绍的方法。其中实验1、2的间隔数 $n=10$;实验3的角度间隔 $\Delta\theta=1^\circ$,投影间隔数 $n=20$ 。从实验结果可以看出,实验1在整体上达到了比较高的识别率,主要错误在于将数字0、1识别成8,原因是这两个数字在间隔比率的分布上都比较均匀。实验2基于投影间隔变化的特征提取方法,其识别率不是很高,主要误识数字是1、2、3,因为它们像素数的间隔变化上不是很明显,而对于具有突变的数字7、9,识别率则比较高。实验3表明,增加数字的特征提取方式和维度可以提高数字的识别率,但同时增加了计算量,消耗了更多时间。从信息论的角度看,全方位、高维度的特征提取保留了数字的更多信息,其中包括对提高识别率不利的负面信息。因此需要在计算复杂度和识别效果之间进行权衡,在保证识别率的前提下,尽量剔除负面信息的影响,提取尽可能多的有利信息,并减少计算量。

对于上述3种特征提取方法,计算复杂度取决于参数投影间隔和旋转角度的选取。投影间隔数越多,间隔划分就越密集;角度间隔越小,旋转的投影基准线就越多,计算复杂度也越大。通过实验可以发现,数字识别率并没有随计算复杂度呈线性增长。在特征维度较低的时候,增加维度可以更全面地提取信息,因此识别率有较快的提高;当特征维度达到一定大小以后,维度的再增加所提取的特征大多是数字的细节信息,如端点毛刺、小环、连接断裂等,而这些负面信息减弱了能有效反映数字特征的有利信息的影响,是基于统计方法所要忽略的。图3(a)、(b)、(c)分别表示上述3种特征提取方法的参数选择和识别率之间的关系。

我们称能使数字识别率达最大并且具有最小计算复杂度的那个参数为最佳参数。而在实际应用中,为达到最大识别率往往需要足够的信息量,因而所选取的最佳参数常包含较多的特征值,具有较大的计算量,所以需要寻找一个在识别率和计算复杂度上都可以接受的参数。如图3(a)、(b)所示,对第1.1节和1.2节的特征提取方法,当选取的间隔数比较少时,通过增加间隔数能有效地提取数字的统计特征,而当间隔数增加到某一数值以后,间隔数的再增加所产生的负面信息会使识别率有所下降,因此常常选取使识别率达极大时的参数。第2节的特征提取方法有旋转角度间隔 $\Delta\theta$ 和投影间隔数 n 两个参数,它们共同决定着识别率。对其进行拟合,得到参数与识别率的关系曲面,如图3(c)所示。曲面存在一条明显的拐线(图中黑粗线所示),数字的识别率先迅速提高而后

基本趋于稳定。由于随机噪声的存在,曲面在稳定后出现较小的波动。拐线近似平行于 $\Delta\theta$ 轴,考虑到最小计算复杂度和随机误差,取拐线在投影间隔数 n 轴上的坐标范围 $n_3 \in \{8, 9\}$ 为实验3的一个参数。由于曲面在 $\Delta\theta$ 轴方向的变化很小,噪声所产生的影响比较大,因此无法通过该曲面比较准确地取得另一个参数 $\Delta\theta$ 。我们从比较3种特征提取方法的角度,在宏观上通过计算得到另一个参数的大小。

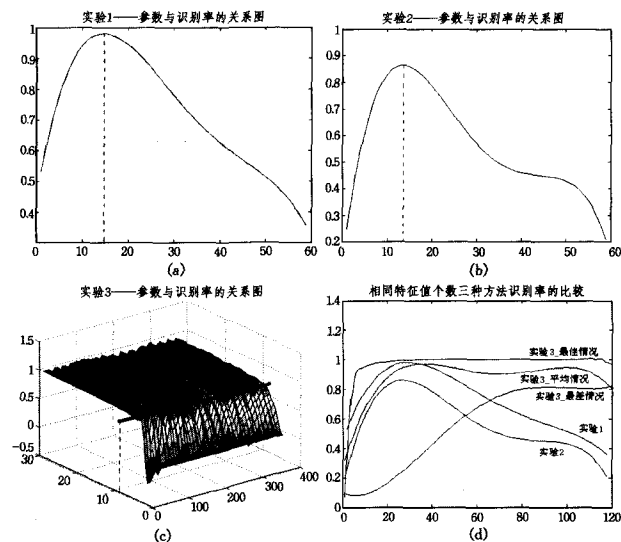


图3 参数选择和识别率的关系

对上述3种特征提取方法的评价,主要考虑识别率和计算复杂度两个方面,而计算复杂度主要体现在特征值个数上。图3(d)为3种方法的特征值个数与识别率的关系曲线图。实验3对于给定的特征值个数 n ,参数 $(n_3, \Delta\theta)$ 的组合可能有很多种,每个组合都对应一个识别率,从而产生了最高、平均、最低识别率3种情况。从图中可以看出,对于相同个数的特征值,基于旋转投影的方法所提取的数字有效信息最多,识别率也最高。从信息论角度来看,通过旋转的方法使特征向量之间尽可能正交,减少了冗余和相关性,能更全面地提取特征信息。同样,对相同的识别率,基于旋转投影的方法所需要的特征值个数最少,计算复杂度也最小。由图3(d)可知,对可接受的计算复杂度(特征值个数 $n \leq 120$),实验3最差情况的识别率比较低,不予考虑。而对于最佳情况和平均情况,识别率达到局部极大所对应的特征值个数 n 分别为27、32。代入式(13)即可算出另一参数 $\Delta\theta$ 的取值范围:

$$\frac{n}{n_3} = \left\lfloor \frac{2\pi}{\Delta\theta} \right\rfloor \quad (13)$$

式中, n/n_3 表示旋转的角度间隔数。由于旋转的次数一样,在 $\Delta\theta \in [2\pi \times n_3 / (n + n_3), 2\pi \times n_3 / n]$ 中选取的角度间隔所得到的识别率相近。因此我们选取3种方法的参数分别为(括号中为相应的特征值个数): $n_1 = 14(28)$, $n_2 = 13(26)$, $(n_{31}, \Delta\theta_1) = (9, 96^\circ)(27)$, $(n_{32}, \Delta\theta_2) = (9, 93^\circ)(27)$, $(n_{32}, \Delta\theta_2) = (8, 83^\circ)(32)$ 。对其实验得到表2的结果。

由表2可知,基于投影旋转的方法在特征值个数相同甚至更少的情况下,其识别率高于前两者,所消耗的测试时间反而最少。

(下转第302页)

- [12] Park H-H, Grings A, Santos M V D. Parallel hybrid evolutionary computation; automatic tuning of parameters for parallel gene expression programming[J]. Applied Mathematics and Computation, 2008, 201: 108-120
- [13] 邓松, 王汝传, 张羽, 等. 基于并行基因表达式编程的网格资源分配算法[J]. 电子学报, 2009, 37(2): 272-277
- [14] 都志辉. 高性能计算并行编程技术——MPI并行程序设计[M]. 北京: 清华大学出版社, 2001
- [15] Cuvillo J D, Zhu W, Gao G R. Landing OpenMP on cyclops64: an efficient mapping of OpenMP to a Many-core System on a Chip [C]//Proceedings of CF'06, Ischia, Italy, 2006: 41-60
- [16] 王小平, 曹立明. 遗传算法: 理论、应用与软件实现[M]. 西安: 西安交通大学出版社, 2002
- [17] 王文义, 秦广军, 王若雨. 自适应的多种群并行遗传算法研究[J]. 计算机工程与应用, 2006, 15: 78-81
- [18] 赖鑫生, 张明义. 基于渗透原理迁移策略的并行遗传算法[J]. 计算机学报, 2005, 28(7): 1146-1152
- [19] Blake C L, Merz C J. UCI repository of machine learning databases[OL]. <http://www.ics.uci.edu/~mllearn/MLRepository.html>, University of California, Irvine, Department of Information and Computer Sciences, 1998

(上接第 270 页)

- [2] Lin Bo, Fang Bin, Li Dong-hui. Character Recognition of License Plate Image Based on Multiple Classifiers[C]//Proceedings of the 2009 International Conference on Wavelet Analysis and Pattern Recognition, Baoding, July 2009: 12-15
- [3] Hsu C-W, Lin C-J. A Comparison of Methods for Multiclass Support Vector Machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2)
- [4] Wang Shen-zheng, Lee H-J. A Cascade Framework for a Real-Time Statistical Plate Recognition System[J]. IEEE Transactions on Information Forensics and Security, 2007, 2
- [5] Guo Jing-ming, Liu Yun-fu. License Plate Localization and Character Segmentation with Feedback Self-learning and Hybrid Binarization Techniques[J]. IEEE Transactions on Vehicular Technology, 2008, 57(3)
- [6] 罗鑫, 吴炜, 杨晓敏, 等. 一种基于 PCA 的多模板字符识别[J]. 电子测量技术, 2007, 30(1)

- [7] 邹明明, 卢迪. 基于改进模板匹配的车牌字符识别算法实现[J]. 研究与开发, 2010, 19(1)
- [8] 甘龙, 高隼, 梁栋, 等. 基于分形维数的车牌字符识别[J]. 中国公路学报, 2002, 15(4)
- [9] Duda R O, Hart P E, Stork D G. Pattern Classification[M]. China Machine Press, 2009: 94-95
- [10] Theodoridis S, Koutroumbas K. Pattern Recognition(Third Edition)[M]. Publishing House of Electronics Industry, 2008: 177-180
- [11] Otsu N. A Threshold Selection Method from Gray-level Histograms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62-66
- [12] 邹岚, 张桂林. 主成分分析方法(PCA)在车辆牌照识别中的应用[J]. 计算机与数字工程, 2007, 35(3)
- [13] 杨晓敏, 周强, 吴炜, 等. 基于支持向量机的车牌字符识别[J]. 四川大学学报: 自然科学版, 2009, 46(4)

(上接第 281 页)

表 2 推荐参数对应的识别率及消耗时间

	0(%)	1(%)	2(%)	3(%)	4(%)	5(%)	6(%)
$n_1=14$	100	90.00	100	96.67	100	100	93.33
$n_2=13$	93.33	76.67	66.67	70.00	93.33	96.67	100
(8,83)	100	100	100	100	95.00	100	100
(9,93)	100	95.00	100	100	100	100	100
(9,96)	100	100	100	100	95.00	100	100
	7(%)	8(%)	9(%)	识别率 (%)	训练时间 (s)	测试时间 (s)	
$n_1=14$	100	90.00	100	97.00	2.275	5.518	
$n_2=13$	96.67	83.33	93.33	87.00	2.226	4.428	
(8,83)	100	95.00	100	99.00	4.619	0.908	
(9,93)	100	95.00	100	99.00	4.084	1.000	
(9,96)	100	100	100	99.50	4.055	0.990	

结束语 本文介绍了基于统计分析理论的手写数字识别, 首先提出了基于投影间隔比率和间隔变化的特征提取方法, 然后从信息论的角度分析特征提取, 提出了基于旋转投影的识别方法。最后理论分析和实验证明了旋转投影在特征数量相同的情况下具有更高的识别率, 并且解决了倾斜数字的识别问题。在今后的工作中, 将对几何结构和统计分析相结合的数字识别方法作进一步研究。

参考文献

- [1] Gonzalez, Woods. Digital Image Processing[M]. New Jersey:

Prentice Hall Publisher, 2008

- [2] 求是科技. Visual C++ 数字图像处理典型算法及实现[M]. 北京: 人民邮电出版社, 2006
- [3] 盛骤, 谢式千. 概率论与数理统计(第 4 版)[M]. 北京: 高等教育出版社, 2008
- [4] 付庆铃, 韩力群. 基于人工神经网络的手写数字识别[J]. 北京工商大学学报: 自然科学版, 2004, 122(13): 44-45
- [5] 耿西伟, 张猛, 沈建京. 基于结构特征分类 BP 网络的手写数字识别[J]. 计算机技术与发展, 2007, 17(1): 130-132
- [6] 罗成, 孙越恒. 基于加强贝叶斯分类的手写数字识别[J]. 微处理机, 2009(03): 77-79
- [7] 宫淑兰. 手写数字识别的研究与应用[D]. 青岛: 山东大学, 2006
- [8] Duda R O, Hart P E, Stork D G. Pattern Classification(Second Edition)[M]. New York: John Wiley & Sons, 2001
- [9] Webb A R. Statistical Pattern Recognition(Paperback)(2nd Edition)[M]. New York: John Wiley & Sons, 2002
- [10] Duda R O, Hart P E. Pattern Classification and Scene Analysis [M]. New York: John Wiley & Sons, 1973
- [11] (美) Mitchell T. 机器学习[M]. 曾华军, 张银奎, 译. 北京: 机械工业出版社, 2003
- [12] Bloch G, Lauer F, Ching Y, et al. A trainable feature extractor for handwritten digit recognition[J]. Pattern Recognition, 2007, 40