

基于距离分布直方图的数字识别算法

吴少泓*, 王云宽, 孙涛, 李兵

(中国科学院自动化研究所, 北京 100190)

(* 通信作者电子邮箱 wushaohong0910@163.com)

摘要:由于自由字体与手写体数字形态的多变性,以往研究中具有较高准确率的算法往往牺牲了快速性,而具有实时性优势的算法却容易引起错误率的上升。针对这个问题,提出一种适用于快速数字识别的特征描述子——距离分布直方图(DDH),并在形状上下文的基础上提出一种既易于实现又具一定鲁棒性的描述子——形状累积直方图(SAH);然后将上述两个特征与其他改进后的拓扑特征相结合,组成最终的多特征矢量,由于其子矢量是由不同方法提取出的独特的特征,因此具有互补性;与此同时,算法中用三种组合特征训练了三个支持向量机来作分类器,综合它们给出的结果和自信度来给出最后的分类结果。在自建数据集、MNIST 和 USPS 数据集上的实验结果显示,平均正确率最高达到了 99.21%,证明了算法的高效性和鲁棒性。

关键词:特征提取;距离映射;相对链码;形状上下文;支持向量机

中图分类号: TP391.41 **文献标志码:** A

Digit recognition based on distance distribution histogram

WU Shao-hong*, WANG Yun-kuan, SUN Tao, LI Bing

(Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Due to the mutability of unstrained or handwritten digits, most algorithms in previous study either forfeited easy implementation for high accuracy, or vice versa. This paper proposed a new feature descriptor named Distance Distribution Histogram (DDH) and adapted Shape Accumulate Histogram (SAH) feature descriptor based on shape context which was not only easy to implement, but also was robust to noise and distortion. To make hybrid features more comprehensive, some other adapted topological features were combined. The new congregated features were complementary as they were formed from different original feature sets extracted by different means. What's more, they were not complicate. Meanwhile, three Support Vector Machine (SVM) with different feature vector were used as classifier and their results were integrated to get the final classification. The average accurate rate of several experiments based on self-established data sets, MNIST and USPS is as high as 99.21%, which demonstrates that the proposed algorithm is robust and effective.

Key words: feature extraction; distance map; relative chain code; shape context; Support Vector Machine (SVM)

0 引言

数字识别是光学字符识别的一个分支,现今被广泛应用到各行各业,比如核对银行支票、机读邮政编码和物流行业。数字识别算法是数字识别领域的核心技术,而特征提取是模式识别算法中至关重要的一步,特征的好坏直接影响到后续步骤的复杂度以及分类器的性能。在过去的研究中,研究人员提出了许多特征提取的方法,大体可以分为三类:结构或形状有关的特征、变换域中的特征以及统计学上的特征。

方向特征、傅里叶描述子^[1]和侧面投影图都属于第一类,因为这些特征都是对形状的描述,它们都有直观上的意义。然而,这类描述子与形状的联系过于紧密,以至于它们不能应对类内变化的情形,而在手写体和多字体数字识别的场合,类内差异无法避免。Belongie 等^[2]提出了一种叫作“形状上下文”的新形状描述子来匹配形状。这种描述子具有缩放平移不变性,对小的几何形变和遮挡也具有一定的鲁棒性。但是直接将这种描述子用于数字识别,会涉及许多复杂的后续处理步骤,比如二分图匹配,这本身是一个非常耗时的优化难

题;并且“形状上下文”在大的形变的情况下也不是很有效。其他的结构特征还有链码、游程编码、特征段、笔画特征、奇异点等^[3-4]。提取这些特征,比如特征笔画和奇异点,需要精准的定义和缜密的编程实现。

第二类是对原图像经过一系列变换,在变换域中表现的特征。比如几何矩、傅里叶变换域特征、小波特征、Harr 特征、Gabor 特征和 KL 变换特征等^[5-7]。这些特征的共同点就是把注意力从图像域转移到抽象的变换域,期望在其中找到更潜在的异同。

最后一类是统计特征,比如概率分布、块分布。这些特征一般利用了直方图技术,经过适当的处理,归一化后的特征不受图像平移和小形变的影响。但在大多数情况下,仅用这些特征难以获得高的分类正确率。

除了以上这些基于传统特征提取的识别方法外,Simard 等^[8]提出了切线距离和切线传播来做模式匹配,利用新的距离度量空间^[9]来解决形状识别问题,这种方法对某些变换群具有不变性。一篇手写体识别的文章声称应用这个理论的实验结果达到了 3.26% 的错误率^[9],但该理论涉及到切线距离

收稿日期:2012-02-15;修回日期:2012-03-21。 基金项目:国家自然科学基金资助项目(60973005)。

作者简介:吴少泓(1986-),女,江西景德镇人,博士研究生,主要研究方向:机器视觉、智能控制;王云宽(1966-),男,山西忻州人,研究员,博士生导师,主要研究方向:机器视觉、智能控制;孙涛(1984-),男,辽宁沈阳人,硕士研究生,主要研究方向:机器视觉、智能交通;李兵(1982-),男,黑龙江哈尔滨人,博士,主要研究方向:机器视觉、智能控制。

计算,是一个繁琐的优化问题,使整个系统的实用性降低。

在这些特征提取和分类技术中,通常具有高准确率的同时也伴随着实现的复杂和低效。对于这种高效性和鲁棒性难以兼备的问题,本文提出了两种易于提取且鉴别力强的新特征,并与其他改进后的特征结合,组成多特征矢量,以训练一组基于支持向量机(Support Vector Machine, SVM)的多类分类器。

1 数字特征矢量的提取

数字的特征向量很多,每类特征向量对字体特征的凸显性也各不相同,为了数字识别算法的鲁棒性,适应字体、高度、宽带等不同,本文吸取了多特征融合的思想,提出了如下四种数字图像特征。

1.1 距离分布直方图(DDH)

1.1.1 距离变换

图像经过距离变换后得到一幅新图像,其各个像素值是原图对应位置的点到离它最近前景像素的距离值,原图可以是边缘图或细化后的二值图。对于输入图像 P ,假设背景像素值为0,前景像素值为1, P 的距离变换输出图像是与它相同尺寸的图像 T ,定义如式(1)所示,对于 T 中的每个像素:

$$T(p) = \min_q \{d(p, q) : P(q) = 1, q \in P\}; \forall p \in T \quad (1)$$

其中 $d(p, q)$ 代表一种距离度量,不同的距离度量引导出不同的距离变换,除了像图1中的“棋盘度量”外,还有欧几里得度量和城市街区度量。

图1是一幅真实图像的距离变换图(为了方便显示,此图中的值已缩放至(0,255))。

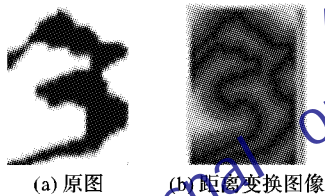


图1 距离变换的例子

1.1.2 距离分布直方图

从图1中可以看出距离图能反映出原图的形状特点,但是这样也出现对形状畸变敏感的问题。本文利用直方图的统计特性来弥补这点不足。

因为距离映射后的结果本身会因图像的缩放而变化,为了消除这种形变带来的影响,使用了归一化方法,经过处理的距离值将会有尺寸不变性。假设图像 T 是上一节距离变换后的图, NT 是按式(2)得到的归一化距离图,距离分布直方图定义如式(3)。

$$NT(p) = T(p)/T_{\max} \quad (2)$$

其中 T_{\max} 是归一化因子。

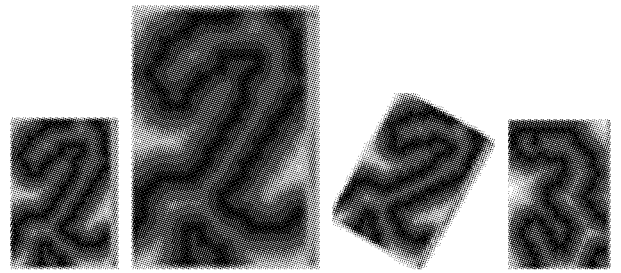
$$DH^k(NT) = \sum_p B^k(p); p \in NT, 0 \leq k \leq K_d, \quad (3)$$

$$B^k(p) = \begin{cases} 1, & T(p) \in bin(k) \\ 0, & \text{其他} \end{cases} \quad (3)$$

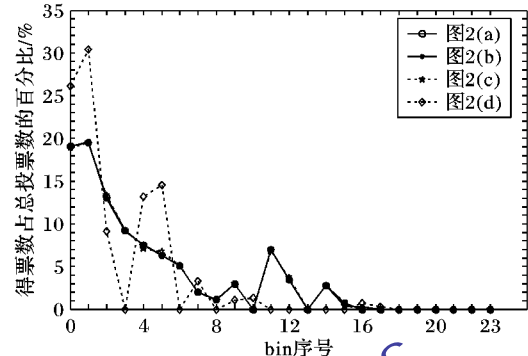
其中 K_d 是距离分布直方图的bin个数。关于参数的选择以及bin的分布将在2.4节参数选择一节中说明。

图2(b)、(c)分别是图(a)放大至1.5倍和经过旋转30°后的图,图(d)是另一个数字,图(e)是距离分布直方图。从图中可以看出,图(a)~(c)对应的直方图分布相似度远远大于它们与图(d)的直方图分布相似度。这种类间距离远大于类

内距离的优点证明了此特征的有效性和鲁棒性。



(a) 数字“2” (b) 图(a)放大至1.5倍 (c) 图(a)旋转30° (d) 数字“3”



(e) 图(a)~(d)的归一化距离分布直方图

图2 距离分布直方图特性说明

1.2 改进的相对链码直方图

Freeman链码^[10]能很有效地表现形状。统计图像的轮廓Freeman链码表示后各种链码值出现的频率,得到了该轮廓的链码直方图。这种直方图较好的特性之一是原图的旋转变换对应着直方图的周期性平移。为了得到理论上的旋转不变性,同时考虑到实际情况下存在无法预料的噪声,需用中值滤波先对轮廓的链码做平滑处理,然后依式(4)计算出轮廓各点的相对链码,最后如式(5)拓展成相对链码直方图。

$$c_i = G(f_i) - G(f_{i-1}) \quad (4)$$

$$h(m) = \#\{c_i = m\}; -7 \leq m \leq 7 \quad (5)$$

其中: f_i 是 C 上像素 p_i 的Freeman链码, G 是中值滤波函数。

1.3 形状累积直方图描述子

考虑一个由某点作为原点,到形状轮廓上所有除此点以外的点形成的矢量集合。这些矢量其实就表达了整个形状相对于这个参考点的结构信息。因为同属一个类别的形状会因真实采样的随机性而得到不同的实例,所以Belongie等^[2]将相对位置的概率分布视为一种鲁棒性更强、更紧凑,而且更有分辨力的描述子。对于形状上 I 的一点 p_i ,依式(6)计算其余 $n-1$ 个点的相对坐标的粗略直方图 h_i :

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in bin(k)\}; \quad 0 \leq k \leq K_s, K_s = K_\theta * K_\rho \quad (6)$$

这个直方图定义为点 p_i 的形状上下文。直方图的bin是在极对数坐标空间均匀分布的,使得该描述子受相对位置较近采样点的影响比较远的更大。其中, K_s 是直方图的bin的总个数, K_θ 是方向 θ 用的bin个数,参数 K_ρ 是对数极坐标 $\log r$ 用的bin个数。

文献[2]中利用此形状上下文前,必须先找到点与点的对应关系,但是获取这种对应关系是一个复杂的问题,本文放弃了该策略,提出了把每个形状上下文直方图归一化的方法。把每个上下文中对应维的数求和,依式(7)计算:

$$AH(k) = \sum_{p_i \in I} h_i(k); p_i \in I \quad (7)$$

得到一个积分直方图,即形状累积直方图AH(Shape

Accumulate Histogram, SAH)。该直方图包括了全局信息,对干扰和形变更具有鲁棒性,同时简化了实现,提高了效率。

1.4 改进的块分布特征

块分布特征提取思想描述如下:首先对输入图像进行特征区域划分和组合,然后统计各个特征区域中的特征值(即该特征区域中代表数字笔画的前景像素个数),各特征区域中的特征值的组合为特征向量,归一化的特征向量表征数字特征。

本文提出的区域划分和组合方法如图3所示。在图3中,首先将整幅图像均分为 3×3 个矩形特征区域,然后将特征区域1,2和3的组合命名为特征区域10,特征区域4,5和6的组合命名为特征区域11,特征区域7,8和9的组合命名为特征区域12,特征区域1,4和7的组合命名为特征区域13,特征区域2,5和8的组合命名为特征区域14,特征区域3,6和9的组合命名为特征区域15,依次统计特征区域1~15上的特征值 $\alpha_i (i = 1, 2, \dots, 15)$,将归一化后的特征值组合成特征向量 η ,即:

$$\eta_i = \alpha_i \cdot \left(\sum_{i=1}^{15} \alpha_i \right)^{-1}; i = 1, 2, \dots, 15 \quad (8)$$

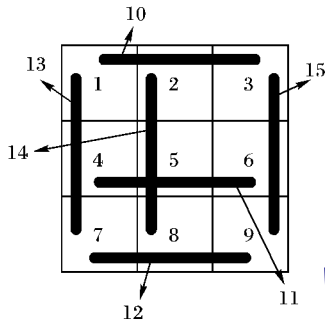


图3 数字图像的区域划分和组合

2 本文的识别算法

对于一个复杂的模式识别问题,一个分类器想要取得较高的准确率是困难的,特别是像手写体数字识别这种具有很大大手风格差异性的情形。文献[11]提到了几种降低错误率的方案:一种方案是在预测分类的后面加一级校验模块,这意味着需反复迭代,“识别—核实—识别”这种循环策略将会极其耗时;另一种是并行地组合多个分类器。

因为不同方法提取的特征矢量很大程度上是互补的,所以在解决分类问题上,多特征融合比单特征更具可靠性和鲁棒性。鉴于K-最近邻算法(K-Nearest Neighbor, KNN)需存储所有的训练向量,而多层感知器网络(Multi-Layer Perceptron, MLP)的训练耗时太长,本文训练了三个概率型支持向量机来作分类器。将不同方法提取的特征矢量并行地交给这三个分类器,综合各自的置信度,给出待识别数字的最终类别,具体的识别流程如图4所示(假设输入原图像是只含有一个待识别数字的二值图),实验结果证明了本文方法的有效性。

2.1 预处理

先找到输入图像中前景像素的最小包围盒 $Box(I)$,将这个盒作为感兴趣区(Region of Interest, ROI);然后用这个包围盒的倾斜角来校正此ROI,并归一化为 $W \times H$ 的图像,记为 $NormI$ 。

2.2 特征提取

首先提取各个子特征。 FV_1 是距离分布直方图:

$$EdgeI = \varphi(NormI) \quad (9)$$

$$FV_1 = DH(EdgeI) \quad (10)$$

其中: $\varphi(\cdot)$ 是Canny算子, $DH(\cdot)$ 是式(3)定义的对 $EdgeI$ 提取的距离分布直方图,记这个24维的向量为特征 FV_1 。

FV_2 和 FV_3 是块分布概率:

$$FV_2 = \eta(NormI) \quad (11)$$

$$ThinI = Thin(NormI) \quad (12)$$

$$FV_3 = \eta(ThinI) \quad (13)$$

其中: $Thin(\cdot)$ 是细化算子, $\eta(\cdot)$ 是式(8)介绍的块分布特征提取算法。

FV_4 是按1.3节中提到的方法计算 $EdgeI$ 的形状累积直方图,再对其归一化后得到的结果。即:

$$FV_4(k) = AH(k) \cdot \left(\sum_k AH(k) \right)^{-1} \quad (14)$$

FV_4 的维数就是此直方图的维数。

FV_5 是相对链码直方图。由于最后要得到的是相对链码的直方图,所以计算相对链码的起始参考点的选择可以是任意的。相对链码值的范围是-7到+7,所以归一化后的 FV_5 是15维。

$$C = Contour(NormI) \quad (15)$$

其中: $Contour(\cdot)$ 是轮廓跟踪算法,并把 C 作为式(4)、(5)的输入。

$$C(m) = h(m), -7 \leq m \leq 7 \quad (16)$$

FV_6 是有关数字的拓扑特征。先计算每行和每列的黑白跳变次数。10个数字的拓扑大体上是一定的,所有的跳变值不可能大于4,所以当计算得到大于4的值时视为噪声而将其饱和至4。具体得到65维的归一化直方图的方法如式(17)~(20)所示。

$$h1(k) = \{0.2 * p(k); p(k) = \text{第} k \text{行的跳变数}\} \quad (17)$$

$$h2(j) = \{0.2 * p(j); p(j) = \text{第} j \text{列的跳变数}\} \quad (18)$$

$$h(i) = \#\{\text{统计满足条件“跳变数} = i\text{”的行或列的总数}, i = 0, 1, 2, 3, 4\} \quad (19)$$

$$FV_6 = \{h1(0), h1(1), \dots, h1(23), h2(0), h2(1), \dots, h2(35), h(0) * \alpha, h(1) * \alpha, \dots, h(4) * \alpha\} \quad (20)$$

其中 $\alpha = 1/60$ 是归一化因子。

f_1, f_2 和 f_3 是分别送入三个支持向量机的三个复合特征矢量:

$$f_1 = FV_1 \oplus FV_2 \oplus FV_3 \oplus FV_4 \oplus FV_5; f_1 \in \mathbf{R}^{129} \quad (21)$$

$$f_2 = FV_1 \oplus FV_2 \oplus FV_3 \oplus FV_4; f_2 \in \mathbf{R}^{114} \quad (22)$$

$$f_3 = FV_1 \oplus FV_2 \oplus FV_3 \oplus FV_6; f_3 \in \mathbf{R}^{119} \quad (23)$$

其中 \oplus 是向量连接算子。

2.3 SVM训练和测试

选择高斯径向基函数 $K(x, y) = \exp(-\gamma \|x - y\|^2)$ 作为支持向量机的核函数。在训练阶段,依次通过2.1和2.2节介绍的方法提取特征得到三个复合特征向量 f_1, f_2 和 f_3 分别训练三个分类器;在测试阶段,测试样本经过特征提取后交予三个分类器,分类器给出各自的预测结果以及相应的预测自信度,根据规则来决定原图中的数字最终类别。关于支持向量机的实现细节,请参看文献[12]。

2.4 参数选择

考虑到这10种数字基本属于窄长型,故2.1节中的图像尺寸归一化参数选择为:图宽 $W = 24$,图高 $H = 36$ 。

对于1.1.2节式(2)、(3)中提到的参数 T_{\max} 和 K_d ,考虑到图中有数字的轮廓线,最大距离值不会大于图像最小包围盒长边的1/2,所以 $T_{\max} = 0.5 \times \min\{\text{NormIwidth}, \text{NormIheight}\}$;实验表明 K_d 的选取对结果的影响不大,在此选择在实验中得到最佳效果时的 K_d 的值, $K_d = 24$ 。

计算1.3节中提到的形状累积直方图用到的参数 K_θ, K_p 时,按文献[12]所建议的,对数极坐标 $\log \gamma$ 用5个bin, θ 用12个bin,即 $K_\theta = 12, K_p = 5$ 。详细原因请参考文献[12]。

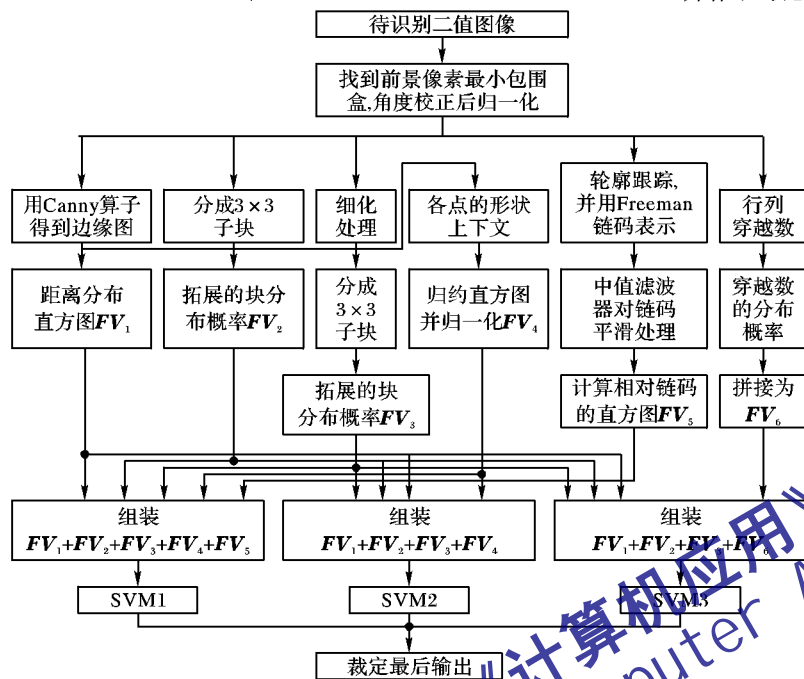


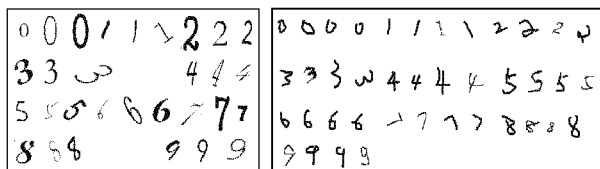
图4 本文识别算法流程图

3 实验结果与分析

3.1 总体性能测试

3.1.1 在自建数据集上的测试

本文用于论证算法的数据集由6436张图片组成,每张图片的大小不一,方向各异,如图5。把数据集分成两部分,其中的1096张作为训练集,剩下的全部作为测试集来用。表1给出了整个数据集的组成情况及本文算法的正确识别率。为了给支持向量机选择较好的参数,先用训练集训练,做5次交叉验证找到最好的 c 和 γ 参数,然后用得到的最好参数再次重新训练。



(a) 训练集中的样例

(b) 测试集中的样例

图5 数据集样例

表1 数据集的分布情况及本文算法的正确识别率

类别	训练样本数	测试样本数	正确识别率/%	类别	训练样本数	测试样本数	正确识别率/%
0	80	400	100.00	5	138	600	98.50
1	80	400	99.75	6	135	500	99.80
2	149	600	99.66	7	80	500	99.80
3	116	600	98.66	8	104	540	99.81
4	134	600	97.16	9	80	600	99.66

从表1可以发现,自建训练集远小于其他文献中所采用的数据集,但最终在测试集上得到的性能仍是令人满意的。经三个分类器后得到分类的平均准确率达到了99.21%,这也证实了算法的有效性。从中还可以看出,不同数字类别识别率不一样,有几类的类内形态差异很大,比如数字“4”,“5”,正是由于这种差异造成了其识别率的降低。

除了上述实验,还进一步将本文算法进行旋转校正拓展,并作了对比测试,看拓展策略能否提升性能。

由于在归一化时选择用的校正角是包围盒的长边与水平轴的夹角,这样即使是两幅图相似,当它们旋转的角度差异太大时,归一化后也可能成为水平或垂直镜像图,所以可以采用另一个办法来降低因角度变化而产生的误差。每个归一化后的测试图历经4次分类,逐次增加旋转 90° ,最后在这4次中选出自信度最高的那次预测结果作为待识别数字的类别。图6是每个数字类用原算法和拓展算法的结果对比。可以看出,实际上拓展算法并没有带来准确度的提升,有时完全的方向不变性是不希望的,这会带来更多的混淆或是模棱两可的结果,如识别“6”和“9”就是这个情况,图6的实验结果也证实了这点;而且拓展算法大大增加了处理时间。性能并没有因额外旋转试探而提升的这个事实也进一步证明了原算法本身就有很好的旋转鲁棒性。在以下的所有对比实验中,本文均忽略拓展而直接采用原算法。

另外还针对不同的支持向量机组合方式做了对比实验,结果如表2所示,从中可以看出本文选取的支持向量机组合方式能取得更好的结果。

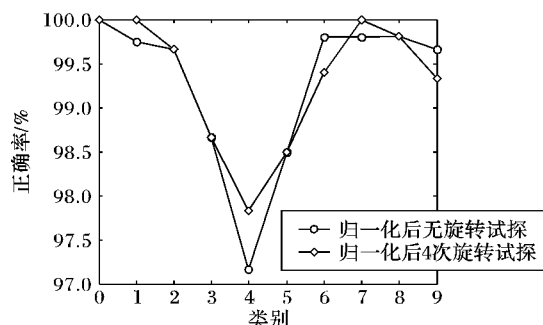


图6 归一化后无旋转和有4次旋转试探的效果比较

表2 不同架构的SVM的分类正确率对比

向量机组合	正确率/%
SVM1	95.22
SVM2	90.12
SVM3	89.99
SVM1 + SVM2	96.21
SVM2 + SVM3	94.43
SVM1 + SVM3	97.72
SVM1 + SVM2 + SVM3	99.21

所有实验都是在英特尔酷睿2的惠普笔记本上测试,主频1.8 GHz,本文算法的提取特征平均时间每幅是18 ms,识别用时小于1 ms。

3.1.2 在 MNIST 和 USPS 数据集上的测试

在 MNIST 数据集上选用 60 000 张样本图作训练集,其余的作测试集;USPS 数据集中每类数字有 1 100 个,随机从中选 800 个组成训练集,其余的用来测试。采用了三种分类器,分别是 SVM、BOOST 和 MLP;特征向量采用的是算法中 119 维的 f_3 向量;MLP 有 2 层隐含层,节点数分别是 500 和 300,再加上有 119 个节点的输入层以及 10 个节点的输出层;BOOST 使用 1 000 棵树。实验结果如表 3 所示。在这个数据集上的测试平均正确率不如文献[2,8]中所提到的那么高,但文献[2,8]的算法耗时且复杂,相比之下,本文算法更快,更简单,而且占用空间较少。正如在自建数据集上的实验结果所示,本文算法处理有旋转的情况更具优势,然而 MNIST 中旋转变化的不多,所以该优势不能在此数据集中很好地表现出来。

从测试结果的对比来看,在 MNIST 和 USPS 上的实验结果没有在建数据集上的好。但原因并不在于数据集的难易程度。从图 5 也可以看到,自建数据集里字体种类更多,任意方向也存在,甚至还有人造的扭曲形变。存在不同性能的原因可能是算法实现时参数的调整 and 选择上,比如归一化图像的大小(起先自建数据集考虑的是一般长宽不等,偏窄长型的数字,而后两个公共集中都是正方形数字),直方图的 bin 个数以及它们在整个范围内的分割,适合自建数据集的参数也许并不是后面两种数据集上的最优参数。

表 3 在 MNIST 和 USPS 数据集上的平均正确率比较

分类器	MNIST	USPS
SVM	93.52	94.61
BOOST	91.34	91.12
MLP	92.22	93.30

3.2 不同形变情况下的比较实验

为了证明本文算法对各种形变的鲁棒性,进行以下几组实验。首先,用软件生成 60 幅 32×32 的模板图片,各种数字都有 6 种不同字体的样本;然后把每个样本都旋转 45° , 90° , 135° ,总共得到 240 张样本图作为训练样本。

在测试尺度变化对算法性能影响时,测试样本是经过训练样本缩放后生成的。每个原始样本都缩放成 20 种不同的大小,从 12×12 到 52×52 (除了原始的 32×32),尺度间隔是 2×2 ,因此每个数字种类都有 480 个测试样本;同时也把训练图片旋转不同的角度来生成另一组测试样本,旋转角度从 0° 到 45° ,每隔 3° 进行旋转(不考虑从 -45° 到 0° 的旋转范围的

原因是避免“6”和“9”的混淆)。

为了证明算法对扭曲变形的鲁棒性,使用软件加入人造变形处理,每种数字都经历 20 种不同的形变。

考虑更复杂点的变形情况,即让原图经历不少于 2 种的复合变形。具体来说,产生 4 组测试样本,分别经历缩放与旋转的复合变换(S+R),缩放和扭曲的复合变换(S+D),旋转和扭曲的复合变换(R+D),以及缩放、旋转和扭曲的复合变换(S+R+D)。每种原始样本都产生 20 种复合形变后的测试样本,复合形变参数是随机选取的。图 7 是各种变形情况下的实验结果对比,可以看出本文算法对形变具有较强的鲁棒性。

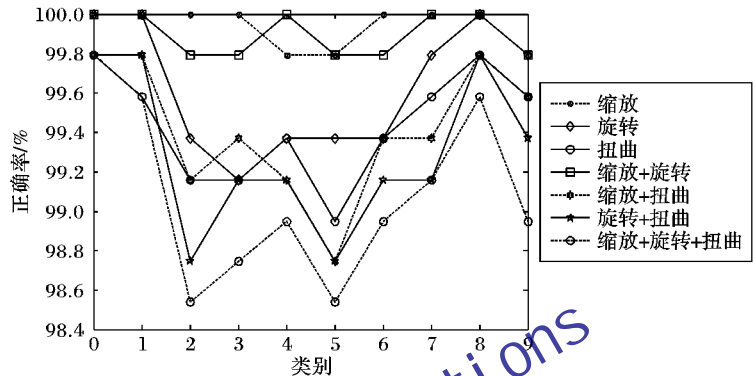


图 7 各种变形情况下的实验结果对比

3.3 与其他算法的性能比较

为了更清楚展示本文算法的优势,与其他文献中的算法在同等条件下进行测试比较。由于不是所有文献都提供了算法实现的源代码,为了能做比较实验,依照各自文献中的描述来实现这些算法,然后在同等的硬件平台上运行测试。虽然文献中有各自的实验结果数据,但基于的测试数据样本不能统一,所以为了使比较有意义,在此没有使用文献中的原始结果数据,而是采用我们对这些算法的实现,在选定的三个同样的数据集上进行比较。这三个测试数据集分别是前文提到的自建数据集、MNIST 数据集和 USPS 数据集。表 4 是各算法的在本文自建数据集上的正确率比较,图 8 是在 MNIST 和 USPS 上的正确率对比,图 9 是所有算法的平均耗时比较(图 8,9 中的算法序号 1~8 分别表示本文算法、文献[13]算法、文献[2]算法、文献[14]算法、文献[5]算法、文献[16]算法、文献[3]算法、文献[8]算法)。

从各组实验结果来看,在多数测试情况下,本文算法都能取得更优或相当的性能。虽然在某些情况下,算法的正确率没有文献[2,8]中的高,但本文算法的耗时远远少于它们。从图 9 中可以看到,本文算法的耗时平均稍多于 20 ms,然而有些更高准确率的算法平均耗时会超过 90 ms。

表 4 在自建数据集上的各种算法准确率比较

类别	本文算法	文献[13]算法	文献[14]算法	文献[8]算法	文献[3]算法	文献[5]算法	文献[16]算法	文献[2]算法
0	100.00	96.75	99.75	100.00	93.00	94.00	99.75	100.00
1	99.75	94.25	98.50	99.25	92.25	93.25	95.50	99.50
2	99.66	87.16	93.50	98.33	88.16	90.50	94.50	99.50
3	98.66	86.66	93.50	98.00	88.83	92.00	94.16	98.83
4	97.16	92.30	92.30	96.16	87.66	92.83	93.50	96.66
5	98.50	88.83	93.16	95.50	89.66	91.83	93.16	98.66
6	99.80	92.16	95.60	98.80	93.00	92.40	94.60	99.40
7	99.80	92.80	93.20	99.20	92.40	93.40	95.20	99.40
8	99.81	93.80	96.48	97.96	90.90	96.48	93.51	99.25
9	99.66	93.30	95.83	99.16	91.50	93.30	94.33	99.66

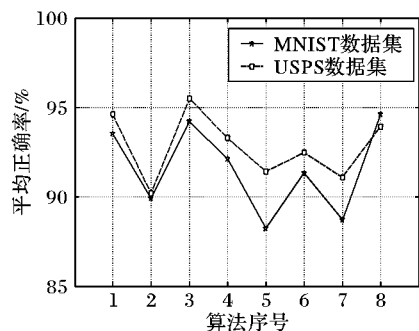


图8 在MNIST和USPS上的各种算法准确率比较

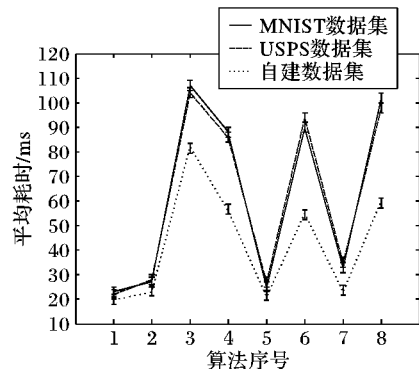


图9 在各数据集上的各种算法平均耗时比较

4 结语

针对文献中高准确率与快速性难以兼得的问题,本文提出了一种高效的特征描述子——距离分布直方图。此特征不仅保留了距离映射与原形状的紧密联系,又利用了直方图对一定形变有较好鲁棒性的特点。为了避免利用形状上下文寻找最优点点匹配这个复杂问题,同时也为了增强特征的鲁棒性,本文还提出了一种新的形状上下文的应用方法——形状累积直方图描述子。同时改进了表现拓扑特性的块分布和相对链码这些特征的提取方法:在特征提取时,对应不同阶段采用不同的归一化方法来避免不同原图像尺寸、不同方向的影响,并用相对链码取代链码来使之具有一定的旋转不变性。最后综合利用这三个新的复合型特征向量来用于自由体数字的识别,使用三个不同的复合特征训练了三个基于支持向量机的概率型多类分类器,最终输出是鉴于各分类器自信度的决策结果。实验结果证明了所提出算法是快速、可靠和鲁棒的。

参考文献:

- [1] 王涛,刘文印,孙家广,等.傅立叶描述子识别物体的形状[J].计算机研究与发展,2002,39(12):1714-1719.
- [2] BELONGIE S, MALIK J, PUZICHA J. Shape matching and object recognition using shape contexts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522.
- [3] 李丽洁,李金,宋阳,等.基于LBP和FSVM的视频文字定位方法[J].计算机工程,2011,37(24):144-146.
- [4] 李忠海,申为峰.基于方向码的旋转不变模板快速匹配方法[J].计算机工程,2010,36(16):210-212.
- [5] LIU GANG, WU DI, ZHANG HONG-GANG, et al. A new feature extraction method based on Fourier transform in handwriting digits recognition[C]// Proceedings of 2002 International Conference on Machine Learning and Cybernetics. Piscataway: IEEE, 2002, 3: 1227-1231.
- [6] WANG XUEWEN, DING XIAOQING, LIU CHANGSONG. Gabor filters-based feature extraction for character recognition[J]. Pattern Recognition, 2005, 38(3): 369-379.
- [7] OUYANG W, ZHANG R, CHAM W-K. Fast pattern matching using orthogonal Haar transform[C]// CVPR 2010: IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE, 2010, 3050-3057.
- [8] SIMARD P Y, CUN Y A L, DENKER J S, et al. Transformation invariance in pattern recognition — tangent distance and tangent propagation[C]// Proceedings of Neural Networks: Tricks on the Trade. Berlin: Springer-Verlag, 1998: 239-274.
- [9] HUTTENLOCHER D P, KLANDERMAN G A, RUCKLIDGE W J. Comparing images using the Hausdorff distance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(9): 850-863.
- [10] ULLAH F, KANEKO S. Using orientation codes for rotation-invariant template matching [J]. Pattern Recognition, 2004, 37(2): 201-209.
- [11] LISBOA P J G, PERANTONIS S J. Invariant digit recognition by Zernike moments and third-order neural networks [C]// Second International Conference on Artificial Neural Networks. Piscataway: IEEE, 1991: 82-85.
- [12] CHANG C-C, LIN C-J. LIBSVM: A library for support vector machines [EB/OL]. [2011-11-05]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [13] YU DECONG, MA LIHONG, LU HANQING. Lottery digit recognition based on multi-features [C]// SIEMIS 2007: IEEE Conference on Systems and Information Engineering Design Symposium. Piscataway: IEEE, 2007: 1-4.
- [14] TEOW L-N, LOE K-F. Robust vision-based features and classification schemes for off-line handwritten digit [J]. Pattern Recognition, 2002, 35(11): 2355-2364.
- [15] LIU C-L, SAKO H. Class-specific feature polynomial classifier for pattern classification and its application to handwritten numeral recognition [J]. Pattern Recognition, 2006, 39(4): 669-681.
- [16] NISHIDA H. Curve description based on directional features and quasi-convexity/concavity [J]. Pattern Recognition, 1995, 28(7): 1045-1051.
- [17] LIU C-L, NAKASHIMA K, SAKO H, et al. Handwritten digit recognition investigation of normalization and feature extraction techniques [J]. Pattern Recognition, 2004, 37(2): 265-279.
- [18] 董卫军,周明全,耿国华.基于形状特征实现人脸身份识别[J].计算机科学,2011,38(4):275-278.
- [19] SADRIA J, SUENA C Y, BUIB T D. A genetic framework using contextual knowledge for segmentation and recognition of handwritten numeral strings [J]. Pattern Recognition, 2007, 40(3): 898-919.
- [20] 史婷婷,刘卫华,伍春晖.一种新的图像空间特征提取方法[J].计算机工程,2012,38(3):218-220.
- [21] FREDRIKSSON K, MAKINEN V, NAVARRO G. Rotation and lighting invariant template matching [C]// LATIN 2004: Proceedings of 6th Latin American Symposium on Theoretical Informatics, LNCS 2976. Berlin: Springer-Verlag, 2004: 39-48.
- [22] 李忠海,李申,崔建国,等.基于快速SIFT特征提取的模板匹配算法[J].计算机工程,2010,37(24):222-224.
- [23] ZHANG P, BUI T D, SUEN C Y. Extraction of hybrid complex wavelet features for the verification of handwritten numerals [C]// IWFHR'04: Ninth International Workshop on Frontiers in Handwriting Recognition. Washington, DC: IEEE Computer Society, 2004: 347-352.
- [24] 张新明,李双群,郑延斌.矩不变调整的二维Shannon熵图像分割及其快速实现[J].计算机科学,2012,39(1):276-280.
- [25] 李文珺,丁岳伟.一种验证码倾斜自动矫正算法[J].计算机工程,2011,37(15):205-207.
- [26] CHEN G Y, BUI T D, KRZYSAK A. Rotation invariant pattern recognition using ridgelets, wavelet cycle-spinning and Fourier features [J]. Pattern Recognition, 2005, 38(12): 2314-2322.