

基于 Nyström 低阶近似的半监督流形排序图像检索

傅向华¹ 李坚强¹ 王志强¹ 杜文峰¹

摘要 实际图像检索过程中, 用户提供的相关反馈有限, 但存在大量未标记图像数据. 本文在前期半监督流形图像检索工作的基础上, 提出一种基于 Nyström 低阶近似的半监督流形排序图像检索方法. 通过采用半监督的流形正则化框架, 将图像数据嵌入到低维流形结构中进行分类排序, 以充分利用大量未标记数据, 并兼顾分类误差、数据分布的几何结构以及分类函数的复杂性. 针对半监督学习速度缓慢的问题, 基于 Nyström 低阶近似对学习过程进行加速. 在较大规模的 Corel 图像数据集上进行了检索实验, 实验结果表明该方法能获得较好的效果.

关键词 图像检索, 流形学习, Nyström 近似, 半监督学习

DOI 10.3724/SP.J.1004.2011.00787

Semi-supervised Manifold-ranking-based Image Retrieval with Low-rank Nyström Approximation

FU Xiang-Hua¹ LI Jian-Qiang¹ WANG Zhi-Qiang¹ DU Wen-Feng¹

Abstract In the real image retrieval process, there are abundant unlabeled images whereas there only exist few labeled images. To address this issue, based on our previous work of semi-supervised manifold image retrieval, this paper proposed a novel learning method named semi-supervised manifold ranking based image retrieval (S²MRBIR). The images are assumed always embedded in low-dimensional sub-manifolds. In particular, S²MRBIR adopts the manifold regularization framework to rank the retrieved images while regarding the relevant feedback process of image retrieval as an online learning process and treating the image retrieval as a classification problem. The manifold regularization framework is capable of taking account of both labeled and unlabeled data, the classification performance, the geometric structures of the data distribution, and the complexity of the classifier. Moreover, an accelerating algorithm based on Low-rank Nyström approximation was proposed to improve the computing procedure of S²MRBIR (NA-S²MRBIR). Experimental results on Corel image database demonstrated the effectiveness of S²MRBIR.

Key words Image retrieval, manifold learning, Nyström approximation, semi-supervised learning

基于内容的图像检索 (Content based image retrieval, CBIR) 是图像检索的主要方法, 其通过提取图像的颜色、纹理、形状以及对象的空间关系等低层视觉特征进行相似查询和检索^[1-2]. 然而, 由于图像低层视觉特征与高层检索语义之间存在着较大的“语义鸿沟”, 使得 CBIR 检索精度受限. 为解决这一问题, 近年来一些研究人员或者通过提取高层语义特征进行基于语义的图像检索^[3], 或者在图像检索过程中整合用户的相关反馈^[4]. 相关反馈可视作监督学习问题, 要获得较好的性能, 则需提供较多的训练样例进行学习, 而实际图像检索过程中, 用户所能提供的相关反馈样例有限, 因此研究人员考虑

将半监督学习引入到图像检索中, 以利用大量未标记的样例来改善监督学习的分类性能, 如 Zhou 等将半监督学习和主动学习整合到相关反馈中, 利用大量未标记图像来提高相关反馈的性能^[5].

此外, 由于图像数据维度很高, 检索过程中一般需进行维数约简. 研究结果表明在表示数据时, 基于局部几何特性的流形结构比基于全局特性的欧几里德结构更具表现力^[6]. 因此一些研究人员将流形学习方法用于图像检索, 其假设内蕴图像空间为嵌入在特征空间之中的低维流形, 通过在图像流形中计算图像间的相似度, 对检索的图像进行排序. 流形排序 (Manifold rank) 算法的思想最早由 Zhou 等^[7] 提出, 随后被用于图像检索, 如 He 等提出一种基于流形排序的图像检索 (Manifold ranking based image retrieval, MRBIR) 方法, 其根据内蕴流形结构计算待查询图像与查询图像的相似度, 以此对数据库中的图像进行排序, 结果表明该方法优于主成分分析 (Principal component analysis, PCA) 等线性维数约简的方法^[8]; Wang 等提出利用核最大间隔投影来发现图像的局部流形结构, 然后通过将图像投

收稿日期 2010-01-21 录用日期 2010-11-24

Manuscript received January 21, 2010; accepted November 24, 2010

国家自然科学基金 (60903114, 61003271), 广东省自然科学基金 (7301329), 深圳市科技计划项目 (JC201005280463A) 资助

Supported by National Natural Science Foundation of China (60903114, 61003271), Natural Science Foundation of Guangdong (7301329), and Science and Technology Foundation of Shenzhen City (JC201005280463A)

1. 深圳大学计算机与软件学院 深圳 518060

1. College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060

影到低维子空间以改善图像检索的性能^[9]. Lin 等提出的增量关系嵌入 (Augmented relation embedding, ARE) 提出学习考虑用户喜好的语义流形, 其构造三个关系图, 一个描述相似关系, 另两个通过使用用户提供的相关反馈编码有关和无关关系^[10]; 李杰等在流形排序的基础上, 提出区域匹配图的方法^[11].

与此不同的是, 本文作者前期提出基于流形正则化框架的半监督图像检索方法^[12], 但该方法在半监督学习过程中计算量较大. 针对这一问题, 本文进一步提出基于 Nyström 低阶近似的半监督流形排序图像检索, 对其进行加速.

1 CBIR 中的相关反馈

在带相关反馈的图像检索中, 相关反馈过程可视为监督学习过程^[4-5]. 设查询图像为 \mathbf{q} , 待检索的图像数据集合为 $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, 图像检索则为确定能反映任意图像 \mathbf{x} 和查询 \mathbf{q} 之间相关关系的函数 $f(\cdot)$, 一般采用某种距离度量或相似性度量, 定义为: $f(\mathbf{x}, \mathbf{q}) = \text{sim}(\mathbf{x}, \mathbf{q})$, 其中 $\text{sim}(\cdot)$ 为相似度函数. 但如果从监督学习的视角, $f(\cdot)$ 可看作一个分类函数, 即: 若 \mathbf{x} 与 \mathbf{q} 相关, 则 $f(\mathbf{x}) > 0$, 否则 $f(\mathbf{x}) < 0$. 而检索过程中, 用户所提供反馈样例可用作训练样本, 对分类函数 $f(\cdot)$ 进行训练.

若检索过程中用户所提供的反馈样本集合为 $\{(\mathbf{x}_i, y_i)\}_{i=1}^l$, 其中 y_i 为 \mathbf{x}_i 的标签, 定义如下:

$$y_i = \begin{cases} 1, & \text{when } \mathbf{x}_i \text{ is relevant with } \mathbf{q} \\ -1, & \text{when } \mathbf{x}_i \text{ is irrelevant with } \mathbf{q} \end{cases}$$

那么, 可利用反馈样本集合作为训练数据进行监督学习, 以确定函数 $f(\cdot)$, 使得损失函数 $\mathcal{L}(\cdot)$ 取得最小值. 损失函数一般采用平方误差和, 即: $\mathcal{L}(\mathbf{x}, f(\mathbf{x}), y) = \sum_{i=1}^l (f(\mathbf{x}_i) - y_i)^2$. 从而自动分类的训练问题转换为式 (1) 的最优化问题:

$$\arg \min_f \mathcal{L}(\mathbf{x}, f(\mathbf{x}), y) = \sum_{i=1}^l (f(\mathbf{x}_i) - y_i)^2 \quad (1)$$

通过求解式 (1), 可得最优分类函数 $f^*(\cdot)$. 该分类函数可用于判断任意图像 \mathbf{x} 是否与查询图像 \mathbf{q} 相关.

2 基于流形正则化的半监督图像检索

2.1 半监督的流形正则化学习框架

给定数据集 $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l, \mathbf{x}_{l+1}, \dots, \mathbf{x}_n\} \subset \mathbf{R}^D$, 标签集合 $\mathcal{Y} = \{1, \dots, c\}$, 其中 \mathcal{X} 前 l 个数数据点 $\mathcal{X}_l = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\}$ 具有标签 $\{y_1, \dots, y_l\} \in \mathcal{Y}$, 后 u 个数数据点 $\mathcal{X}_u = \{\mathbf{x}_{l+1}, \mathbf{x}_{l+2}, \dots, \mathbf{x}_{l+u}\}$ (其中 $l+u=n$) 没有标签. 为利用未标注数据, Belkin 等^[13] 利用图拉普拉斯矩阵作为数据内蕴流形结构

的近似, 并将其作为惩罚项与核学习的正则化框架相结合, 提出半监督的拉普拉斯流形正则化框架. 该框架可转化为式 (2) 的最优化问题:

$$f^* = \arg \min_{f \in \mathcal{H}_K} \frac{1}{l} \sum_{i=1}^l \mathcal{L}_i(\mathbf{x}_i, y_i, f) + \gamma_A \|f\|_K^2 + \frac{\gamma_I}{(u+l)^2} \sum_{i,j=1}^{l+u} (f(\mathbf{x}_i) - f(\mathbf{x}_j))^2 \mathbf{W}_{ij} \quad (2)$$

其中, $\mathcal{L}_i(\mathbf{x}_i, y_i, f)$ 为样本 (\mathbf{x}_i, y_i) 导致的损失, \mathcal{H}_K 为与某个 Mercer 核 $K: \mathcal{X} \times \mathcal{X} \rightarrow \mathbf{R}$ 相关的可再生希尔伯特空间 (RPHS), $\|f\|_K$ 表示 \mathcal{H}_K 上定义的范数, 表征分类函数的复杂性, 决定分类函数 f 的泛化能力, γ_A 为正则化因子; \mathbf{W}_{ij} 为由所有数据点组成的数据邻接图中边的权值, 而 $(f(\mathbf{x}_i) - f(\mathbf{x}_j))^2$ 为数据点 \mathbf{x}_i 和 \mathbf{x}_j 的标签之间的差异, 显然, 若原始特征空间中邻近的点的标签相差很大, 那么 $(f(\mathbf{x}_i) - f(\mathbf{x}_j))^2 \mathbf{W}_{ij}$ 会变得很大, 因此, $\sum_{i,j=1}^{l+u} (f(\mathbf{x}_i) - f(\mathbf{x}_j))^2 \mathbf{W}_{ij}$ 可理解为保持数据几何结构的一致性, 即原始特征空间中邻近的点, 经 f 分类后, 应具有相似的标签.

若令 $\mathbf{f} = [f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_{l+u})]^T$, $L = D - W$, D 为对角矩阵, 且 $D_{ii} = \sum_{j=1}^{l+u} \mathbf{W}_{ij}$, 那么式 (2) 可写为

$$f^* = \arg \min_{f \in \mathcal{H}_K} \frac{1}{l} \sum_{i=1}^l \mathcal{L}_i(\mathbf{x}_i, y_i, f) + \gamma_A \|f\|_K^2 + \frac{\gamma_I}{(u+l)^2} \mathbf{f}^T L \mathbf{f} \quad (3)$$

根据表示定理, 式 (3) 的最优化问题的解形式为: $f^*(\mathbf{x}) = \sum_{i=1}^{l+u} \alpha_i K(\mathbf{x}, \mathbf{x}_i)$, 其中 $K(\cdot)$ 为 RPHS 中的核函数. 若样本 (\mathbf{x}_i, y_i) 的损失取 $\mathcal{L}_i(\mathbf{x}_i, y_i, f(\mathbf{x}_i)) = (y_i - f(\mathbf{x}_i))^2$, 可得最小二乘正则化 (LapRLS) 算法; 若样本 (\mathbf{x}_i, y_i) 的损失取 $\mathcal{L}_i(\mathbf{x}_i, y_i, f(\mathbf{x}_i)) = \max(0, 1 - y_i f(\mathbf{x}_i))$, 可得支持向量机的正则化 (LapSVM) 算法^[13].

2.2 基于流形正则化的半监督流形排序图像检索

考虑将上述基于流形正则化的半监督学习框架用于图像检索, 以对待查询的任意图像 \mathbf{x} 进行分类排序. 由于分类排序基于图像的内蕴流形结构进行, 故称其为流形排序图像检索.

流形排序图像检索在特征提取预处理之后, 进行两次排序, 如图 1 所示. 第一次排序时, 对于给定的查询图像 \mathbf{q} 和待查询图像集合 \mathcal{X} , 首先根据 L1 距离度量计算 \mathcal{X} 中的图像数据与查询图像 \mathbf{q} 的相似度, 并根据相似度值进行排序, 返回排序结果. 第二次排序时, 首先由用户从第一次排序结果中选择部分

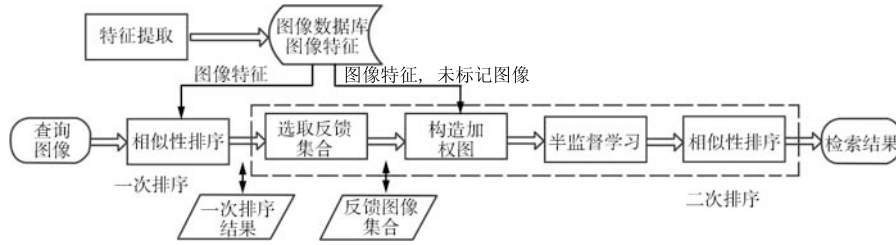


图 1 半监督流形排序图像检索过程

Fig. 1 Image retrieval process based on semi-supervised manifold ranking

图像作为反馈样例; 然后将反馈样例作为已标注样本, \mathcal{X} 中的其他图像作为未标注样本, 基于流形正则化框架进行半监督学习, 计算最优参数 α^* 和分类函数 $f^*(\mathbf{x}) = \sum_{i=1}^{l+u} \alpha_i^* K(\mathbf{x}_i, \mathbf{x})$; 最后根据训练好的分类函数 $f^*(\cdot)$ 对 \mathcal{X} 中的图像 \mathbf{x} 再次进行分类排序.

该检索过程可以分为嵌套的两部分: 1) 为图 1 中的虚线框部分, 其基于流形正则化框架, 利用已标记的反馈图像和大量未标记的数据库图像进行半监督学习, 训练可用于对图像数据库中的图像进行分类排序的分类模型; 2) 为给定查询, 返回结果图像的整个图像检索过程, 其需调用式 (1) 中的半监督学习过程. 分别用基于流形正则化的半监督学习算法 (MRBS²L)^[13] 和半监督流形排序图像检索算法 (S²MRBIR) 进行描述:

1) 基于流形正则化的半监督学习算法 (MRBS²L)

输入. l 个已标注图像数据 $\{(\mathbf{x}_i, y_i)\}_{i=1}^l$, u 个无标注图像数据 $\{\mathbf{x}_j\}_{j=l+1}^{l+u}$, 正则化参数 γ_A 和 γ_I .

输出. 分类函数 $f^*: \mathbf{R}^n \rightarrow \mathbf{R}$.

步骤 1. 利用 $l+u$ 个数据结点, 构建数据邻接图 G , 图 G 中边的权值 W_{ij} 采用热核权值, 即若 $i \neq j$, 则 $W_{ij} = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2)$, 否则, $W_{ii} = 0$;

步骤 2. 选择某个核函数 $K(\cdot)$, 计算 $l+u$ 个图像数据的 Gram 矩阵 $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$;

步骤 3. 计算邻接图 G 的拉普拉斯矩阵: $L = D - W$, 其中 D 为对角矩阵, 其对角元素 $D_{ii} = \sum_{j=1}^{l+u} W_{ij}$;

步骤 4. 利用 LapRLS 或 LapSVM 计算 α^* ;

步骤 5. 输出函数 $f^*(\mathbf{x}) = \sum_{i=1}^{l+u} \alpha_i^* K(\mathbf{x}_i, \mathbf{x})$.

在用户进行相关反馈之前, 用于分类排序的分类器尚未进行训练. 因此, 流形排序图像检索中, 需要进行两次排序: 1) 对于提交的查询图像 q , 首先根据 L1 距离计算图像集合 \mathcal{X} 中的图像与 q 的相似度, 对 \mathcal{X} 中的图像进行排序并返回结果, 由用户从返回的图像中选择 l 个图像进行相关反馈; 2) 将用户反馈的 l 个图像作为半监督学习算法的已标注样本数据, 并利用数据库中其他未标注的图像进行半监督

学习训练, 然后再利用训练好的半监督分类器对图像数据进行分类排序, 返回查询结果. 整个半监督流形排序图像检索算法 (S²MRBIR) 算法描述如下:

2) 流形排序图像检索算法 (S²MRBIR)

输入. 查询图像 q , 图像集合 \mathcal{X} .

输出. 检索结果集合.

步骤 1. 对查询图像 q 和待查询图像集合 \mathcal{X} 进行特征提取, 计算 q 与 \mathcal{X} 的特征空间;

步骤 2. 基于 L1 距离度量, 计算 q 与 \mathcal{X} 中图像的相似度, 根据相似度结果由大到小返回 k 个图像;

步骤 3. 由用户从返回的结果图像中选择 l 个图像, 对其进行标记, 形成反馈样例集合;

步骤 4. 选择 \mathcal{X} 中剩余的 $n-l$ 个未标记图像, 形成未标记的图像集合;

步骤 5. 调用 MRBS²L 算法, 训练分类函数 f^* ;

步骤 6. 根据分类函数 f^* , 对 \mathcal{X} 中的图像进行分类排序, 返回结果集合.

其中, 半监督学习过程耗时较多. MRBS²L 算法中, 涉及到 Gram 矩阵的求逆, 其时间复杂度为 $O((l+u)^3)$, 即 $O(n^3)$. 而 S²MRBIR 算法的时间复杂度由 MRBS²L 算法决定, 故其时间复杂度亦为 $O(n^3)$.

3 基于 Nyström 方法的核低阶近似

上述 MRBS²L 和 S²MRBIR 算法中, 当数据集规模增大时, 会导致核矩阵 K 增加, 从而使训练时间大大增加. 为改进半监督学习的计算过程, 以加快流形排序图像检索的速度, 采用改进的 Nyström 方法对核矩阵进行低阶近似^[14].

对由样本集合 $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 构成的核矩阵 K 进行特征分解, 则:

$$K = \Phi_{\mathcal{X}} \Lambda_{\mathcal{X}} \Phi_{\mathcal{X}}^T \quad (4)$$

其中, $\Lambda_{\mathcal{X}} = \text{diag}\{\lambda_i^{(F)}\}_i$, $\lambda_1^{(F)} \geq \lambda_2^{(F)} \geq \dots \geq 0$, $\Phi_{\mathcal{X}}$ 为正规化矩阵. Nyström 方法从样本集合 \mathcal{X} 中随机选择包含 m 个样本的子集 $\mathcal{Z} = \{\mathbf{z}_i\}_{i=1}^m$, 利用其来近似核矩阵 K , 则式 (4) 的特征系统可表示为

$$\frac{\Phi_{\mathcal{X}} \simeq \sqrt{\frac{m}{n}} E \Phi_{\mathcal{Z}} \Lambda_{\mathcal{Z}}^{-1}, \Lambda_{\mathcal{X}} \simeq n}{m \Lambda_{\mathcal{Z}}} \quad (5)$$

其中, $E \in \mathbf{R}^{n \times m}$, $E_{ij} = K(\mathbf{x}_i, \mathbf{z}_j)$, 而 $\Phi_{\mathcal{Z}}, \Lambda_{\mathcal{Z}} \in \mathbf{R}^{m \times m}$ 包含了 $H \in \mathbf{R}^{m \times m}$ 的特征向量和特征值, 其中 $H_{ij} = K(\mathbf{z}_i, \mathbf{z}_j)$. 将式 (5) 代入式 (4), 则得到核矩阵 K 的 Nyström 低阶近似基本公式:

$$K \simeq \left(\sqrt{\frac{m}{n}} E \Phi_{\mathcal{Z}} \Lambda_{\mathcal{Z}}^{-1} \right) \left(\frac{n}{m} \Lambda_{\mathcal{Z}} \right) \left(\sqrt{\frac{m}{n}} E \Phi_{\mathcal{Z}} \Lambda_{\mathcal{Z}}^{-1} \right)^T = EH^{-1}E^T = \tilde{K} \quad (6)$$

并且, 式 (6) 的近似误差可表示为 K 与其近似值差值的 Frobenius 范数, 则

$$\varepsilon = \|K - EH^{-1}E^T\|_F \quad (7)$$

根据式 (7), 显然误差值越小, 近似值越准确. 对于常用的核, 影响近似质量的最重要因素是: \mathcal{X} 中每个采样点与 \mathcal{Z} 中离其最近的界标 (Landmark) 点之间的量化误差. 如果这一量化误差为 0, 那么对核矩阵的 Nyström 低阶近似是完全准确的. 而由于 K-means 聚类可获得量化误差的局部最小值, 因此可以采用 K-means 聚类所获得的中心点作为界标点, 设 m 为样本子集 \mathcal{Z} 所期望的界标点数目, 则 m 越大, Nyström 近似越精确, 然而计算量也会越大.

上述基于 K-means 的 Nyström 核矩阵低阶近似算法^[14] 描述如下:

输入. 样本集合 \mathcal{X} , 界标点数目 m .

输出. 低阶近似矩阵 \tilde{K} .

步骤 1. 对样本集合 \mathcal{X} 运行 K-means 聚类算法, 获得 m 个聚类中心点;

步骤 2. 将 m 个中心点作为样本子集 \mathcal{Z} , 根据样本集合 \mathcal{X} 和样本子集 \mathcal{Z} , 计算矩阵 E ;

步骤 3. 计算 m 个中心点确定的核矩阵 H 和逆阵 H^{-1} ;

步骤 4. 根据公式 $\tilde{K} = EH^{-1}E^T$ 计算核近似矩阵 \tilde{K} .

在该近似算法中, 步骤 1 聚类的时间复杂度为 $O(tmn)$, t 为迭代的次数; 步骤 2 计算矩阵 E 的时间复杂度为 $O(mn)$; 步骤 3 计算核矩阵 H 的时间复杂度为 $O(m^2)$, 计算逆阵 H^{-1} 的时间复杂度为 $O(m^3)$; 步骤 4 计算 \tilde{K} 的时间复杂度为 $O(mn)$. 由于一般 $m \ll n$, $t \ll n$, 故基于 K-means 的 Nyström 核矩阵低阶近似算法的时间复杂度为 $O(mn)$.

可以利用近似核 \tilde{K} 代替第 2.2 节中 MRBS²L 算法的核矩阵 K , 以加速整个半监督流形排序图像检索的处理过程, 本文称其为基于 Nyström 低阶近

似的半监督流形排序图像检索 (NA-S²MRBIR), 其对应的时间复杂度为 $O(mn^2)$.

4 实验结果与分析

实验对本文所提的 NA-S²MRBIR, S²MRBIR 方法与 MRBIR, SVM 以及 L1 距离法进行比较. 为验证本文算法的有效性, 选择在较大规模的 Corel 图像数据库中进行仿真实验. 实验环境为 CPU Pentium 4 3.00 GHz, 内存 1 G, 操作系统为 Windows XP.

4.1 实验设置与评价指标

实验所采用的 Corel 图像数据库共包括 9908 张图片, 包含 96 个类别, 涉及自然景观、动植物、历史建筑、城市生活等多个方面, 每个类别包含 85 ~ 120 个图像, 每幅图像大小为 128 像素 \times 85 像素或 128 像素 \times 96 像素. 实验分为两部分, 首先在 Corel 图像子集上验证本文所提半监督流形排序方法的有效性, 然后在整个 Corel 图像数据库上验证 Nyström 低阶近似中 m 取值对查准率的影响.

实验时, 每次从待测图像集合中任意选择一幅图像作为查询图像, 剩余的其他图像作为待查询图像, 然后对检索结果进行统计评价. 检索性能的评价指标采用查准率范围曲线 (Precision-scope curve)^[15], 其中范围表示由用户指定的在检索返回结果中排在最前面的图像数目 N , 而查准率表示用户指定的范围 N 中相关图像的比率. 查准率范围曲线描述了不同范围时的查准率, 可以用于评价算法的整体性能. 考虑到实际应用, 并方便与已有的其他检索方法比较, 实验中在选择图像特征时只考虑颜色、纹理等全局特征, 各特征及维数如表 1 所示.

表 1 实验所采用的图像特征及维数
Table 1 Image features and dimensions used in the experiments

名称	描述维数
颜色相关图	HSV 空间, 144 维
颜色直方图	HSV 空间, 32 维
颜色矩	HSV 空间, 9 维
Tamura 纹理特征	粗糙度、对比度、方向度, 20 维
金字塔结构的小波纹理特征	24 维

4.2 实验结果及分析

实验过程分成两部分, 首先验证本文所提半监督流形排序方法的有效性. 该部分实验中, 将 NA-S²MRBIR, S²MRBIR, MRBIR, SVM 以及 L1 距离法进行比较. 由于计算量很大, 参考已有研究文

文献的实验设置情况, 每次选择 Corel 图像数据库的子集进行实验. 整个实验分 4 组进行, 每组实验从 Corel 图像数据库中选择 10 个特征比较明显的类别, 每个类别包含具有相同内容的 100 幅图像, 共 1 000 幅图像, 形成一个 Corel 图像子集. 其中一组子集每个类别的典型图像如图 2 所示.



图 2 图像数据库中 10 类图像的典型图像
Fig. 2 Typical images of ten categories in Corel image database

相关反馈中, 用户标记 10 幅图像, 即 $l = 10$. 参照文献 [13–14] 中的实验结果, MRBIR 方法的参数分别设为: $k = 200$, $\alpha = 0.99$, $\sigma_l = 0.05$, 迭代次数设为 50 次. 而 S^2 MRBIR 方法中, 邻接图的边权值 W_{ij} 采用热核权值, 经过多次实验, 正则化参数设置为 $\gamma_A l = \gamma_l l / (u + l)^2 = 0.005$. NA- S^2 MRBIR 的参数 m 取值 200, 其他参数取值与 S^2 MRBIR 方法相同. SVM 的参数采用自动优化确定的参数. 分别对 4 组 Corel 图像子集进行无标记数据的无相关反馈实验和有部分标记的有相关反馈实验, 并计算相同参数时 4 组 Corel 图像子集上的实验结果平均值, 得到无相关反馈时的实验结果如图 3 所示, 有相关反馈的实验结果如图 4 所示.

图 3 的结果表明, 在无相关反馈情况下, 虽然 NA- S^2 MRBIR, S^2 MRBIR, MRBIR 方法的结果优于 L1 方法, 但检索准确度提高不是很大, 在 Scope 为 10 时的准确率约为 42%, 在 Scope 为 20 时的准确度约为 35%, 而随着返回结果的增多, 准确度下降. 与图 3 相比, 图 4 中有反馈的图像检索准确度明显改善, 在 Scope 为 10 时的准确率均达到 67% 以上, 在 Scope 为 20 时 NA- S^2 MRBIR, S^2 MRBIR, MRBIR 方法均达到 58% 以上的准确度, 说明相关反馈对于改善图像检索准确度很有效; 并且, 从图 4 也可以看出, 采用流形排序的 NA- S^2 MRBIR, S^2 MRBIR, MRBIR 的方法要明显优于 SVM 方法, 而本文所提出的采用半监督学习的 NA- S^2 MRBIR, S^2 MRBIR 方法结果又要稍优于 MRBIR 方法, 说

明利用大量未标记数据集可以改善检索结果.

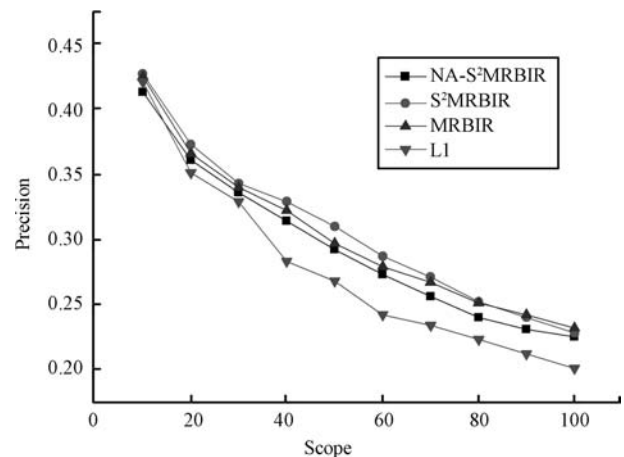


图 3 无相关反馈时的图像检索实验结果
Fig. 3 Retrieval experiment results without relevance feedback

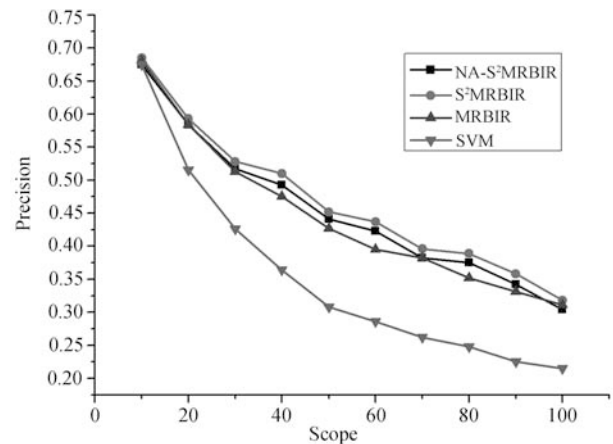


图 4 带相关反馈的图像检索实验结果
Fig. 4 Retrieval experiment results with relevance feedback

不同方法在 4 组 Corel 图像子集上的平均运行时间如表 2 所示, 包括无反馈时的图像检索时间和反馈的时间. 由表 2 可以看出, 在所有的方法中, L1 距离法的处理时间最快, 其次为 SVM 方法, 而基于流形的方法 (MRBIR, S^2 MRBIR, NA- S^2 MRBIR) 由于计算量较大, 处理时间都相对较长, 尤其是 S^2 MRBIR 方法, 因为增加了半监督学习过程, 处理时间远大于其他方法; 而采用 Nyström 低阶近似后的 NA- S^2 MRBIR 方法相对于 S^2 MRBIR 方法, 无论是无反馈还是有反馈, 处理时间都大大减少, 即使与 MRBIR 方法相比时间也少很多. 综合图 3 和图 4 的实验结果, 说明 NA- S^2 MRBIR 方法能够在几乎不影响检索准确度的情况下, 有效提升图像检索的处理速度.

第二部分实验是验证 Nyström 低阶近似中, m 取值对查准率的影响. 实验中分别选择 $m = 10, 20, 50, 100, 200, 500, 1000$, 在由所有图像组成的 Corel 数据库中进行实验, 实验结果如图 5 所示. 从实验结果可以看出, 当 m 较小时, 增加 m 的值, 会提高查准率; 但当 m 到一定的值以后, 如 $m = 200$ 和 $m = 500, m = 1000$ 的查准率相比, 差别并不大, 这说明, m 并不需取很大的值, 就可以采用基于 K-means 聚类的 Nyström 低阶近似方法较好地近似原来的 Gram 矩阵, 从而提高原问题的处理速度.

表 2 不同方法的平均处理时间

Table 2 Comparison of average processing time

时间 (s)	L1	MRBIR	SVM	S ² MRBIR	NA-S ² MRBIR
无反馈	0.012	0.768	-	1.635	0.276
反馈	-	0.783	0.087	1.627	0.282

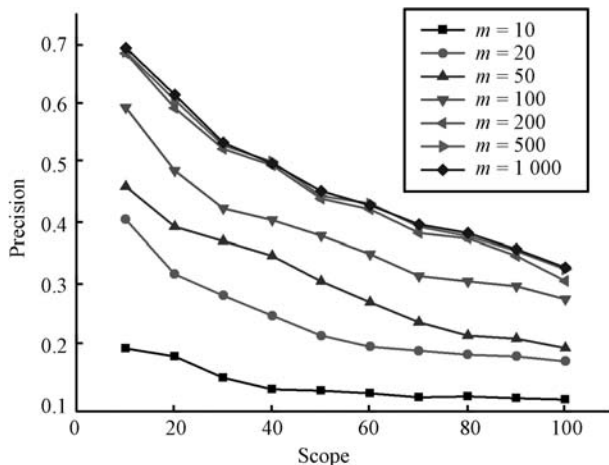


图 5 NA-S²MRBIR 方法 m 取不同值时的实验结果比较
Fig. 5 Experiment results of the NA-S²MRBIR method with different m values

5 结论

本文在前期半监督流形图像检索工作的基础上, 提出一种基于 Nyström 低阶近似的半监督流形排序图像检索方法, 通过采用半监督的流形正则化框架对检索图像进行分类排序, 充分利用未标记数据, 并基于 Nyström 低阶近似对计算过程进行加速. 该方法既支持无反馈的图像检索, 也支持有反馈的图像检索. 在 Corel 图像数据集上的实验结果表明, 该方法能获得较好的效果.

不过, 由于用户提供的相关反馈经常动态变化的, 因此如何使流形正则化算法能够利用已有的学

习结果进行增量学习, 值得进一步研究. 而且, 对于半监督学习过程的加速, 也可以考虑其他基于采样的近似计算方法.

References

- 1 Smeulders A W M, Worring M, Santini S, Gupta A, Jain R. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(12): 1349–1380
- 2 Liu Qing, Xu Lu-Ping, Ma Yi-De, Wang Yong. Image NMI feature extraction and retrieval method based on pulse coupled neural networks. *Acta Automatica Sinica*, 2010, **36**(7): 931–938
(刘勃, 许录平, 马义德, 王勇. 基于脉冲耦合神经网络的图像 NMI 特征提取及检索方法. *自动化学报*, 2010, **36**(7): 931–938)
- 3 Liu Y, Zhang D S, Lu G J, Ma W Y. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 2007, **40**(1): 262–282
- 4 Arevalillo-Herraez M, Ferri F J, Domingo J. A naive relevance feedback model for content-based image retrieval using multiple similarity measures. *Pattern Recognition*, **43**(3): 619–629
- 5 Zhou Z H, Chen K J, Dai H B. Enhancing relevance feedback in image retrieval using unlabeled data. *ACM Transactions on Information Systems*, 2006, **24**(2): 219–244
- 6 Luo Si-Wei, Zhao Lian-Wei. Manifold learning algorithms based on spectral graph theory. *Journal of Computer Research and Development*, 2006, **43**(7): 1173–1179
(罗四维, 赵连伟. 基于谱图理论的流形学习算法. *计算机研究与发展*, 2006, **43**(7): 1173–1179)
- 7 Zhou D Y, Weston J, Gretton A, Bousquet O, Scholkopf B. Ranking on data manifolds. *Advances in Neural Information Processing Systems*. Cambridge: The MIT Press, 2004. 169–176
- 8 He J R, Li M J, Zhang H J, Tong H H, Zhang C S. Generalized manifold-ranking-based image retrieval. *IEEE Transactions on Image Processing*, 2006, **15**(10): 3170–3177
- 9 Wang C, Zhao J, He X F, Chen C, Bu J J. Image retrieval using nonlinear manifold embedding. *Neurocomputing*, 2009, **72**(16–18): 3922–3929
- 10 Lin Y Y, Liu T L, Chen H T. Semantic manifold learning for image retrieval. In: *Proceedings of the 13th Annual ACM International Conference on Multimedia*. New York, USA: ACM, 2005. 249–258
- 11 Li Jie, Cheng Yi-Min, Ge Shi-Ming, Zhang Ling. Integrated manifold ranking and region matching for image retrieval. *Journal of Chinese Computer Systems*, 2008, **29**(3): 511–515
(李杰, 程义民, 葛仕明, 张玲. 结合流形排序和区域匹配的图像检索. *小型微型计算机系统*, 2008, **29**(3): 511–515)

- 12 Wang Zhi-Qiang, Fu Xiang-Hua, Zhao Liang-Hui, Du Wen-Feng. Manifold based semi-supervised learning for content-based image retrieval. *Geomatics and Information Science of Wuhan University*, 2009, **34**(8): 928–931, 935
(王志强, 傅向华, 赵良辉, 杜文峰. 基于内容的半监督流形排序图像检索. 武汉大学学报信息科学版, 2009, **34**(8): 928–931, 935)
- 13 Belkin M, Niyogi P, Sindhvani V. Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. *The Journal of Machine Learning Research*, 2006, **7**: 2399–2434
- 14 Zhang K, Tsang I W, Kwok J T. Improved Nyström low-rank approximation and error analysis. In: *Proceedings of the 25th International Conference on Machine Learning*. New York, USA: ACM, 2008. 1232–1239
- 15 Huijismans D P, Sebe N. How to complete performance graphs in content-based image retrieval: add generality and normalize scope. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(2): 245–251



傅向华 深圳大学计算机与软件学院副教授. 2005 年获西安交通大学电信学院计算机系博士学位. 主要研究方向为数据挖掘, 信息检索. 本文通信作者.

E-mail: fuxh@szu.edu.cn

(**FU Xiang-Hua** Associate professor at the College of Computer Science and Software Engineering, Shenzhen University. He received his Ph. D. degree from Xi'an Jiaotong University in 2005. His research interest covers data mining and information retrieval. Corresponding author of this paper.)

He received his Ph. D. degree from Xi'an Jiaotong University in 2005. His research interest covers data mining and information retrieval. Corresponding author of this paper.)



李坚强 深圳大学计算机与软件学院讲师. 2008 年获华南理工大学自动化学院博士学位. 主要研究方向为自动化控制和机器人. E-mail: lijq@szu.edu.cn

(**LI Jian-Qiang** Lecturer at the College of Computer Science and Software Engineering, Shenzhen University. He received his Ph. D. degree from South

China University of Technology in 2008. His research interest covers automatic control and robots.)



王志强 深圳大学计算机与软件学院教授. 1992 年获湖北工业大学工学硕士学位. 主要研究方向为多媒体信息处理. E-mail: wangzq@szu.edu.cn

(**WANG Zhi-Qiang** Professor at the College of Computer Science and Software Engineering, Shenzhen University. He received his master degree

from Hubei University of Technology in 1992. His main research interest is multimedia information processing.)



杜文峰 深圳大学计算机与软件学院讲师. 2007 年获华中南大学计算机学院博士学位. 主要研究方向为无线网络. E-mail: duwf@szu.edu.cn

(**DU Wen-Feng** Lecturer at the College of Computer Science and Software Engineering, Shenzhen University. He received his Ph. D. degree from Central South University in 2007. His main research interest is wireless network.)

He received his Ph. D. degree from Central South University in 2007. His main research interest is wireless network.)