

基于半监督线性近邻传递的相关反馈方法

黄传波, 金 忠

(南京理工大学计算机科学与技术学院, 江苏 南京 210094)

摘 要: 提出了一种半监督线性近邻传递的相关反馈方法 FSLNP (feedback semi-supervised linear neighborhood propagation). 该算法不仅能够保持正、负例约束信息, 而且能够保持图的局部以及全局相关性结构信息. 采用相关反馈的有标签和未知标签图像点, 找到比较好的表示图像相关性的一个图结构, 来揭示图像点的语义间结构关系. 实验结果表明: 该算法可以提高检索的准确度, 而且在经过长期学习后可以获得一个优化相关性的图结构.

关键词: 相关反馈; 半监督学习; 图像检索; 线性近邻传递

中图分类号: TP118

文献标识码: A

文章编号: 1002-0411(2011)-03-0289-07

Relevance Feedback Algorithm Based on Semi-supervised Linear Neighborhood Propagation

HUANG Chuanbo, JIN Zhong

(College of Computer Science & Technology, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: A feedback semi-supervised linear neighborhood propagation method (FSLNP) is proposed. FSLNP method can not only preserve the positive and negative constraints but also preserve the local and global relevance structure information of the whole graph. With both labeled and unlabeled images in relevance feedbacks, a better structure for relevance representation among images is found to reveal the semantic structure. Experimental results show that FSLNP can effectively improve retrieval accuracy, and after long term learning, an optimal relevance graph space can be obtained.

Keywords: relevance feedback; semi-supervised learning; image retrieval; linear neighborhood propagation

1 引言 (Introduction)

基于内容的图像检索 (CBIR) 的目的就是根据图像的视觉内容从大型图像库中检索出用户需要的图像^[1]. 用户的需求可以归结为高层语义方面, 而图像的视觉内容通常用低层特征表示, 如: 颜色、纹理和形状等. 相关反馈 (RF) 技术试图通过用户的反馈信息在高层语义概念和低层视觉特征之间建立联系, 此领域的研究主要是为了弥补目前图像检索系统中底层视觉特征表示和图像高层语义之间的鸿沟. 从 20 世纪 90 年代将 CBIR 引入到相关反馈技术中, 发展到现在已出现了许多方法, 如启发式方法、优化学习、聚类、机器学习等, 这些方法使 RF 技术得到了快速的发展^[2].

在很多实际应用中, 随着数据采集技术和存储技术的发展, 获取大量的无标签样本已变得非常容易, 而获取有标签样本通常需要付出很大的代价, 因此, 半监督学习已成为近年来机器学习领域的一

个研究热点. 半监督学习^[3]是一种同时利用有标签和无标签数据的方法, 其基本思想是利用一些分布上的假设, 对未标记数据的标记进行预测, 然后合并到标记的数据中去. 目前, 半监督学习研究涵盖范围非常广泛, 一方面不断修改或者扩展各种传统有监督、无监督学习算法以提高它们在半监督情况下学习的能力; 另一方面, 不断有新的数学方法引入到半监督学习中来. 同时, 半监督学习的研究对象已经由单纯的半监督数据的训练扩展到半监督数据的流形分析、半监督数据和图模型的关系等. 出现了很多半监督学习方法, 如: 半监督 EM 算法^[4-5]、半监督归纳算法^[6]、协同训练 (co-training) 算法^[7]、Tri-Training 算法^[8]、直推式支持向量机^[9-10]以及支持向量机的主动学习方法^[11]等. EM 是获取大量未标记的图像, 利用这些未标记的图像构造一个用于判别分类的生成模型, 然后再用 EM 方法来估计该模型的最大后验概率参数. Wu 等人^[4]首

基金项目: 国家自然科学基金资助项目 (60873151, 60973098, 90820306).

通讯作者: 黄传波, huangjunfengcq@126.com 收稿/录用/修回: 2010-02-08/2010-04-30/2010-08-09

先将 EM 算法用于图像分类中, 根据用户的反馈信息并利用 EM 算法来估计目标图像的分布. Sia 等人^[5]利用 EM 算法估计用户查询图像每一维向量的空间分布. 鲁珂等人^[6]提出的半监督归纳图像检索反馈算法, 是通过 SVM 和 LPP (locality preserving projection) 方法相结合得到一个分类器, 能够使距离差值最大化并很好地保留数据的局部特征, 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. 协同训练算法^[7], 此类算法隐含地利用聚类假设或流形假设, 它们使用两个或多个学习器, 在学习过程中, 这些学习器挑选若干个置信度高的未标记实例进行相互标记, 从而使模型得以更新. Tri-Training 算法^[8], 使用三个分类器, 不仅可以简便处理标记置信度估计问题以及对未标记实例的预测问题, 还可以利用集成学习来提高泛化能力. 该算法首先对有标记实例集进行可重复取样以获得三个有标记训练集, 然后从每个训练集产生一个分类器. 在对未标记实例进行预测时, Tri-Training 算法不再像以往算法那样挑选一个分类器来使用, 而是使用集成学习中经常用到的投票法集成三个分类器实现对未标记实例的预测. 直推式支持向量机算法^[9], 在训练过程中, 该算法不断修改 SVM 的划分超平面并交换超平面两侧某些未标记实例的可能标记, 使得 SVM 在所有训练数据 (包括有标记和未标记实例) 上最大化间隔, 从而得到一个既通过数据相对稀疏的区域又尽可能正确划分有标记实例的超平面. 支持向量机的主动学习方法^[11], 是依据越靠近 SVM 分类面的数据点, 其被正确分类的可能性越小这一特点, 选择那些靠近分类面的无标签数据点提供给用户标注以实现两个类之间超平面的最大改进. 在 CBIR 过程中, 即使将用户在相关反馈过程中提供的信息考虑进来, 有标签图像数目仍然是比较少的, 图像库中通常存在大量的未标记图像, 在查询之前无法判断它们是否与查询相关. 显然, CBIR 任务是典型的有标签实例很少、未标记实例非常多的任务. 因此, 基于内容的图像检索是利用未标记实例学习技术的很好试验场, 另一方面, 通过引入这些学习技术可能有助于突破 CBIR 的技术瓶颈^[12].

通过以上分析可以看出, 在图像检索实际应用中, 对基于半监督学习相关反馈算法的研究是十分必要的. 最近, 提出了一种新的标签传递算法——LNP (linear neighborhood propagation)^[13], 建立在统计学习理论基础上的 LNP 方法, 在有限的样

本下有良好推广能力. 基于此, 本文通过距离测度的变换引入先验信息形成一种新的半监督学习算法, 即半监督线性近邻传递算法 SLNP (semi-supervised linear neighborhood propagation), 并将其用于 CBIR 的相关反馈过程中, 形成基于半监督线性近邻传递的相关反馈 FSLNP. 在图像检索实验中, 获得了较好效果.

2 LNP 简介 (LNP introduction)

LNP 算法^[13]首先是用一系列重叠的线性近邻块来逼近整个图, 在每个块中边权值是用标准二次规划过程来求解的, 然后合并所有边权值形成整个图的权重矩阵.

2.1 近邻图的构造

假设 $X = \{x_1, x_2, \dots, x_l, x_{l+1}, \dots, x_n\}$ 表示 R^d 空间的 n 个图像点的集合, $Y = \{1, -1\}$ 是标签集合, 前 l 个点 $x_i \in X$ ($i \leq l$) 是有标签点, 标签表示为 y_i , $y_i \in Y$, 其余点 $x_u \in X$ ($l+1 \leq u \leq n$) 是无标签的点.

LNP 不像传统基于图的方法那样考虑成对点之间的关系, 而是利用每个点的近邻信息来构建图 G . 借助 LLE (locally linear embedding) 方法, 假定每个图像点能通过它邻近点的线性组合来优化重构. 因此, 可以将其转化为极小化式

$$\min \sum_{i=1}^n \left\| x_i - \sum_{j: x_j \in N_k(x_i)} w_{ij} x_j \right\|^2 \quad (1)$$

其中, $N_k(x_i)$ 表示 x_i 的 k 个近邻点的集合, x_{i_j} 是 x_i 的第 j 个近邻点, w_{i_j} 是 x_{i_j} 对 x_i 的贡献. 添加两个限制条件, $\sum_{x_j \in N_k(x_i)} w_{i_j} = 1$, $w_{i_j} \geq 0$, 显然, x_{i_j} 与 x_i 越相似, w_{i_j} 值就越大. 因此, w_{i_j} 可以用来度量 x_{i_j} 与 x_i 的相似性.

在计算出所有重构权重后, 就可以构造一个权重矩阵 W .

$$W = (w_{ij})_{n \times n} \quad (2)$$

2.2 标签传播

假定 F 表示定义在 X 上的分类函数集, $\forall f \in F$ 能对每个 x_i 指定 $f_i = f(x_i)$ 的实际值, 无标签图像点 x_u 的标签可以通过 $f_u = f(x_u)$ 决定.

在每次迭代中, 让每个图像点从它的近邻吸收标签信息的权值并保留它先前状态的一些标签信息^[14], 那么在第 $t+1$ 次迭代后 x_i 的标签为

$$f_i^{t+1} = \alpha \sum_{j: x_j \in N_k(x_i)} w_{ij} f_j^t + (1-\alpha) y_i \quad (3)$$

其中, $0 < \alpha < 1$ 是 x_i 从它的近邻获得的标签信息的权值. $\mathbf{y} = (y_1, y_2, \dots, y_n)^T, y_i \in Y, (i \leq l), y_u = 0 (l+1 \leq u \leq n), \mathbf{f}^t = (f_1^t, f_2^t, \dots, f_n^t)^T$ 是在 t 次迭代后的预测标签向量, 令 $\mathbf{f}^0 = \mathbf{y}$, 这样可以重新将迭代方程写为

$$\mathbf{f}^{t+1} = \alpha \mathbf{W} \mathbf{f}^t + (1 - \alpha) \mathbf{y} \quad (4)$$

可以用此式更新每个图像点的标签直到稳定状态.

3 SLNP 的相关反馈 (Relevance feedback based on semi-supervised linear neighbourhood propagation)

3.1 SLNP 算法

LNP 算法是借助 LLE 方法, 假定每个图像点通过它邻近点的线性组合优化重建来构建 G . 在此, 计算距离时, 可以将已知标签的正例点、负例点以及未知标签点这种先验知识加入其中, 在不改变局部拓扑结构的前提下, 尽量增大源数据中的有效信息量. 据以上分析, 提出 SLNP 半监督局部线性近邻扩展算法, 其思想是将监督学习引入 LNP 算法, 应用图像点的已知标签信息调整图像点之间的距离, 使用调整后的距离来实现线性近邻重构.

针对 LLE 未能利用样本的类别信息的缺点, Dick 等提出一种有监督的 LLE 算法^[15]. 传统的 LLE 算法在第一步时是根据样本点间的欧氏距离来寻找 k 个近邻点, 而有监督 LLE 在处理这一步时, 增加了样本点的类别信息. 因此, 为增大原数据的有效信息量, 可以借鉴近两年提出的监督式局部线

性嵌入算法, 构建 k 近邻的距离表示方法^[16]:

$$D(x_i, x_j) = \begin{cases} \sqrt{1 - e^{-\frac{d^2(x_i, x_j)}{\beta}}}, & x_i \text{ 与 } x_j \text{ 有相同的标签} \\ \sqrt{e^{\frac{d^2(x_i, x_j)}{\beta}}}, & x_i \text{ 与 } x_j \text{ 有不同的标签} \end{cases} \quad (5)$$

其中, $d(x_i, x_j)$ 为 x_i 与 x_j 之间的欧氏距离, 参数 β 用于防止在 $d(x_i, x_j)$ 较大时, $D(x_i, x_j)$ 增长过快. 当 β 为 4 时, 式 (5) 的函数曲线如图 1 所示.

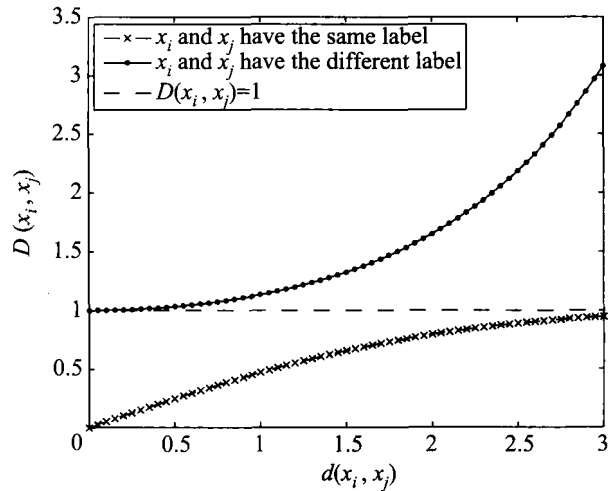


图 1 式 (5) 中 $D(x_i, x_j)$ 随 $d(x_i, x_j)$ 的变化
Fig.1 Change of $D(x_i, x_j)$ with $d(x_i, x_j)$ in formula (5)

基于这种想法, 引入先验信息进行半监督学习, 进一步改进距离求解过程. 当未知两图像点 x_i 与 x_j 是否同类, 但 x_i 与 x_j 互为 k 近邻时, 它们之间的距离大于它们同类时的距离, 而小于它们之间关系为其它情况时的距离, 这样距离调整公式如下:

$$D(x_i, x_j) = \begin{cases} \sqrt{1 - e^{-\frac{d^2(x_i, x_j)}{\beta}}}, & x_i \text{ 与 } x_j \text{ 有相同的标签} \\ \sqrt{1 - e^{-\left(\frac{d(x_i, x_j)}{\beta}\right)^2}} + \sqrt{e^{\left(\frac{d(x_i, x_j)}{\beta}\right)^2}} / 2, & x_i \text{ 与 } x_j \text{ 标签未知, 且 } x_i \in N_k(x_j) \text{ 或 } x_j \in N_k(x_i) \\ \sqrt{e^{\frac{d^2(x_i, x_j)}{\beta}}}, & \text{其它} \end{cases} \quad (6)$$

式中, $N_k(x_i)$ 表示 x_i 的 k 近邻.

当 β 为 13 时, 式 (6) 的函数曲线如图 2 所示. 从图 2 可以看出, 当 $d(x_i, x_j)$ 相同时, 具有相同标签的类内距离小于 1, 不同类间距离大于 1, 未知是否为同一类的近邻图像点之间的距离位于两者之间, 该方法保证了类间距离比类内距离大. 由以上分析, 可以用 $D(x_i, x_j)$ 距离构建 LNP 算法中图的权重矩

阵.

SLNP 算法的具体步骤如下:

Step1 利用欧氏距离 $d(x_i, x_j)$ 计算每个图像点的 k 近邻点 $N_k(x_i)$.

Step2 结合样本的已知标签信息和近邻信息, 利用式 (6), 将图像点间的距离计算调整为 $D(x_i, x_j)$.

Step3 运用 $D(x_i, x_j)$ 重新计算每个点的 k 近

邻, 用式 (1) 通过 x_i 的近邻点计算 x_i 的线性重构权重 w_{ij} , 由式 (2) 得到一个权重矩阵 \mathbf{W} .

Step4 根据式 (4), 使任意图像点从其相邻图像点获取标签信息.

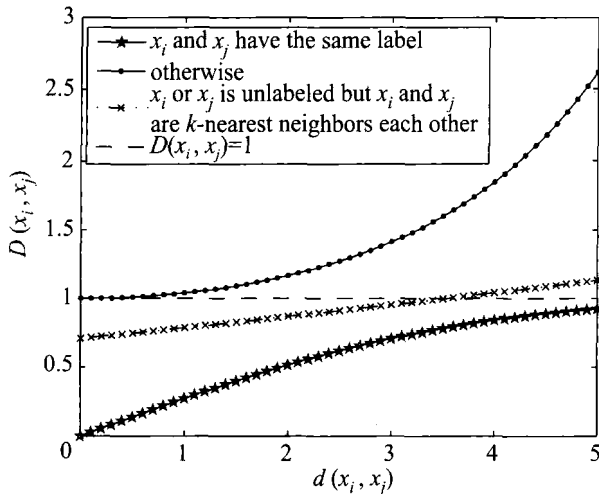


图2 式 (6) 中 $D(x_i, x_j)$ 随 $d(x_i, x_j)$ 的变化

Fig.2 Change of $D(x_i, x_j)$ with $d(x_i, x_j)$ in formula (6)

3.2 基于 SLNP 的相关反馈 FSLNP

FSLNP 是把用户每次检索结果标注的标签信息作为下次标签扩展的已知标签集, 用它们来对图的结构重新作一些调整, 使各图像点之间的相似关系可以更接近用户的语义理解, 从而使得检索结果更加准确.

3.2.1 FSLNP 模型建立

FSLNP 就是将 SLNP 算法应用于图像检索的相关反馈机制. FSLNP 的模型如下:

假设 q 表示查询输入图像, 在 h 次反馈后用户提供的正例图像集为 $A = \{x_1, x_2, \dots, x_m\}$, 反例图像集为 $B = \{x_{m+1}, x_{m+2}, \dots, x_l\}$. l 个已知标签的图像构成有标签图像集, 其余 $n-l$ 个为未知标签的图像集. 此时的目标就是根据 A 和 B 预测出未知标签图像的关系类别标注. 为了在学习过程中有效地结合上述的未知标签图像和有标签图像的信息, 定义了一个图 $G = (V, E)$, 在这个图中顶点集合 V 代表图像库中各个图像点, 任意一个图像点 x_i 和它的近邻图像点 x_j 相连的边 E 为其相似性, 以权重值 w_{ij} 来量化, 用式 (1) 计算. 任意一个图像点通过 w_{ij} 从它近邻图像点获取标签信息, 依次循环地传递到其它图像点, 最终达到全局稳定状态, 图像点之间边的权重越大, 图像点越相似, 它们拥有相同标签的可能性就越大 (全局一致性的假设), 从而导出未知标签图像点的标注信息.

于是, 反馈机制就可以看作是在这个图上估计一个标注函数 f . 它充分结合了未知标签图像点和

有标签图像点的信息, 为了实现全局一致性假设, 未知标签图像的标注问题最终被形式化为: 相关图像点类别的标签信息, 根据其近邻图像点的相似度在图上进行全局传递.

3.2.2 FSLNP 描述

FSLNP 的具体步骤如下:

Step1 使用欧氏距离度量相似性的方法进行检索, 在初始检索结果中选择 N 幅图像提供给用户, 用户标注与检索目标相关的图像, 得到用于反馈的正例集合 I_p^0 和反例集合 I_n^0 . 更新集合 I_p 和 I_n : $I_p = I_p \cup I_p^0$, $I_n = I_n \cup I_n^0$, 初始时正例和反例集合 I_p 和 I_n 都为零.

Step2 将 I_p 和 I_n 作为已知标签信息的图像点 (x_i, y_i) , $x_i \in I_p \cup I_n$, $y_i = \begin{cases} +1, & x_i \in I_p \\ -1, & x_i \in I_n \end{cases}$, 运用 SLNP 算法进行标签传递.

Step3 根据每个图像的标签分值, 按照分值由大到小将图像库中的所有图像进行排序, 取最前 N 幅图像反馈给用户.

Step4 用户对反馈的图像重新进行标注. 得到用于反馈的正例集合 I_p^1 和反例集合 I_n^1 , 用此再进行集合 I_p 和 I_n 的更新, $I_p = I_p \cup I_p^1$, $I_n = I_n \cup I_n^1$.

Step5 转到 **Step2**, 依据新的有标签信息重新进行标签扩展, 并继续进行检索直到找到足够数量的相似图像或者相似图像数目不再增加为止.

3.2.3 FSLNP 分析

FSLNP 在每一次循环传递标签的迭代中, 每一个图像点都接收了两方面的信息: 其一是来自相邻节点的信息, 其二是自身的初始标签信息. 然后, 再通过参数 α 来指定两者信息的权重, 通过该方式将标签信息扩散开来. 有标签图像点的信息在每一次标签的全图传递后都将被填回, 因此, 这些标签信息就像是固定资源一样被循环推开, 直到整个系统达到一个稳定状态. 所有已知标签图像点的初始标签信息在整个算法的标签传递中得到了最充分的利用.

FSLNP 利用用户反馈迅速优化图像点之间的相似关系, 当前反馈处理将对未来查询产生作用, 它一方面直接从大量的标注和未标注图像中用 LLE 学习的方法从近邻节点提取标签信息; 另一方面又利用用户提供的相关性标注来改进图结构, 因而可以得到更加准确的检索效果.

但是, 本方法的不足在于那些有标签图像点的初始标签信息的作用在扩散过程中因为参数 α 的限

制而被淡化. 此外, 如何自动指定恰当的参数 α 值也是一个难点.

4 实验与分析 (Experiment and analysis)

通过对比实验, 说明本文所提出的 FSLNP 算法在图像检索中的有效性.

4.1 图像特征

图像低层视觉特征抽取是基于内容的图像检索中的关键问题, 一般视觉特征包括颜色、纹理和形状等. 事实上, 如果低层视觉特征足够准确, 低层特征空间中的欧氏距离能准确反映图像间的语义关系, 那么在低层特征空间能很简单实现最近邻检索且能够保证检索的实现. 但是, 在低层视觉特征和高层语义之间不易提取明显的强相关性, 为了发现数据中语义结构不得不凭借用户的交互作用. 为验证本文所提出反馈算法的优越性, 用表 1 所示的 274 维图像低层特征进行检索实验, 它包含了颜色、纹理和形状信息.

为消除量纲等问题的影响, 在数据分析前对数据进行预处理是必不可少的. 因此, 本文采用此方法进行特征预处理. 具体如下:

$$x_{ij}^* = dx_{ij} - \min_{1 \leq i \leq h} x_{ij} R_j \quad (7)$$

其中, x_{ij} ($i = 1, 2, \dots, h; j = 1, 2, \dots, n$), h 为特征的维数, n 为图像库中图像数量. $R_j = \max_{1 \leq i \leq n} x_{ij} - \min_{1 \leq i \leq n} x_{ij}$ 表示极差.

表 1 实验中所用的图像特征

Tab.1 Image feature used in the experiment

特征名称	维数
颜色直方图 ^[17]	256
颜色矩 ^[18]	9
边缘直方图 ^[17]	5
边缘方向直方图 ^[19]	4
总计	274

4.2 FSLNP 的参数确定

参数的选择是大多数学习问题中的关键环节, 在一些情况下, 学习性能会随着参数的不同选择而快速变化, 因此, 必须合理地选择参数值, 以取得较好的算法性能.

在 SLNP 算法中, 有 3 个参数: β 、 α 和 k . β 用于调整改进距离 $D(x_i, x_j)$ 随对应欧氏距离 $d(x_i, x_j)$ 增长的速度. α 是图像点从它的 k 个近邻点获得标签信息的权值. k 是取的近邻点数. 根据文 [13,16], α 取 0.99, β 取图像库中所有图像点间欧氏距离的平

均值, 在此情况下, FSLNP 对于不同 k 值第一次反馈后的结果 (precision@top 25 和 precision@top 36) 如图 3 所示.

由此可知, 由于算法中参数 k 具有刻画图像数据空间局部几何和鉴别结构信息的作用, 随着 k 值的增大, 检索精度呈下降趋势, k 取为 4 时精度最好.

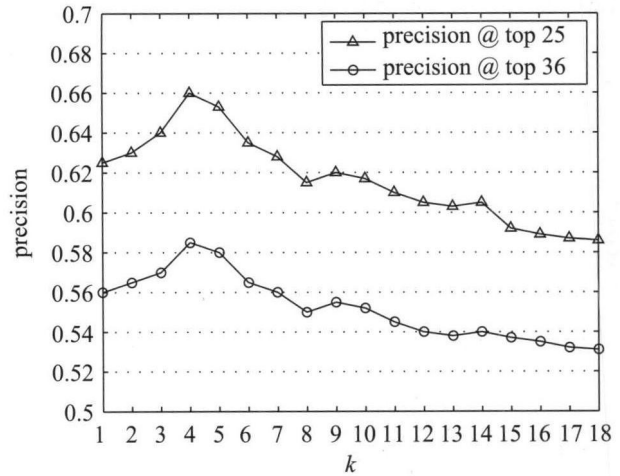


图 3 FSLNP 中检索精度随参数 k 值的变化情况
Fig.3 Retrieval precision versus different values of the parameter k in FSLNP

4.3 算法性能分析

为了说明 FSLNP 的有效性, 本文将 FSLNP 与 LNP 算法^[13] 和支持向量机的主动学习算法^[11] 进行比较实验. Corel 图像库广泛使用在基于内容的图像检索系统, 如文 [20-21]. 为了便于评价, 选择这个库中的图像用于测试. 选择 30 个语义类的彩色图像, 包括大象、雪山、鸟、海浪、汽车、花等, 每个语义类由 100 幅图像组成. 采用 precision-scope curve 和 precision@top N 作为评判指标.

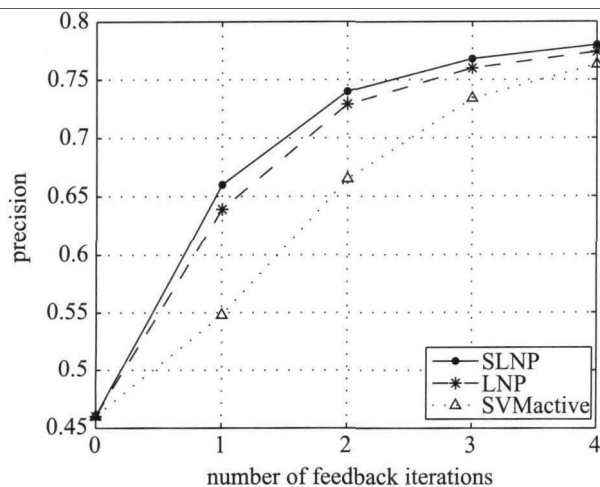
实验 I: 为分析对比三种算法对不同反馈次数的性能变化, 采用 precision@top 25 和 precision@top 36 评价方法对其性能进行比较实验, 其结果如图 4 所示.

图中开始点为没有反馈信息的原始检索结果, 可以看出, 用户进行前两轮相关反馈后, 其检索精度有大幅度提高, 此后, 检索精度的提升则渐趋平缓. 本文提出的算法在总体范围上要优于其它两种方法, 三种方法都明显优于无反馈算法, 这显示出用户提供的相关反馈信息对于提高检索性能是非常有益的, 随着反馈次数增加能够发掘出更多的相关图像, 因而反馈机制可以明显地提高查准率, 上述实验结果也证明了这一点.

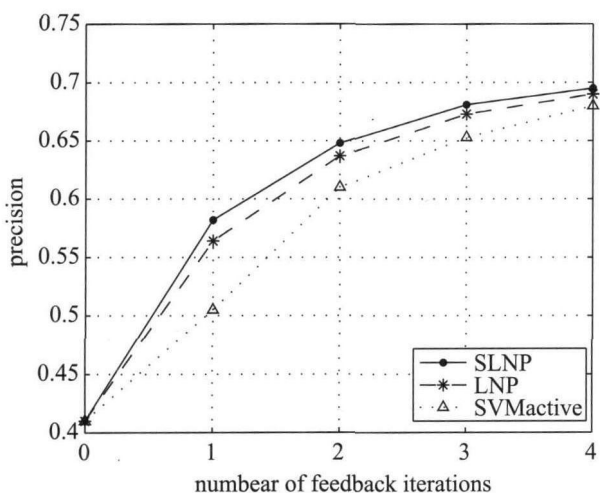
实验 II: 为直观显示本文算法效果, 将无反馈

时和用本文方法进行一次反馈后的前 25 幅图显示出来,如图 5 所示.

图 5(a) 中,第 1 行第 5 幅,第 2 行第 3 幅,第 3 行第 3 幅,第 4 行第 3 幅,第 5 行第 1~5 幅是负例检索图像,其查准率为 64%;图 5(b) 中,只有第 5 行的第 4、5 幅是负例检索图像,其查准率为 92%. 显然,经过第一次反馈其查准率得到了大幅提高,说明提出的 FSLNP 算法性能是比较好的.



(a)



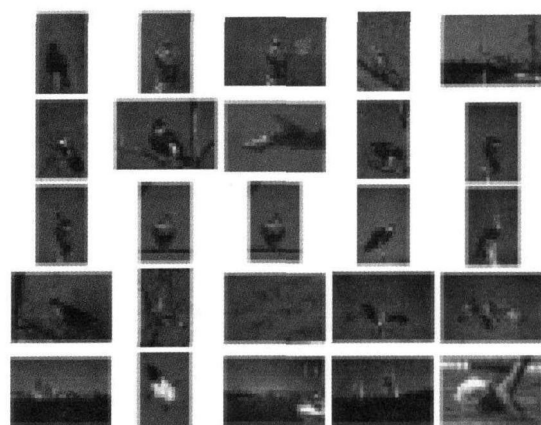
(b)

图 4 取前 25 幅 36 幅图时三种算法的检索性能

Fig.4 Performance of the three algorithms for image retrieval in precision@top 25 and precision@top 36

从以上实验结果可以看出, FSLNP 算法取得了比 LNP 算法及支持向量机的主动学习算法更好的性能. 其主要原因是: 这种方法可以充分利用图的结构来平滑未知标签图像点的标签信息, 因而, 未知标签图像点的标签信息不仅由附近的有标签图像点决定, 而且还由附近的未知标签图像点决定. 而对于支持向量机的主动学习算法, 因为有标签图像点太少而不足以揭示关系类别的分类边界, 因而无

法相对准确地学习出分类超平面, 从而导致运行效果不佳. 本文算法在构建图时, 基于其结构充分利用先验知识, 通过距离变换增加了有效信息量. 所以, 本文提出的 FSLNP 方法在图像检索中能获得更好的结果.



(a)



(b)

图 5 无反馈和一次反馈后的结果

Fig.5 The results when no feedback is provided and those after the first feedback is provided

5 结束语 (Conclusion)

如何利用大量的未知标签图像点改善学习性能已成为当前图像检索研究中备受关注的问题. 本文的贡献就是利用未知标签图像点信息来半监督学习各图像点间的相关性, 解决有标签图像点不足的问题, 主要有以下两个方面: (1) 提出用图来建立图像库中各图像点间的相关性模型, 当有标签图像点不足时, 依靠未知标签图像点的信息能够尽可能准确地挖掘出图结构内部潜在的信息来辅助检索实现. (2) 通过改进 k 近邻的构造方法, 降低样本点分布对检索结果的影响, 增加了先验知识, 增大了有效信息量.

参考文献 (References)

- [1] Smeulders A W M, Worring M, Santini S, et al. Content-based image retrieval at the end of the early years[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(12): 1349-1380.
- [2] Zhou X S, Huang T S. Relevance feedback in image retrieval: A comprehensive review[J]. *Multimedia Systems*, 2003, 8(6): 536-544.
- [3] Chapelle O, Scholkopf B, Zien A. *Semi-supervised learning*[M]. Cambridge, MA, USA: MIT Press, 2006.
- [4] Wu Y, Tian Q, Huang T. Discriminant-EM algorithm with application to image retrieval[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ, USA: IEEE, 2000: 222-227.
- [5] Sia K C, King I. Relevance feedback based on parameter estimation of target distribution[C]//*Proceedings of the International Joint Conference on Neural Networks*. Piscataway, NJ, USA: IEEE, 2002: 1974-1979.
- [6] 鲁珂, 赵继东, 叶娅兰, 等. 一种用于图像检索的新型半监督学习算法[J]. *电子科技大学学报*, 2005, 34(5): 669-671.
- Lu K, Zhao J D, Ye Y L, et al. Algorithm for semi-supervised learning in image retrieval[J]. *Journal of University of Electronic Science and Technology of China*, 2005, 34(5): 669-671.
- [7] Blum A, Mitchell T. Combining labeled and unlabeled data with co-training[C]//*Proceedings of the Annual ACM Conference on Computational Learning Theory*. New York, NJ, USA: ACM, 1998: 92-100.
- [8] Zhou Z H, Li M. Tri-training: Exploiting unlabeled data using three classifiers[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2005, 17(11): 1529-1541.
- [9] El-Yaniv R, Pechyony D, Vapnik V. Large margin vs. large volume in transductive learning[J]. *Machine Learning*, 2008, 72(3): 173-188.
- [10] 薛贞霞, 刘三阳, 刘万里. 基于SVDD的渐进直推式支持向量机器学习算法[J]. *模式识别与人工智能*, 2008, 21(6): 721-727.
- Xue Z X, Liu S Y, Liu W L. SVDD based learning algorithm with progressive transductive support vector machines[J]. *Pattern Recognition and Artificial Intelligence*, 2008, 21(6): 721-727.
- [11] Tong S, Chang E. Support vector machine active learning for image retrieval[C]// *Proceedings of the ACM International Conference on Multimedia*. New York, NJ, USA: ACM, 2002: 107-118.
- [12] Zhou Z H, Chen K J, Dai H B. Enhancing relevance feedback in image retrieval using unlabeled data[J]. *ACM Transactions on Information Systems*, 2006, 24(2): 219-244.
- [13] Wang F, Zhang C S. Label propagation through linear neighborhoods[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2008, 20(1): 55-67.
- [14] Zhu X, Ghahramani Z. *Learning from labeled and unlabeled data with label propagation*[R]. Pittsburgh, PA, USA: Carnegie Mellon University, 2002.
- [15] 王和勇, 郑杰, 姚正安, 等. 基于聚类和改进距离的LLE方法在数据降维中的应用[J]. *计算机研究与发展*, 2006, 43(8): 1485-1490.
- Wang H Y, Zheng J, Yao Z A, et al. Application of dimension reduction on using improved LLE based on clustering[J]. *Journal of Computer Research and Development*, 2006, 43(8): 1485-1490.
- [16] Geng X, Zhan D C, Zhou Z H. Supervised nonlinear dimensionality reduction for visualization and classification[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2005, 35(6): 1098-1107.
- [17] Manjunath B S, Ohm J R, Vasudevan V V, et al. Color and texture descriptors[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(6): 703-715.
- [18] Stricker M A, Orengo M. *Similarity of color images*[C]//*Proceedings of SPIE - The International Society for Optical Engineering*. Bellingham, WA, USA: Society of Photo-Optical Instrumentation Engineers, 1995: 381-392.
- [19] Pinheiro A M G. Image description using scale-space edge pixel directions histogram[C]//*Proceedings of the Second International Workshop on Semantic Media Adaptation and Personalization*. Piscataway, NJ, USA: IEEE, 2007: 211-218.
- [20] Lin Y Y, Liu T L, Chen H T. Semantic manifold learning for image retrieval[C]//*Proceedings of the Annual ACM International Conference on Multimedia*. New York, NJ, USA: ACM, 2005: 249-258.
- [21] Yu J, Tian Q. Learning image manifolds by semantic subspace projection[C]//*Proceedings of the Annual ACM International Conference on Multimedia*. New York, NJ, USA: ACM, 2006: 297-306.

作者简介:

黄传波 (1972-), 男, 博士生. 研究领域为模式识别与图像处理等.

金忠 (1961-), 男, 博士, 教授, 博士生导师. 研究领域为模式识别, 图像处理等.