

利用 Radon 变换实现手写数字识别的新方法

杨志华

YANG Zhi-hua

广东商学院 信息学院, 广州 510320

Information Science School, Guangdong University of Business Studies, Guangzhou 510320, China

YANG Zhi-hua. Novel handwritten numeral recognition based on Radon transform. Computer Engineering and Applications, 2008, 44(30): 13-15.

Abstract: A novel rotation-invariant handwritten numeral recognition is presented based on Radon transform and BP neural network. A Radon transform is firstly implemented to a handwritten numeral image to extract the rotation-invariant features and then, the BP network classifier is employed to classify it. Experiments show encouraging results.

Key words: handwritten numerals recognition; feature extraction; Radon transform; neural network

摘要: 提出了一种新的手写数字识别方法, 利用传统的 Radon 变换, 找到了一种新的旋转不变特征, 最后采用 BP 神经网络分类器进行分类。实验表明, 该方法不仅具有 93.89% 的高识别率, 而且对字符旋转具有很好的鲁棒性。

关键词: 手写数字识别; 特征提取; Radon 变换; 神经网络

DOI: 10.3778/j.issn.1002-8331.2008.30.004 **文章编号:** 1002-8331(2008)30-0013-03 **文献标识码:** A **中图分类号:** TP391

1 引言

手写字体的识别一直以来都是模式识别研究领域的一个热点问题。近 20 多年来, OCR (Optical Character Recognition) 获得了迅猛的发展。目前, 印刷体字符识别技术已基本成熟, 识别率已经达到了商业应用的要求。联机手写字体的识别也被成功地应用于手机等识别率要求并不太高的产品中。而脱机手写字体识别, 由于其本质的困难, 一直没有得到很好的解决。手写数字是手写字体识别的一个重要内容。高性能的手写数字识别技术不仅可以直接应用于邮政编码的自动识别、银行支票自动处理等系统中, 也可以为其它手写字体识别技术提供直接的借鉴。近几十年来, 国内外研究者陆续提出了一些识别方法, 这些方法大致可以分为两类, 即基于结构特征的方法^[1-7]和基于统计特征的方法^[8-9]。两类方法各有优劣, 迄今为止, 还没有一种被公认的好的方法。

本文提出了一种新的手写数字识别方法, 利用传统的 Radon 变换, 找到了一种新的旋转不变特征, 最后采用 BP 神经网络分类器进行分类。实验表明, 本文方法不仅具有 93.89% 的高识别率, 而且对字符旋转具有很好的鲁棒性。

2 Radon 变换

Radon 变换是 J.Radon 于 1971 年提出的, 但是, 在快速 Fourier 提出之前, 它几乎没有引起人们的兴趣。现在, Radon 变换已成为医学成像和遥感成像的主要工具^[10]。

Radon 变换实际上是一个投影变换。如图 1 所示, 设 $f(x, y)$ 为二维平面 (x, y) 上的一个函数, 则其 Radon 变换可以写成

$$P_{\alpha}(u) = \int_{-\infty}^{\infty} f(x, y) dv \quad (1)$$

式中, α 为 u 轴与 x 轴的夹角, 如图 1 所示, x, y 与 u, v 的关系为

$$\begin{cases} x = u \cos \alpha - v \sin \alpha \\ y = u \sin \alpha + v \cos \alpha \end{cases} \quad (2)$$

$f(x, y)$ 的二维 Fourier 变换为

$$F(\Omega_1, \Omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j(\Omega_1 x + \Omega_2 y)} dx dy \quad (3)$$

那么, 平面 (Ω_1, Ω_2) 中过原点、倾角为 α 的直线上的函数 $F(\Omega_1, \Omega_2)$ 的值为

$$F_{\alpha}(\Omega_1, \Omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j\Omega(x \cos \alpha + y \sin \alpha)} dx dy \quad (4)$$

式中, $\Omega_1 = \Omega \cos \alpha$, $\Omega_2 = \Omega \sin \alpha$ 。

将积分的平面由 (x, y) 转换到 (u, v) , 则式(4)为

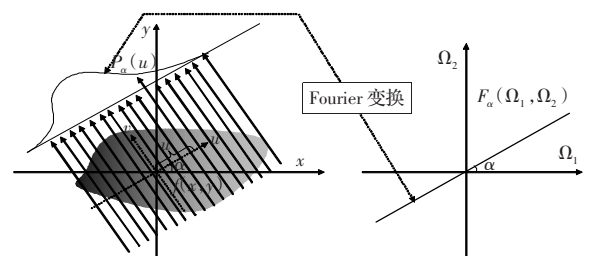


图1 Radon 变换示意图

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60475042, No.10631080)。

作者简介: 杨志华(1964-), 男, 博士, 副教授, 主要研究领域为信号分析、模式识别和图像处理。

收稿日期: 2008-06-17 修回日期: 2008-07-28

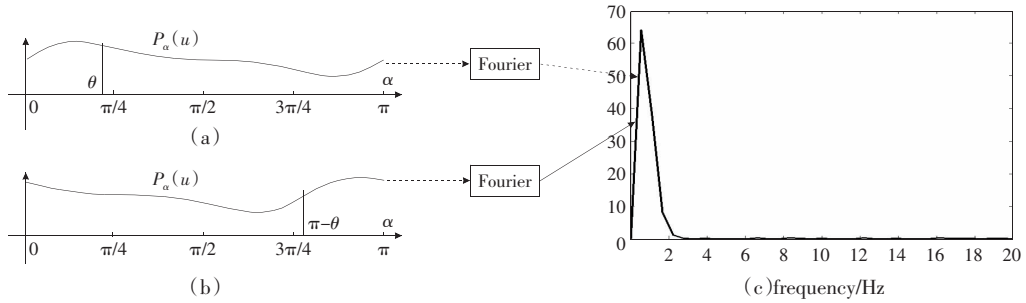


图2 给定 u 的 Radon 变换及其 Fourier 变换

$$F_{\alpha}(\Omega_1, \Omega_2) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(u \cos \alpha - v \sin \alpha, u \sin \alpha + v \cos \alpha) dv \right] e^{-j\Omega u} du = \int_{-\infty}^{\infty} P_{\alpha}(u) e^{-j\Omega u} du \quad (5)$$

式(5)说明, $F_{\alpha}(\Omega_1, \Omega_2)$ 可由原函数倾角为 α 的 Radon 变换 $P_{\alpha}(u)$ 通过 Fourier 变换求得。如果将 α 从 0° 变到 180° , 就可得到整个平面的 $F(\Omega_1, \Omega_2)$ 值, 而对 $F(\Omega_1, \Omega_2)$ 做 Fourier 反变换便可重构原函数 $f(x, y)$ 。

3 特征提取与分类方法

3.1 数字图像的 Radon 变换

对于一幅数字字符图像 A , 由式(1)得到的 Radon 变换事实上是一个关于 α 和 u 的二元函数, 因此, $P_{\alpha}(u)$ 又可写成二元函数的形式 $P(\alpha, u)$ 。对于给定的 u , $P(\alpha, u)$ 是一个关于 α 的以 π 为周期的一元周期函数。假设对于某个给定的 u , $P(\alpha, u)$ 的函数曲线如图 2(a) 所示, 那么, 当图像 A 旋转 θ 后, $P(\alpha, u)$ 的函数曲线将如图 2(b) 所示, 即图像旋转 θ 以后, $P(\alpha, u)$ 的函数曲线只是平移 θ 个单位, 它们的 Fourier 变换保持不变, 如图 2(c)。

这说明, 对于给定的 u , $P(\alpha, u)$ 的 Fourier 变换是关于图像 A 旋转不变的。下面, 将利用这一特点来实现旋转不变的手写数字图像识别。

3.2 特征提取

对于给定的数字字符图像 A , 在 $0 \sim \pi$ 选取等距的 N 个角度做 Radon 变换, 得到二维矩阵 $P(\alpha, u)$, 然后对每一个确定的 u , 计算 $P(\alpha, u)$ 的 Fourier 变换, 记为 $FP_u(\omega)$ 。一般地, $FP_u(\omega)$ 的主要能量集中在低频段, 也就是说 $FP_u(\omega)$ 的前 K 个频率含量包含了图像 A 的大部分信息。例如, 图 3(a) 是一幅大小为 24×24 的数字字符图像。取角度 $\alpha = 0^{\circ}, 11.25^{\circ}, 22.5^{\circ}, \dots, 168.75^{\circ}$ 做 Radon 变换, 然后直接做 Radon 反变换, 得到的图像如图 3(b) 所示。图 3(c) 是对每一个 u , 计算 $P(\alpha, u)$ 的 Fourier 变换 $FP_u(\omega)$, 然后取 $FP_u(\omega)$ 的前三项做 Fourier 反变换, 再做 Radon 反变换

后还原的图像。可以看到, 它基本保留了原图像的主要信息。因此, 可以以它们作为识别数字字符的特征, 而且, 这样的特征是旋转不变的。

最后, 得到如下的特征向量

$$V = (FP_{u_1}(\omega_1), \dots, FP_{u_1}(\omega_K), \dots, FP_{u_M}(\omega_1), \dots, FP_{u_M}(\omega_K)) = (v_1, v_2, \dots, v_{K \times M}) \quad (6)$$

式中, K 表示对每一个 $FP_u(\omega)$ 所取的低频分量的个数, K 越大, 特征向量包含的信息越丰富, 但是特征维数也越大; M 是在 (u, v) 平面上图像 A 在 u 轴上投影的最大 u 值, 它取决于图像 A 本身的大小, 设 A 为 $m \times n$ 的图像, 则 $M = \sqrt{m^2 + n^2}$ 。

3.3 分类方法

分类器的设计是模式识别中的一个关键问题, 除了经典的基于距离的分类方法和统计决策方法外, 近年来, 神经网络分类器获得了人们足够的重视和广泛的研究。其良好的非线性分类能力, 使得其分类效能大大提高^[11]。由于本文的研究重点不是分类器设计, 因此, 本文采用经典的 BP 神经网络分类器来进行分类。关于 BP 神经网络分类器的详细介绍请参阅文献^[11]。

4 实验结果与分析

为了测试本文方法的有效性, 采用 CENPAMI (Concordia University, CAN) 数字字符数据库进行训练和测试。该数据库分 3 个数据集, 分别为 Set A, Set B 和 Set T, 各有 0~9 的样本 200 个。实验中, 以 Set A 为训练集, Set T 为测试集。对于每一幅手写数字图像, 首先将其规范化为 24×24 的图像。考虑 Radon 变换的计算量, 只取角度 $\alpha = 0^{\circ}, 11.25^{\circ}, 22.5^{\circ}, \dots, 168.75^{\circ}$ 做 Radon 变换 $P(\alpha, u)$, 这样, 虽然会丢失许多信息, 但是, 实验表明, 丢失的信息对于数字字符的结构影响并不大, 如图 3(b) 所示。对每一个 u , 计算 $P(\alpha, u)$ 的 Fourier 变换 $FP_u(\omega)$, 然后取 $FP_u(\omega)$ 的前三项作为特征, 形成特征向量。由于图像大小为 24×24 , 其 u 的最大取值为 34, 对每一个 u , 取 $FP_u(\omega)$ 的前三项, 因此, 总的特征维数为 $34 \times 3 = 102$ 维。最后, 用 BP 神经网络分类器进行分类。

应该注意的是, 最终的分类结果很大程度上依赖于 BP 神经网络的隐层神经元的个数和训练样本的个数。隐层神经元个数太多和太少, 都得不到好的分类结果。训练样本个数太少, 会导致神经网络训练不够, 产生所谓欠学习, 而个数太多, 则可能导致所谓过学习。因此, 适当的训练次数和隐层神经元个数是



图3(a) 数字字符图像 图3(b) 经过 Radon 逆变换还原的图像 图3(c) 取 Fourier 变换低频成分并经 Radon 逆变换后还原的图像

获得高识别率的基础。图 4 是分别取隐层神经元个数为 30,35,40,45,训练样本个数分别为 10,20,30,40,50,60,70,80 时的识别率,可以看到,当训练样本个数为 50,隐层神经元个数为 40 或 45 时,可以获得最高的识别率,此时的识别率为 93.89%。

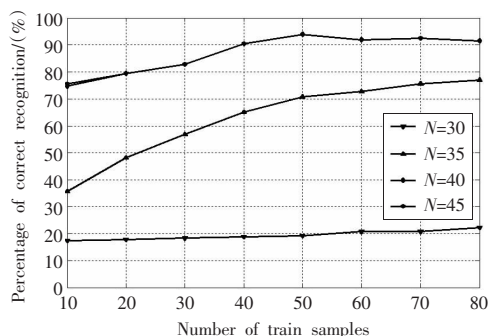


图 4 不同隐层神经元个数和训练样本个数的识别率

取隐层神经元个数为 45,训练样本个数为 50,对神经网络进行训练。用 CENPAMI 数字字符数据库的 Set T 进行测试,实验结果如表 1。表中每类样本的测试样本个数均为 200 个。值得注意的是,由于本文的方法具有旋转不变性,而在旋转不变意义下字符“6”和“9”是没有区别的,也即本文方法不能区分这两个字符,所以,表 1 只给出了字符 0~8 的测试结果。

表 1 测试结果

类别	0	1	2	3	4	5	6	7	8	总计
正确识别数	191	195	183	188	192	179	186	190	186	1 690
识别率/(%)	95.5	97.5	91.5	94.0	96.0	89.5	93.0	95.0	93.0	93.89

为测试本文方法对字符旋转的鲁棒性,在 CENPAMI 数据库的 Set T 中任意抽取 100 个字符,将字符旋转任意角度,每旋转一次得到一个测试样本,每个字符旋转 10 个角度,这样总共产生 1 000 个测试样本。按上述相同方法进行测试,测试结果如表 2。

表 2 字符旋转任意角度后的测试结果

类别	0	1	2	3	4	5	6	7	8	总计
抽取字符数	12	8	13	11	11	12	10	10	13	100
总字符数	120	80	130	110	110	120	100	100	130	1 000
正确识别数	114	76	114	104	105	106	94	93	121	927
识别率/(%)	95.0	95.0	87.8	94.5	95.5	88.3	94.0	93.0	93.1	92.7

从测试结果来看,本文方法确实对字符旋转有很好的鲁棒性,平均识别率仍达到 92.7%。

表 3 给出了现有文献报道的其它几种主要方法对同一数

表 3 与已有方法的比较

方法	文献[2]	文献[3]	文献[5]	文献[6]	文献[7]	文献[8]	本文
识别率/(%)	86.05	97.10	93.10	98.70	99.00	97.40	93.89

据库进行测试的结果。值得注意的是,尽管本文方法就识别率来说并不太高,但本文方法对字符的旋转不变性有很好的鲁棒性,这对于手写数字识别来说是非常重要的。

5 结论

手写字符识别一直是模式识别中一个重要而困难的问题,本文从一个新的角度给出了一种手写数字识别方法,为手写数字识别提供了一条全新的思路,是对手写字符识别的有益的探索。已有的许多方法是以数字字符的孔洞数、凹陷区、轮廓、投影等局部信息为特征的,而手写数字随意性大,经常出现连笔、断笔等现象,因此,在提取这些特征之前需要做大量的预处理,而这本身就是非常困难的。本文方法算法简单,避免了困难的预处理工作,更重要的是本文方法对字符旋转具有很好的鲁棒性。

参考文献:

- [1] Suen C Y, Bertoh M, Mori S. Automatic recognition of hand-printed characters—the state of the art[C]//Proc of the IEEE, 1980, 68(4): 469–487.
- [2] Nadal C, Legault R, Suen C Y. Complementary algorithms for the recognition of to tally unconstrained handwritten numerals[C]//Proc the 10th International Conference Pattern Recognition, Paris, 1990: 434–449.
- [3] Lee S W. Offline recognition of to tally unconstrained hand-written numerals using multi-layer cluster neural network[J]. IEEE Trans on Pattern Recognition and Machine Intelligence, 1996, 18(6): 648–652.
- [4] Cheng D, Yan H. Recognition of handwritten digits based on contour information[J]. Pattern Recognition, 1998, 31(3): 235–255.
- [5] 娄震, 胡钟山, 杨静宇, 等. 基于轮廓分段特征的手写体阿拉伯数字识别[J]. 计算机学报, 1999, 22(10): 1065–1072.
- [6] Teow L N, Loe K F. Robust vision-based features and classification schemes for off-line handwritten digit recognition[J]. Pattern Recognition, 2002, 35(11): 2355–2364.
- [7] 苗夺谦, 张红云, 李道国, 等. 基于主曲线的脱机手写数字识别[J]. 电子学报, 2005, 33(9): 1639–1643.
- [8] Frader P D, Khabou M A. Automatic feature generation for hand-written digit recognition[J]. IEEE Trans on Pattern Recognition and Machine Intelligence, 1996, 18(12): 1256–1261.
- [9] Lee S W. Offline recognition of to tally unconstrained hand-written numerals using multi-layer cluster neural network[J]. IEEE Trans on Pattern Recognition and Machine Intelligence, 1996, 18(6): 648–652.
- [10] 张贤达, 保铮. 非平稳信号分析与处理[M]. 北京: 国防工业出版社, 1998.
- [11] 阎平凡, 张长水. 神经网络与模拟进化计算[M]. 北京: 清华大学出版社, 2000.