

基于 LVQ 神经网络改进算法的人民币号码识别方法¹

赵晶晶¹, 葛庆平¹, 房超²

¹ 首都师范大学信息工程学院, 北京 (100037)

² 中国大恒(集团)有限公司北京图像视觉技术分公司, 北京 (100080)

E-mail: jingjing518_510@163.com

摘 要: 光学字符识别是模式识别领域中最经典也是得到最广泛应用的方向之一,而人民币号码识别系统是应用光学字符识别技术的典型系统之一。本文提出一种基于 LVQ 神经网络的人民币号码识别算法。在该算法中,提取字符的网格特征,采用自适应增加新神经元的聚类技术和自适应学习速率调整技术来提高竞争神经元的利用率并缩短了训练速度。在人民币图片库中测试表明,该算法识别一个号码平均耗时 0.16ms, 识别正确率达到 99.59%, 能够满足人民币号码检测系统的需求。

关键词: 人民币号码; 神经网络; LVQ 改进算法;

中图分类号: TP 249; TP 391.4

1. 引 言

作为一种特殊的产品,人民币每一个产环节都有很严格术标准要求。在人民币号码的印制过程中,需要对号码的正误以及印刷质量做出判定,人工检查存在缺乏客观性、精度欠高、视觉容易疲劳、速度缓慢、不能及时反馈等问题,随着计算机、信息技术的发展,开发基于现代图像信息技术的人民币号码在线检测系统可以为人民币的生产工艺提供必要保障,而号码识别技术是号码在线检测系统的重要技术基础。

号码识别的基本流程是:对采集到的号码图像,经过图像预处理后,形成规一化后的二值字符图像,利用字符的点阵特征或网格特征设计分类器,将二值化后的字符图像以数组的形式输入到分类器中,输出字符识别结果。分类器可以采用模板匹配,或者利用大量的样本训练的神经网络来做分类器也是很有效的方法^[2-5]。模板匹配算法的识别效率高,但是模板的学习建立过程过于依赖主观判断,检测过程对图像质量的要求高,鲁棒性较差。

神经网络在本质上是一种非线性动力系统,通过学习获得知识,逼近目标,分辨模式。学习矢量量化网络(Learning Vector Quantization, LVQ)是一种混和网络,由输入层、竞争层和线性输出层组成。通过有监督及无监督的学习进行分类。比起普遍使用的 BP 网络,有其独特的优势。BP 网络的缺点在于收敛速度问题、局部极小点问题以及网络瘫痪等问题。LVQ 网络的优点是算法实现简单,模块化强、训练时间少、识别效率高。

人民币号码在线检测系统要求 120 个字符/秒的处理速度和和大于 99% 的准确率,其中 OCR 处理时间不能超过总检测时间的 30%, 需要为号码印刷质量的检测留下足够的处理时间,因此实时性和高准确率是设计 OCR 算法的基本准则。本文基于 LVQ 神经网络的方法进行了人民币号码识别算法的研究,在 LVQ 改进算法的基础上采用训练后自动增加新聚类神经元技术和自适应学习速率调整等技术,来提高竞争神经元的利用率、号码识别率和缩短训练速度。

¹ 本课题得到北京市教委项目(KM200710028018)的资助。

2. LVQ 神经网络工作原理

2.1 LVQ 算法

学习向量量化(简称 LVQ)是 Kohonen 在 1989 年提出^[6,7],基本思想是在给定初始权值(码字)的基础上,使用有类别属性的训练样本,通过监督式自适应学习的方法来校正这些权值。经过若干次训练后, LVQ 网络中神经元的权值点就基本反应了训练样本点的统计分布。尽管是一个有监督训练方法, LVQ 采用了无监督数据聚类技术,通过对数据的处理,获得了聚类中心。LVQ 的网络结构除了每个输出单元与一类相关外,与竞争学习网络的结构极为相似。其拓扑结构如图 1 所示,包括一个输入层,一个 Kohonen 层和一个输出层,其中 Kohonen 层也称为竞争层。

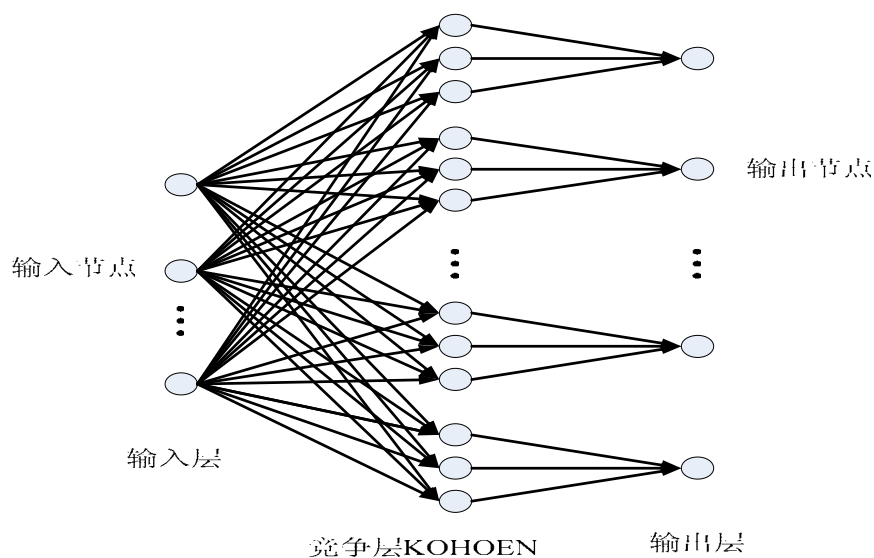


图 1 LVQ 神经网络的拓扑结构

Fig.1 Topol structure of LVQ neural network

LVQ 学习算法实质上是一种根据训练样本的特征,进行“奖一惩”的迭代学习算法。具体说就是,对分类正确的权值进行“奖励”,对分类错误的权值进行“惩罚”。经过若干次训练后,所得到的权值矢量不再明显变化,说明网络已收敛。令 X 为样本集, W 为神经网络的神经元集, C 代表获胜神经元, L 代表输出类别集。LVQ 神经网络的基本实现过程是:首先确定各网络参数,选择学习样本,然后选择一些具有代表性的样本对权值进行初始化,这样可以大大地加快学习的过程。接着就可以寻求获胜神经元 C :

$$\|X-W_c\| = \min_i \|X-W_i\|, \quad i=1,2,\dots,M \quad (1)$$

用 L_{wc} 代表与获胜神经元权值向量相关联的类,用 L_{wt} 代表与输入向量相关联的类。

$$\begin{aligned} W_c(n+1) &= W_c(n) + \eta(n)[X - W_c(n)] & L_{wt} &= L_{wc} \\ W_c(n+1) &= W_c(n) - \eta(n)[X - W_c(n)] & L_{wt} &\neq L_{wc} \end{aligned} \quad (2)$$

其中, $\eta(n) = \eta(0)(1 - (n/N))$ 是学习速率。后来 Kohonen 提出了几种改进算法: LVQ2、LVQ3 等,改进算法的主要思想是通过引入一个次获胜神经元来增加获得权值训练的神经元个数,从而缓解学习矢量量化的内在矛盾,使网络实现正确分类的可能性增大。对于 LVQ3 网络,将输入向量到获胜神经元 C 与次获胜神经元 R 之间的距离分别定义为 d_C 与 d_R , 用一个窗来

表示, 定义为:

$$\begin{aligned} d_C/d_R &= 1-\varepsilon \\ d_R/d_C &= 1+\varepsilon \end{aligned} \quad (3)$$

其中, ε 为给定阈值。窗条件如下:

$$\min[d_{C1}/d_{C2}, d_{C2}/d_{C1}] > (1-\varepsilon)/(1+\varepsilon) \quad (4)$$

当满足窗条件时, 权值按下式规则更改:

$$W_C(n+1) = W_C(n) + \beta(X - W_{C1}(n)) \quad (5)$$

其中, $\beta = m\varepsilon$, m 的范围是 $0.1 < m < 0.5$, m 可以选择为常数, 也可以随 ε 的变化而变化, 使得窗比较窄时学习速率 β 也较小。

2.2 LVQ 算法中学习速率的调整

LVQ 算法中学习速率 $\eta(t)$ 是个很重要的参数, 它影响算法的稳定性和权值收敛的速度, 是 LVQ 神经网络训练过程中需要重点考虑的参数。在定义学习速率的时候要贯彻快速稳定的原则, 其通常有三种定义方法:

a) 定学习速率: 简单的计算中不考虑训练时间, 用定学习速率最为简单, 其选取和数据的分布与数量有关。应用中一般取 $q=0.01-0.02$ 较为合适, 它可以保证学习的速率和收敛的稳定性。

b) 最优化学习速率: 最优化学习速率是 kohonen 提出的带有最优化思想的学习速率。它是在每次训练结束后, 根据该时刻训练样本在该时刻神经网络中识别结果的正确或错误来调整学习速率, 其定义如下:

$$\eta(t) = \eta(t-1) / [(1+s(t)) \eta(t-1)] \quad (6)$$

若在 $t-1$ 时刻的训练样本被正确分类, 在 t 时刻, $s(t)=+1$;

若在 $t-1$ 时刻的训练样本被错误分类, 在 t 时刻, $s(t)=-1$ 。

最优化学习速率中所需的循环次数较少, 但每次计算学习速率之前都要判断上一时刻的训练样本在该时刻神经网络中是否被正确分类。因此学习过程比较麻烦, 耗时较多; 对于具体问题还需要设定学习速率上限值, 以防止学习过程中学习速率过大而导致学习过程不稳定, 即当 $\eta(t) > \eta_0$ 时, $\eta(t) = \eta_0$ 。

c) 自适应学习速率: 一般来说, 学习速率 $\eta(t)$ 的选择需要考虑的是, 在迭代的初始时刻选择较大的值, 然后, 随着迭代的进行, 线性或非线性地降到 0。

$$\eta(t) = \eta(0)(1-t/T) \quad (7)$$

在初始阶段, $\eta(t)$ 通常取较大的值, 表示算法迅速修正较大的误分类权值, 随着时间的进行, $\eta(t)$ 越来越小, 表明使用较小的学习系数(缓慢的自适应)来修正较小的误分类权值, 以至在学习完成后, 误分类的样本数得到最少。使用自适应学习速率, 神经网络的权值在刚开始快些收敛, 结束时保证权值相对稳定, 因此它比定学习速率有较快的收敛速率和更高的稳定性。

从学习速率的选取来说, 自适应学习速率最好, kohonen 学习速率虽然有很好的理论效果, 但实际结果很差并且容易使算法发散。

3. 图像预处理

图像的预处理是进行基于 LVQ 算法的号码识别前的一种必要技术手段。每张人民币的号码是由 2 个英文字母（冠字 A—Z）和 8 个为阿拉伯数字（0—9）构成，进行预处理的目的是要获得 10 个号码在图像上的准确位置。根据获得图像样品质量和算法需求，首先通过滤波技术消除成像系统的光照和随机干扰等，然后利用图像分割技术确定每个号码的实际区域，最后将每个字符进行归一化处理成面积为 20 像素*40 像素的二值化样本，为后续的号码识别做准备。预处理流程及其分割结果分别如图 2 和图 3 所示：

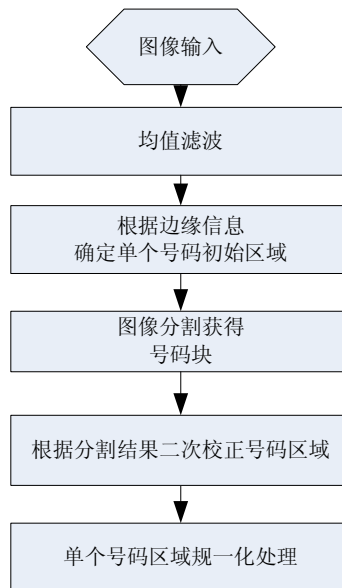


图 2：人民币号码图像预处理流程

Fig.2 the figures preprocessing flow chart of Rmb picture

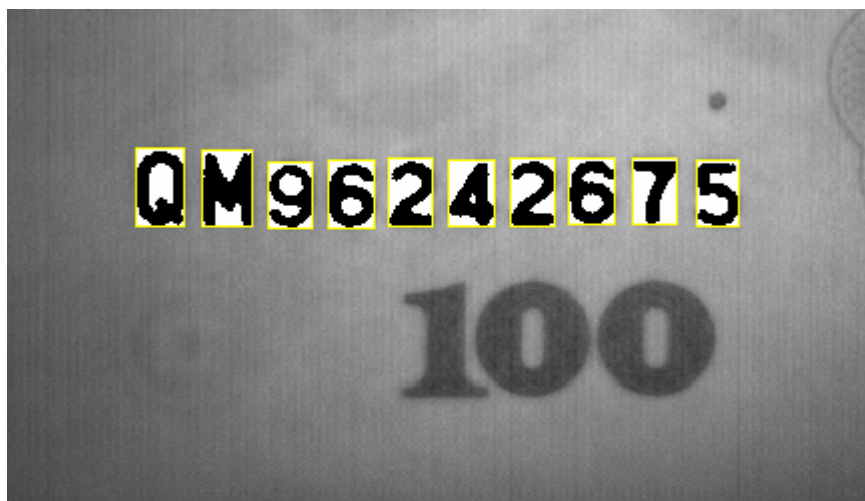


图 3 人民币样本图像号码区域分割结果图

Fig.3 the figures region picture of Rmb picture

3.1 号码结构特征的提取

对于号码的自动识别来说,必须通过输入字符的特征量来实现自动识别。特征量是对字符自身性质的一种描述,好的模式特征应该具备以下几个条件:(1)具有较好的类内一致性和类间区分度;(2)稳定性好,具有较好的抗躁能力;(3)具有较好的平移不变性、旋转不变性和尺度不变性等。^[9]人民币号码具有笔画少,结构简单的特点,同时由于人民币上英文字母和阿拉伯数字印制的位置顺序是预知的,所以识别的过程就被转化为一组英文字母的识别和一组阿拉伯数字的识别,两组之间无交叉,这样降低了识别特征量的选择的难度。

点阵特征提取是比较常用的一种二值图像的特征提取方法,二值字符图像的点阵结构反映了一个字符的整体结构特征,如果待识别的字符字体变换不大并且字符不是很大,不经过任何处理的点阵特征向量能够作为特征用于识别^[8]。

相对于点阵特征,网格特征提取能够更有效的反映字符的局部信息^[8]。如图4,对归一化后的字符进行分块,然后统计每一个块内前景像素的个数,最后以每块前景点像素个数与整个字符前景点像素个数的比值的归一化结果作为统计特征,组成 $\text{PreWidth}/\text{BLOCKSIZEH} * \text{PreHeight}/\text{BLOCKSIZEV}$ 维向量(其中, BLOCKSIZEH 和 BLOCKSIZEV 表示水平和竖直方向划分块的尺寸)。这种方法是对点阵结构的分区,也在一定程度上缩小了特征向量的维数。

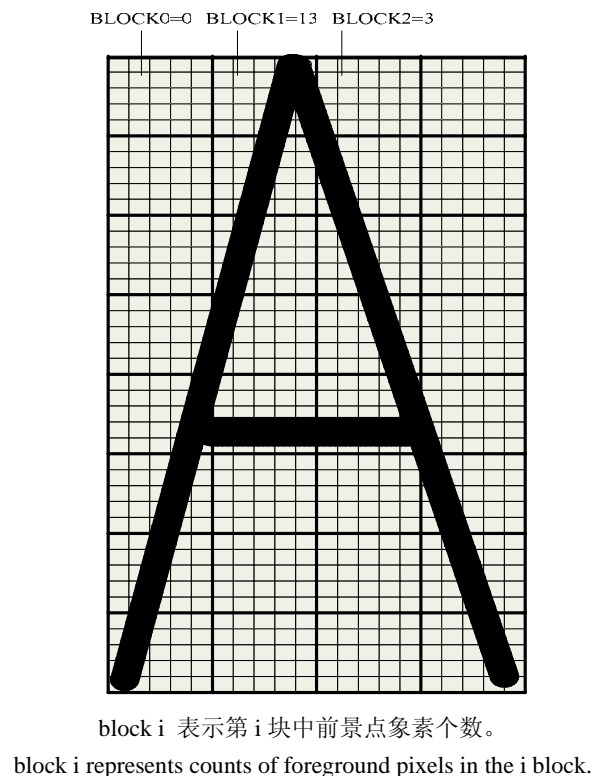


图4 网格特征提取

Fig.4 Grid feature extract

考虑到识别速度和算法的鲁棒性,本文采用网络的统计量信息作为基于LVQ算法的号码识别的特征向量。需要注意的是,个别结构相近字符,在采集图像和预处理时如果不能完全消除干扰的影响,统计量信息可能也会出现相似的情况,例如{“O”,“D”}, {“8”,“9”}等相似字符组。因此系统在LVQ分类识别后,对个别相近字符组利用拐角点等形态信息了细节

特征的二次判断, 确保系统的稳定性。

采用自适应增加新聚类的聚类技术和自适应学习速率调整技术来提高竞争神经元的利用率和缩短训练速度。

4. LVQ 神经网络的学习算法

本文对每个归一化处理后的字符区域进行 5*5 网格分块, 从而提取 25 个统计信息特征量作为输入。在训练前, 将所有样本与权值(Kohonen 层特征向量)都进行归一化, 使每一输入样本向量的模为 1, 这样做的目的不但可以提高网络的准确性, 而且可以大大加快整个训练的过程。

4.1 基于最小学习误差增量的神经元自动生成算法

对于同一代同一面值的人民币, 其号码形状是固定的。但是, 不同代不同面值的人民币, 其号码形状是不同的。用固定的 36 个输出样本制约了系统的兼容性和扩展性, 针对这个问题, 本文提出在号码检测系统中使用基于最小学习误差增量的神经元自动生成算法。

算法的实现过程如下:

1: 在训练前, 将所有样本与权值(Kohonen 层特征向量)都进行归一化, 使每一输入样本向量的模为 1, 这样做的目的不但可以提高网络的准确性, 而且可以大大加快整个训练的过程。

2: 预先设定主获胜神经元和次获胜神经元相应的学习误差上界 MAXERROR1 和 MAXERROR2。

3: 在训练的过程中, 当产生与获胜神经元和次获胜神经元均不聚类(获胜样本和训练样本的向量的欧式距离超过误差上界)的样本时, 系统新产生一个 Kohonen 层神经元组。

4: 如果 Kohonen 层产生了新的神经元, 则同时更新权值数量和输出神经元数量, 每一步都要对新更新和产生的权值进行归一化处理。否则根据“奖一惩”迭代式对获胜神经元的权值进行更新。保证 Kohonen 层每一神经元组与一组输入特征结构对应, 输出层根据现场数据的需求将不断变化。

4.2 样本训练参数的选择

神经网络的学习过程是离线进行的, 号码检测实时系统只需要自动更新学习之后的权值数量和输出神经元的数量, 因此学习速率并不是系统首要考虑的问题。尽管如此, 综合比较了三类学习速率的定义方法, 系统选择使用自适应学习速率, 基于以下原因:

- (1) 初始参数选择容易, 学习过程简单
- (2) 较高的收敛稳定性
- (3) 较快的收敛速率

具体的参数初始设置为: 输入层神经元数为 25; 竞争层神经元数为 36*25; 输出层神经元数为 36, 每个神经元对应一个单独的类别(0-9, A-Z)。采用自适应学习速率的技术: 初始学习率为 $\eta(0)=0.01$; 最大学习率为 $\eta_0=0.1$; 相对学习率: $m=0.3$; 窗口宽度: $\varepsilon=0.2$; 当训练样本分类均属于正确分类时停止训练。这些参数是在反复测试的基础上得出来的一组优化参数。

5. 实验结果及分析

本文的算法用 VC6.0++实现, 对传统 LVQ3 算法和改进后的 LVQ3 (训练后自动增加新聚类神经元技术和自适应学习速率调整技术) 从算法效率和准确率上进行了比较。

在 P4 XXGHZ, XXM RAM 的 PC 平台下, 在测试样本集中的 1248 张人民币样图像(即 12480 个号码)中, 提取 36 张人民币样本(共含有 360 个字符,其中包括 72 个冠字和 288 个数字)作为特征训练样本, 输入 LVQ 网络中, 得到训练好的神经网络。剩余的 1212 张人民币样本(共含有 12120 个字符,其中包括 2424 个冠字和 9696 个数字)作为待识别样本进行 LVQ 神经网络识别, 得到结果如表 1、2、3 所示。

表 1 不同学习方式的 LVQ3 算法识别率及训练时间对比

Tab1 Contrast recognition rate and train time between different LVQ3

主要算法	LVQ3	LVQ3 改进(本文)
Kohonen 层神经元是否自动生成	否	是
训练 360 个样本所需时间(s)	32	29
训练迭代次数	1000	1000
学习速率调整方式	自适应学习速率	自适应学习速率
样本识别率	12270/12480=98.32%	12432/12480=99.62%

表 2 不同学习方式的 LVQ3 算法的对比

Tab 2 Contrast among different learn rate adjustment

主要算法	LVQ3 改进 1	LVQ3 改进 2	LVQ3 改进 (本文)
Kohonen 层神经元是否自动生成	是	是	是
训练 360 个样本所需时间(ms)	61	74	56
训练迭代收敛次数	6	6	6
学习速率调整方式	定学习速率	最优化学习速率	自适应学习速率
样本识别率	99.57%	99.60%	99. 59%

表 1 和表 2 给出了采用不同学习方式下的实验对比数据, 通过表 1 可以看出, 相同算法情况下, 在训练过程中自动产生新的 Kohonen 层神经元组不但可以提高识别的准确率, 也可以在一定程度上缩短样本训练的时间; 由表 2 可以看出, 在相同算法的其它参数也相同的情况下, 自适应学习速率调整不但可以保证很高的识别率, 同时训练时间也有所缩短。

表 3 改进后的算法识别结果表

Tab3 Distribution of recognition results

	冠字符 (个) (A-Z)	数字字符 (个) (0-9)	待识别样本 (个)
总个数	2424	9696	12120
正确识别个数	2383	9687	12070
错误识别个数	41	9	50
识别率	98.31%	99.91%	99. 59%

表3列出了利用本文设计的LVQ算法的识别率,数据表明英文冠字的识别率相对较低,这主要是因为相似字符{“D”,“O”}的统计特征向量相似,导致总体识别率降低。利用在第3节介绍的方法从细节上再次校正相似字符,英文冠字错误识别的个数降低至4个。

实验过程中,对号码识别效率进行了测算,单个号码识别平均耗时0.16ms,这表明,本文设计的算法在识别效率上与传统的模板匹配的方法相似,但未进行二次校正前的识别率相对于传统的模版匹配方法有明显改善,达到了在线系统的实时性和识别率的要求。

6. 结论

通过研究,利用LVQ网络进行人民币号码识别是一种较好的途径。首先,基于LVQ神经网络的人民币号码识别方法在训练与分类的耗时上较以往的基于BP网络的分类方法大大减少,在分类速度上也能适应现场的实时检测;再者,本文通过改进LVQ算法,采用了训练后自动增加新聚类神经元的技术和自适应学习速率调整技术,有效的缩短了训练速度,并且提高了竞争神经元的利用率和号码识别率。

参考文献

- [1] 吴贵芳,徐科,徐金梧.基于LVQ神经网络的冷轧带钢表面缺陷分类方法.北京科技大学学报,2005(6): 27
- [2] Ballard DH,Brown CM.Computer vision[M].New Jersey:Prentice-Hall,1982:65-70.
- [3] JPratt W K.Digital image processing[M].New York:John Wiley&Sons,1991.
- [4] Lui H C,Lee C M,Fang G.Neural network application to container number recognition[C].Computer Software and Applications C-onference,1990:190-195.
- [5] Egmont P M,Ridder D,Handels H.Image processing using neural networks-a review[J].Pattern Recognition,2002:2279-301
- [6] Kohonen T. New Developments of learning vector quantization and the self-organizing map. In: Symposium on Neural Net-works. Osaka: Alliances and Perspectives in Senri, 1992
- [7] 吴简彤,王建华神经网络技术及其应用.哈尔滨:哈尔滨工程大学出版社,1998
- [8] 徐为.一种字符识别算法在自动识别系统中的应用.东北农业大学学报,2006(4):37
- [9] J Trier O D,Jain A K, Taxt T. Feature extraction methods for character recognition-A survey[J]. Pattern Recognition 1996,9(4): 641~662.

Application of LVQ Clustering Corrective Algorithm to Identification of RenMinBi Number

Zhao Jingjing¹, Ge Qingping¹, Fang Chao²

¹ Department of Information Engineering College, Capital Normal University, Beijing (100037)

² China Daheng Group, Inc. Beijing Image Vision Technology Branch, Beijing (100080)

Abstract

Optical character recognition(OCR) is one of classical fields in pattern recognition and becomes one of most successful applications. The RMB number auto recognition system is one of classical systems using OCR. LVQ is a method of neural network, In the paper, we designed LVQ clustering corrective algorithm, extract the the grid features, and adopt the self-adapting to increase new nerve cell, adjusting learn rate used the self-adapting, which improved the number of available competition nerve cell and cut down the speed of train. The result show that recognition one number average cost 0.16ms, the accuracy rate up to 99.59%, the algorithm can meet the needs of RMB number auto recognition system.

Keywords: RMB number, neural network, LVQ clustering corrective algorithm

作者简介:

赵晶晶(1984—),女,工学硕士研究生,主要从事机器视觉和图像处理研究。

葛庆平(1951—),男,副教授,主要从事机器视觉和图像处理研究。

房超(1972—),男,中国大恒(集团)有限公司北京图像视觉技术分公司视觉算法部经理。