

# 基于改进的多流形算法的人脸图像集识别<sup>①</sup>

李文媛, 于威威, 张 燕

(上海海事大学 信息工程学院, 上海 201306)

**摘 要:** 一个图像集由大量变化不一的图像组成, 而且这些图像都表示同一个人. 现实中的图像集数据是非线性的, 造成这些现象的因素有人脸的角度不同、光线的明暗等, 因此图像集中的每幅图像都是变化的, 如果近似的将一个图像集建模为线性子空间, 而忽略了集合中数据结构的变化, 很显然是不合理的, 这也必然会影响到最后的识别率. 受流形理论知识的启发, 可以将图像集建模为一个流形, 这与传统的将图像集建模为子空间的方法有着本质区别. 本文在基于流形的人脸图像集识别方法的基础上进行改进, 提出新的计算子空间距离方法, 最后采用所有最短子空间距离的平均值作为流形之间的距离, 称为改进的多流形方法(Improved multi-manifold method, IMM). IMM 方法在 CMU PIE 数据库上进行实验, 结果表明该方法相比其他方法具有更高识别率.

**关键词:** 图像集; 流形; 线性子空间; 主角

## Improved Multi-Manifold-Based Method for Face Image Set Recognition

LI Wen-Yuan, YU Wei-Wei, ZHANG Yan

(College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China)

**Abstract:** An image set consists of a large number of different images, and these images represent the same person. In the real life, the dates of image sets are nonlinear due to the variation from viewpoint, emotion and illumination. Inspired by the manifold theory knowledge, we assume that modeling each image set as a manifold will be more efficient compared with the traditional method of modeling an image set as a subspace intrinsically. Because the images in an image set are different from each other, it is unreasonable to model an image set as a linear subspace without ignoring the data structure in the set, which may decrease the recognition rate ultimately. In the paper, we introduce a method of improved multi-manifold based face recognition for image sets, and propose a new method for computing the distance between two subspaces. In addition, in order to get the minimal manifold distance, we choose the mean value of closest pair of subspaces as the manifold distance. We call this new method, an improved multi-manifold (IMM). Experimental results on the public available face databases, CMU PIE, demonstrate that the new method outperforms the competing methods

**Key words:** image set; manifold; linear subspace; principal angles

## 1 引言

传统人脸识别过程是对标准的单幅图像进行识别, 但是在现实生活中, 一幅标准的图像很难获取. 随着科技的发展, 存储大量的图像成为事实. 与单幅图像相比, 每个集合里有许多图像, 这些图像很自然地覆盖了人脸的外貌变化, 如: 姿势变化, 表情不一, 灯光明暗等. 所以, 通过充分利用集合的信息, 寻求一种具有更高鲁棒性的人脸识别方法.

在过去的十几年中, 基于图像集识别的研究已经取得了一些成果. 图像集识别方法一般分为两种: 基于参数方法和基于非参数方法. 前者是概率建模方法, 代表性的方法有流密度散度, 该方法是用参数分布函数来表示一个图像集, 并用 K-L 散度来衡量两个集合分布的相似性. O. Arandjelović 等人<sup>[1]</sup>提出用单一的高斯模型和混合高斯模型(GMM)<sup>[2]</sup>, Wen Wang 等人<sup>[3]</sup>在此基础上提出了基于高斯分布的黎曼流形判别分析方法,

① 收稿时间:2016-04-09;收到修改稿时间:2016-05-26 [doi:10.15888/j.cnki.csa.005520]

同样也是将图像集建模为高斯混合模型。

基于非参数方法是假设数据分布情况,代表性的算法是将图像集建模为一个线性子空间. Hamm 等人<sup>[4]</sup>提出格拉斯曼流形判别学习方法(GDA),该方法是在欧式空间将图像集建模为一个线性子空间,在格拉斯曼流形中代表一个点,通过非线性映射将数据点映射到新的格拉斯曼流形上,最后在新的格拉斯曼流形深进行分类.但是该方法没有充分考虑数据分布的集合结构. Yamaguchi 等人<sup>[5]</sup>提出互子空间方法(MSM), Fukuid 等人<sup>[6]</sup>在此基础上提出了约束的互子空间方法(CMSM).该方法认为前几个特征向量更具有判别性,用该特征向量张成子空间,然后进行标准化和正交化.随着流形的研究<sup>[7]</sup>,流形的建模方法也被用到图像集识别当中.例如: Jinwen Lu 等人<sup>[8]</sup>提出的多流形深度学习分类方法. Lei Zhang 等人<sup>[9]</sup>提出了多流形判别分析方法,该方法是在 fisher 判别分析框架下引入图嵌入学习,这和 GDA 方法类似.

Li 等人<sup>[10]</sup>认为在光照和姿态变化的情况下,人脸图像在高维空间是非线性的,所以不同的人会形成不同的子流.鉴于这个理论,我们认为一个图像集中所有的数据都分布在一个非线性流上,在流形上构建局部线性子空间,图像集合之间的相似性用流形与流形之间的距离测度来衡量<sup>[20]</sup>.由于流形是非线性的,目前没有给出统一地定义流形与流形之间的距离公式.本文采用所有最短子空间距离的平均值作为流形与流形之间的距离.

## 2 多流形的分类算法的分析

流形是非线性的,这和图像集中数据的非线性具有一致性<sup>[21]</sup>.基于非线性流形模型是由许多局部线性模型构成的理论知识<sup>[7,22]</sup>,一个流形上的图像数据可以再分成局部线性子空间,多流形分类的方法就转化为计算子空间之间距离,并联合这些子空间距离作为最终的流形之间的距离.基于多流形的图像集分类方法如图1所示.

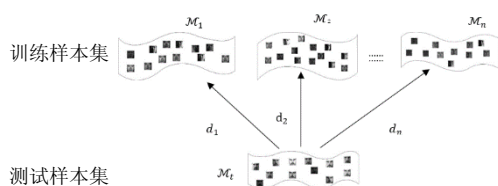


图1 图像集分类模型

### 2.1 构建子空间

目前从非线性流中提取局部线性模型的方法很多,典型的有聚类方法 k-means<sup>[11,12]</sup>, 分级聚类(HAC)<sup>[13]</sup>等,这些都是以一个简单的聚类方法构建局部线性模型,因此不能保证局部子空间的线性性质.所以 Wang 等人<sup>[14]</sup>利用欧式距离和测地线距离的比值,式(1)来规约最大线性块,即局部线性模型.该方法解决了基于迭代的聚类方法中存在的问题,使得构建的子空间具有线性性质.本文也采用该方法在非线性流上构建局部线性子空间.

$$R(x, y) = \frac{D_C}{D_E} \quad (1)$$

定义. 一个图像集  $X = \{x_1, x_2, \dots, x_N\}$ ,  $x_i \in R^D$ ,  $R^D$  是  $D$  维向量空间,  $N$  是该集中的样本数量. 该图像集可以改写为:

$$X = \bigcup_{i=1}^m S_i \quad (2)$$

$$S_i \cap S_j = \emptyset, \text{ 仅当 } i \neq j, \text{ 且 } i, j = 1, 2, \dots, m.$$

$$S_i = \{x_1^{(i)}, x_2^{(i)}, \dots, x_{N_i}^{(i)}\}$$

其中,  $m$  是局部线性模型数量,  $N_i$  表示每个模型中的样本数量.  $m$  的值由一次性算法<sup>[14]</sup>获得. 从式(2)可以看出,非线性流可以表述为许多局部线性空间的集合.如图2所示,一个流形上代表一个人,但是受到光照和拍摄角度影响,这些图像之间的变化还是很大的,所以将一个图像集合张成一个子空间显然是不合理的,但是一个人在姿势和光线变化不大的情况下,我们近似的认为该组图像是具有线性的,因此可以张成一个线性子空间.

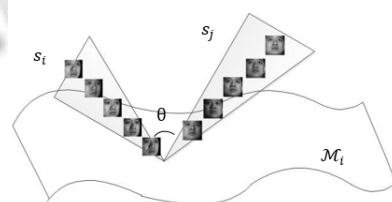


图2 一个流形上张成两个子空间

式(1)中的矩阵  $R(x_1, x_2)$  反映了一个局部线性模型的非线性程度,例如,在任一个已构建成的局部线性模型  $C_i$ , 它的非线性分数可以用式(3)表述

$$\delta_i = \frac{1}{N_i^2} \sum_{m=1}^{N_i} \sum_{n=1}^{N_i} R(x_m^i, x_n^i) \quad (3)$$

### 2.2 局部模型距离度量

#### 2.2.1 主角

主角是衡量两个局部模型相似度的关键. 有两个

图像集合,  $X=\{x_1, x_2, \dots, x_N\}$ ,  $Y=\{y_1, y_2, \dots, y_M\}$ ,  $x_i, y_j \in R^D$ .  $X_1, Y_2$  分别是  $X, Y$  一个线性子集, 其张成的子空间是  $S_1 = span\{X_1\}$ ,  $S_2 = span\{Y_1\}$ . 则  $S_1, S_2$  之间的主角  $0 \leq \theta_1 \leq \theta_2 \leq \dots \leq \pi/2$  [16].

定义

$$\cos \theta_k = \max_{u_k \in S_1} \max_{v_k \in S_2} u_k^T v_k \quad (4)$$

满足  $u_i^T u_i = v_i^T v_i = 1, u_i^T u_k = v_i^T v_j = 0$ . 其中,  $i, j = 1, 2, 3, \dots, k-1$ .

式(4)中,  $u_k$  和  $v_k$  是第  $K$  个典型向量对,  $\cos \theta_k$  的值称为典型相关系数. 计算主角的方法很多, 但是具有数值稳定的算法[15]是通过奇异值分解(SVD)来计算主角. 如两个子空间  $S_1, S_2$ , 它们的正交基分别是  $P_1 \in R^{D \times d_1}$ ,  $P_2 \in R^{D \times d_2}$ , 则  $P_1^T P_2$  的奇异值分解定义为:

$$P_1^T P_2 = Q_{12} \Lambda Q_{21}^T, \text{ 且满足 } \Lambda = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r) \quad (5)$$

其中,  $r = \min(\dim(S_1), \dim(S_2))$ ,  $Q_{12}$  和  $Q_{21}$  是正交矩阵,  $\Lambda$  矩阵中的对角线上的值是主角的  $\cos \theta_i$ , 即

$$\cos \theta_k = \sigma_k, k=1, 2, \dots, r$$

所以, 相关的典型向量  $U = P_1 Q_{12} = [u_1, u_2, \dots, u_{d_1}]$ ,  $V = P_2 Q_{21} = [v_1, v_2, \dots, v_{d_2}]$ . 从式(6)可以直观地看出, 如果最大的主角还是很小, 则这两个子空间就非常相似. 所以说通过主角来衡量两个子空间的相似性是有理论依据的.

### 2.2.2 子空间距离

两个子空间之间的相似度可以通过主角来反映. 经典的 MSM 方法利用最小主角来计算两个子空间距离:

$$D_{\max}(S_1, S_2) = \sqrt{1 - (\cos \theta_1)^2} = \sin \theta_1 \quad (7)$$

因此我们可以通过主角来计算两个子空间的距离. 图3所示为在三维坐标下模拟示意两个子空间的距离.

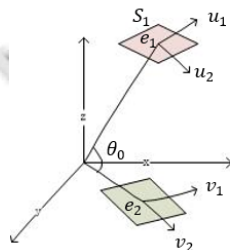


图3 子空间  $S_1, S_2$  之间的相似距离

其中  $e_1$  和  $e_2$  是两个局部空间的样本均值, 其主角为  $\theta_0$ . 传统的方法是采用局部模型的样本均值来表示局部模型特征. 由于原始数据的维度一般都很大,

为了便于后面高效的计算, 所以先通过 PCA 对每个局部模型进行降维, 得到 PCA 子空间, 然后在该子空间内计算局部模型的样本均值.

根据图3定义两个子空间的距离:

$$D_m(S_1, S_2) = \sqrt{1 - (\cos \theta_0)^2} \quad (8)$$

用样本均值的标准正交基表示:

$$D_m(S_1, S_2) = \frac{\|e_1\| \cdot \|e_2\|}{e_1^T e_2} \quad (9)$$

可以用距离的倒数来表示其相似度, 即将式(9)可以改写为相似度量:  $\frac{e_1^T e_2}{\|e_1\| \cdot \|e_2\|}$ .

很显然仅仅用样本均值不足以表现现实数据, 不同的人脸在姿势和光照变化不大的环境中得到的样本均值有可能很相似.

在经过 SVD[18]过程后, 发现不同的人脸变化方向和变化大小都不同. 所以将变化的因素考虑进去, 使其更具有判别性. 文献[19]中介绍了投影度量的概念, 两个子空间的投影距离可以表示为子空间的正交投影矩阵  $P_{S_1}, P_{S_2}$  的二范数值:

$$D_p(S_1, S_2) = \|P_{S_1}, P_{S_2}\|_2 \quad (10)$$

通过式(5)得到奇异值矩阵  $\Lambda$ , 该对角线上的值表示典型相关系数, 有着特殊的几何意义, 其值分别表示了在该维度上的变化程度, 且大小为  $\cos \theta_k = \sigma_k$ . 重新改写投影距离:

$$\begin{aligned} D_p(S_1, S_2) &= \sqrt{\sum_1^r (\sin \theta_i)^2} \\ &= \sqrt{r - \sum_1^r (\cos \theta_i)^2} \\ &= \sqrt{r - \sum_1^r (\sigma_i)^2} \end{aligned} \quad (11)$$

两个子空间的距离最终表示为:

$$D(S_1, S_2) = D_m(S_1, S_2) + D_p(S_1, S_2) \quad (12)$$

有了确定的子空间距离, 那么流形之间的距离也容易定义了.

### 2.3 流形与流形之间的距离

判断两个集合是否是属于同一个类别, 我们参照传统的距离度量来衡量两个流形之间的相似性. 上面我们已经讨论了线性子空间可以表示局部模型, 而且流形可以看作是线性局部空间的集合.

如两个流形  $M_1$  和  $M_2$ , 由线性局部模型表示:  $M_1 = \{S_1, S_2, \dots, S_m\}$ ,  $M_2 = \{S'_1, S'_2, \dots, S'_n\}$ , 用  $W_i$  表示流形  $M_1$  中第  $i$  个子空间到  $M_2$  上的最短距离, 同理,  $W_j$  表

示  $M_2$  中第  $j$  个子空间到  $M_1$  的最短距离. 定义如下:

$$W(i) = \arg \min_j D(S_i, S'_j) \quad (13)$$

$$W(j) = \arg \min_i D(S_i, S'_j) \quad (14)$$

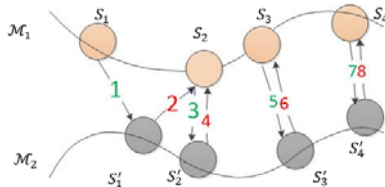


图 5 流形之间的距离

图中绿色数字连接的子空间对属于  $W_i$ , 红色数字连接的子空间对属于  $W_j$ , 为了更准确表述流形上的信息, 我们利用所有最短子空间对距离来计算流形距离, 所以两个流形之间的距离可以定义为:

$$D(M_1, M_2) = \frac{1}{m+n} \left( \sum_1^m D(S_i, S'_{w(i)}) + \sum_1^n D(S'_{w(j)}, S_j) \right) \quad (15)$$

### 2.4 算法实现步骤

综合上面的讨论, 总结下该方法的实现步骤, 算法如下:

训练过程:

(1) 给出训练集  $X_{train} = \{X_1, X_2, \dots, X_M\}$ , 共有  $M$  个图像集, 每个集合代表一个识别目标, 共有  $m$  个样本图像. 先计算每个图像集里  $m \times m$  的测地线距离和欧式距离以及  $R$  值.

(2) 用 PCA 对每个局部线性模型进行降维(保留 95% 的贡献率), 得到 PCA 子空间.

测试过程:

(1) 测试集  $X_{test} = \{X_1, X_2, \dots, X_N\}$ , 共有  $N$  个图像集, 每个集合有  $n$  幅图像, 构建局部线性模型

(2) 计算待测的图像集中子空间和训练数据中子空间之间的距离, 即  $D(S_i, S'_j)$ .

执行 1-NN 分类, 计算识别率.

## 3 实验及结果

为了验证该方法的有效性, 我们在 CMU\_PIE 人脸数据库进行实验, 并且与以下四种识别算法进行对比实验.

LLE +K-means<sup>[12]</sup>: 该方法是将流形学习中典型的局部线性嵌入算法和 K-均值聚类算法相结合. 在用 K-means 方法进行样本聚类时没有设置确定的 K 值, K 值选取完全依赖于实验经验.

MSM 算法: 该方法是基于子空间方法, 采用最小主角计算子空间距, 即选择子空间的极大距离. 而且该方法没有进行特征提取直接执行最近邻分类;

均值距离方法和投影距离方法指的是本文 2.2 章节讨论的基于局部子空间样本的均值距离和子空间的投影距离.

### 3.1 实验数据库

CMP PIE 人脸库由美国卡耐基梅隆大学创建, 包含 68 个人, 约有 41368 幅图像, 每个人大约有 170 幅图像, 包括 13 种姿态(如: 抬头, 低头, 侧脸等), 43 种不同光照和 4 种表情. 这些图像充分反映出人脸的位置、光照和表情等的变化, 实验过程中将每幅图像统一剪裁大小为  $32 \times 32$ , 每个人取 170 幅图像. 部分人脸图像如图 6 所示.



图 6 CMU\_PIE 部分人脸图像

### 3.2 对比实验结果

(1) IMM 算法与其他算法的对比实验

实验参数设置: 本实验中作对比实验的 LLE+k-means 方法中  $K=5$ , 所以算法的训练图像集数量为 50, 测试图像集数量为 120. 最后通过 1-NN 分类器来评价 IMM 算法和其他方法的性能. 表一给出了这几个方法在 CMU\_PIE 数据库上的结果.

表 1 不同算法在 CMU\_PIE 数据库上的识别率

| 算法          | 识别率(%) |
|-------------|--------|
| LLE+k-means | 82.235 |
| MSM         | 76.471 |
| 均值方法        | 77.941 |
| 投影方法        | 77.941 |
| IMM         | 92.647 |

从表 1 可以看出 IMM 方法的识别率远远高于其他四种方法, 表明本文提出的 IMM 方法是可行有效的, 在提高识别率方面有了很大的突破.

(2) 训练样本数量对识别率的影响

对于 CMP PIE 数据库, 取 68 个人, 即共有 68 个类别, 每个人 170 幅图像, 分别取每个人前 50, 100, 150 幅作为训练样本, 相应地剩下 120, 70, 20 幅图像作

为测试样本. 这三组实验的识别率情况如图 7 所示. 为了进一步分析, 求出三组训练样本数量和测试样本数量差值, 分别为 30,70,130. 图 8 所示这三种差值情况下的识别率.

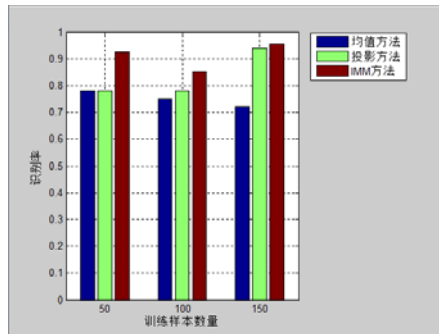


图 7 不同训练样本下的识别率

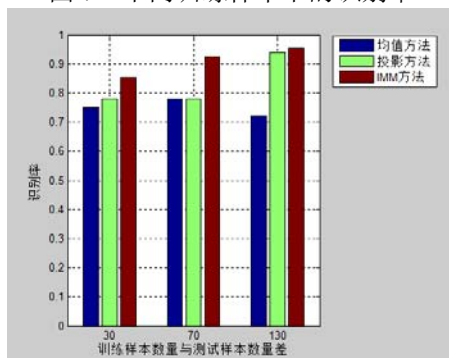


图 8 测试与训练样本差下的识别率

从图 7 来看, 输入的训练样本数对最后的识别率没有直接反映出规律性的影响, 但是从图 8 来看, 随着测试样本数量和训练样本的数量之差越大, 识别率也相应的越来越高, 而且相比另外两种方法, IMM 方法的识别率更高.

### (3) 参数 K 分析

本文中通过 K-NN 算法计算测地线距离. 选择训练图像集数量为 50, 测试图像集数量为 120. 图 9 所示 K 值在 5,10,15,20 下的识别率.

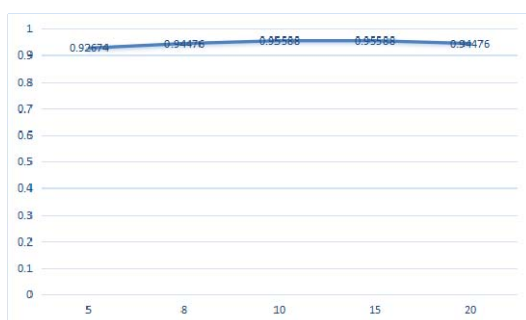


图 9 不同 K 值下的识别率

由图 9 看到  $k$  的值对于最后的识别率影响很小, 由此说明在流形上利用最大线性模型来构建线性子空间的方法比 LLE+K-means 方法优越.

## 4 结语

通过对比实验, 提出的 IMM 方法有较高的识别率. IMM 方法是将一个图像集建模为一个流形, 在该模型下构建局部线性模型, 利用主角的概念将局部子空间的均值距离和投影距离之和来衡量两个子空间的距离, 然后将所有最小子空间的距离取平均作为两个流形的相似度量.

训练样本参数的实验表明输入的训练样本数量对最后的识别率没有直接反映出规律性的影响, 但是随着测试样本数量和训练样本的数量之差越来越大, 识别率也相应的越来越高, 而且相同差值情况下, 相比均值方法和投影方法, IMM 方法的识别率仍然更高.

本文采用所有最小子空间对距离的平均值作为最终的流形距离, 这存在一定的缺陷, 因为没有考虑到全局数据分布, 仅仅将所有最近子空间对距离平均化, 无法避免出现较大变化的情况. 当出现变化较大时, 平均距离则无法更真实的衡量两个流形之间的相似性. 这个是以后要研究解决的问题.

## 参考文献

- Arandjelović O, Shakhnarovich G, Fisher J, Cipolla R, Darrell T. Face recognition with image sets using manifold density divergence. Proc. Comput. Vision Pattern Recog. Conf. 2005. 581–588.
- Shakhnarovich G, Fisher JW, Darrell T. Face recognition from long-term observations. Proc. Eur. Conf. Comput. Vision. 2002. 851–868.
- Wang W, Wang RP, Huang ZW, Shan SG, Chen XL. Discriminant analysis on riemannian manifold of Guassian distributions for face recognition with image sets. CVPR. 2015.
- Hamm J, Lee DD. Grassmann discriminant analysis: A unifying view on subspace-based learning. Proc. of the 25th International Conference on Machine Learning(ICML). 2008.
- Yamaguchi O, Fukui K, Maeda K. Face recognition using temporal image sequence. Proc. of the 3rd. Int. Conf. on Face & Gesture Recognition. Washington D.C., USA. IEEE

- Computer Society. 1998. 318.
- 6 Fukui K, Yamaguchi O. Face recognition using multi-viewpoint patterns for robot vision. *Int. Symp. of Robotics Res.* 2003. 192–201.
- 7 詹宇斌. 流形学习理论与方法及其应用研究[博士学位论文]. 长沙: 国防科学技术大学, 2011.
- 8 Lu JW, Wang G, Deng WH, Moulin P, Zhou J. Multi-manifold deep metric learning for image set classification. *CVPR.* 2015.
- 9 Yang W, Sun C, Zhang L. A multi-manifold discriminant analysis method for image feature extraction. *Pattern Recognition*, 2011, 44(8): 1649–1657.
- 10 Li SZ, Jain AK. *Handbook of Face Recognition.* New York: Springer-Verlag, 2005.
- 11 Kim TK, Arandjelović O, Cipolla R. Boosted manifold principal angles for image set-based recognition. *Pattern Recog.*, 2007, 40(9): 2475–2484.
- 12 Hadid A, Pietikinen M. From still image to video-based face recognition: An experimental analysis. *Proc. IEEE 6th Int. Conf. Autom. Face Gesture Recog.* 2004. 813–818.
- 13 Zhou S, Krüeger V, Chellappa R. Probabilistic recognition of human faces from video. *Comput. Vision Image Underst.*, 2003, 91(1): 214–245.
- 14 Wang R, Shan S, Chen X, Gao W. Manifold-manifold distance with application to face recognition based on image set. *Proc. Comput. Vision Pattern Recog. Conf.* 2008. 2940–2947.
- 15 Björck Å, Golub GH. Numerical methods for computing angles between linear subspaces. *Mathematics of Computation*, 1973, 27(123): 579–594.
- 16 Hotelling H. Relations between two sets of variates. *Biometrika*, 1936, 28: 321–372.
- 17 Tenenbaum J, Silva V, Langford J. A global geometric framework for nonlinear dimensionality reduction. *Science*, Dec. 2000, 290(22): 2319–2323.
- 18 冯欣. 奇异值分解 SVD 的理论分析及在图形学中的应用. *信息与电脑: 理论版*, 2010, (5).
- 19 Fan W, Yeung DY. Locally linear models on face appearance Manifolds with application to dual-subspace based classification. *IEEE Conference on Computer Vision & Pattern Recognition.* 2006. 1384–1390.
- 20 魏莱, 王守觉. 基于流形距离的半监督判别分析. *软件学报*, 2010, 21(10): 2445–2453.
- 21 肖睿. 高维空间模式鉴别分析及多流形学习[博士学位论文]. 上海: 上海交通大学, 2012.
- 22 朱林, 雷景生, 毕忠勤, 等. 一种基于数据流的软子空间聚类算法. *软件学报*, 2013, 24(11): 2610–2627.