

基于 Gabor 变换和 LMBP 神经网络的车牌汉字字符识别

赖树雨, 李向明, 段 涛

(武汉理工大学 自动化学院, 湖北 武汉 430070)

摘要: 字符识别是汽车牌照自动识别系统中的关键环节, 汉字字符识别是其中的难点。提出用 Gabor 滤波器对灰度汉字图像抽取横、竖、撇、捺的 4 幅能量特征图像的方法, 同时对 Gabor 滤波器组输出值进行非线性变换, 使其适应于不同亮度和低质量灰度车牌字符图像的识别, 最终采用网络法提取 4 幅能量特征图像的特征, 用改进的 BP 神经网络作为车牌汉字字符的识别器, 提高车牌识别率。

关键词: 车牌汉字识别; Gabor 变换; 非线性变换; LMBP 神经网络分类器

中图分类号: TP311

文献标识码: A

文章编号: 1673-9833(2008)05-0094-03

Recognition of License Plate Chinese Character Based on Gabor Filters and BP Neural Networks

Lai Shuyu, Li Xiangming, Duan Tao

(School of Automation, Wuhan University of Technology, Wuhan 430070, China)

Abstract: Character recognition is the key of recognition systems for license plate, and Chinese character recognition is the most difficult point. A kind of method that the gray character image is transformed into horizontal, vertical, charge-offs, Na energy feature image by Gabor filters is proposed. Non-linear transformation is designed to regulate the outputs of Gabor filters adaptively which is used to improve the performances for low quality images. Finally we rely on network method for feature extraction and advanced BP neural network as the classification to enhance the rate of the recognition of license plate Chinese character.

Key words: recognition of license plate Chinese character; Gabor Filters; non-linear transformation; BP neural network

车牌字符识别中的汉字识别是车牌识别的一个瓶颈。本文采用已经分割好的汉字灰度图像, 直接运用 Gabor 小波变换对其进行特征提取来提高车牌识别率。在模式分类器的选择上, 由于标准 BP 算法基于最速下降法, 具有最速下降法的固有缺点(如易陷入局部极小, 收敛速度慢和引起振荡效应等^[1]), 所以本文采用了改进的 BP 算法——LMBP 算法。LM (LevenberMarquart) 算法在最优化理论中是一种非常有效的优化设计方法, 从收敛速度和收敛性来看, 它是牛顿法和最速下

降法的一种折中算法。

1 Gabor 变换原理

Gabor 滤波器能最好的兼顾信号在时域和频域中的分辨能力, 充分描述图象的纹理信息, 它的小波特性的说明了 Gabor 滤波结果是描述图象局部灰度分布的有力工具。因此, 图象的纹理信息可用 Gabor 滤波来抽取。Gabor 滤波结果可以描述不同方向上灰度的分布信息, 并可以容忍图象有一定的平移、旋转、亮度

收稿日期: 2008-07-22

作者简介: 赖树雨(1983-), 女, 江西赣州人, 武汉理工大学硕士研究生, 主要研究方向为智能控制与智能自动化。

不均、尺度变化等情况。

二维 Gabor 滤波器的冲激响应函数的形式^[2]为:

$$g(x, y, \sigma, \theta, l) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2}\left[\frac{R_1^2 + R_2^2}{\sigma}\right]\right\} * \exp\left(i * \frac{2\pi R_1}{l}\right),$$

其中: $R_1 = x \cos \theta + y \sin \theta$; $R_2 = -x \cos \theta - y \sin \theta$ 。

把上式展开可得 Gabor 函数的实部和虚部分别为:

$$\begin{cases} G_R(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \cos \frac{2\pi(x \cos \theta - y \sin \theta)}{l}; \\ G_I(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \sin \frac{2\pi(x \cos \theta + y \sin \theta)}{l}. \end{cases} \quad (1)$$

从式 (1) 可知波长参数 l 及方向参数 θ 的选取在 Gabor 特征提取中很重要。

根据如下原则设置 Gabor 滤波器的参数: 1) 字符笔画的平均宽度为 W 。Gabor 滤波器组对宽度为 $\sigma/2$ 的中心线条最敏感, 所以有 $W = \sigma/2$, 即 $\sigma = 2W$; 2) 由于汉字主要由横撇竖捺 4 种笔画构成, 投影在空间域上相邻 2 种笔画之间恰好相差 45° , 所以 $\theta = [0^\circ, 45^\circ, 90^\circ, 135^\circ]$; 3) Gabor 滤波器组在空间上的抽样间隔应满足 $D \leq \sigma/2$; 4) 频率的范围根据所要识别的对象尺寸来选定, 以便所有的局部信息能够通过小波表示出来。

若有 1 组归一化为 64×64 大小的车牌汉字灰度图像, 参数设置为: $\sigma = 4.5$; $D = 2$; 参数 $l \in (3.2, 3.9)$; $\theta = [0^\circ, 45^\circ, 90^\circ, 135^\circ]$ 。

l 和 θ 不同的取值可以得到不同的 Gabor 小波矩阵, 再将灰度均衡化和归一化好的车牌汉字灰度图像 $F(x, y)$ 与小波矩阵的实部和虚部进行卷积, 得到字符图像的实特征和虚特征。本文采用 Gabor 滤波器输出能量计算特征, 能量特征可用下列函数表达:

$$Q(x, y) = \left[(G_R(x, y) * F(x, y))^2 + (G_I(x, y) * F(x, y))^2 \right]^{1/2}。$$

表 1 为“鄂”字通过 Gabor 滤波器得到的能量特征。

表 1 “鄂”字 Gabor 能量特征

Table 1 Gabor energy features of character “鄂”

波长 参数 l	方向参数 θ			
	0	$\pi/4$	$3\pi/4$	$\pi/2$
10				
5.0				
3.9				
3.3				

从表 1 可知, 当 l 取值较大时, 获得的是图像的全局信息; 当 l 取值较小时, 获得的是图像的细节信息。通过改变 l 和 θ 值, 可以获得不同尺度和不同方向上的图像信息。

2 对滤波器组输出进行非线性变换

在现实中由于存在光照不同, 图像摄取过程中图像的亮度、对比度不同, 和同一个字符图像中, 各笔画可能因不同均匀光照而亮度不同, 且图像中的干扰会产生一定的干扰输出^[3], 因此, 为了抑制亮度变化和干扰输出, 采用 sigmoid 函数 $f(t) = (e^{2t} - 1) / (e^{2t} + 1)$ 对 Gabor 滤波器组输出的幅值进行非线性变换得到 $G'(x, y)$, 再把它与灰度图像 $F(x, y)$ 卷积, 得到所求的能量特征,

$$G'_R(x, y) = f(G_R(x, y)) = \frac{e^{2G_R(x, y)} - 1}{e^{2G_R(x, y)} + 1},$$

$$G'_I(x, y) = f(G_I(x, y)) = \frac{e^{2G_I(x, y)} - 1}{e^{2G_I(x, y)} + 1}。$$

图 1 为“鄂”字经过非线性变换的能量特征图。



图 1 “鄂”字的非线性变换的能量特征图

Fig. 1 The Gabor energy features picture of “鄂” based on non-linear transformation

从图 1 知, “鄂”字经过非线性变换的能量特征更明显, 非线性变换使 Gabor 滤波器有较强的抗噪能力。

3 特征提取

经过非线性变换后的 Gabor 滤波器组的输出, 可以将 4 个 $N \times N$ (64×64) 的输出平面分割成为 $M \times M$ (8×8) 的区域, 在每个区域 $R(X, Y)$ 中, 分别求出模的和, 即: $S_X(x, y) = \sum_{(m, n) \in R(X, Y)} Q(m, n)$, 这样, 对于每个方向的 Gabor 滤波器输出可得到 M^2 维特征矢量, 再将 4 个方向上的特征矢量合并成一个维数为 $4M^2$ 的矢量。

4 LMBP 神经网络原理

图像识别分类中应用最广泛的是 BP 神经网络, 而在 BP 网络中, 又以单隐层 BP 网络为最多。LM 算法克服了标准 BP 算法的缺点, 是一种较为有效的优化设计方法。从收敛速度和收敛性来看, 它是牛顿法和最速下降法的一种折中算法。

设 $W(k)$ 表示第 k 次迭代的网络权值向量, 维数为 M , 新的权值向量 $W(k+1)$ 可根据下面的规则求得:

$$W(k+1) = W(k) + \Delta W(k)。$$

设误差指标函数为:

$$E(W) = \frac{1}{2} \sum_{i=1}^N e_i^2(W) = \frac{1}{2} \sum_{i=1}^N (t_i - o_i),$$

式中: t_i 为期望输出的误差;

o_i 为实际输出的误差;

N 为输出向量维数。

$e(W) = [e_1(W), e_2(W), \dots, e_n(W)]^T$, L-M 算法:

$$\Delta W = [J^T(W)J(W) + \mu I]^{-1} J^T(W)e(W),$$

式中: 比例系数 μ 大于 0, 为常数; I 是单位矩阵; $J(W)$ 为 Jacobian 矩阵。

在许多情况下, 训练 BP 网络采用 L-M 算法可以获得比其它任何一种算法更小的均方差。

5 算法 MATLAB 实现

算法采用 MATLAB6.0 开发, 运算实现对每幅图像的 4 次 Gabor 变换、输出幅值的非线性变换和 256 维特征向量提取, 并利用样本特征训练 LMBP 分类器, 最后测试分类器的正确识别率。该程序的功能实现由 5 部分组成: 1) 双线性插值算法处理车牌汉字图像归一化, 大小为 64×64 , 可由函数 `imresize(I, [64,64], 'bilinear')` 实现; 2) 特征提取: 对样本的每幅图像进行 4 次 Gabor 变换, 并对 Gabor 滤波器的输出进行非线性变换, 再提取出每个车牌汉字的 256 维特征, 最终对所有样本特征矢量结合成 256×100 维的矩阵并保存; 3) 训练 BP 分类器: 读出训练样本特征和目标矢量, 训练函数选用 `trainlm`, BP 网络的组建函数为 `net=newff(pr, [30 6], {'logsig' 'purelin'}, 'trainlm', 'learnngdm')`; 4) 优化 BP 网络性能: 设置合适的隐层数目和隐层结点数, 使训练误差和学习时间达到所要求; 5) 识别车牌汉字图像和检测网络的识别精度。

5.1 LMBP 神经网络的结构设计

BP 网络的识别性能主要取决于网络结构和训练算法, 而网络结构的确定主要是指它的隐含层数目和相对应的结点数, 训练算法已经有最速下降法, 加入动量因子、学习率可变、变梯度算法、拟牛顿算法和 LM 算法^[5]。按照标准 BP 神经网络设计方法, 当汉字图像归一化处理为 64×64 字符图像, 得到 256 维的车牌汉字字符的特征矢量。所以, 网络的输入节点数为 256 个, 隐层传输函数为 sigmoid 函数, 输出层函数为 purelin 函数, 输出模式采用二进制编码, 车牌汉字字数将近 30, 所以输出结点为 5。

5.2 隐层结点对网络训练的影响

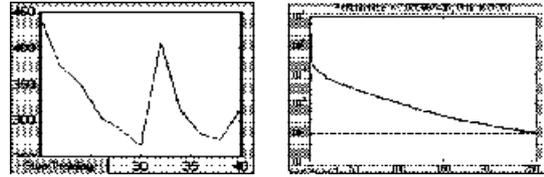
隐层结点数将是影响网络识别性能的主要原因, 欠结点数会影响网络正确识别车牌汉字图像, 过多的结点数会使网络复杂化, 使网络训练需要更多的学习时间。图 2 表示训练步数 echoes 与隐层结点数的关系。

由图 2 可知, 隐层的结点 $n=30$ 比较合适, 此时的误差为 0.000 997 435。

5.3 测试结果

样本选取针对宋、黑、仿宋和楷体字体, 以 18、20、24、28 号字用画图工具按 Bmp 格式形成 300 幅带干扰

的车牌汉字字符图片。整理出 30 组样本, 随机取 20 组来训练 BP 网络, 另外 10 组用来检测 BP 网络识别率。试验中发现, 只要测试字符样本格式包含在训练字符样本格式中, 识别率可达 96%。因此, 只要车牌汉字字符训练样本库足够大, 基于 Gabor 变换和 LMBP 神经网络的车牌汉字字符识别系统是可取的方法。



a) 训练步数与隐层结点数关系 b) $n=30$ 时训练误差曲线

图 2 隐层结点数与网络训练性能的关系

Fig. 2 The relationship between the number of hidden node and the performance of net training

6 结语

本文研究基于 Gabor 变换和 LMBP 神经网络的车牌汉字字符识别方法, 实验表明: Gabor 小波变换可以抽取字符图像局部空间中最重要和稳定的笔画方向信息; 对 Gabor 滤波器组输出值进行非线性变换, 对字符的不同亮度、大小、倾斜等干扰因素有较强抗噪能力。而神经网络具有很强的泛化能力, 这 2 种方法结合可有效提高车牌汉字识别率。

参考文献:

- [1] 王建梅, 覃文忠. 基于 L-M 算法的 BP 神经网络分类器[J]. 武汉大学学报: 信息科学版, 2005, 30(10): 928-930.
- [2] 曾姝彦, 张广军, 李秀智. 基于 Gabor 滤波器的图像目标识别方法[J]. 北京航空航天大学学报, 2006, 30(8): 954-957.
- [3] 王学文, 丁晓青, 刘长松. 基于 GABOR 变换的高鲁棒汉字识别新方法[J]. 电子学报, 2002, 30(9): 1317-1321.
- [4] 周开利, 康耀红. 神经网络模型及其 MATLAB 仿真程序设计[M]. 北京: 清华大学出版社, 2006: 89-100.
- [5] 吕俊, 张兴华. 几种快速 BP 算法的比较研究[J]. 现代电子技术, 2003, 167(24): 96-99.
- [6] Suthaj, Ramaraj. Neural Network Based-on Offline Tamil Handwritten Character Recognition System[C]// International Conference on Computational Intelligence and Multimedia Applications. [s.l.]: [s.n.], 2007: 940-947.
- [7] 王家文, 曹宇. 图形图像处理[M]. 北京: 国防工业出版社, 2001.

(责任编辑: 张亦静)