

基于内容检索的一种中间表达机制*

白雪生 徐光祐 史元春

(清华大学计算机科学与技术系 北京 100084)

(清华大学智能技术与系统国家重点实验室 北京 100084)

E-mail: xsbai@vision.cs.tsinghua.edu.cn

摘要 在对多媒体信息进行基于内容的检索时,灵活有效的检索表达机制是影响系统性能的重要因素.现有系统都采用十分简单的表达机制,难以处理复杂的检索要求.通过分析,文章提出了将检索请求映射为特征空间上的相似度分布函数,并提供了一种组合机制——CbExpr (content based expression)来使底层的单特征检索近似此相似度分布函数,从而完成对于复杂检索基于内容检索要求的处理.文章阐述了CbExpr的基本思想和示例,并给出了在Internet/浏览器平台上的实现.

关键词 基于内容检索,特征空间,相似度分布函数,检索表达,结果集.

中图法分类号 TP311

在对图像、视频等多媒体信息进行基于内容的检索时,灵活有效的检索表达机制是影响检索效率的重要因素.由于媒体对象的内容是由表现形式为高维向量的特征构成的,因此,在对巨容库进行检索时,要求在底层针对特征(如颜色、纹理等)建立 k -最近邻查找索引,以提高检索速度.因此,如何将检索要求转换为底层索引的查找操作,就成为基于内容检索表达机制的研究所要解决的问题.

现有系统的表达机制都十分简单,仅对多特征检索融合进行了一些探讨. IBM的QBIC(query by image content)系统对多特征的复合检索采用了单特征检索结果取交集的方法^[1],这种方法缺乏用户对特征重要性进行调节的灵活性. UCSD(University of California, San Diego)的Virage系统通过对不同单特征空间中的距离赋以权重,构成了一个总特征空间中的复合距离,通过调节权重和检索的特征值,来表达不同的检索请求^[2-4].

在实际检索中,往往需要表达相当复杂的检索要求.如,Chabot计划中对于“日落”图像的检索,其对应的描述图像主颜色可分为红、黄或紫^[5].在此情况下,采用上述机制,用户的检索要求就难以通过一次提交来完成.可见,现有检索表达机制的能力难以满足复杂检索要求的需要.

在本文中,我们对基于内容检索的表达问题进行了系统的分析,提出所有检索要求都可映射为特征空间上的一个相似度函数.在此基础上,我们对现有的检索表达方法进行了讨论,提出了CbExpr (content based expression)——一种基于内容检索的表达机制.作为一种组合机制,CbExpr通过对单特征检索结果的组合,在内容上近似地来描述空间上的相似度函数,从而表达复杂的检索要求.

第1节给出基于内容检索的基本概念.第2、3节分别介绍CbExpr的基本思想和它所包括的操作与函数.第4节给出了一些典型的检索表达,对CbExpr作进一步的说明.第5节给出了实验结果.

1 基于内容检索表达的基本概念

为了便于对基于内容的检索表达机制问题进行系统地分析,我们需要对一些概念作一下说明.

* 本文研究得到国家863高科技项目基金和清华大学211学科建设项目基金资助.作者白雪生,1972年生,博士生,主要研究领域为计算机视觉,基于内容检索.徐光祐,1940年生,教授,博士生导师,主要研究领域为计算机视觉,多媒体技术.史元春,女,1967年生,在职博士生,副教授,主要研究领域为分布式多媒体信息处理,CSCW.

本文通讯联系人:白雪生,北京100084,清华大学计算机科学与技术系信息教研组

本文1997-11-12收到原稿,1998-04-17收到修改稿

定义 1. 在基于内容的检索中,任何被检索对象(图像、视频镜头)都可被一个用于存取的 oid(一般为长整数)所唯一标识.所有被检索对象的集合构成了对象库 OB,即 $OB = \{oid\}$.

定义 2. 对被检索对象,采用特征提取方法 E_i (如颜色直方图、纹理)可得此对象的特征描述 $f_i(oid)$, $f_i(oid)$ 为高维空间 F_i 中的向量,其坐标值 $(x_{i,1}, x_{i,2}, \dots, x_{i,d_i})$ 分别对应于对象在该方向上的特征分量. F_i 称为对应特征 E_i 的子特征空间, d_i 为 F_i 的维数.所有子特征空间的笛卡尔积构成了对象的特征空间 F , $F = F_1 \times F_2 \times \dots \times F_N$, N 为特征的数目.从 OB 到 F 的映射 $f: OB \rightarrow F$ 给出了对象的特征描述 $f(oid)$, $f(oid)$ 在 F_i 上的坐标等于对应于 $f_i(oid)$ 的坐标.

在上述定义的基础上,我们假设对象内容可用特征空间 F 中的特征来充分描述,则给定检索要求,对于特征空间中任一点 X ,就存在一个数值反映特征描述为 X 的媒体对象对此检索要求的满足程度——相似度,从而得到一个定义在特征空间 F 上的相似度分布函数.任何检索要求都与这样一个相似度函数相对应.我们给出其定义,并进而对基于内容检索建立形式化描述如下.

定义 3. 在建立特征空间后,任何检索要求都可用其在特征空间上的相似度分布函数 $s(X)$ 来表征.相似度分布函数 $s(X)$ 定义为从特征空间 F 映射到 $[0, 1]$ 区间的函数,即有 $s: F \rightarrow [0, 1]$. $s(X)$ 的函数值表示特征空间中坐标为 X 的对象与此概念的相似度,1 表示相似度最大,0 为最小.

定义 4. 对于对象库 OB,已知检索要求 R 在特征空间上的相似度分布函数 $s(X)$,则对于 R 的基于内容检索,即在库 OB 中寻找相似度最大的一组对象及其对应的相似度,一般指定返回对象个数 (M) 或相似度的阈值 (要求返回大于某一指定阈值 th 的所有对象),即寻找

$$\{(oid, s(f(oid))) | oid \in OB \text{ and } s(f(oid)) \in \max_{j \in OB} \{s(f(j))\}\}, \quad (1)$$

或

$$\{(oid, s(f(oid))) | oid \in OB \text{ and } s(f(oid)) > th\}. \quad (2)$$

2 检索表达的灵活性(flexibility)

2.1 底层的单特征索引技术

实际应用中对象库一般都是巨容库,为了便于实时检索,需采用快速索引结构和查找算法.由于特征是高维向量,较为有效的索引技术是 k -最近邻搜索,亦称相似检索.这种方法将特征空间中两点的相似度与其间的距离对应起来,并根据此距离度量建立索引结构.利用此索引结构,给定检索特征,可在很小的查找范围内快速得到与检索特征距离最近(亦最相似)的一组对象.

相似度与距离的对应关系可用多种函数形式来描述,一般说来,相似度作为距离的函数 $sf(d)$,只要满足 3 个条件即可:(a) 在距离为 0 处为 1,即 $sf(0) = 1$;(b) 是距离的单调递减函数;(c) 在无穷远处为 0,即 $sf(+\infty) = 0$.在本文中,我们将采用 $sf(d) = e^{-d}$ 形式的指数型相似度-距离对应函数.

2.2 现有检索表达的相似度函数

根据上述定义,单特征检索的相似度分布函数在对应子特征空间上为一单峰函数,检索特征处取峰值为 1,无穷远处取值为 0,等相似度曲面为一超球面.在整个特征空间 F 上,相似度函数仅取决于子特征空间上的值,呈“岭”状分布,而等相似度曲面为一超柱面.

QBIC 系统采用了单特征检索结果取交集实现多特征复合检索的方法,这种方法未反映相似度函数分布的概念,但其本质仍是复合特征空间中的一个单峰函数. Virage 系统通过对不同单特征空间中的距离赋以权重,构成总特征空间中的复合距离来进行检索,对应相似度函数是相关子特征组合空间中的单峰函数,等相似度曲面为一超椭球面.可见,上述方法仅可表达单峰相似度函数的检索要求,而难以充分接近特征空间复杂(多峰)相似度分布函数的检索要求,因此表达能力有限.

2.3 检索表达中的可调节因素

显然,基于内容表达机制的目的,即提供一种方法,使其可利用底层单特征检索的单峰相似度函数的组合来接近任一检索请求的相似度函数.通过分析,我们给出构造复杂相似度分布函数中的几种可调节因素:(1) 单特

征检索中的距离尺度变换,导致相应相似度函数的峰的“宽度”发生变化;(2)对相似度函数的比例变换,导致相似度函数的“高度”发生比例变化;(3)多相似度函数的最大值函数;(4)多相似度函数的最小值函数;(5)多相似度函数的加权函数.

根据这5个因素,我们提出了一种基于内容的检索表达机制 CbExpr,通过上述变换来实现用底层检索对应的单峰函数来近似任意分布相似度函数.

3 CbExpr 的基本操作和函数

CbExpr 包括 1 个基本的单特征检索操作(Query)、6 个复合操作(Union, Intersect, Weight, Truncate, Threshold, Multiply)和两个函数(GetNumber, GetMinthreshold). 单特征检索是对根据某一特征建立的距离索引进行检索来获得最基本的单特征检索结果,而通过复合操作和函数,可以将不同检索的结果组合起来,构成具有复杂相似度分布函数的检索.

3.1 结果集

在 CbExpr 中,所有操作都是围绕对检索结果的处理来定义的,因此,首先需要对检索结果进行形式化的描述.为此,我们引入结果集的概念.结果集 $RSet$ 是所有操作(包括单特征检索检索请求和复合操作)所返回的结果,其形式为 $RSet = \{(i, v)\}$,其中 $i \in OB$,为对象标识; $v = s(f(i))$,为相似度; (i, v) 按照相似度 v 由大到小的顺序排列.

3.2 单特征检索 Query

单特征检索是产生结果集的最基本操作,它对应于对某一特征的距离索引进行的底层查找.我们借用 D. A. White 和 R. Jain 对相似检索的定义,并给出单特征检索的形式如下:

$$RSet = Query(fname, f_v, n, th, e), \tag{3}$$

其中 $fname$ 是特征名称,对应特定索引文件; f_v 为高维向量,代表检索特征值. n 为一整数,代表要求返回的对象个数; th 为 $[0, 1]$ 区间上的实数,为相似度阈值,要求检索返回对象相似度不小于此阈值. n 与 th 同时作用. e 是近似检索引入的用于界定检索误差的近似参数,为非负实数,反映了近似检索所允许的与精确检索的误差.当 $e = 0$ 时,近似检索就退化成了精确检索.

3.3 并(union)

并(union)是一种复合操作,是对传统集合并操作的一种扩充,其定义如下:

$$Union(RSet_1, RSet_2, \dots, RSet_m) = \{(i, v) | (\exists j(i, v) \in RSet_j) \wedge (\forall j \forall u(i, u) \in RSet_j \rightarrow v \geq u)\}. \tag{4}$$

即在返回结果集中,对于 $RSet_1, RSet_2, \dots, RSet_m$ 中具有相同对象标识 i 的元素 $(i, u_1), (i, u_2), \dots, (i, u_n)$,仅包括具有最大相似度值的元素.

3.4 交(Intersect)

交(Intersect)是对传统集合交操作的一种扩充,其定义为

$$Intersect(RSet_1, RSet_2, \dots, RSet_m) = \{(i, \min_j(u_{ij})) | \exists j(i, u_{ij}) \in RSet_j\}, \tag{5}$$

其中 u_{ij} 的定义为:若 $\exists v(i, v) \in RSet_j, u_{ij} = v$; 否则 $u_{ij} = 0$. 显然,在返回结果集中,只有出现在所有作用对象结果集 $RSet_1, \dots, RSet_m$ 中的对象具有非零相似度,其他对象的相似度均为 0.

3.5 权(Weight)

权(Weight)操作同样作用于多个结果集,其定义为

$$Weight(RSet_1, a_1, RSet_2, a_2, \dots, RSet_m, a_m) = \{(i, \prod_{j=1}^m (u_{ij})^{a_j}) | \exists j(i, u_{ij}) \in RSet_j\}. \tag{6}$$

将权操作转换到距离空间来讨论,在返回结果集中,对象 i 相似度对应的“距离” $Dist_i$ 为

$$Dist_i = -\ln(\prod_{j=1}^m (u_{ij})^{a_j}) = \sum_{j=1}^m a_j \cdot (-\ln(u_{ij})) = \sum_{j=1}^m a_j \cdot Dist_{ij}, \tag{7}$$

即在返回结果集中,对象 i 的“距离” $Dist_i$ 等于各 $RSet_j$ 中的对象所对应的“距离” $Dist_{ij}$ 的加权和.显然,权操作具

有 Virage 系统融合不同子特征空间距离构成“复合距离”的作用。对于仅有一个结果集作为作用对象的情况,即 $Weight(RSet_1, a_1)$, 由定义可知,其作用相当于距离尺度变换,在第 2 节中的距离尺度变换可以通过这个操作来完成。

3.6 截断(Truncate)

截断操作的作用对象是一个结果集,此外,它还带一个整数形式的截断参数 M 。其定义为

$$Truncate(RSet, M) = \{(i, v) | (i, v) \in RSet, v \in \text{Max}_M(u, \exists j(j, u) \in RSet)\}. \quad (8)$$

即截断操作返回 $RSet$ 中具有最大相似度的 M 个元素。当 $RSet$ 中元素数目小于或等于 M 时,此操作无效,返回的结果集与 $RSet$ 相同。

3.7 阈值(Threshold)

与截断操作类似,阈值操作的作用对象也是一个结果集,此外,它还带一个取值在 $[0, 1]$ 区间上的阈值参数。其作用是对此结果集取阈值,仅保留那些相似度不小于此阈值参数的元素,即

$$Threshold(RSet, th) = \{(i, v) | (i, v) \in RSet \& v \geq th\}. \quad (9)$$

3.8 乘数(Multiply)

乘数操作的作用对象也是一个结果集,此外,还带一个实数形式的乘数因子 a 。其定义如下:

$$Multiply(RSet, a) = \{(i, v) | v = 1, \text{if } (a \cdot u) > 1; v = a \cdot u, \text{otherwise } (i, u) \in Rset\}. \quad (10)$$

即其作用是对 $RSet$ 中所有对象的相似度同时乘以因子 a 。若相乘后相似度大于 1,则取值为 1。第 2 节中讨论的相似度函数峰的“高低”的改变可以通过这个操作来完成。

3.9 取元素个数函数(GetNumber)

GetNumber 是一个作用于单一结果集的函数,它返回此结果集中元素的个数,即

$$GetNumber(RSet) = |RSet|. \quad (11)$$

GetNumber 返回的函数值可用于截断操作中截断参数的设置。

3.10 取最小阈值函数(GetMinthreshold)

GetMinthreshold 也是一个作用于单一结果集的函数,它返回此结果集中最小的相似度值(最后一个元素的相似度),即

$$GetMinthreshold(RSet) = \text{Min}\{s(f(i)), (i, s(f(i))) \in Rset\}. \quad (12)$$

GetMinthreshold 返回的函数值可用于阈值操作中截断参数的设置。

4 CbExpr 的检索表达能力示例

利用 CbExpr 的操作和函数,我们可将其作为一种检索语言来完成复杂概念的检索。在本节中,我们将给出一些用 CbExpr 检索的实例,以对其表达能力和表达方法加以说明。

4.1 传统交操作

我们在前面分析中曾指出, QBIC 采用的对单特征索引结果求交的操作,难以与相似度函数相对应。但利用 CbExpr,可以实现传统的交操作。考虑结果集 $RSet_1$ 和 $RSet_2$, 根据 Intersect 的定义,在对 $RSet_1$ 和 $RSet_2$ 进行 Intersect 操作得到的结果集中,仅在一个结果集中出现的对象相似度为 0。在 $RSet_1$ 和 $RSet_2$ 不存在相似度为 0 的对象的情况下,以下的 CbExpr 检索可实现传统的交操作。

$$\begin{aligned} & Union(Threshold(Intersect(RSet_1, RSet_2), GetMinthreshold(RSet_1)), \\ & Threshold(Intersect(RSet_1, RSet_2), GetMinthreshold(RSet_2))). \end{aligned}$$

4.2 多特征融合

基于内容的检索常常要融合不同特征进行检索。一种灵活的方法是利用 Weight 操作,通过权重来调节不同特征的重要性。一个对图像进行颜色、纹理多特征融合检索的 CbExpr 检索如下:

$$Weight(Query(color, f_c, N_c, 0.0, 0.0, 0.0), 1.0; Query(texture, f_t, N_t, 0.0, 0.0, 0.0), 1.0).$$

其中 color, texture 分别为对应的特征名称; f_c 和 f_t 分别为检索特征向量; N_c 和 N_t 为检索要求返回的对象个

数. 阈值都被设为 0.0, 使阈值约束无效. 近似参数也都取 0.0, 即为精确检索. 各单特征检索后的系数代表相应的权重(此处均为 1.0), 可以根据需要进行调节.

4.3 多峰逼近

对多峰形状相似度函数逼近的 CbExpr 表达如下:

$$Union(RSet_1, RSet_2, \dots, RSet_n),$$

其中各 $RSet_j$ 分别为对应的单峰函数. 实际上, 通过使用足够多、足够窄的峰, 可以对任意复杂形状给出相当的近似. 但在实际应用中, 往往只需给出粗略的近似就可以满足要求.

5 实现与结论

我们在 Internet/浏览器平台上实现了一个针对图像的基于内容检索系统, 并把 CbExpr 作为该系统的中间表达机制. 用户可直接输入和编辑 CbExpr 的检索表达式, 其界面如图 1 所示. 通过界面上的 Feature Button, 用户可激活相应的单特征(颜色、纹理等)特征值输入界面(用 Java 编写).

在图 1 中, CbExpr 表达式所显示的是两个颜色单特征检索操作的 Union 操作. 其形式为

$$Union(Query(color = 255, 0, 0 \quad Truncate_number = 16 \quad Threshold_value = 0.0), \\ Query(color = 0, 0, 255 \quad Truncate_number = 16 \quad Threshold_value = 0.0)).$$

即表达的是检索含有红($R=255, G=0, B=0$)或蓝($R=0, G=0, B=255$)主颜色特征的图像, 每个单特征检索要求最多返回 16 个媒体对象($Truncate_Number=16$), 返回对象的相似度大于 0.0($Threshold_value=0.0$, 实际上, 即无相似度阈值限制). 图 2 给出了检索提交后的返回结果, 共 32 幅图像, 每幅图像下的数值为此图像对于检索要求的相似度值. 返回图像按相似度的大小顺序排列, 可见红色主颜色的图像(如图中花果等图像)和蓝色主颜色的图像(如图中水纹等图像)根据相似度交错排列, 很好地完成了检索要求.



图 1 CbExpr 输入和编辑界面

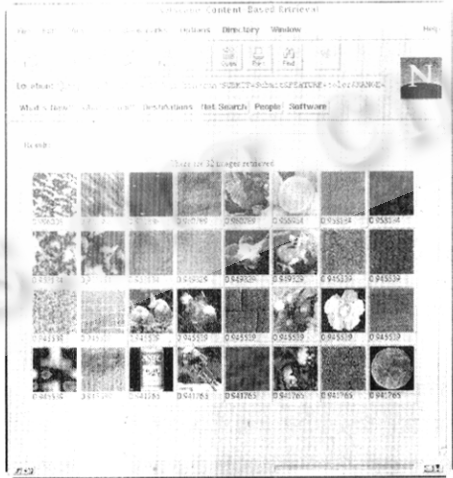


图 2 检索提交返回结果

参考文献

- 1 Lee D, Flickner M, Barber R *et al.* Indexing for complex queries on a query-by-content image database. In: IAPR (International Association of Pattern Recognition) ed. Proceedings of the 12th IAPR International Conference on Pattern Recognition—Conference A: Computer Vision and Image Processing. Los Alamitos, CA: IEEE Computer Society Press, 1994. 142~146
- 2 Bach J R, Fuller C, Gupta A *et al.* The virage image search engine: an open frame work for image management. In: Sethi I K, Jan R C eds. Proceedings of SPIE: Storage and Retrieval for Still Image and Video Databases IV, Vol 2670. Bellingham: SPIE, 1996. 76~87

- 3 White D A, Jain R. Similarity with the SS-tree. In: Su SYW ed. Proceedings of the 12th IEEE International Conference on Data Engineering. Los Alamitos, CA: IEEE Computer Society Press, 1996. 516~523
- 4 White D A, Jain R. Similarity indexing: algorithms and performance. In: Sethi I K, Jain R C eds. Proceedings of the SPIE: Storage and Retrieval for Image and Video Databases IV, Vol 2670. Bellingham: SPIE, 1996. 62~75
- 5 Ogle V E, Stonebreaker M. Chabot, retrieval from a relational database of images. Computer, 1995,28(9),40~48

An Intermediate Expressing Mechanism of Content-based Retrieval

BAI Xue-sheng XU Guang-you SHI Yuan-chun

(Department of Computer Science and Technology Tsinghua University Beijing 100084)

(State Key Laboratory of Intelligent Technology and Systems Tsinghua University Beijing 100084)

Abstract Flexible expressing mechanism is an important factor influencing system performance in content-based retrieval, the simple mechanisms current systems employed can not express complex content queries. In this paper, an intermediate query expressing mechanism—CbExpr is proposed. The mechanism first maps each content query to a similarity distribution function in feature space, then approximates this function using basic single feature queries. Basic thoughts and examples are given, also with implementation on Internet/Navigator platforms.

Key words Content-based retrieval, feature space, similarity distribution function, query expression, result set.