

保局投影算法在图像检索中的应用研究^{*}

鲁珂, 赵继东, 叶娅兰, 曾家智

(电子科技大学 计算机科学与工程学院, 四川 成都 610054)

摘要: 对国际上近来提出的保局投影(LPP)算法在图像检索中的最新应用研究进行了详细介绍;分析指明了几种基于LPP的图像检索算法的特点;设计并完成了基于LPP的图像检索算法图像检索效果的比较实验;最后根据实验结果总结了各类算法的优缺点。

关键词: 保局投影; 图像检索; 流形学习

中图法分类号: TP317.4 文献标识码: A 文章编号: 1001-3695(2006)12-0056-03

Research on Algorithms of LPP for Image Retrieval

LU Ke, ZHAO Ji-dong, YE Ya-lan, ZENG Jia-zhi

(College of Computer Science & Engineering, University of Electronic Science & Technology of China, Chengdu Sichuan 610054, China)

Abstract: In this paper, several algorithms of Locality Preserving Projections(LPP) for image retrieval are introduced, the basic characteristic of algorithms based on LPP are analyzed and indicated. Several experiments to evaluate the effectiveness of these algorithms are performed, our conclusion according to experimental results are proposed.

Key words: LPP(Locality Preserving Projections); Image Retrieval; Manifold Learning

由于数字图像应用的飞速增长,对有效图像管理工具的需求也与日俱增,因此,基于内容的图像检索(CBIR)技术正在成为国际范围的研究热点。近来很多研究人员致力于解决由于图像特征空间的高维数所带来的问题。一般说来,图像特征空间的维数由几十到几百不等,如一个颜色直方图可能包含256个颜色子区间(Bin)。高维数给图像检索带来了一些问题:

由于维数灾难(随着维数的增长,所需要的样本数呈指数级增长)的原因,如果样本必须依赖高维表示的话,从样本中学习在计算上是不可行的,因此,降维是可学性的必要条件。在多媒体数据库中,高维数的图像表示在计算上的开销很大(计算密集),而绝大部分用户并不想为搜索结果等得太久,对于提高存储量和效率来说,进行降维是必需的。

主成分分析(PCA)是进行降维的一种经典技术。它通过保留原始特征(n 维)协方差矩阵的 m ($m < n$)个占显著贡献的特征向量(这些特征向量是由原始特征线性变换得到的)来构建新的 m 维空间,将原始的 n 维数据映射到新的 m 维空间中进行降维。对于线性的内嵌流形,PCA可以确保发现流形的维数并产生一个基于正交主成分因子形式的紧凑数据表示,然而当数据位于非线性子流形时,PCA将不能找到内嵌的架构。

我们研究使用保局投影(Locality Preserving Projects, LPP)来进行图像检索。LPP是一种最近提出的用于流形学习的算法,当数据位于非线性子流形时,LPP可以很好地保留数据的局部信息。近来有很多研究表明,图像数据可以被表示为非线性的流形结构,因此,当采用基于LPP的算法进行图像检索时,它可以很好地保留图像的局部信息。在许多图像检索任务

中,这些局部信息比全局信息更加有效,因此对于图像检索来说,基于LPP的算法可以获得最佳的效果。本文分析几种基于LPP的图像检索算法的特点,然后根据实验结果总结了这几种算法的优缺点。

1 LPP 简介

LPP提出的最初目的是用于非线性流形的学习和分析。与PCA算法提取最具有代表性的特征来进行降维不同,LPP算法可以提取最具有判别性的特征来进行降维,因此在保留局部特征时具有明显的优势。下面简单介绍一下LPP算法的原理。假定有一个相似性矩阵 S 按下面的方法定义:

- (1) 当 x_i 是 x_j 的 K -近邻,或者 x_j 是 x_i 的 K -近邻时, $S_{ij} = e^{-\frac{x_i - x_j}{t}}$;
- (2) 否则, $S_{ij} = 0$ 。

这里的 t 是一个常量,LPP可以通过解下面的最小值问题来实现:

$$a_{opt} = \arg \min_a \sum_{i=1}^m (a^T x_i - x_i^T a)^2 S_{ij} = \arg \min_a a^T X L X^T a \tag{1}$$

同时满足条件: $a^T X D X^T a = 1$

这里的 $L = D - S$ 是拉普拉斯算子^[1], $D_{ij} = \sum_j S_{ij}$ 表示 x_i 附近的局部密度,LPP的基底函数即是矩阵 $(X D X^T)^{-1} X L X^T$ 对应于那些选取较小特征值的特征向量。因为矩阵 $(X D X^T)^{-1} X L X^T$ 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,将 $A_k = (a_1, \dots, a_k)$ 作为转换矩阵,则降维空间中两个数据点之间的欧几里得距离可以用下面的方法计算:

$$\text{dist}(y_i - y_j) = \|y_i - y_j\| = \|A^T x_i - A^T x_j\| = \sqrt{(x_i - x_j)^T A A^T (x_i - x_j)}$$

如果 A 是正交矩阵, $AA^T = I$, 则度量结构可以保持不变。

2 基于 LPP 的几种图像检索算法

2.1 基于 LPP 的半监督归纳算法

半监督学习近年来得到了广泛的关注。过去, 大量的该类研究集中在如何进行转导设置来预测输入观测数据的类别。现在, 很多应用领域都在尝试引入渐进式学习的方法, 如图像检索和文档分类领域。下面介绍一种结合 LPP 和支持向量机 (SVM) 的半监督学习算法。

用 $X = [x_1, \dots, x_m]$ 表示数据点的一个集合, 数据点共有 c 个类别。假设有 t 个数据点是已标注的, 剩下的 $(m - t)$ 个点是未标注的。用下面的方法构建一个权图: 对于任意两点, 如果彼此都是最近邻点, 则在该两点间作一条连线。对应的权按下面的方法定义:

- (1) 当 x_i 是 x_j 的 K -近邻或者 x_j 是 x_i 的 K -近邻时, $S_{ij} = e^{-\frac{\|x_i - x_j\|^2}{t}}$;
- (2) 否则, $S_{ij} = 0$ 。

注意到: 上面的定义反映出了数据空间的固有流形结构的特征。

再回过头来看 SVM 的机制, SVM 着眼于使二分类之间的空白最大化。SVM 的损害函数可以表示为

$$L(x_i, y_i, x) = \frac{2}{w}, y_i(x_i \times w + b) - 1 \geq 0, \forall_i$$

该损害函数的缺点是只考虑了标注后的数据点, 而忽略了未标注的数据点。一个自然的想法是应该结合未标注的数据点并保留权图的结构。另外, 我们希望如果两点之间有连线, 则它们应该属于同一类别。

假设 f 是分类器函数, 则 $f(x_i)$ 是 x_i 的估计标注。因此, 将下面的损害函数最小化:

$$G(\{x\}_i, f) = \sum_{ij} (f(x_i) - f(x_j))^2 S_{ij}$$

假定 f 是线性的, 可得到 $f(x) = w^T x$, 通过简单的代数变换可以得到

$$G(\{x\}_i, f) = w^T X L X^T w$$

其中, L 是拉普拉斯算子。

最佳的分类器可以由下面的方法得到

$$w^{opt} = \arg \min_w \|w\|^2 - \alpha w^T X L X^T w$$

这种算法的优点在于可以使二分类之间的空白最大化并保留数据的局部信息; 可以有效地使用未标注数据, 因而可以提高检索的准确度。然而, 这种算法的应用还有几个尚未确定的问题: 如在哪种情况下, 局部信息是特别重要的; 在 K -近邻搜索中, 如何确定参数 K , 这些问题都有待于进一步的研究。

2.2 基于 LPP 的迭代学习算法

LPP 算法的问题之一在于: 最近相邻图并不能确保很好地模仿流形的局部结构。下面介绍一种新型的流形学习算法——迭代保局投影 (Iterative Locality Preserving Projections, ILPP)。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, 然后根据这种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 投影

将图像 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, 同时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, 其中第 i 个节点对应图像矢量 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, 即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

用一个权矩阵 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

定义一个对角权矩阵 D , 其元素就是 S 的行 (或列, 因为 S 是对称矩阵) 元素的数值和。进一步定义 $L = D - W$, 该矩阵称为拉普拉斯矩阵^[1]。假设 $\{a_1, a_2, \dots, a_n\}$ (n 是当前图像空间的维数) 是下面这个特征问题的广义特征向量的集合:

$$X L X^T a = \lambda X D X^T a \tag{2}$$

假定 $W_{LPP} = [a_1, \dots, a_{n-1}]$, 可以将图像投影到一个 $n - 1$ 维的子空间中:

$$x \leftarrow W_{LPP}^T X$$

其中 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 (2) 更新最近邻图, 算法继续。

该算法的优点在于迭代地更新最近的相邻图, 从而使之能更好地模仿流形的局部结构, 因此用于图像检索时可以获得更好的效果。

3 实验分析

3.1 实验设计

下面设计了几个实验来评价上述算法用于图像检索时的性能。实验中使用的图像数据库是从 Corel 图库中选出的包括 79 个语义类的 10 000 个图像, 构成一个大型的多种类的图集。在系统中使用了三种颜色特征和三种纹理特征, 特征空间的维度是 435。我们设计了一种自动的反馈方案来模拟真实用户操作下的检索过程。在每一次循环中, 系统从前 100 个匹配结果中选出前三个不正确的结果作为不相关的样例, 同时选出最多三个正确结果作为相关的样例 (上一次循环选出的相关样例在这次排除在选择之外)。这些自动产生的反馈被加入到求检样例集中, 从而使检索结果最佳。为了评价算法的性能, 定义了检索准确度:

$$\text{检索准确度} = (\text{前 } N \text{ 个检索结果中相关图像的个数}) / N$$

3.2 使用半监督归纳算法的比较实验

图像检索本质上是一个分类问题, 给出一个待查询图像, 我们需要找到一个最佳的分类器, 它能区分相关和无关的图像。

在图 1 中, 显示了前 20 个返回结果的准确度, 可以看出,

半监督归纳算法的检索效果要优于标准 SVM 和转导 SVM。经过八轮反馈,在前 20 个返回结果中,标准 SVM 的准确度可以达到 51%,转导 SVM 的准确度可以达到 51%,而半监督归纳算法的准确度可以达到 63%。

3.3 使用 ILPP 迭代学习算法的比较实验

下面我们对采用 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 2 中可以看到前 20 个返回结果的准确度。分析实验结果,作一个大致的比较:在七次反馈后,对于前 20 个返回结果,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。

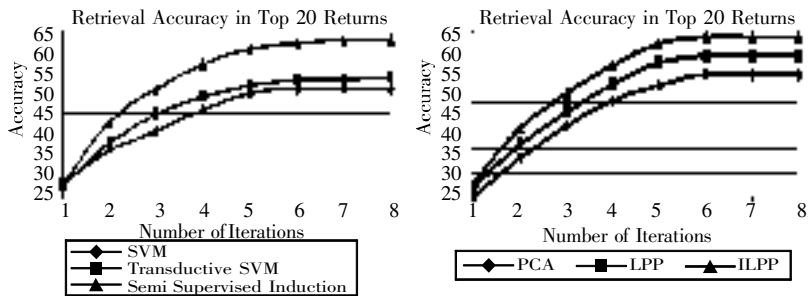


图 1 半监督归纳算法的检索准确度比较图

图 2 ILPP 算法的检索准确度比较图

4 总结

LPP 是近来国际上的一个研究热点,相对于广泛应用的 PCA 算法,LPP 在应用于图像降维时可以获得更好的效果。基

于 LPP 的半监督归纳算法是在 LPP 算法基础上,结合了 SVM 算法的机制而提出的一种新型算法。该算法的优点在于可以使二分类之间的空白最大化并保留数据的局部信息,因而可以提高检索的准确度。ILPP 算法是在 LPP 算法的基础上,采用迭代机制更新最近相邻图,从而使之能更好地模仿流形局部结构,因此用于图像检索时可以获得更好的效果。实验结果显示,基于 LPP 的半监督归纳算法的分类准确度明显优于 SVM 算法;ILPP 算法在图像检索中的效果明显优于 PCA 算法,比 LPP 的效果也要强一些。总之,基于 LPP 的半监督归纳算法和 ILPP 算法均是在 LPP 算法基础上进行的一些优化研究,目前这些研究取得了一些成果,提高了 LPP 算法的检索效率。

参考文献:

[1] Xiaofei He, Partha Niyogi. Locality Preserving Projections[C] . Vancouver: Advances in Neural Information Processing Systems, 2003. 291-299.

[2] Xiaofei He, Oliver King, Wei-Ying Ma, *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.

[3] Ke Lu, *et al.* Image Retrieval Using Dimensionality Reduction[C] . CIS, 2004. 775-781.

作者简介:

鲁珂(1974-), 男, 讲师, 博士研究生, 主要研究方向为网络多媒体; 赵继东(1976-), 男, 讲师, 博士研究生, 主要研究方向为计算机网络; 叶娅兰(1978-), 女, 博士研究生, 主要研究方向为网络体系结构; 曾家智(1939-), 男, 教授, 博导, 主要研究方向为计算机网络与通信。

(上接第 47 页)

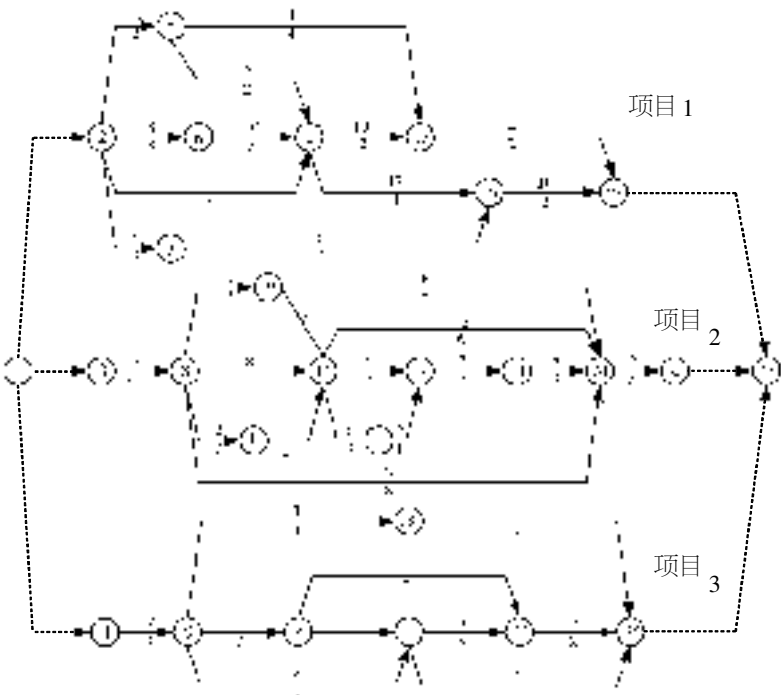


图 1 实例网络计划图

表 1 资源均衡优化结果

任务编号	开始时间	任务编号	开始时间	任务编号	开始时间	任务编号	开始时间
2-5	0	8-10	4	12-15	8	18-25	13
2-6	0	8-13	6	12-20	8	19-22	13
2-12	0	8-11	4	13-24	12	19-25	16
2-7	0	8-24	6	13-16	9	20-23	12
3-8	1	9-18	8	13-17	9	21-24	15
4-9	4	9-14	7	14-22	15	22-25	18
5-15	2	9-19	7	14-19	10	24-26	18
5-12	3	10-24	13	15-23	11		
6-12	3	10-13	7	16-17	12		
7-20	3	11-13	7	17-21	13		

4 结束语

单项目资源均衡优化虽然能改善单个项目的资源分布状况,但并不能实现整个企业内的资源均衡,因而本文提出了多项目资源均衡优化的概念,并建立了多项目资源均衡优化问题的模型。针对多项目资源均衡问题,给出了一种遗传算法的求解方法。在算法中,按照网络图的拓扑排列顺序,将任务开始时间编码为染色体,满足了任务间的逻辑约束关系,能减少遗传操作过程中非法个体的修复计算量,加快算法的收敛速度。多项目资源均衡优化可以有效地实现企业内的资源均衡配置,在工程中有一定的应用推广价值。

参考文献:

[1] 罗时鹏,王乘,杨彪.基于浮动工时率的资源均衡问题研究[J] .华中科技大学学报,2003,31(3) : 78-80.

[2] Hiyassat M A S. Applying Modified Minimum Moment Method to Multiple Resource Leveling[J] . Journal of Construction Engineering and Management, 2001, 127(3) : 192-198.

[3] 毕玉峰,王选仓,孙立军.公路工程中多种资源均衡优化的遗传算法设计[J] .中国公路学报,2003,17(1) : 86-89.

[4] 张连营,骆刚,鹿丽宁.遗传算法在工程项目资源优化中的应用[J] .天津大学学报,2001,34(2) : 188-192.

[5] 王祖和,亓霞.多资源均衡的权重优选法[J] .管理工程学报,2002,16(3) : 91-93.

[6] Rudolph G. Convergence Analysis of Canonical Genetic Algorithms [J] . IEEE Trans. on Neural Networks, 1994, 5(1) : 96-101.

作者简介:

王为新(1980-), 男, 山东郓城人, 硕士研究生, 主要研究方向为计算机辅助设计、项目网络计划优化等; 李原(1964-), 女, 陕西西安人, 教授, 博士, 主要研究方向为计算机辅助技术、并行工程与虚拟制造技术等。