

基于内容图像检索系统的性能改进方法^{*}

方艳梅¹, 杨 灿², 肖自美¹, 陈跃峰¹

(1. 中山大学电子与通信工程系, 广东 广州 510275;

2. 广东省信息中心, 广东 广州 510000)

摘 要: 在通用型的CBIR原型系统的基础上, 首先讨论CBIR系统的性能指标, 并对CBIR系统的检索性能改进方法进行了研究; 然后重点讨论了异种特征的组合查询和相关反馈两种性能改进技术, 并使用原型系统对以上两种技术进行了测试, 使图像检索的查准率和和查到率得到明显改善, 并有效地增强了CBIR系统的灵活性.

关键词: 基于内容的图像检索; 相关反馈; 组合特征查询

中图分类号: TP919.81 **文献标识码:** A **文章编号:** 0529-6579 (2001) 05-0039-05

随着多媒体技术和 Internet 网络技术的迅速发展, 数字图像的使用日益广泛, 成为信息社会中的主要信息资源之一. 基于内容的图像检索 (content-based image retrieval, CBIR) 技术为图像检索提供了一种有效而便捷的方法, 是当前多媒体技术中重要的研究课题. 图像数据具有难以用符号化的方法描述的信息线索, 如图像的颜色、图像中对象的形状等. 如何使系统直接从图像中获取信息线索, 并将这些线索用于数据库中的图像检索操作, 帮助用户从数据库中检索出合适的图像信息对象, 这就是 CBIR 的主要内容.

目前已经推出一些产品. IBM 公司的 QBIC (query by image content) 是第一个商业性的 CBIR 系统^[1]. Virage 是 Virage 公司的基于内容图像搜索引擎^[2]. 新加坡大学和 MIT 媒体实验室合作开发了 Photobook 系统等. 但是由于 CBIR 系统的检索性能有待于更进一步的研究. 本文对 CBIR 技术的检索性能进行了深入研究, 并重点讨论了两种改进 CBIR 检索性能的技术: 异种特征的组合查询和相关反馈. 在我们已开发的通用型的 CBIR 原型系统上对以上两种技术进行了测试. 实验结果表明, CBIR 系统的检索性能得到明显改善.

1 CBIR 系统的性能指标

CBIR 技术基于图像的视觉特征来进行图像检索, 它能在图像集合中找出符合指定视觉特征

描述的图像. 不同与一般字符检索的精确匹配, 基于内容的图像检索采用相似匹配技术, 系统首先分析出用户查询中包含的视觉特征, 然后通过图像特征的匹配计算, 检索出与查询特征最相似的若干图像, 并将其按相似程度由大到小排列, 返回给用户.

衡量 CBIR 系统性能好坏的最重要的指标是检索结果中相关图像 (即符合用户查询期望的图像) 的比例. 可以用查准率 (precision) 和查到率 (recall) 来表示^[3]. 前者反映出查到的图像是否都是相关的图像, 可以用于衡量特征相近而语义无关的错误率; 后者指是否所有语义相关图像都被检索出来, 可以用来表示语义相关而特征相左的错误率.

令 I_t 代表检索结果的图像总数, I_r 为检索出的相关图像数目, I_v 是整个图像库中所有相关图像的总数, 则有:

$$Pr = \frac{I_r}{I_t}; Re = \frac{I_r}{I_v}$$

2 通用型 CBIR 原型系统及其基本性能

对于通用型的 CBIR 系统来说, 为同时支持各种类型的图像, 系统多使用颜色、形状、纹理等通用特征. 通用特征的通用性使其语义描述能力受到影响, 其检索性能要差于专门为某一类图像设计的专用特征.

^{*} 基金项目: 广东省自然科学基金资助项目 (963037)

收稿日期: 2000-12-07; 作者简介: 方艳梅 (1966-), 女, 工程师; 通讯联系人: 肖自美.

我们开发了一个通用型 CBIR 原型系统来测试 CBIR 技术的检索性能, 该系统分别支持了颜色、颜色结构、主颜色、形状和纹理这五种通用的图像特征. 我们从图像素材光碟中选择了各类图像共 435 幅组成原型系统的图像库, 其中包括车辆、人物、贝壳、鱼类、珊瑚、树木、风景、建筑、古迹、食物、纹理图案、抽象画等图像. 本文的测试是基于这些图像来进行的, 每次检索系统返回图像的数目 $k = 15$.

首先测试单个通用特征对各种图像的检索能力, 具体算法在文献 [4] 描述. 表 1 和表 2 列出了单个特征检索各类图像的查准率, 黑体数字表示某类图像获得的最大查准率, 其对应的特征就是对这类图像的最佳特征. 表 1 上半列中列出

的各类图像有较好的检索结果, 这是因为表 1 中的图像不但语义简单明确 (基本上都可以注释为单个物体), 而且构图上都比较简单. 如摩托、汽车、人物、贝壳、鱼类等均是简单背景下的单个物体的图像, 而树木、山脉、河流等图案基本上由大色块构成. 从构图简单的图像中提取出来的特征描述通常能较好地对应图像中的语义, 对于这些图像, 基于某个特征的检索通常有较为理想的结果. 相反, 表 1 下半列的图像普遍含有较为复杂的语义内容 (如一个场景或一组物体), 构图上也相当复杂或抽象, 其特征内容与语义内容之间存在较明显的语义断层, 故其查准率普遍较小.

表 1 单个特征检索的查准率 ($k = 15$)

Tab.1 The precision for retrieving single feature ($k = 15$)

特征	摩托	汽车	人物	贝壳	鱼类	树木	山脉	河流	纹理图案
颜色	0.40	0.37	0.56	1.00	0.60	0.73	0.93	0.78	0.08
颜色结构	0.55	0.20	0.23	1.00	1.00	0.40	0.73	0.61	0.21
主颜色	0.53	0.08	0.09	1.00	0.93	0.41	0.80	0.60	0.07
形状	1.00	0.87	0.61	0.63	0.68	0.33	0.40	0.34	0.53
纹理	1.00	0.97	0.81	0.74	0.55	0.53	0.53	0.55	0.98
	建筑	古迹	雪景	河滩	珊瑚	食物	抽象画		
颜色	0.20	0.37	0.36	0.40	0.54	0.32	0.33		
颜色结构	0.25	0.53	0.40	0.33	0.32	0.20	0.20		
主颜色	0.23	0.58	0.19	0.23	0.09	0.23	0.21		
形状	0.20	0.33	0.41	0.42	0.10	0.25	0.08		
纹理	0.40	0.21	0.34	0.45	0.22	0.22	0.09		

3 CBIR 系统的性能改进

为了改进通用型 CBIR 原型系统的检索性能, 我们在系统中引入异种特征的组合查询 (composite query) 和相关反馈 (relevance feedback) 的技术, 取得了有意义的结果.

3.1 异种特征的组合查询

不同的图像特征有各自的矢量表示方法和相似性测度算法. 因此, 基于内容的检索通常是采用单独的一种特征来进行的. 单个通用特征的特征和语义之间存在相当的语义断层, 某些语义上完全不相关的图像的特征矢量在特征空间中的位置却相当接近, 这样看起来不相似的图像基于这个特征却会得到很好的相似性测度. 当单个特征不能够为检索提供足够的区别性信息时, 系统会给出一些错误的检索结果.

同时利用多种特征进行异种特征的组合查询可以使上述情况有所改善. 这是因为, 两个语义上不相关的图像虽然会在某个特征上有好的相似性测度, 但同时和其它特征上也取得好的相似性测度的几率是相当小的. 另外, 异种特征的组合查询使用户可以更灵活、更有效地表达他们的查询要求. 比如, 用户可能希望查找在颜色和形状上都与查询示例相似的图像, 这时, 用户就可以向系统提交一个组合查询.

异种特征的组合查询^[6]的过程如图 1 所示. 系统先要将组合查询分解为若干基于单个特征的查询, 单独进行检索. 对所有由单个特征检索返回的结果图像, 系统依照特定算法把它们在每个特征上的相似性测度进行合并^[2], 并按合并后的综合相似性测度对这些图像重新排序, 从中选出 k 个具有最大相似性的图像作为检索结果序列返

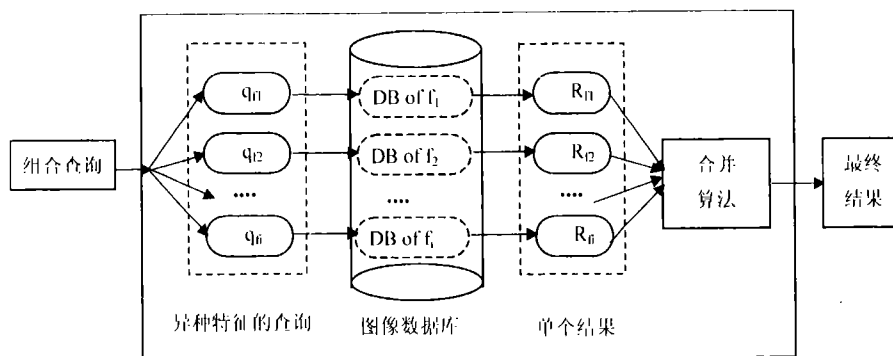


图1 组合查询的过程

Fig.1 The procedure of composite query

回给用户。

在我们的 CBIR 原型系统中使用线性组合的算法来合并单个检索的结果序列。它通过对各种特征的相似性测度进行加权的线性运算来生成综合相似性测度。

合并前要先对单个特征的相似性测度进行特征间归一化(inter-feature normalization)处理。这是因为不同特征的相似性测度算法得出的相似性测度的值域不一定相同。某些特征的相似性测度(如颜色直方图)的值域为区间[0,1],而另一些特征(如纹理分布的欧氏距离)则不是。很明显,必须统一各特征的相似性测度结果的取值范围。

在归一化处理后,在组合查询中各特征的地位是客观平等的。在此基础上,我们可以进一步为各特征分配不同的权重值,称为特征间加权(inter-feature weights)。特征权重的分配可以从两个因素来考虑。

客观上,不同特征的检索性能有差别,例如一般情况下,基于颜色的查询会比基于形状的查询有更高的准确率。当然,各种特征的检索能力的差别最终取决于它们应用在何种领域的图像上。系统的设计者基于数据库中保存的图像预先测试各种特征的检索能力,然后为准确率高的特征分配大的权重。

主观上,提交查询的用户对各特征有不同的重视程度。例如,一个基于纹理和颜色的组合查询,用户可能更关心纹理信息的匹配,在纹理信息的前提下,用户希望对颜色进行一些考虑,在整体的考虑上,纹理的份量达到80%,而颜色只占20%。这时,用户可以指定纹理特征的权

重为8而颜色特征的权重为2。可见,特征间加权可以使用户更精确地描述其信息需求。

我们可以用数学方法来描述线性组合。设 Q 是一个查询图像而 I 是数据库中的一个图像,图像有 N 种特征表示。是 Q 与 I 之间在第 i 种特征上的规一化的相似性测度,是分配给第 i 种特征的权重值,则在进行 N 种特征的组合查询时, Q 与 I 之间的综合相似性测度根据以下公式来计算:

$$S_i = \sum_{i=1}^N w_i S_i / \sum_{i=1}^N w_i$$

下面给出应用异种特征组合查询的实验结果。

我们使用两种不同方法来进行异种特征的组合。一种方法是将颜色、颜色结构、主颜色、形状和纹理这5种特征以相同的权重值进行等权组合。另一种方法是基于表1的数据,选出对该类图像的查准率最大的两个特征作为该类图像的较优特征,将其进行组合,并按照它们的查准率的比值来分配其权重值。以汽车图像为例:纹理和形状特征的查准率分别为0.97和0.87,是其5种查准率中最大的2个值,所以汽车图像的较优特征组合为纹理与形状特征按权重值97:87的组合。表2将这2种特征组合方法的查准率与最优的单个特征查准率进行了比较。可以看出,5个特征的等权组合方法,无论对简单图像或复杂图像,都基本不能提高查准率,对某些图像来说,其查准率甚至比单个特征所能获得的查准率更小。而2个较优特征的加权组合方法则能够明显地改善对复杂图像的检索(见表2下半列),在用于简单图像检索时查准率也略有提高(见表2上半列)。

表2 不同特征组合模型的查准率 ($k=15$)Tab.2 The precision for heterogenous features of query ($k=15$)

特征	摩托	汽车	人物	贝壳	鱼类	树木	山脉	河流	纹理图案
最优的单个特征	1.00	0.97	0.81	1.00	1.00	0.73	0.93	0.78	0.98
5个特征的等权组合	1.00	0.85	0.82	1.00	1.00	0.85	0.87	0.84	0.71
2个特征的加权组合	1.00	0.97	0.89	1.00	1.00	0.74	0.92	0.95	0.97
特征	建筑	古迹	雪景	河滩	珊瑚	食物	抽象画		
最佳的单个特征	0.40	0.58	0.41	0.45	0.54	0.32	0.33		
5个特征的线性组合	0.42	0.53	0.43	0.45	0.56	0.60	0.22		
2个特征的加权组合	0.67	0.58	0.79	0.73	0.65	0.78	0.51		

从测试数据可以得出以下结论:

(1) 使用异种特征组合查询时,关键在于如何分配特征的权重值.只有正确地设定各特征的权重值,才能达到改进检索性能的目的.否则,会导致检索结果更差.

(2) 当某个特征就能够较有效地检索图像时,将这种特征与其它特征进行组合查询的效果不大.如果对某类图像来说,各种特征的检索能力都不突出时,将这些特征按正确的权重进行组合后,检索性能有明显的改进.

3.2 相关反馈

通过异种特征的组合查询,用户可以根据自己关心何种图像内容来选择检索时要使用的特征种类以及特征的组合方式.但通用CBIR系统的用户很可能并不知道怎样将检索目标对应为一个组合特征模型,特别是在特征种类较多时,用户很难评估特征权重的不同比例对查询结果的影响.另外,在很多情况下,用户的查询目标很难通过一幅或几幅示例图像来精确表达.检索系统需要一种手段以更大程度地利用用户的知识来辅助查询.信息检索领域的研究已经证明:让用户交互式地参与检索的整个过程,使系统充分利用用户的知识能力,可以明显地改进检索的性能^[5~7].

相关反馈就是一种由用户指引计算机去找到相关信息的机制.其过程是,系统按用户的查询要求检索出一组图像返回给用户.用户浏览结果集合,挑选出满意的和不满意的结果,并把对结果的评价反馈给系统,系统从用户的标注中分析出用户的查询需要,并自动调整查询特征,形成一个新的查询.如此循环,直到用户满意查询结果为止.可见,它是一个逐步求精的过程,通过特征调整、重新匹配的循环过程来定位用户所期

望的图像.用户不必负责分配特征的权重值,只要告诉系统对检索到的图像满意或不满意即可.系统通过分析用户的反馈来了解当前用户最感兴趣的是哪些特征,并在检索过程中动态调整特征的权重值来精化查询.

下面将讨论如何根据用户的相关反馈来调整特征的权重.设用户的组合查询中组合了 r 种特征.系统基于单个特征 f_j 检索出 k 个最相似图像的集合: $RT^j = [RT_1^j, \dots, RT_i^j, \dots, RT_k^j]$,然后合并得到最终的检索结果: $RT = [RT_1, \dots, RT_i, \dots, RT_k]$ 返回给用户.用户对 RT 中每个图像的相关性进行判断,将其分成:相关图像、不相关图像、部分相关(即并非完全无关但又不是很相似)图像.我们按以下公式定义图像 RT_i 的相关系数 $Score_i$:

$$Score_i = \begin{cases} 1, & \text{相关} \\ 0, & \text{部分相关} \\ -1, & \text{无关} \end{cases}$$

用户的反馈信息 $Score = [Score_1, \dots, Score_i, \dots, Score_k]$ 返回给系统.系统要根据用户反馈重新计算特征 f_j 的权重值 w_j ,该值应该反映出特征 f_j 的检索与用户的检索需求之间的对应程度.

系统首先求出 RT_j 中的第 i 个图像 RT_i^j 的相关系数 $Score_i^j$:如果 RT_i^j 出现在 RT 中,则 $Score_i^j$ 就等于用户为该图像指定的相关系数;如果 $Score_i^j$ 落在 RT 之外,我们认为它是部分相关的,其 $Score_i^j$ 为0.特征 f_j 的权重值 w_j 按下面公式来计算:

$$w_j = \begin{cases} \sum_{i=1}^k Score_i^j, & \sum_{i=1}^k Score_i^j > 0 \\ 0, & \sum_{i=1}^k Score_i^j \leq 0 \end{cases}$$

可以看出，当同时出现 RT^i 在和 RT 这两个图像集合中的相关图像数目越多时，特征 f_j 的权重值 w_j 就会越大。也就是说，特征越能反映用户的信息需求，在检索过程中的重要性就越大。

系统按照计算出来的权重值重构查询特征，进行下一次的检索。重复以上过程，查询被逐渐精化而接近用户的期望。

表 3 列出了运用相关反馈技术的测试数据。表中的数据是不同图像类型的查准率的平均值，反映了系统整体上的检索能力。相关反馈能够提高系统检索的准确性，查准率随着反馈的次数的增多而增大。反馈的次数没有限制，原则上是直至用户满意为止。从表中数据可以看出，第一次使用相关反馈的效果最明显，随着反馈次数的增多，查准率的增加速度逐渐放缓。

表 3 相关反馈的测试结果 ($k = 15$)

Tab.3 Test results of relevance feedback ($k = 15$)

	初始查询	1 次反馈	2 次反馈	3 次反馈
查准率	0.75	0.81	0.83	0.84

4 结 论

本文研究了 CBIR 系统的性能改进方法。实验证明，使用异种特征的组合查询和相关反馈技术能够有效地增强 CBIR 系统的灵活性，改进系统的检索性能。使用异种特征的组合查询的关键在于正确设定各特征的权重，该方法对复杂图像

的检索性能有明显的改善。当用户使用不恰当的特征权重值来组合其查询时，系统能够通过相关反馈技术来调整特征权重值，更有效地组合异种特征，来提高系统检索的准确性。反馈的次数没有限制，原则上直到用户满意为止，但一般使用 3 次反馈就已接近能获得的最优性能。

参考文献：

[1] NIBLACK W, BARBER R, EQUITZ W. The QBIC project: Querying images by content using color, texture and shape[J]. Storage and Retrieval for Image and Video Database, 1996, 173 – 187.

[2] PENTLAND A, PICARD R W, SCLAROFF S. Photobook: Tools for content-based manipulation of image databases [J]. Storage and Retrieval for Image and Video Database II, 1996, 34 – 47.

[3] RUI Y, HUANG T, MEHROTRA S. A relevance feedback architecture for content-based multimedia information systems[J]. Workshop on Content Based Access of Image and Video Libraries, Porto Rico, June 1997.

[4] 陈跃峰, 肖自美, 杨灿. 基于内容的图像检索系统的关键技术及其实现[J]. 计算机工程与应用, 2000 (10): 17 – 21.

[5] Simone Santini, Ramesh Jain. Beyond query by example [C]. ACM Multimedia'98, Bristol, UK, 345 – 350.

[6] SATON G, MCGILL M J. Introduction to modern information retrieval[M]. McGraw-Hill Book Company, 1983.

[7] COX I. PicHunter: Bayesian relevance feedback for image retrieval, Proceedings of 13th Int. Conf. on Pattern Recognition, Vienna, Austria, 1996.

A Performance Modified Method of Content-Based Image Retrieval System

FANG Yan-mei¹, YANG Can², XIAO Zi-mei¹, CHEN Yue-feng¹

(1. Department of Electronic and Communication, Zhongshan University, Guangzhou 510275, China;

2. Information Center of Guangdong Province, Guangzhou 510000, China)

Abstract: The specification of general CBIR system is first introduced, then the method of retrieval performance of CBIR system is improved, especially, the composite query and retrieval feedback method. Experiments showed that the techniques of composite query and retrieval feedback can improve the performance of our prototype system obviously.

Keywords: content-based image retrieval (CBIR); composite query; relevance feedback