

# 基于最佳鉴别变换和分类器组合的人脸自动识别

赵海涛 金 忠 杨静宇

(南京理工大学计算机科学系, 南京 210094)

**摘 要** 人脸识别技术在商业和法律上有广泛的应用前景, 在安全监控中也大有用武之地. 其主要任务是利用已有的人脸图象库, 识别静止的或视频图象中的一张或多张人脸. 从抽取具有统计不相关的模式特征着手, 通过基于小波变换的图象分解和 KL 变换等处理, 避开人脸识别的小样本集的局限, 并通过运用具有统计不相关性的最佳鉴别变换, 来抽取人脸的有效鉴别特征. 同时, 利用多特征多分类器组合的方法对图象进行识别. 该方法在 ORL 人脸图象库上进行实验, 得到识别错误率为 2% 的实验结果, 这是目前在此人脸图象数据库上所得到的最好的实验结果. 而且本方法对人脸的姿态、表情等条件具有一定的不敏感性.

**关键词** 人脸识别 最佳鉴别变换 小波变换 KL 变换 多分类器组合

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2000)09-0769-04

## The Automatic Recognition of Human Faces Based on the Uncorrelated Optimal Discriminant Transformation and Multiclassifier Combination

ZHAO Hai-tao, JIN Zhong, YANG Jing-yu

(Dept. of Computer Sci., Nanjing Univ. of Sci. & Tech. 210094)

**Abstract** Face recognition technology (FRT) has numerous commercial and law enforcement applications, especially in video surveillance. The primary task at hand, given still or video images, requires the identification of one or more persons using a database of stored face images. Based on Fisher discriminant criterion, In order to extract features by using the uncorrelated discriminant transformation, we use orthogonal wavelet transformation and KL transformation to process the face images at first. According to peoples' recognition experience, we use multi-feature and multi-classifier combination to give out the classification results. Experiments on ORL database obtained an error rate of 2%, which is the best result on this database up to now. Experimental results also show that this method does not sensitive to the pose and expression of human faces.

**Keywords** Face recognition, Optimal discriminant transformation, Wavelet transformation, KL transformation, Combining classifier

## 0 引 言

人脸识别是模式识别研究领域的重要课题, 也是一个目前非常活跃的研究方向. 它一般可描述为: 给定一静止或动态图象, 利用已有的人脸数据库来确认图象中的一个或多个人. 近年来, 人脸识别技术

研究非常活跃, 其中基于 KL 变换的特征脸方法和以奇异值特征为代表的代数特征方法等得到了新的发展, 而且人工神经网络、小波变换等新思想、新方法也在人脸识别研究中得到了很广泛的应用. 由于图象的代数特征反映了图象的本质, 尽管它的物理意义并不十分直观, 但是它对模式的分类仍起着重要的作用, 因此, 对人脸图象进行代数变换, 通过抽

取反映人脸图象内在属性的代数特征来进行人脸识别,这是目前人脸识别研究很有潜力的一个方向.

在模式识别理论中,特征抽取的一般原则是要求所抽取的特征之间的统计相关性越小越好,最好是抽取不相关的特征,可是在Foley-Sammon最佳鉴别矢量集当中,虽然任意两个最佳鉴别方向是正交的,但投影所得到的两个鉴别特征却是统计相关的.因此本文运用了一种具有统计不相关性的最佳鉴别变换<sup>[1]</sup>来处理图象.近年来,分类器组合方法已成为模式识别领域的前沿研究课题,并在模式识别的很多应用方面,如字符识别、目标识别等领域,取得了较好的应用效果.由于不同性质的特征往往反映物体的不同方面,因而在一种特征空间很难区分的两种模式,有可能在另一种特征空间上很容易分

开,而对应于同一特征的不同分类器又从不同的角度将该特征映射到决策空间上,因此利用不同性质特征和不同分类器的组合就可能全面区分出一个物体,从而得到一个较好的分类结果.

## 1 ORL人脸图象数据库

ORL人脸图象数据库由40人,每人10幅不同表情和姿态的 $92 \times 112$ 象素大小的图象所组成,其中有些图象是拍摄于不同时期,其人脸脸部表情与脸部细节均有变化,例如,笑或不笑、眼睛睁着或闭着、带或不带眼镜等,人脸姿态也有变化,其深度旋转与平面旋转可达到 $20^\circ$ ;人脸的尺度也有最多10%的变化.图1是ORL人脸图象数据库中的一组图象.



图1 ORL人脸图象数据库中的一组图象

## 2 人脸图象的特征抽取

### 2.1 具有统计不相关性的最佳鉴别矢量集

设模式特征 $\mathbf{X}$ 为 $n$ 维矢量,模式类别有 $c$ 个,模式类内散布矩阵为 $\mathbf{S}_w$ ,类间散布矩阵为 $\mathbf{S}_b$ ,总体散布矩阵为 $\mathbf{S}_t$ .Fisher鉴别准则函数定义如下

$$F(\boldsymbol{\varphi}) = \frac{\boldsymbol{\varphi}^T \mathbf{S}_b \boldsymbol{\varphi}}{\boldsymbol{\varphi}^T \mathbf{S}_w \boldsymbol{\varphi}} \quad (1)$$

其中 $\boldsymbol{\varphi}$ 为 $n$ 维列向量.

使式(1)Fisher鉴别准则函数取到最大值的向量 $\boldsymbol{\varphi}$ 是Fisher最佳鉴别方向,其物理意义是模式样本集在 $\boldsymbol{\varphi}$ 方向上的投影使得模式具有最小的类内散布和最大的类间散布. $\boldsymbol{\varphi}$ 是下列广义本征方程中与最大的本征值对应的本征向量

$$\mathbf{S}_b \boldsymbol{\varphi} = \lambda \mathbf{S}_w \boldsymbol{\varphi} \quad (2)$$

求出Fisher最佳鉴别方向 $\boldsymbol{\varphi}$ .在求出 $r$ ( $r \geq 1$ )

个最佳鉴别方向 $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_r$ 后,第 $r+1$ 个最佳鉴别方向是满足式(3)共轭正交条件的,使Fisher鉴别准则函数取到最大值的向量

$$\boldsymbol{\varphi}_{r+1}^T \mathbf{S}_t \boldsymbol{\varphi}_i = 0, \boldsymbol{\varphi}_{r+1}^T \boldsymbol{\varphi}_i = 1, i = 1, 2, \dots, r \quad (3)$$

由上述条件可得,具有统计不相关性的最佳鉴别矢量集的任意两个分量 $y_i$ 和 $y_j$ 是统计不相关的.即

$$E[(y_i - E y_i)(y_j - E y_j)] = \boldsymbol{\varphi}_i^T \mathbf{S}_t \boldsymbol{\varphi}_j = 0 \quad (4)$$

关于具有统计不相关性的最佳鉴别方向的求解有如下的结论:

具有统计不相关性的最佳鉴别矢量集的第 $r+1$ 个最佳鉴别方向 $\boldsymbol{\varphi}_{r+1}$ 是下列广义本征方程中与最大的本征值对应的本征向量

$$\mathbf{P} \mathbf{S}_b \boldsymbol{\varphi}_{r+1} = \lambda \mathbf{S}_w \boldsymbol{\varphi}_{r+1} \quad (5)$$

式中

$$\mathbf{P} = \mathbf{I} - \mathbf{S}_t \mathbf{D}^T (\mathbf{D} \mathbf{S}_t \mathbf{S}_w^{-1} \mathbf{S}_t \mathbf{D}^T)^{-1} \mathbf{D} \mathbf{S}_t \mathbf{S}_w^{-1} \quad (6)$$

这里: $\mathbf{I}$ 为单位矩阵,而且

$$D = [\mathcal{Q}_1, \mathcal{Q}_2, \dots, \mathcal{Q}_n]^T = \begin{pmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \\ \vdots \\ \mathcal{Q}_n \end{pmatrix} \quad (7)$$

在求解基于 Fisher 鉴别准则函数的最佳鉴别矢量集的问题时, 使 Fisher 鉴别准则函数取到最大值的 Fisher 最佳鉴别方向  $\mathcal{Q}$  是下列广义本征方程中与最大的本征值对应的本征向量

$$S_w \mathcal{Q} = \lambda S_m \mathcal{Q} \quad (8)$$

在大样本集的情况下, 类内散布矩阵  $S_w$  为非奇异矩阵, 因此能精确求解式(8), 即可以得到最佳鉴别矢量, 但在小样本集的情况下,  $S_w$  可能为奇异矩阵, 因而使得最佳鉴别矢量的求解问题变得相当困难. 另外, 由于只有在类内散布矩阵  $S_w$  为非奇异矩阵时, 才能计算具有统计不相关性的最佳鉴别矢量集, 因此对于 ORL 图象库中的人脸图象, 如果通过实验来直接求解统计不相关的最佳鉴别矢量集或改进的最佳鉴别矢量集, 就会由于图象的维数高, 且训练样本少, 因而这一方面将致使运算量太大, 另一方面可能根本不能求解出最佳鉴别矢量集, 因此本实验采用下述 3 种方法先将图象降到 42 维.

如今小波分析以其优良的视频局部性能已成为有力的信号分析处理工具. 因为小波分析是用一组不同的带通滤波器对信号进行滤波, 然后将信号分解到不同的频带上, 进行分析处理, 而图象信息可以看成是一类复杂的非线性、非平稳的信号源, 所以小波变换在图象处理中, 也有着广泛的应用. 同时由于正交小波变换既能保留数字图象的原有信息, 又具有分解的唯一性和去相关作用, 即是一种能量保持型的可逆变换, 所以我们选择由多尺度分析获得的标准正交小波基来对人脸图象做离散的小波变换. 而且由于在变换后得到的子图象中, 其低频子图象包含了原始图象的主要描述信息, 它的分类效果接近原始图象的分类效果, 因此, 我们略去 3 个高频子图象, 而只把低频子图象作为降维后的图象用于以后的进一步处理.

本文采用的方法第一种首先是运用基于小波变换的图象分解方法将图象降到  $12 \times 14 = 168$  维, 然后再运用 KL 变换将图象降到  $6 \times 7 = 42$  维; 第二种是直接运用基于小波变换的图象分解方法, 将图象降到  $6 \times 7 = 42$  维; 第三种是运用  $15 \times 15$  的窗口对图象进行扫描, 并计算窗口中所有象素点的均值, 即窗口每次向右移动 15 个象素, 当窗口不能再向右移

动时, 再向下移动 15 个象素, 并同时窗口移动到图象的最左侧, 然后依次扫描整个图象, 将图象降到  $6 \times 7 = 42$  维的方法.

## 2.2 维数问题

在模式识别理论中, 有一个最佳维数问题, 即一般认为, 对一个模式识别问题来说, 任何模式特征都存在一个最佳维数, 当模式特征的维数等于最佳维数时, 即可以获得最好的识别率. 由于 ORL 图象库的类别数为 40, 因此根据 Fukunaga 维数定理, 由 42 个人脸图象矢量, 可抽取 39 维的人脸有效鉴别特征, 即可由 42 个人脸图象矢量空间得到 39 维的人脸鉴别特征空间.

对降维后的矢量, 实验运用每类的前五个矢量作为训练样本, 后五个作为测试样本.

## 3 单分类器实验及结果

通过对 ORL 人脸图象库的 400 幅图象(40 人每人 10 幅图象) 识别实验(200 个样本为训练样本, 200 个样本为测试样本), 其识别错误率实验结果见表 1.

表 1 单分类器的识别结果

分类方法	第 1 种降维方法		第 2 种降维方法		第 3 种降维方法	
	神经网络 分类器	最近邻 分类器	神经网络 分类器	最近邻 分类器	神经网络 分类器	最近邻 分类器
识别错误数目	18	12	24	14	22	10
识别错误率	9%	6%	12%	7%	11%	5%

## 4 多分类器组合识别实验及结果

目前较常用的多分类器组合方法是投票法. 因为投票法简单, 且易于实现, 它采用的是少数服从多数的原则, 即支持某一决策的分类器个数若最大, 则这一决策就是最终的结果. 对上述表 1 中得到的识别结果再用投票法进行实验, 其识别错误率为 2.5%, 拒识率为 1%.

下面的实验就运用了一种基于多级分类的投票法. 本文所说的多级分类是指从上一级的所有类别中挑选出与测试样本最为相似的那些类别, 作为一个候选子集, 然后再进一步分类、识别. 其步骤如下:

第 1 步, 对于表 1 中几种方法得到的识别结果, 若其中有 4 个或 4 个以上相同, 实验就将此结果作为分类的最终结果, 否则拒识;

第2步,对于每一个拒识的样本,只将表1中 $n$ 种方法得到的识别结果作为第3步分类时该样本可能的识别结果;

第3步,对于每一个拒识的样本,在第2步形成的类别范围内,进行二次识别,并运用投票法按少数服从多数的原则进行分类.

运用上述的方法对 ORL 图象数据库进行分类,第1步的实验结果为15个样本拒识,有两个样本识别错误;经过二次识别,最终的实验结果为4个样本识别错误,识别错误率仅为2%.分别考虑3个最近邻分类器和3个BP神经网络分类器,它们的平均识别错误率依次为6%和10.7%;而6种降维识别方法总的平均识别错误率为8.3%.

Lawrence 等曾分别用特征脸方法与神经网络方法对 ORL 人脸图象数据库作过分类识别实验,在每人的训练样本数取5、测试样本数也取5的情况下,特征脸方法的识别错误率为10.5%,而神经网络方法的最低识别错误率为2.5%,平均识别错误率为3.8%<sup>[4]</sup>.

图2是一个分类错误的例子,其中右侧的图象是一个被错误分类的图象,分类器将其划分为与左侧的图象一类.图3是一个正确识别的例子,分类器将它们正确地归为一类.从正确识别的例子可以看出,本方法对人脸的姿态和表情的确有一定的不敏感性.



图2 错误分类的例子



图3 正确分类的例子

## 5 结束语

本文算法在一定程度上提高了识别率,对于人脸的姿态、表情等条件也具有一定的不敏感性.运用本方法对 ORL 图象库进行的多特征、多分类器组合分类实验时,识别的正确率为98%,达到了目前在 ORL 图象数据库上所得到的最好的结果.

总之,通过对图象进行最佳鉴别变换,并在此基础上进行识别是一个很有应用潜力的方法,其应用也不仅仅限于人脸识别.

## 参考文献

- 1 Sammon J W, Jr., An optimal discriminant plane, IEEE Trans. Comput., 1970, 19(9): 826~ 828.
- 2 Foley D H, J W Sammon, Jr., An optimal set of discriminant vectors, IEEE Trans. Comput., 1975, 24(3): 281~ 289.
- 3 Longstaff I D. On extensions to fisher's linear discriminant function, IEEE Trans. Pattern Anal. Machine Intell., 1987, 9(2): 321~ 324.
- 4 Lawrence S, Giles C L, Tsoi A C *et al.* Back, face recognition: a convolutional neural-network approach, IEEE Trans. Neural Network, 1997, 8(1): 98~ 113.
- 5 彭辉, 张长水, 荣钢等. 基于 K-L 变换的人脸自动识别方法. 清华大学学报(自然科学版), 1997, 37(3): 67~ 70.
- 6 金忠, 杨静宇, 陆建峰. 一种具有统计不相关性的最佳鉴别矢量集. 计算机学报, 1999, 22(10): 1~ 4.
- 7 程正兴. 小波分析算法与应用. 西安: 西安交通大学出版社, 1998.



**赵海涛** 现在南京理工大学计算机系攻读博士学位. 感兴趣的研究方向为图象处理、人脸识别等.

**金忠** 南京理工大学应用数学系教师, 主要的研究方向为概率统计、图象处理、人脸识别、手写数字识别等.

**杨静宇** 南京理工大学计算机系教授、博士生导师, 主要研究方向为图象处理、模式识别等.