

基于多特征融合和 Boosting RBF 神经网络的人脸识别

嵇新浩 丁荣涛

(浙江商业职业技术学院信息技术系 杭州 310053)

摘要 提出一种多特征信息融合的人脸识别方法。应用 Zernike 矩方法和非负矩阵分解法(NMF)分别提取具有旋转不变性的人脸几何特征和人脸子空间投影系数特征,将这两种具有一定互补性的特征串行融合,得到一个分类能力更强的特征。在此基础上,采用 RBF 神经网络进行人脸识别。为了提高神经网络的分类准确率和泛化能力,采用 Boosting 方法进行网络集成。实验结果表明,提出的算法利用较少样本数据即可快速地进行人脸识别。

关键词 人脸识别, Zernike 矩, 非负矩阵分解法, Boosting 方法, RBF 神经网络

Face Recognition Based on Feature Fusion and Boosting RBF Neural Network

Ji Xin-hao DING Rong-tao

(Information Technology Department, Zhejiang Vocational College of Commerce, Hangzhou 310053, China)

Abstract A new face recognition method based on feature fusion was proposed. Firstly, face projection coefficient features were extracted by NMF methods. Then these features were combined with the face rotation invariance Zernike moments features so as to get a new feature which had higher discriminating power. With this new feature, RBF neural network is used to class the face. In order to improve the precision of the RBF neural network for face recognition, boosting algorithm is used to build an integration-neural network. The experiment results show that the algorithm can recognize face fleetly using a few samples.

Keywords Face recognition, Zernike moments, Non-negative matrix factorization(NMF), Boosting algorithm, RBF neural network

1 引言

人脸识别是当前人工智能和模式识别的研究热点。它可以用于身份认证、公共场合对人的监视、图像数据库的检索、提高人与计算机的交互能力等。长期以来,如何利用计算机对人脸图像进行自动检测与识别,一直是图像处理与模式识别的研究热点与难点。人脸识别有两种典型方法^[1-3]:一种是基于人脸几何特征的识别,通过提取出局部轮廓信息及灰度信息来设计具体识别算法;另一种是基于代数特征的自动人脸识别方法,即子空间分析方法,其本质是一种特征提取和选择的方法,主要思想是在原空间(样本空间)中寻找合适的子空间(特征空间),通过将高维样本投影到低维子空间上,在子空间上进行分类。

本文提出一种新颖的基于 Boosting RBF 神经网络的人脸识别方法。提取人脸图像 Zernike 矩特征和人脸子空间特征作为 RBF 神经网络分类器的输入。为了进一步提高神经网络的泛化能力,采用 Boosting 方法^[4,5]进行网络集成。在 RBF 神经网络分类器的学习过程中,Boosting 方法对训练失败的训练样本赋予较大的权值,使得在下一个网络学习中集中对比较难的训练样本进行学习。同时,Boosting 方法给每个神经网络赋予对应的权值,所设计的分类器以加权投票方式进行分类决策。实验结果表明,提出的算法利用较少样本数据即可快速地进行人脸识别,具有潜在的应用价值。

2 人脸特征提取

2.1 人脸的 Zernike 矩特征

矩在统计学中用于表征随机量的分布。若把二值图或灰

度图看作是二维密度分布函数,就可把矩技术应用于描述一幅图像的特征。Zernike 矩方法是源于 Yeague 提出的正交矩思想,能够构造出任意的高阶矩,且具有旋转不变性,有非常强的图像表示能力,可以利用得到的 Zernike 不变矩进行图像重建,因此矩特征所含信息量丰富。

设二维离散图像函数用 $f(x, y)$ 表示, $n + m$ 阶 Zernike 不变矩表示为

$$A_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y) [V_m(x, y)]^* dx dy \quad (1)$$

实质上,它是一种映射,将图像函数变换到一组正交基函数上。

图像 $f(x, y)$ 的 Zernike 矩是该图像在一组正交多项式 $\{V_m(x, y)\}$ 上的投影。这里所谓正交,是指 $\{V_m(x, y)\}$ 在单位圆内 $\{x^2 + y^2 \leq 1\}$ 满足下列条件:

$$\iint_{x^2+y^2 \leq 1} [V_m(x, y)]^* V_p(x, y) dx dy = \begin{cases} \frac{\pi}{n+1}, & \text{if } n = p, m = q \\ 0, & \text{if } n \neq p, m \neq q \end{cases} \quad (2)$$

式中, $V_m(x, y) = R_m(x, y) e^{jm \tan^{-1} \frac{y}{x}}$, 其中 n 是正整数或零; m 是正或负整数,且满足 $n - |m| = \text{偶数}$; $R_m(x, y)$ 是径向多项式,即

$$R_m(x, y) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)! (x^2 + y^2)^{s(\frac{n-s}{2})}}{s! [\frac{n+|m|}{2} - s]! [\frac{n-|m|}{2} - s]!} \quad (3)$$

假设 $f(r, \theta)$ 是图像目标在极坐标下的函数表示,相应的傅氏级数形式为

$$f(r, \theta) = \sum_{n=-\infty}^{\infty} f_n(r) e^{jn\theta} = \sum_{n=-\infty}^{\infty} f_n(r, \theta) \quad (4)$$

这里

$$f_n(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \theta) e^{-jn\theta} d\theta, f_n(r, \theta) = f_n(r) e^{jn\theta} \quad (5)$$

这样 Zernike 矩变为

$$A_{nm} = \frac{n+1}{n} \int_0^1 \int_0^{2\pi} f(r, \theta) V_m^* r dr d\theta = \frac{n+1}{n} 2\pi \int_0^1 f_n(r) R_m(r) r dr \quad (6)$$

离散形式为

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_m^*(r, \theta) \quad (7)$$

其中 V_m^* 是 V_m 的共轭, $x^2 + y^2 \leq 1, V_m(r, \theta) = R_m(r) e^{jn\theta}$ 。

由于 Zernike 矩的定义中引入了完全正交的基函数集, 取代了原点矩定义中的非正交基函数集, 因此在理论上 Zernike 矩比原点矩具有更好的数学性质。

对人脸图像进行大小调整后, 求取 Zernike 矩的幅值作为特征, 得到一个 15 维的人脸矩特征向量。虽然 Zernike 矩可以抵抗少量噪声干扰, 但是当人脸图像受到其他因素(如光照)的影响时, Zernike 矩特征识别能力会大大降低。因此, 本文采用提取人脸子空间特征与之融合的方法提高特征的识别能力, 进而提高人脸的识别率。

2.2 NMF 提取人脸子空间特征

非负矩阵分解法(NMF, Non-negative Matrix Factorization)和 K-L 变换等方法一样, 都是一种矩阵分解方法, 它们的不同之处在于分解过程中所受的约束不同。正因如此, 在人脸识别中, 非负矩阵分解克服了其他分解方法中在特征脸空间上投影得到的投影系数可能出现正负相互抵消的情况。非负矩阵分解中, 无论是特征图像、投影系数以及重建图像, 都能保证是非负的。具体的分解方法如下:

对于任何由 Num 幅图像、每一幅图像均为 $n \times 1$ 维矩阵组成的矩阵 $X = [x_{ij}] (n * Num \text{ 维})$, 其每一列均由一幅人脸图像的非负灰度值所组成, 我们都可以将其分解为两个矩阵 B, H 的乘积:

$$X \approx BH \quad (8)$$

其中, B 是一 $n \times m$ 维矩阵, H 是一 $m \times Num$ 矩阵, 代表系数或权重; $m \leq n$, 因此该分解达到了降维的目的。

在这种矩阵分解中, PCA 分解能够保证矩阵的正交性, 而 NMF 分解主要能够保证矩阵的非负性。因此, B 和 H 均是非负的。

2.2.1 NMF 分解的价值函数

NMF 分解中, 定义 X 和 Y 距离为

$$C(X, Y) = \|X - Y\|^2 = \sum_{i,j} (X_{i,j} - Y_{i,j})^2 \quad (9)$$

其中, i, j 分别表示矩阵 X 和矩阵 Y 的行标和列标, $0 \leq i < n, 0 \leq j < Num$ 。

我们用该距离作为衡量分解 $X \approx BH = U = [y_{ij}]$ 的标准: 该距离越小, 该分解越接近原矩阵, 最后该距离接近于 0。

因此, NMF 分解的定义如下:

$$\min_{B, H} (\|X - BH\|^2) \quad (10)$$

对于任意的 j , 满足

$$\sum_i b_{ij} = 1$$

其中, $B, H \geq 0$, 即所有的 B, H 均是非负的。

2.2.2 NMF 分解算法

根据上面的分析, 给出下面的算法实现过程。

① 初始矩阵 B, H , 其元素均为非负随机数;

② 更新 B, H , 直到 $\|X - BH\|^2$ 收敛:

(a) 更新 H 矩阵一行元素, $H_{i,j} = H_{i,j} \frac{(B^T X)_{ij}}{(B^T B H)_{ij}}$

(b) 更新 B 矩阵相应的一列元素, $B_k = B_k \frac{(X H^T)_k}{(B^T H H^T)_{k,k}}$

(c) 用(a)、(b)中刚刚更新的行与列更新 $B^T X, B^T B H, X H^T, B^T H H^T$ 中的相应元素。设 \bar{x} 是所有训练样本的平均图像, 对于每一个人脸训练样本 x_i , 投影到特征子空间, 得到如下特征:

$$h_i = B^{-1}(x_i - \bar{x}) \quad (11)$$

对于任一待识样本 q , 将其投影到特征子空间, 得到如下特征:

$$h_q = B^{-1}(q - \bar{x}) \quad (12)$$

用非负矩阵分解法提取了 16 维人脸子空间基向量, 即特征脸, 将人脸图像在子空间的 16 维投影系数 h 作为人脸的特征。将 15 维 Zernike 矩特征和 16 维子空间投影系数特征串行组合到一起, 得到一个 31 维的人脸特征向量 X 。

3 RBF 神经网络

径向基函数 RBF 神经网络是一种局部逼近网络, 能以任意精度逼近任一连续函数, 与传统的 BP 神经网络相比有众多的优越性。BP 网络收敛速度慢, 容易陷入局部极值点。RBF 网络训练速度快, 对于每个输入输出数据对, 只有少量权值需要调整。RBF 神经网络由输入层、隐含层和输出层组成, 其结构如图 1 所示。输入层节点只传递到隐含层, 隐含层节点由基函数构成, 输出层节点通常是线性的。隐含层节点通过径向基函数对输入信号产生一个局部响应, RBF 神经网络的输出层节点是隐含层基函数的输出进行线性加权组合, 即输出层的输出 y_i 为

$$y_i = \sum_{j=1}^m w_{ij} R_j(x), j=1, 2, \dots, p \quad (13)$$

式中 w_{ij} 为隐含层第 i 个节点到输出层第 j 个节点之间的连接权系数, m, p 分别为隐含层和输出层节点个数。 $R_i(x)$ 为径向基函数, 我们选高斯函数:

$$R_i(x) = \exp(-\|x - c_i\|^2 / 2\sigma_i^2), i=1, 2, \dots, m \quad (14)$$

其中 x 是 n 维输入向量; c_i 是第 i 个基函数的中心, 它与 x 的维数相同; σ_i 是第 i 个基函数的宽度; 范数 $\|x - c_i\|$ 表示 x 和 c_i 间的距离, 通常取欧式距离。

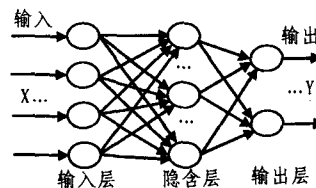


图 1 RBF 神经网络结构

4 基于 Boosting RBF 神经网络的人脸识别

以 Zernike 矩特征和 NMF 提取人脸子空间特征串行组合到一起, 得到一个 31 维的人脸特征向量 X 来表示人脸样本, 并以此作为神经网络分类器的输入, 利用 Boosting 方法进行分类, 如图 2 所示。 S 是训练样本, D_t 是第 t 轮学习样本的权向量, α_t 是第 t 个神经网络的权值, X 是测试样本的特征向量, Y 是其类别。

(1) 基于 Boosting RBF 神经网络的人脸识别方法如下。

Step1 初始化。初始赋予每个样本 $(x(1), y'(1))$, $(x$

(2), $y'(2)$, ..., $(x(n), y'(n))$ 相等的权重 $D_1(i) = \frac{1}{n}$ 。

Step2 For $t=1, 2, \dots, T$ Do

①在 D_t 下, 用 K 均值聚类方法和梯度下降法进行学习, 得到神经网络 C_t ;

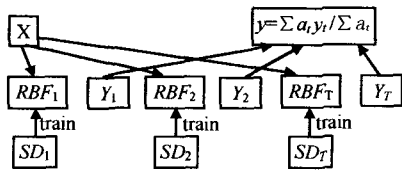


图2 基于 boosting RBF 神经网络的人脸识别

1) 用 K 均值法确定聚类中心 c_i :

1. 随机选择初始中心 $c_i(0), i=1, \dots, m$;

2. 根据 Boosting 方法, 在第 t 个网络的学习过程中, RBF 神经网络中心更新式

$$c_{iq}(k) = c_{iq}(k-1) + \alpha(k) D_t(k) [x(k) - c_{iq}(k-1)] \quad (15)$$

2) 用梯度下降法来确定 w_{ij} :

1. 随机选择初始权重系数 $w_{ij}(0)$;

2. 用梯度下降法得到 w_{ij} 的更新式:

$$w_{ij}(k+1) = w_{ij}(k) + \eta R_i(k) D_t(y_j(k) - y'_j(k)) \quad (16)$$

其中 w_{ij} 为第 t 个神经网络隐层第 i 个节点到输出层第 j 个节点之间的连接系数, η 为学习步长, $y_j(k)$ 和 $y'_j(k)$ 分别是网络输出层第 j 个节点的实际输出和理想输出。

②计算该神经网络的错误率 E_t :

$$E_t = \sum_{k=1}^N \sum_{j=1}^p D_t^2(k) (y_j(k) - y'_j(k))^2 \quad (17)$$

$$\beta_t = E_t / (1 - E_t) \quad (18)$$

Step3 根据错误率更新样本的权重;

$$1 - \frac{|y(k) - y'(k)|}{\max(|y(k) - y'(k)|_{z, D_{t+1}(k) = D_t(k) \cdot \beta_t})} \quad (19)$$

Step4 计算网络的权重为

$$\alpha_t = \ln(1/\beta_t) \quad (20)$$

(2) 人脸识别

以 Zernike 矩和 NMF 提取人脸子空间特征得到的 31 维特征向量 X 作为各个 RBF 神经网络的输入, 得到各个 RBF 网络的输出 y_t , 集成神经网络的输出为

$$y = \sum a_t y_t / \sum a_t \quad (21)$$

其中 $y = (y_1, y_2, \dots, y_{10})$, 若 $y_i = \max\{y_j\}, j=1, \dots, 10$, 则该人脸属于第 i 类。

5 实验结果分析

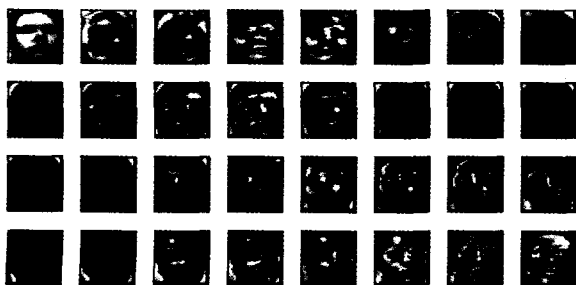


图3 NMF 方法得到的特征脸子空间

实验以国际上广泛使用的 ORL 标准人脸库为对象, 对文中算法的识别性能进行测试, 并与传统算法进行比较。ORL 人脸库由 10 人, 每人 10 幅共 100 幅图像组成。取每人

前 5 幅图像作为训练样本, 后 5 幅作为测试样本, 这样训练样本和测试样本总数各为 50 且互不重叠。对每幅图像进行旋转、剪裁、缩小和放大变换, 得到 32×32 统一大小的校准图像, 再对图像做灰度拉伸等预处理工作, 以改善图像质量。然后分别采用 NMF 方法得到特征脸子空间图像, 部分特征脸图像如图 3 所示。

以 Zernike 矩特征和 NMF 提取人脸子空间特征串行组合到一起, 得到一个 31 维的人脸特征向量 X 来表示人脸样本, 作为 Boosting 神经网络分类器的训练样本, 训练神经网络分类器。神经网络的输入结点数为特征的维数(实验中取 31)。隐层结点数为输入层的 1.5 倍。输出层节点数为人脸的类别数 $N, N=10$ 。

实验 1 比较 2 种不同的分类方法: 即单独的 RBF 神经网络方法^[1]和基于 Boosting RBF 神经网络与 3 种不同的特征结合的识别率, Boosting 训练次数取 25。结果如表 1 所示。

表 1 Boosting 算法与神经网络实验结果比较

	RBF 神经网络识别率	boosting RBF 神经网络识别率
Zernike 不变矩	84%	90%
NMF 提取子空间投影系数	82%	88%
融合特征	92%	98%

由表 1 可见, Zernike 不变矩特征对提取人脸特征是一种比较合适的方法; Boosting 神经网络分类器的分类能力优于单一 RBF 神经网络的效果, 说明 Boosting 算法加强了神经网络的分类能力; Zernike 不变矩特征与 NMF 子空间人脸特征融合后, 人脸识别效果明显变优。

实验 2 Boosting 训练次数对识别率和识别时间的影响, 结果如表 2。

表 2 不同 Boosting 训练次数对分类结果的影响

Boosting 训练次数	识别率	识别时间(s)
15	86%	0.34
20	92%	0.36
25	98%	0.42

由此可见, 随着 Boosting 训练次数增加, 分类器的性能更好。但可以观察到, 当训练次数为 20 到 30 之间, 提升的比率逐渐减小。因此, 如果在实际应用中, 需要在训练时间和性能之间做个权衡。

结束语 人脸识别是非常重要的研究课题。本文提出了一种新颖的基于 Boosting RBF 神经网络的人脸识别方法, 该方法提取人脸图像 Zernike 矩特征和人脸子空间特征作为 RBF 神经网络分类器的输入。为了提高神经网络的泛化能力, 采用 Boosting 方法对 RBF 神经网络进行集成。分类器以加权投票方式进行分类决策。实验结果说明了该方法的正确性和有效性。

参考文献

- [1] 张俊, 何昕, 李介谷. 基于面部几何特征点提取的人脸识别方法. 红外与激光工程, 1999, 28(4)
- [2] 范宏深, 倪国强, 申会堂. 利用几何特性及神经网络进行人脸探测技术的研究. 光学技术, 2002, 28(2)
- [3] Lee D, Seung H. Learning the parts of objects by nonnegative matrix factorization. Nature, 1999, 401: 788-791
- [4] Freund Y. Boosting a Weak Algorithm by Majority. Information and Computation, 1995, 121(2): 256-285
- [5] Freund Y, Schapire R E. A Decision-Theoretic Generalization of OnLine Learning and an Application to Boosting. Journal of Computer and System Sciences, 1997; 55(1): 119-139