保局投影算法的优化研究

赵继东，鲁 珂，吴 跃

(电子科技大学计算机科学与工程学院 成都 610054)

【摘要】保局投影算法的基础是构造一个模拟图像局部结构的最近相邻图，但该最近相邻图并不总能够准确表示图像的流形结构，该文提出了一种基于保局投影的迭代保局投影优化算法。该方法可以不断地迭代更新保局投影算法的最近相邻图，最近邻图的构成直接影响到保局投影算法的性能，因此，优化后的最近相邻图可以更好地表示出图像的流形结构。从而可以得到更佳的降维映射。对该算法与PCA及LPP的图像检索效果进行实验比较，结果表明，该算法可以获得更好的效果。

关 键 词 图像检索； 迭代保局投影算法； 保局投影； 流形学习

中图分类号 TP391.4 **文献标识码** A

Research on the Optimization of Locality Preserving Projections

ZHAO Ji-dong, LU Ke, and WU Yue

(School of Computer Science and Engineering, University of Electronic Science and Technology of China Chengdu 610054)

Abstract Locality Preserving Projections (LPP) is based on a nearest neighbor graph which models the local geometrical structure of the image manifold. However, this graph can not always accurately estimate the intrinsic manifold structure. A novel algorithm called Iterative locality preserving projections (ILPP) is preposed. ILPP iteratively updates the nearest neighbor graph, so that it can better model the intrinsic manifold structure. Experimental results comparison show that our algorithm outperforms PCA and LPP for image retrieval.

Key words image retrieval; iterative locality preserving projections algorithm; LPP; manifold learning

随着数字图像容量的快速增长，对于图像管理工具效率的要求也越来越高。因此，提高基于内容的图像检索(content based image retrieval，CBIR)的效率也就成为了当前的研究热点。

图像检索的一个基本问题是如何对图像进行适当的表示。通常，图像的表示是通过它的一些特征(如颜色、纹理、轮廓等)来实现的，一般是用高维欧几里德空间中的一些特征矢量来表示。本文把图像特征矢量组成的集合定义为“图像空间”，尽管图像空间的总的维数相当高，但它的内在维数却可能要低一些。另外，有些研究也表明图像空间可能是一种流形[1-2]。

传统的图像流形的学习算法有主成分分析(principal component analysis，PCA)[3]和保局投影(locality preserving projections，LPP)[4]等。PCA沿着特征变化最大的方向投影图像，该算法从重构误差的角度来说是有优势的。当图像流形是线性的情况下，PCA可以有效地得到流形的表示，但在多数情况下，图像流形却是高度非线性的。

近年来，国内开展了一些非线性数据降维的研究[5-6]，而且针对LPP算法也有一些研究[7-9]。LPP是一种试图找到流形的局部结构的线性算法，它通过构造一个最近相邻图来模仿流形的局部结构，根据该图最终获得图像的低维映射[10]。然而，LPP算法的问题在于，最近相邻图并不能确保很好地模仿流形的局部结构。本文提出了一种新型的流形学习算法，即迭代保局投影(iterative locality preserving projections，ILPP)算法。该算法迭代地更新最近相邻图，从而使之能更好地模仿流形的局部结构，然后根据随时更新的最近相邻图来获得图像的投影，最终在投影得到的子空间中进行图像检索。

1 PCA及LPP简介

PCA是一种普遍采用的用于子空间学习算法，它的基本思想是提取最具有代表性的特征，并使重构误差最小化。假定有一个图像集$\{x_1,x_2,\cdots,x_n\}\subset R^n$，设$\boldsymbol{a}$为转换向量，$\boldsymbol{y}_i=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}_i$，其中$\boldsymbol{x}_i$、$\boldsymbol{y}_i$亦为向量。如均值$\overline{y}=\frac{1}{m}\sum\boldsymbol{y}_i$、$\overline{x}=\frac{1}{m}\sum\boldsymbol{x}_i$，

收稿日期：2007－04－24；修回日期：2007－12－26

基金项目：国家自然科学基金(60702072)；四川省应用基础研究基金(2006JB-67)

作者简介：赵继东(1976－)，男，在职博士生，讲师，主要从事网络与信息系统方面的研究；鲁 珂(1974－)，男，副教授，主要从事网络多媒体技术方面的研究；吴 跃(1958－)，男，教授，博士生导师，主要从事网络与信息系统方面的研究.

PCA的目标函数可以表示如下：

$$\boldsymbol{a}_{\rm opt}=\arg\max_{\boldsymbol{a}}(\boldsymbol{y}_i-\overline{y})^2=$$
$$\arg\max_{\boldsymbol{a}}\boldsymbol{a}^{\rm T}\left(\frac{1}{m}\sum_i(\boldsymbol{x}_i-\overline{x})(\boldsymbol{x}_i-\overline{x})^{\rm T}\right)\boldsymbol{a}=$$
$$\arg\max_{\boldsymbol{a}}\boldsymbol{a}^{\rm T}\boldsymbol{C}\,\boldsymbol{a}$$

该函数满足条件$\boldsymbol{a}^{\rm T}\boldsymbol{a}=1$，其中，$\boldsymbol{C}$为协方差矩阵。该目标函数的解就是矩阵$\boldsymbol{C}$的特征向量，因为$\boldsymbol{C}$是对称的，所以PCA的基底函数是直交的。

与PCA提取最具有代表性的特征不同，LPP提取最具有判别性的特征。假定有一个相似性矩阵$\boldsymbol{S}$定义为：(1) 当$\boldsymbol{x}_i$是$\boldsymbol{x}_j$的K-近邻或者$\boldsymbol{x}_j$是$\boldsymbol{x}_i$的K-近邻时，$S_{ij}={\rm e}^{\frac{\|x_i-x_j\|^2}{t}}$，$t$为一常量；否则，$S_{ij}=0$。LPP可以通过解下面的最小值问题来实现：

$$\boldsymbol{a}_{\rm opt}=\arg\min_{\boldsymbol{a}}\sum_{i=1}^{m}(\boldsymbol{a}^{\rm T}\boldsymbol{x}_i-\boldsymbol{x}_i{}^{\rm T}\boldsymbol{a})^2S_{ij}=$$
$$\arg\min_{\boldsymbol{a}}\boldsymbol{a}^{\rm T}\mathbf{XLX}^{\rm T}\boldsymbol{a} \qquad (1)$$

但需满足条件$\boldsymbol{a}^{\rm T}\mathbf{XDX}^{\rm T}\boldsymbol{a}=1$。本文定义一个对角权矩阵$\boldsymbol{D}$，它的元素就是对称矩阵$\boldsymbol{S}$的行(或列)元素的数值和。进一步定义矩阵$\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{W}$，该矩阵也称为拉普拉斯矩阵或拉普拉斯算子[1]，$D_{ii}=\sum_j S_{ij}$表示$\boldsymbol{x}_i$附近的局部密度，LPP的基底函数即是矩阵$(\mathbf{XDX}^{\rm T})^{-1}\mathbf{XLX}^{\rm T}$的对应于那些选取的较小特征值的特征向量。因为矩阵$(\mathbf{XDX}^{\rm T})^{-1}\mathbf{XLX}^{\rm T}$通常并不是对称的，所以LPP的基底函数不是直交的。

一旦特征向量被计算出来后，设$\boldsymbol{A}_k=(a_1,a_2,\cdots,a_k)$作为转换矩阵，则降维空间中两个数据点之间的欧几里德距离为：

$$\mathrm{dist}(\boldsymbol{y}_i-\boldsymbol{y}_j)=\|\boldsymbol{y}_i-\boldsymbol{y}_j\|=\|\boldsymbol{A}^{\rm T}\boldsymbol{x}_i-\boldsymbol{A}^{\rm T}\boldsymbol{x}_j\|=$$
$$\sqrt{(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\rm T}\boldsymbol{A}\boldsymbol{A}^{\rm T}(\boldsymbol{x}_i-\boldsymbol{x}_j)}$$

如果$\boldsymbol{A}$是正交矩阵，$\boldsymbol{A}\boldsymbol{A}^{\rm T}=\boldsymbol{I}$，则度量结构可以保持不变。

2 迭代保局投影算法

上面已经提到，LPP的一个问题是：最近相邻图并不能够总是确保准确地表示出图像的流形结构。在LPP的基础上，本文对LPP的算法进行改进，改进的算法可以迭代地更新最近相邻图，从而使得对图像的表示更加准确。本文把改进优化后的学习算法称为“迭代保局投影算”。下面是该算法的主要步骤。

(1) PCA投影。把图像$\boldsymbol{x}_i$投影到PCA子空间中(该子空间通过去除零特征值对应的特征向量获得)，通过PCA投影，可以使提取的特征之间是不相关的，并且可使新数据矩阵的阶数与特征个数(维数)相同。

(2) 创建最近邻图。假设G为一个具有n个节点的图，其中的第i个节点对应图像矢量$\boldsymbol{x}_i$。如果节点i是节点j的K−近邻或者节点j是节点i的K−近邻，则在节点i与j之间用一条边连接。如果图像属于不同的语义类，则每个类别的图像会构成各自类别的最近邻图。也就是说：只有当两个节点属于同一类别时，才能连接它们。

(3) 确定权重。与LPP算法一样，用相似性权矩阵$\boldsymbol{S}$模拟图像的局部结构。如果节点i与j之间是相连的，取$S_{ij}=e^{\frac{\|x_i-x_j\|^2}{t}}$，其中$t$是一个常量；否则，取$S_{ij}=0$。

(4) 计算保局投影。若n是当前图像空间的维数，设$\{a_1,a_2\cdots,a_n\}$是特征问题$\mathbf{XLX}^{\rm T}\boldsymbol{a}=\lambda\mathbf{XDX}^{\rm T}\boldsymbol{a}$的广义特征向量的集合，假定$\boldsymbol{W}_{\rm LPP}=[a_1,a_2,\cdots,a_{n-1}]$，可以将图像投影到一个$n$−1维的子空间$\boldsymbol{x}\leftarrow\boldsymbol{W}_{\rm LPP}^{\rm T}\boldsymbol{X}$中。本文中，$\boldsymbol{x}$是经过更新的$n-1$维的图像表示。

(5) 终止条件。如果当前图像空间的维数(n−1)等于图像的类别数，迭代停止；否则，回到步骤(2)，更新最近邻图，算法继续。

3 理论分析

用最小割图分解的方法同样可以推导出式(1)。一般地，一个图G分解成两个独立子图G_1和G_2可以表示为$G=G_1\cup G_2$，且$G_1\cap G_2=\phi$。最小割图的标准可以表示为$\min\limits_{G_1,G_2}\sum\limits_{i\in G_1}\sum\limits_{j\in G_2}S_{ij}$。

对于图中每一个节点(数据点)$\boldsymbol{x}_i$，本文用一个数据$\boldsymbol{y}_i$来表示$\boldsymbol{x}_i$属于子图G_1($\boldsymbol{y}_i$ = 1)或者属于G_2($\boldsymbol{y}_i$ = −1)，$\boldsymbol{y}_i\in\{-1,1\}$。如果$\boldsymbol{x}_i$和$\boldsymbol{x}_j$属于同一子图，则$\boldsymbol{y}_i-\boldsymbol{y}_j=0$；否则，$|y_i-y_j|=2$。因而，该最小割问题可以简化为求出目标函数$\min\limits_{\boldsymbol{y},\boldsymbol{y}_i\in\{-1,1\}}\sum\limits_{i,j}(\boldsymbol{y}_i-\boldsymbol{y}_j)^2S_{ij}$的最优解$\boldsymbol{y}=(y_1,y_2,\cdots,y_m)^{\rm T}$。

因为$\boldsymbol{y}_i$只能为1或者−1，使得上面问题的求解比较复杂。为了降低复杂度，一个办法是使用光谱松弛法来近似地估计最优解。可以先放宽条件$\boldsymbol{y}_i\in\{-1,1\}$，使$\boldsymbol{y}_i$可以为任意实数；进而，如果限制从$\boldsymbol{x}_i$到$\boldsymbol{y}_i$的映射是线性的，即$\boldsymbol{y}_i=\boldsymbol{a}^{\rm T}\boldsymbol{x}_i$，则可以得到式(1)。用光谱图分解理论知，当子空间的维数等于图像的种类数时，可以获得最佳的效果。因此，在本文提出的ILPP算法中，最终将图像投影到一个k

维子空间中，k为图像的种类数。

4 实验分析

下面，对采用PCA算法、标准LPP算法和ILPP算法在降维空间中进行检索的实验效果作出比较。在图1中，可以看到前20个返回结果的准确度；在图2中，可以看到前50个返回结果的准确度。分析实验结果可以作出一个大致的比较：在7次反馈后，对于前20个返回结果(图1)，PCA算法可以达到56%的准确度，LPP算法可以达到60%的准确度，而ILPP算法可以达到64%的准确度。

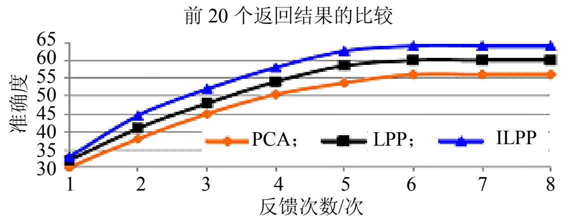

图1 前20个返回结果的检索准确度

对于前50个返回结果(图2)，PCA算法可以达到48%的准确度；LPP算法可以达到50%的准确度；ILPP算法可以达到55%的准确度。可以看出，ILPP算法的检索效果要优于PCA算法和标准LPP算法。

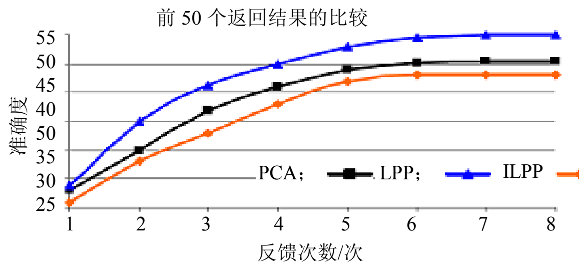

图2 前50个返回结果的检索准确度

5 总 结

本文提出了一种基于LPP的新型迭代学习算法，该算法迭代地更新最近相邻图，相对LPP算法可以更好地表示出图像的流形结构。本文的理论分析证明，当子空间的维数等于图像的种类数时，图像检索可以获得最佳的效果。因此，在本文提出的ILPP算法中，最终将图像投影到一个k维子空间中。实验结果显示，该算法可以有效地提高检索的准确度。

参 考 文 献

[1] VAPNIK V. The nature of statistical learning theory[M]. New York: Springer Verlag, 1995.

[2] JOACHIMS T. Transductive inference for text classification using support vector machines[C]//16th International Conference on Machine Learning. San Francisco, USA: [s.n.], 1999.

[3] HE Xiao-fei, KING O, MA Wei-ying, et al. Learning a semantic space from user's relevance feedback for image retrieval[J]. IEEE Trans on Circuit and Systems for Video Technology, 2003, 13(1): 39-48.

[4] HE Xiao-fei, Niyogi P. Locality preserving projections[C]// Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2004: 327-334.

[5] 何 力, 张军平, 周志华. 基于放大因子和延伸方向研究流形学习[J]. 计算机学报, 2005, 28(12): 2000-2009.
HE Li, ZHANG Jun-ping, ZHOU Zhi-hua. Investigating manifold learning algorithms based on magnification factors and principal spread directions[J]. Chinese Journal of Computers, 2005, 28(12): 2000-2009.

[6] 张振跃, 查宏远. 线性低秩逼近与非线性降维[J]. 中国科学A辑(数学), 2005, 35(3): 273-285.
Zhang Zheng-yue, Zha Hong-yuan. Linear low-rank approximation and onlinear dimensionality reduction[J]. Science in China (Series A: Mathematics), 2005, 35(3): 273-285.

[7] LU Ke, HE Xiao-fei. Image retrieval using dimensionality reduction[J]. Lecture Notes in Computer Science, 2004, 3314: 775-781.

[8] LU Ke, HE Xiao-fei. Image retrieval based on incremental subspace learning[J]. Pattern Recognition. 2005, 38(11): 2047-2054.

[9] HE Xiao-fei, CAI Deng, MIN Wan-li. Statistical and computational analysis of locality preserving projection[C]// Proceedings of the 22nd International Conference on Machine Learning. Bonn, Germany: ACM Press, 2005: 281-288.

[10] LU Ke, ZHAO Ji-dong, CAI Deng. An algorithm for semi-supervised learning in image retrieval[J]. Pattern Recognition, 2006, 39(4): 717-720.

编 辑 熊思亮