

分组排序多特征融合的图像检索方法

刘胜蓝 冯 林 孙木鑫 刘 洋

(大连理工大学电子信息与电气工程学部 辽宁大连 116024)

(大连理工大学创新实验学院 辽宁大连 116024)

(liusl@mail.dlut.edu.cn)

Group and Rank Fusion for Accurate Image Retrieval

Liu Shenglan, Feng Lin, Sun Muxin, and Liu Yang

(Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, Liaoning 116024)

(School of Innovation Experiment, Dalian University of Technology, Dalian, Liaoning 116024)

Abstract Single feature is not discriminative to describe informational content of an image, which is always a shortcoming in traditional image retrieval task. It can be seen that one image can be described by different but complemented features. So multi-feature fusion ranking methods should be further researched to improve the ranking list of query in image retrieval. Most existing multi-feature fusion methods only focus on the nearest neighbors in image retrieval. When the ranking result of each neighbor graph is poor, it is hard to get ideal image retrieval result after graph fusion. In order to solve the problem, this paper proposes a novel multi-feature fusion method for image retrieval—GRF (group ranking fusion). The proposed method divides similar images of a data set into the same group, and improves the retrieval result of neighbor graph through similar images group. The GRF method expands the fusion scope in the premise of guaranteeing retrieval precision premise. At last, the experimental results on three standard data sets demonstrate that GRF can effectively utilize multi-feature graph to improve the performance of image retrieval.

Key words multi-feature fusion; content-bases image retrieval; minimizing normalized cut; graph learning; retrieval re-rank

摘 要 在图像检索中,多特征图融合方法大多仅对最近邻域进行融合.当每个特征的近邻图排序结果较差时,融合后的新图难以得到理想的检索效果.为了解决该问题,提出一种新的多特征图融合图像检索方法——分组排序融合(group ranking fusion, GRF),该方法将数据集中的相似图片划分为图片组,利用相似图片组对近邻图的检索结果进行改进,在保持精度的前提下扩充了融合范围.最后,在3个标准数据集上的实验结果表明:多特征融合方法能够有效地利用多特征图提高图像检索效果.

关键词 多特征融合;基于内容的图像检索;规范最小割;图学习;检索重排

中图法分类号 TP391.41

收稿日期:2015-10-29;修回日期:2016-07-05

基金项目:国家自然科学基金项目(61602082,61672130,61370200);中国博士后科学基金面上资助基金项目(2015M581331)

This work was supported by the National Natural Science Foundation of China (61602082, 61672130, 61370200) and the General Program of the Postdoctoral Science Foundation of China (2015M581331).

近年来随着互联网和移动终端上图像数据的急剧增长,如何快速有效地从图像库中检索用户需要的图像已成为计算机视觉领域的热点问题.图像检索^[1]可分为3类:基于文本^[2]、图像内容^[3]和图像语义^[4]的图像检索.基于文本和基于语义的检索效果尚不理想;而基于内容的图像检索(CBIR)技术近些年取得了很多成果,其检索效果也较为理想.因此,CBIR成为图像检索的主流方式.CBIR主要研究图像特征提取和图像排序两方面的工作,图像特征提取的方法和图像排序都对图像检索的效果有很大的影响,好的图像特征配合与之相适应的图像排序方法才能得到理想的图像检索效果.

在图像排序方面,已有很多优秀的成果.如刘广海等人^[5]发现1范数排序比欧氏距离排序效果更好.然而1范数排序存在排名方式独立的问题.因此,在后续的研究中,图学习在图像检索中起着至关重要的作用.He等人^[6]将流形排序(manifold ranking,MR)应用于图像检索,取得了较好效果.He等人^[6]进一步提出了广义的MR,利用K-NN初始化得分向量,该方法在数据分布不复杂的情况下可以获得比MR更优的结果,但容易因K-NN选取不当造成误差扩散.王斌等人^[7]将图结构中的K-NN利用样本节点的度和权值的最小化改进为

K-RNN,提升图像检索的效果.在图结构方面,Huang等人^[8]利用超图实现MR,节点之间的权值由概率决定,由于不同节点对有多个权值,所以在图像检索中可以取得比较好的排序结果.除了超图,多图也是样本点间多权值的一种实现手段.Zhao等人^[9]利用图像的不同特征构建多图加强样本点之间的关系实现图像检索,Zhang等人^[10]利用Graph PageRank和Graph Density方法将多图进行融合,实现图像的排序.

Graph Density^[10]方法为目前主流的图学习方法之一.该方法通过融合差异性大的特征组来提高检索结果的准确性.但Graph Density方法在多特征图融合上依然存在限制.Graph Density方法要求至少一个特征对待检索图片表现良好,才能取得较理想的检索结果如图1^[10]所示.若单特征的检索结果不理想,则Graph Density方法难得得到理想的效果.

例如,在局部特征和全局特征都没有得到理想检索结果的情况下,如图2所示的Graph Density方法的检索结果不会得到改善.在图2(a)中,对待检图片而言,不同特征的检索结果均不理想,导致Graph Density方法没有获取到充分的信息,改善待检图片的检索结果如图2(b)所示.

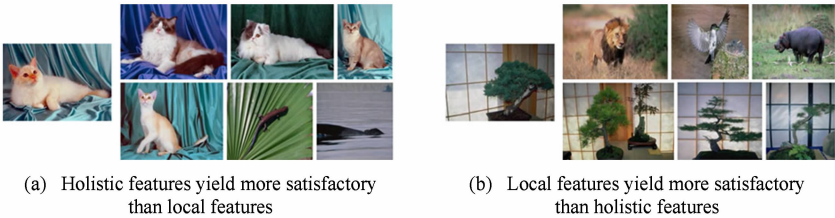


Fig. 1 The performance of holistic features and local features on the same data-set
图1 全局特征和局部特征在同一个数据集上的检索结果

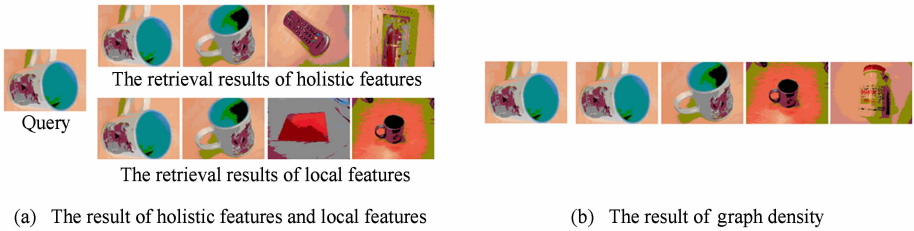


Fig. 2 Retrieval result of one query images in UK-bench
图2 UK-bench数据集中一张图片的检索结果

为了获取足够的信息,本文将待检图片的检索结果视为新的待检图片组,利用该图片组的检索结果获取原始待检图片的信息,将检索结果的范围扩

大到数据集中与检索图片相似的图片组,若图片组中能够获取足够的信息,就可以提升检索效果.

为了从相似图片组中获取足够的信息解决 Graph

Density 方法对特征要求过高的问题, 本文提出分组排序的图融合 (group ranking fusion, GRF) 方法. 该方法主要分为 4 个步骤:

- 1) 将图像检索数据集, 获取特征的排序结果 (第 2 节), 例如 HSV^[11], BOW^[12];
- 2) 建立权重图 (第 3 节), 权重图中节点为数据集中的图片, 其中节点之间权重和图像之间的相似度成正比, 对于不同特征的权重图采用图融合 (3.1 节) 的方法;

3) 将权重图中节点分为 n 个独立的集合 (集合大小等于图像检索输出图像的个数), n 个大小相同的集合满足规范最小割^[13];

4) 将最近邻域中与检索图片不在同一个图片组中的近邻删除, 余下的近邻就是本文方法的检索结果, 具体过程如图 3 所示.

权重图的规范最小割是一个 NP 完全问题^[13], 为了求解规范最小割, 本文 (3.2 节) 设计一种求解规范最小割极值的随机算法.

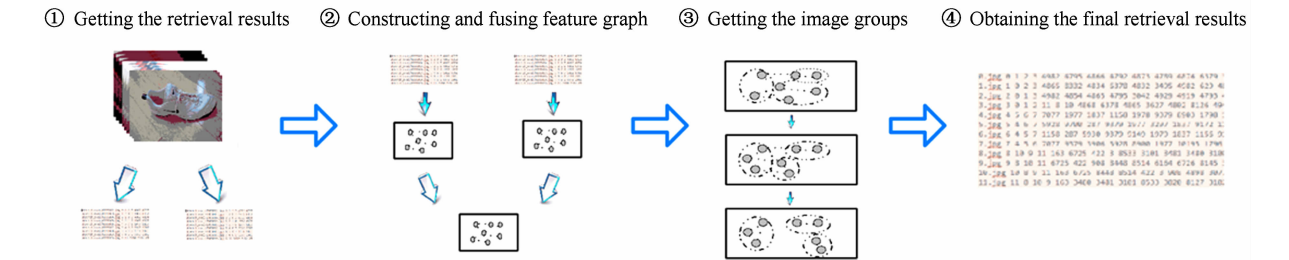


Fig. 3 Group and rank fusion method
图 3 分组排序的图融合方法

1 图像特征提取

图像特征提取很大程度上决定了图像检索的效果, 也是多特征图融合方法的重要步骤. 本节主要简单介绍实验中 GRF 使用到的图像特征提取方法.

1.1 全局图像特征

全局特征主要是通过图片整体特性来描述图片所携带的信息. 例如统计颜色信息的颜色直方图 (HSV^[11])、统计纹理信息的局部二值模式 (LBP) 等.

HSV 为统计颜色信息直方图, 它能够描述一幅图像中颜色的全局分布, 即不同色彩在整幅图像中所占的比例, 特别适用于描述那些难以自动分割的图像和不需要考虑物体空间位置的图像, 但是它无法描述图像中颜色的局部分布及每种色彩所处的空间位置, 即无法描述图像中的某一具体的对象或物体.

LBP 为统计纹理信息的局部二值模式, 它也描述了图像或图像区域所对应景物的表面性质, 但是它无法完全反映高层次图像内容.

除此之外还有许多全局图像方法, 例如基于局部敏感度 Hash (LSH^[14]), 以及在 LSH 上改进马氏 (Mahalanobis) 距离^[15] 等. 本文的实验主要使用 HSV 特征.

1.2 局部图像特征

局部图像特征通过描述图片的区域来描述图片

特征, 其中图片区域的描述子通常满足不变性 (鲁棒性) 和可区分性, 例如尺度不变性特征 (SIFT^[16-17], SURF^[18]).

SIFT 是应用广泛的局部特征, 它对尺度、旋转以及一定视角和光照变化等图像变化都具有不变性, 并且 SIFT 具有很强的可区分性.

SURF 是对 SIFT 的改进版本, 它利用 Haar 小波来近似 SIFT 方法中的梯度操作, 同时利用积分图技术进行快速计算, SURF 的速度是 SIFT 的 3~7 倍, 大部分情况下它和 SIFT 的性能相当.

近期很多二值特征描述子被提出, 例如 ORB^[19], BRISK^[20], FREAK^[21] 等. 局部特征描述子结合词汇树^[22] (VOC) 或词袋^[12] (BOW) 转化为图像的局部特征. 本文的实验主要使用 SIFT 特征 + BOW 方法以及 CNN 特征.

2 特征图构建

要在排序阶段实现不同特征的图融合, 首先需要构建一个好的特征图. 一般的想法是构建出来图的边的权重是图片之间为相似图片的概率, 但它是不可计算的. 同本文类似的替代方案是文献 [10, 23] 等, 本节将会介绍 GRF 算法特征图的构建方式.

为了方便说明本文特征图的构建方法, 定义 q 为当前要创建出边的图片, i 为数据集中将要创建入边的图像, d 和 i' 为数据集中的图片, $N_k(i) =$

$\{n_1^i, n_2^i, \dots, n_k^i\}$ 为检索图像 i 出现的前 k 个候选图片组成的集合. 对于任意的检索结果 (如 HSV^[11], VOC^[22] 等), 创建有向图 $G=(V, E, \omega)$ (V 为点集, E 为边集, ω 为边权函数). 为了求出有向图 $G=(V, E, \omega)$ 的边权函数 ω , 需要定义 i, i' 的杰卡德 (Jaccard) 相似系数^[24]:

$$J(i, i') = \frac{|N_k(i) \cap N_k(i')|}{|N_k(i) \cup N_k(i')|}. \quad (1)$$

边权函数 ω 定义为

$$\omega'(q, i, i') = \begin{cases} 1, & (J(i, i') > \alpha) \wedge (i' \in N_k(i)), \\ 0, & \text{else,} \end{cases} \quad (2)$$

$$\omega(q, i) = \begin{cases} \sum_{i' \in V} \omega'(q, i, i'), & i \in N_k(q), \\ 0, & \text{else,} \end{cases} \quad (3)$$

其中, α 为常量 ($\alpha \in [0.1, 0.3]$). 基于式 (2) (3), 对于任意的检索结果可以构建出有向图 $G=(V, E, \omega)$. 由式 (2) (3) 可以得出, 节点 i 和 i' ($i' \in N_k(i)$) 的 $J(i, i')$ 大于常数 α 的数量越多, 则 $\omega(q, i)$ 值越大.

为了利用相似图片组的检索信息, 本文将图像检索数据集划分为 n 个集合 $S=\{S_1, S_2, \dots, S_n\}$, 集合 S_i 的大小为 k (一般为检索出的相似图像的个数), 且集合之间满足关系:

$$(\forall i \forall j) S_i \cap S_j = \emptyset, \quad (4)$$

$$\bigcup_{1 \leq i \leq n} S_i = V, \quad (5)$$

$$|S_i| = k. \quad (6)$$

图像检索数据集的图片个数若不满足 n 种图片每种图片的个数恰好为 k 个, 则随机选择 $k \cdot n - |V|$ 张图片进行一次复制, 这样可保证每个集合得到的图片数目相同. 同时, 定义 $c(i)$ 为图片 i 所在集合的编号. 为了求得最优的划分集合 $S=\{S_1, S_2, \dots, S_n\}$, 定义如下函数.

定义 1. 函数 $f(i, i')$ 为

$$f(i, i') = \begin{cases} 1, & c(i) = c(i'), \\ 0, & \text{else.} \end{cases}$$

由定义 1 可知, 对于数据集 V 的划分 $S=\{S_1, S_2, \dots, S_n\}$ 的准确率 (*precision*) 期望为

$$\begin{aligned} E(\text{precision}) &= \\ E\left(\sum_{1 \leq j \leq n} \sum_{i \in S_j \wedge i' \in S_j} P(f(i, i'))\right) &= \\ E\left(\sum_{i \in V \wedge i' \in V} P(f(i, i'))\right) &- \\ \sum_{1 \leq j \leq n} \sum_{i \in S_j \wedge i' \notin S_j} P(f(i, i')) &= \end{aligned}$$

$$\begin{aligned} &\sum_{i \in V \wedge i' \in V} E(P(f(i, i'))) - \\ &\sum_{1 \leq j \leq n} \sum_{i \in S_j \wedge i' \notin S_j} E(P(f(i, i'))). \end{aligned} \quad (7)$$

其中, $\sum_{i \in V \wedge i' \in V} E(P(f(i, i')))$ 是基于检索数据集的常量与集合划分无关. 因此, 对于给定的图像检索集合, 本文的目标是寻找使 $E(\text{precision})$ 值最大的集合划分:

$$S = \arg \max_s E(\text{precision}) =$$

$$\arg \min_s \left(\sum_{1 \leq j \leq n} \sum_{i \in S_j \wedge i' \notin S_j} E(P(f(i, i'))) \right). \quad (8)$$

从式 (8) 可以得出, 求任意的划分集合 $S=\{S_1, S_2, \dots, S_n\}$ 对应的 $E(\text{precision})$ 关键是 $P(f(i, i')=1)$, 由于 $P(f(i, i')=1)$ 的准确值是不可知的. 因此, 如果 $\omega(q, i)$ 越大, 则 $P(f(i, i')=1)$ 越大, 此时, 可用 $\omega(q, i)$ 近似代替 $P(f(i, i')=1)$, 求出在不同划分集合条件下 $E(\text{precision})$ 的大小关系.

3 分组排序的图融合

分组排序的图融合 (GRF) 方法主要包括: 多特征图融合方法以及分组排序方法. 在本节中, 将分别对以上内容进行详细介绍.

3.1 图融合

为了获取图像特征互补信息, 提高图像检索的准确率, 需要设计多种图像特征融合方法. 现有的多特征融合方法主要有: 在特征提取阶段的融合方法和在排序阶段的融合方法. 在特征提取阶段的融合方法在效率和可靠性方面表现一直不佳, 尤其是融合特征越多效率越难以保证. 对于要求高效的图像检索, 在特征提取阶段的融合方法存在严重缺陷. 因此, 在排序阶段的融合方法成为图像检索中多特征融合的主要研究方向. 目前在排序阶段的融合方法主要是 Dwork 等人^[25] 提出的图融合方法, 但该方法仍然存在一定的限制. 为了避免上述多特征融合出现的问题, 本文进行如下融合方式, 本文方法同文献^[10, 26-27]相似.

对不同特征计算出来的图像候选集合可以获得不同的特征图 $G^m=(V^m, E^m, \omega^m)$, 融合后的特征图记为 $G=(V, E, \omega)$, 它满足: 1) $V = \bigcup V^m$; 2) $E = \bigcup E^m$; 3) $\omega(i, i') = \sum_m \omega^m(i, i')$. 图 4 说明了本文多特征图融合的方式. 由 GRF 算法在不同特征图下的融合方式可知: 融合之后的图 $G=(V, E, \omega)$ 中节点 i 和 i' ($i' \in N_k(i)$) 的 $J(i, i')$ 大于常数 α 的数量越多, 则 $\omega(q, i)$ 值越大.

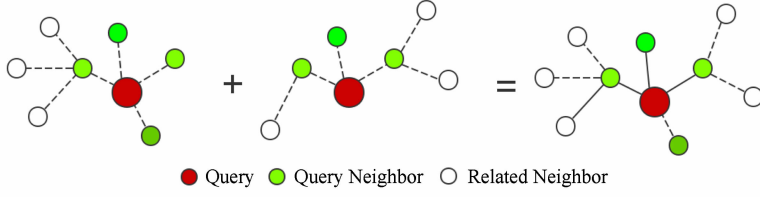


Fig. 4 Fusion of two graphs

图 4 2 种不同的 $G^m = (V^m, E^m, \omega^m)$ 融合成 $G = (V, E, \omega)$

3.2 分组排序方法

为了求解较优的集合划分 $S = \{S_1, S_2, \dots, S_n\}$ (求解规范最小割为 NP 完全问题), 本文设计一种简单的优化方法. 使 $E(\text{precision})$ 在 $S = \{S_1, S_2, \dots, S_n\}$ 下达到较优值. 每次迭代寻找 2 个不同集合的点, 如果这 2 个点交换所在集合之后, 新的集合划分 $S' = \{S'_1, S'_2, \dots, S'_n\}$ 的 $E'(\text{precision})$ 增大则交换 2 个点.

通过上述优化方法, 可保证集合划分 $S = \{S_1, S_2, \dots, S_n\}$ 的 $E(\text{precision})$ 逐渐增大, 当不存在交换 2 个点能够使 $E(\text{precision})$ 增大时, 则认为集合划分 $S = \{S_1, S_2, \dots, S_n\}$ 的 $E(\text{precision})$ 达到极值点. 迭代的具体方式如下:

算法 1. 主要迭代过程.

输入: 结构图 $G = (V, E, \omega)$;

输出: 划分集合 $S = \{S_1, S_2, \dots, S_n\}$.

任意给定初始集合 $S^0 = \{S_1^0, S_2^0, \dots, S_n^0\}$;

While True

 If $(\exists i, \exists i') \nabla E(\text{precision}) > 0$

 Swap (i, i') ;

 Else Return;

End If

End While

Return

不难发现, 每次迭代需要计算集合的总权值的变化情况. 总权值的计算速度直接影响算法的运行效率, 对于给定节点 i 和 i' , 由式(8)可得:

$$\begin{aligned}
 E(\text{precision}) &= \sum_{u \in V, v \in V} E(P(f(u, v))) - \\
 &\sum_{1 \leq j \leq n} \sum_{u \in S_j \wedge v \notin S_j} E(P(f(u, v))) = \\
 &\sum_{u \in V, v \in V} P(f(u, v) = 1) - \\
 &\sum_{1 \leq j \leq n} \sum_{u \in S_j - \{i, i'\}} \sum_{v \notin S_j - \{i, i'\}} P(f(u, v) = 1) + \\
 &\sum_{(i, u) \in E \wedge \exists j (i \in S_j \wedge u \notin S_j)} P(f(i, u) = 1) + \\
 &\sum_{(i', u) \in E \wedge \exists j (i' \in S_j \wedge u \notin S_j)} P(f(i', u) = 1). \quad (9)
 \end{aligned}$$

同理, 可得交换之后:

$$\begin{aligned}
 E'(\text{precision}) &= \sum_{u \in V, v \in V} E(P(f(u, v))) - \\
 &\sum_{1 \leq j \leq n} \sum_{u \in S_j \wedge v \notin S_j} E(P(f(u, v))) = \\
 &\sum_{u \in V, v \in V} P(f(u, v) = 1) - \\
 &\sum_{1 \leq j \leq n} \sum_{u \in S_j - \{i, i'\}} \sum_{v \notin S_j - \{i, i'\}} P(f(u, v) = 1) + \\
 &\sum_{(i, u) \in E \wedge \exists j (i \in S_j \wedge u \notin S_j)} P(f(i, u) = 1) + \\
 &\sum_{(i', u) \in E \wedge \exists j (i' \in S_j \wedge u \notin S_j)} P(f(i', u) = 1), \quad (10)
 \end{aligned}$$

由式(9)(10)可知交换之后:

$$\begin{aligned}
 \nabla E(\text{precision}) &= \\
 E(\text{precision}) - E'(\text{precision}) &= \\
 &\sum_{(i, u) \in E \wedge \exists j (i \in S_j \wedge u \notin S_j)} P(f(i, u) = 1) + \\
 &\sum_{(i', u) \in E \wedge \exists j (i' \in S_j \wedge u \notin S_j)} P(f(i', u) = 1) - \\
 &\sum_{(i, u) \in E \wedge \exists j (i \in S_j \wedge u \notin S_j)} P(f(i, u) = 1) - \\
 &\sum_{(i', u) \in E \wedge \exists j (i' \in S_j \wedge u \notin S_j)} P(f(i', u) = 1), \quad (11)
 \end{aligned}$$

由式(11)可以得出 $\nabla E(\text{precision})$ 的计算方法.

算法 2. $\nabla E(\text{precision})$.

输入: $i, i' (i, i' \in V)$;

输出: d .

$d = 0$;

For $(i, v) \in E$

 If $v \in S_j \wedge i \in S_j$

$d := d + \omega(i, v)$;

 Else If $v \in S_j \wedge i' \in S_j$

$d := d - \omega(i, v)$;

 End If

End For

For $(i', v) \in E$

 If $v \in S_j \wedge i \in S_j$

$d := d - \omega(i', v)$;

 Else If $v \in S_j \wedge i' \in S_j$

$d := d + \omega(i', v)$;

```
End If
End For
Return
```

本文已给出最大化 $E(\textit{precision})$ 的迭代方法, 并且可以通过本文方法对任何候选集合进行优化, 进而得到更加合理的候选图像集合.

4 实验结果分析

本节首先说明数据集和实验的方法, 然后详细说明在每一个数据集上的实验结果, 对每个数据集的结果进行分析.

4.1 数据集

本文评估采用 3 个标准数据集: UK-bench, Corel-10K, Corel-1K. 其中, UK-bench 具有图片旋转和尺度变化、图片类别多和数据集合相对较大等特点; Corel-10K 具有每类图片个数多和数据集合相对较大等特点; Corel-1K 具有每类图片个数多和数据集合大相对较小等特点. 通过以上数据集, 验证本文算法的有效性. 数据集详细参数如表 1 所示:

Table 1 Attributes of Experimental Data Set
表 1 实验数据集的属性

Database	Image Size	# of Class	# of Each Class	Total Image
UK-bench	640×480	2 550	4	10 200
Corel-1K	384×256	10	100	1 000
Corel-10K	Vary	100	100	10 000

UK-bench^[28] 包含 2 550 个不同的数据种类的图片, 每类图像包含 4 张相似的图片. 10 200 张图片同时包含数据集合和询问集合. 每张图片返回 4 张候选图片, 采用 N-S(改成斜体)评估.

Corel-1K^[29] 包含 10 个不同的数据种类的图片, 每类图像包含 100 张图片. 1 000 张图片包含数据集合和询问集合. 每张图片寻找 20 张候选图片, 并评估候选图片的准确率.

Corel-10K^[29] 包含 100 个不同的数据种类的图片, 每类图像包含 100 张图片. 1 000 张图片包含数据集合和询问集合. 每张图片寻找 12 张候选图片, 并评估候选图片的准确率.

4.2 特征选取及融合方法

基础的图像检索方式包括: VOC, HSV, BOW 等. 本文将会应用单特征优化方法对 G^{VOC} , G^{HSV} , G^{BOW} 进行优化, 并应用图融合方法对 G^{VOC} , G^{HSV} 和 G^{BOW} 进行融合优化.

1) VOC. 本文采用基于图像检索的词汇树, 每张图片使用 VLFeat-library 提取出 2500 SIFT 特征. 词汇树包含 7 层, 每层 10 个分支, 在 image-net 上进行训练, 最后应用到 UK-bench 数据集上.

2) HSV. 本文采用对于检索图像的每个像素点计算出 11 维向量: 黑、蓝、棕、灰、绿、橙、粉、紫、红、白、黄. 对于检索图像, 本文以平均向量为颜色描述符作为图片的描述向量, 最后应用到 UK-bench, Corel-1K, Corel-10K 数据集上.

3) BOW. 本文采用基于图像检索的词袋, 每张图片使用 VLFeat-library 提取出密集 SIFT 特征. 词袋采用 1000 个词汇, 进行 kmeans 聚类, 最后应用到 Corel-1K 和 Corel-10K 数据集上.

4.3 UK-bench 数据集

为了说明 GRF 方法在多特征图融合上的效果. 在 UK-bench 数据集上, GRF 方法采用 VOC 和 HSV 方法作为基础特征, 进行多特征图融合, 并同目前的主流的其他多特征图融合方法进行比较, 比较结果详如表 2 所示:

Table 2 Comparisons of N-S Score on the UK-bench Data Set
表 2 在 UK-bench 数据集多特征图融合方法的 N-S 值结果

Method	N-S Value
SVM Fusion ^[10]	3.17
MSD ^[39]	3.23
CNN ^[38]	3.56
Ref ^[23]	3.28
Ref ^[10]	3.67
Graph PageRank ^[10]	3.76
Graph Density ^[10]	3.77
VOC ^[30]	3.54
GRF VOC	3.73
GRF Fusion	3.83
GRF Iterative	3.86

结果较 VOC(N-S 值为 3.54)偏低. 经过不同的图融合方法都得到不同的提高. 本文的单特征优化, 使 HSV 的 N-S 值提高了 0.12, VOC 结果的 N-S 值提高了 0.29, 造成 HSV 和 VOC 差异的主要原因是: VOC 和 HSV 的 $P(f(i, i')=1)$ 的大小不同, 导致 VOC 的特征图相对于 HSV 的特征图更接近于真实的概率. 因此, GRF 算法对 VOC 较 HSV 的 N-S 值效果提高更大.

本文的图融合方法可以将 UK-bench 数据集的 N-S 值提高到 3.83. 通过 HSV, VOC 采用文献^[31]的方法进行迭代. 在迭代之后 HSV(N-S 值

为 3.35)和 VOC(N-S 值为 3.73),进行 GRF 方法的多特征图融合可以将 UK-bench 数据集的 N-S 分数提高到 3.86,GRF 算法的迭代效果如图 5 所示:

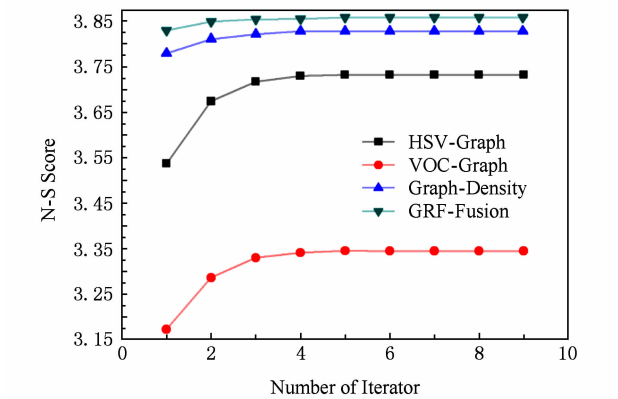


Fig. 5 N-S Score on the UK-bench data set after iterative building graphs

图 5 UK-bench 数据集上不同方法迭代之后的 N-S 值

Table 3 Comparison of Different Image Retrieval Methods on Corel-1K

表 3 图融合以及其他算法在 Corel-1K 数据集上前 20 幅图片的准确率

Method	African	Beach	Building	Bus	Dinosaur	Elephant	Flower	Horse	Mountains	Food	Avg
ODII ^[32]	84.7	45.4	67.8	85.3	99.3	71.1	93.3	95.8	49.8	80.8	77.3
BEP ^[33]	51	90	58	78	78	100	84	100	84	38	78.3
GMM ^[34]	72.5	65.2	70.6	89.2	100	70.5	94.8	91.8	72.25	78.8	80.57
PCM ^[35]	84.9	35.6	61.6	81.8	100	59.1	93.1	92.8	40.4	68.2	71.7
CTCHIPS ^[36]	68.3	54	56.2	88.8	99.3	65.8	89.1	80.3	52.2	73.3	72.7
CNN ^[38]	65.05	63.35	55.85	47.45	49.4	47.15	45.35	60.9	58.6	86.6	57.97
MSD ^[39]	76.2	42.05	83	91.25	99.85	66.9	91.85	92.95	49.35	88.95	78.235
GRF Fusion	85.3	57.15	76.25	86.5	100	86.45	86.1	97.55	81.15	77.25	83.37

Table 4 Comparison of Different Image Retrieval Methods on Corel-10K

表 4 不同算法在 Corel-10K 数据集上的准确率 %

Method	Precision
Ri-HOG ^[37]	52.13
HOG ^[40]	33.29
CNN ^[38]	46.81
MSD ^[39]	45.62
HSV	46.33
BOW	41.25
GRF Fusion	54.69

表 3 和表 4 的准确率结果说明,GRF 方法利用全局图像检索信息使 BOW 和 HSV 方法的结果得到了较大的提升.由于不同方法的融合更加准确地

4.4 Corel-1K 和 Corel-10K 数据集

为了说明 GRF 在不同数据集上具有良好效果,本文在 Corel-1K 和 Corel-10K 数据集上采用 BOW 和 HSV 方法作为基础特征进行实验,并同目前在这 2 个数据集上具有较好效果的其他方法进行比较.在 Corel-1K 和 Corel-10K 数据集中每类图片包含 100 幅图片.在 Corel-1K 数据集中,本文比较前 20 幅图片的准确率;在 Corel-10K 数据集上,本文比较前 12 幅图片的准确率;采用 BOW 和 HSV 进行特征融合.相对于基本 BOW 和 HSV 准确率提高了 11.70%,具体的准确率如表 3 所示.在 Corel-10K 数据集上,本文采用 BOW 和 HSV 进行特征融合前 12 幅图片的准确率可以提升 8.36%.相对于 Corel-1K 提升幅度差异主要由于 $P(f(i,i')=1)$ 的大小,导致提升幅度减少,具体 GRF 算法的准确率如表 4 所示.

求出 $w(i,i')$,使本文方法在图融合方面表现出了很大的优势.对于前 k 幅图片准确率的提升并不相同,当 k 接近 50 时在 Corel 数据集上的效果最明显,在 Corel-1K 数据集上可以提高 15% 的准确率,在 Corel-10K 数据集上可以提高 9% 的准确率,不同的 k 值的准确率如图 6 所示.

4.5 时间复杂度分析

在本文实验中, k 根据实际要求的候选图片数量取值为 4~30;分组方法固定迭代 5 次,每次迭代的时间复杂度为 $O(k^2|V|)$,如果数据集较大可以采用离线的方式预先计算结果;每张检索图像重排方法的时间复杂度为 $O(k \log k)$.本文方法计算检索图像排序结果的时间复杂度明显小于提取图像特征的时间复杂度.尤其当 k 的取值较小时($k=4$),本文方法的检索时间小于 1 ms,由此可见本文方法不影

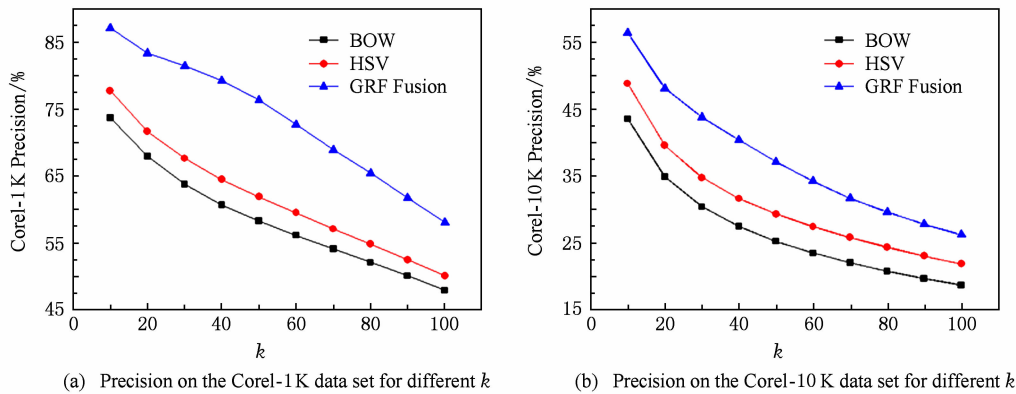


Fig. 6 Precision on the Corel data set for different k

图6 Corel数据集的前 k 幅图片准确率情况

响图像检索的时间复杂度. 在本文实验的数据集上重排时间如表5所示:

Table 5 The Re-rank Time on the Test Datasets	
表5 不同数据集上的重排时间	
Dataset	Re-rank Time/ms
UK-bench	0.674
Corel-1K	31.269
Corel-10K	15.889

5 结 论

本文利用全局候选图片的信息,提出GRF图像排序融合方法.该方法利用不同特征图,将图像检索数据集划分为相似图片组以提升图像检索的效果.从实验结果可以看出,本文多特征图融合方法优于已有的方法;此外,通过融合图像的局部和全局特征,能够更全面地刻画图像的颜色、尺度及旋转等信息,可以较原特征(全局或局部特征)获得更高的精度,且原特征的效果越优,多特征图融合的检索效果越好.当所有特征对该图片检索结果存在较大偏差时,GRF方法依然能够提升图像检索的效果.该方法还存在一些不足,当相似图片组对检索图像整体没有较好的检索结果时,GRF方法难以提升图像检索的效果.在今后的工作中,可以通过研究此问题的解决办法,进一步提升多特征图融合方法对图像检索效果.

参 考 文 献

[1] Datta R, Joshi D, Li Jia, et al. Image retrieval: Ideas, influences, and trends of the new age [J]. ACM Computing Surveys, 2008, 40(2): No. 5

[2] Zhang Chen, Joyce Y C, Rong Jin. User term feedback in interactive text-based image retrieval [C] //Proc of the 28th Annual Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2005: 51-58

[3] Li Xiangyang, Zhuang Yueting, Pan Yunhe. Content-based image retrieval techniques and systems [J]. Journal of Computer Research and Development, 2001, 38(3): 344-354 (in Chinese)
(李向阳, 庄越挺, 潘云鹤. 基于内容的图像检索技术与系统 [J]. 计算机研究与发展, 2001, 38(3): 344-354)

[4] Wang Huifeng, Sun Zhengxing, Wang Jian. Semantic image retrieval: Review and ResEarch [J]. Journal of Computer Research and Development, 2002, 39(5): 513-523 (in Chinese)
(王惠锋, 孙正兴, 王箭. 语义图像检索研究进展 [J]. 计算机研究与发展, 2002, 39(5): 513-523)

[5] Liu Guanghai, Yang Jingyu. Content-based image retrieval using color difference histogram [J]. Pattern Recognition, 2013, 46(1): 188-198

[6] He Jingrui, Li Mingjing, Zhang Hongjiang, et al. Generalized manifold-ranking-based image retrieval [J]. IEEE Trans on Image Processing, 2006, 15(10): 3170-3177

[7] Wang Bin, Pan Feng, Hu Kaimo, et al. Manifold-ranking based retrieval using k -regular nearest neighbor graph [J]. Pattern Recognition, 2012, 45(4): 1569-1577

[8] Huang Yuchi, Liu Qingshan, Zhang Shaoting, et al. Image retrieval via probabilistic hyper graph ranking [C] //Proc of the 23rd IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2010: 3376-3383

[9] Zhao Sicheng, Yao Hongxun, Yang You, et al. Affective image retrieval via multi-graph learning [C] //Proc of the ACM Int Conf on Multimedia. New York: ACM, 2014: 1025-1028

[10] Zhang Shaoting, Yang Ming, Cour T, et al. Query specific rank fusion for image retrieval [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2015, 37(4): 803-815

[11] Zheng Liang, Wang Shengjin, Liu Ziqiong, et al. Packing and padding: Coupled multi-index for accurate image retrieval [C] //Proc of the IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2014: 1939-1946

- [12] Christian W, Matthijs D, Hervé J. Bag-of-colors for improved image search [C] //Proc of the 19th ACM Int Conf on Multimedia. New York: ACM, 2011: 1437-1440
- [13] Shi Jianbo, Jitendra M. Normalized cuts and image segmentation [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905
- [14] Mayur D, Nicole I, Piotr I, et al. Locality-sensitive hashing scheme based on p-stable distributions [C] //Proc of the 20th Annual Symp on Computational Geometry. New York: ACM, 2004: 253-262
- [15] De Maesschalck R, Jouan-Rimbaud D, Massart D L. The mahalanobis distance [J]. Chemometrics and Intelligent Laboratory Systems, 2000, 50(1): 1-18
- [16] Lowe D G. Object recognition from local scale-invariant features [C] //Proc of the 7th IEEE Int Conf on Computer Vision, Vol2. Piscataway, NJ: IEEE, 1999: 1150-1157
- [17] Lindeberg T. Scale invariant feature transform [J]. Scholarpedia, 2012, 7(5): 10491-10491
- [18] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359
- [19] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF [C] //Proc of IEEE Int Conf on Computer Vision (ICCV). Piscataway, NJ: IEEE, 2011: 2564-2571
- [20] Leutenegger S, Chli M, Siegwart R Y. BRISK: Binary robust invariant scalable keypoints [C] //Proc of IEEE Int Conf on Computer Vision (ICCV). Piscataway, NJ: IEEE, 2011: 2548-2555
- [21] Alahi A, Ortiz R, Vandergheynst P. Freak: Fast retina keypoint [C] //Proc of IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2012: 510-517
- [22] Chen D, Tsai S S, Chandrasekhar V, et al. Robust image retrieval using multiview scalable vocabulary trees [C/OL] //Proc of SPIE Electronic Imaging, 2009[2014-08-18]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.142.8937&rep=rep1&type=pdf>
- [23] Jegou H, Schmid C, Harzallah H, et al. Accurate image search using the contextual dissimilarity measure [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2010, 32(1): 2-11
- [24] Niwattanakul S, Singthongchai J, Naenudorn E, et al. Using of Jaccard coefficient for keywords similarity [C] //Proc of the Int Multi-Conf of Engineers and Computer Scientists, 2013 [2014-10-11]. http://www.iaeng.org/publication/IMECS2013/IMECS2013_pp380-384.pdf
- [25] Dwork C, Kumar R, Naor M, et al. Rank aggregation methods for the Web [C] //Proc of the 10th Int Conf on World Wide Web. New York: ACM, 2001: 613-622
- [26] Jing Yushi, Baluja S. Visualrank: Applying pagerank to large-scale image search [J]. IEEE Trans on Pattern Analysis and Machine Intelligence. 2008, 30(11): 1877-1890
- [27] Duygulu P, Barnard K, De Freitas J F G, et al. Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary [C] //Proc of the 7th European Conf on Computer Vision. Berlin: Springer, 2002: 97-112
- [28] Nister D, Stewenius H. Scalable recognition with a vocabulary tree [C] //Proc of IEEE Computer Society Conf on Computer Vision and Pattern Recognition. Los Alamitos, CA: IEEE Computer Society, 2006: 2161-2168
- [29] Li Jia, Wang J Z. Automatic linguistic indexing of pictures by a statistical modeling approach [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(9): 1075-1088
- [30] Wang Xiaoyu, Yang Ming, Timothee C, et al. Contextual weighting for vocabulary tree based image retrieval [C] //Proc of the 13th IEEE Int Conf on Computer Vision. Piscataway, NJ: IEEE, 2011: 209-216
- [31] Wang Meng, Li Hao, Tao Dacheng, et al. Multimodal graph-based reranking for Web image search [J]. IEEE Trans on Image Processing, 2012, 21(11): 4649-4661
- [32] Guo Jingming, Heri P, Su Huaisheng. Image indexing using the color and bit pattern feature fusion [J]. Journal of Visual Communication and Image Representation, 2013, 24(8): 1360-1379
- [33] Walia E, Pal A. Fusion framework for effective color image retrieval [J]. Journal of Visual Communication and Image Representation, 2014, 25(6): 1335-1348
- [34] Zeng Shan, Huang Rui, Wang Haibing, et al. Image retrieval using spatiograms of colors quantized by Gaussian mixture models [J]. Neurocomputing, 2016, 17: 673-684
- [35] Yu F X, Luo Hao, Lu Z M. Colour image retrieval using pattern co-occurrence matrices based on BTC and VQ [J]. Electronics Letters, 2011, 47(2): 100-101
- [36] Lin C, Chen R, Chan Y. A smart content-based image retrieval system based on color and texture feature [J]. Image and Vision Computing, 2009, 27(6): 658-665
- [37] Chen Jinhui, Toru N, Tetsuya T, et al. Content-based image retrieval using rotation-invariant histograms of oriented gradients [C] //Proc of the 5th ACM on Int Conf on Multimedia Retrieval. New York: ACM, 2015: 443-446
- [38] Babenko A, Slesarev A, Chigorin A, et al. Neural codes for image retrieval [C] //Proc of European Conf on Computer Vision. Berlin: Springer, 2014: 584-599
- [39] Liu Guanghai, Li Zuoyong, Zhang Lei, et al. Image retrieval based on micro-structure descriptor [J]. Pattern Recognition, 2011, 44(9): 2123-2133
- [40] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] //Proc of IEEE Computer Society Conf on Computer Vision and Pattern Recognition. Los Alamitos, CA: IEEE Computer Society, 2005: 886-893



Liu Shenglan, born in 1984. PhD. Post doctor at Dalian University of Technology. His main research interests include pattern recognition and machine learning.



Sun Muxin, born in 1989. Master candidate at Dalian University of Technology. His main research interests include pattern recognition and machine learning.



Feng Lin, born in 1969. Professor and PhD supervisor at Dalian University of Technology. His main research interests include intelligent image processing, robotics, data mining, and embedded systems.



Yang Liu, born in 1992. PhD candidate at Dalian University of Technology. His main research interests include video analysis and machine learning.

《计算机研究与发展》征订启事

《计算机研究与发展》(Journal of Computer Research and Development)是中国科学院计算技术研究所和中国计算机学会联合主办、科学出版社出版的学术性刊物,中国计算机学会会刊. 主要刊登计算机科学技术领域高水平的学术论文、最新科研成果和重大应用成果. 读者对象为从事计算机研究与开发的研究人员、工程技术人员、各大专院校计算机相关专业的师生以及高新企业研发人员等.

《计算机研究与发展》于 1958 年创刊,是我国第一个计算机刊物,现已成为我国计算机领域权威性的学术期刊之一. 并历次被评为我国计算机类核心期刊,多次被评为“中国百种杰出学术期刊”. 此外,还被《中国学术期刊文摘》、《中国科学引文索引》、“中国科学引文数据库”、“中国科技论文统计源数据库”、美国工程索引(Ei)检索系统、日本《科学技术文献速报》、俄罗斯《文摘杂志》、英国《科学文摘》(SA)等国内外重要检索机构收录.

国内邮发代号:2-654;国外发行代号:M603

国内统一连续出版物号:CN11-1777/TP

国际标准连续出版物号:ISSN1000-1239

联系方式:

100190 北京中关村科学院南路 6 号《计算机研究与发展》编辑部

电话: +86(10)62620696(兼传真); +86(10)62600350

Email: crad@ict. ac. cn

http://crad.ict. ac. cn