基于KFD-Isomap的人脸识别

李睿凡[1] 郝红卫[2] 涂序彦[2] 王枞[1]
（1 北京邮电大学信息工程学院 北京 100876 2 北京科技大学信息工程学院 北京 100083）

摘 要： 神经生理学近来强调视觉以流形方式感知外部对象并提出 Isomap 的流形学习方法。该方法在数据描述等问题上取得良好效果，但从模式分类观点看并非最优。以此提出 KFD-Isomap 算法，用核 Fisher 判别替代 Isomap 中的经典多维尺度分析，并应用于人脸识别问题。两个人脸数据库的识别实验表明了该算法的有效性。

关键词： 人脸识别 流形 Isomap 核 Fisher 判别

Face Recognition Based on KFD-Isomap

Li Ruifan[1] Hao Hongwei[2] Tu Xuyan[2] Wang Cong[1]
([1]Beijing University of Posts and Telecommunications, Beijing, 100876)
([2]University of Science and Technology Beijing, Beijing, 100083)

Abstract: Recently, neuroscientists emphasized the manifold ways of perception and proposed Isomap for manifold learning. Favorable results have been achieved using Isomap for data description and visualization. However, since the Isomap is developed based on minimizing the reconstruction error with multidimensional scaling, it may not be optimal from the perspective of pattern classification. Therefore, an improved version of Isomap, namely KFD-Isomap, is proposed using kernel Fisher discriminant for face recognition. Experiments on two face databases have shown its feasibility.

Keywords: Face recognition; manifold; isomap (isometric feature mapping); kernel Fisher discriminant (KFD).

1 引 言

人脸识别是计算机视觉与模式识别的研究热点，最新的综述见文献[1]。子空间方法试图提取、分析高维人脸图像空间中潜在的低维子空间解决人脸识别问题。其典型方法有基于线性主成分分析的特征脸和应用 Fisher 线性判别的 Fisher 脸。进而，通过核函数的引入，这两种方法扩展成核特征脸、核 Fisher 脸方法。实验表明，核函数的引入能够在一定程度上克服人的姿态差异、外界光照强度与方向变化等负面因素，进而提高了人脸识别的准确率[2]。

近来，神经生理学研究认为人的视觉以流形的方式感知外部对象[3]。流形，简言之，是一个局部欧式的拓扑空间。Isomap(Isometric Feature Mapping)[4]是一种非线性流形学习的方法。该方法在全局坐标下用局部几何度量学习非线性的嵌套流形，从而达到维数约简的目的。其基本点在于，测地线距离(沿流形表面的距离测度)反映了潜在流形的真实几何特征，如图 1 所示。从而，Isomap 算法用点间的欧氏距离估计点间的测地线距离，然后使用经典的多维尺度分析保持数据内在几何性质。

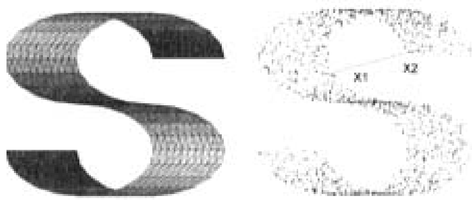

图 1 S 形曲面图：测地线距离真实反映了潜在流形的特征

具体地，Isomap 算法首先计算输入空间中任意点间的欧式距离，以该测度确定点的邻域关系，构造出一个以点间欧式距离为权值的图。邻域点的选择有两种方法：将空间中某点与其它点间欧式距离最近的 K 个点作为该点的邻域点；定义一个以某点为球心，以 ε 为半径的超球，将该超球内的点作为该点的邻域。邻域参数为该算法的唯一参数。然后以欧式距离估计流形上的测地线距离，更新图矩阵：对于邻域内的点，直接用欧式距离替代测地线距离；对于非邻域内的点，用图中连接两点的最短距离作为测地线距离的估计。最后用经典的多维尺度分析

处理图矩阵，得到保持几何结构的内嵌流形。算法的实现细节见文献[4]。

尽管 Isomap 的非线性流形学习方法在数据的维数约简与数据可视化等数据描述问题上表现出良好的性能。但是，从模式分类的观点看，其数据处理方式并非是最优的。首先，它目标是保持最小的数据重建误差，而非区分不同类别；其次，它要求良好采样的样本数据集，而实际模式分类应用的很多场合却是数据的稀疏，这在人脸识别问题上较为突出。最后，其获取的从输入空间到嵌套空间映射为隐式表达，因此与数据可视化相比，Isomap 对于模式分类问题还需要额外的步骤，即用广义回归神经网络[5]等非线性拟合方法获得映射的显式表达。基于第一点的考虑，Yang[6]提出了 Ext-Isomap 算法，将 Fisher 线性判别的方法引入到 Isomap 中，替代多维尺度分析寻找最佳投影方向。

本文提出将基于核函数的 Fisher 判别方法引入 Isomap 算法。与 Isomap、Ext-Isomap 相比，该算法对于测地线距离的估计采用相同的方法，然后将测地线距离矩阵的列向量作为核 Fisher 判别的输入特征向量以寻找最佳投影方向。最后使用两个人脸数据库，比较 KFD-Isomap、Isomap、Ext-Isomap、特征脸、Fisher 脸等识别算法，实验结果表明了该方法的有效性。

2 KFD-Isomap 算法

一般地，考虑集合 $X=\bigcup_{i=1}^{c}X_i$，c 为子集数，即样本类别；$\sum_{i=1}^{c}n_i=n$，n_i 为每类样本数，n 为样本总数；$x_i\in R^N$，每个样本属于 N 维空间。首先用欧式距离度量 $d_X(x_i,x_j)=\|x_i-x_j\|_2$ 计算样本集中任意两点之间的欧式距离，用以确定任意一个样本点 x_i 的邻域。K-Isomap 取与点 x_i 较近的 K 个点作为该点的邻域；ε-Isomap 将以该点为球心，以 ε 为半径的超球内的点作为邻域。接着，可以将点间的邻域关系表示为带权值的图 G，即，如果 x_i 与 x_j 有邻接关系，则矩阵 G 的元素 $G_{ij}=d_X(x_i,x_j)$；否则 $G_{ij}=\infty$。

然后，依据邻接关系图 G 估计嵌套流形 M 上的测地线距离 $d_M(x_i,x_j)$。对于邻近点 x_i 和 x_j，其测地线距离即为 $d_G(x_i,x_j)$；对于非邻域点，可以采取 Floyd 算法，即：

$$d_G(x_i,x_j)=\min(d_G(x_i,x_j),d_G(x_i,x_k)+d_G(x_k,x_j))$$

进一步，将图矩阵 G 记为 $\{g_1,g_2,\cdots,g_n\}$，g_i 为对应于样本 x_i 的特征向量。

以上的处理与 Isomap 算法一致。接着，Isomap 算法以图 G 作为经典的多维尺度分析的输入，找到嵌套的本征结构；Ext-Isomap 则以每个样本点对应的测地线距离矢量作为 Fisher 线性判别的输入，寻找最佳的投影矩阵。本文采用核 Fisher 判别对测地线距离矢量进行处理，先将其映射到高维核空间，然后用 Fisher 线性判别寻找最佳投影矩阵，使其在分类意义上最优，利用核函数将高维空间中的点积运算转换为输入空间中函数求值简化计算。

假定非线性映射 $\phi:R^g\to R^f$，R^f 为高维特征空间，有 $f\geq g$。那么，定义高维特征空间中的类内离散度矩阵与类间离散度矩阵分别如下：

$$S_w^\phi=\sum_{i=1}^{c}\sum_{j=1}^{n_i}(\phi(g_j)-\mu_i^\phi)(\phi(g_j)-\mu_i^\phi)^T \quad (1)$$

$$S_b^\phi=\sum_{i=1}^{c}n_i(\mu_i^\phi-\mu^\phi)(\mu_i^\phi-\mu^\phi)^T \quad (2)$$

其中，u_i^ϕ 与 u^ϕ 分别为映射空间上样本类 i 的均值与总体均值且 $u_i^\phi=\frac{1}{n_i}\sum_{j=1}^{n_i}g_j$，$u^\phi=\frac{1}{n}\sum_{i=1}^{n}g_i$。

如果 S_w^ϕ 为非奇异矩阵，则最优的投影 W_{opt} 可以由使得类间离散度矩阵与类内离散度矩阵的行列式比值最大的正交列向量构成，即最佳投影矩阵为：

$$W_{opt}=\arg\max_W\frac{|W^TS_b^\phi W|}{|W^TS_w^\phi W|} \quad (3)$$

W_{opt} 由 S_b^ϕ 和 S_w^ϕ 的广义特征向量组成，其列向量 w 为较大特征值 λ 对应的特征向量，即广义特征方程，$S_b^\phi w=\lambda S_w^\phi w$。

依据核的复制理论[7]，任意的 w 都是样本特征向量的线性组合，即：

$$w=\sum_{i=1}^{n}\alpha_i\phi(g_i)=\Phi\alpha \quad (4)$$

其中，$\Phi=[\phi(g_1),\phi(g_2),\cdots,\phi(g_n)]$，系数列向量 $\alpha=[\alpha_1,\alpha_2,\cdots,\alpha_n]^T$。

图 2　ORL 人脸数据库中部分图像

将 $\phi(g_j)$ 投影到 w 上，利用(4)有：

$$\begin{aligned} w^T\phi(g_j) &= (\Phi\alpha)^T\phi(g_j) \\ &= \alpha^T\Phi^T\phi(g_j) \\ &= \alpha^T[k(g_1,g_j),\cdots,k(g_n,g_j)]^T \\ &= \alpha^T\zeta_j \end{aligned} \tag{5}$$

这里，$\zeta_j=[k(g_1,g_j),k(g_2,g_j),\cdots,k(g_n,g_j)]^T$；同时，核函数 $k(x,y)=<x\cdot y>$，$<x\cdot y>$ 为点积运算。其中的核函数可以为高斯径向基函数 $k(x,y)=\exp(-\|x-y\|^2/c)$，或多项式核函数 $k(x,y)=(x\cdot y)^d$，其中的 c、d 均为正常数。

同理可得：

$$w^T\mu_j^\phi=\alpha^T m_j \tag{6}$$

$$w^T\mu^\phi=\alpha^T m \tag{7}$$

其中，$m_j=\frac{1}{n_j}\left[\sum_{i=1}^{n_j}k(g_1,g_i),\cdots,\sum_{i=1}^{n_j}k(g_n,g_i)\right]^T$

$$m=\frac{1}{n}\left[\sum_{i=1}^{n}k(g_1,g_i),\cdots,\sum_{i=1}^{n}k(g_n,g_i)\right]^T$$

考虑(3)的分子、分母可以得到：

$$w^TS_b^\phi w=\alpha^TK_b\alpha$$

$$w^TS_w^\phi w=\alpha^TK_w\alpha$$

其中，$K_b=\sum_{i=1}^{c}n_i(m_i-m)(m_i-m)^T$

$$K_w=\sum_{i=1}^{c}\sum_{j=1}^{n_i}n_i(\zeta_j-m_i)(\zeta_j-m_i)^T$$

那么，有如下的广义特征值问题，

图 3　Yale 人脸数据库中部分图像

$$A=\arg\max_{A}=\frac{\left|A^TK_bA\right|}{\left|A^TK_wA\right|}=[\alpha_1,\alpha_2,\cdots,\alpha_{c-1}] \tag{8}$$

对于类内的离散度矩阵 K_w 的奇异性，可以采用两种方式进行处理：给 K_w 的每个对角线元素加上一小正数，改变其为非奇异矩阵，即 $K_w=K_w+\sigma I$，σ 为一个足够小的正数；或者采用主成分分析的方法，见文献[8]。

如果有待测试样本 y，那么其在 w 方向上的投影可以计算如下：

$$\begin{aligned} w^T\phi(y) &= \alpha^T\Phi^T\phi(y) \\ &= \alpha^T[k(g_1,y),k(g_2,y),\cdots,k(g_n,y)]^T \end{aligned}$$

至此，KFD-Isomap 算法用于模式分类的流程可小结如下：

构建邻域图：计算输入空间任意点间的欧式距离，定义每个点的邻域，邻域图 G 的权值为点间距离。

1) 计算最短路径：用 Floyd 算法求取图中点的最短距离。
2) 确定投影向量：将每个点表达为图矩阵的特征向量，使用核 Fisher 判别求解广义特征值问题。

分类测试样本：构造一个广义回归神经网络，映射测试样本到权值图，并用核 Fisher 判别投影矩阵将特征映射到子空间，最后用最近邻分类器确定其类别。

3 人脸识别实验

为检验改进算法的有效性，本文用 KFD-Isomap、Isomap、Ext-Isomap、特征脸、Fisher 脸等对 ORL 和 Yale 两个人脸数据库进行了人脸识别实验。ORL 人脸数据库包含 36 位男性和 4 位女性，每人 10 幅图像，共 400 幅面部图像，每幅图像为 112×92 象素，256 灰度级。图像包括光照、表情（眼睛的开闭、是否微笑）和面部细节（是否佩戴眼镜）等成像条件的轻微变化。其图像实例如图 2。为减少计算量，图像被重采样为 56×46 象素

大小。Yale 人脸数据库包括 15 个人，大小为 320×243 的 165 幅人脸图像。图像包含光照、面部表情、及是否佩戴眼镜等成像条件的较大变化。为减少背景图像的干扰，图像被裁剪为 192×236 大小，然后重采样为 48×59 大小以减少计算量。图 3 为部分裁剪后的样本图像。

实验中，将每个图像看成高维空间中的点，每幅图像表示为一个列向量。先进行数据的归一化处理，即将数据变换为均值为零，方差为一的列向量

表 1 ORL 数据库实验结果

方法	邻域参数	降维空间	错误率(%)
Eigenface	NA	35	3.00
Fisherface	NA	39	1.75
Isomap	ε=12	45	3.50
Ext-Isomap	ε=12	39	1.25
KFD-Isomap	ε=15	39	0.75

表 2 Yale 数据库实验结果

方法	邻域参数	降维空间	错误率(%)
Eigenface	NA	35	26.67
Fisherface	NA	14	9.70
Isomap	ε=15	40	26.06
Ext-Isomap	ε=12	14	19.39
KFD-Isomap	ε=15	14	8.48

集合，然后分别用上述五种算法作为特征提取的步骤，在子空间上采用最近邻分类器。数据集合的组织采取 leave-one-out 策略，即每次将 $n-1$ 个样本作为训练集，剩余的一个样本作为测试集。算法中的参数由实验确定。文献[7]的人脸实验表明，ε 邻域定义比 K 邻域定义有较好的实验结果，所以，文中只考虑 ε 的邻域定义。另外，多项式核函数采用平方次 d 为 2 核函数。

两个人脸数据库的实验结果如表一、二。KFD-Isomap 在几种方法中表现出最好的性能。由于基于主成分分析的特征脸与基于多维尺度分析的 Isomap 方法都从重建误差最小的角度出发，所以它们具有类似的识别效果。KFD-Isomap 与 Ext-Isomap 方法中判别分析的引入，使得 Isomap 方法的识别率有较大提高。

4 结论

提出了一种用核 Fisher 判别函数改进的 Isomap——KFD-Isomap 进行特征提取，并用于人脸识别的方法。该方法用测地线距离矩阵的列向量作为特征，并用核 Fisher 判别替代多维尺度分析建立最佳投影方向。人脸识别实验中，以基于 Isomap 的算法作为特征提取手段，并使用最近邻分类器与特征脸、Fisher 脸算法进行了比较。两个人脸数据库的实验表明该方法的性能优于实验采用的其它方法。此外，该方法还可应用于其它模式分类问题。

参考文献

[1]W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Computing Surveys, vol. 35, pp. 399-458, 2003.

[2]M.-H. Yang, "Kernel Eigenfaces vs. kernel Fisherfaces: face recognition using kernel methods," presented at Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition, 14-16 April 2002, 2002.

[3]H. S. Seung and D. D. Lee, "The manifold ways of perception," Science, vol. 290, pp. 2268-2269, 2000.

[4]J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, vol. 290, pp. 2319-2323, 2000.

[5]D. F. Specht, "A general regression neural network," IEEE Transactions on Neural network, vol. 2, pp. 568-576, 1991.

[6]M. H. Yang, "Extended isomap for pattern classification," presented at 18th National Conference on Artificial Intelligence (AAAI-02), Jul 28-Aug 1 2002, Edmonton, Alta., Canada, 2002.

[7]S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. R. Muller, "Fisher discriminant analysis with kernels," presented at Neural Networks for Signal Processing - Proceedings of the IEEE Workshop, Madison, WI, USA, 1999.

[8]P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 711-720, 1997.

作者简介

李睿凡：男，1975 年生。1998 年，2001 年毕业于华中理工大学分获工学学士、工学硕士学位。现北京邮电大学攻读信号与信息处理博士学位，主要研究方向为：模式识别、机器学习与计算机视觉。