

基于径向基概率神经网络的工程图纸图形符号识别

翟传敏¹, 杜吉祥^{2,3}, 黄 飞¹

(1. 合肥学院 机械工程系, 安徽 合肥 230022; 2. 国立华侨大学 计算机系, 福建 泉州 362021; 3. 中国科学技术大学 信息科学技术学院, 安徽 合肥 230026)

摘 要: 基于径向基概率神经网络, 提出一种扫描工程图纸图像分割后的图形符号识别方法. 针对已分割的扫描工程图纸图形符号图像, 首先进行二值化处理, 然后对二值图形符号图像进行 Hu 不变矩特征提取, 再使用一种新型的径向基概率神经网络进行分类, 从而实现图像识别. 为加快径向基概率神经网络的收敛速度, 采用递归最小二乘算法进行训练. 实验结果表明, 径向基概率神经网络在识别性能与速度等方面非常适合于工程图纸的图形符号识别.

关键词: 径向基概率神经网络; 图形符号; 工程图纸识别

中图分类号: TP31 **文献标识码:** A **文章编号:** 1673-4785(2006)01-0088-04

Graphic symbol recognition of engineering drawings based on radial basis probabilistic neural networks

ZHAI Chuan-min¹, DU Ji-xiang^{2,3}, HUANG Fei¹

(1. Department of Mechanical Engineering, Hefei University, Hefei 230022, China; 2. Department of Computer, National Huaqiao University, Quanzhou, 362021, China; 3. School of information Science and technology, University of Science and Technology of China, Hefei, 230026, China)

Abstract: A novel graphic symbol recognition approach of engineering drawings based on radial basis probabilistic neural networks (RBPNN) is proposed. The Hu invariant moment method is applied to extract the shape features of the segmented graphic symbol image of scanned engineering drawings. The experimental results show that the RBPNN achieves a higher recognition rate and better classification efficiency with respect to radial basis function neural networks (RBFNN) and multi-layer perceptron networks (MLPN) for the graphic symbol recognition task.

Keywords: radial basis probabilistic neural network; graphic symbol; engineering drawings recognition

工程图纸是工程技术人员描述设计对象、表达设计思想的主要工具. 据估计, 在全世界现有的 80 亿张工程图纸中, 85 % 以上是人工绘制的. 因此, 实现工程图纸的自动识别与解释一直是 CAD 技术追求的目标^[1]. 工程图纸识别与理解是以图像分析技术、计算机图形学及人工智能技术为理论基础发展起来的一门应用基础技术. 工程图纸识别技术是当今研究的一个非常活跃的课题, 在建筑、机械、交通等应用领域具有重要的意义. 从 20 世纪 80 年代初开始, 日本、美国、加拿大等工业发达国家对图纸的识别与理解技术的各个环节都开展了大量的研究,

如 CAD Overlay、Raserex、GTX、VPStudio 等. 国内从 20 世纪 80 年代中期起, 华中科技大学、清华大学、中国科学院自动化所、西北工业大学及浙江大学等单位也相继推出了一些图纸识别与理解系统, 如华中科技大学的 EDIS 软件、清华大学的 ANNO 等. 然而, 由于该问题涉及的学科多、难度大, 目前的研究结果普遍存在处理速度慢、智能程度低以及适应性能差等缺点^[1].

工程图扫描图像所表示的内容主要由 3 部分组成: 图线、文字和图形符号. 在工程图纸的自动识别研究中, 除了要图线进行矢量化的重构, 还要能够识别与其相关的文字、尺寸及其他标注符号. 目前这三方面的准确识别技术都未能得到很好的解决. 因此, 扫描工程图纸的图形符号自动识别是工程图纸识别和工程量自动计算中的重要组成部分. 目前, 对

收稿日期: 2006-02-15.

基金项目: 国家自然科学基金资助项目(60405002); 合肥学院自然科学基金研究基金资助项目(05ky013zr).

于工程图纸图形符号的识别,常采用的方法是根据图形符号特征建立识别函数库来进行匹配识别,因此灵活性较差.如果能够提取有效的特征,则识别的关键就在于设计具有良好分类能力的分类器.

作为分类器的一种,人工神经网络技术由于其自学习、自适应和高度并行性等优点而受到广泛的关注.迄今为止,BP 算法是应用最普遍的前馈神经网络学习算法之一,基于 BP 算法训练的多层感知机网络是目前应用最为广泛的一种神经网络模型. BP 网络虽然在理论上是一种可实现全局优化的网络,但由于算法本身的制约,比较容易陷入局部极小值^[2].而径向基函数神经网络(radial basis function neural networks, RBFNN)既有生物学背景,又与函数逼近理论相吻合,因此在模式识别和非线性函数逼近领域得到广泛应用.概率神经网络(probabilistic neural networks, PNN)则以其训练速度快的优点,已广泛用于模式识别领域^[2].在对比 RBFNN 和 PNN 2 种网络模型性能的基础上,文献[3]提出了一种新型的前馈神经网络模型—径向基概率神经网络(radial basis probabilistic neural networks, RBPNN)模型,这种新的模型既能利用 RBFNN 和 PNN 2 种模型的优点,又可避免它们的缺点^[2-5].

因此,文章针对工程图纸扫描图像分割后的图形符号,利用径向基概率神经网络的分类辨识特性进行分类识别研究.

1 图形符号特征提取

对于工程图纸这一特定研究对象,所要识别的图形符号相对于要识别的文字是有限的,而且一般是比较工整的手写体或印刷体.这些图形符号主要包括 10 个尺寸数字 0~9,以及其他常用标注符号 +、-、±、R、°、∠等,共 40 类^[6].

由于设计人员的习惯不同,以及图纸幅面的不同,导致这些符号的风格不同、大小不一、工整程度也不同.同时,由于图纸比例不同及扫描分辨率不同,所表现出的图形符号大小也存在差异,因此图形特征矢量的提取应具有平移、尺度和旋转不变性.

M. K. Hu 使用图像的几何矩的非线性组合推导出了 7 项具有旋转、平移和尺度不变性的不变矩,并且证明了矩对于信息的保留性,即矩序列由函数 $f(x, y)$ 唯一确定,反之矩序列也能完整准确地重建函数 $f(x, y)$,此后不变矩作为图像的特征被广泛地应用于二维图像识别领域当中^[7].对于 Hu 不变矩的计算方法有很多,在这里采用了 Chen 提出的改进矩算法^[8]来计算这些 Hu 不变矩,具体算法描述如下:

首先 Chen 的 $p + q$ 级改进矩可定义如下:

$$M_{pq} = \oint_C x^p y^q ds.$$

(1)

式(1)表示是沿着闭合轮廓的线积分.在实际应

用中采用离散化公式来近似计算:

$$M_{pq} = \sum_{(x,y) \in C} x^p y^q.$$

(2)

同样中心距 μ_{pq} 也就可以近似计算为

$$\mu_{pq} = \sum_{(x,y) \in C} (x - \bar{x})^p (y - \bar{y})^q,$$

(3)

$$\bar{x} = \frac{M_{10}}{M_{00}}, \bar{y} = \frac{M_{01}}{M_{00}}.$$

(4)

为了得到尺度不变性 μ_{pq} 可进一步正则化为

$$\rho_{pq} = \mu_{pq} / \mu_{00}.$$

(5)

式中:正则化因子为: $= p + q + 1$. 利用正则化中心距 ρ_{pq} 就可以计算出 7 个 Hu 不变矩:

$$Hu_1 = \rho_{20} + \rho_{02},$$

(6)

$$Hu_2 = (\rho_{20} - \rho_{02})^2 + 4\rho_{11}^2,$$

(7)

$$Hu_3 = (\rho_{30} + 3\rho_{12})^2 + (\rho_{03} - 3\rho_{21})^2,$$

(8)

$$Hu_4 = (\rho_{30} - 12\rho_{12})^2 + (\rho_{30} + 21\rho_{21})^2,$$

(9)

$$Hu_5 = (\rho_{30} - 3\rho_{12})(\rho_{30} + 12\rho_{12})[3(\rho_{30} + 12\rho_{12})^2 - 3(\rho_{03} + 21\rho_{21})^2] + (3\rho_{21} - \rho_{03})(\rho_{21} + \rho_{03})[3(\rho_{30} + 12\rho_{12})^2 - (\rho_{03} + 21\rho_{21})^2],$$

(10)

$$Hu_6 = (\rho_{20} - \rho_{02})[(\rho_{30} + 12\rho_{12})^2 - (\rho_{03} + 21\rho_{21})^2] + 4\rho_{11}(\rho_{30} + 12\rho_{12})(\rho_{03} + 21\rho_{21}),$$

(11)

$$Hu_7 = (3\rho_{21} - \rho_{03})(\rho_{30} + 12\rho_{12})[3(\rho_{30} + 12\rho_{12})^2 - 3(\rho_{03} + 21\rho_{21})^2] + (3\rho_{12} - \rho_{30})(\rho_{21} + \rho_{03})[3(\rho_{30} + 12\rho_{12})^2 - (\rho_{03} + 21\rho_{21})^2].$$

(12)

2 径向基概率神经网络分类器设计

2.1 径向基概率神经网络模型

径向基概率神经网络是一种新型的前馈神经网络模型,由径向基函数神经网络(RBFNN)和概率神经网络(PNN)发展而来^[2].因此,RBPNN 充分利用了 RBFNN 的样本集中模式交错影响以及 PNN 实时训练的优点,具有快速的收敛性能,可以在网络的输出端进行输入样本的概率密度估计等等.在某种程度上,RBPNN 也减少了 RBFNN 和 PNN 的缺点,如计算复杂度的降低等等. RBPNN 的网络结构如图 1,在拓扑结构上该网络有 4 层结构:即 1 个输入层、2 个隐层和 1 个输出层.第 1 隐层是非线性处理层,主要是由训练集中优选出的样本(又称隐中心矢量)组成.第 2 隐层根据其隐中心矢量的类别属性对第 1 隐层的输出进行选择性地求和.网络的输出层最终完成非线性映射,如分类、函数逼近和预测等.

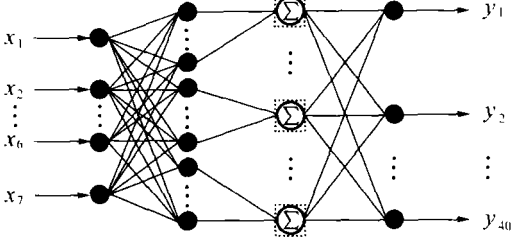


图 1 径向基概率神经网络结构

Fig. 1 The structure scheme of the RBPNN

假定训练样本集中包含 N 个样本,且这些样本归属于 M 个模式类别 ($N = M$),对于第 t 个输入样本 x_t ,则 RBPNN 输出层的第 j 个单元的 outputs 为

$$y_j(x_t) = \sum_{i=1}^M w_{ji} h_i(x_t) = h(x_t) w_j, j = 1, 2, \dots, M; \quad (13)$$

$$h_i(x_t) = \sum_{l=1}^{n_l} \Phi(x_t - c_l^i) = \sum_{l=1}^{n_l} \Phi(x_t - c_l^i - F), i = 1, 2, \dots, M. \quad (14)$$

式中: $w_j = [w_{j1}, w_{j2}, \dots, w_{jM}]^T$, $h(x_t) = [h_1, h_2, \dots, h_M]$, $h(x_t)$ 还可以表示为 $h(t)$, 其对应的 $h_j(x_t)$ 或者 h_{ij} 为网络第 2 隐层第 i 个节点的输出; $v_j(x_t)$ 或者 v_{ij} 为样本 x_t 输入时网络输出层第 i 个单元的 outputs; w_{ji} 是连接第二隐层第 i 个节点与输出层第 j 个节点的权重; $\Phi(\cdot)$ 是高斯核函数; c_l^i 是与第 2 隐层第 i 个节点相连的、在第 1 隐层对应的第 l 个隐中心矢量; n_i 是第 1 隐层第 i 个模式的隐中心矢量数; 是核函数控制参数; F 是矩阵的 F -范数。

2.2 径向基概率神经网络的 RLS 训练算法

从结构上来看, RBPNN 只有 1 层权重需要训练,即第 2 隐层与输出层间的连接权重,而且输出层的每个节点的 outputs 是第 2 隐层输出的线性叠加,因此, RBPNN 的训练和测试速度将远远快于 RBFNN 与 MLPN. 并且,对 RBPNN 来说,由于第 2 隐层将第 1 隐层的 outputs 按类别求和,即使某一训练样本 x_t 与所有隐中心矢量的距离都很远,即对应于 x_t 的第 1 隐层所有神经元的 outputs 都很小,但由于第 2 隐层求和的作用,使得第 2 隐层所对应的 outputs 都有较大的值,并保证相互间能够明显区分,这样就不会出现在输出端“拒识”该训练样本的情况,从而保证 RBPNN 具有较强的分辨能力. 在这一点上 RBPNN 明显要比 RBFNN 优越。

对于 RBPNN 的权重训练,有很多方法可供选择,如矩阵伪逆算法、ROLSA^[5]、RLS^[4] 等等. 其中前 2 种算法能够利用成熟的矩阵算法,权重学习无需大规模的迭代计算过程,具有较好的收敛精度等优点,但也存在不足,如在计算过程中需要求取矩阵的逆、计算量较大等等. 而 RLS 算法利用成熟的自适应滤波(卡尔曼滤波器)原理来获得网络联接权值的递推迭代,该算法的基本原理为:在每一时刻的迭代是由输入训练样本形成的协方差矩阵,递推求得权值的精确解(误差梯度为 0 时的解). 所以,该算法具有迭代的方向性强、收敛速度快等优点. 因此文章使用 RLS 算法来训练 RBPNN,期望能够充分利用 RBPNN 权重的固有特性,在训练速度、收敛精度以及对大规模样本集的快速训练方面,发挥算法的优

势.

类似于其他网络训练方法, RLS 算法也是基于误差代价函数为优化目标函数的训练算法,与传统的误差代价函数的区别在于, RLS 每一步迭代的误差代价是基于前面迭代误差的加权累加,因此,又称为加权误差代价函数,即

$$J(k) = \frac{1}{2} \sum_{t=1}^k \sum_{j=1}^M d_j^2(t) = \frac{1}{2} \sum_{t=1}^k \sum_{j=1}^M (d_j(t) - y_j(t))^2. \quad (15)$$

式中: $d_j(t)$ 是对应于输入样本 x_t 网络输出层的第 j 个节点的误差值; $d_j(t)$ 与 $y_j(t)$ 分别是对应于输入样本 x_t 网络输出层第 j 个节点的期望信号和实际 outputs; 是加权遗忘因子,其作用是淡化过去样本对当前估值参数的影响,尽可能反映当前输入样本的特性,其取值为 $0 < \lambda < 1$. 一般来说, λ 的取值越大,则反映当前估值参数对过去样本的依赖性越大. 根据文献[3],其对应的权重更新过程为

$$w_j(k) = w_j(k-1) + g(k) [d_j(k) - h(k) w_j(k-1)], \quad (16)$$

$$g(k) = \frac{P(k-1) h^T(k)}{1 + h(k) P(k-1) h^T(k)}, \quad (17)$$

$$P(k) = [P(k-1) - g(k) h(k) P(k-1)] / \lambda. \quad (18)$$

详细的训练算法描述如下

- 1) 随机初始化权重矩阵 $w_j(0) = R^{M \times 1} (1 \leq j \leq M)$ 、逆相关矩阵 $P(0) = R^{M \times M}$, 设定核函数控制参数、停止条件及迭代数 $k = 1$;
- 2) 选择隐中心矢量;
- 3) 对于每个输入样本 x_t , 计算第 2 隐层的 outputs 向量 $h(t)$, $1 \leq t \leq N$;
- 4) 分别计算 $g(k)$ 及逆相关矩阵 $P(k)$;
- 5) 计算网络输出层第 k 步迭代时实际误差向量 $j(k) = d_j(k) - h(k) w_j(k-1)$, $1 \leq j \leq M$;
- 6) 更新权重矩阵 $w_j(k) = w_j(k-1) + g(k) j(k)$, $1 \leq j \leq M$;
- 7) 判断是否满足停止条件,如果满足,则停止迭代;否则, $k = k + 1$, 转至 4)。

3 实验结果

为验证所提出的算法的有效性,对常见的 40 类图形符号的图像进行分类与识别. 实验用的样本图像库来源于本实验室,均是从扫描的 0~4 号工程图纸中分割提取的独立图像,基本上可以覆盖各种常见模式的信息,共 1 320 幅图像. 为了让训练样本能在整个特征空间总体分布具有代表性,每种图形符号最少有 20 个样本数据. 每一次实验中,对于每一类图形符号,随机选取 10 个图像作为训练数据,剩余的图像作为测试数据. 所有算法都是在

MSVC++6.0开发环境中实现.

首先,选择所有的 400 个训练图像的特征矢量作为 RBPNN 第 1 隐层的中心矢量,第 2 隐层的节点数设置为 40,第 1 隐层的控制参数按照经验公式设置为 0.2,训练算法采用文章 2.2 节的 RLS 方法,加权遗忘因子.该实验重复进行 20 次,平均的正确识别率为 92.2%.

接着,为比较 RBPNN 的性能,选择相同的训练数据、测试数据和训练算法,使用具有单隐层和双隐层的多层感知机网络(MLPN),以及径向基函数网络(RBFNN)进行比较.其中,具有单隐层的 MLPN 的隐层节点数设置为 50;具有双隐层的 MLPN 的第 1 隐层节点数设置为 50,第 2 隐层节点数设置为 50;径向基函数网络第 1 隐层的设置(节点数与控制参数)完全与 RBPNN 相同. MLPN 与 RBFNN 的训练算法均采用 RLS 方法.同样该实验重复进行 20 次,其平均的识别结果如表 1 所示.从表中可以看出,RBPNN 的分类性能要好于 BPNN、MLPN 和 RBFNN.另外,对于一幅图像的识别,RBPNN 花费的平均训练时间和识别时间要低于其他分类器.

表 1 分类器性能比较

Table 1 Classification performance comparison between the RBPNN and other classifiers

| 分类器 | 平均正确 识别率/% | CPU 训练 | 时间/s |
|-----------|---------------|-----------|-------|
| RBPNN | 92.2 | 0.127 | 0.006 |
| RBFNN | 90.7 | 0.193 | 0.007 |
| MLPN(单隐层) | 88.8 | 0.323 | 0.008 |
| MLPN(双隐层) | 87.3 | 0.454 | 0.010 |

4 结束语

文章基于径向基概率神经网络(RBPNN)和 Hu 不变矩特征,提出了一种扫描工程图纸图形符号识别的方法.实验结果表明,对于 40 类常见的图形符号的识别,径向基概率神经网络是一种新型、高效的分类器,平均正确识别率可达到 92%,在识别性能、速度等方面要优于多层感知机网络和径向基函数网络等传统的神经网络模型,非常适合于工程图纸的图形符号识别.

参考文献:

[1]董玉德,赵 韩,王 平,等.工程图纸识别与理解的研究现状分析[J]. 合肥工业大学学报(自然科学版),2005,28(1):29-33.
DONG Yude, ZHAO Han, WANG Ping, et al. Analysis of research status of engineering drawings recognition and interpretation [J]. Journal of Hefei University of Technology, 2005, 28(1):29-33.
[2]黄德双. 神经网络模式识别系统理论[M]. 北京:电子工

业出版社,1996.

HUANG Deshuang. Systematic theory of neural networks for pattern recognition [M]. Beijing: Publishing House of Electronic Industry of China, 1996.

[3] HUANG Deshuang. Radial basis probabilistic neural networks: model and application [J]. International Journal of Pattern Recognition and Artificial Intelligence, 1999, 13(7): 1083-1101.
[4]赵温波,杨鹭怡,王立明.径向基概率神经网络的混合结构优化算法[J]. 系统仿真学报,2004,16(10):2175-2180.
ZHAO Wenbo, YANG Luyi, WANG Liming. The hybrid structure optimization algorithms of radial basis probabilistic neural networks[J]. Journal of System Simulation, 2004, 16(10): 2175-2180.
[5] HUANG Deshuang, ZHAO Wenbo. Determining the centers of radial basis probabilistic neural networks by recursive orthogonal least square algorithms [J]. Applied Mathematics and Computation, 2005, 162: 461-473.
[6]梁德本,叶玉驹.机械制图手册3版[M].北京:机械工业出版社,2005.
LIANG Deben, YE Yuju. Manual of mechanical drawing:3rd ed[M]. Beijing: Publishing House of Mechanical Industry of China, 2005.
[7]冈萨雷斯.数字图像处理:第2版[M].阮秋琦,阮宇智,等译.北京:电子工业出版社,2003.
GONZALEZ C. Digital image processing:2nd ed [M]. Beijing: Publishing House of Electronic Industry of China, 2003.
[8]CHEN C C. Improved moment invariants for shape discrimination [J]. Pattern Recognition, 1993, 26(5):683-686.

作者简介:



翟传敏,女,1977年生,讲师,2002年在合肥工业大学获工学硕士学位.主要从事模式识别、机械CAD相关研究工作.
E-mail:cmzhai@hfuu.edu.cn



杜吉祥,男,1977年生,博士后,2005年在中国科学技术大学获工学博士学位.主要研究方向为模式识别与图像处理等.已在国际会议和国际杂志上发表论文10余篇,大部分被SCI、EI收录.



黄 飞,男,1978年生,讲师,2002年在合肥工业大学获工学硕士学位.主要从事CAD/CAM、测控技术及仪器科学技术的研究.