

基于奇异值分解和径向基函数神经网络的人脸识别算法研究

何正风¹ 孙亚民²

(佛山科学技术学院基础教育系 佛山 528000)¹ (南京理工大学计算机科学与技术学院 南京 210094)²

摘要 提出一种基于奇异值分解和径向基函数神经网络的人脸特征提取与识别方法,来解决人脸识别中的高维、小样本问题。该方法采用奇异值分解、奇异值降维压缩、奇异值矢量标准化和奇异值矢量排序,最后得到用于识别的奇异值特征矢量。运用基于径向基函数神经网络分类器进行人脸分类识别。在 ORL 数据库上进行实验和数据分析表明,该方法无论是在分类的错误率上还是在学习的效率上都能表现出极好的性能。

关键词 RBF 神经网络,奇异值分解,特征提取,人脸识别

中图分类号 TP301.6 **文献标识码** A

Research on Face Recognition Algorithm Based on SVD and RBF Neural Network

HE Zheng-feng¹ SUN Ya-min²

(Department of Basic Education, Foshan University, Foshan 528000, China)¹

(School of Computer and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)²

Abstract To solve the problem of high dimension and small sample in face recognition, this paper proposed a facial feature extraction and recognition method based on singular value decomposition and radial basis function neural network. By singular value decomposition, dimension reduction and compression, vector standardization and vector sorting, we finally got the feature vectors of singular value used to identify. It used the neural network classifier based on radial basis function to classify and recognize face. The experiments and data analysis on ORL database show that, this method has good performance whether in the error rate of classification or in the learning efficiency.

Keywords Radial basis function neural network, Singular value decomposition, Feature extraction, Face recognition

提出了一种利用奇异值分解(SVD)来获取人脸特征的方式。作为一种代数特征,奇异值具有代数和几何不变性,并且对噪声不敏感。该方法利用人脸图像的奇异值作为特征,利用秩分解的方法进行降维以提取人脸图像的奇异值。径向基函数网络是一种性能良好的前馈型三层神经网络,具有全局逼近性质和最佳逼近性能,训练方法快速易行,RBF 函数还具有局部响应的生物合理性。在隐层参数固定的条件下,由线性最小二乘法计算隐层和输出层之间的连接权值,由梯度下降法调整隐层神经元的中心和宽度。这种混合学习算法,能使 RBF 网络逼近 Moody 准则下的最优结构。

1 奇异值分解和径向基函数网络理论分析

1.1 基于奇异值的特征提取

文献[1]论证了奇异值特征是识别图像的有效特征。我们可以从矩阵的奇异值分解及其性质出发,找到采用奇异值特征描述人脸图像信息特征的依据。

定理 1(奇异值 SVD 分解) 若矩阵 $A=R^{m \times n}$,则存在正交矩阵 $U=[u_1, \dots, u_m] \in R^{m \times m}$, $V=[v_1, \dots, v_n] \in R^{n \times n}$ 使得 $U^T A V = \text{diag}[\sigma_1, \dots, \sigma_p] = W$, $p = \min(m, n)$,即 $A = U W V^T$,则称该式为 A 的奇异值分解,其中 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$, $\sigma_i (i = 1, 2, \dots, p)$ 为 A 的奇异值,是 $A A^T$ 或 $A^T A$ 的特征值 λ_i 的平

方根,即 $\sigma_i = \sqrt{\lambda_i}$ 。

若矩阵 A 的 SVD 由定理 1 给出,则

$$\|A\|_F = \sqrt{\sum_{i=1}^p \sigma_i^2}, \|A\|_2 = \sigma_1 \quad (1)$$

表明矩阵的度量特征与矩阵的奇异值密切相关。由此可知,矩阵的 F 范数等于该矩阵所有奇异值的平方和,矩阵的 2-范数等于该矩阵的最大奇异值^[1]。

文献[1]证明了图像的奇异值特征具有以下特性。

定理 2(奇异值的稳定性) 假设 $A, B \in R^{m \times n}$, A, B 的奇异值分别有 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$, $\tau_1 \geq \tau_2 \geq \dots \geq \tau_p$, 其中

$$p = \min(m, n) \quad (2)$$

则 $|\sigma_i - \tau_i| \leq \|A - B\|_2$ 。

定理 2 意味着当矩阵 A 有微小扰动时,奇异值的变化不大于扰动矩阵的 2-范数。这就意味着图像的奇异值特征对于图像噪声、图像光照条件等不同引起的变换具有不敏感的特性^[1,2]。

定理 3(奇异值的比例不变性) 假设 $A, B \in R^{m \times n}$, A 的奇异值为 $\sigma_1, \sigma_2, \dots, \sigma_k$, 矩阵 $a \times A$ 的奇异值为 $\sigma_1^*, \sigma_2^*, \dots, \sigma_k^*$, 则

$$|a|(\sigma_1, \sigma_2, \dots, \sigma_k) = (\sigma_1^*, \sigma_2^*, \dots, \sigma_k^*) \quad (3)$$

定理 3 说明如果对矩阵 A 与 $a \times A$ 的奇异值进行标准化

本文受广东省自然科学基金(S2011020002719, 10152800001000016)资助。

何正风(1957-),男,讲师,主要研究方向为数学建模与仿真, E-mail: hzfg09@126.com。

处理,则可实现奇异值的比例不变性。即对图像做行或列的置换运算,图像的奇异值特征不变^[1,3]。

定理 4(奇异值的旋转不变性) 假设 $A, B \in R^{m \times n}$, 矩阵 A 的奇异值分别为 $\sigma_i (i=1, 2, \dots, p), p=\min(m, n)$, 若 p 为酉矩阵, 则矩阵 PA 的奇异值与矩阵 A 的奇异值相同, 即

$$|AA^H - \sigma_i^2 I| = |PA(PA)^H - \sigma_i^2 I| = 0 \quad (4)$$

定理 4 说明对图像做旋转运算和置换运算, 图像的奇异值特征不变。由定理 1 至定理 4 可知, 奇异值固有的稳定性、比例不变性和旋转不变性, 能有效地反映矩阵特征, 因此, 在人脸识别中将图像矩阵的奇异值作为识别特征不但必要而且有效^[5]。

1.2 奇异值降维压缩

奇异值降维压缩是为了降低神经网络的负担, 减少运算量, 以达到实时运算的要求。由定理 1 知, $m \times n$ 维图像矩阵奇异值个数为 $p=\min(m, n)$ 。如果直接对 P 个奇异值的特征进行识别, 由于 P 较大, 势必增加神经网络的冗余性和结构复杂性, 从而影响网络的计算效率和收敛率^[6]。为此需要进行奇异值降维压缩, 保留图像矩阵的较大的奇异值, 忽略较小奇异值, 经过奇异值降维压缩后, 原始图像的特性是否改变了? 原始图像矩阵与奇异值降维压缩后新得到的图像矩阵的关系如何? 下面通过两个定理加以说明^[7]。

定理 5 假设 $A \in R^{m \times n}$, A 的 SVD 由定理 1 给出, 且 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_p = 0$, 则 $\text{rank}(A) = r$,

$$A = \sum_{i=1}^r \sigma_i u_i v_i^T = U_r W_r V_r^T \quad (5)$$

定理 5 将矩阵表示成了秩 r 矩阵的和, 同时将矩阵中向量组的相关性问题转化为矩阵的非零奇异值个数问题。即矩阵的秩等于非零奇异值个数^[1,8]。

定理 6 假设 $A \in R^{m \times n}$, A 的 SVD 由定理 1 给出, 且 $\text{rank}(A) = r \geq s (m, n, r, s$ 均为正整数), 若

$$W_s = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_s), A_s = U \begin{vmatrix} W_s & 0 \\ 0 & 0 \end{vmatrix} V^T = \sum_{i=1}^s \sigma_i u_i v_i^T$$

因此

$$\text{rank}(A_s) = \text{rank}(W_s) = s$$

$$\|A - A_s\|_F = \min \|A - B\|_F \mid B \in R^{m \times n} \quad (6)$$

定理 6 表明, 在 F 范数意义下, W_s 是在空间 $R_s^{m \times n}$ (秩为 s 的 $m \times n$ 维矩阵构成的线性空间) 中 A 的一个降秩最佳逼近^[1,5]。也就是说, 在 F 范数意义下, A_s 是 A 中所有秩为 s 矩阵的一个最佳逼近。由定理 5 和定理 6 可知, 奇异值降维压缩是舍弃某些较小奇异值的过程。将人脸图像矩阵进行奇异值分解, 再通过奇异值降维压缩可获得图像矩阵的有效秩, 这里, 有效秩为所保留的奇异值个数。由保留奇异值生成的新的图像矩阵是原始图像矩阵在 F 范数意义下的一个降秩最佳逼近^[1,9]。

1.3 人脸特征提取的算法实现

本文的特征提取步骤为:

步骤 1 首先求出图像 A 的 $A^T A$ 。然后计算 $A^T A$ 的特征值 $\lambda_i (i=1, 2, \dots, l)$, 并按照从大到小的顺序排列, 特征值分别开平方 $\sigma_i = \sqrt{\lambda_i}$ 作为图像 A 的特征向量 $X, X = (\sigma_1, \sigma_2,$

$\dots, \sigma_l)$ 。奇异值矢量标准化的输出特征在 $[0, 1]$ 之间, 而且矢量中的特征均为从大到小排列, 即所有类别样本的奇异值矢量均具有结构相似的模式特征^[10]。

步骤 2 按照 $\sum_{i=1}^k \lambda_i / \sum_{i=1}^{M-1} \lambda_i \geq \alpha$ 进行奇异值特征向量的降维压缩。 M 为总共得到的特征向量, 可以选择 $\alpha = 80\%$ 。这样说明样本在前 K 个轴上的能量占整个能量的 80% 以上;

步骤 3 按照 $Y = (X - x_{\min}) / (x_{\max} - x_{\min})$ 进行奇异值特征向量的标准化。 $X \in R^d$ 为奇异值输入矢量, $Y \in R^d$ 为输出矢量, $X \in R^d$ 为奇异值输入矢量, x_{\max}, x_{\min} 分别为输入矢量 X 的最大值和最小值。实现奇异值特征的比例不变性, 由定理 3 可知, 奇异值具有比例不变性, 但必须经过标准化处理^[1,11]。

1.4 RBF 神经网络分类器的设计

RBF 神经网络的结构如图 1 所示。径向基神经网络的实质, 即寻找一个基于 $R^r \rightarrow R^s$ 的映射函数 $s \ll r$, 其中 r 是输入空间的维数, s 是输出空间的维数, u 是隐层节点。

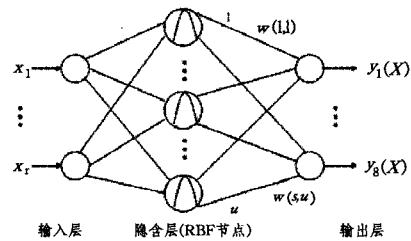


图 1 RBF 神经网络结构

一个具有 r 个输入和一个输出的 RBF 神经网络, 该网络可以看成如下形式的映射 $f: R^r \rightarrow R^s$:

$$y(X) = w_0 + \sum_{i=1}^u w_i R_i(\|X - C_i\|) \quad (7)$$

式中, $X \in R^r$ 是输入向量, $R_i(\cdot)$ 是基函数, $\|\cdot\|$ 表示输入空间上的欧氏范数, $w_i (0 \leq i \leq u)$ 是权值, $C_i \in R^r (1 \leq i \leq u)$ 是 RBF 的中心, u 是 RBF 的单元数。为了定义上的简单, 这里仅仅考虑了单输出的情形, 但很容易推广到多输出的情形^[1]。

在 RBF 神经网络中, $R_i(\cdot)$ 的选择有多种, 在各种类型的基函数中, 只有高斯函数是可分解的, 并随着与中心距离的增大而单调递减, 同时形成一种局部可调的处理单元。通常, 高斯函数是 RBF 神经网络的标准选择。高斯函数不仅适合于全局映射而且适合于细化局部的特征, 在已学习的映射上无需太多的改动就能达到快速学习的目的^[1,12]。

如果采用高斯函数而不考虑偏置量, 则式(1)可以写为如下的函数:

$$y(X) = \sum_{i=1}^u w_i R_i(-\frac{\|X - C_i\|^2}{\sigma^2}) \quad (8)$$

如果把各高斯函数的输出归一化, 则 RBF 网络可以产生如下归一化的输出响应。

$$y(X) = \frac{\sum_{i=1}^u w_i R_i(-\frac{\|X - C_i\|^2}{\sigma^2})}{\sum_{i=1}^u R_i(-\frac{\|X - C_i\|^2}{\sigma^2})} \quad (9)$$

由于高斯函数具有各向同性, 因此以高斯函数作为激活函数的神经网络称为径向基(RBF)神经网络。文献[2]对

RBF神经网络和多层感知器(MLPs)的逼近能力进行了深入的探讨。为度量神经网络的逼近能力,作者提出了“最佳逼近”的概念,这个概念大致描述如下:设 K 表示一个度量空间中的一个聚集,对于该空间中的每一个点 p ,在 K 中相应地存在一个与 p 距离最短的点^[9]。文献[2]的结论指出:RBF神经网络具有最好的逼近特性,而 MLPs 则没有。RBF神经网络分类器的输出具有线性判别函数的特性,它们在输出空间产生线性决策边界。因此,径向基神经网络分类器的性能严格依赖于 u 维空间中类的可分离性,这个 u 维空间是 u 个 RBF 单元通过非线性变换产生的。正因 RBF 神经网络有如此良好的特性,所以本文的研究基于动态径向基神经网络分类器。

1.5 动态 RBF 神经网络结构确定和参数选择

从人脸识别角度来看,不同类之间的最优边界应该用 RBF 神经网络来估计。反过来,从 RBF 神经网络的角度来看,神经网络可看作特征空间到类别的映射^[6]。每个输入样本用实向量来表示,每个输出代表一个类别。因此,对确定的 RBF 神经网络做如下设置:

- (1)输入的数量等于它的特征数,即输入空间的维数。
- (2)输出的数量等于类的量数。

隐含节点的确定比较麻烦,目前还是个尚未解决的问题。这里,我们提出的方法如下:

- (1)首先,让 RBF 的单元数量等于输出的数量, $u = s$, 即假定每个类只有一个子类。
- (2)对于每一个 RBF 单元 $k, k = 1, 2, \dots, u$, 把均值选为属于这个类的样本数据的中心,即:

$$C^k = \frac{1}{n^k} \sum_{i=1}^{n^k} P^k(i), k = 1, 2, \dots, u \quad (10)$$

式中, $P^k(i)$ 为属于类 k 的第 i 个样本, n^k 为类 k 中训练样本的总数^[7]。

- (3)对于任意类 k , 计算从平均值 C^k 到类 k 中最远点 $P^k(f)$ 之间的欧氏距离 d_k 为:

$$d_k = \| P^k(f) - C^k \| \quad (11)$$

- (4)对于任意类 k :

①按下式计算类 k 的平均值和其他类平均值之间的距离 $dc(k, j)$ 为:

$$dc(k, j) = \| C^k - C^j \|, j = 1, 2, \dots, s, j \neq k \quad (12)$$

- ②寻找

$$d_{\min}(k, l) = \arg \min_j (dc(k, j)), j = 1, 2, \dots, s, j \neq k \quad (13)$$

- ③检查 $d_{\min}(k, l)$ 和 d_k, d_l 之间的关系

情况 1 无重叠:如果 $d_k + d_l \leq d_{\min}(k, l)$, 那么类 k 同其他类没有重叠(见图 2(a))。

情况 2 有重叠:如果 $d_k + d_l > d_{\min}(k, l)$, 那么类 k 已经和其他类有重叠,这种情况下可能会产生错误分类。图 2(b)表示了这种情况, $d_k + d_l > d_{\min}(k, l)$, $|d_k - d_l| < d_{\min}(k, l)$ 。图 2(c)描述了 $d_k + d_l > d_{\min}(k, l)$, 且 $|d_k - d_l| \geq d_{\min}(k, l)$ 这种情况^[8]。

- (5)分裂准则:

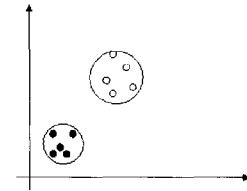
包含准则:如果类 k 被 l 类完全包含,即 $d_k + d_l > d_{\min}(k, l)$ 且 $|d_k - d_l| \geq d_{\min}(k, l)$, 那么类 l 将被分裂成两个聚类。

错分准则:如果类 k 包含多个其他类的数据,那么类 k 将分裂成两个聚类。

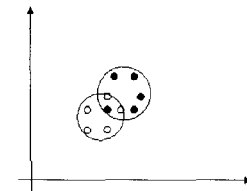
如果类 k 满足上面两准则中的一个条件,类 k 就将分裂成两个聚类,此时的中心基于相应的样本数据并根据式(7)计算。

(6)重复(2)–(5),直到所有的训练样本满足上面的两个准则。

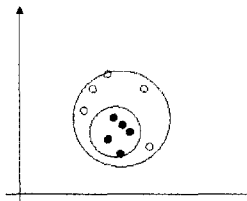
提出一种基于训练样本的类别信息的新的聚类算法,使所有同类的数据能被聚集在一起,尽量减少不同类数据混杂在一起,同时使 RBF 神经网络分类器结构尽可能紧凑^[9]。



(a) $d_k + d_l \leq d_{\min}(k, l)$



(b) $d_k + d_l > d_{\min}(k, l)$, $|d_k - d_l| < d_{\min}(k, l)$



(c) $d_k + d_l > d_{\min}(k, l)$ 且 $|d_k - d_l| \geq d_{\min}(k, l)$

图 2 聚类 and 样本分布

2 实验结果与分析

用 ORL 人脸数据库中的部分图像来验证所提算法。ORL 人脸数据库共有 400 幅人脸图像,它们是在英国剑桥大学的 Olivetti Research Laboratory 拍摄的,照片拍摄时间是在 1992 年 4 月到 1994 年 4 月。人脸姿态也有相当程度的变化,深度旋转和平面旋转可达 20° ;人脸的尺度也有多达 10% 的变化。图 3 是 ORL 人脸库中的 10 幅图像。



图 3 ORL 人脸数据库的 10 幅图像

学习之前,基于不同特征维数在3次仿真过程中的错误聚类数目列于表1中。由表1可以看出,当特征维数为39时,只有5个类属于如图2(a)所示的完全分离的分布,有235个类属于如图2(b)所示的分布。随着特征维数的减少,数据的分离特性将越来越好,相应地,聚类的性能也更好。

表1 学习前3次仿真结果

特征维数	数据的分布		
	聚类数	较好分开的类的数量	重叠的类的数量
39	40	5	235
25	40	96	144
10	40	187	53

在实验中,取ORL人脸图像库中的前5张作为训练图像,后5张作为待测图像。同时比较了文献[1]中基于奇异值向量特征的方法与本文方法的识别精度,实验结果如表2所列。

表2 不同人脸识别算法的实验结果

人脸识别方法	识别率
文献[1]的方法	21.25%
本文的方法	86.75%

实验结果发现,基于奇异值分解的秩矩阵逼近的人脸识别方法识别精度优于基于奇异值向量的方法。而本文提出的识别方法既利用了左右奇异值向量的信息,也利用了贡献率最大的奇异值的信息,它们的组合对图像的重建贡献也最大。本文提出的算法克服了传统奇异值特征识别方法的弱点,因此可以获得较好的识别效果。

文献[10]报道,利用多层前馈网络(MLN)分类器,分别采用PCA和SOM(自组织特征映射)方法来提取ORL数据库人脸特征时,其错误率分别为41.3%和39.8%。我们提出的方法不同于MLN,特别是充分利用了样本的信息进行聚类分析和确定RBF神经网络结构与参数,因此,识别效果比MLN、CNN^[10]和文献[7]的RBF方法都要好很多。

采用平均错误率作为指标来评判性能。所谓平均错误率 E_{ave} 的定义由下式表示^[10,11]:

$$E_{ave} = \frac{\sum_{i=1}^q n_{mis}^i}{qn_{tot}} \quad (14)$$

式中, q 为实验运行的次数,每次运行把数据库随机分成两个集合; n_{mis}^i 是第 i 次运行中错误分类的数量; n_{tot} 是每次运行时总的测试样本数量。基于 E_{ave} 准则,在同一个ORL数据库上比较CNN^[10]、NFL^[11]以及M-PCA^[12]的结果如表3所列。

表3 性能识别比较

方法	仿真次数	$E_{ave}/\%$
M-PCA	4	2.6
NFL	4	3.124
CNN	3	3.82
本文的方法	3	1.90

这里,CNN最佳 E_{ave} 值是基于3次实验得到的,对于NFL,最佳的错误率是当特征维数为39时得到的,平均错误率基于4次实验估计出来;对于M-PCA,总的性能是在平均

4次实验后得到的;对于本文提出的方法,最佳的错误率是基于3次运行得到的,特征维数为30,人脸特征维数提取方法与文献[11]完全相同,而划分训练样本和测试样本的方法也完全同文献[10]和文献[11]。

结束语 利用奇异值分解的方法来提取人脸代数特征并进行一系列的奇异值降维、压缩,结果表明这种方法减少了计算量,是行之有效的。然后,通过结构尽可能紧凑的RBF神经网络分类器来处理高维、小样本的分类问题。提出一种新颖的有监督的聚类方法,来确定RBF神经网络分类器的结构和初始参数。最后,提出了一种混合的学习算法来训练RBF神经网络,使得在梯度下降寻优算法中大大降低了搜索空间的维数。在ORL数据库上进行的仿真结果表明,这个方法无论是在分类的错误率上还是在学习的效率上都能表现出极好的性能。

参考文献

- [1] Hong Z Q. Algebraic feature extraction of image for recognition [J]. Pattern Recognition, 2010, 24(3): 211-219
- [2] Cheng Y, et al. Human Face recognition Method Based on the Statistical Model of Small Sample Size [A]// SPIE Proceedings on Intelligent Robots and Computer Vision [C]. Boston Massachusetts, USA, 2009: 85-95
- [3] 王蕴红,等. 基于奇异值分解和数据融合的脸像鉴别[J]. 计算机学报, 2009, 23(6): 649-653
- [4] Lawrence S, et al. Face recognition: a convolution neural-network approach [J]. IEEE Transactions on Neural networks, 2007, 8(1): 98-113
- [5] Nebauer C. Evaluation of convolution neural networks for visual Recognition [J]. IEEE Trans, Neural Networks, 2008, 9(4): 685-69
- [6] Turk M A, Pentland A P. Eigenfaces for Recognition [J]. J. Cognitive Neuroscience, 2009(3): 71-86
- [7] Fukunaga K. Introduction to Statistical Pattern Recognition(2nd ed)[M]. San Diego, CA: Academic Press, 2010
- [8] Swets D L, Weng J. Using Discriminant Eigenfeatures for Image Retrieval [J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2006(18): 831-836
- [9] Lotlikar R, kothari R. Fractional-Step Dimensionality Reduction [J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2010(22): 623-627
- [10] Lawrence S, Giles C L, Tsoi A C, et al. Face Recognition: A Convolutional Neural-Network Approach [J]. IEEE Trans. Neural Networks, Special Issue on Neural Networks and Pattern Recognition, 2007(8): 114-132
- [11] Li S Z, Lu J. Face Recognition Using the Nearest Feature Line Method [J]. IEEE Trans. Neural Networks, 2009(10): 439-443
- [12] Brenna V, Principe J. Face Classification Using a Multiresolution Principal Component Analysis [C]// Proc. IEEE Workshop Neural Networks for Signal Processing. 2008: 506-515