

# 基于内容的图像检索中的相关反馈技术发展

徐建军 吴玲达

(国防科技大学多媒体研究开发中心 长沙410073)

**摘要** 早期的基于内容的图像检索系统以图像处理技术为中心,研究重点集中在视觉特征的选择和提取方面,而没有充分考虑到人们在视觉方面的主观性和广泛使用的高层次语义概念和低层次视觉特征之间的相关性。为了解决上述问题,相关反馈这项技术被引入到基于内容的图像检索中来。本文介绍了相关反馈的发展,着重阐述了相关反馈技术的各种算法以及其在CBIR中的应用,并对相关反馈的发展方向进行了讨论。

**关键词** 基于内容图像检索,相关反馈,贝叶斯分类,SVM(支持向量机)

## The Development of the Relevance Feedback Technology in Content-Based Image Retrieval

XU Jian-Jun WU Ling-Da

(Multimedia Research and Development Center, National Univ. of Defence Technology, Changsha 410073)

**Abstract** The early content-based image retrieval system is image processing technology-centric, and the focus of the study is selection and extraction of visual features, which don't take the correlation between the visual subjectivity and high-level semantic concepts into account fully. To solve it, the relevance feedback technology is introduced to the content-based image retrieval system. This paper introduces the development of relevance feedback, and we put emphasis on illuminating various algorithms and their applications in CBIR. Finally, the prospect of this technology is also discussed.

**Keywords** Content-based image retrieval, Relevance feedback, Bayesian classifier, SVM

## 1 引言

随着多媒体技术以及 Internet 的迅速发展,信息的形式逐渐由单一的文本方式改变为图形、图像、动画、视频等多媒体信息为主的表现方式。人们对图像数据的查询需求日益增长,一些图像检索系统 QBIC<sup>[10]</sup>, Photobook<sup>[11]</sup>, VisualSE-EK<sup>[12]</sup>, Chabot<sup>[13]</sup>以及 UIUC 的 MARS 等相继被推出。

基于图像内容的查询,其内容包括图像的颜色、形状、纹理和语义等特征。其中图像的颜色、形状以及纹理被认为是同一层次上的特征,具有相对直观的特点。语义内容是较高层次上的特征,具有主观抽象的特点。因此从查询的层次可以分为以下三类:

1. 基于原始数据的查找。每幅图像是像素点简单有序的集合体,查找时相似性的度量准则是点对点的比较。这个层次的比较是非常具体的,只有在使用相对精确匹配计算时才使用。

2. 基于特征的查找。在基于特征的层次上描述图像。图像的特征就是一幅图像与其他图像不同的原始特性或根本属性。有的特征与图像的视觉外观相对应,具有原始性,如亮度、形状描述子、灰度纹理;还有一些特征缺少自然的对应性,如灰度直方图、颜色直方图、空间频谱图。不同的特征被划分在不同的特征矢量组。图像的区域划分基本上是根据矢量的不同特性。同一区域内的特征矢量具有相同的属性。数据库内的图像文件的特征矢量集以及抽取的方法事先抽取并保存。查询时,使用者给出的模板图像按照系统的特征抽取方法提取出目标特征矢量与图像数据库中的特征矢量进行相似性的匹

配。

3. 基于语义的查找。可以看作是基于对象的查找。查找对象包括的具体物体,发生的场景,图像所描述的情感色彩等都属于这个层次的查找。

这些系统中在检索过程存在着许多问题。其中一个问题是检索过程以计算机为中心,使得一些查询结果并不能完全满足用户的要求。主要原因是计算机视觉技术还不够完善,从图像中自动提取的特征主要还是低层次特征。另一个问题是在很多情况下,在以人为主的高层语义和以计算机为主的低层次图像特征之间存在着巨大的鸿沟,无法找到直接的映射关系。人的主观性常常被忽略,因为不同的人对同一事物的感知是不一样的,即使是同一个人不同的环境下的感知也是不一样的。图像检索系统的最终用户是人,因此从心理学的角度来捕捉人对图像内容的认知是相当重要的。为了把用户模型嵌入到图像检索系统中,并且增进人机之间的交互,最近几年在CBIR中引入了相关反馈机制。

## 2 相关反馈技术的引入

相关反馈在文本信息检索(TIR)中已经被广泛使用,主要是通过一种人机交互机制能够使计算机能够不断了解用户对查询结果的满意程度,并通过不断学习把输出调整到符合用户期望的状态。

相关反馈技术的运用至少希望达到下面的目的:(1)询问扩展(Query Expansion):可以对用户不够具体的初始查询进行细化、具体化,将已经检索的文档集不断缩小,最终得到相关度最大的文档集;(2)调整词权(Term Re-weighting):用户

的查询目的并不是十分明确时,系统可根据用户对初次查询结果的相关性判断通过调整词权改变词语在询问中的作用,从而改变检索效果。

### 2.1 经典的相关反馈算法

常用的方法就是 Rocchio<sup>[23]</sup>的算法,公式如下:

$$Q = \alpha Q + \beta \left( \frac{1}{N_{R_i \in D_R}} \sum D_i \right) - \gamma \left( \frac{1}{N_{N_i \in D_N}} \sum D_i \right)$$

$D_R$  是一组相关的文档,  $D_N$  是一组不相关的文档,  $\alpha, \beta, \gamma$  是常数,  $N_R$  和  $N_N$  分别是  $D_R$  和  $D_N$  中的数量。这项技术被应用到 MARS 系统中,并取得了很好的效果。

### 2.2 CBIR 中相关反馈的过程

在基于内容的图像检索中,一个典型的相关反馈的过程如下:

Step1) 对于用户提供的样本图片,系统根据相似矩阵和样本图片及图像数据库中图片的图像特征,给出一个基于相似度的排序列表,数据库中图片和样本图片如果在特征空间中越相似,则排序越靠前。

Step2) 用户进行相关反馈,在该列表中选择一组符合该次检索内容的正反馈图片和一组不符合检索内容的负反馈图片,然后提交给检索系统。

Step3) 系统根据用户提交的信息,来优化相似矩阵并提交给用户新的检索结果;返回 step2), 系统重新进行查询。

因此,在相关反馈中最重要的问题就是如何有效地利用用户提供的交互信息,并通过优化检索向量和相似矩阵来提高检索的准确度。所以从本质上来说,相关反馈在信息检索中是一种指导性学习的技术,用以提高系统的检索能力。

## 3 CBIR 中相关反馈的主要技术

最近几年美国 MIT、美国 Illinois 大学、荷兰 Amsterdam 大学等相继开展了相关反馈在基于内容的图像检索系统中的应用,并取得相当多的学术成果。相关反馈的主要研究成果可以主要分为以下几个方面。

### 3.1 基于修改查询向量或者距离判别标准

这种方法是通过调整低层各个特征、特征各维的权重来优化检索结果。Illinois 大学的 Yong Rui 等人将相关反馈的方法用于图像的检索中<sup>[1,4]</sup>。其基本思路是对在特征空间中对特征分布标准差大的特征降低权重,反之则增加权重,来提高检索的效果。为此他们提出了一种多层的图像模型,他们假设一个图像对象  $o$  有  $i$  个特征,用  $f_1, f_2, \dots, f_i$  表示,这些特征之间的权值分别为  $w_1, w_2, \dots, w_i$ ; 对于每个权值有可能有多种表示,每种特征  $f_i$  表示方法之间又存在着一定的权值  $w_{i1}, \dots, w_{ij}$ , 此外,对于每种特征的表示往往是高维的,每维之间也存在着一定的权值,它们可以用  $W_{i1}, \dots, W_{in}$  来表示。对于一个检索结果来说,比较的过程是从最低层特征表示的每一维开始,然后进行权加和,直到计算到最高层的每种特征之间的加权和。用户对查询返回的结果有5种评判选择:很相关,相关,没有意见,不相关,很不相关,这些选择对应于有一定的评分,分别为3, 1, 0, -1, -3, 考虑到这些评分后对每个权值进行动态更新。权值更新分为特征权值的更新和特征向量之间权值的逐级更新。这种方法的特点是通过动态调整权值的方法来缩短用户高层语义概念和低层特征之间的差距,经过多次更新后就可比较准确地得到用户的查询意图。

在文[22]中作者仍然基于 Rui 的多层图像模型,不同的是在系统中引入了查询时刻的概念,即相关反馈是一个不断

迭代的过程,其具有时序特性。文章基于的假设是针对同一个查询主题,不同用户在不同时刻提供的查询实例会有所不同,但在某种程度上这些查询实例间会有一些的相关性,例如  $t$  时刻提供的正例和  $t-1$  时刻提供的正例很可能在某些视觉特征上具有一定的共性。

虽然基于矢量模型已经实现了多种方法,但其固有的问题就是低层特征并不能很好地完整表达出图像的语义内容,而用户关心的恰恰就是语义内容(如图像中的一个物体或某个区域而不是背景),因此反馈图像可能在语义内容上是非常类似的,但是可能由于背景或其它因素而在低层特征上差别很大,因此仅仅使用低层特征在表示用户的反馈和描述其意图时并不是很有效。

### 3.2 调整图像数据库的分量和类间关系

因此文[7,8]把语义内容加入到图像检索的相关反馈中。文[7]提出的框架尝试用相关矩阵把语义信息嵌入到低层特征中。采用聚类分析的方法,首先在原有的特征空间中对图像进行聚类,根据用户对训练样本的评价,把特征空间进行划分,形成不同的图像分类,使得相关图像落入同一类别或者距离相近的类中。在图像查询过程中根据类与类之间的距离确定图像间的相似程度。在此框架中,在图像聚类中的语义相关性从用户的反馈中学习,并用来改善检索性能。但是问题是在进行图像的聚类时该如何选取图像的粒度,即应该基于整幅图像还是基于区域,如果是基于区域,那么该如何对图像进行有效的分割,这都给系统的实现增加了一定的困难。

文[9]中构造了一个语义网络,每幅图像都以不同的关键字和权重加以描述。关键词用于表达图像的语义,每幅图可以有多个关键字描述,图像语义越简单,则用以描述的关键字越简单,关键词的权重反映了该词对图像语义描述的准确程度;在此基础上实现语义相关反馈,然后将该算法与视觉特征相结合,实现两个层次上的相关反馈,构造出一个统一的相似性度量方程。但是该方法的前提是对搜集来的存储在数据库中的图像进行预先标注或是一些已经分好类的图像数据或是网络上带有 TAG 标记的图像,之后还要对语义网络进行训练,这就给系统和用户带来了不必要的麻烦和开销。

### 3.3 系统是否具有记忆功能

经典的相关反馈方法的一个缺点就是在相关反馈过程中获得的知识不能被记忆来改善检索性能,即用户不得不过相同的经常是枯燥的过程才能得到相同的反馈结果,知识不能不积累,给用户的查询造成了很大的麻烦。为了克服这些缺陷,人们使用学习方法来记忆用户在相关反馈中的主观性,文[7]中从用户反馈中以图像聚类之间语义相关的形式来学习用户的主观性并记忆在一个相关矩阵中。实验表明使用从以前查询中学习到的知识可以在接下来的查询中有效地减少迭代的次数。文[14]文中提出一个基于半监督的模糊 C 平均聚类方法来从先前的检索经验中来学习类分布(元知识),使用模糊规则,用元知识合并到基于统计特征的相关反馈中来改善检索性能,并取得了良好的效果。

这种方法的问题就是不能保证积累下来的知识都对检索有帮助,必须要考虑一些特殊的情况,如用户在多次反馈后没有找到自己想要的图像,在标记中失去耐心而胡乱标记一番,那么得到的这种经验知识就必须舍弃。但是如何对用户的行为模式进行研究,还有待于研究。

### 3.4 基于模式识别中分类的方法

以前的相关反馈方法大多把其看作是一个机器学习的过

程,Rocchio 的公式就是一个简单的学习方法,此外还有决策树学习、基于核的学习,以及基于神经网络的学习等一些方法。然而用户在查询过程的耐心是有限的,经常不愿意提供大量的反馈样本,这样可供训练的样本就很少,因此现在紧迫的问题就是如何在非常高维特征空间中从小训练样本中进行学习,这就使得许多的学习方法如决策树学习以及人工神经网络都不再适合 CBIR。

此方法的基本思想是根据用户反馈的信息(样本),进行统计推断。从本质上认为图像检索系统也是一个分类问题,把交互的过程看作是一个训练过程,用户认为比较满意和不满意的图片可以分别看做正例训练数据和反例训练数据,而进一步查询的过程是一个分类问题。

Cox 最早把这一理论运用到图像检索系统 PicHunter<sup>[2]</sup>中并取得了良好的检索效果,此系统的缺点是没有考虑反例起到的作用。Vasconcelos<sup>[3]</sup>等研制的图像检索系统中采用了基于区域的贝叶斯的方法,由于采用了高斯混合模型估计先验概率,因此计算复杂度比较高,尤其是当图像数据库不断增大的时候需要重新估计高斯混合模型的参数。文[5]中采取的假设和 Cox 的相反,认为该图片与之对应的所有正反馈图片构成了一个拥有共同语义含义的图像集合,并通过对于该图片集的高斯拟合构造出该检索样本图片在特征空间的贝叶斯分类器。该分类器的参数反映了语义类在特征空间中的分布信息。同时,利用用户的正反馈信息来不断更新分类器的参数使得检索效果不断提高。对于负反馈信息,系统中假设负反馈样本间是彼此独立的,通过构造一个简单的混合模型来拟合出负反馈样本在特征空间的分布,并利用一种类似于“挖洞”的过程将负反馈信息用于检索过程。

此外由于类器可以把所有的例子分为两类:相关组和非相关的一组,这样就可以把许多模式识别领域的方法借鉴到相关反馈中来,如线性分类器<sup>[20]</sup>、NN 分类器<sup>[21]</sup>、贝叶斯分类器、支持向量机<sup>[15,17]</sup>等。值得一提的是,基于支持向量机(SVM)的分类方法具有较好的学习性能,SVM 并不像其它方法那样假设样本的分布情况,它基于结构风险最小化原则,能够对小样本情况下的学习问题给出最优解,并且在样本趋于无穷时可以保持良好的一致性,SVM 就是通过核函数在原空间或经投影后在高维空间中构造最优分类面,将给定的属于两个类别的训练样本分开,构造超平面的依据是两类样本离超平面的距离最大化。从数学形式来看,采用高斯核函数的 SVM 分类器具有 RBF 分类网络的形式,这种模型可以较好地描述混合高斯分布情况下的样本分布特点。

**总结** 在本文中,我们简要介绍了相关反馈技术的发展,对为什么要在 CBIR 系统引入相关反馈机制进行了探讨,详细介绍了 CBIR 中的相关反馈技术并对其进行了分类论述。虽然相关反馈技术在 CBIR 中取得很多的成果,但在此领域中仍然还有很多急需解决的问题,例如如何从小样本的训练数据中进行快速的学习并对图像进行检索,如何在用户的反馈过程中积累知识并对知识的有效性进行判断等等,总之,这些问题还有待于解决,这也是众多学者们努力研究的方向。

## 参 考 文 献

1 Rui Y, Huang T S, Ortega M, et al. Relevance feedback: A power tool for interactive content-based image retrieval. IEEE

- Trans on Circuits and Video Technology, 1998,8(5)
- 2 Cox I, Miller M, Omohundro S M, et al. Pichunter: Bayesian relevance feedback for image retrieval system. In: Int'l Conf on Pattern Recognition. Vienna, Austria, 1996. 361~369
- 3 Vasconcelos, Lippman A. Bayesian representations and learning mechanisms for content based image retrieval. In: SPIE Storage and Retrieval for Media Databases. San Jose, CA, 2000
- 4 Rui Y, Huang T S. A Novel Relevance Feedback Technique in Image Retrieval. ACM Multimedia, 1999
- 5 Su Zhong, Zhang Hongjiang. Using Bayesian Classifier in Relevant Feedback of Image Retrieval
- 6 Qian Fang, Li Mingjing, Zhang Lei. Gaussian Mixture model for relevance feedback in image retrieval
- 7 Lee C, Ma W Y, Zhang H J. Information Embedding Based on User's relevance Feedback for Image Retrieval: [ Technical Report HP Laps]. 1998
- 8 Wood M E J, Campbell N W, Thomas B T. Iterative refinement by relevance feedback in content-based digital image retrieval. In: Proc. of the 6th ACM Int'l Conf on Multimedia'98. Bristol, England, 1998
- 9 朱兴全,张宏江,刘文印,等. iFind:一个结合语义和视觉特征的图像相关反馈检索系统. 计算机学报, 2002,25(7)
- 10 Flickner M. et al. Query by Image and video content: the QBIC system. IEEE Computer, 1995,28(9):23~32
- 11 Pentland A, Picard R, Sclaroff S. Photobook: Content-based manipulation of image databases. International Journal of Computer Vision, 1996,18(3): 233~254
- 12 Simth J, Chang S-F. VisualSEEK: A fully automated content-based image query system. In: Proc. of the 4th ACM Multimedia Conf. Boston, 1996. 87~98
- 13 Ogle V, Stonebraker M. Chabol: Retrieval from a relational database of Images. IEEE Computer, 1995, 28(9): 40~48
- 14 Concepts Learning with Fuzzy Clustering and Relevance Feedback
- 15 张磊,林福宗,张钹. 基于支持向量机的相关反馈图像检索算法. 清华大学学报(自然科学版),2002,42(1)
- 16 张磊,林福宗,张钹. 基于前向神经网络的图像检索相关反馈算法设计. 计算机学报,2002,25(7)
- 17 Tong S, Chang E. Support Vector Machine Active Learning for Image Retrieval. ACM Multimedia 2001, Ottawa, Canada
- 18 MacArthur S D, Brodley C E, Shyu C-R. Relevance feedback decision trees in content-based image retrieval. IEEE Workshop on Content-based Access of Image and Video Libraries, 2000. 68~72
- 19 Laaksonen J, Koskela M, Oja E. PicSOM: Self-Organizing Maps for content-based Image Retrieval. In: Proc. of Intl. Joint Conf. on NN, July, 1999
- 20 Wu Y, Tian Q, Huang T S. Discriminant EM algorithm with Application to Image Retrieval. IEEE CVPR, South Carolina, 2000
- 21 Wu P, Manjunath B S. Adaptive Nearest Neighbor Search for Relevance Feedback in Image Database. ACM Multimedia conference, Ottawa, Canada, 2001
- 22 段立娟,高文,林守勋,等. 图像检索中的动态相似性度量方法. 计算机学报,2001,24(11)
- 23 Rocchio Jr J J. Relevance Feedback in Information Retrieval. In The SMART Retrieval System: Experiments in Automatic Document Processing (Salton, G. eds), Prentice-Hall, 1971. 313~323