Foley-Sammon 鉴别矢量集理论分析及优化模型

徐　勇　杨　强　杨静宇
（南京理工大学计算机科学系　南京210094）

The Theory Analysis on FSDVS and an Optimal Model

XU Yong　YANG Qiang　YANG Jing-Yu
(Department of Computer Science, Nanjing University of Science and Technology, Nanjing 210094)

Abstract The attribute of the maximum of $R(\xi)$ in arbitrary subspace of R^n is discussed dedicatedly. The theory analysis indicates that every F-S discriminant vector is better than respective vector in other discriminant vectors sets, which consist of eigenvectors of $s_b\xi=\lambda s_w\xi$. But, the fact that the F-S vectors are statistically correlated degrade the F-S vectors set. Two ways are used to obtain "good" F-S vectors set. The experiments on Concordia University CENPARMI handwritten numeral database suggest that the new vectors set is better than the origin. A new problem model on F-S discriminant vectors set is proposed in this paper. It's easily understanding that the problem model is superior to the original F-S discriminant vectors set problem model.

Keywords Pattern recognition, F-S discriminant vectors set, Feature extraction, Optimal model

1 引言

Fisher 线性变换方法在模式识别中有着重要的作用，其基本思想是寻找一个“最佳”投影方向，使模式投影到该方向后具有最大类间距离与最小类内距离，该方向对应的矢量称为 Fisher 最佳鉴别矢量。Sammon 在模式识别研究中使用的鉴别平面是这一思想的发展[1]。组成鉴别平面的第一个鉴别矢量即为 Fisher 最佳鉴别矢量，第二个鉴别矢量为与 Fisher 最佳鉴别矢量正交且使 Fisher 准则函数具有极大值的矢量。在此基础上，Foley 与 Sammon 进一步提出了最优鉴别矢量集的思想[2]，使用多个正交鉴别矢量来进行两类问题的识别。Duchene 与 Leclercq 于1988年得到了多类识别问题的正交鉴别矢量集的分析解[3]，习惯上称这样的矢量的集合为 F-S 鉴别矢量集。F-S 鉴别矢量集在线性变换方法中占据着重要的位置，近年来，F-S 鉴别矢量集一直吸引着模式识别领域研究者的关注[4~6]。

2 F-S 鉴别矢量集

设有 c 个模式类别 $\omega_1,\omega_2,\cdots,\omega_c$，模式类内散布矩阵 s_w，模式类间散布矩阵 s_b，及总体散布矩阵 s_t 分别为：

$$s_w=\sum_{i=1}^{c}p(\omega_i)E[(X-m_i)(X-m_i)^T|\omega_i] \tag{1}$$

$$s_b=\sum_{i=1}^{c}p(\omega_i)(m_i-m_0)(m_i-m_0)^T \tag{2}$$

$$s_t=E[(X-m_0)(X-m_0)^T]=s_b+s_w \tag{3}$$

其中，m_i 为第 i 类模式的期望，$p(\omega_i)$第 i 类模式的先验概率，一般地取 $p(\omega_i)=1/c$，m_0为模式总体的期望，X 为 n 维列向量，ξ_i 为 F-S 鉴别矢量集中第 i 个鉴别矢量。求解问题可描述为如下数学模型

$$\begin{cases}\max R(\xi)\\ \xi^T\xi_j=0 \quad j=1,2,\cdots,i-1\end{cases} \tag{4}$$

其中，$\xi_1,\xi_2,\cdots,\xi_{i-1}$为已求得的鉴别矢量，

$$R(\xi)=\frac{\xi^Ts_b\xi}{\xi^Ts_w\xi} \tag{5}$$

关于模型(4)的求解问题，有如下定理

定理1[3]　第 i 个 F-S 鉴别矢量 ξ_i 为如下广义特征方程中最大特征值对应的特征向量

$$M\xi=\lambda\xi \tag{6}$$

其中，

$$M=(I-s_w^{-1}[D^{(i-1)}]^T[S^{(i-1)}]^{-1}D^{(i-1)})s_w^{-1}s_b \tag{7}$$

$$D^{(i-1)}=\begin{bmatrix}\xi_1^T\\ \xi_2^T\\ \vdots\\ \xi_{i-1}^T\end{bmatrix} \tag{8}$$

$$S^{(i-1)}=[s_{mn}^{(i-1)}] \quad m,n=1,2,\cdots,i-1 \tag{9}$$

$$s_{mn}^{(i-1)}=\xi_m^Ts_w^{-1}\xi_n \quad m,n=1,2,\cdots,i-1 \tag{10}$$

金忠、杨静宇[7]还给出了如下第 i 个 F-S 鉴别矢量 ξ_i 解的另一形式。

定理2　如下广义特征方程的最大特征值对应的特征向量即为第 i 个 F-S 鉴别矢量 ξ_i

$$ps_b\xi=\lambda s_w\xi \tag{11}$$

其中，

$$p=I-D^T[Ds_w^{-1}D^T]^{-1}Ds_w^{-1} \tag{12}$$

D 同(8)式的 $D^{(i-1)}$

3 广义 Rayleigh 商问题

(5)式中的 $R(\xi)$为广义 Rayleigh，因此，F-S 鉴别矢量的求解即为在约束函数 $\xi^T\xi_j=0$条件下广义 Rayleigh 商极值的求解问题。而极值的大小标志着单个 F-S 鉴别矢量对模式的鉴别能力。F-S 方法是在子空间 $L(\xi_1,\cdots,\xi_{i-1})$的正交补空间中求取第 i 个 F-S 鉴别矢量。无人讨论这种求解方法所得的广义 Rayleigh 商的极值性质。为了透彻了解这个问题，该节专门对这一问题进行理论上的分析。

徐　勇　博士研究生，讲师，研究方向：模式识别、图像处理。**杨　强**　博士研究生，研究方向：模式识别、图像处理。**杨静宇**　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。

引理1[8] 若 $s_b, s_w \in R^n$，s_w 正定，广义特征方程 $s_b\xi=\lambda s_w\xi$ 的特征值为 $\lambda_1 \geqslant \lambda_2 \cdots \geqslant \lambda_n \geqslant 0$，则存在完备标准 s_w 共轭特征向量 $X_1, X_2, \cdots, X_n$ 满足

$$X_i^T s_w X_j=\begin{cases}0(i\neq j)\\1(i=j)\end{cases},\quad X_i^T s_b X_j=\begin{cases}0(i\neq j)\\\lambda_i(i=j)\end{cases} \tag{13}$$

定理3 若广义特征方程 $s_b\xi=\lambda s_w\xi$ 的特征值为 $\lambda_1 \geqslant \lambda_2 \cdots \geqslant \lambda_n \geqslant 0$，则(5)式定义的广义 Rayleigh 商 $R(\xi)$ 对 $\forall\ \xi \in R^n$，有 $\lambda_n \leqslant R(\xi) \leqslant \lambda_1$

证明：由引理1，可令

$$\xi=l_1X_1+l_2X_2+\cdots+l_nX_n \quad \forall\ \xi \in R^n$$

则

$$R(\xi)=\frac{\lambda_1 l_1^2+\lambda_2 l_2^2+\cdots+\lambda_n l_n^2}{l_1^2+l_2^2+\cdots+l_n^2}$$

显然，有

$$\lambda_n \leqslant R(\xi) \leqslant \lambda_1$$

定理4 若 s_w 正定，广义特征方程 $s_b\xi=\lambda s_w\xi$ 的特征值为 $\lambda_1 \geqslant \lambda_2 \cdots \geqslant \lambda_n \geqslant 0$，$X_1, X_2, \cdots, X_n$ 为标准 s_w 共轭特征向量，$s_{k-1}=L(X_1, X_2, \cdots, X_{k-1})$，则 $\max\limits_{\xi \in s_{k-1}^{\perp}} R(\xi)=\lambda_k$

证明：因 $s_{k-1}=L(X_1, X_2, \cdots, X_{k-1})$，

所以 $s_{k-1}^{\perp}=L(X_k, X_{k+1}, \cdots, X_n)$

令

$$\xi=l_kX_k+l_{k+1}X_{k+1}+\cdots+l_nX_n \quad \forall\ \xi \in s_{k-1}^{\perp}$$

则

$$R(\xi)=\frac{\lambda_k l_k^2+\lambda_{k+1} l_{k+1}^2+\cdots+\lambda_n l_n^2}{l_k^2+l_{k+1}^2+\cdots+l_n^2}$$

显然

$$\max_{\xi \in s_{k-1}^{\perp}} R(\xi)=\lambda_k$$

定理5 若 s_w 正定，广义特征方程 $s_b\xi=\lambda s_w\xi$ 的特征值为 $\lambda_1 \geqslant \lambda_2 \cdots \geqslant \lambda_n \geqslant 0$，则

$$\min_{\dim S=n-k+1}\{\max_{\xi \in S} R(\xi)\}=\lambda_k$$

其中，S 为 R^n 的任意 $n-k+1$ 维子空间，$1<k\leqslant n$。

证明：令 $X_1, X_2, \cdots, X_n$ 为 $s_b\xi=\lambda s_w\xi$ 的标准 s_w 共轭特征向量，记

$$s_k=L(X_1, X_2, \cdots X_k)$$

对 $\forall\ S \in R^{n-k+1}$

由于 $\dim S=n-k+1$，$\dim s_k=k$

所以 $\dim S+\dim s_k>n$，$S \cap s_k$ 必然非空

设 $\xi^0 \in S \cap s_k$，可令

$$\xi^0=l_1X_1+l_2X_2+\cdots+l_kX_k$$

则

$$R(\xi^0)=\frac{\lambda_1 l_1^2+\lambda_2 l_2^2+\cdots+\lambda_k l_k^2}{l_1^2+l_2^2+\cdots+l_k^2}$$

显然 $R(\xi^0) \geqslant \lambda_k$

而 $\max\limits_{\xi \in S} R(\xi) \geqslant R(\xi^0)$

故 $\max\limits_{\xi \in S} R(\xi) \geqslant \lambda_k$

由 S 的任意性

$$\min_{\dim S=n-k+1}\{\max_{\xi \in S} R(\xi)\} \geqslant \lambda_k$$

另一方面，若 $S=s_{k-1}^{\perp}$，由定理4，可知

$\forall\ \xi \in S \quad \max R(\xi)=\lambda_k$

所以，由 S 的任意性

$$\min_{\dim S=n-k+1}\{\max_{\xi \in S} R(\xi)\} \leqslant \lambda_k$$

故 $\min\limits_{\dim S=n-k+1}\{\max\limits_{\xi \in S} R(\xi)\}=\lambda_k$

由定理5，很容易得出下述推论：

推论1 F-S 方法求出的 F-S 鉴别向量 ξ_i 具有 $R(\xi_i) \geqslant \lambda_i$，$i \geqslant 2$ 且 $R(\xi_1)=\lambda_1$ 的性质

而定理4表明特征方程 $s_b\xi=\lambda s_w\xi$ 的特征向量 ξ_i^0 具有 $R(\xi_i^0)=\lambda_i$ 的性质，也就是说，就单个鉴别向量而言，ξ_i 比 ξ_i^0 更优。因此，单个 F-S 鉴别向量优于特征方程 $s_b\xi=\lambda s_w\xi$ 相应各特征向量，稍后的实验结果证实了我们的理论分析。

4 基于 F-S 鉴别向量的特征抽取

求得 k 个 F-S 鉴别矢量后，令

$$\phi=\begin{bmatrix}\xi_1^T\\\xi_2^T\\\vdots\\\xi_k^T\end{bmatrix} \tag{14}$$

对训练集中的每一个模式 X 作如下线性变换

$$Y=\begin{bmatrix}y_1\\y_2\\\vdots\\y_k\end{bmatrix}=\phi X \tag{15}$$

的方法称为基于 F-S 鉴别向量的特征抽取。变换空间的维数小于原始模式的维数。模式识别在变换空间中进行。变换空间中如下定理成立。

定理6[9] 由 F-S 鉴别向量集得到的任意两个特征 y_i 与 y_j 具有统计相关性。

证明：$E[(y_i-Ey_i)(y_j-Ey_j)]=\xi_i^T s_t \xi_j$

在模型(4)中约束函数 $\xi^T\xi_i=0$ 成立的条件下可知，一般而言，

$$\xi_i^T s_t \xi_j \neq 0 \tag{16}$$

亦即等式

$$E[(y_i-Ey_i)(y_j-Ey_j)]=0 \tag{17}$$

不能成立，故 y_i 与 y_j 具有统计相关性，也称 F-S 鉴别矢量具有相关性。

5 F-S 鉴别矢量的优选及实验

5.1 鉴别矢量的优选

特征抽取的一般性原则是要求抽取的特征之间统计相关性越小越好。在这一方面，传统的 F-S 鉴别矢量集并不具有最优性质，因为 F-S 鉴别矢量不具有统计无关性，而且，在求解过程中也未对鉴别矢量间的相关性加以任何约束。从这个角度看来，传统 F-S 鉴别矢量集具有改进或重新组合的必要。在这种改进中，应把所抽取的特征是否具有较小的统计相关性作为重要指标加以考虑。为了进一步讨论的方便，定义如下鉴别矢量的相关系数：

定义1 任意 F-S 鉴别矢量 ξ_i, ξ_j 间的相关性由如下相关系数度量

$$\rho(\xi_i, \xi_j)=\frac{\xi_i^T s_t \xi_j}{\sqrt{\xi_i^T s_t \xi_i}\sqrt{\xi_j^T s_t \xi_j}} \tag{18}$$

这种相关性度量本质上反映了变换空间 y_i 与 y_j 的相关性，因为有下式成立

$$\frac{E[(y_i-Ey_i)(y_j-Ey_j)]}{\sqrt{E[(y_i-Ey_i)(y_i-Ey_i)]}\sqrt{E[(y_j-Ey_j)(y_j-Ey_j)]}}=\frac{\xi_i^T s_t \xi_j}{\sqrt{\xi_i^T s_t \xi_i}\sqrt{\xi_j^T s_t \xi_j}} \tag{19}$$

下面考虑对顺序求取的 k 个 F-S 鉴别矢量按一定规则进行优选（重新组合），将优选后的鉴别矢量用于特征抽取。在 F-S 鉴别矢量的优选中，认为满足如下两点的矢量 ξ 是较优的：

i)既具有较大的 $R(\xi)$ 值；

ii)也具有与其它鉴别矢量相关性较小的性质。

为此，提出如下的"评价函数"$f(\xi)$：

$$f(\xi)=R(\xi)-\sum_{j\neq 1}|\rho(\xi,\xi_j)| \tag{20}$$

一般地，"评价函数"$f(\xi)$ 值越大，说明矢量 ξ 的性能越优。如果挑选部分 $f(\xi)$ 值较大的矢量组成新的鉴别矢量集，则这样的矢量集中不仅单个鉴别矢量具有较好的性质，而且整个鉴别矢量集还具有较小统计相关性的优点，这样的优选鉴别矢量集应具有更好的整体性能。设计如下两种鉴别矢量的优选方案：

方案 Ⅰ ξ_1 作为优选矢量集的第一个鉴别矢量，其它鉴别矢量按 $f(\xi)$ 值从大到小依次选取。

方案 Ⅱ 所有鉴别矢量按 $f(\xi)$ 值从大到小依次选取。

上述两种方案均允许选出至多 k 个鉴别矢量组成优选集。

5.2 实验结果及分析

选用 Concordia 大学 CENPARMI 手写体阿拉伯数字库进行实验。该数字库训练集为4000个样本，测试集为2000个样本。研究者常使用该数字库256维 Gabor 变换特征[10]与121维 Legendre 矩特征[11]进行研究。我们使用这两种特征分别进行 F-S 鉴别矢量与优选的 F-S 鉴别矢量的识别实验。对这两个特征，均计算出其前30个 F-S 鉴别矢量，在此基础上进行鉴别矢量的优选。实验中采用最小距离分类器。整个实验在 matlab 上编程实现。

表1 单个共轭特征向量 ξ_i^0 与 F-S 鉴别矢量 ξ_i 性能对比及 $R(\xi_i^0)$ 与 $R(\xi_i)$ 值

	i	1	2	3	4	5	6	7	8	9	10	20	30
X^G	$R(\xi_i^0)$	3.91	2.65	1.87	1.63	1.38	0.96	0.69	0.54	0.50			
	ξ_i^0 错误率	0.67	0.71	0.72	0.74	0.78	0.79	0.78	0.80	0.80			
	$R(\xi_i)$	3.91	3.82	3.70	3.58	3.38	3.19	2.94	2.77	2.61	2.53	1.67	1.29
	ξ_i 错误率	0.67	0.66	0.66	0.66	0.69	0.69	0.71	0.72	0.72	0.69	0.72	0.75
X^L	$R(\xi_i^0)$	4.85	2.61	2.18	1.63	1.02	0.96	0.69	0.43	0.43			
	ξ_i^0 错误率	0.62	0.71	0.72	0.76	0.79	0.78	0.77	0.81	0.80			
	$R(\xi_i)$	4.85	4.68	4.42	4.30	4.06	3.88	3.70	3.47	3.31	3.12	2.09	1.64
	ξ_i 错误率	0.62	0.61	0.62	0.62	0.63	0.63	0.63	0.65	0.64	0.65	0.71	0.74

表1说明：由于使 $R(\xi_i^0)$ 值为0的共轭特征向量不具有对模式的鉴别能力，因此，只进行对应 $R(\xi_i^0)$ 值不为0的9个共轭特征向量的实验。

表2 顺序求取的 F-S 鉴别矢量集的识别错误率

鉴别矢量个数	F-S 鉴别向量		优选 F-S 鉴别向量			
	X^G	X^L	X^G 方案 Ⅰ	X^G 方案 Ⅱ	X^L 方案 Ⅰ	X^L 方案 Ⅱ
1	0.674	0.616	0.674	0.788	0.616	0.691
2	0.673	0.617	0.563	0.630	0.601	0.547
3	0.672	0.616	0.491	0.528	0.585	0.551
4	0.670	0.617	0.426	0.423	0.562	0.483
5	0.669	0.617	0.412	0.397	0.552	0.439
6	0.663	0.617	0.398	0.329	0.530	0.446
7	0.661	0.616	0.361	0.339	0.523	0.392
8	0.659	0.616	0.355	0.316	0.509	0.401
9	0.568	0.615	0.336	0.301	0.499	0.394
10	0.559	0.615	0.327	0.294	0.495	0.373
11	0.537	0.614	0.317	0.284	0.492	0.366
15	0.479	0.590	0.300	0.287	0.478	0.251
20	0.396	0.535	0.283	0.264	0.474	0.241
25	0.318	0.498	0.286	0.254	0.469	0.247
30	0.291	0.463	0.291	0.291	0.463	0.463

实验说明：表中 X^G 代表 Gabor 特征，X^L 代表 Legendre 特征。上述实验结果显示，单个 F-S 鉴别矢量有优于共轭特征向量的性能。但是，F-S 鉴别矢量集作为一个整体来讲，随着鉴别矢量个数的增加，识别错误率的减小较缓慢。而基于方案Ⅰ和方案Ⅱ得到的 F-S 鉴别矢量优选集有如下两个特点：

1)随鉴别矢量个数的增加，错误率很快降低，鉴别矢量数目较少的优选集可得到与整个 F-S 鉴别矢量集相当的识别结果。

2)甚至，方案Ⅱ在 Legendre 特征上的最优识别结果明显优于原 F-S 鉴别矢量集的最优结果！

关于相关系数绝对值的对比分析表明，F-S 鉴别矢量集中很多矢量间的相关性较大，而优选集中处于集合前面位置的矢量具有与其它矢量相关性较小的性质。正是这种性质使得基于优选集中较小个数的鉴别矢量可以得到相对高的识别率。因此，实验明确说明了相关性的大小的确是影响鉴别矢量集整体性能的重要因素。从信息论的角度分析，F-S 鉴别矢量集中存在较大的信息冗余，由于这种信息冗余的存在，使得识别性能并不随鉴别矢量数目的增加而较快提高。实验还启示我们，若不能得到完全不相关的矢量集，可以考虑寻求彼此间相关性较小的矢量集。

表3 X^G 第一、第二 F-S 矢量与其它矢量间相关系数及方案Ⅱ优选所得第一、第二矢量(ξ_{25},ξ_{24})与其它矢量间相关系数绝对值

	i	1	2	3	4	5	6	7	8	9	10	20	30
F-S	ξ_1	1.00	1.00	0.99	0.99	0.98	0.97	0.96	0.95	0.29	0.92	0.80	0.21
	ξ_2	1.00	1.00	1.00	0.99	0.99	0.98	0.96	0.95	0.32	0.91	0.79	0.19
优选	ξ_{25}	0.15	0.15	0.17	0.18	0.20	0.21	0.21	0.18	0.07	0.18	0.11	0.45
	ξ_{24}	0.26	0.24	0.25	0.25	0.23	0.27	0.25	0.19	0.33	0.33	0.41	0.06

结论 本文从理论上分析了在 n 维空间的任意维子空间广义 *Rayleigh* 商 $R(\xi)$ 的极值性质，分析表明，单个 F-S 鉴别矢量有较优的分类性能。但由于 F-S 鉴别矢量间统计相关性

(下转第182页)

```
        Occupancy == vacant | occupied_by(occupant:Dhip)
    value
        fits:Ship × Berth → Bool
end
object T:TYPES
```

下面定义系统的抽象规范，其中可以利用前面已经得到的一致性条件，使用已经定义的类型。具体的方法可以依据RSL中所规定的步骤[2]。

于是可以得到初始规范 A_HARBOUR0（由于篇幅关系，此处略）。

3.4 精化初始化规范并形成精化后的UML模型

对应初始化规范，可以从以下几个方面验证是否符合初始需求的要求：(1) 船只到达能够被正确记录；(2) 存在合适的空闲泊位时船只能够停泊；(3) 停靠的船只可以离开；(4) 船只只能停靠在合适的泊位。

在初始规范 A_HARBOUR0规范中，并没有考虑 pool 和 berth 的具体性质，因此状态改变函数是抽象定义，下面通过定义 pool 和 berth 对象，精化初始规范得到新的形式化初始规范 A_HARBOUR1[1]。

要保证精化的有效性，需要证明 A_HARBOUR1是 A_HARBOUR0的精化，也就是要证明 A_HARBOUR0的所有 axiom 在 A_HARBOUR1的函数中都满足。

为了说明该证明过程，从 A_HARBOUR0中抽取出船只是否能够停靠的形式化初始需求如下：

```
[waiting_docks]
∀ h:Harbour, s1, s2:T.Ship, b:T.Berth•
    waiting(s2, docks(s1,b,h)) ≡ occupancy(b,h)
            pre can_arrive(s,h)                          (1)
```

在 A_HARBOUR1中 waiting 函数被精化为：

```
waiting: T.Ship × Harbour → Bool
waiting(s,(ws, bs)) ≡ P.Isin(s, ws)                      (2)
```

若 A_HARBOUR1是 A_HARBOUR0的精化，则(2)式应使(1)式成立，证明如下：

```
[waiting_docks]
waiting(s2,docks(s1,b,h) ≡ s1 ≠ s2 ∧ waiting(s2,h) =〉
P.Isin(s2,docks(s1,b,h) ≡s1 ≠ s2 ∧ waiting(s2,h) =〉
P.Isin(s2,(P.remove(s1,ws),B.change(T.indx(h),T.occupied_by
(s1),bs) =>
P.Isin(s2,(p.remove(s1,ws),B.change(T.indx(h),T.occupied_by
    (s1),bs))) ≡
    s1 ≠ s2 ∧ P.Isin(s2,h) =>
P.Isin(s2,(ws \ s2, bs † s1) ≡ s1≠s2 ∧ P.Isin(s2,h) =>
若 s1=s2 P.Isin(s2,ws \ s1) ≡ false
        s1≠s2 ∧ P.Isin(s2,h) ≡ false
若 s2≠s2 (ws \ s1, bs † s1) ≡ h
        => P.Isin(s2,h) ≡ s1≠s2 ∧ P.Isin(s2,h)
```

□

其余证明都可以按类似方法完成。

最后，利用 UML 模型和 RSL 模型的对应关系进行模型转换，这里需要注意转换过程中对 scheme、type、object 的处理。然后可以根据形式化规范建立顺序图、协同图，生成代码后，利用前面的规范具体编写完整代码[1]。

参 考 文 献

1 周彦晖.基于面向对象模型的形式化规约和 FDOOM 开发方法：[硕士论文].2002

2 The RAISE Method Group: C. George, A. E. Haxthausen, S. Hughes, R. Milne, S. Prehn, J. S. Pedersen. The RAISE Development Method. TERMA Elecktronik AS, Denmark, 1999

3 Andrews D J, Groote J F, Middelburg C A, et al. Semantics of Specification Languages. Workshops in Computing. Spring-Verlag, 1993

4 Bruun P M,et al. RAISE Tools Refernce Manual. Technical Report LACOS/CRI/DOC/17, CRI: Computer Resources International, 1995

5 Fowler M. UML Distilled (Second Edition). Addison-Wesley, 2000

6 Albir S. UML in a nutshell. O'Reilly,1998

7 Booch G. Rumbaugh, J.Jacobson. The Unified Modeling Language User Guide. Addison-Wesley,1999

8 周彦晖,张为群.软件形式化与可视化软件模型的转换.计算机科学,2003,30(7)

(上接第66页)

的存在，按顺序求取的 F-S 鉴别矢量集整体性能并非"最优"。在提出的综合反映单个鉴别矢量优劣与鉴别矢量间彼此相关性大小的评价函数 $f(\xi)$ 的基础上，设计出两种 F-S 鉴别矢量的优选方案，并在 Concordia 大学 CENPARMI 手写数据库上进行了实验。实验表明，优选出的鉴别矢量集中位置靠前的鉴别矢量由于彼此间相关性较小，基于这些矢量作特征抽取的识别结果优于相同数目的按顺序排列的 F-S 鉴别矢量集的性能。实验验证了鉴别矢量间相关性大小对特征抽取的重要性。

进一步，我们认为对于 F-S 方法求解第 i 个鉴别矢量的一个优化模型应为如下多目标规划问题模型

$$\begin{cases}\max R(\xi) \\ \min g(\xi) \\ \xi^T\xi_j=0\end{cases} \tag{22}$$

其中，

$$g(\xi)=\sum_j|\rho(\xi,\xi_j)| \quad j=1,2,\cdots,i-1 \tag{23}$$

依据最优化理论，模型(4)所得的最优解只是模型(22)的弱有效解。

参 考 文 献

1 Sammon J W. An optimal discriminant plane. IEEE Trans Comput,1970,19(9):826~829

2 Foley D H,Sammon J W. An optimal set of discriminant vectors. IEEE Trans Comput,1975,24(3):281~289

3 Duchene J,Leclercq S. An optimal transformation for discriminant and principal component analysis. IEEE Trans on pattern analysis and machine intelligence,1988,10(6):978~983

4 Hamamoto Y,Matsuura Y,Kanaoka T,et al. A note on the orthonormal discriminant vector method for feature extraction. Pattern Recognition,1991,24(7):681~684

5 Liu K,Cheng Y Q,Yang J Y. A generalized optimal set of discriminant vectors. Pattern Recognition,1921,25(7):731~739

6 Guo Y F,Shu T T,Yang J Y,et al. Feature extraction method based on the generalized Fisher Discriminant criterion and face recognition. Pattern Analysis & Application, 2001, 4(1): 61~66

7 Jin Z,Yang J Y,Hu Z S,et al. Face recognition based on the uncorrelated discriminant transformation. Pattern Recognition, 2001,34:1405~1416

8 丁学仁,蔡高厅.工程中的矩阵理论.天津:天津大学出版社,1985

9 Jin Z,Yang J Y,Tang Z M,et al. A theorem on the uncorrelated optimal discriminant vectors. Pattern Recognition,2001,34:2041~2047

10 Yoshihiko H,et al. Recognition of handwritten numerals using Gabor features. In: Proc. of the Thirteenth ICPR. 250~253

11 Liao S X,Pawlak M. On image analysis by moments. IEEE Trans Pattern Analysis and Machine Intelligence,1996,18(3):254~266