

发掘相关反馈日志中关联信息的图像检索方法*

张亮⁺, 施伯乐, 周向东, 刘莉, 张琪

(复旦大学 计算机与信息技术系, 上海 200433)

Enhancing Content-Based Image Retrieval by Exploiting Relevance Feedback Logs

ZHANG Liang⁺, SHI Bai-Le, ZHOU Xiang-Dong, LIU Li, ZHANG Qi

(Department of Computing and Information Technology, Fudan University, Shanghai 200433, China)

+Corresponding author: Phn: +86-21-65641867, E-mail: zhangl@fudan.edu.cn, <http://www.cit.fudan.edu.cn>

Received 2002-09-04; Accepted 2003-03-04

Zhang L, Shi BL, Zhou XD, Liu L, Zhang Q. Enhancing content-based image retrieval by exploiting relevance feedback logs. *Journal of Software*, 2004,15(1):41~48.

<http://www.jos.org.cn/1000-9825/15/41.htm>

Abstract: Relevance feedback (RF) has been successfully used in content-based image retrieval (CBIR). However, most CBIR systems seldom reuse the latent semantic correlation among images revealed by RF log to guide retrieving across sessions. In this paper, concurrence of images in a RF record is regarded as a kind of semantic homogeneity in certain context and the image-retrieving problem is cast as an authority-image-finding task. Records in RF logs first extend the result from traditional CBIR systems. This produces a relevant graph of images related to the query with multiplex contexts. Then, a modified HITS algorithm is applied to it to distill consensus about semantic relevance. As a result, both visual content and semantic relevance can be maintained in image retrieval and the efficiency is much improved compared with traditional CBIR methods. Experimental results demonstrate its superiority in both objective criteria and semantic clustering capability against the Corel database with 60 000 images.

Key words: image retrieval; relevance feedback logs; image content; semantics in image; HITS algorithm

摘要: 相关反馈日志蕴含着丰富的对象语义关联信息,但大多数基于内容的图像检索(CBIR)方法却缺乏对它们的重用.提出一种发掘反馈日志中图像关联信息的自动化图像检索方法,将反馈事例中图像的共生现象视为一定上下文中的图像分类.检索时,结合CBIR的检索结果和多种上下文中的图像分类实例,借鉴HITS算法的思想从中提炼图像的本质性关联,获得综合内容和语义的图像检索结果.对6万幅Corel图像数据库的实验表明,该方法可以显著改善查全率和查准率,且检索结果能够更好地满足用户的语义检索需求.

关键词: 图像检索;相关反馈日志;图像内容;图像语义;HITS算法

* Supported by the National Natural Science Foundation of China under Grant No.69933010 (国家自然科学基金)

作者简介: 张亮(1963—),男,湖北武汉人,博士,教授,主要研究领域为支持多媒体应用的数据库技术及信息集成;施伯乐(1935—),男,教授,博士生导师,主要研究领域为数据库理论及应用;周向东(1969—),男,博士,讲师,主要研究领域为数据库,信息检索;刘莉(1978—),女,硕士生,主要研究领域为数据库,信息检索;张琪(1979—),女,硕士生,主要研究领域为数据库,信息检索.

中图法分类号: TP391 文献标识码: A

图像作为一种重要的媒体类型,其检索技术研究受到了广泛的重视.基于内容的图像检索(content-based image retrieval,简称 CBIR)已成为该领域的主流技术^[1].CBIR 自动提取图像的视觉特征,如色彩、纹理和形状,并以此作为图像内容的表达和图像相关程度的判别要素.

由于图像语义的丰富性和“相关”判断的主观性,以语义为特征的检索需求和 CBIR 的视觉特征之间往往难以匹配,特征相似(similar)难以保证图像相关(relevant),因而导致检索效率的恶化.为了解决这一问题,近期的 CBIR 大都采用了相关反馈(relevance feedback)机制,分析用户对部分返回结果的相关性判断,调整检索样本和/或修正相似性度量,力图使系统逐步贴近用户的语义检索需求^[1-3].值得注意的是,大多数 CBIR 方法仅限于在一次检索活动(session)中使用相关反馈技术,没有对检索样本与反馈事例之间的对应关系以及反馈事例之间的关联关系加以保存、发掘和利用.每次获得新的检索样本后,系统不得从头开始学习,需要较长的学习过程才能揣摩到用户的检索意图,给出较为合理的结果.现有方法的另一个问题是,系统不可能利用多个反馈事例中图像的关联,缺乏长期的学习技能,无法返回语义内容相关但视觉特征迥异的检索结果.因此,对相关反馈信息的长期学习和利用,以及从不同上下文的关联信息中提炼与当前检索图像本质相关图像的新型图像检索方法值得关注.

近年来,人们已逐步认识到对相关反馈事例进行长期学习的必要性.例如,Lu 等人^[4]在 CBIR 图像内容组织方式之外,又添加了一层语义网络,借助用户的反馈信息对图像进行语义分类.其主要特色是将用户相关反馈的历史信息明确地作用于语义分类之中,提供了基于语义检索图像的可能性,但它以文字表征语义进而建立图像关联的策略有可能面临自然语言表达多义性的问题.之后,Yang 等人^[5]尝试利用 Wordnet 来解决这一问题.可是,这两种方法在处理带有不同上下文背景的关联关系问题时采用的机制过于简单,难以从中提炼与检索样本本质性相关的图像.此外,Bartolini 等人^[6]的 FeedbackBypass 记录以往检索样本及其“最优”检索参数(通过相关反馈调整后的样本点和相似性度量),并给出了高效的反馈事例组织结构.但其方法依然局限于视觉特征能够完整反映语义内容的假设条件之中,无法返回语义内容相关但视觉特征迥异的检索结果.

综上所述,用户相关反馈日志蕴含着丰富的对象分类信息,有助于提高图像检索效率.如何发掘这些信息,综合 CBIR 的视觉特征和分类信息的语义内容,提炼与检索样本本质相关的图像集合,排除主观因素形成的反馈“噪声”,同时改善短期检索的效果和长期学习能力是一个值得研究的课题.这也正是本文的研究内容.

本文第 1 节简要介绍 CBIR 和结构分析算法 HITS,它们是本文工作的基础.第 2 节详细描述发掘相关反馈日志的图像检索方法.第 3 节从多个方面分析新方法的特点.第 4 节报告对 6 万幅 Corel 图像库的实验结果.从这两节可以发现,新方法是对现有 CBIR 方法的扩展,不仅利用了 CBIR 的成果,明显改善了查全率和查准率,还使检索结果更好地满足用户的语义检索需求.第 5 节是研究结论.

1 CBIR 和 HITS 概要

作为新方法的基础,下面我们简要介绍一些相关概念和算法.更为详尽的材料请参考文献[1~7].

1.1 CBIR

CBIR 是当前图像检索的主流技术^[1].其中所谓的内容泛指可自动处理的图像视觉特征,如色彩、纹理或形状.CBIR 以这些特征或其组合近似图像的语义信息并在它们的基础上判断图像的相关程度.

与大多数 CBIR 方法类似,本文采用向量模型.对于图像库 Ω 中的图像 p ,用底层视觉特征(如颜色和纹理)来表示图像,记为向量 $p=[p_1, \dots, p_N]^T$,其中 p_k 是特征分量的表征值, $k=1, 2, \dots, N$.同样地,检索样本表示为向量 $q=[q_1, \dots, q_N]^T$.两幅图像的相关程度通过对应向量的相似度 $Sim(p, q)$ 来计测.本文采用余弦度量.CBIR 检索算法如下:

CBIR 检索算法.

/* Given an image database Ω , both sample image q and image $p \in \Omega$ are represented by feature vectors */

NAME: CBIR

INPUT: a query sample image q and required number of images in result Num

OUTPUT: a set of images ranked by similarity in regard to q

1. Calculate the feature vector of q
2. For all $p \in \Omega$ /* retrieving */
Calculate the similarity $Sim(p, q)$
3. Ranking all $p \in \Omega$ by $Sim(p, q)$ in a descend order
4. Fetch the top Num images to R
5. If user is satisfied
Return R
Otherwise /* relevance feedback */
 - a) Display images in R
 - b) Get user Feedbacks F
 - c) Adjust q and parameters according to F
 - d) Go to 2

为缩小语义与特征之间的差异,近期的 CBIR 大都采用了相关反馈机制^[1-3]。算法第 5 步中的相关反馈是一个人机交互过程。系统根据当前检索样本和用户反馈事例修正检索样本和/或相似性度量参数,以反映用户脑海中的语义概念向系统内特征集合的映射关系。这一过程不断反复,直至用户提交新的检索样本或结束检索为止。目前,相关反馈大都局限于一次检索活动中,反馈结果对新样本图像的检索不提供帮助。

1.2 HITS

HITS^[7]是一个面向 Web 的页面精选算法。它基于这样一种假设,即页面间的超链接隐含了页面权威性的质量信息。更明确地说,页面 p 中设计了指向页面 q 的超链接,意味着赋予了 q 某种程度的权威性。进一步而言,相对于某个共同主题 σ 的权威页面之间还应该彼此密集链接。一些称为集散(hub)页面的网页将同一主题的权威(authority)页面“黏着”成一团。两者之间存在着相互强化的关系:一个好的集散页面一定指向一组好的权威页面;一个好的权威页面一定被一组好的集散页面所指向。

参照文献[7]的符号约定,HITS 算法如下。它由两个相继的过程 Subgraph 和 Iterate 构成。前者负责对由搜索引擎得到的结果根据页面链接线索扩展成与主题相关的图 $G=(V, E)$;后者通过迭代从 G 中提炼权威页面和集散页面。Iterate 赋予 G 每个节点以权威值和集散值,形成图 G 权威向量 x 和集散向量 y 。Iterate 迭代施加

$$I \text{ 操作: } x_i^{(p)} = \sum_{q:(q,p) \in E} y_i^{(q)}, \quad O \text{ 操作: } y_i^{(p)} = \sum_{q:(q,p) \in E} x_i^{(q)}$$

于向量 x 和向量 y 。文献[7]已证明,对于足够大的 k , x 和 y 将收敛于不动点。实际应用中,可在 k 次迭代之后,从 x_k 中提取前 Num 个最大分量值所对应的页面作为关于主题 σ 的权威页面,在 y_k 中提取前 Num 个最大分量值所对应的页面作为关于主题 σ 的集散页面。

HITS 算法。

NAME: Subgraph

- INPUT: 1. σ : a query string
2. E : a text-based search engine
3. t, d : natural numbers

OUTPUT: a collection of linked pages

1. Let $R\sigma$ be the top t results of E on σ
 2. Set $S\sigma = R\sigma$
 3. For each page $p \in R\sigma$
 - a) Let $\Gamma^+(p)$ be the set of all pages p points to
 - b) Let $\Gamma^-(p)$ be the set of all pages pointing to p
 - c) Add all pages in $\Gamma^+(p)$ to $S\sigma$
 - d) If $|\Gamma^-(p)| \leq d$ Add all pages in $\Gamma^-(p)$ to $S\sigma$
- Otherwise Add d pages from $\Gamma^-(p)$ to $S\sigma$

4. Return $S\sigma$

NAME: Iterate

- INPUT: 1. $G: S\sigma$ returned from Subgraph, denoted as $G=(N, E), n=|N|$
2. k : a natural number

OUTPUT: a collection of linked pages

1. Let z be the vector $[1 \ 1 \ \dots \ 1]^T \in R^n$
2. Set $x_0 = z, y_0 = z$
3. For $i=1, 2, \dots, k$,

- a) Apply the I operation: $x_i^{(p)} = \sum_{q:(q,p) \in E} y_{i-1}^{(q)}$

- b) Apply the O operation: $y_i^{(p)} = \sum_{q:(q,p) \in E} x_i^{(q)}$

- c) Normalize x'_i , obtaining x_i

- d) Normalize y'_i , obtaining y_i

4. Return (x_k, y_k)

2 发掘相关反馈日志的图像检索方法

本文希望设计一个新的图像检索方法,以实现下列目标:第 1,发掘和利用以往用户反馈中图像的关联信息;第 2,将图像检索从基于视觉特征提升到综合特征和语义的水平;第 3,提炼图像间的关联,在大量反馈中消除不和谐噪声;第 4,用图像集合隐含地表示样本所体现的概念,回避以自然语言表达概念所带来的额外问题。

2.1 发掘和利用图像间的关联

用户的一次相关反馈定义了特定语义环境中反馈图像与样本图像的某种语义相关性.为此,我们将用户的检索样本 q 连同用户指定的反馈图像构成一个反馈事例,按 $q':\{f_1, \dots, f_k\}, q', f_i \in \Omega$ 的格式保存于相关反馈日志中.其中,如果 $q \in \Omega$,则 $q'=q$;否则, q' 是以 q 作为样本检索时系统返回的最近似图像的编号.之后检索时,一旦发现反馈事例中的某个图像再次出现,就可以带出记录中的其他图像,实现了粗略的语义扩展。

应当承认,反馈事例中的图像关联并不等同于它们之间的语义相关.每一例相关反馈都依赖于当时用户特定的上下文(context).然而,世界上大多数人对事物评判有着基本的共同点,因此大量反馈事例将蕴含着部分图像间语义划分本质上的共性.这可以用下面的方法去发掘和提炼。

2.2 关联的传播

一个反馈事例一般不可能包含所有语义相关的图像,需要根据当前检索目标加以收集.仿照 HITS,提取 CBIR 结果中排名靠前的 t 个图像作为根集(root set),依据用户相关反馈日志,扩充成基集(base set).工作中,我们

取 $t=5$,而不是 HITS 中的 $t=200$.这是因为,在 HITS 原本发展起来的文本信息检索领域,前期的检索基本上都是基于检索词的某种评价指标(如 tfidf)完成的.与语义更为丰富、检索工具更为简陋、检索精度更为低下的图像相比,文字的语义较为准确,因此检索结果中排名较高的 t 个文档与检索词 σ 相关的内容占据绝大多数,然而对于 CBIR,情况就可能完全不同.对于大规模的图像库,没有用户参与反馈的检索结果在很多情况下只含有少数几个相关图像,甚至只有检索样本自己。

为了弥补 t 太小带来的困难,对于把握性较大的对象,应迅速增大它们的关联范围.一般来说,检索样本、用户本次检索过程中反馈的图像都可以认为是与用户当前检索需求密切相关的.我们称其为核集(kernel set).通过计算核集按反馈日志中关联的传递闭包,能够发现经过多步关联才能联系在一起的图像。

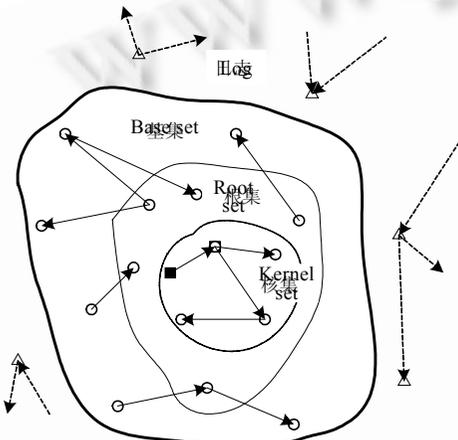


Fig.1 Kernel set, root set, base set, RF log and their relationship

图1 核集、根集、基集、RF 日志及其关系

核集、根集、基集与 RF 日志的关系如图 1 所示。

2.3 本质性关联的提炼

用户的归类带有其主观性,日志中每一条用户相关反馈记录是在当时特定条件下用户对若干图像的语义分类,用户甲和用户乙的观点可能不同,一个用户过去与现在所理解的概念可能也会有所改变.因此,根集经过扩展之后,引入了语义异构性,基集所代表的图像分类还需要进一步提炼,形成最为本质的分类信息。

借助 HITS 的思想,在由检索样本和反馈日志共同决定的基集中,本质相关图像应该是其间密集指向的图像集合.一幅图像是否为满足检索需求的相关图像,取决于它是否曾被多个重要的样本图像所指向.类似地,重要的集散图像是那些指向多个相关图像的、曾经作为检索样本的图像.图像的本质相关性应通过迭代得到凸现。

与 HITS 略有区别的是,相关反馈日志中图像共生的次数也是鉴别它们相关程度的一个重要指标.这体现了众多用户在类别概念上的趋同性,需要在算法中得到体现.另外,图像的视觉特征相似度也应该作为相关图像的排名参考因素.特别是在日志中没有包含检索样本的反馈记录,而前 t 个检索结果的扩展又没有构成连通子图

的时候,算法返回结果的排名应该是 CBIR 检索结果的排名.

2.4 反馈日志的组织

随着检索活动的推进,用户相关反馈日志将逐步增大.有效管理反馈事例记录是保证算法可伸缩的重要保证.目前,我们以图像邻接矩阵的形式组织反馈事例.矩阵元素 a_{pq} 表示日志中以图像 p 为样本图像的反馈事例中包含图像 q 的次数.该矩阵极度稀疏,因此以适应稀疏矩阵的数据结构保存于文件中,分别按行和列索引.

2.5 图像检索算法CBIR/LM

发掘用户相关反馈日志中关联信息的图像检索算法 CBIR/LM(CBIR with log mining)如下.

/* Given an image database Ω , and a relevance feedback log file L */

NAME: CBIR/LM

INPUT: a query sample image q and required Number of images in result Num

OUTPUT: a set of images ranked by similarity in regard to q

1. Set $H=\emptyset$ /* feedback in this session */
 2. $R=CBIR(q,NUM)$ /* traditional CBIR without relevance feedback */
 3. Let G be the top t images in R /* $t=5$ in the experiments */
 4. Calculate the base set S
 - a) Set $S=(\{q\}\cup H)\cup G$ /* $\{q\}\cup H$ is the kernel set */
 - b) Make a real unitary matrix W in size of $|S|\times|S|$
 - c) Update W according to L , for each record $s:\{f_1,\dots,f_{k_s}\}\in L$
 - i) If $s\in S$, increment $W[s,f]$ by 1 for $f\in\{f_1,\dots,f_{k_s}\}$
 - ii) For $g\in S\cap\{f_1,\dots,f_{k_s}\}$, increment $W[g,f]$ by 1 for $g\in\{f_1,\dots,f_{k_s}\}\cup\{s\}\setminus\{f\}$
 5. Make a vector z in size of $|S|$, such that $z_i=1$ if $i\in\{q\}\cup H$; $sim(q,i)$ otherwise
 6. Normalize z
 7. Set vectors $x=z$, and $y=z$
 8. Distill authority images, iterate following Operation k times

/* k is determined by system, $k=10$ in the experiments */

 - a) Apply the I operation: $x'=W^T y$ /* refer to §1.2 */
 - b) Apply the O operation: $y'=W x'$
 - c) Normalize x' , obtaining x
 - d) Normalize y' , obtaining y
 9. Set $n=\min(|S|, Num)$
 10. Fetch n images in x with largest values, denote the sequence by A
 11. If $(Num-n)>0$, concatenate $(Num-n)$ images in $R\setminus A$ to A , denote it by T
 12. If user is satisfied,
 - a) Update L by $L\cup\{q:H\}$ /* refer to §2.1 */
 - b) Return T
- Otherwise /* relevance feedback */
- a) Display images in T
 - b) Get user Feedbacks F
 - c) Adjust q and parameters according to F
 - d) Set H by $H\cup F$
 - e) Go to 2

算法由 4 大部分组成.第 1 部分(步骤 2)与通常的 CBIR 一致,只是去掉了相关反馈部分.第 2 部分(步骤 3)起到了 HITS 算法中 Subgraph 的作用,根据相关反馈日志 L 扩张 CBIR 的检索结果,为后一阶段做准备.第 3 部分(步骤 8)利用迭代提炼本质上相关的图像,即增强那些在 L 中具有频繁联系图像的地位.这部分的核心依然是 HITS 的 Iterate,表述上略有不同,用矩阵与向量的乘法替代了求和,这更便于结合本方法对标准 HITS 的扩充,即将过去的邻接矩阵扩展为带权邻接矩阵(步骤 4b)和 4c)),权重 $W[g,f]$ 是 L 中图像 g 指向 f 的次数.第 4 部分(步骤 9~步骤 12)处理新情况下的结果形成与相关反馈处理.由于提炼过程产生的相关图像(即权威图像)数可能少于

用户需要返回的数量,所以算法将 CBIR 结果中没有在权威图像集合里出现的、但与检索样本 q 视觉相似的图像拼接在权威图像之后显示(步骤 9~步骤 11).这种情况主要出现在 L 中还没有蓄积足够多的有关 q 及其视觉上最相关的 t 幅图像信息的场合.对向量 z 初始化(步骤 5)考虑了视觉相似在提炼过程中的作用.如果图像是检索样本或在本次检索前面几个回合的用户相关反馈中出现,就认为它肯定满足用户的需要,否则,将对应元素初始化为该图像与检索样本的视觉相似度.当然,经过提炼,上面的因素不会在相关性分析中占据主要地位,但当无法进行链接分析(如极端情况下 L 中没有有关 q 和它视觉上最相关的 t 幅图像信息)的场合,CBIR 的检索结论将占据绝对重要的地位.在处理本次检索的相关反馈问题时,步骤 12d)扩展本次的反馈图像 H ,待用户退出检索时,更新相关反馈日志(步骤 12a)).此外,算法步骤 3a)使用了第 2.2 节介绍的关联传播方法.针对图像检索的特殊应用,调整了 HITS 算法中的 Subgraph.

需要指出的是,算法中带权邻接矩阵 W 与组织日志文件的邻接矩阵是完全不同的对象.后者记录整个数据库中图像在反馈日志中的共生信息,而前者只是基集里图像在反馈日志中的共生信息,其规模与检索样本代表的主题密切相关,远远小于组织日志文件的邻接矩阵.这正是 HITS 的得意之处,即以小规模子图(与检索词相关)的代价获得近似全局(整个 Web)的最优结果.

3 方法的效果分析

本节参照第 2 节设定的 4 个目标,定性分析 CBIR/LM 的效果.实验及其分析在下一节给出.

CBIR/LM 可以发掘和利用以往用户反馈中图像的关联信息.从算法的步骤 4a)可以看出,如果在相关反馈日志 L 中存在涉及检索样本 q 、本次检索相关反馈图像(H 的元素)或 CBIR 前 t 个结果的任何记录,CBIR/LM 都将会对此加以利用和扩展.吸收进来的信息通过步骤 8 的提炼过程,形成最终的相关图像.

CBIR/LM 将图像检索从基于视觉内容提升到综合内容和语义的水平.这主要得益于对用户相关反馈历史记录利用.用户的反馈建立在语义相关的基础上,用它们来扩展检索候选的相关图像集合,就把这种语义信息带入了新算法的整个检索过程和检索结果之中.算法的提炼过程(步骤 8)也主要是基于相关反馈信息来工作的,因此,链接分析不会动摇语义相关的基础.算法中根集 I 的形成(步骤 3)以及初始向量的设计(步骤 5)考虑了图像视觉内容对检索结果排名的影响.

CBIR/LM 提炼图像间的关联,剔除反馈事例中隐藏的噪声.步骤 8 的迭代过程,是在前面扩展的图像关联图中寻找权威节点的过程.基于第 2.3 节的定义,图像的相关性(权威性)和集散性存在相互增强的关系,那些偶尔曾被反馈但极少与基集 S 中其他图像关联的噪声图像不可能获得足够的权威值和集散值,因而最终将被算法抛弃.

CBIR/LM 用图像集合表示概念,回避以自然语言表达概念所带来的额外问题.算法对用户提供的图像分类信息,不论是提炼前还是提炼后,所有的类别都是用外延的方式表示的,即刻画为图像的集合,不依赖任何文字描述.因此,无须考虑同名异义或同义异名的问题.语义维护也相对简单.

4 实验分析

为了验证 CBIR/LM 对 CBIR 的性能改进,对 Corel 60 000 幅图像数据库(<http://jzw.stanford.edu/IMAGE/download/corel1m.60k.tar>.由 Dr.James WANG 提供.对此表示衷心的感谢)进行了图像检索实验.

在 CBIR 视觉特征方面,实验系统采用色彩与纹理特征.首先,将 RGB 色彩空间转换为标准 u - v 色彩空间:

$$u = \frac{R}{R+G+B}, \quad v = \frac{G}{R+G+B}.$$

然后,在 u, v 上各取 16 个 bin,构成一个 32 维的向量,再对图像的亮度分量进行 3 次小波变换得到 10 个子段,取每个子段的标准差合成后得到一个 10 维的向量.色彩特征和纹理特征加权求和形成图像特征.

实验中,CBIR/LM 算法和参数设定与第 2.5 节中的叙述相同.检索效果评价指标是查全率(recall)和查准率(precision).

实验 1. CBIR/LM 对多种类型图像改进效果的对比.选择 3 个有代表性的图像类别:国旗、海洋鱼类和野外大象图片.国旗类图像较为简单,对象形状较为规整,背景噪声小,但不同国家的旗帜有色彩、纹理变化;海洋鱼类图像复杂度居中,形状变化较大,背景较为复杂;野外大象图片最为复杂,背景噪声大,对象形状、纹理、色彩发生很大的变化.对每类图像随机抽取 16 幅图像作为检索样本,当前相关反馈事例不加入反馈日志库 L .使用 CBIR/LM 之前, L 各有同类图像 30%左右的反馈案例.基于查全率和查准率的对比检索结果如图 2~图 4 所示.为便于观察,各次实验结果按普通 CBIR 检索效果的升序排布.其中图 2 中从左向右:查全率-查准率综合效果改善情况、各次实验查全率的改善情况以及各次实验查准率的改善情况.6 万幅图库中语义相关图像 100 幅,检索结果限定前 32 幅图像.图 3 中,从左向右分别为查全率-查准率综合效果改善情况、各次实验查全率的改善情况以及各次实验查准率的改善情况.6 万幅图库中语义相关图像 85 幅,检索结果限定前 32 幅图像.图 4 中,从左向右分别为查全率-查准率综合效果改善情况、各次实验查全率的改善情况以及各次实验查准率的改善情况.6 万幅图库中语义相关图像 100 幅,检索结果限定前 32 幅图像.

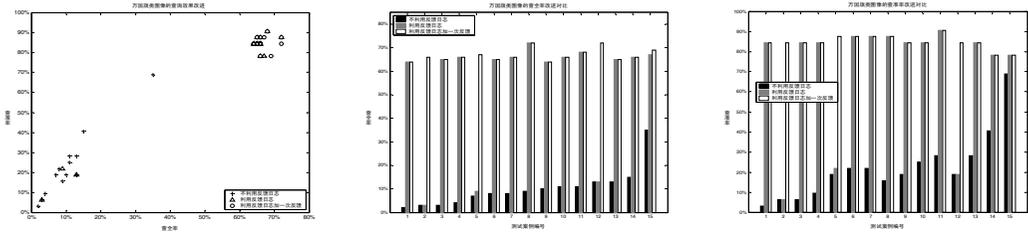


Fig.2 Performance improvement against flags data set
图 2 对万国国旗照片的检索效果图

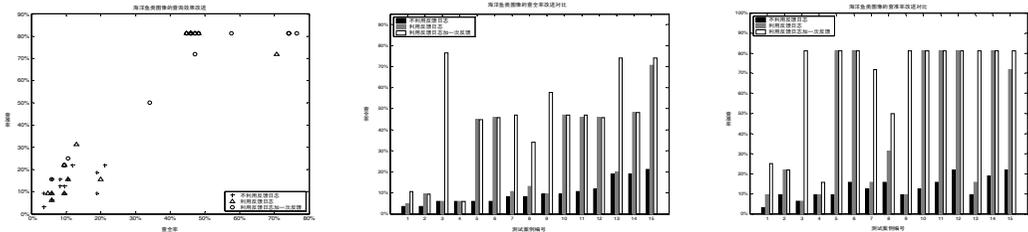


Fig.3 Performance improvement against ocean fish data set
图 3 对海洋鱼类照片的检索效果图

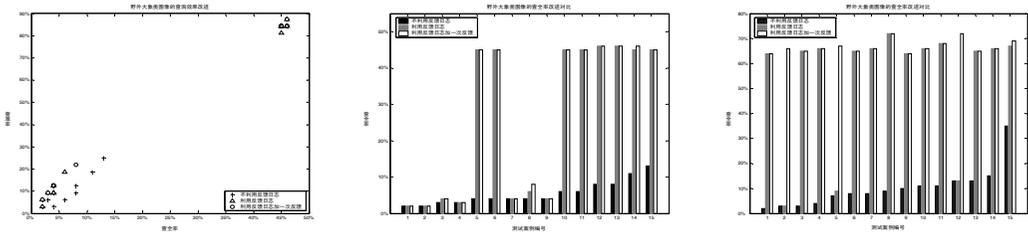


Fig.4 Performance improvement against field elephant data set
图 4 对野外大象照片的检索效果图

通过第 1 列的查全率-查准率(recall-precision)可以发现综合效果的改善,第 2 列和第 3 列分别是查全率和查准率效果对比图,从中能直接发现单个测试案例在相应评价指标上的变化.应该指出的是,由于综合效果图中每个测试案例没有编号,原本 CBIR 检索效果就不好,经过 CBIR/LM 虽然有所改善,但总体检索效果不高的实验结果容易引起对新方法改进效果的误解.这一点在后两幅图中被揭示出来.总的来看,使用 CBIR/LM,各项检索指标都得到了很大范围的提升,改善后的结果也明显优于其他利用相关反馈机制的 CBIR 方法.

实验 2.观察用户相关反馈日志规模对检索效果的影响.选择国旗图像类.分别设定反馈日志 L 中记录数量占同类图像数量的 0%,1%,5%,10%,15%,20%,25%和 30%,考察检索效果的变化.结果如图 5 所示.中图和右图中的实线代表不同样本的 10 次实验,带方块标记的加粗虚线是它们的平均性能随日志规模的变化情况.图 5 中,

从左向右:查全率-查准率综合效果改善情况、各次实验查全率的改善情况以及各次实验查准率的改善情况.中图和右图中带有方块标记的加粗虚线表示 10 次实验的平均查全率和查准率随日志规模的变化情况.可以发现,不论是综合效果,还是单项指标,随着用户反馈日志规模的不断扩大,检索效果都会有所改善.反馈日志规模为 0% 时对应着普通 CBIR 的检索效果.在维持了 30% 左右的反馈案例之后, CBIR/LM 的平均检索效果比它提高了 3~4 倍.这表明,用户的相关反馈案例确实是一种不可忽视的宝贵财富,合理地加以利用,将会带来明显的效果.同时还表明, CBIR/LM 将是一个自适应系统,随着时间的演化,其性能将会越来越好.

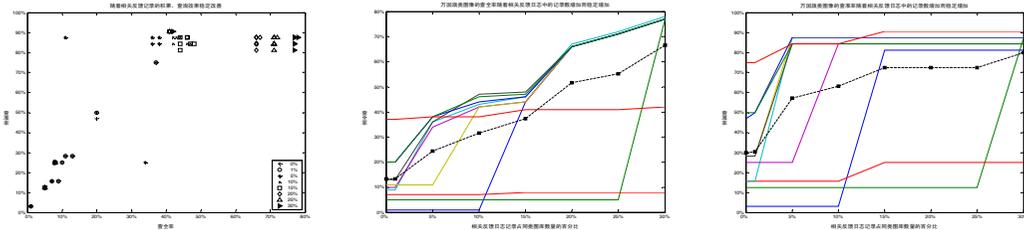


Fig.5 Retrieval performance versus log scale
图 5 检索性能与日志中反馈记录数的关系图

5 结论

用户的一次相关反馈定义了特定语义环境下相关图像的一个集合.由于世界上大多数人对事物评判有着基本的共同点,多次的反馈事例尽管存在由于语义评价主观性和评价指标不同带来的差异,但可以认为众多由相同元素参与的反馈事例蕴含着本质上的共性.因此,检索活动中形成的用户相关反馈日志蕴含着丰富的对象分类信息,值得加以开发和利用.但现行 CBIR 方法却普遍忽视了此类信息.

本文提出了一种发掘用户相关反馈日志的图像检索方法,它将以往用户反馈中图像的关联信息用于扩充候选相关图像集合,将图像检索从基于视觉内容提升到综合内容和语义的水平.借鉴 HITS 以局部视图代价获得全局优选结果的思想,该方法通过链接分析,提炼图像间的关联,发现本质上与样本图像语义相关的图像集合.由于采用了以图像集合表示概念的策略,因此避免了以自然语言表达概念所带来的额外问题.经过定性分析和对 6 万幅 Corel 图像库的检验,证实新方法可以明显改善查全率和查准率,检索结果能够更好地满足用户的语义检索需求.

References:

- [1] Rui Y, Huang TS, Chang SF. Image retrieval, current techniques, promising directions, and open issues. *Journal of Visual Communication and Image Representation*, 1999,10(1):39-62.
- [2] Rui Y, Huang TS, Mehrotra S. Content-Based image retrieval with relevance feedback in MARS. In: *Proc. of the IEEE Int'l. Conf. on Image Processing, II*. IEEE Computer Society, 1997. 815-818. <http://ieeexplore.ieee.org/iel5/4998/13789/00638621.pdf>
- [3] Rui Y, Huang TS. A novel relevance feedback technique in image retrieval. In: *Proc. of the 7th ACM Int'l. Conf. on Multimedia*. ACM Press, 1999. 67-70. <http://citeseer.nj.nec.com/rui99novel.html>
- [4] Lu Y, Hu CH, Zhu XQ, Zhang HJ, Yang Q. A unified framework for semantics and feature based relevance feedback in image retrieval systems. In: *Proc. of the 8th ACM Int'l. Conf. on Multimedia*. New York: ACM Press, 2000. 31-38.
- [5] Yang J, Liu WY, Zhang HJ, Zhung YT. Thesaurus-Aided approach for image browsing and retrieval. In: *Proc. of the IEEE Int'l. Conf. on Multimedia and Expo*. 2001. 1135-1138. <http://ieeexplore.ieee.org/iel5/8766/27769/01237927.pdf>
- [6] Bartolini I, Ciaccia P, Waas F. FeedbackBypass: A new approach to interactive similarity query processing. In: *Proc. of the 27th Int'l. Conf. on Very Large Data Bases*. Roma: Morgan Kaufmann Publishers, 2001. 201-210.
- [7] Kleinberg JM. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 1999,46(5):604-632.