



交互式图像检索中的相关反馈技术研究进展^{*}

谭晓阳^{1,2**}, 孙正兴¹, 张福炎¹

(1. 南京大学多媒体计算机研究所, 南京, 210093;

2. 南京航空航天大学计算机系, 南京, 210016)

摘要: 相关反馈是近年来交互式图像检索领域研究的重要方向. 首先提出了基于相关反馈的图像检索系统框架, 并在此基础上从机器学习的角度分析了相关反馈学习的算法模型、样本获取、分布密度估计, 及其在其特定应用背景下的困难和挑战. 进而对图像相关反馈技术的研究现状进行调查总结, 从聚类和分类两个方面对各种相关反馈算法在基于内容的图像检索中的应用进行了较为深入地研究和比较. 最后对相关反馈技术发展趋势进行了展望, 指出了该技术与图像语义抽取、用户模型建立以及软计算技术之间存在的密切关系.

关键词: 相关反馈, 基于内容的图像检索, 信息检索

中图分类号: TP 39

Relevance Feedback in Content-based Image Retrieval: the State of the Art

Tan Xiao-Yang^{1,2}, Sun Zheng-Xing¹, Zhang Fu-Yan¹

(1. Institute of Multimedia Computing Technology, Nanjing University, Nanjing, 210093, China;

2. Department of Computer Science and Engineering, Nanjing,
University of Aeronautics and Astronautics, Nanjing, 210016, China)

Abstract: Motivated by the fast growth of image databases, content-based image retrieval (CBIR) has received widespread research interest recently, where a typical user query is represented by a dynamic combination of visual and semantic descriptions of the desired image or class of images. Unfortunately, users often have difficulty specifying such descriptions. To alleviate those problems, users' queries are often refined interactively through mining the relevance of the feedback images, with the parameters combining the features automatically adjusted to adapt to the users' original needs.

The aim of this paper is to clarify some of the issues raised by this new technology by reviewing the current capabilities and limitations of the relevance feedback (RF) techniques from the viewpoint of machine learning. A

* 基金项目: 国家自然科学基金(69903006), 教育部高等学校骨干教师资助项目[教技司(2000)65号], 中国博士后科学基金(中博基金[1997]11号)

收稿日期: 2003-09-23

** 通讯联系人: E-mail: txy@ultratech.com.cn

concept CBIR framework with RF techniques embed in is proposed first to facilitate the description, and the basic components of a RF-based CBIR system are illustrated as well.

By using the proposed model, the paper proceeds with reviewing various RF techniques existing in the literatures from three aspects. The data which can be used as the source of RF-mining is investigated firstly. To analysis the obtained feedbacks, some assumptions about the distribution of the data must be made. We describe two kinds of probable distribution assumptions commonly seen in literatures, i. e., Gaussian distribution and mixture models, and the advantages and shortcomings of both assumptions are compared. Thirdly, we identify six challenges one may encounter in the practice of applying RF technique in CBIR context and point out that there is not such perfect algorithm that can deal with all the mentioned difficulties well at the same time.

Next, we focus on the classical RF methods, which mainly originate from the pattern recognition or machine learning area. Roughly saying, we survey the typical RF algorithms from two branches, i. e. the clustering-based and classifying-based algorithms. The main difference between the two kinds of algorithms lies in that the clustering-based algorithms care more about the understanding of the "query point", which is usually considered as the semantic representation of user's query in his or her mind, while the classifying-based ones try to find the class boundary directly from the image database using some prior knowledge about the statistical structures of the feedback data. However, it is worth mentioning that the boundary between the two branches of algorithms is soft in nature. The merits and shortcomings of the classical RF methods are also discussed.

Finally, several typical CBIR systems, either commercial or academic, with RF-techniques embed in are surveyed. The surveyed systems include such famous systems as QBIC, FourEyes, PicHunter, etc. The relevance feedback techniques adopted in those systems are identified and emphasized.

In summary, we think that human is a key factor in the running of the whole CBIR system and it would play a more important role in the operation of the system than now. RF techniques, as a way to incorporate man in loop, will continue to receive wide interests from researchers. The paper then suggests several future promising research directions through analyzing the close relationship between relevance feedback technique and the abstracting of image semantic, user modeling and soft computing.

Key words: relevance feedback, content based image retrieval, information retrieval

多媒体技术和 Internet 的发展给人们带来巨大的多媒体信息海洋,并进一步导致了超大型图像信息库的产生,如何有效的帮助人们快速、准确地找到所需要的信息,成了图像检索系统所要解决的核心问题.20 世纪 90 年代初,基于内容的图像检索(content-based image retrieval,CBIR)应运而生.CBIR 主要利用图像的低层特征(如颜色、纹理、形状、空间布局等)、导出特征(即从低层特征中导出的初级语义特征)及抽象特征(高级语义特征)来进行检索^[1],其中基于后两种特征的检索称为基于语义的图像检索.CBIR 尽管在某些特定的领域取得了成功,但由于低层特征往往无法完全把握人在图像相似性上的感受;图像检索需要用户更大程度的主动参与,从而对用户的应用水平提出新的要求^[2],这些原因给 CBIR 系统的推广和应用造成极大障碍.

尽管用户不易清楚表达其所需信息,但系统一旦向用户展示某些结果时,即使不十分准确,用户仍能轻易从中辨认出感兴趣的部分,如果用户能将其观察结果告知系统(而不必描述自己具体信息需求),信息检索系统就可以充分利用这些信息来提高检索效率,这就导致

了相关反馈技术的产生. 相关反馈技术能够屏蔽许多构造查询的细节, 使用户在不知道信息库内容构成和搜索环境的情况下, 也能构造出有用的查询. 检索过程被分割为更小的检索步骤序列, 便于逐步逼近用户感兴趣的主体区域. 同时, 用户对查询的控制能力也得到增强.

这种技术最初应用在文本检索领域, 已有 30 余年的历史, 然而直到 20 世纪 90 年代中期才开始得到研究者的重视. 1996 年, Minka 和 Picard 等人开始研究如何从用户的反馈中进行学习, 系统学习用户给出的正例和反例将图像特征空间重组为多个分离的特征组^[3]; 1998 年, Rui 等形式化地提出了图像检索中相关反馈的体系结构^[4], 从而正式揭开了图像相关反馈研究的序幕.

本文试图对图像相关反馈技术的研究现状进行调查总结, 重点讨论图像相关反馈的处理模型、新的算法与技术、及实现方法等方面的内容, 并对未来的发展趋势进行展望.

1 图像相关反馈的基本工作模型

为便于后面讨论, 首先提出一个图像相关反馈的概念模型(图 1). 图中实线代表信息流向, 虚线代表可能的系统参数调整. 一个典型的相关反馈过程一般分 4 个步骤: ①形成查询: 用户首先提交一个查询(该查询可能来自图像库, 也可能是其他方式, 如文本、草图等). ②查询处理: 系统先将用户查询转变为中间形式, 然后与图像库中的图像进行相似性比较, 根据相似程度或相似概率对查询结果排序, 并将前 k 个结果显示给用户. ③反馈循环: 用户对检索结果进行相关性评估, 并将评估结果提交系统. 系统的相关反馈模块根据用户的相关信息对系统进行相应调整(如修改查询表示、修改相似性计算函数参数、修改图像中间表示等), 然后用调整后的系统重新进行检索, 并再次将检索结果提交用户, 进入下一轮循环. ④终止: 当一定数量的循环后, 用户或者得到需要的信息, 或者失去耐心, 使得查询终止.

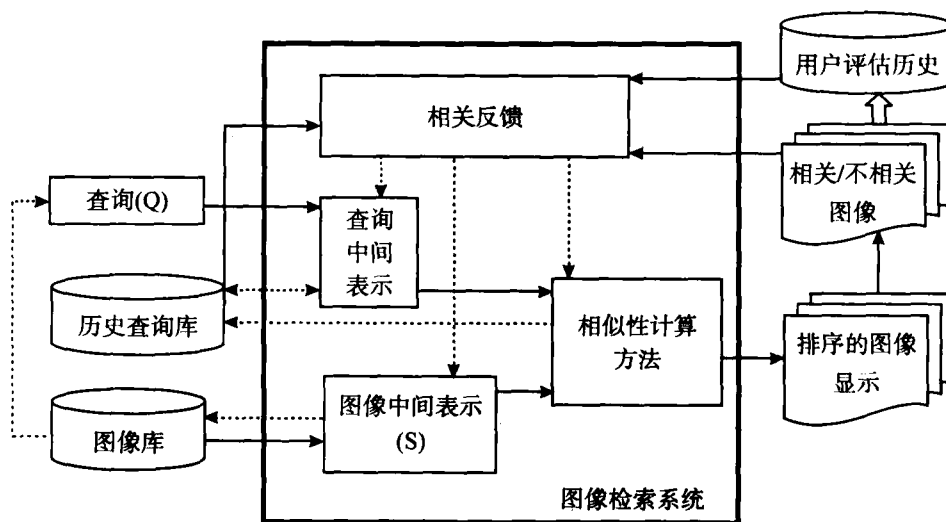


图 1 图像相关反馈的基本模型

Fig. 1 Basic image retrieval model with relevance feedback

为了提高相关反馈和查询的效率,图中用两个数据库分别保存用户的评估历史或用户个性化信息(profile)以及成功的查询历史,前者可用于挖掘用户的当前查询会话或历史会话的兴趣点^[5],从而使相关反馈技术能够用于个性化检索及信息过滤.后者可用于加快查询,例如已经成功的查询的有关参数可直接从库中调出,而不必重新计算.

2 图像相关反馈算法概述

图 1 中的相关反馈模块本质上是一个机器学习机制,用户的反馈是学习样本,系统根据对该样本集的学习,试图发现用户的真实查询意图,并根据不同的学习方法表示为不同的知识表示,如决策树、依赖网络、关联规则、线性回归模型等,用于指导下一轮的检索.

由于相关反馈学习机制的样本、学习任务及学习方法的不同,出现了各种相关反馈学习算法.本节基于上文的基本模型,从机器学习的角度对当前文献中出现的各类相关反馈算法进行调查和总结,希望能够找出这些不同算法之间的内在联系,从而探讨这些算法之间进一步相互组合和扩展的可能性.

2.1 训练样本的获取和种类 图像相关反馈中的学习样本主要是指用户的反馈以及从这些反馈中提取的各种统计信息.一般而言,用户的反馈的形式主要有如下几种:

2.1.1 仅给出正例(相关图像),常用于学习相关图像的分布密度^[6].

2.1.2 不仅给出正例(相关图像),同时还标明反例(不相关图像),这时常用于学习分类器,进而对图像库进行分类^[7].对是否考虑反例的问题,一直存在争论,支持者认为当图像类的密度部分重叠时利用反例有利于分清用户感兴趣的图像类,另外,当用户给出所有的正例后,唯有采用反例才能避免陷入检索空间的局部最小^[8].反对者认为,反例通常代表性不大、语义分散难以控制,使用反例只能徒增用户负担及计算复杂性,而对性能提高不大.

2.1.3 给出多个不同等级的主观标记,如指明某图像高度相关、相关、无意见、不相关、高度不相关等^[4].

2.1.4 给出一种比较,即某图像比某图像更相关或更不相关.[9]认为用户作出比较比挑选相关图像更为容易.

2.1.5 未标注样本.在图像检索中用户的标注只是少数,受 EM 算法思想的启发,一些研究者开始考虑如何利用图像库中大量未标注的样本进行学习,从而试图克服上文提到的小样本问题^[10].

2.1.6 算法主动请求用户标注.前面的方法都是用户标注,算法被动分析.[11]采用了主动学习的思想,挑选离 SVM 决策边界近的点作为“信息量最丰富点”让用户进行标注,这种方法可以加快对特征空间的搜索.

2.1.7 给出一种结果图像集的“布局”.Santini 从图像的浏览方式中受到启发,提出了一种新的用户界面,用户可以拖动图像进行聚类,来体现用户对图像之间相似性的理解,系统可通过用户给出的新的图像显示位置来学习相似计算函数的参数,从而逼近用户的“感觉相似性”^[12].

2.1.8 其他可供学习的用户信息.例如系统可利用用户选择相关图像的先后顺序、个数、彼此之间一致性程度等方面来学习用户的检索需求^[5].

2.2 图像/特征分布假定 各种学习算法一般都对查询或相关图像所代表的目标图像类或目标特征的分布模型都进行了强度不同假定.主要有如下几种:

2.2.1 单一正态分布模型 即假定图像特征之间彼此独立,目标图像服从以查询向量为中心,以训练样本的协方差的高斯分布.这是最常见的一种假定^[4,8,13],然而这种模型在许多情况下过于简单,同一语义的图像往往分布在特征空间的多处位置.[13]取消了图像特征独立性的假设,用SVD来解决样本的协方差矩阵不可逆问题.

2.2.2 混合高斯模型 这种模型考虑到了图像的局部分布特征,能较好反映图像在特征空间中的实际分布情况.[6]用Parzen窗法把每个相关图像都作为中心来估计分布,实际上是走向混合高斯模型的一个极端,缺点是模型过于复杂,从而影响其泛化能力.[14]采用了一个类似混合高斯模型的前向神经网络来表示正例图像分布:根据用户的反馈,用多个超球来逼近正例图像在特征空间中的分布.获取混合模型参数的标准方法是EM算法.然而在图像集分布未知时,EM算法反而可能使分类器的准确度下降,因此可以采取一些预处理措施来加以避免^[10].

2.3 相关反馈学习的挑战 无论采用何种学习机制,都受到信息检索这一特定环境的约束,这些约束给相关反馈中的学习问题带来巨大的挑战,主要表现在以下几个方面:

2.3.1 模型的不确定性 信息检索的不确定性本质,使得相关反馈的研究对象存在严重的不确定性,所谓的不确定性是指,一,模型未知或知之甚少;二,模型的结构和参数可能在很大范围内变化.目前大多数方法均假设正例图像符合单一正态分布,这种模型在很多情况下过于简单.

2.3.2 样本高维度问题 在机器学习中,小样本是针对样本集的大小和样本的维数之比而言,如果比值小于1,则称该训练样本集为小样本.研究表明,一次交互中用户反馈的图像数一般不超过20个,而图像数据的维度数一般都在几十到几百个,因此,用户给出的样本往往不足以代表正例的真实分布,这使得很多机器学习算法得不到稳定的和有推广能力好的解.

2.3.3 训练样本的非对称性问题 两个相似的特征值在特征空间上体现为它们在一个空间点附近;而与一个特征值不相似的值,可能位于除了这个特征值所在空间位置附近以外的整个特征空间上,这种正例和反例的不对称关系使得相关反馈不能简单的看作传统的二分类问题.[7]将正例和反例进行了不同的处理,对那些距离负反馈样本很近的图像进行惩罚(增加其与目标类的距离),并用多个方差较小的混合高斯模型来表示这些反例,从而减小反例对正例的影响.[15]将相关反馈问题看作一个“偏向单类问题”(Biased Classification Problem),即尽管用户反馈的类别数未知,然而用户总是偏向正例图像代表的单分类.从这一思想出发,他提出了一个类似于Fisher判别分析准则的优化准则,令反例相对于正例尽量分散,而正例内部应尽量聚集.

2.3.4 增量学习 信息检索中的学习不是一次完成的,而需要根据用户的上下文随时进行调整,同时,用户的反馈信息一般是多次输入,因此算法应当在当前反馈信息和历史信息之间取得平衡,这往往要求算法具有增量学习的能力^[16].

2.3.5 实时性问题 一方面,交互式信息检索要求对用户的响应必须及时,另一方面,由于用户的主观意图可能随时间、环境的改变而改变,因此只能在线学习,同时要求算法的收敛速度不能太慢.以上两方面对相关反馈学习算法的计算复杂度提出了要求,即必须在算法的准确度和算法的复杂性之间进行平衡.

在后面的讨论中将会看到,不同的相关反馈算法从不同侧面对上述问题进行了研究,然而还没有一个“完美”的算法能同时解决上面所有问题.

3 经典的图像相关反馈算法

从本节开始将更深入的对各种相关反馈算法进行分析讨论.在以下的讨论中,本文假定图像特征已预先抽取完成.经典的图像相关反馈算法将相关反馈问题看作一个聚类问题或分类问题.

3.1 聚类方法 多数的相关反馈算法都将相关反馈看成一个以查询向量为聚类中心,计算其 k 个最近邻居的聚类问题.一般假定查询向量点代表了用户的查询意图,而相关图像的选择代表了用户对图像的“感觉相似性”评价,因此相关反馈的任务就是:①查询点移动,使其尽量靠近相关图像,同时远离不相关图像;②修改查询,即调整特征空间中各图像特征的权重,使相似性计算尽量与用户的感受一致.这种方法明显带有文本检索中的相关反馈技术的痕迹.

3.1.1 查询点移动 查询点移动的方法大多数来自著名的 Rocchio 公式及其变形,这种方法的基本原理是,图像被表示为特征向量空间中的一个向量(点),系统根据用户反馈的正例和反例调整查询向量各分量的权重,使其向正例方向移动,同时远离反例集合,直觉上相当于捕捉用户查询意图.查询向量如下调整:

$$Q' = \alpha Q + \beta \left(\frac{1}{N_{R'}} \sum_{i \in D_{R'}} D_i \right) - \gamma \left(\frac{1}{N_{N'}} \sum_{i \in D_{N'}} D_i \right)$$

上式中 α 、 β 和 γ 是经过适当选取的常数, $N_{R'}$ 和 $N_{N'}$ 分别是相关反馈中正例 $D_{R'}$ 和反例 $D_{N'}$ 的个数, Q' 表示修改后的查询,随着反馈次数的增加, Q' 不断向最优的查询 Q_{opt} 逼近.

这种方法首先被 MARS 系统从文本检索移植到图像检索中^[4],随后的 MindReader 系统也采用了这种方法^[13],该系统以 Mahalanobis 距离为相似度算法,并从理论上证明了新查询点的最优化值为所有相关图像的加权平均.

3.1.2 调整特征向量权重 上面的方法没有考虑不同视觉特征在查询中所起的作用,如果能分析到这一信息,给重要的特征分配较高的权重,应该能提高检索效果.为了分析某个特征的相对重要性,直观想法是,应当给那些使正例良好聚类,而使正例和反例区分开的特征以较高的权重.因此需要分析所有正例在该特征对应的特征轴上的特征分量值的分布情况,方差越大,则正例在该特征上的不相似度越大,反之相似程度越大.越能代表用户的查询需求,这启发我们用方差的倒数来衡量特征权重(如 MARS 系统^[4]).随后, Ishikawa 等进一步将特征向量权重的调整问题形式化为一个优化学习的目标函数,使各正例图像距离查询向量的距离和最小化,通过求解这一优化问题, Ishikawa 等证明了最优化的权重调整方法应当与方差的平方的倒数成正比^[13].

3.1.3 聚类法存在的问题及其改进 上述的调整特征权重的方法尽管实现简单,具有一定的数学理论支持,然而在图像检索的具体环境下,仍然存在很多有待研究的问题:

①最近邻法只能检索出那些离查询向量较近的点,而对特征空间中语义相似而距离查询点较远的点则无能为力.

②基于最近邻算法的相关反馈技术往往需要对图像集进行穷尽搜索,而真实的图像集往往具有图像数量大、图像维度高的特点,这给系统的实时性能带来严峻的考验.解决这一

问题的一般思路是采用离线处理和在线学习相结合的办法,在离线阶段对图像集进行充分预处理,在检索时则根据实际情况动态调整^[4,17,18]。

3.2 分类方法 相关反馈问题看作一个二分类问题是另外一种较为普遍的学习模式。这种方法的基本思路是在检索过程中根据用户对图像的标注(相关或不相关),动态学习一种分类器,用于将图像集分为相关和不相关两类,并将相关部分作为结果展现给用户。具体构造何种分类器取决于人们对先验信息的了解程度,不同的目标图像类分布的假定导致了不同的分类算法。

3.2.1 概率分类器 在基于概率的分类器构造中,一般需要首先估计目标类密度函数(类概率密度的各种假定参前面 3.2 节),然后利用贝叶斯公式构造描述分类边界的判别函数。设用户标记的相关图像为 $R' = (R_1, R_2, \dots, R_k)$, 则目标类的类概率密度 $P(x|R)$ 可用训练样本来估计为: $P(R'|x) = \frac{P(x|R')P(R')}{P(x)}$, 式中 $P(R')$ 表示相关类的先验概率。这实际是一种“软分类”方法,后验概率被当做测试样本与某个分类的相似性的度量^[7]。

3.2.2 几何分类器 几何分类器的主要优点是无需对样本分布作任何假定,而利用某些启发规则来对目标空间进行迭代搜索。^[19]用 C4.5 算法递归地划分特征空间,直到每个划分内的向量点均属于同一分类为止,这样实际上构造了一棵决策分类树,可用于对图像集划分出相关图像子集,并在此候选集中用 KNN 算法求出与查询点最接近的 K 张图像作为结果返回。^[20]利用了层次自组织神经网络的拓扑保持特性来建立相关性模型。近年来不少研究者试图在相关反馈过程中采用支持向量机(SVM)来构造分类器^[11,15,21],SVM 基本上是一种二类分类器,其主要优点主要体现在能够从大量训练样本中挑选出相对较少的训练样本(支持向量),在高维空间中训练出推广能力强的非线性分类器^[16]。然而在相关反馈中,仅仅被动依赖用户提供的少量样本来进行判决往往效果并不如意,因此可以考虑主动请求用户标注那些靠近 SVM 边界的点(即信息含量最丰富的点),从而缓解小样本问题^[11]。SVM 法尽管取得了较好的实验效果^[11,15],然而其计算时间相对较长,因此是否适合直接用于相关反馈还有待研究。

3.2.3 关于分类器的讨论 不同的分类器代表了对事物的不同看法,对同一训练集训练多个不同的分类器不仅可以提高整体分类性能,而且不同的分类器在其各自的特征子空间中性能能够达到最优,因此还能够体现出较好的局部区分能力,这与混合分布模型的思想十分一致。基于这一思想,^[20]用一个算法自动为每个图像生成超过 45 000 个所谓“高度选择”的特征,然后用正例和反例基于高斯假设学习出多个弱二类分类器(约 20 个),最后用“AdaBoost”算法将这些弱分类器组合为一个强分类器,这种方法展现了组合的分类器的强大功能。

监督学习与非监督学习的结合能够互相弥补不足,也是一个重要的研究方向,即所谓的“半监督学习”(semi-supervised learning)。^[10]实际上就是这一思想的一个例子,先用监督学习法得到一个较弱的分类器,然后用非监督学习的方法,利用未标注图像来学习混合高斯模型的参数。另外一个例子是^[18],先用非监督学习得到图像聚类,然后在线利用监督学习法进行类别调整。

4 典型的基于相关反馈的图像检索系统

自 20 世纪 90 年代初以来, CBIR 日渐成为一个热门的研究领域, 人们建立了许多图像检索系统^[1], 为了增加检索性能, 不少系统都支持相关反馈技术, 其中较为典型的有:

4.1 QBIC^[22] 它是第一个商业 CBIR 系统, 其系统框架和采用的技术对以后的图像检索系统具有深远的影响. QBIC 支持基于示例的图像检索, 当用户在获得系统响应后, 可以进一步对检索图像的颜色、纹理、形状等做出反馈, 系统则根据用户的反应进行调整. 不足的是, QBIC 系统对用户的要求较高, 并非每个用户都清楚自己在颜色等方面的查询需求.

4.2 FourEyes^[3] 是 MIT 多媒体实验室开发的一套交互式图像浏览和检索工具. 考虑到没有任何单一的特征能够满足所有查询需求, 并且人类感觉具有相当强的主观性, 因此 Picard 等人提出一种“society of model”的方法, 将人的因素纳入图像标注和检索循环之中. 实验证明该方法在进行交互式图像标注时非常有效.

4.3 PicHunter^[9] 该系统试图对用户需求进行建模, 在假定用户检索的类型是“目标检索”的前提下, 采取主动学习的方式, 每次向用户提交的不是“最相关图像”, 而是能够使后续比较次数最少的图像, 让用户进行标注, 由此快速逼近用户需求. 该系统的特点是没有采用查询求精的方法, 而是利用贝叶斯规则, 基于概率分布的估计, 对可能的目标图像进行预测.

4.4 MARS^[4] 是 Illinois 大学开发的一套图像检索原型系统. MARS 系统的一个主要特色在于, 它并不试图发现一个单一的“最佳”特征表示, 而是关注如何将多种不同视觉特征组合为一个有效的检索体系, 使之能够根据不同的应用和用户进行自适应调整. MARS 系统的一个主要贡献在于形式化地提出了基于相关反馈的图像检索系统框架.

4.5 iFind^[23] iFind 系统将图像的语义特征和视觉特征进行了无缝集成, 它通过用户对检索的结果图像的评估, 不仅调整查询向量的权重, 而且动态修改正负反馈图像与对应关键字的关联权值. 初始的关键字可以通过外部信息源(如 Web)得来, 更多的关键字则是以半自动的方式从与用户的交互中获取, 随着交互的增多, 系统的词汇库便会不断增加, 同时关键字与图像之间的映射就越来越接近真实情况.

5 结束语

基于内容的图像检索技术经过多年来的发展, 已经取得一定成果, 然而要真正实现图像、语义、用户三者的融合, 还有很长的路要走. 作为 CBIR 中重要一环, 相关反馈技术代表了用户的参与, 符合“human in loop”的思想, 日益引起研究者的重视. 本文对当前图像检索领域的相关反馈技术研究状况进行了回顾和总结, 对一些有代表性的方法进行了较为详细的介绍和分类, 指出了目前这一领域已有的技术局限性和一些需要解决的关键技术问题. 在我们开发的“visual engine”图像检索中, 已经开始对其中一些关键问题进行了初步的探讨和实验.

展望未来的发展, 我们认为图像相关反馈技术在以下几方面的研究值得关注:

5.1 结合语义的图像相关反馈算法 图像的语义提取和图像检索是两个紧密相关的领域. 相关反馈是一种重要的向系统提供用户语义信息的手段, 在目前图像分割技术不是十分成熟的情况下, 可以将语义(如图像标注)和相关反馈进行充分的结合以促进检索效率的提高^[23~25].

5.2 结合用户模型的图像相关反馈算法 信息检索(information retrieval, IR)的本质就是交互式的,在这个过程中对用户的要求不可小视,一个理想的图像检索系统应该在为用户提供信息检索功能的同时,尽量减轻用户负担,使用户的操作更为简单和容易.因此有必要对用户模型进行深入研究,以建立更为有效的相关反馈模型^[5,9].

5.3 基于软计算方法的图像相关反馈算法 软计算技术允许在计算过程中出现不精确、不确定甚至部分准确的计算过程,这与图像检索的不确定性本质不谋而合,目前已逐步出现将软计算方法与相关反馈技术进行结合的趋势,包括模糊集理论^[26]、基于概率论的推理^[8]、神经网络^[27]在内的软计算方法已被应用于相关性建模.

5.4 更高效的相关反馈算法 随着互联网上图像的不断增多,使得 Web 上的图像检索成为一个重要的研究和应用领域,一些基于 Web 的 CBIR 系统,如 WebSeek^[28], ImageRover^[29]也纷纷采用相关反馈技术来提高检索性能,这促使我们研究更为快速、高效的相关反馈算法.

总之,可以预计,随着技术的不断更新和应用领域的不断扩展,相关反馈这种交互式的人机协同工作机制必将在图像检索和图像理解领域发挥更大的作用.

References

- [1] Amold W M, Marce W, Simone S, *et al.* Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000,22(12): 1 349~1 379.
- [2] Stefano M. Relevance:The whole history. *Journal of the American Society for Information Science*, 1997, 48(9): 810~832.
- [3] Minka T P, Picard R W. Interactive learning using a "Society of Models". *Proceedings of IEEE CVPR*, 1996: 447~452.
- [4] Rui Y, Huang T S, Ortega M, *et al.* Relevance feedback: A power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Video Technology*, 1998,8(5): 1~13.
- [5] Ruthven I. On the use of explanation as a mediating device for relevance feedback. *Proceedings of the 6th European Conference On Digital Libraries. ECDL*, 2002: 338~345.
- [6] Meilhac C, Nastar C. Relevance feedback and category search in image databases. *IEEE International Conference on Multimedia Computing and Systems*, 1999(1): 512~517.
- [7] Su Z, Zhang H J, Ma S P. Relevance feedback using a Bayesian classifier in content-based image retrieval. *Proceedings of SPIE Int Soc Opt Eng*, 2001,4 315: 97~105.
- [8] Vasconcelos N, Lippman A. Bayesian relevance feedback for content-based image retrieval. *IEEE Workshop CBAIVL*, South Carolina, 2000: 63~67.
- [9] Cox I J, Miller M L, Minka T P, *et al.* The Bayesian image retrieval system, PicHunter: Theory, implementation and psychological experiments. *IEEE Transactions on Image Processing*, 2000,9(1): 20~37.
- [10] Wu. Y, Tian Q, Huang T S. Discriminat EM algorithm with application to image retrieval. *IEEE Conference of Computer Vision and Pattern Recognition(CVPR)'2000*. South Carolina: Hilton Head Island, 2000: 1 222~1 227.
- [11] Tong S, Chang E. Support vector machine active learning for image Retrieval. *ACM Multimedia*, 2001: 107~119.

- [12] Santini S, Jain R. Integrated browsing and querying for image database. *IEEE Transactions on Multimedia*, 2000,7(3): 26~39.
- [13] Ishikawa Y, Subramanya R, Faloutsos C. Mindreader: Query databases through multiple examples. *Proceedings of the 24th VLDB conference*, 1998: 218~227.
- [14] Zhang Lei, Lin Fu-Zong, Zhang Bo. A neural network based self-learning algorithm of image retrieval. *Journal of Software*, 2001,12 (10): 1 479~1 485.
- [15] Zhou X S, Huang T S. Small sample learning during multimedia retrieval using BiasMap. *IEEE Conference of Computer Vision and Pattern Recognition*, 2001(1): 11~17.
- [16] Xiao R, Wang J C, Sun Z X, *et al.* An approach to incremental SVM learning algorithm. *Journal of Nanjing University(Natural Sciences)*,2002,3(2):152~157. (萧 嵘,王继成,孙正兴等.一种 SVM 增量学习算法. *南京大学学报(自然科学)*,2002,3(2): 152~157).
- [17] Wu P, Manjunath B S. Adaptive nearest neighbor search for relevance feedback in large databases. *Proceedings of ACM International Multimedia conference*, 2001: 202~209.
- [18] Lee C, Ma W Y, Zhang H J. Information embedding based on user's relevance feedback for image retrieval. *SPIE Photonic East*, 1999: 20~22.
- [19] MacArthur S D, Brodley C E, Shyu C. Relevance feedback decision trees in content-based image retrieval. *IEEE Workshop CBAIVL*, 2000: 68~72.
- [20] Tieu K, Viola P. Boosting image retrieval. *IEEE Conference of Computer Vision and Pattern Recognition*, 2000: 228~235.
- [21] Chen Y, Zhou X S, Huang T S. One-class SVM for learning in image retrieval. *IEEE International Conferences on Image Processing*, 2001: 7~10.
- [22] Niblack W, Barber R, Equitz W, *et al.* The QBIC project: Querying images by content using color, texture, and shape. *Proceedings of SPIE Electronic Imaging: Science and Technology*. CA: San Jose, 1993: 173~187.
- [23] Lu Y. A unified framework for semantics and feature based relevance feedback in image retrieval systems. *Proceedings of ACM MM2000*. Los Angeles: ACM Press, 2000: 31~38.
- [24] Chang S F, Chen W, Sundaram H. Semantic visual templates: linking visual features to semantics. *Proceedings of IEEE International Conference of Image Processing(ICIP'98)*3. Chicago, 1998: 531~535.
- [25] Carson C, Belongie S, Greenspan H, *et al.* Region-based image querying. *Proceedings of CVPR'97 Workshop on Content-based Access of Image and Video Libraries*. 1997: 42~49.
- [26] Freigui H. Interactive image retrieval using fuzzy sets. *Pattern Recognition Letters*, 2001,22: 1 021~1 031.
- [27] Laaksonen J, Koskela M, Oja E. PICSOM: Self-organizing maps for content-based image retrieval *INNS-IEEE International Joint Conference on Neural Networks*, 1999.
- [28] Smith J R, Chang S-F. Visually searching the web for content. *IEEE Multimedia*, 1997,4(3): 12~20.
- [29] Sclaroff S, Taycher L, Cascia M L. ImageRover: A content-based image browser for the world wide web. *Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries*. San Juan: Puerto Rico, 1997.