

基于 2DPCA 的手写数字识别

王军平¹, 赵振华²

(1. 咸阳职业技术学院 电子信息系, 陕西 咸阳 712000; 2. 兰州理工大学 电气工程与信息工程学院, 甘肃 兰州 730050)

摘要: 手写字符由于书写风格和习惯的不同,造成字符模式的不稳定。针对这一问题,本文首先对字符图像进行图像预处理,统一字符笔画的粗细,改善局部特征,随后利用二维主分量分析法(2DPCA)直接对字符图像矩阵进行变换,抽取字符特征,建立字符的特征矩阵及重构模型,利用最邻近方法和重构误差法进行字符识别。最后通过美国国家邮政局 MINIST 手写数字库中进行识别实验,验证了算法的准确性和鲁棒性。

关键词: 特征矩阵; 重构; 最邻近法; 2DPCA

中图分类号: TP391

文献标识码: A

文章编号: 1674-6236(2012)21-0058-04

Handwritten digit recognition based on 2DPCA

WANG Jun-ping¹, ZHAO Zhen-hua²

(1. Department of Electronic and Information, Xianyang vocational & Technical College, Xianyang 712000, China;

2. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: To overcome the problem of handwritten numerals models instability caused by different writing styles, a robust method is presented in this paper. First, the numeral image was pre-processed to be of similar thickness and improved for its local features. Then 2-dimension principal component analysis (2DPCA) is directly used for character image to extract character features and establish character's feature matrixes and reconstruction models. Character recognition is conducted based on a nearest neighbor classifier and reconstructed models' error method respectively. Finally, the algorithm proposed in this paper is tested on all the characters in MINIST database and a comparison is made with one dimension principal component analysis, and the experimental results validate the accuracy and robustness accuracy of the proposed algorithm.

Key words: feature matrix; reconstruct; the nearest neighbour method; 2DPCA

手写数字识别研究有着广泛的应用背景和重要的理论意义^[1-3]。如今在邮政、财政、税务等工作中都需要进行手写数字的识别,识别结果的好坏直接影响到工作的效率。另外,由于数字识别的类别较小,有助于作深入分析及验证一些新的理论。例如支持向量机的提出,就是先在手写体数字识别领域里进行验证,然后推广到了其他的领域。

目前,对手写数字识别的研究依据特征的提取可分为两大类:基于字符统计规律和基于字符结构特征。基于统计规律的方法是利用字符样本库,找出 0 到 9 中每类字符空间分布的统计规律,构成分类器进行识别。基于字符结构特征的方法是分析字符笔画的构造如圈、端点、交叉点、轮廓等来构造分类器进行识别。两类方法各有优势,总体而言,统计方法能更好地描述一类模式的本质特征,对于与给定训练集差别不大的字符具有较高的识别率;基于字符结构特征的方法精确地描述了字符的细节特征,对书写结构较规范的字符有较高的识别率。

在字符的特征提取中,主元分析(PCA)是一种十分有效的方法。PCA 的思想是将高维样本空间的样本投影到某个低

维子空间,使得在该子空间中,投影样本的类间方差最大,类内方差最小。在字符识别中,PCA 方法通常是将样本库中的每个字符图像矩阵转换为一维向量,然后求出样本总体的协方差矩阵,计算出该矩阵的特征值及特征向量,根据特征值及对应的特征向量确定子空间的基向量。子空间的这些基向量又称为字符图像的特征图,每个字符图像都可以由特征图的不同加权和重构出来,其与原图的均方误差是没有选取的那些特征值之和。由于选取的特征向量是对应于特征值较大的那些向量,一般远少于特征向量总数,这样就实现了对原始样本空间的降维。

虽然 PCA 能有效地降低样本空间的维数,但在实现过程中 2D 的图像矩阵必须首先转换为 1D 的图像向量,其所产生的图像向量空间的维数就很高,例如在 MINIST 字符库中,每个字符图像为 28×28 像素,转换为一维图像向量是 1×784,这些图像向量所构成的空间的维数就为 784。这样就很难精确计算相应的协方差矩阵。为了克服这一困难,就产生了二维主元分析(2DPCA)方法。相对于 PCA,2DPCA 是基于二维图像矩阵而非一维向量,即不需将图像转换成一维的向量,取代 PCA 中样本总体协方差矩阵的是图像协方差矩阵,它是直

收稿日期:2012-07-21

稿件编号:201207132

作者简介:王军平(1966—),男,陕西乾县人,副教授。研究方向:图像处理、模式识别。

接从字符图像矩阵中构造出来的。这样得到的图像协方差矩阵较 PCA 要小很多。2DPCA 有两个明显的优点: 首先容易精确计算协方差矩阵, 其次确定相应的特征向量所耗的时间要少的多。

文中首先介绍 2DPCA 的原理及其算法的描述, 随后简要说明识别过程要用到的一些图像预处理, 第 3 节讲述两种分类器的构造, 第 4 节是实验及结果分析, 最后做出总结。

1 二维主元分析

1.1 原理及算法

用 X 表示一个归一化的 n 维列向量, A 为某个 $m \times n$ 的图像矩阵, 通过下面的线性变换将 A 投影到 X :

$$Y = AX \quad (1)$$

这样就得到了一个 m 维的向量 Y , 称之为图像 A 的投影特征向量。为了确定最佳的投影向量 X , 引入投影到 X 上的样本的整体散度来度量 X 的分辨能力。投影样本的总体散度可以用投影特征向量的协方差矩阵的迹来确定。由此得到下面的判据:

$$J(X) = \text{tr}(S_x) \quad (2)$$

其中 S_x 表示训练样本的投影特征向量所构成的协方差矩阵, $\text{tr}(S_x)$ 表示的迹。通过对 (2) 式的最大化来找出 X 的某个投影方向, 投影到该方向的样本总体散度最大。协方差矩阵 S_x 可表示为

$$S_x = E(Y - EY)(Y - EY)^T = E[AX - E(AX)][AX - E(AX)]^T = E[(A - EA)X][(A - EA)X]^T \quad (3)$$

所以有

$$\text{tr}(S_x) = X^T E[(A - EA)^T (A - EA)] X \quad (4)$$

现在定义下面的矩阵

$$G_i = E[(A - EA)^T (A - EA)]$$

称矩阵 G_i 为图像协方差(散度)矩阵。容易验证 G_i 为 $n \times n$ 非负定矩阵。 G_i 可以利用训练图像样本直接计算出来。假定共有 M 个训练样本, $A_j (j=1, 2, \dots, M)$ 表示第 j 个训练图像, 大小为 $m \times n$ 的矩阵, \bar{A} 表示所有训练图像的均值。则 G_i 为

$$G_i = \frac{1}{M} \sum_{j=1}^M (A_j - \bar{A})^T (A_j - \bar{A}) \quad (5)$$

则判据 (2) 可表示为

$$J(X) = X^T G_i X \quad (6)$$

其中 X 为归一化的列向量。使这个判据最大化的归一化向量 X 被称为最优投影轴。直观上讲, 它意味着在图像投影到 X 后, 所得的投影样本的总体散度是最大的。

最优投影轴 X_{opt} 是最大化 $J(X)$ 的归一化向量, 即 G_i 的特征向量中对应于最大特征值的那个特征向量。通常需要选取一组正交的投影轴, X_1, \dots, X_d , 来最大化 $J(X)$, 即

$$\begin{cases} \{X_1, \dots, X_d\} = \arg \max J(X) \\ X_i^T X_j = 0, i \neq j, i, j = 1, \dots, d \end{cases}$$

事实上, 最优的投影轴, X_1, \dots, X_d , 是 G_i 的前 d 个最大特征值所对应的正交特征向量。

1.2 特征提取

已经得到了 2DPCA 的最优投影向量 X_1, \dots, X_d , 就可以用这些向量进行特征提取。对图像样本 A , 定义:

$$Y_k = AX_k, \quad k=1, 2, \dots, d \quad (7)$$

这样得到一组特征向量 Y_1, \dots, Y_d , 称之为样本图像 A 的主成分(向量)。注意 2DPCA 的每个主成分是向量, 而 PCA 的主成分是标量。

用这些主成分向量构成一个 $m \times d$ 矩阵 $C = [Y_1, \dots, Y_d]$, 称其为图像样本 A 的特征矩阵或特征图。

1.3 基于 2DPCA 的图像重构

在 PCA 方法中, 重构图像是用主成分和特征向量(特征图)结合在一起完成的。2DPCA 可以用下面类似的方法实现对图像的重构。

设图像的协方差矩阵为 G_i , 其前 d 个最大的特征值所对应的特征向量为 X_1, \dots, X_d , 这些特征向量是正交的。将图像样本投影到这些向量轴上, 生成主成分向量, $Y_k = AX_k, (k=1, 2, \dots, d)$ 。设

$$V = [Y_1, \dots, Y_d], U = [X_1, \dots, X_d], \text{ 则有} \quad V = AU \quad (8)$$

因为 X_1, \dots, X_d 是正交的, 由 (8) 可以得到样本 A 的重构图像:

$$\bar{A} = VU^T = \sum_{k=1}^d Y_k X_k^T \quad (9)$$

设 $\bar{A}_k = Y_k X_k^T (k=1, \dots, d)$, 其和图像 A 的大小是一样的, 它表示了 A 的重构子图。可以将这 d 个子图相加来近似表示 A 。当主成分向量的数量 $d=n$ (n 为 G_i 的特征向量总数) 时, 则有 $\bar{A}=A$, 即用图像 A 的主成分向量完整地将其重构, 并且没有损失任何细节。反过来, 若 $d < n$, 则重构图像 \bar{A} 是 A 的一个近似表达。

2 字符识别方法

用 2DPCA 方法对字符图像进行变换后, 就可以利用图像的主成分向量来构造分类器进行字符的识别。作为研究, 文中采用了两种识别方法。第一种是最邻近法, 另一为重构误差法, 下面分别介绍。

2.1 最邻近法

设 $C_i = [Y_1^{(i)}, \dots, Y_d^{(i)}], C_j = [Y_1^{(j)}, \dots, Y_d^{(j)}]$ 为任意两个特征矩阵, 它们之间的距离定义为

$$d(C_i, C_j) = \sum_{k=1}^d \|Y_k^{(i)} - Y_k^{(j)}\|_2 \quad (10)$$

其中 $\|Y_k^{(i)} - Y_k^{(j)}\|_2$ 表示两个主成分向量 $Y_k^{(i)}$ 和 $Y_k^{(j)}$ 之间的欧氏距离。

现在假定训练样本为 C_1, C_2, \dots, C_N (N 为训练样本的总数), 并且每个样本都指定了类别 γ_k , 对某个测试样本 C , 若 $d(C, C_i) = \min d(C, C_j)$, 且 $C_i \in \gamma_k$, 则有 $C \in \gamma_k$ 。

2.2 重构 i 误差法

对于 0 到 9 之间每一类字符的训练集, 用 2DPCA 进行分

基于2DPCA的手写数字识别

作者: 王军平, 赵振华

作者单位: 王军平(咸阳职业技术学院电子信息系, 陕西咸阳, 712000), 赵振华(兰州理工大学电气工程与信息工程学院, 甘肃兰州, 730050)

刊名: 电子设计工程

英文刊名: Electronic Design Engineering

年, 卷(期): 2012, 20(21)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_dzsjgc201221023.aspx