

基于小波特征的非线性鉴别特征抽取技术

陈才扣^{***} 高秀梅^{*} 郭丽^{*} 杨静宇^{*}

^{*}(南京理工大学计算机系 南京 210094)

^{**}(扬州大学信息工程学院 扬州 225009)

摘要: 该文提出了一种基于小波特征的非线性鉴别特征抽取方法, 即在进行非线性映射之前, 首先利用小波变换对原始输入图像进行预处理, 获取低频平滑、水平细节和垂直细节等 3 个子图的小波特征, 然后在频域上, 对它们分别进行核 Fisher 鉴别分析。对最终获得的 3 组鉴别特征设计了一种特征融合的方法。在 ORL 标准人脸库上的试验结果表明所提方法不仅在识别性能上优于现有的核 Fisher 鉴别分析方法, 而且, 在 ORL 人脸库上的特征抽取速度提高了近 13 倍。

关键词: 人脸识别, 核 Fisher 鉴别分析, 小波变换, 特征抽取

中图分类号: TN391.4 **文献标识码:** A **文章编号:** 1009-5896(2005)02-0290-04

Wavelet Feature-Based Nonlinear Feature Extraction Technique

Chen Cai-kou^{***} Gao Xiu-mei^{*} Guo Li^{*} Yang Jing-yu^{*}

^{*}(Dept of Computer Science, Nanjing University of Science and Technology, Nanjing 210094, China)

^{**}(Information and Engineering College, Yangzhou University, Yangzhou 225001, China)

Abstract The paper develops a novel nonlinear feature extraction method based on wavelet features. Its main idea is that wavelet transform is first employed to preprocess the original training images before the nonlinear mapping and three groups of wavelet features: lowest frequency subimage, horizontal detail and vertical detail, are derived respectively. What follows, Kernel Fisher Discriminant Analysis(KFDA) is performed on three classes of wavelet features. Three final discriminant feature vectors are obtained, from which a feature fusing method is developed. Finally, The experimental results on ORL face databases indicate that the proposed method is more effective than the current KFDA. And, more importantly, its consumed time in feature extraction is only one thirteenth of that of KFDA. Moreover, the experiments also demonstrate that this method is robust in uncontrolled lighting condition.

Key words Face recognition, Kernel Fisher discriminant analysis, Wavelet transform, Feature extraction

1 引言

近年来小波分析由于其优良的时频域局部特性使其成为一种有效的信号/图像分析和处理方法^[1-3]。高西奇等人^[1]提出了基于小波变换的主元分析人脸图像识别算法, 该方法与直接对图像进行主分量分析相比, 不仅可以提高运算速度, 而且提取的特征更加反映各类人脸之间的差异。Sergent^[2]指出低频域和高频域有不同的作用, 低频部分刻画的是图像的整体(形状), 高频部分包含了相当数量的细节信息。Liu 等人^[3]提出人脸的形状(shape)图像与纹理(texture)图像所包含的鉴别信息是互补的。其实, 分辨率较高的原始图像包含了大部分的纹理信息, 而分辨率较低的平滑图像包含了相当数量的形状信息。

基于核的非线性特征抽取技术^[4-9]最近受到了广泛关注,

正成为当前模式识别领域中一个迅猛发展的新方向。它的主要思想最初是由 Vapnik^[9]提出并应用于支持向量机(SVM)中。其中, 最著名的两种基于核的非线性特征抽取技术分别是核主分量分析(KPCA)和核 Fisher 鉴别分析(KFDA)。KPCA 首先由 Scholkopf 等人^[4]提出并应用于特征抽取中。之后, Mika 等人^[5], Baudat 和 Anouar^[6]以及 Roth 和 Steinhage^[7]利用核方法将 Fisher 线性鉴别分析进一步推广到非线性情形, 提出了 KFDA 方法。Mika 等人的工作主要解决两类模式的分类问题, 而 Baudat 他们的算法可应用于多类模式分类问题。Yang^[8]将 KPCA 与 KFDA 相结合应用于人脸识别中, 提出了核特征脸(kernel eigenfaces)和核 Fisher 脸(kernel Fisherfaces)。尽管核方法在图像等模式识别领域应用的非常广泛和成功, 但也存在如下一些问题: (1) 构造特征空间 H

中的核矩阵 K 所耗费的计算量非常大。因为核函数 $k(x_i, x_j)$ 需要计算两个向量间的内积, 当向量的维数很高时 (如人脸识别中的训练样本的维数 n 普遍高达上万维), 计算 K 所需的计算量为 $O(nN^2)$ (N 是训练样本的数目); (2) 当 N 很大时, 会使得核矩阵的存储空间急剧增加。

本文提出了一种基于小波特征的非线性鉴别特征抽取方法, 即在进行非线性映射之前, 首先利用小波变换对原始输入图像进行预处理, 抽取图像的小波特征; 然后在频域上对低分辨率的小波特征进行核 Fisher 鉴别分析。在 ORL 标准人脸库上的试验结果验证了所提出方法的有效性。

2 二维离散小波变换

2.1 二维离散小波变换的基本概念

假定在 $L^2(R)$ 上有一个多分辨率分析(MRA) $\{V_j\}_{j \in Z}$ 及相应的尺度函数 $\varphi(x)$, 定义 j 尺度下的二维尺度空间 $V_j^2 = V_j \otimes V_j$, 其中符号 \otimes 表示空间相乘, 则由 $\varphi_{j,m}(x) = 2^{-j/2} \varphi(2^j x - m)$ 是 V_j 的标准正交基可知, $\{\varphi_{j,m}(x) \varphi_{j,m}(y)\}_{n,m \in Z}$ 一定是 V_j^2 的标准正交基。用 W_{j+1}^2 表示 V_{j+1}^2 在 V_j^2 中的正交补空间, 用 $\psi(x)$ 表示与 $\varphi(x)$ 相应的一维小波, 由此可以得到 $L^2(R^2)$ 中 3 个小波函数 $\Psi^1 = \varphi \otimes \psi = \varphi(x) \psi(y)$, $\Psi^2 = \psi \otimes \varphi = \psi(x) \varphi(y)$, $\Psi^3 = \psi \otimes \psi = \psi(x) \psi(y)$, 使得 $\{\Psi^i(2^j x_1 - k, 2^j x_2 - l)\}_{j,k,l \in Z, i=1,2,3}$ 构成 $L^2(R^2)$ 的规范正交基。

图 1 给出了 ORL 人脸数据库中的一幅人脸的二层小波分解后各子图像的频率空间位置图以及一个分解示意图。其中, 子图像 LL2 表示低频分量, 为原图的平滑图像; 子图像 LH1 是水平方向低频、垂直方向高频的分量, 表现原图的水平边缘; 子图像 HL1 是水平方向为高频、垂直方向为低频的分量, 表现原图的垂直边缘; 子图像 HH1 是高频分量, 表现原图的斜边缘。因此, 经过 n 层分解后, 图像的低频部分只有原图大小的 $1/4^n$, 而其概貌信息却基本保持不变, 因为图像的大部分信息集中在低频部分。显然, 利用分解后的低频尺度系数矩阵提取特征, 可以大大降低图像特征的维数, 从而节省存储空间, 提高计算效率。

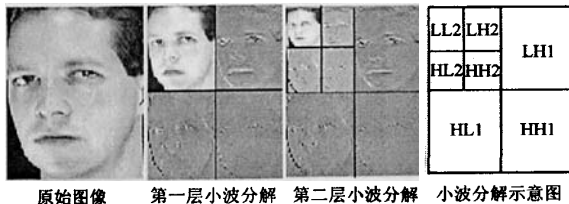


图 1 二维离散小波分解

2.2 小波特征的预处理

统计表明, 原始图像的能量绝大部分都集中在最高层的平滑子图像 LL1 上, 水平细节子图像和垂直细节子图像也包含了部分数量的分类信息, 而对角高频子图像包含的有用分

类信息相对很少。对于图像识别任务的原始图像 (如分辨率为 112×92 的 ORL 人脸数据库中的原始人脸图像) 较多地反映了面部的局部细节信息, 其中包含了较多的光照、遮罩、旋转扭曲和面部表情等局部细节信息, 这些信息体现了模式类内的变化, 对人脸鉴别来说是不利的。所以, 我们对原始图像进行三层小波分解, 选用第三层的低频分量、水平方向和垂直方向的细节分量作为小波特征, 这样既保留了面部的全局形状信息, 同时又淡化了局部细节。

由于小波变换后能力集中在低频平滑子图中, 相比水平和垂直两个细节子图, 其均值要大的多, 而且后两者中的值绝大多数为零, 这种情况对后面的最优鉴别特征抽取是不利的。为了消除这 3 组特征在数值上的非均衡性造成的不利影响, 需要对它们进行如下的标准化预处理: $y = (x - \mu) / \sigma$, 其中, x 为 m 维特征样本, μ 为训练样本的均值向量, $\sigma = \sqrt{\sum_{j=1}^n \sigma_j^2}$, σ_j 为训练样本的第 j 个特征分量的标准差。

3 核 Fisher 鉴别分析与特征融合

3.1 核 Fisher 鉴别分析

核 Fisher 鉴别分析的主要思想就是首先将原始输入空间线性不可分的训练样本通过某非线性映射 ϕ (由预先选定的核函数隐式地确定) 变换到某一高维特征空间 H , 使得变换后的训练样本在新的特征空间上具有更好的可分性, 然后在征空间 H 上作线性鉴别分析(LDA)。为了避免直接显式地处理映射后的样本, 引入支持向量机中的核函数, 即著名的核技巧 (kernel trick), 这些核函数可以计算某些特征空间上的内积, 即 $k(X, Y) = \langle \phi(X), \phi(Y) \rangle$ 。在支持向量机中常用的核函数有高斯 RBF 核函数 $k(X, Y) = \exp(-\|X - Y\|^2 / \sigma)$, 多项式核函数 $k(X, Y) = (X \cdot Y + c)^d$, 其中 σ, c 和 d 都是常数。经过非线性映射之后, 则映射后的训练样本在高维特征空间 H 上的类内散布矩阵 S_w^ϕ , 类间散布矩阵 S_b^ϕ 和总体散布矩阵 S_t^ϕ 分别为

$$S_b^\phi = \sum_{i=1}^c \frac{N_i}{N} (m_i^\phi - m^\phi)(m_i^\phi - m^\phi)^T \quad (1)$$

$$S_w^\phi = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} (\phi(X_j) - m_i^\phi)(\phi(X_j) - m_i^\phi)^T \quad (2)$$

$$S_t^\phi = S_b^\phi + S_w^\phi = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} (\phi(X_i) - m^\phi)(\phi(X_i) - m^\phi)^T \quad (3)$$

其中 $m^\phi = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} \phi(X_j)$, $m_i^\phi = \frac{1}{N_i} \sum_{j=1}^{N_i} \phi(X_j)$, $i = 1, \dots, c$, c 是训练样本的类数。

在特征空间 H 上相应的 Fisher 鉴别准则函数定义为

$$J_f^\phi(w) = w^T S_b^\phi w / (w^T S_w^\phi w) \quad (4)$$

其中 w 是特征空间内的任意非零向量。

为了在特征空间 H 中寻找最优的线性鉴别向量 (对应于原始空间中的非线性鉴别向量), 需要得到仅包含内积或核

函数的鉴别准则函数形式。根据再生核理论 (Theory of reproducing kernel) [9], 任何解向量 w 一定位于由 $\{\phi(X_1), \dots, \phi(X_N)\}$ 张成的子空间内, 即

$$w = \sum_{i=1}^N \alpha_i \phi(X_i) = \Phi \alpha \quad (5)$$

其中 $\Phi = [\phi(X_1), \dots, \phi(X_N)]$, $\alpha = [\alpha_1, \dots, \alpha_N]^T \in R^N$ 。

将式(1)、式(2)、式(3)和式(5)代入式(4), 经过一定的矩阵变换, 则得到 $K_b = \sum_{j=1}^c N_j (\frac{1}{N_j} K_j I - \frac{1}{N} K I) (\frac{1}{N_j} K_j I - \frac{1}{N} K I)^T$, $K_w = \sum_{j=1}^c K_j (I - \frac{1}{N_j} 11^T) K_j^T$, $K_f = K (I - \frac{1}{N} 11^T) K^T$, 其中, $K = (K_{ij})_{i,j=1,\dots,N}$, $K_{ij} = k(X_i, X_j)$; $K_j = (K_{jk})_{k=1,\dots,N}$, $(K_{jk})_k = k(X_b, X_k) = \phi(X_k)^T \phi(X_j)$; $\mathbf{1}$ 表示其元素都为 1 的向量, I 是单位阵。

于是得到特征空间 H 上的 Fisher 鉴别准则函数的等价形式:

$$J_f^*(\alpha) = \frac{w^T S_b^* w}{w^T S_w^* w} = \frac{\alpha^T K_b \alpha}{\alpha^T K_w \alpha} \quad (6)$$

由广义 Rayleigh 商的极值性质[10], 使得准则式(6)极大化的最优解向量集为以下广义特征方程 $K_b \alpha = \lambda K_w \alpha$ 的 d 个最大特征值所对应的特征向量 $\alpha^1, \alpha^2, \dots, \alpha^d$ 。

3.2 特征融合

将上节获得的 3 组小波特征分别进行核 Fisher 鉴别分析, 抽取特征空间上的最优鉴别特征。设 X 为一原始训练样本, 设 W_1, W_2 和 W_3 为 3 个小波变换矩阵, A_1, A_2 和 A_3 为 3 组小波特征执行 KFDA 后的变换矩阵, 则有

$$Y_1 = A_1^T \phi(W_1^T X) \quad (7)$$

$$Y_2 = A_2^T \phi(W_2^T X) \quad (8)$$

$$Y_3 = A_3^T \phi(W_3^T X) \quad (9)$$

然后利用如下的串行组合形式构造新的鉴别特征:

$$Y = \begin{pmatrix} \beta_1 Y_1 \\ \beta_2 Y_2 \\ \beta_3 Y_3 \end{pmatrix} \quad (10)$$

其中 $\beta_1 + \beta_2 + \beta_3 = 1$, 为了强化全局形状信息, 淡化局部细节, 要求 β_1 的值远大于 β_2 和 β_3 。并且当 β_2 和 β_3 等于 0 时, 融合特征向量 Y 等价于单特征 Y_1 。

4 算法实现

- (1) 利用二维离散小波变换对原始输入训练样本 $x_i \in R^n$ ($i=1, \dots, N$) 进行 3 层小波分解, 提取 3 组小波特征;
- (2) 对每组小波特征分别进行核 Fisher 鉴别分析, 获得特征空间上的 3 组鉴别特征;
- (3) 利用式(10)计算每个原始样本的最终鉴别特征向量;
- (4) 利用最近邻分类器或最小距离分类器进行分类识别。

下面给出本算法的结构框图如图 2 所示。

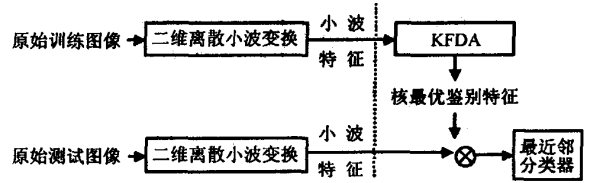


图 2 人脸识别的算法结构图

5 试验与分析

ORL 标准人脸库由 40 人, 每人 10 幅 112×92 图像组成, 其中有些图像是拍摄于不同时期的; 人的脸部表情和脸部细节有着不同程度的变化, 比如, 笑或不笑, 眼睛或睁或闭, 戴或不戴眼镜; 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。

本试验中, 我们每次随机选取每人的 5 幅图像作为训练样本, 剩余的 5 幅作为测试样本, 这样训练样本和测试样本总数均为 200。算法中小波变换使用 Haar 小波函数, KFDA 选用参数 $\sigma = 1.5 \times 10^8$ 的高斯核函数。第 1 步对 200 个原始训练图像进行 3 层小波分解; 第 2 步, 作为比较, 对第 1 层、第 2 层和第 3 层的 3 组小波特征 (为简化起见, 这 3 组特征分别表示为 1-waveletface、2-waveletface 和 3-waveletface) 分别进行核 Fisher 鉴别分析, 经特征融合后, 最终分别得到 39 个最优鉴别向量。作为比较, 针对相同的原始样本, 同样数量的训练样本和测试样本, 不进行小波分解, 直接利用 KFDA 对原始训练样本进行最优非线性鉴别特征抽取, 也得到 39 个最优鉴别向量。然后在每个投影空间内, 分别利用最近邻分类器、最小欧氏距离分类器进行分类识别。为了验证本文方法对样本选取的鲁棒性, 分别对本文方法和 KFDA 进行 100 次试验, 将 100 次的识别结果进行平均, 作为两种方法最终的识别率。识别结果分别见表 1、表 2 和表 3。

从表 1, 我们看到, 与 KFDA 相比, 本文方法识别率非常的稳定, 最低为 95%, 最高为 99%, 并且大部分集中在 96%~97.5%。而 KFDA 方法的波动比较大, 最低为 93.0%, 最高 98.5%, 从实验结果可以初步地说明本文方法对样本的选取具有鲁棒性, 至于其原因部分地可以从下面的分析中找到, 至于其理论依据将是我们下一步研究的内容。从表 2 可见, 本方法识别性能优于 KFDA 方法, 这验证了降维后的小波特征降低了模式内的变化, 更加反映各类人脸之间的差异, 提高了图像的鉴别能力。并且我们可以看到, 对于 ORL 人脸数据库, 第 3 层小波特征达到最优的识别性能, 因为继续分解将会丢失大量有用的鉴别信息。而且本方法对于光照、表情等模式内变化具有良好的鲁棒性。表 3 可看出, 我们的方法 (如 3-waveletface+KFDA) 的特征抽取速度是 KFDA 的近 13 倍。这是因为第 3 层的小波特征的维数是 $n = 14 \times 12 = 168$, 与原始的人脸图像的维数 $n(n=10304)$ 相比, 降低了近 2 个数量级, 而构造特征空间上的散布矩阵如 K_w 所耗费的时间复杂度为 $O(n^2)$, 故本工具具有突出的速度优势。

表 1 3-waveletface+KFDA 和 KFDA 在最近邻分类器上的 100 次试验不同识别率出现的次数

识别率 (%)	93.0	93.5	94.0	94.5	95.0	95.5	96.0	96.5	97.0	97.5	98.0	98.5	99.0
KFDA	1	2	1	2	5	5	11	15	22	24	7	5	0
3-wavelet+KFDA					2	3	2	5	22	29	20	10	7

表 2 ORL 人脸库上 Waveletface+KFDA 和 KFDA 在两个分类器上的正确识别率

	KFDA	1-waveletface+KFDA	2-waveletface+KFDA	3-waveletface+KFDA
最近邻	0.967	0.962	0.969	0.975
最小距离	0.956	0.935	0.961	0.969

表 3 ORL 人脸库上 Waveletface+KFDA 和 KFDA 特征抽取和分类时间对照表

	特征抽取时间(s)	分类时间(s)	总时间(s)
KFDA	650.79	25.16	675.95
1-waveletface+KFDA	308.21	25.04	330.25
2-waveletface+KFDA	69.09	25.11	94.20
3-waveletface+KFDA	52.21	25.12	77.33

6 结论

利用小波变换提取的特征更加反映各类人脸之间的差异, 而核方法可以抽取非线性特征。本文结合两者的优点, 提出了一种新的有效的非线性鉴别特征抽取方法, 即首先利用小波变换对原始输入图像进行预处理, 抽取图像的小波特征, 然后在频域上进行核 Fisher 鉴别分析。在 ORL 标准人脸库上的试验结果表明所提出的方法不仅在识别性能上优于现有的核 Fisher 鉴别分析方法, 而且, 特征抽取的速度提高了近 13 倍。

参 考 文 献

[1] 高西奇, 周洪祥, 何振亚. 基于小波变换的主元分析人脸图像识别. 东南大学学报, 1996, 26(2): 137 - 141.

[2] Sergent J. Microfenesis of face perception. in: H.D.Ellis. Aspects of Face Processing, Nijhoff, Dordrecht, 1986: 80 - 104.

[3] Liu C J, Wechsler Harry. A shape- and texture-based enhanced Fisher classifier for face recognition. *IEEE Trans. on Image Processing*, 2001, 10(4): 598 - 608.

[4] Schölkopf Bernhard, Smola Alexander, Muller K R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 1998, 10: 1299 - 1319.

[5] Mika Sebastian, Ratsch Gunnar, Weston Jason, Schölkopf Bernhard, et al.. Fisher discriminant analysis with kernels. In

IEEE workshop on Networks for Signal Processing, 1999: 41 - 48.

[6] Baudat G, Anouar F. Generalized discriminant analysis using a kernel approach. *Neural Computation*, 2000, 12(10): 2385 - 2404.

[7] Roth Volker, Steinhage Volker. Nonlinear discriminant analysis using kernel functions. In Proc. of Neural Information Processing systems, Denver, Nov. 1999: 568 - 574.

[8] Yang M H. Kernel eigenfaces vs. kernel fisherfaces: face recognition using kernel methods, Proc. of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition(RGR'02), Washington DC, May 2002: 215 - 220.

[9] Vapnik V. The Nature of Statistical Learning Theory. New York: Springer-Verlag, 1995: 302 - 309.

[10] 程云鹏. 矩阵论. 西安: 西北工业大学出版社, 1999: 294 - 302.

陈才扣: 男, 1967 年生, 讲师, 在职博士生, 研究方向为模式识别、人脸检测与识别、基于内容的图像检索。

高秀梅: 女, 1968 年生, 讲师, 在职博士生, 研究方向为模式识别、人脸检测与识别、图像处理。

郭 丽: 女, 1975 年生, 讲师, 博士, 研究方向为模式识别、基于内容的图像检索。

杨静宇: 男, 1941 年生, 教授, 博士生导师, 主要研究方向为计算机视觉、信息融合、模式识别、智能机器人。