

基于内容检索中算法的层次化体系组织*

唐波,刘雨,孙茂印,汤晓安

(国防科技大学 电子科学与工程学院,湖南 长沙 410073)

摘要 基于内容的检索中,检索算法的常规实现是面向单个特征完成特征提取和相似性匹配,再对单特征检索结果加权完成多特征检索,这种方法计算量大,缺乏灵活性。针对于此,提出了一种检索算法的层次化组织体系,能够灵活组织面向不同特征的检索算法,通过传输因子、优先级因子和复杂度因子的赋值,获得逐步精确的检索结果。实验证明,在查全率和查准率基本一致的前提下,与常规检索体系相比,新检索体系能够减少近 40% 的检索时间。

关键词 层次化体系;基于内容检索;特征提取;相似性测量

中图分类号:TP391.4 文献标识码:A

Hierarchical Architecture of Algorithms in Content-based Retrieval

TANG Bo, LIU Yu, SUN Mao-yin, TANG Xiao-an

(College of Electronic Science and Engineering, National Univ. of Defense Technology, Changsha 410073, China)

Abstract In the content-based retrieval, the conventional approach to the retrieval algorithms is to conduct the feature extraction and similarity match on the basis of single feature and then to conduct multi-feature extraction under the condition of weighting the result of single feature extraction. Such an approach incurs large amount of calculation and lacks flexibility. Therefore, a new architecture to organize different retrieval algorithms in content-based retrieval is proposed. It appears as a hierarchical architecture, which is able to acquire a gradually accurate result. The transfer factor, priority factor and complexity factor are defined in this architecture, which can help to combine different feature-oriented retrieval algorithms flexibly. Experiment shows that this architecture improves the retrieval efficiency, resulting in a reduction of 40% of the retrieval time.

Key words hierarchical architecture; content-based retrieval; feature extraction; similarity measuring

本文所关注的内容检索主要是基于内容的图像检索(Content-based Image Retrieval, CBIR)和基于内容的视频检索(Content-based Video Retrieval, CBVR)。CBIR 通常研究三类特征:颜色特征、形状特征和纹理特征, CBVR 则还进一步考虑了运动特征。近年来,这方面的研究取得了许多成果^[1],研究者深入分析了诸如颜色直方图、颜色相关向量^[2](Color Coherence Vectors, CCV)、同态纹理^[3](Homogeneous Texture)、方向细节直方图^[4]、统计纹理特征^[5]、全局运动信息^[6]等许多特征表现形式,并且基于这些特征形式开发了许多检索算法。检索算法的原理在于测量特征或特征向量的相似性,但是针对同一对象使用不同的特征、不同的相似性测度可能会产生完全不同的结果。于是,研究者采用多特征联合检索的方式来减少这种由于选用特征的差异带来的结果差异。目前,尽管多特征描述机制目前已经有了广泛地研究^[7],但是常规的检索算法仍然依赖于单特征,往往只是在多个单特征检索结果之后进行加权或其他合成方法决定最终结果,这种方式能够克服单特征检索时的弊端,但其计算强度大、时间开销大,并且进行多特征综合的方式一经算法固定就无法由用户更改。换句话说,目前的多特征机制尚不够灵活,不能让用户自由地选择检索特征和检索算法,它适用于描述、管理图像或视频对象^[8]而非检索对象。

本文提出了一种通用的层次化结构来组织与特征相关的不同的检索算法。在这种体系结构中,定

* 收稿日期:2006-08-25

基金项目:国家 863 高技术研究计划资助项目(2003AA134016)

作者简介:唐波(1972—),男,副教授,博士。

义了优先级因子来反映相关特征的重要性,定义复杂度因子来描述相关算法的计算强度,定义传输因子来刻画检索结果的置信区间。结构中上一层的输出结果集作为下一层的输入集,并且每一层都相对独立,便于用户选择合适的检索算法。实验证明,这种结构能够提高检索效率。

1 检索算法的常规实现机制

在内容检索的常规实现机制中,检索算法需要计算样本对象与原始对象的相似性。样本对象是指用于在数据集中进行样本检索(Query by Example, QBE)的样本图像或视频,原始对象是指数据集中存储的图像或者视频,从中查询与样本相似的结果。这种常规结构如图1所示,为了实现一个样本的检索,不得不对所有的原始对象提取特征,然后基于这些特征计算相似度并且进行比较与合成。

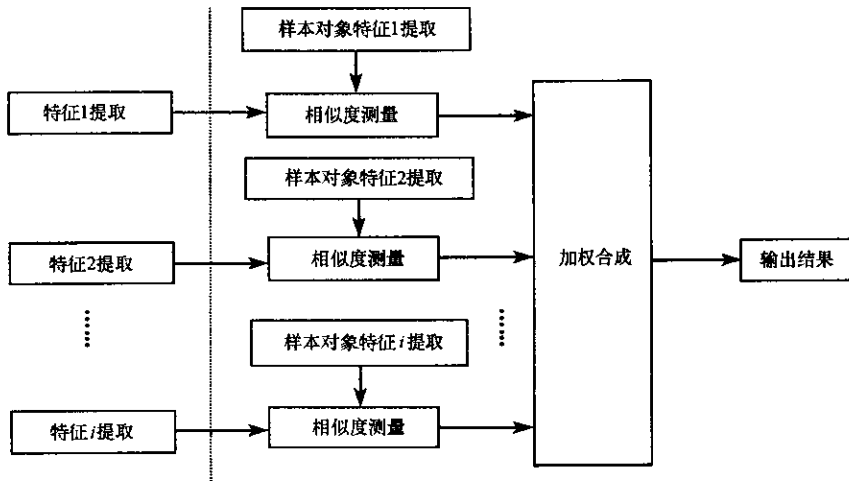


图1 检索算法的常规组织结构

Fig.1 Regular scheme of retrieval algorithm

假设 N 为数据集中原始对象的数目,对一个样本对象比较三个特征,设 te_i 为样本对象第 i 个特征的特征提取时间($i=1, 2, 3$),设 tc_i 为样本对象与原始对象第 i 个特征进行相似度测量所用的时间($i=1, 2, 3$)。至于原始对象的特征提取,由于可以事先完成,因此在计算检索时间时可以不予考虑。于是,常规体系下基于内容检索的时间耗费 T_1 如公式(1)所示:

$$T_1 = N \times (te_1 + tc_1) + N \times (te_2 + tc_2) + N \times (te_3 + tc_3) = N \times \sum_{i=1}^3 (te_i + tc_i) \quad (1)$$

2 层次化体系结构

层次化的检索体系结构是基于这样的考虑:我们可以使用相对简单的特征排除明显不匹配的对象,而仅仅只对潜在的相似对象提取和比较相对复杂的特征。图2展示了一个3层的层次结构。

我们定义了传输因子(TF)、优先级因子(PF)和复杂度因子(CF)来详细描述这种结构。传输因子是0到1之间的一个数,用来表示结果集中允许从当前层传输到下一层的百分比。它反映了当前层检索结果的置信度,即有多大比例的检索结果是具有潜在相似性的结果,值得在下一层去处理。传输因子的设置相对独立,每一层的传输因子都可以与其他层不同。

我们将检索算法的划分为5个优先级,分别用数字1~5表示,数字越大代表的优先级越高,这就是优先级因子。优先级因子使得用户能够强调某些特别的检索特征和检索算法,为检索算法的人机交互提供了途径,这在实际应用中非常普遍。

复杂度因子用来描述检索算法的计算复杂度。表1给出了层次体系结构中复杂度因子与算法复杂度的对应关系。具有较低复杂度的检索算法通常安排到层次结构的较高层,而在底层的算法通常相对复杂。这样,在传输因子的作用下,较复杂的算法只需处理较少的对象,从而检索效率得到提高。

表1 复杂度因子与算法复杂度的关系
Tab.1 Relationship between complexity factor and algorithm complexity

复杂度因子	算法复杂度数量级
1	$\ln N$ or lower
2	N
3	$N \ln N$
4	N^2
5	$N^2 \ln N$
6	N^3
7	2^N or higher

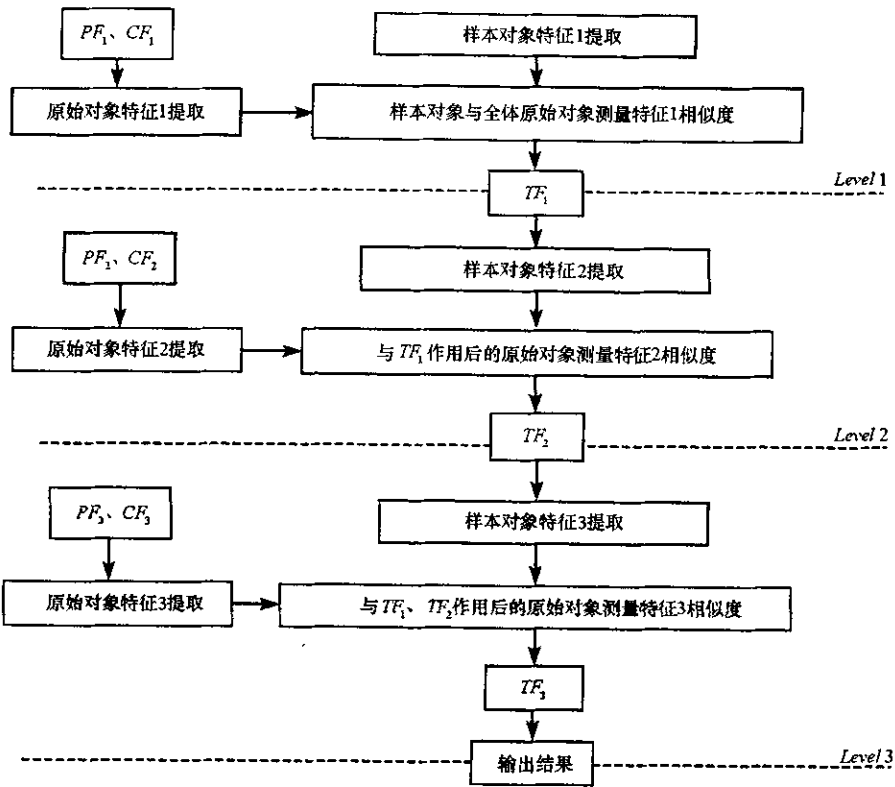


图2 基于内容检索的层次化体系结构

Fig.2 Hierarchical architecture for CBR

层次结构的每一层均设置了传输因子,同时根据样本对象特征确定优先级因子和复杂度因子的取值。传输因子实质上用作相似性测度的标准,只有相似度测量值大于传输因子的原始对象才能够传递到下一层。这些测量值在相似性测量算法中已经被归一化,这样,在相似性比较的同时,就能够决定哪些对象可以传递到下一层,而下一层就可以在上层继续比较其他对象的同时开始本层的比较工作,这就使得整个检索体系能够并行实现,提高检索效率。优先级因子和复杂度因子决定相关的检索算法位于哪一层,优先级越高、复杂度越低意味着算法在整个体系中的处于相对较高的层次。这就确保了在相对重要的特征上与样本对象相似的原始对象不会在一开始就被排除;其次,较复杂的算法仅仅需要处理较少的对象,从而体现出这种层次结构的优越性。在这种层次结构下,时间消耗的分析如下。

假设在层次体系中同样比较3个特征,则层次结构有3层,各层的传输因子分别是 TF_i ($i=1, 2, 3$), n , te_i , tc_i 的定义同前,则整体时间消耗 T_2 为:

$$T_2 = N \times (te_1 + tc_1) + N \times TF_1 \times (te_2 + tc_2) + N \times TF_1 \times TF_2 \times (te_3 + tc_3) \quad (2)$$

比较1)式和2)式,由于 $TF_1 < 1$ 并且 $TF_2 < 1$,于是有 $T_2 < T_1$,即层次体系结构的时间开销要少于常

规组织形式。

3 实验验证及其结果

实验研究验证的指标主要关注的是检索时间和基于内容检索中两个最常使用的参数:查全率 (Recall)和查准率 (Precision)。查全率和查准率的定义如下:

$$recall = \frac{N_{correct}}{N_{correct} + N_{miss}}, \quad precision = \frac{N_{correct}}{N_{correct} + N_{false}} \quad (3)$$

$N_{correct}$ 是指检索得到的正确结果的数目, N_{miss} 是指漏检的正确结果的数目, N_{false} 是指检索得到的错误结果的数目。一般而言,这两个参数是相互影响的,但是对于一个自动检索系统,我们更关注的是查全率,尽可能确保正确的结果都能被检索出来。

我们构造了一个3层的层次结构,使用的检索算法为:分块颜色直方图算法、CCV算法^[2]以及同态纹理算法^[3]。前两种算法使用的是颜色特征,第三种算法使用的是纹理特征,检索中我们优先考虑颜色特征,其次考虑纹理特征。层次体系中每一层的传输因子均设置为0.5,按照上述优先顺序的思路,三种算法的PF取值分别依次设置为5、5、4,根据对各算法复杂度的分析,CF值依次为4、5、6。这样,在层次体系中,分块颜色直方图算法位于顶层,CCV算法位于中间层,同态纹理算法位于底层。检索使用的图像来自于美国华盛顿大学的groundtruth图像数据库(<http://www.cs.washington.edu/research/imagedatabase/groundtruth/>)。我们指定每个图像集的第一幅图像为样本图像,在整个图像集中检索与样本相似的原始图像。

实验结果如表2所示。通过对比,可以发现在两种检索方式下,查准率和查全率基本保持一致,区别只在于检索时间。本文提出的层次检索体系相对于常规检索机制,检索时间减少了将近40%,效率的提高非常可观。

表2 检索性能比较

Tab.2 Retrieval performance comparison

图像集	检索时间(ms)	常规检索机制		检索时间(ms)	层次检索体系	
		查准率	查全率		查准率	查全率
Arboregreens	436	66.7%	92.2%	285	65%	92.3%
Football	478	52.9%	100%	292	53.5%	100%
Leaflesstrees	452	69.4%	97.5%	288	68.6%	97.8%
Sanjuans	417	86.0%	92.5%	273	87.0%	93.3%
Columbiagorge	722	50.8%	94.5%	420	51.4%	94.7%
Greenland	1823	53.8%	97.8%	874	54.2%	98.2%
Greenlake	424	56.5%	96.7%	265	58%	96.5%

4 结论

本文针对基于内容检索中检索算法的实现机制,提出了一种层次化的算法体系结构,在这种体系结构中,针对不同特征的检索算法能够灵活地组合起来,在传输因子、优先级因子和复杂度因子等参数的作用下,实现高效检索。实验结果证明,与常规的检索机制项比较,层次检索体系中的查准率和查全率与前者基本一致,但是能够减少将近40%的检索时间。

参考文献:

- [1] Ngo C W, Zhang H J, Pong T C. Recent Advances in Content Based Video Analysis[J]. International Journal of Image and Graphics, 2001, 1(3): 445-468.
- [2] Pass G, Zabih R, Justin M J. Comparing Images Using Color Coherence Vector[A]. Proc. of ACM Multimedia 1996[C], Boston MA USA, 1996: 65-73.
- [3] Ro Y M, Kim M, Kang H K, et al. MPEG-7 Homogeneous Texture Descriptor[J]. ETRI Journal, 2001, 23(2): 41-51.
- [4] Andrououtsos D, Plataniotis K N, Venetsanopoulos A N. Direction Detail Histogram for Image Retrieval[J]. Electronic Letters, 1997, 33(6): 1935-1936.
- [5] 李晓华, 沈兰荪. 基于小波压缩域的统计纹理特征提取方法[J]. 电子学报, 2003, 31(12A): 2123-2126.
- [6] Yu T, Zhang Y. Retrieval of Video Clips Using Global Motion Information[J]. Electronics Letters, 2001, 37(5): 893-895.
- [7] Punpiti P, Nikitas A, Sanan S, et al. Multi-feature Content Based Image Retrieval[A]. Int. Conf. on Computer Graphics and Imaging (CGIM98)[C], Halifax, Canada, 1998.
- [8] 刘宇驰, 谢毓湘, 吴玲达, 等. 一种开放式视频管理框架[J]. 国防科技大学学报, 2006, 28(1): 73-76.

