

文章编号:1009-3087(2016)05-0135-08

DOI:10.15961/j.jsuese.2016.05.019

基于最佳路径森林分类的贪婪学习方法在 CBIR 系统的应用

孙 挺^{1,2},耿国华²

(1. 周口师范学院 网络工程学院,河南 周口 466001;2. 西北大学 可视化研究所,陕西 西安 710069)

摘 要:针对一般相关反馈的基于内容图像检索(CBIR)方法不能有效处理相关图像和非相关图像的问题,提出了一种基于最佳路径森林分类(OPF)的贪婪学习方法(GL-OPF),该方法旨在返回每次迭代查询的最相关图像。首先,查询图像和数据集图像通过 Gabor 小波变换提取特征向量;然后,通过 GL-OPF 主动学习方法获得图像关联性反馈,生成标记训练集;最后,标记训练集通过 OPF 分类器进一步评估形成相关性和非相关性原型集,每次迭代都会返回查询的最相关图像。3 个公开图像数据集 Caltech101、Corel 和 Pascal 上的实验验证了本文方法的有效性。实验结果表明,在 3 个数据集中,迭代 8 次时,GL-OPF 的查询精度比其他 3 种方法均有较大提高,此外,GL-OPF 的迭代运行时间和查询时间与 OPF 几乎相同,很大程度改进了 OPF 方法。

关键词:基于内容图像检索;最佳路径森林分类;贪婪学习;Gabor 小波;相关性反馈

中图分类号:TP391

文献标志码:A

Application of Greedy Learning Based on Optimum-path Forest Classification in CBIR System

SUN Ting^{1,2}, GENG Guohua²

(1. College of Network Eng., Zhoukou Normal Univ., Zhoukou 466001, China;

2. Visualization Inst., Northwestern Univ., Xi'an 710069, China)

Abstract: In order to deal with related images and non-related images effectively in content-based image retrieval (CBIR), a method of greedy learning based on optimum-path forest classification (OPF), named as GL-OPF, was proposed. Firstly, feature vectors of query images and the images in database were extracted by Gabor wavelet transform. Then, the relevance feedback of images was obtained by GL-OPF active learning, generating training set of tags. Finally, prototype sets of relevance and unrelated were formed by further evaluation of OPF classifier of mark sets, and the most relevant query images would return after every iteration. The effectiveness of proposed method was verified by experiments on the three image databases Caltech101, Corel and Pascal. The experimental results showed that in eight iterations, the query precision of GL-OPF rises more than that of other three methods. In addition, the running and query time of GL-OPF is almost the same as that of OPF.

Key words: content-based image retrieval; optimum-path forest classification; greedy learning; Gabor wavelet; relevance feedback

基于内容的图像检索(content-based image retrieval, CBIR)是机器视觉和数据挖掘领域最热门的研究课题之一^[1-2]。CBIR 系统根据用户给定的查询条件返回数据集中的相关图像。由于问题的不确定性,针对不同用户的同一查询可能会改变相关性的意义。而从现实角度看,最小化响应时间、有效标

记的图像数量和最大化用户满意度是 CBIR 的最主要挑战^[3-4]。

已经有很多研究者对 CBIR 进行深入研究,在众多方法中,基于智能算法的 CBIR 研究最广。例如,文献[5]将粒子群优化算法与用户的反馈(reviewer feedback, RF)过程结合,提出了一种基于粒子群的

收稿日期:2015-09-14

基金项目:国家重点基础研究发展计划前期研究专项资助(2011CB311802);河南省科技厅科技发展计划科技攻关项目资助(122400450356);河南省科技厅科技发展计划软科学项目资助(132400410927)

作者简介:孙 挺(1972—),男,副教授,博士。研究方向:图像处理;智能信息处理;科学可视化。E-mail:suntingzk@126.com

网络出版时间:2016-9-19 9:43 网络出版地址: <http://www.cnki.net/kcms/detail/51.1596.T.20160919.0943.002.html>

<http://jsuese.scu.edu.cn>

CBIR 方法, 简称为 PSO-RF 算法, 该方法可以避免初始检索对用户认知的影响以及对反馈效果造成的局限性^[5], 但容易出现标注样本数不足的问题。

文献[6]通过合并未标记图像提出了一种基于改进二进制重力搜索的 CBIR 方法(improved binary gravitational search algorithm, IBCSA)。通过最小化已标记数据的观测误差函数来选择综合性能最好的回归函数, 这样可以兼顾图像的语义特征及图像空间的几何结构^[6], 然而, 该方法依然存在视觉特征和语义信息的“语义鸿沟”问题。文献[7]提出了一种基于最佳路径森林分类(optimum path forest, OPF)的 CBIR 方法。其中, 数据集图像按照相关类进行分类, 图像通过正反例间的距离进行排序, 该距离在前驱迭代中通过用户标注的图像获得。该方法可以弱化检索问题的不确定性, 但其通过简单的贪婪算法^[8]和基于支持向量机^[9](support vector machine, SVM)的方法所获得的分类效果比较差, 且在给定应用中也无法估计合适的迭代次数。

文献[10]提出了排名分数与排名位置的合并策略, 通过图像多样性惩罚贪婪函数。在逐次迭代消减排序中, 计算排在前面图像的相似度分数^[10], 使图像排序结果具有较高的相关性和多样性, 但对于小样本训练集容易产生局部最优, 影响检索结果。文献[11]是一种图像检索的集成方法, 利用色矩(color metric, CM)来代表颜色特征, 利用局部二进制模式(local binary patterns, LBP)来代表纹理, 该方法非常注重目标检索图像的表达形式, 然而, 没有考虑后续检索效率。文献[12]是一种基于最近邻的改进相关反馈算法, 根据检索图像和相关后验概率定义随机变量, 以获得相关性正确概率, 利用正则化得到的平滑估计解决小样本问题。然而, 没有解决初始检索对用户认知的影响和相关反馈造成的局限性。

文献[7-12]都倾向于处理非相关图像, 而本文提出了一种基于 OPF 分类器的贪婪算法(GL-OPF)。在预测过程中, 使用相关或非相关图像进行最佳路径代价排序; 减少被动原型点的数量, 使主动原型点数量多于被动原点; 在几次迭代过程中返回最大信息量的图像, 并从返回结果中找出最相关图像。提出的方法在查询精度和查全率方面有明显提高, 且运行和查询时间与原 OPF^[7]相差不大, 是 OPF^[7]在 CBIR 中的有效改进。

1 图像特征表征

Gabor 小波变换具有多分辨率性和多方向性, 2 维

Gabor 函数 $g(x, y)$ 是正弦调制高斯函数的结果, 正弦频率 ω 和标准差 σ_x, σ_y 决定 $g(x, y)$, 其定义如下:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{[-(1/2)((x^2/\sigma_x^2) + (y^2/\sigma_y^2)) + 2\pi j\omega x]} \quad (1)$$

式中, $g(x, y)$ 的扩张和旋转所得的 Gabor 小波表示为:

$$g_{mn}(x, y) = a^{-m} g(x', y') \quad (2)$$

式中: $x' = a^{-m}(x\cos\theta + y\sin\theta)$, $y' = a^{-m}(-x\sin\theta + y\cos\theta)$ 表示尺度, $m \in \{0, 1, \dots, M-1\}$; $\theta = n\pi/N$ 表示方向, $n \in \{0, 1, \dots, N-1\}$; a 的定义如下:

$$a = (U_h/U_l)^{(-\frac{1}{M-1})} \quad (3)$$

式中, U_h 和 U_l 分别为设计频段的上界和下界, 详细信息参考文献[13]。本文的 Gabor 变换选择了 2 个尺度和 4 个方向。

一般而言, 图像检索最终解释为相关性的计算问题, 而特征表示是相关性计算的前提, 不同的特征会得到不同的相关性, 如 2 幅图像中形状的相关性、颜色特征的相关性以及纹理特征的相关性可能有很大差异。对于图像检索, 图像的纹理和形状是区别图像类别最主要的特征。Gabor 小波提取的特征能够在空间域和频域同时取得最优局部化的特性, 因此能很好地描述空间频率(尺度)和空间位置等局部结构信息。因此, 本文的图像特征利用 Gabor 小波提取, 具体是通过 2 维 Gabor 小波变换获取高频和低频系数。对于数据集中的每幅图像, 在提取 Gabor 小波前, 先将其转换为灰度图像, 且构建的特征向量包含归一化的 Gabor 小波系数的前 2 个矩, 其中, 小波系数的标准差和均值用于计算 Gabor 小波系数; 然后将标准差和均值存储在特征向量中, 这些元素可以表征图像的尺度信息和空间位置结构信息。本文的系统框架如图 1 所示。

2 最佳路径森林分类的贪婪学习与检索

假定图像数据集为 Z , 图像 $t \in Z$, 图像的特征用 $v(t)$ 表示, 由特征提取函数 v 计算得到。利用距离函数 $d(s, t)$ 测量 2 幅图像 $s, t \in Z$ 间的相似度, 数据对 (v, d) 称为描述子。本文方法使用单描述子 (v, d) 进行算法描述和评估。

大多数分类表是基于贪婪方法的相关性反馈, 图 2 给出贪婪方法^[14]的 3 种实例。图 2(a)是将第 1 次迭代后的正例(相关性图像)用于移除下一查询点(位于特征空间的几何中心)。该方法已经被成功地应用于 CBIR 系统^[15], 其他 2 种方法使用相关图像作为下一查询点, 且依赖于查询点集间的距离, 查询点集和正例在特征空间中形成不同的 iso-曲面

(图2(b)、(c))。

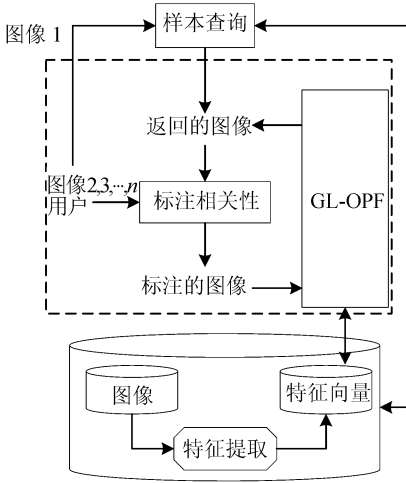


图1 本文CBIR系统框架图

Fig.1 Framework of the proposed CBIR system

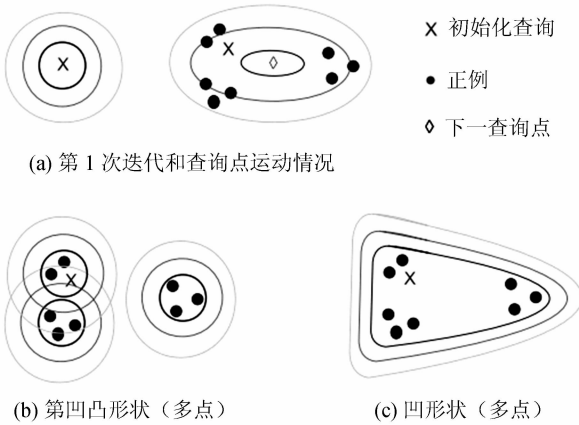


图2 简单贪婪方法改变的查询形状

Fig.2 Query shape change by simple greedy method

2.1 采用最佳路径森林的学习和分类

通常,用户意图列表 X 中包含了相关和非相关图像,经过几次迭代后,主动学习方法可以获得图像关联性反馈。生成标记训练集 T ,每次迭代后,该集合从 $T \leftarrow T \cup X$ 生成的相关性反馈结果中获取新的元素。本文将集合 T 用于设计最佳路径森林分类器 OPF,该过程包括第一步评估得到 T 中用于每类图像的代表性样本(原型),形成相关性和非相关性原型集: $S_R \subset T$ 和 $S_I \subset T$ 。训练过程需要一幅完全图,图中的结点都是数据集 T 中的元素,权重集 $arcs(s, t)$ 由 $d(s, t)$ 度量得到。图中的每条路径都有代价,通过计算从 $S_R \cup S_I$ 到每个结点 $t \in T$ 的代价得到最小代价路径。随后,该分类器评估集合 $Z \setminus T$ 中的图像,按递增顺序计算得到从原型集 $S_R \cup S_I$ 到每个结点 $t \in Z \setminus T$ 的代价,从而得到最佳路径。并且当最佳路径根植于 S_R 时,在相关性减小的集合 $Y \in Z \setminus T$ 中

插入结点 t 。

在给定的完全图中,图中结点 t 是 T 中的元素。路径 π_t 由不重复的结点序列 $\{t_1, t_2, \dots, t_n = t\}$ 表示。从起点 $R(\pi_t) = t_1$ 到终点 $t_n = t$ 的长度(最弱链)即度量函数 $d(t_i, t_{i+1})$ 的最大值 $C(\pi_t)$,计算方法定义如下:

$$C(\pi_t) = \max_{i=1,2,\dots,n-1} \{d(t_i, t_{i+1})\} \quad (4)$$

当 $C(\pi_t) \leq C(\pi_l)$ 时,以 t 结尾的路径 π_t 最佳;对 T 中的每一结点 t ,计算起点为 $R(t) \in S_R \cup S_I$,终点为 t 的最佳路径。该方法主要是为了获取 T 的最佳路径划分,这样 S_R 和 S_I 中所有相关与非相关原型便会完全分隔,且每个训练结点 $t \in T$ 会被分配到与该结点所属原型具有最强链接的类。该划分被推断为最佳路径森林图 P - 非循环图。在前驱图 P 中存储了所有最佳路径,即式(4)用于分配每个结点 $t \in T \setminus S_R \cup S_I$ 在最佳路径集 $S_R \cup S_I$ 中的前驱结点 $P(t)$;或者当 $t \in S_R \cup S_I$ 时,标记 $P(t) = nil$ 。

算法1是OPF算法,阐述了最佳路径森林 P 中路径代价函数 C (式(4))的计算过程。式(4)为每个结点 $t \in T$ 输出以 t 为终点的最佳路径的最小代价 $C(t)$ 、根结点 $R(t) \in S_R \cup S_I$ 、在最佳路径中的前驱结点 $P(t)$ 。按代价 $C(t)$ 的递增序列输出训练结点列表 T' ,用于加速分类。

算法1 OPF 算法

输入:训练集 T 、原型集 $S_R \subset T$ 和原型集 $S_I \subset T$ 、描述子 (v, d) 。

输出:最佳路径森林 P (前驱映射表)、路径代价映射表 C 、根映射表 R 和训练结点序列列表 T' 。

辅助标记:优先级队列 Q 和成本变量 cst 。

```

1 for(每个结点  $s \in T \setminus S_R \cup S_I$ )
2   do 设置  $C(s) \leftarrow +\infty$ 
3 end
4 for(每个结点  $s \in S_R \cup S_I$ )
5   do 设置  $C(s) \leftarrow 0, P(s) \leftarrow nil$ ,
6      $R(s) \leftarrow s$ , 将  $s$  插入  $Q$ 
7 end
8 while( $Q$  非空)
9   do 从  $Q$  中移除最小代价结点  $s$ , 将  $s$  插入  $T'$ 。
10    for(每个结点  $t \in T, C(t) > C(s)$ )
11      计算  $cst \leftarrow \max\{C(s), d(s, t)\}$ 
12      if  $cst < C(t)$  then
13        if  $C(t) \neq +\infty$  then
14          从  $Q$  中移除结点  $t$ 
15        end

```

```

16       $P(t) \leftarrow s, R(t) \leftarrow R(s), C(t) \leftarrow cst$ 
17      将  $t$  插入  $Q$  中
18      end
19  end
20 end

```

根据函数 C 给定的最佳路径,趋向于选择最靠近的结点生成路径上的每一条链路。在相关或非相关类中,原型集 S_R 和 S_I 必须从最相近的样本中选取。这些原型集用于避免来自对立类的最佳路径造成的误分类。通过计算点集 T 构成的完全图中的最小生成树(minimum spanning tree, MST)得到原型集^[16]。对于 MST 中的每条弧 $arc(s, t)$, 如果 $\lambda(s) \neq \lambda(t)$, 则 s 和 t 被标记为原型。若 $\lambda(s)$ 是相关的, 则 $\lambda(t)$ 是非相关的, 则 s 插入 S_R , t 插入 S_I 。函数 $\lambda(\cdot)$ 为分类的相关性标记函数。

OPF 算法遵循动态规划方案。第 1 ~ 7 行是初始化代价、前驱和根轨迹图。迫使 $S_R \cup S_I$ 中的最佳路径开始, 在 Q 中插入根结点。在一个代价非增长序列中, 主循环计算从 $S_R \cup S_I$ 到每个结点 $s \in T$ 的最佳路径(第 8 ~ 20 行)。在每一次迭代中, 当移除 Q 中的最后一个结点后获得 P 中的最小代价路径 $C(t)$, 保存其在 T' 中的次序(第 9 行); 剩下的算法伪代码用于评估通过邻结点 $s(t \neq s)$ 的路径是否比当前以 t 为终点的路径代价更小, 依次更新 Q 、 $C(t)$ 和 $P(t)$ 。

在分类过程中, 每个样本 $t \in Z \setminus T$; 以 t 为终点的最佳路径可以很容易地通过查找训练结点 $s^* \in T$ 中最小值(由式(5) 计算得到) 来识别。

$$C(t) = \min_{\forall s \in T} \{ \max \{ C(s), d(s, t) \} \} \quad (5)$$

结点 s^* 是以 t 为终点的最佳路径的前驱映射表 $P(t)$ 中的点, 图像 t 通过 $\lambda(R(s^*))$ 分类。

上述算法中, T' 的作用是加速式(5) 的评估速度。当 $\max \{ C(s), d(s, t) \} < C(p)$, 且 T' 中结点 p 的位置优于 s 时, 则算法停止^[17]。

2.2 使用最佳路径森林分类的贪婪学习方法

算法 2 是使用 OPF 的贪婪主动学习方法, 以相关性递减顺序返回图像列表 $R \in Z$ 。

算法 2 GL-OPF 算法

输入: 待查询图像 q 、描述子 (v, d) 、每次迭代返回 N 幅图像、图像数据集 Z 。

输出: 以相关性递增排序的检索图像列表 $R \subset Z$ 。

辅助标记: 训练图像集 T , 基于 OPF 分类器的映射表 (R, P, C, T') , 原型集 $S_R \subset T$ 和原型集 $S_I \subset T$; 相关图像集 $Y \in Z \setminus T$; 每次迭代返回的 N 幅图像序

列表 $X \in Y$ 。

```

1  设置  $R \leftarrow \emptyset$  且  $T \leftarrow \emptyset$ 
2  计算列表  $X: d(q, t)$  递增的  $N$  幅最相近图像集
3  while(用户不满意)
4  do 用户标记相关(非相关) 图像, 生成训练集  $T \leftarrow T \cup X$ , 每幅图像都具有相关性标记  $\lambda(t)$ 
5  在相关图像集  $X$  中插入  $R$ 
6  计算  $T$  中原型集  $S_R$  和  $S_I$ , 并调用算法 1 执行  $(R, P, C, T') \leftarrow \text{OPF}(T, S_R, S_I, v, d)$ 
7  用  $(R, P, C, T')$  和  $\lambda(t)$  (用于  $t \in S_R \cup S_I$ ) 对图像  $Z \setminus T$  进行分类, 创建相关候选图像集  $Y$ 
8  计算新列表  $X$ : 在  $\bar{d}(t, S_R, S_I)$  递增序列中, 计算  $N$  幅最相近图像  $t \in Y$ 
9  end
10 返回  $R \leftarrow R \cup X$ 

```

其中: 第 1 行为列表 R 和训练集 T 被初始化为空集合。对于给定的查询图像 q 和图像数据集 Z , 首先按 $d(q, t)$ 递增序列, 返回包含 N 幅最接近图像 $t \in Z$ 的列表 X (第 2 行)。在主循环中执行学习过程(第 3 ~ 9 行)。第 4 行为用户标记 X 中的相关图像, 剩下的图像可认为是非相关的。该算法从本质上创建了一个训练集 $T \leftarrow T \cup X$, 该集合中的图像 $t \in T$ 都具有关联标签 $\lambda(t)$ 。 X 中的相关图像被插入到输出列表 R (第 5 行)。OPF 训练的计算过程见第 6 行: 生成相关原型集 $S_R \in T$ 和非相关原型集 $S_I \in T$; 得到 OPF 属性、前驱映射表 $P(t)$ 、代价 $C(t)$ 和结点 $t \in T$ 的根映射表 $R(t)$; 按最优代价 $C(t)$ 递增排序的训练图像集 T' 用于更快地分类。第 7 行为使用式(5) 对 $Z \setminus T$ 中的图像进行分类, 生成相关图像集 Y 。在归一化的平均距离 $\bar{d}(t, S_R, S_I)$ (即 t 与 2 个原型集间的距离, 如式(6) 所示) 递增序列中, 生成包含 N 幅最相近图像 $t \in Y$ 的新列表 X 。

$$\bar{d}(t, S_R, S_I) = \frac{\bar{d}(t, S_R)}{\bar{d}(t, S_R) + \bar{d}(t, S_I)} \quad (6)$$

式中:

$$\bar{d}(t, S_R) = \frac{1}{|S_R|} \sum_{i=1}^{n_R} d(S_{R_i}, t) \quad (7)$$

$$\bar{d}(t, S_I) = \frac{1}{|S_I|} \sum_{i=1}^{n_I} d(S_{I_i}, t) \quad (8)$$

式(7)中, S_{R_i} 表示第 i 个属于相关原型集 S_R 的图像, 其最大数量为 n_R ; 式(8)中, S_{I_i} 表示第 i 个属于非相关原型集 S_I 的图像, 其最大数量为 n_I 。重复主循环, 直至迭代结果令用户满意。随着序列 $\bar{d}(t, S_R, S_I)$ 递

增,生成最后的集 X ,此时得到的检索图像集 R 中的所有图像都是相关的。

3 实验与分析

3.1 实验数据集及对比方法

使用以下3个实验数据集:

1) Caltech101^[18]: 9 144 幅图像被分为 101 个类别,每类大概 40 ~ 800 幅图像,大部分类包含约 50 幅图像;

2) Corel^[4]: 来自 Corel GALLERY Magic-Stock Photo Library 2 的 3 906 幅图像子集,被预先分为如非洲人、大象、花等 85 个图像类,每类拥有 7 ~ 98 幅不同数量的照片;

3) Pascal^[19]: 使用该数据集子集,有 3 448 幅图片,分为 23 类。每一类都由不同数量的图片构成,包含 72 ~ 446 幅子图(感兴趣的区域可能来自同一幅图像)。

3 个数据集的图像大小为 256 像素 × 384 像素和 384 像素 × 256 像素 JPEG 格式,将所有图片统一缩放为 256 像素 × 256 像素。所有实验在 Intel 酷睿 i3 4170 处理器,主频 2.49 GHz,内存 4.0 GB Win7 操作系统的台式机上进行,采用 Matlab2011b 编程实现。图 3 为数据集中的一些图像示例。



图 3 数据集中的图像示例

Fig. 3 Image samples in data set

对比的方法有文献[5]提出的粒子群优化和相关反馈的混合方法(PSO-RF),文献[6]提出的改进二进制重力索引算法(IBGSA)和文献[7]提出的最佳森林分类方法(OPF),这些方法都容易实现,文献[5,7]源代码可以在 github 上找到。假设这些方法均使用相同的初始查询图片 q ,每次迭代返回 N 幅图像(每次迭代反馈相关图像数),取 $N = 30$,每次初始化查询 q ,都从列表 $X \subset Z$ 开始,该列表包含 N 幅按 $d(q, t)$ 递增顺序初始化图像,其中 $t \in Z$ 。不断重复主循环(算法2),直到用户满意时终止。但是为了比较方便,用户可以修改迭代次数。本文方法在主循环迭代 I 次后终止,在 Corel 数据集中,当 I 等于 3、5 和 8 次迭代时,评价各方法;在其他数据集中,当 I

等于 3 和 8 时,评价各方法。

3.2 评估函数

CBIR 算法的查准率(精度)和查全率(召回率)是检索中广泛使用的性能度量,也是 CBIR 算法评估最常见的标准^[3],查准率 Pr 定义为正确检索图像数与检索到的图像总数的比值,定义如下:

$$Pr = \frac{Ac}{Re} \quad (9)$$

式中, Ac 为准确检索图像数, Re 为检索到的图像总数。

查全率 $Recall$ 是检索到的图像总数与被检索图像总数的比值,定义如下:

$$Recall = \frac{Re}{Tol} \quad (10)$$

式中, Tol 表示数据集中的被检索图像总数。通常情况下,查全率很高时,查准率很低;而当查准率很低时,查全率很高。查全率和查准率具有互逆关系。

3.3 实验结果

精度、查全率曲线用于测量在第 I 次迭代后,返回集 R 中所用方法的效果。每个图像数据集,用于模拟用户的行为,使用每幅图像 $q \in Z$ 作为初始化查询点,并标记每次迭代后返回 X 中的 N 幅图像的相关点,即来自相同的查询类的图像。

图 4 是 Corel 数据集中,当迭代次数为 $I = 3, 5, 8$ 时,各方法的平均精度-查全率曲线。由图 4 可以观察到,在任意次的迭代中,本文方法优于其他方法。OPF 和 GL-OPF 从一开始就表现良好,但是,GL-OPF 比 OPF 演化得更快,因为 GL-OPF 在此过程的学习需要迭代的次数更少。

从图 4(a)可以看出,在 3 次迭代中,查全率在(0.1, 0.9)区间范围,GL-OPF 明显优于其他方法。对比图 4(b)、(c)可以看出,随着迭代次数的增加,IBG-SA 似乎在迭代 3 次就表现出“饱和”状态,其他 3 种方法随着迭代增加表现出查询精度增加的趋势,其中,GL-OPF 表现最优,贪婪算法的“本质”得到体现,排序时使用的相关性预测,已排除了大部分错误。

图 5 和 6 是在 Corel 数据集中使用本文方法和 OPF 方法检索的 2 个实例。实例 1 中基于 OPF 的 CBIR 检索结果中有 26 幅与查询样本相关,查全率为 0.26,而本文方法的检索结果中有 41 幅与查询样本相关,查全率为 0.41;实例 2 中基于 OPF 的 CBIR 检索结果中有 33 幅与查询样本相关,查全率为 0.33,本文方法的检索结果中有 43 幅相关图像,查全率为 0.43。因此,本文方法(GL-OPF)明显提

升了 OPF 的检索能力。

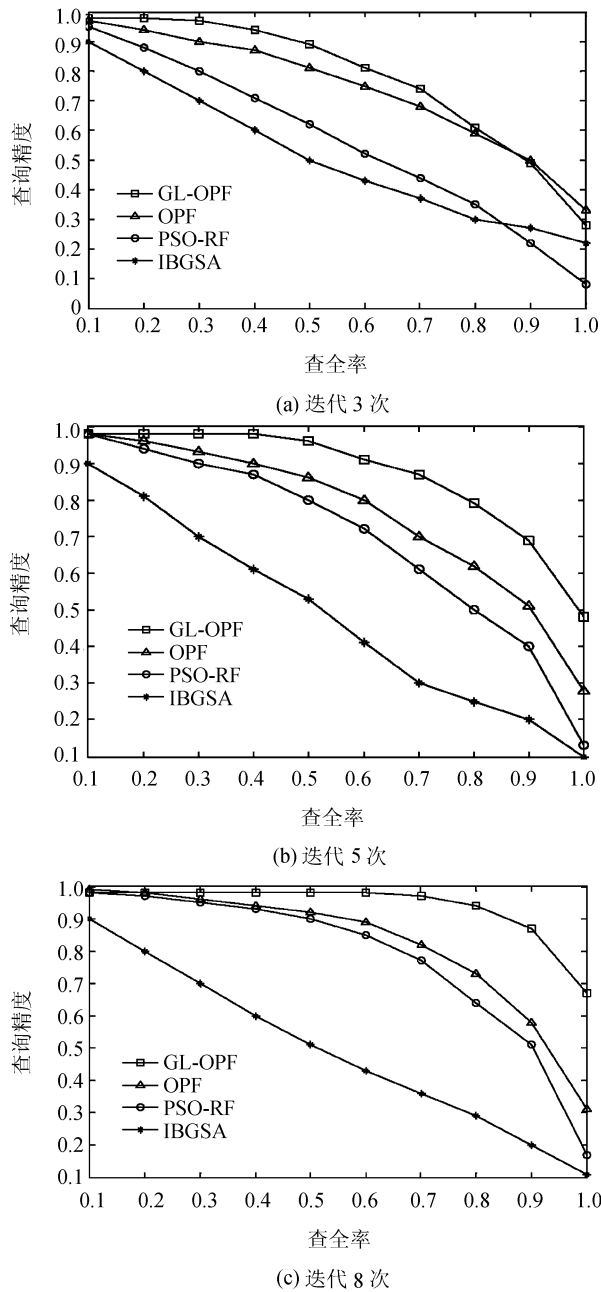


图 4 Corel 数据集中各方法的平均精度 - 查全率曲线
Fig.4 Average precision-recall curves of each method in Corel data set

图 7 和 8 给出了图像数据集 Caltech101 和 Pascal 中的平均精度 - 查全率曲线,这里只进行 3 次和 8 次迭代。

由图 7、8 可以观察到:Caltech101 数据集中,本文方法的优势更大,而在 Pascal 数据集中,3 次迭代时 IBGSA 优于 PSO-RF。实际上,PSO-RF 通常需要更多的迭代次数来提高性能。而本文方法可以有效地评估相关和非相关图像集。



图 5 图像检索实例 1
Fig.5 Example 1 of image retrieval

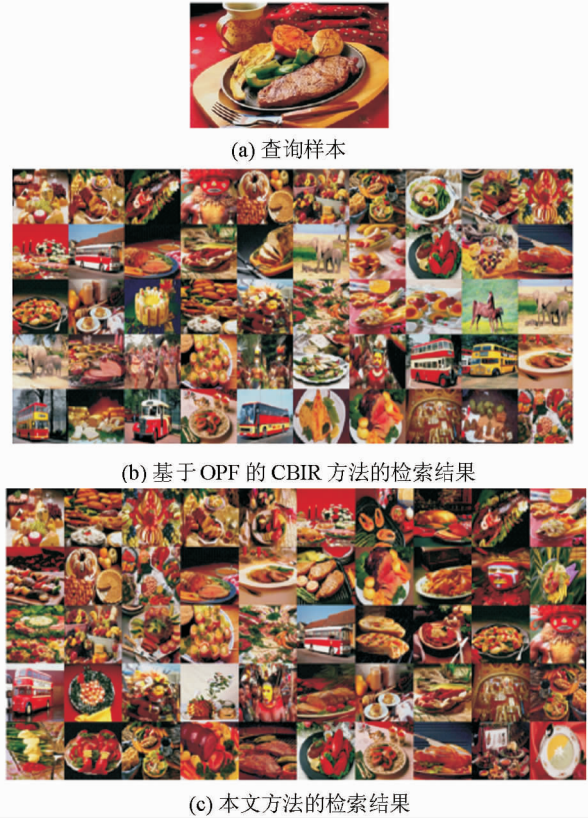


图 6 图像检索实例 2
Fig.6 Example 2 of image retrieval

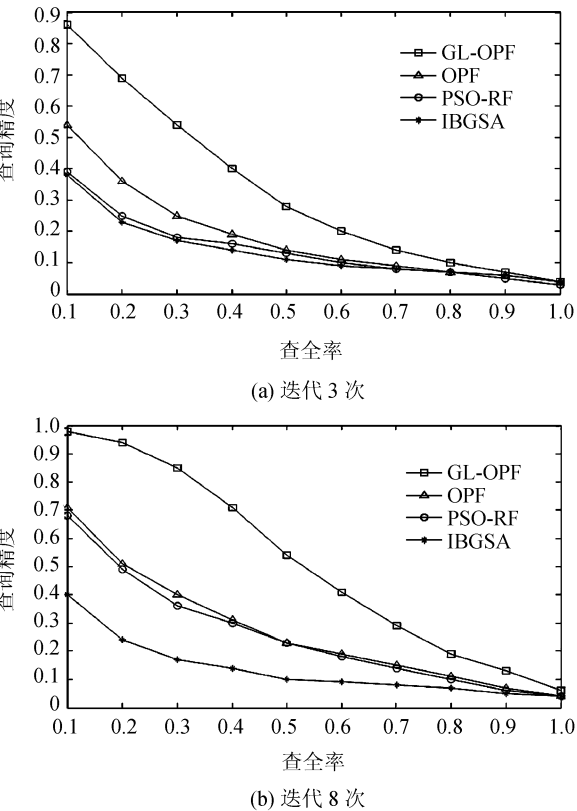


图7 Caltech101 数据集中各方法的平均精度 - 查全率曲线
Fig.7 Average precision-recall curves of each method in Caltech101 data set

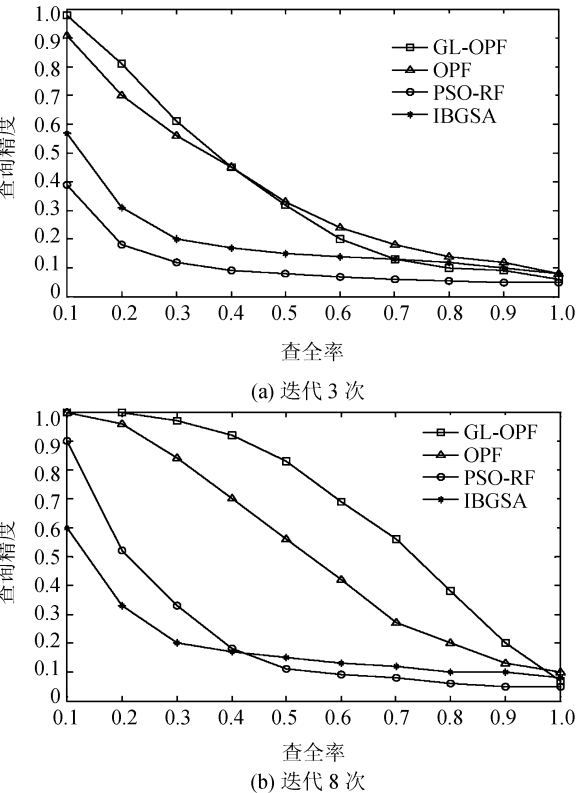


图8 Pascal 数据集中各方法的平均精度 - 查全率曲线
Fig.8 Average precision-recall curves of each method in Pascal data set

3.4 运行时间比较

为了更好地体现本文方法在运行效率方面的优越性,比较了各方法的运行时间。表1 为不同数据集中8次迭代的总运行时间。表2 为不同数据集中8次迭代后查询的平均时间。由表1、2 可以看出: PSO-RF 的运行时间最多,这可能是由于 PSO 有的时候很难找到满足要求的解。本文方法的实验结果比较令人满意。在 Corel 数据集中,8 次迭代后 GL-OPF 和 OPF 的平均代价为 0.06 s,而且其每次查询的平均运行时间与 OPF 是一样的。这充分说明了本文方法(GL-OPF)在运行效率方面的优势。

表1 8次迭代的总运行时间

s				
数据集	IBGSA	PSO-RF	OPF	GL-OPF
Caltech101	169.5	199.8	178.0	178.0
Corel	54.7	112.3	37.4	37.5
PASCAL	25.7	604.6	35.7	35.9

表2 8次迭代后查询的平均运行时间

s				
数据集	IBGSA	PSO-RF	OPF	GL-OPF
Caltech101	0.14	5.93	0.13	0.13
Corel	0.41	6.46	0.06	0.06
PASCAL	0.05	6.17	0.07	0.07

4 结论与展望

提出了基于最佳路径森林分类器的贪婪主动学习方法 GL-OPF,该方法是基于 OPF 方法的有效改进。本文实验包含合理数量的用户交互,设置了少量的响应图像和迭代次数。实验结果验证了本文方法的有效性。此外,运行时间与原 OPF 基本相同。

未来的研究工作为:如何使用关联多描述子对本文方法进行改进和扩展,关联描述子的相关技术如贝叶斯框架、遗传算法等智能算法等。

参考文献:

[1] Rashedi E, Nezamabadi-Pour H, Saryazdi S. Long term learning in image retrieval systems using case based reasoning[J]. Engineering Applications of Artificial Intelligence, 2014, 35(10): 26-37.

[2] Bugatti P H, Kaster D S, Ponciano-Silva M, et al. PRoS-Per: Perceptual similarity queries in medical CBIR systems through user profiles[J]. Computers in Biology & Medi-

- cine, 2014, 45(2): 8 – 19.
- [3] Li Zongmin, Tang Zhihui. Books image retrieval based on grid space frame[J]. Journal of Image and Graphics, 2013, 18(3): 325 – 329. [李宗民, 唐志辉. 九宫格空间框架的图书图像检索[J]. 中国图象图形学报, 2013, 18(3): 325 – 329.]
- [4] Wang Mingsan. Research on key techniques of content based image retrieval[D]. Chengdu: University of Electronic Science and Technology of China, 2014. [王明三. 基于内容的图像检索关键技术研究[D]. 成都: 电子科技大学, 2014.]
- [5] Xu Xiangli, Zhang Libiao, Liu Xiangdong, et al. Image retrieval relevance feedback algorithm based on particle swarm optimization[J]. Acta Electronica Sinica, 2010, 38(8): 1935 – 1940. [许相莉, 张利彪, 刘向东, 等. 基于粒子群的图像检索相关反馈算法[J]. 电子学报, 2010, 38(8): 1935 – 1940.]
- [6] Rashedi E, Nezamabadi H. Feature subset selection using improved binary gravitational search algorithm[J]. Journal of Intelligent and Fuzzy Systems, 2014, 26(3): 1211 – 1221.
- [7] Silva A D, Xavier A, Magalhães L P. A new CBIR approach based on relevance feedback and optimum path forest classification[J]. Journal of WSCG, 2010, 27(3): 73 – 80.
- [8] Fang Hong, Yang Hairong. Greedy algorithm and compressed sensing theory[J]. Journal of Automation, 2011, 37(12): 1413 – 1421. [方红, 杨海蓉. 贪婪算法与压缩感知理论[J]. 自动化学报, 2011, 37(12): 1413 – 1421.]
- [9] Tong S, Chang E Y. Support vector machine active learning for image retrieval[C]//Proceedings of the Ninth ACM International Conference on Multimedia. Ottawa: IEEE, 2001: 107 – 118.
- [10] Wang Chuanchuan. Research on ordering of image retrieval in network community[D]. Jinan: Shandong University, 2011. [王川川. 网络社区图像检索中的排序研究[D]. 济南: 山东大学, 2011.]
- [11] Choudhary R, Raina N, Chaudhary N, et al. An integrated approach to content based image retrieval[C]//Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference). Greater Noida: IEEE, 2014: 2404 – 2410.
- [12] Yi Tangtang, Huang Lihong. Improved relevance feedback algorithm based on nearest neighbor approach in CBIR[J]. Application Research of Computers, 2015, 32(8): 2326 – 2330. [易唐唐, 黄立宏. CBIR 中一种基于最近邻的改进相关反馈算法[J]. 计算机应用研究, 2015, 32(8): 2326 – 2330.]
- [13] Jian M, Guo H, Liu L. Texture image classification using visual perceptual texture features and Gabor wavelet features[J]. Journal of Computers, 2009, 4(8): 763 – 770.
- [14] Liu D, Hua K A, Vn K, et al. Fast query point movement techniques for large CBIR systems[J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(5): 729 – 743.
- [15] Datta R, Joshi D, Li J, et al. Image retrieval: Ideas, influences, and trends of the new age[J]. ACM Transactions on Computing Surveys, 2008, 40(2): 1 – 66.
- [16] Zhi Lijia, Zhang Shaomin, Zhao Dazhe, et al. Minimum spanning tree fusing multi-feature point information for medical image registration[J]. Journal of Computer Research and Development, 2011, 48(3): 501 – 507. [支力佳, 张少敏, 赵大哲, 等. 融合多种特征点信息的最小生成树医学图像配准[J]. 计算机研究与发展, 2011, 48(3): 501 – 507.]
- [17] Cormen T, Leiserson C, Rivest R. 算法导论[M]. 4 版. 北京: 人民邮电出版社, 2012.
- [18] Papa J P, Cappabianco F A M, Falcao A X. Optimizing optimum-path forest classification for huge datasets[C]//2010 International Conference on Pattern Recognition. Istanbul: IEEE, 2010: 4162 – 4165.
- [19] Everingham M, van Gool L, Williams C K I, et al. The Pascal visual object classes (VOC) challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303 – 338.