

## 核不相关空间算法及其在人脸识别中的应用

林玉娥 顾国昌 刘海波 沈晶

(哈尔滨工程大学大学计算机科学与技术学院 哈尔滨 150001)

**摘要:** 不相关空间算法是求解不相关鉴别矢量集的快速算法,但是将其应用在人脸识别中将遇到小样本问题,并且算法只是一种线性的特征提取方法。该文提出一种核不相关空间算法,该方法的关键是高维特征空间中不相关空间的计算,对此提出一种简单的计算方法,即根据 eigenface 中将高阶矩阵计算转化成低阶矩阵计算的思想,将高维特征空间中不相关空间的计算仍归结为标准的特征值分解问题。所提出的算法能够有效地解决小样本问题。在 ORL 人脸库上的实验结果验证了所提出的算法的可行性和有效性。

**关键词:** 人脸识别; 不相关空间算法; 小样本问题; 核不相关空间算法

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2009)08-1812-04

## Kernel Uncorrelated Space Algorithm and Its Application to Face Recognition

Lin Yu-e Gu Guo-chang Liu Hai-bo Shen Jing

(College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China)

**Abstract:** Uncorrelated space algorithm is a fast method for the uncorrelated discriminant vectors extraction, but it may encounter the small size samples problem when it is applied to face recognition task. In addition, it is only a linear feature extraction technique. In this paper, kernel uncorrelated space algorithm is proposed. The key of the proposed algorithm is to how to compute the uncorrelated space in the higher dimensional feature space. As to this problem, a very simple and easy method is proposed, which originates from the eigenface that transforms the computation of the high order matrix into the computation of the low order matrix, and then the actual computation of the uncorrelated space in the higher dimensional feature space is reduced to a standard eigenvalue problem. In addition, the proposed algorithm can effectively overcome small size samples problem. The numerical experiments on facial databases of ORL show that the proposed method is effective and feasible.

**Key words:** Face recognition; Uncorrelated space algorithm; The small size samples problem; Kernel uncorrelated space algorithm

### 1 引言

不相关线性鉴别分析算法<sup>[1,2]</sup>是对经典 LDA 的发展,并被认为是线性特征提取的最有效方法之一。而文献[2]的方法不仅理论简单,且具有较快的求解速度,本文为了将其与文献[1]算法区别,称之为不相关空间算法(uncorrelated space algorithm)。然而不相关空间算法,是假设总体散布矩阵非奇异时给出的一种快速算法,而在实际应用中,特别是在人脸识别中,经常会遇到高维小样本问题,使得总体散布矩阵是奇异的,对此文献[2]并没有进行理论上的研究与讨论。另一方面,不相关空间算法属于线性特征提取方法,因而最终获得的只是线性特征。但是现实中许多问题都是线性不可分的,而基于核

函数的学习方法<sup>[3-6]</sup>,由于可以使得线性不可分的问题变得线性可分,因此已成为模式识别领域一个新的发展方向。

据此,本文提出了一种核不相关空间算法(kernel uncorrelated space algorithm)。与文献[7]提出的核不相关鉴别矢量集算法不同,本文的核不相关空间算法不需假设数据是中心化处理的,也不需对高维解进行事先假设,因而算法更具有一般性,且理论更加简单。新算法的关键一步是如何求出高维特征空间中的不相关空间,对此提出了一种新的求解方法,该方法来源于 eigenface<sup>[8]</sup>求解高阶协方差矩阵所采用的思想,也就是对于小样本问题,由于协方差矩阵是对称矩阵,所以可以将高阶协方差矩阵的求解转化成低阶协方差矩阵的求解,根据这一思想,将高维特征空间中的不相关空间计算仍归结为标准的特征值分解问题,同时所提出的算法能

2008-06-02 收到, 2009-03-30 改回

中国博士后基金(20060400809)和黑龙江省青年科技基金(QC06C022)资助课题

够有效地解决小样本问题。

## 2 Fisher 准则与不相关空间算法

首先将类内散布矩阵,类间散布矩阵和总体散布矩阵分别记为:  $S_w$ ,  $S_b$  和  $S_t$ , 则 Fisher 鉴别准则函数定义为

$$J(\phi) = \frac{\phi^T S_b \phi}{\phi^T S_t \phi} \quad (1)$$

不相关空间算法是一种求解不相关鉴别矢量集的快速算法,它是以广义的 Fisher 鉴别准则函数和一组共轭正交条件作为求解模型,其基本思想是,假设  $S_t$  可逆,已求出  $r-1$  个鉴别矢量,则对于鉴别矢量  $\phi_r$ ,可按下面的准则函数求出:

$$\begin{aligned} \max J(\phi_r) &= \frac{\phi_r^T S_b \phi_r}{\phi_r^T S_t \phi_r}, \\ \text{s.t.} \quad &\begin{cases} \phi_r^T S_t \phi_r = 1 \\ \phi_r^T S_t \phi_t = 0, t=1, \dots, r-1 \\ \phi_r \in \mathbf{R}^n \end{cases} \end{aligned} \quad (2)$$

对于这组最佳鉴别矢量  $\phi_1, \phi_2, \dots, \phi_r$  的求解,为了避免计算逆阵的迭代求解方法,不相关空间算法提出了首先构造不相关空间  $\mathbf{V}$ ,即求出  $\mathbf{V}$  使得  $\mathbf{V}^T S_t \mathbf{V} = \mathbf{I}$ ,因此最终将求解不相关鉴别矢量集转化成为在不相关空间里,求解使类间散布矩阵到极大值的一组正交鉴别矢量集。具体步骤请参见文献[2]。

## 3 一种新的核不相关空间算法

### 3.1 关于总体散布矩阵的奇异问题讨论

在应用不相关空间算法时,实际上只有两种情况,一种是当矩阵  $S_t$  非奇异时,这时直接应用不相关空算法就可求出不相关鉴别矢量集;另一种是当矩阵  $S_t$  奇异时,此时不能直接应用不相关空算法,而基于核的非线性方法,首先要将原始训练样本通过一个非线性映射  $\phi$  变换到某一高维(可能是无限维)特征空间  $\mathbf{H}$  中,所以采用非线性变换后也存在这一问题,因此在将不相关空间算法推广成为非线性方法之前,首先对  $S_t$  奇异情况进行理论分析与讨论。

**定义1** 设  $\alpha_1, \alpha_2, \dots, \alpha_n$  表示  $S_t$  的标准正交的特征矢量集,令  $\mathbf{Z} = \{\mathbf{x} | S_t \mathbf{x} = 0, \mathbf{x} \in \mathbf{R}^n\}$ ,  $\mathbf{Z}_\perp$  代表  $\mathbf{Z}$  的正交补空间,则有  $\mathbf{Z}_\perp = \text{span}\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ ,  $\mathbf{Z} = \text{span}\{\alpha_{m+1}, \alpha_{m+2}, \dots, \alpha_n\}$ , 其中  $m = \text{rank}(S_t)$ 。

**引理1**<sup>[9]</sup> 设  $\mathbf{A}$  为非负定阵,  $\mathbf{x}$  为  $n$  维向量,则  $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$ , 当且仅当  $\mathbf{A} \mathbf{x} = 0$ 。

**定理1** 假设  $S_t$  是奇异的,  $\phi_1$  为满足基于 Fisher 准则函数  $J(\phi)$  的第1个统计不相关鉴别矢量,则有

$\phi_1 \in \mathbf{Z}_\perp$ 。

**证明** 用反证法来证明。由于  $\phi_1$  为满足  $J(\phi)$  的第1个统计不相关鉴别矢量,即  $\phi_1$  满足下式:

$$\max J(\phi) = \frac{\phi_1^T S_b \phi_1}{\phi_1^T S_t \phi_1}, \text{ s.t. } \phi_1^T S_t \phi_1 = 1 \quad (3)$$

首先设  $\phi_1 \notin \mathbf{Z}_\perp$ , 且令  $\phi_1 = \mathbf{w}_1 + \eta_1$ , 其中  $\mathbf{w}_1 \in \mathbf{Z}_\perp$ ,  $\eta_1 \in \mathbf{Z}$ , 则有

$$\begin{aligned} \phi_1^T S_t \phi_1 &= (\mathbf{w}_1 + \eta_1)^T S_t (\mathbf{w}_1 + \eta_1) = \mathbf{w}_1^T S_t \mathbf{w}_1 + \mathbf{w}_1^T S_t \eta_1 \\ &\quad + \eta_1^T S_t \mathbf{w}_1 + \eta_1^T S_t \eta_1 \end{aligned}$$

由引理1和  $\eta_1 \in \mathbf{Z}$  可得  $\eta_1^T S_t \eta_1 = 0$ , 和  $\mathbf{w}_1^T S_t \eta_1 = 0$ , 而  $\eta_1^T S_t \mathbf{w}_1 = (\mathbf{w}_1^T S_t \eta_1)^T = \eta_1^T S_t \mathbf{w}_1 = 0$  因此可得  $\phi_1^T S_t \phi_1 = \mathbf{w}_1^T S_t \mathbf{w}_1$ ; 同理可得:  $\phi_1^T S_b \phi_1 = \mathbf{w}_1^T S_b \mathbf{w}_1$ ; 因而可得

$$J(\phi_1) = \frac{\phi_1^T S_b \phi_1}{\phi_1^T S_t \phi_1} = \frac{\mathbf{w}_1^T S_b \mathbf{w}_1}{\mathbf{w}_1^T S_t \mathbf{w}_1} = J(\mathbf{w}_1)$$

而对于限制条件可得  $\phi_1^T S_t \phi_1 = \mathbf{w}_1^T S_t \mathbf{w}_1 = 1$ 。

综上所述,可知  $\mathbf{w}_1$  就是满足式(3)的特征向量,因此可得:  $\phi_1 = \mathbf{w}_1 \in \mathbf{Z}_\perp$ 。证毕

**定理2** 假设  $S_t$  是奇异的,则对于满足式(2)的任意一个不相关鉴别矢量  $\phi_t$ , 都有  $\phi_t \in \mathbf{Z}_\perp$ , 其中  $t=1, \dots, C-1$ 。

**证明** 用归纳法来证明。由定理1知  $\phi_1 \in \mathbf{Z}_\perp$ , 则假设已求出  $\phi_t = \mathbf{w}_t \in \mathbf{Z}_\perp$  ( $t=1, \dots, r-1$ ), 则  $\phi_r$  即是求使式(2)达到第  $r$  极大值的特征向量。设  $\phi_r = \mathbf{w}_r + \eta_r$ , 其中  $\mathbf{w}_r \in \mathbf{Z}_\perp$ ,  $\eta_r \in \mathbf{Z}$ , 则按照定理1的证明过程,不难得到:  $\phi_r^T S_t \phi_r = (\mathbf{w}_r + \eta_r)^T S_t (\mathbf{w}_r + \eta_r) = \mathbf{w}_r^T S_t \mathbf{w}_r$ , 及  $\phi_r^T S_b \phi_r = (\mathbf{w}_r + \eta_r)^T S_b (\mathbf{w}_r + \eta_r) = \mathbf{w}_r^T S_b \mathbf{w}_r$ , 因而可得

$$J(\phi_r) = \frac{\phi_r^T S_b \phi_r}{\phi_r^T S_t \phi_r} = \frac{\mathbf{w}_r^T S_b \mathbf{w}_r}{\mathbf{w}_r^T S_t \mathbf{w}_r} = J(\mathbf{w}_r)$$

而对于限制条件可得  $\phi_r^T S_t \phi_r = \mathbf{w}_r^T S_t \mathbf{w}_r = 1$ , 由于  $\phi_t = \mathbf{w}_t \in \mathbf{Z}_\perp$  ( $t=1, \dots, r-1$ ), 则  $\phi_r^T S_t \phi_t = (\mathbf{w}_r + \eta_r)^T S_t \phi_t = \mathbf{w}_r^T S_t \phi_t + \eta_r^T S_t \phi_t$ , 而由  $\eta_r \in \mathbf{Z}$ , 则可得  $\phi_r^T S_t \phi_t = \mathbf{w}_r^T S_t \phi_t = \mathbf{w}_r^T S_t \mathbf{w}_t = 0$ 。

综上可知  $\mathbf{w}_r$  就是满足式(2)的特征向量,因而可得:  $\phi_r = \mathbf{w}_r \in \mathbf{Z}_\perp$ 。因此,可知对于任意不相关鉴别矢量  $\phi_t$  都有  $\phi_t \in \mathbf{Z}_\perp$ 。证毕

由定理1和定理2知,就 Fisher 准则而言,子空间  $\mathbf{Z}$  中不存在任何有效的鉴别信息,不相关鉴别矢量只能在  $\mathbf{Z}_\perp$  中取。

### 3.2 核不相关空间算法

核不相关空间算法的基本思想:首先将原始训练样本通过一个非线性映射  $\phi$  变换到某一高维(可能是无限维)特征空间  $\mathbf{H}$  中,然后在高维特征空间  $\mathbf{H}$  中完成鉴别矢量集的求解。假设有  $C$  个样本类,

设第  $i$  类有  $N_i$  个样本, 共有  $\sum_{i=1}^C N_i = N$  个样本,

$\mathbf{X} = [\mathbf{x}_1^1 \cdots \mathbf{x}_1^{N_1} \cdots \mathbf{x}_C^{N_1} \cdots \mathbf{x}_C^{N_c}]$  表示全体样本集, 其中  $\mathbf{x}_i^j \in \mathbf{R}^n$  代表第  $i$  类的第  $j$  个样本, 经过非线性映射  $\Phi$  后, 则  $\mathbf{X}$  在高维特征空间  $\mathbf{H}$  中的表示为下式:

$$\Phi(\mathbf{X}) = [\Phi(\mathbf{x}_1^1) \cdots \Phi(\mathbf{x}_1^{N_1}) \cdots \Phi(\mathbf{x}_C^{N_1}) \cdots \Phi(\mathbf{x}_C^{N_c})] \quad (4)$$

上式中  $\Phi(\mathbf{x}_i^j) \in \mathbf{H}$ , 是  $\mathbf{x}_i^j$  所对应的高维空间  $\mathbf{H}$  中的样本矢量, 在高维空间  $\mathbf{H}$  中相应的类内散布矩阵  $\mathbf{S}_w^\Phi$ , 类间散布矩阵  $\mathbf{S}_b^\Phi$  和总体散布矩阵  $\mathbf{S}_t^\Phi$  分别定义如下<sup>[10]</sup>:

$$\mathbf{S}_w^\Phi = \Phi(\mathbf{X})(\mathbf{I} - \mathbf{L})(\mathbf{I} - \mathbf{L})^\top (\Phi(\mathbf{X}))^\top = \mathbf{A}\mathbf{A}^\top \quad (5)$$

$$\mathbf{S}_b^\Phi = \Phi(\mathbf{X})(\mathbf{L} - \mathbf{M})(\mathbf{L} - \mathbf{M})^\top (\Phi(\mathbf{X}))^\top = \mathbf{B}\mathbf{B}^\top \quad (6)$$

$$\mathbf{S}_t^\Phi = \Phi(\mathbf{X})(\mathbf{I} - \mathbf{M})(\mathbf{I} - \mathbf{M})^\top (\Phi(\mathbf{X}))^\top = \mathbf{T}\mathbf{T}^\top \quad (7)$$

式中  $\mathbf{I}$  是  $N \times N$  的单位阵,  $\mathbf{L} = (\mathbf{L}_i)_{i=1, \dots, c}$  是  $N \times N$  的块对角阵, 其中每个对角块  $\mathbf{L}_i$  所有元素均为  $1/N_i$ , 大小为  $N_i \times N_i$  的矩阵, 式中  $\mathbf{M}$  为大小为  $N \times N$  的矩阵, 该矩阵的所有元素均为  $1/N$ 。而此时所求的鉴别矢量  $\phi_r$  即应满足:

$$\left. \begin{aligned} \max J^\Phi(\phi_r) &= \frac{\phi_r^\top \mathbf{S}_b^\Phi \phi_r}{\phi_r^\top \mathbf{S}_t^\Phi \phi_r} \\ \text{s.t.} \quad &\left\{ \begin{aligned} \phi_r^\top \mathbf{S}_t^\Phi \phi_r &= 1 \\ \phi_r^\top \mathbf{S}_i^\Phi \phi_i &= 0, \quad i = 1, \dots, r-1 \\ \phi_i &\in \mathbf{H} \end{aligned} \right. \end{aligned} \right\} \quad (8)$$

核不相关空间算法的关键一步是如何求出高维特征空间中的不相关空间, 而由上一节的定理 1 和定理 2 可知, 有效的不相关鉴别矢量只取自  $\mathbf{S}_t^\Phi$  的非零空间, 因此求解高维特征空间的不相关空间也就是求出  $\mathbf{S}_t^\Phi$  所有非零特征值与所对应的特征向量, 然而由于  $\Phi$  未知, 因而不能直接计算出不相关空间, 对此给出一种新的求解方法, 由文献[8]所提出 eigenface 方法可知, 该方法在计算高维数据的总体散布矩阵时, 使用了对称矩阵的一个性质, 即  $\mathbf{T}\mathbf{T}^\top$  与  $\mathbf{T}^\top\mathbf{T}$  具有相同的特征值, 从而将高阶矩阵的求解转化成为低阶矩阵的求解, 因此将 eigenface 的求解思想用到本文的算法中, 下面将对这一求解方法进行推导。

根据  $\mathbf{T}\mathbf{T}^\top$  所对应的特征值与  $\mathbf{T}^\top\mathbf{T}$  的特征值相同, 所以首先来求  $\mathbf{T}^\top\mathbf{T}$  的特征值与特征向量, 即

$$\begin{aligned} \mathbf{T}^\top\mathbf{T} &= (\mathbf{I} - \mathbf{M})^\top \Phi(\mathbf{X})^\top \Phi(\mathbf{X})(\mathbf{I} - \mathbf{M}) \\ &= (\mathbf{I} - \mathbf{M})^\top \mathbf{K}(\mathbf{I} - \mathbf{M}) \end{aligned} \quad (9)$$

式中的  $\mathbf{K} = \Phi(\mathbf{X})^\top \Phi(\mathbf{X})$  即为核矩阵, 是由训练样本的内积生成。此时的  $\mathbf{K}$  为  $N \times N$  的对称矩阵 ( $N$  为训练样本总数), 因此可得

$$\mathbf{T}^\top\mathbf{T}\mathbf{U} = \Sigma\mathbf{U}$$

$$\mathbf{T}\mathbf{T}^\top\mathbf{T}\mathbf{U} = \mathbf{T}\Sigma\mathbf{U}$$

$$\mathbf{S}_t^\Phi(\mathbf{T}\mathbf{U}) = \Sigma(\mathbf{T}\mathbf{U})$$

其中  $\Sigma = \begin{pmatrix} \Sigma_m & 0 \\ 0 & 0 \end{pmatrix}$ ,  $\Sigma_m$  是  $\mathbf{T}^\top\mathbf{T}$  的非零特征值按降

序排列的对角阵,  $m = \text{rank}(\mathbf{T}^\top\mathbf{T})$ ,  $\mathbf{U}$  是  $\mathbf{T}^\top\mathbf{T}$  的特征向量组成的矩阵, 令  $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2)$ ,  $\mathbf{U}_1 \in \mathbf{R}^{N \times m}$ ,  $\mathbf{U}_2 \in \mathbf{R}^{N \times (N-m)}$ ,  $\mathbf{U}_1$  为  $\mathbf{T}^\top\mathbf{T}$  的所有非零特征值所对应的特征向量, 则可知  $\mathbf{P} = \mathbf{T}\mathbf{U}_1 = \Phi(\mathbf{X})(\mathbf{I} - \mathbf{M})\mathbf{U}_1$  即是  $\mathbf{S}_t^\Phi$  所有非零特征值所对应特征向量, 此时则有

$$\begin{aligned} \mathbf{P}^\top \mathbf{S}_t^\Phi \mathbf{P} &= (\mathbf{T}\mathbf{U}_1)^\top \mathbf{S}_t^\Phi (\mathbf{T}\mathbf{U}_1) = \mathbf{U}_1 \mathbf{T}^\top (\mathbf{T}\mathbf{T}^\top) \mathbf{T}\mathbf{U}_1 \\ &= \mathbf{U}_1 \mathbf{T}^\top (\mathbf{T}\mathbf{U}_1^\top \mathbf{U}_1 \mathbf{T}^\top) \mathbf{T}\mathbf{U}_1 \\ &= (\mathbf{U}_1 \mathbf{T}^\top \mathbf{T} \mathbf{U}_1^\top) (\mathbf{U}_1 \mathbf{T}^\top \mathbf{T} \mathbf{U}_1) = \Sigma_m \Sigma_m \end{aligned}$$

因此可令

$$\mathbf{V} = \mathbf{P}\Sigma_m^{-1} = \mathbf{T}\mathbf{U}_1\Sigma_m^{-1} \quad (10)$$

则有  $\mathbf{V}^\top \mathbf{S}_t^\Phi \mathbf{V} = \mathbf{I}$ ,  $\mathbf{V}$  即是所求的不相关空间, 然后将  $\mathbf{S}_b^\Phi$  投影至  $\mathbf{V}$  上, 得

$$\begin{aligned} \mathbf{V}^\top \mathbf{S}_b^\Phi \mathbf{V} &= (\mathbf{T}\mathbf{U}_1\Sigma_m^{-1})^\top \mathbf{S}_b^\Phi (\mathbf{T}\mathbf{U}_1\Sigma_m^{-1}) \\ &= \Sigma_m^{-1} \mathbf{U}_1^\top \mathbf{K}(\mathbf{L} - \mathbf{M})(\mathbf{L} - \mathbf{M})^\top \mathbf{K}\mathbf{U}_1 \Sigma_m^{-1} \end{aligned} \quad (11)$$

再对式(11)进行特征分解, 求得前  $C-1$  最大特征值所对应的特征向量, 即完成了核不相关鉴别矢量的求解过程, 其具体实现步骤如下:

(1) 首先计算核矩阵  $\mathbf{K}$ ;

(2) 求出式(9)有非零特征值及所对应的特征向量, 即  $\Sigma_m$  和  $\mathbf{U}_1$ , 此时可按式(10)设不相关空间为  $\mathbf{V} = \mathbf{P}\Sigma_m^{-1} = \mathbf{T}\mathbf{U}_1\Sigma_m^{-1}$ ;

(3) 按式(11)将  $\mathbf{S}_b^\Phi$  投影至  $\mathbf{V}$  上, 然后对式(11)进行特征分解, 求得前  $C-1$  最大特征值所对应的特征向量  $\mathbf{\Gamma}$ , 则最终所得的核不相关鉴别矢量集为  $\mathbf{\Psi} = \mathbf{V}\mathbf{\Gamma} = \mathbf{T}\mathbf{U}_1\Sigma_m^{-1}\mathbf{\Gamma}$ 。

## 4 实验

为了验证提出算法的性能, 在ORL人脸库上进行了实验。在ORL人脸库中, 有40个人的400幅图像, 每人10幅, 其中有些图像是拍摄于不同时期, 人脸姿态有相当程度的变化, 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, 每幅图像的分辨率为112×92。实验中采用最近邻方法进行分类, 算法所使用的核函数为高斯核函数,  $k(x, y) = \exp(-\|x - y\|^2 / \sigma)$ , 其中  $\sigma = 5.5e + 7$ 。对核不相关空间算法与不相关空间算法的识别性能进行了比较。

首先在实验中选择每人的前5幅图像作为训练样本, 剩余图像作为测试样本, 对两种算法提取19-39个鉴别矢量所得识别率进行了比较, 结果如

图 1 所示。可以看出核不相关空间算法在 19 个鉴别矢量时, 就取得较好的识别率, 明显优于线性算法, 当提取 39 个鉴别矢量时, 核不相关空间算法达到了

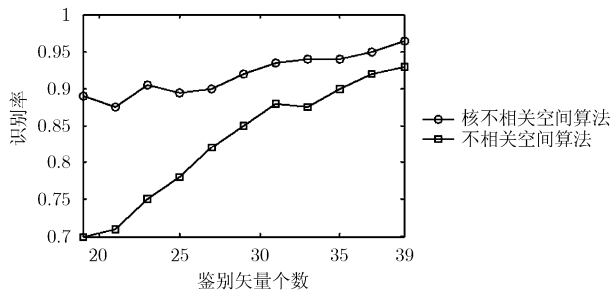


图 1 鉴别矢量个数变化时两种方法的性能比较

最高识别率 0.96, 而线性的算法只有 0.93。

为了进一步验证提出算法的性能, 将本文算法与不相关空间算法, KPCA, Kfisher 和文献[7]的方法进行了比较, 本次实验中, 每人分别随机选取 3-6 幅图像作为训练样本, 剩余的图像作为测试样本, 每组实验重复进行 10 次, 取平均值作为识别结果, 实验结果如表 1 所示。从表 1 的识别结果可以看出, 核不相关空间算法明显优于线性不相关空间算法, 这也说明经过非线性变换后, 确实可以提高算法的识别性能, 而且也优于 Kfisher 和 KPCA, 这是因为本文算法所求得鉴别矢量集, 彼此之间具有不相关的特性, 因而其识别结果优于这两种核方法, 同时也不低于文献[7]的识别性能, 这说明本文的求解方法较文献[7]的求解方法更加有效。

表 1 5 种方法实验结果的比较

算法名称	样本数目及正确识别率(%)			
	3	4	5	6
不相关空间算法	88.35	91.50	92.25	93.63
KPCA	88.86	92.45	94.20	95.12
KFisher	89.21	93.83	95.30	96.25
文献[7]算法	92.86	95.58	97.40	97.63
本文算法	93.07	96.33	97.40	97.75

## 5 结论

本文进一步完善与发展了原有的不相关空间算法, 提出一种核不相关空间算法, 新算法不需事先对数据进行中心化处理, 也不需对高维解进行假设, 因而更简单, 也更具有通用性。核不相关空间算法的关键一步是如何求出高维特征空间中的不相关空间, 对此根据 eigenface 求解高阶协方差矩阵所采用

的思想, 将高维特征空间中的不相关空间计算仍归结为标准的特征值分解问题, 同时所提出的算法能够有效解决小样本问题, 从而为基于核映射的不相关矢量集算法提供了一种新的求解思路。

## 参考文献

- [1] Jin Z, Yang J, and Hu Z, *et al.* Face recognition based on the uncorrelated discrimination transformation [J]. *Pattern Recognition*, 2001, 34(7): 1405-1416.
- [2] 陈绵书, 陈贺新, 刘伟. 一种新的求解无相关鉴别矢量集方法[J]. *计算机学报*, 2004, 27(7): 913-917.  
Chen Mian-shu, Chen He-xin, and Liu Wei. A new method for resolving the uncorrelated set of discriminant vectors [J]. *Chinese Journal of Computers*, 2004, 27(7): 913-917.
- [3] Schölkopf B, Smola A, and Müller K R. Nonlinear component analysis as a kernel eigenvalue problem [R]. Max-Planck-Institut für biologische Kybernetik, Technical Report 44, 1996.
- [4] Bach F R and Jordan M I. Kernel independent component analysis[R]. Department of Computer Science, University of California, Technical Report UCB/CSD-01-1166, 2001.
- [5] Mika S, Ratsch G, and Weston J, *et al.* Fisher discriminant analysis with kernels [C]. Proceedings of IEEE International Workshop on Neural Networks for Signal Processing. Madison, Wisconsin, August 1999: 41-48.
- [6] Yang J, Frangi A, and Yang J Y, *et al.* KPCA plus LDA: A complete kernel Fisher discriminant framework for feature extraction and recognition [J]. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 2005, 27(2): 230-244.
- [7] 梁志贞, 施鹏飞. 核不相关鉴别分析以及它在字符识别中的应用[J]. *计算机研究与发展*, 2006, 43(1): 132-137.  
Liang Zhi-zhen and Shi Peng-fei. Kernel uncorrelated discriminant analysis and its application to handwritten character recognition [J]. *Journal of Computer Research and Development*, 2006, 43(1): 132-137.
- [8] Turk M and Pentland A. Eigenfaces for recognition[J]. *Journal of Cognitive Neuroscience*, 1991, 3(1): 71-86.
- [9] Liu K and Yang J Y. An efficient algorithm for Foley-Sammon optimal set of discriminant vectors by algebraic method[J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 1992, 6(5): 817-829.
- [10] Zheng W M and Zhao L. Foley-Sammon optimal discriminant vectors using kernel approach [J]. *IEEE Transactions on Neural Networks*, 2005, 16(1): 1-9.

林玉娥: 女, 1979 年生, 博士生, 研究方向为模式识别及人工智能。

顾国昌: 男, 1946 年生, 教授, 博士生导师, 主要研究方向为智能机器人软硬件体系结构、任务规划、模式识别和人工智能等。

刘海波: 男, 1976 年生, 副教授, 研究方向为模式识别及智能机器人。

沈晶: 女, 1969 年生, 副教授, 研究方向为模式识别及智能机器人。