一种基于奇异值分解的特征抽取方法

王文胜[*,**]　　陈伏兵[*]　　杨静宇[*]
[*](南京理工大学计算机系　南京　210094)
[**](中国电子科技集团公司第 28 研究所　南京　210007)

摘　要： 特征抽取是模式识别的基本问题之一，Fisher 线性鉴别分析是特征抽取中最为经典和广泛使用的方法之一。该文分析了 Fisher 线性鉴别分析在求解过程中可能存在的问题：鉴别矢量的分量可能是复数；特征值对扰动的敏感性；鉴别矢量之间未必具有正交性。由此提出了均衡散布矩阵的概念，并利用均衡散布矩阵构造了一种新的线性鉴别准则。利用奇异值分解定理，将求取鉴别矢量转化为对矩阵求奇异向量。用该方法进行求解可以有效地避免前述的问题。试验结果表明，该鉴别准则具有良好的鉴别能力。

关键词： 模式识别，特征抽取，线性鉴别分析，奇异值分解，人脸识别

中图分类号： TN391.4　　**文献标识码：** A　　**文章编号：** 1009-5896(2005)02-0294-04

A Method of Feature Extraction Based on SVD

Wang Wen-sheng* **　　Chen Fu-bing*　　Yang Jing-yu*
*(*Dept of Computer Science, Nanjing University of Sci. & Tech., Nanjing* 210094, *China*)
**(28th *Institute, China Electronic Technology Group Corporation, Nanjing* 210007, *China*)

Abstract Feature extraction is primary problem of pattern recognition. As one of the most classic methods in the field of feature extraction, Fisher linear discriminant analysis is applied widely. It may meet several possible problems in finding optimal set of discriminant vectors: the components of these vectors may not be real; the eigenvalue may be sensitive; these vectors may not be orthogonal each other. So the balanced scatter matrix is proposed in this paper. Based on the matrix, a discriminant criterion is formed. The optimal set of discriminant vectors can be acquired througn singular value decomposition theorem. The method can avoid the problems metioned above. The result of face recognition experiment shows that it has powerful ability of feature extraction.

Key words Pattern recognition, Feature extraction, Linear discriminant analysis, Singular value decomposition, Face recognition

1　引言

特征抽取是模式识别研究的基本问题，它的基本任务是找出对分类最有效的特征，通常需要用一个定量的准则衡量特征对分类的有效性。Fisher 线性鉴别分析[1-4]是特征抽取中最为经典和广泛使用的方法之一，它通过将高维空间的特征经过线性变换投影到低维空间中去，选择具有最小类内散布和最大的类间散布的投影方向，抽取出有效鉴别特征。

设模式 $\boldsymbol{X}$ 为 n 维矢量，模式类别有 m 个，模式类内散布矩阵为 $\boldsymbol{S}_w$，类间散布矩阵为 $\boldsymbol{S}_b$。Fisher 鉴别准则函数定义为

$$J_f(\boldsymbol{\varphi})=\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{S}_b\boldsymbol{\varphi}\,/\,\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{S}_w\boldsymbol{\varphi} \tag{1}$$

其中 $\boldsymbol{\varphi}$ 为任一 n 维非零列向量。

在 Fisher 思想的基础上，Wilks[2] 、Duda[3] 等分别提出了鉴别矢量集的概念，即寻找一组鉴别矢量构成子空间，以原始样本在该子空间内的投影矢量作为鉴别特征用于识别。该方法被称为经典的 Fisher 线性鉴别分析方法。该方法旨在通过最优化准则函数式(2)找到一个最优的投影矩阵 $\boldsymbol{W}_{\mathrm{opt}}$。

$$J_c(\boldsymbol{W})=\det(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{S}_b\boldsymbol{W})/\det(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{S}_w\boldsymbol{W}) \tag{2}$$

可以证明[2,4]，当投影矩阵的列向量取为矩阵 $\boldsymbol{S}_w^{-1}\boldsymbol{S}_b$ 的 d 个最大的特征值所对应的特征向量时比值 $J_c(\boldsymbol{W})$ 最大。

利用矩阵 $\boldsymbol{S}_w^{-1}\boldsymbol{S}_b$ 求鉴别矢量集，可能存在的问题为：（1）其特征值和特征向量的分量都有可能是复数；（2）特征值的敏感性：即如果矩阵作很小的扰动，那么其特征值的扰动可能很大[5]；（3）求取的鉴别矢量之间未必是正交的。

对上述问题的一种解决方式是利用广义特征方程 $\boldsymbol{S}_b\boldsymbol{X}=\lambda\boldsymbol{S}_w\boldsymbol{X}$ 进行求解[6]。该方法可避免特征值和特征向量

2003-09-26 收到，2004-02-12 改回
国家自然科学基金(60072034)资助课题

是复数的情况，但不能保证鉴别矢量之间是正交的，同时广义特征值仍然可能存在敏感性问题。

Foley-Sammon 提出采用一组满足正交条件的最佳鉴别矢量集进行特征抽取的方法[7] (简称 FS 方法)。 该方法旨在寻找一组最优鉴别矢量集 $\varphi_1,\cdots,\varphi_d$，它们在最大化 Fisher 准则函数的同时满足以下正交条件：

$$\varphi_i^{\mathrm{T}}\varphi_j = 0,\quad \forall i \neq j,\quad i,j = 1,\cdots,d \tag{3}$$

FS 方法求出的鉴别矢量彼此之间是正交的，但该方法仍然需要通过求解广义特征方程来实现，同样存在敏感性问题。

本文提出一种方法，可以避免上述 3 个方面的问题。

2　均衡散布矩阵

在 Fisher 线性鉴别准则中利用类内散布情况来平衡类间散布情况，使得变换后的类间散布变大的同时类内散布不至于同时变得过大。借鉴该思想，如果 $\boldsymbol{S}_w$ 非奇异，定义：

$$\boldsymbol{S}_b^* = (\boldsymbol{S}_w^{-1}\boldsymbol{S}_b)(\boldsymbol{S}_w^{-1}\boldsymbol{S}_b)^{\mathrm{T}} = \boldsymbol{S}_w^{-1}\boldsymbol{S}_b^2\boldsymbol{S}_w^{-1} \tag{4}$$

用该矩阵来描述经过均衡化后的类间散布程度，其实质就是对类间散布矩阵作加权处理。如果 $\boldsymbol{S}_w$ 奇异，而 $\boldsymbol{S}_t$ 非奇异，将式(4)中的 $\boldsymbol{S}_w^{-1}$ 换成 $\boldsymbol{S}_t^{-1}$，以下的结论仍然成立。称由式(4)定义的矩阵 $\boldsymbol{S}_b^*$ 为均衡散布矩阵。

显然 $\boldsymbol{S}_b^*$ 是对称阵。除此之外，矩阵 $\boldsymbol{S}_b^*$ 还具有下述性质：

性质 1　矩阵 $\boldsymbol{S}_b^*$ 是非负定矩阵；其特征值是非负实数，特征向量的分量都是实数。

将 $\boldsymbol{S}_b^*$ 的特征值按由大到小的顺序排列为

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$

相应的标准正交的特征向量记为：$\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$。

设 $\boldsymbol{\Phi} = \mathrm{span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_r\}$ 是由 $\boldsymbol{S}_b^*$ 的所有非零特征值对应的特征向量生成的空间，$\boldsymbol{\Phi}^{\perp} = \mathrm{span}\{\boldsymbol{u}_{r+1},\cdots,\boldsymbol{u}_n\}$ 是由 $\boldsymbol{S}_b^*$ 的所有零特征值对应的特征向量生成的空间，显然它是 $\boldsymbol{\Phi}$ 的正交补空间。

性质 2　$\forall \boldsymbol{x} \in \boldsymbol{R}^n$，$\boldsymbol{S}_b^*\boldsymbol{x} \in \boldsymbol{\Phi}$。

性质 3　$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{S}_b^*\boldsymbol{x} = 0$ 当且仅当 $\boldsymbol{S}_b^*\boldsymbol{x} = 0$。

性质 4　矩阵 $\boldsymbol{S}_b^*$ 的特征向量是矩阵 $\boldsymbol{S}_w^{-1}\boldsymbol{S}_b$ 的左奇异向量。该性质可以从下面的奇异值分解定理得出。

奇异值分解定理[5,8]：设 $A \in \boldsymbol{R}_r^{m\times n}(r>0)$，则存在 m 阶正交矩阵 $\boldsymbol{U}$ 和 n 阶正交矩阵 $\boldsymbol{V}$，使得 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{V} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$，其中 $\Sigma = \mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_r)$，$\sigma_i(i=1,2,\cdots,r)$ 为矩阵 $\boldsymbol{A}$ 的全部非零奇异值，$\boldsymbol{U}$ 的列向量是 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 的特征向量，$\boldsymbol{V}$ 的列向量是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的特征向量。

为方便起见，称定理中奇异值对应的的矩阵 $\boldsymbol{U}$ 的列向量为左奇异向量。

性质 5　如果矩阵 $\boldsymbol{S}_b^*$ 被一个对称矩阵 $\boldsymbol{E}$ 扰动，则它的特征值变化范围不超过 $\|\boldsymbol{E}\|_2$；如果 $\boldsymbol{S}_w^{-1}\boldsymbol{S}_b$ 被一个矩阵 $\boldsymbol{E}$ 扰动，则它的奇异值变化范围不超过 $\|\boldsymbol{E}\|_2$。这一性质可以从以下的引理 1 得出。

引理 1[5]　若 $\boldsymbol{A}$ 和 $\boldsymbol{A}+\boldsymbol{E}$ 是 $n\times n$ 对称矩阵，则对 $k=1,2,\cdots,n$，有

$$|\lambda_k(\boldsymbol{A}+\boldsymbol{E}) - \lambda_k(\boldsymbol{A})| \leq \|\boldsymbol{E}\|_2$$

其中 $\lambda_k(\boldsymbol{A}+\boldsymbol{E}),\lambda_k(\boldsymbol{A})$ 分别对应矩阵 $\boldsymbol{A}+\boldsymbol{E},\boldsymbol{A}$ 的第 k 个特征值（特征值按从大到小的顺序排列）；

若 $\boldsymbol{A}$ 和 $\boldsymbol{A}+\boldsymbol{E}$ 均属于 $\boldsymbol{R}^{m\times n}, m \geq n$，则对 $k=1,2,\cdots,n$，有

$$|\sigma_k(\boldsymbol{A}+\boldsymbol{E}) - \sigma_k(\boldsymbol{A})| \leq \|\boldsymbol{E}\|_2$$

其中 $\sigma_k(\boldsymbol{A}+\boldsymbol{E}),\sigma_k(\boldsymbol{A})$ 分别对应矩阵 $\boldsymbol{A}+\boldsymbol{E},\boldsymbol{A}$ 的第 k 个奇异值（奇异值按从大到小的顺序排列）。

3　基于均衡散布矩阵的线性鉴别准则

定义　$J(\boldsymbol{\varphi}) = \boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{S}_b^*\boldsymbol{\varphi} / \boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{\varphi}$，我们采用：$\max\limits_{\boldsymbol{\varphi}\neq 0} J(\boldsymbol{\varphi})$ 作为特征抽取有效性的鉴别准则。称使得 $J(\boldsymbol{\varphi}) > 0$ 的非零矢量 $\boldsymbol{\varphi}$ 为有效鉴别矢量。函数 $J(\boldsymbol{\varphi})$ 具有以下性质：

性质 1　$J(\boldsymbol{\varphi}) = 0$，当且仅当 $\boldsymbol{\varphi} \in \boldsymbol{\Phi}^{\perp}$；

性质 2　若 $\boldsymbol{\varphi} \in \boldsymbol{\Phi}$，则 $J(\boldsymbol{\varphi}) > 0$。

由 Rayleigh 商的性质可知，函数 $J(\boldsymbol{\varphi})$ 的梯度为 0 当且仅当 $\boldsymbol{\varphi}$ 为 $\boldsymbol{S}_b^*$ 的一特征向量，$J(\boldsymbol{\varphi})$ 在稳定点的值为 $\boldsymbol{S}_b^*$ 的特征值；并且 $J(\boldsymbol{\varphi})$ 的最大值可在单位球面上达到，故上述准则等价于：

$$\max_{\|\boldsymbol{\varphi}\|_2=1} J(\boldsymbol{\varphi}) \tag{5}$$

因此使得式(5)成立的鉴别矢量 $\boldsymbol{\varphi}$ 必定为 $\boldsymbol{S}_b^*$ 的最大特征值对应的单位特征向量。

在求出第 $i\,(i\geq 1)$ 个最佳鉴别矢量 $\boldsymbol{\varphi}_1,\boldsymbol{\varphi}_2,\ldots,\boldsymbol{\varphi}_i$ 后，第 $i+1$ 个最佳鉴别矢量为满足下述正交条件式(6)并使 $J(\boldsymbol{\varphi})$ 取到最大值的单位向量 $\boldsymbol{\varphi}_{i+1}$：

$$\boldsymbol{\varphi}_{i+1}^{\mathrm{T}}\boldsymbol{\varphi}_j = 0,\quad j = 1,2,\cdots,i \tag{6}$$

为了进一步寻找鉴别矢量，我们需要下述引理 2：

引理 2[8]　设 $\boldsymbol{\varphi} \in \mathrm{span}(\boldsymbol{u}_l,\boldsymbol{u}_{l+1},\cdots,\boldsymbol{u}_m)$, $1 \leq l \leq m \leq n$，则有：$\max\limits_{|\boldsymbol{\varphi}|=1} J(\boldsymbol{\varphi}) = \lambda_l$

下面的定理1和定理2给出了求取最优鉴别矢量集的方法。

定理 1　若前 i 个最优鉴别矢量分别取为 $\varphi_1 = u_1, \cdots, \varphi_i = u_i$，则第 $i+1$ 个最优鉴别矢量 φ_{i+1} 可取为 u_{i+1}；并且按这种方式选取的有效鉴别矢量的最大个数为 r。

证明　若前 i 个最优鉴别矢量分别取为 $\varphi_1 = u_1, \cdots, \varphi_i = u_i$，由正交性条件知，$\varphi_{i+1}$ 只可能从 R^n 的子空间 $\mathrm{span}\{u_1, \ldots, u_i\}$ 的正交补空间 $\mathrm{span}\{u_{i+1}, \cdots, u_n\}$ 中选取，由引理2，$\max\limits_{|\varphi_{i+1}|=1} J(\varphi_{i=1}) = \lambda_{i+1}$。而 $J(u_{i+1}) = \lambda_{i+1}$，因此 φ_{i+1} 可取 u_{i+1}。这样选取的有效鉴别矢量的最大个数为 r 是显然的。

证毕

由定理1和矩阵 S_b^* 的性质4，立即可得：

定理 2　使得 $J(\varphi)$ 最大并满足正交性条件的最优鉴别矢量集可取为矩阵 $S_w^{-1} S_b$ 的所有左奇异向量对应的正交向量集。

4　实验结果

采用ORL人脸图像数据库进行人脸识别实验。ORL人脸图像数据库由40人、每人10幅分辨率为92×112的灰度图像组成。用每人10幅图像样本中的其中5幅作为训练样本，剩下的5幅作为测试样本，这样训练样本和测试样本的总数都是200。为了使得训练样本的抽样具有随机性，我们对每人的10幅图像按1～10进行编号，按表1分配训练样本和测试样本。

表1　训练样本和测试样本的选择

序号	训练样本编号	测试样本编号
1	（1，2，3，4，5）	（6，7，8，9，10）
2	（6，7，8，9，10）	（1，2，3，4，5）
3	（1，3，5，7，9）	（2，4，6，8，10）
4	（2，4，6，8，10）	（1，3，5，7，9）
5	（4，5，6，7，8）	（1，2，3，9，10）

由于每幅人脸图像矢量的维数为92×112＝10304，而样本总数为400，因此本实验属于典型的高维小样本识别问题。在原始图像空间，类内散布矩阵和总体散布矩阵都是奇异的。以总体散布矩阵 S_t 作为KL坐标系的产生矩阵，对原始图像做KL变换，将维数压缩为200－40＝160。在变换后的空间内，类内散布矩阵 S_w 非奇异，类间散布矩阵的秩为 $\mathrm{rank}(S_b) = 39$。

实验选用文献[9]中的求特征向量方法、文献[6]中的求广义特征向量方法、FS方法和本文方法作对比，采用最小距离分类器进行分类。对ORL数据库的平均识别错误率和平均识别时间如表2所示。

表2　不同方法对ORL人脸数据库的平均识别错误率比较（%）

鉴别矢量维数	文献[9]方法	文献[6]方法	FS方法	本文方法
20	9.4	10.5	6.3	5.7
21	9.7	10.0	6.0	5.3
22	9.8	10.0	5.1	5.5
23	9.6	10.5	4.8	5.5
24	9.9	10.4	5.0	5.4
25	10.5	10.4	4.8	5.6
26	11.0	10.0	5.1	5.1
27	10.6	10.6	4.8	4.8
28	10.1	9.3	4.7	4.6
29	10.1	9.7	4.1	4.5
30	9.9	9.1	4.2	4.3
31	9.1	8.9	3.7	4.2
32	8.9	8.9	3.6	4.1
33	8.8	9.0	3.6	4.1
34	8.8	8.2	3.3	3.7
35	8.7	7.8	3.3	3.8
36	8.6	7.9	3.3	3.2
37	8.3	8.0	3.7	3.6
38	7.9	7.5	3.9	3.7
39	7.1	6.9	4.3	3.5
平均识别时(s)	4.507	4.31	42.366	4.287

5　结论

实验结果表明：本文方法的识别效果优于文献[9]的求特征向量方法和文献[6]的求广义特征向量方法，与FS方法的识别效果相近，但是其识别时间比FS方法要少的多，并且识别错误率基本呈下降趋势，而FS方法的识别错误率有上升的趋势。理论和实验都表明本文提出的均衡散布矩阵具有一系列良好的性质，基于均衡散布矩阵的线性鉴别准则在求解鉴别矢量时避免了在引言中提到的各种问题，并巧妙地将选择鉴别矢量的问题转化为求矩阵的奇异值。这些特点使得本文的方法具有良好的应用前景。

参考文献

[1] Fisher R A. The use of multiple measurements in taxonomic problems. *Ann. Eugenics*, 1936, 7: 179－188.

[2] Wilks S S. Mathematical Statistics. New York: Wiley, 1962: 577－578.

[3] Duda R O, Hart P E, Stork D G. Pattern Classification, 2nd Edition. New York: John Wiley & Sons, 2001: 117 – 123.
[4] Fisher R A. The Statistical Utilization of Multiple Measurements. *Ann. Eugenics*, 1938, 8: 376 – 386.
[5] Golub G H, Van Loan C F. 袁亚湘等译. 矩阵计算. 北京：科学出版社, 2001.
[6] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisherfaces - recognition using class specific linear projection. *IEEE Trans. on Pattern Anal. Machine Intell*, 1997, 19(7): 711 – 720.
[7] Foley D H, Sammon J W Jr. An optimal set of discriminant vectors. *IEEE Trans. on Computer*, 1975, 24(3): 281 – 289.
[8] 程云鹏. 矩阵论. 西安:西北工业大学出版社, 1989: 294 – 302.
[9] Martínez A M, Kak A C. PCA versus LDA. *IEEE Trans. on Pattern Anal. Machine Intell*, 2001, 23 (2): 228 – 232.

王文胜：男，1969 年生，博士生，主要研究方向为模式识别、作战辅助决策.
陈伏兵：男，1963 年生，博士生，副教授，主要研究方向为模式识别.
杨静宇：男，1941 年生，教授，博士生导师，主要研究领域为计算机视觉、信息融合、模式识别、智能机器人.

中国计算机学会
第二届全国 Web 信息系统及其应用会议(WISA2005)
征文通知

全国 Web 信息系统及其应用会议（WISA）是中国计算机学会电子政务与办公自动化专委会主办的系列会议。首届会议 WISA2004 于 2004 年 10 月在湖北省武汉市圆满召开，会议共收到 368 篇，其中 54 篇论文在《武汉大学学报（英文）》（EI 源刊）作为正刊专辑发表。WISA2005 将于 2005 年 8 月 5-7 日在沈阳召开。会议将继续这一良好的传统，将在 Web 技术、信息系统、电子政务与办公自动化等方面进行深入广泛的学术交流。会议论文集仍将分两部分出版，录用论文中将选择 50 篇左右高水平论文以英文方式继续由《武汉大学学报（英文）》正刊专辑（Special Issue）出版，中文论文集将由著名计算机核心期刊和中央级出版社出版。会议期间除进行会议论文交流外，还将邀请著名学者作特邀报告。本次会议还将评选大会优秀论文和优秀学生论文。

一、征文范围（包括但不限于）

Web 信息挖掘与检索
语义 Web 与智能 Web
Web 与网格计算
Web 与数据库技术
XML 与半结构化数据管理
Web 信息系统环境与基础
Web 应用框架和体系结构
Web 与信息系统安全性
Web 信息系统开发工具
Web 系统度量与分析技术
Web 站点逆向工程与维护技术
Web 测试与 Web 应用的质量保证
多媒体数据管理
工作流模型
组件与中间件技术
代理技术及信息管理
自动文本索引与分类技术
决策支持与分析技术
电子政务与电子商务框架及应用
电子政务与办公自动化发展现状与趋势

二 来稿要求

1 本次会议只接受 Email 投稿。
2 中英文稿均可，一般不超过 6000 字，为了便于出版论文集，来稿必须附中英文摘要、关键词、资助基金与主要参考文献，注明作者及主要联系人姓名、工作单位、详细通信地址（包括 Email 地址）与作者简介。稿件要求采用 WORD 或 PDF 格式。

三 联系信息

1 论文投稿地址： 中国人民大学信息学院 孟小峰 刘青（qliu@ruc.edu.cn）
2 会务情况： 东北大学信息学院软件所 于戈 赵志滨（zhaozb@mail.neu.edu.cn）
3 大会网站：http://www.neu.edu.cn/wisa2005

四 重要日期

1 征文截至日期：2005 年 4 月 5 日
2 录用通知发出日期：2005 年 4 月 30 日
3 正式论文提交日期：2005 年 5 月 20 日