

# 一种用于图像检索的新型半监督学习算法

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 成都 610054)

**【摘要】**基于支持向量机的理论提出了一种用于图像检索的半监督学习算法。该算法的基本思想是, 如果两点彼此是最近点, 则它们共用一个标注。因此, 该算法可以在具有最大类间空隙和很好保留位置特征的基础上找到一个投影。对该算法和标准支持向量机及转导(transductive)支持向量机的图像检索效果进行了实验比较, 结果表明该算法可以获得更好的效果。

**关键词** 支持向量机; 半监督学习; 保局投影; 图像检索

中图分类号 TP391.3M 文献标识码 A

## Algorithm for Semi-Supervised Learning in Image Retrieval

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

(School of Computer Science and Engineering, UEST of China Chengdu 610054)

**Abstract** In this paper, based on Support Vector Machine (SVM), we introduce a semi-supervised learning algorithm for image retrieval. The basic consideration of the algorithm is that, if two data points are close to each other, they should share the same label. Therefore, it is reasonable to search a projection with maximal margin and locality preserving property. Comparing our algorithm to standard SVM and transductive SVM, Experimental results show efficiency and effectiveness of our algorithm.

**Key words** SVM; semi-supervised learning; LPP; image retrieval

随着数字图像容量的快速增长, 对于图像管理工具的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)的效率成了当前的研究热点<sup>[1]</sup>。

有一些研究试图将半监督学习用于图像检索领域, 因此近几年对半监督学习的关注越来越多。与传统的监督学习只使用标注后的数据不同, 半监督学习同时使用标注后的数据和未标注的数据。通常, 未标注数据能够更好地表示数据空间的固有几何结构, 可以有效地提高分类效果。本文在对支持向量机(Support Vector Machine, SVM)和保局投影(Locality Preserving Projections, LPP)进行研究的基础上<sup>[2-4]</sup>, 提出了一种新型的用于图像检索的半监督学习算法。通过SVM和LPP方法的结合, 可以得到一个分类器, 能够使距离差值最大化并很好地保留数据的局部特征。另外, 本文提出的算法可以很好地辨识数据空间的流形结构(欧几里得空间可以看作是一种特殊的流形)。由于上述特点, 新型半监督学习算法可以在图像检索中获得很好的应用效果。

## 1 SVM及LPP简介

### 1.2 SVM简介

由Vapnik等人首先提出的SVM是一种基于统计学习理论的模式识别方法, 现已在许多领域(如生物信息

学,文本和手写识别等)成功应用。SVM训练算法的基于结构风险最小而不是经验风险最小的思想,导致了在训练多项式、神经网络、径向基函数等领域的许多新方法。SVM没有设定数据的表示方法,因此可以应用于那些尚没有足够的知识来说明输入数据的场合<sup>[2]</sup>。

以下通过一个二分类问题来说明SVM的机制。假定一个由标注后样本组成的数据集  $D = \{x_i, y_i\}_{i=1}^l$ , 其中  $y_i \in \{-1, 1\}$ 。希望在众多的二分该数据集的线性分类面中, 找到一个可以使归纳错误最小化并具有最小上界的分类面。可以发现, 具有上述特性的分类面实际上是使得二类数据间空白最大的一个超平面。如果对一个新的数据点  $x$  进行分类, 将会根据该点与判别边界的关系来分配标注, 对应的判别函数如下:

$$f(x) = \text{sign}\left(\sum_{i=1}^l a_i y_i \langle x, x_i \rangle - b\right)$$

## 1.2 LPP简介

LPP是最近提出的用于线性空间降维的一种算法。对于原始空间中的一个数据点  $x$ , 假设有一个投影  $w \in R^n$  使得  $y = w^T x$ 。LPP的目的是发现固有的, 以欧几里得方式内嵌的流形结构。假设有一个位置相似矩阵  $S$ , 则可以通过求解下面的最小化问题得到最佳的投影为:

$$w_{\text{opt}} = \arg \min_w \sum_{i=1}^m (w^T x_i - w^T x_j)^2 S_{ij} = \arg \min_w w^T X L X^T w$$

式中 附加约束条件为  $w^T X D X^T w = 1$ ,  $L = D - S$  是拉普拉斯算子。  $D_{ij} = \sum S_{ij}$ 。  $D_{ij}$  代表  $x_i$  周围的位置密度,  $D_{ij}$  越大, 就表示点  $x_i$  越重要。可以通过简单的代数变换将上面的问题转化为下面的特征向量问题, 即:

$$X L X^T w = \lambda X D X^T w$$

## 2 半监督归纳算法

近年来受广泛的关注。以往, 大量的半监督学习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, 在很多应用领域(例如图像检索和文档分类领域)都在尝试引入渐进式学习方法。但是大多数现有的半监督学习方法对于新的测试数据不能很容易地进行标注, 而且计算具有相当的复杂性。本节提出了一种有效结合SVM和LPP的半监督学习算法。

用矩阵  $X = [x_1, x_2, \dots, x_m]$  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, 对于任意的两点, 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,  $S_{ij} = e^{-\|x_i - x_j\|^2}$ ;
- 2) 否则,  $S_{ij} = 0$ 。

以上的定义反映了数据空间的固有流形结构的特征。

再来看SVM的的机制, SVM着眼于使二分类之间的空白最大化。SVM的损害函数可以  $G$  表如下:

$$G(x_i, y_i, x) = \frac{2}{w}$$

式中  $y_i(x_i w + b) - 1 = 0, \forall i$ 。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, 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, 希望将下面的损害函数最小化:

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

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:

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

式中  $L$  是拉普拉斯算子。最佳分类器可以由下面的方法得到:

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

也可以将上述算法用核方法扩展到非线性的情况<sup>[2]</sup>。考虑再生核希尔伯特空间(RKHS)中的分类器, 假定核函数为  $K$ , 再生核希尔伯特空间  $H_K$  可以定义为:

$$H_k = \left\{ f \mid f(x) = \sum_{i=1}^m a_i K(x_i, x), a_i \in R \right\}$$

解决上述 $H_k$ 中的最优化问题,可以得到一个非线性分类器。

### 3 实验分析

#### 3.1 实验设计

以下设计了几个实验评价新型半监督学习算法用于图像检索的性能。实验中使用的图像数据库是从Corel图库选出的包括79个语义类的10 000个图像,构成了1个大型的多种类图集。在系统中使用了3种颜色特征和3种纹理特征,特征空间的维度是435,以一种自动反馈方案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3个不正确的结果作为不相关的样例,同时选出最多3个正确结果作为相关的样例,且上一次循环选出的相关样例应排除在选择之外。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,定义检索准确度为:

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

#### 3.2 使用半监督归纳算法的图像检索实验

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

图1中,可以看到前20个返回结果的准确度;图2中,可以看到前50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8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%。对于前50个返回结果的情况,半监督归纳算法的准确度也比标准SVM高8%,比转导SVM高6%。

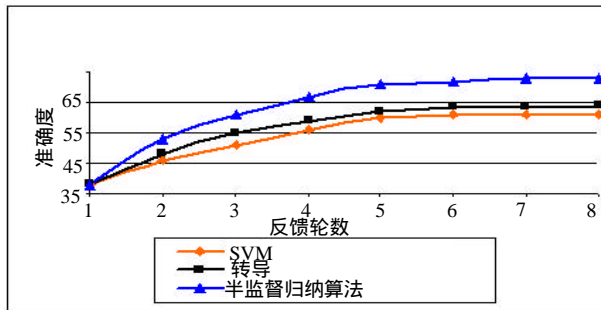


图1 前20个返回结果的检索准确度

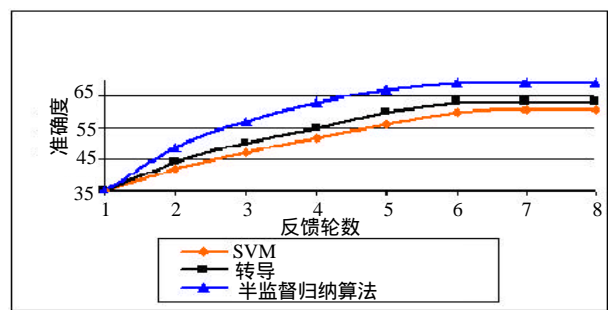


图2 前50个返回结果的检索准确度

### 4 总结

本文提出了一种有效结合SVM和LPP的新型半监督学习算法,该算法可以使两分类之间的空白最大化并保留数据的局部信息。该算法原来主要用于流形结构的学习,本文将其推广应用于图像检索。实验结果显示,该算法可以有效地使用未标注数据,因而可以提高检索的准确度。

#### 参考文献

- [1] He Xiaofei, Oliver King, Ma Wenying, 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
- [2] Vapnik V. The nature of statistical learning theory[M]. New York: Springer Verlag, 1995
- [3] Thorsten J. Transductive inference for text classification using support vector machines[C]. 16th International Conference on Machine Learning, San Francisco, USA, 1999
- [4] He Xiaofei, Partha Niyogi. Locality preserving projections[C]. Advances in Neural Information Processing Systems, Vancouver, Canada, 2003