

基于本体和奇异值分解的图像数据索引查询算法

黄 霞

(三亚航空旅游职业学院基础教学部,海南三亚 572000)

摘 要: 提出一种基于领域本体潜在语义索引和奇异值分解的图像数据查询算法,将查询扩展向量映射到潜在语义空间,根据相似度计算方法计算查询向量与图像文档之间的相似度,并将相似度大于阈值的文档作为检索结果降序排列返回给用户.该算法能更有效地提高图像检索的查准率和查全率.

关键词: 图像索引查询;本体;潜在语义索引;奇异值分解

中图分类号: TN391.41 **文献标识码:** A **文章编号:** 0372-2112(2014)02-0288-04

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2014.02.012

An Image Data Query Algorithm Based on Ontology and Singular Value Decomposition

HUANG Xia

(Department of Basic Courses, Sanya Aviation & Tourism College, Sanya, Hainan 572000, China)

Abstract: An image data query algorithm based on latent semantic indexing in domain ontology and singular value decomposition is proposed. It projects query expansion vector on latent semantic area. A new method to computing similarity between query vectors and query documents is proposed. Users can get accurate and effective retrieval results using the similarity. The experimental results show that the proposed solution can improve image index query recall ratio and precision ratio.

Key words: image index query; ontology; latent semantic indexing; singular value decomposition

1 引言

随着大数据时代的到来,以图像为主的各种多媒体数据的容量呈爆炸性的增长.图像数据库作为图像数据的集合,广泛应用于文字模式识别、医学图像分析、地图数据处理和显微照片分析等领域.图像数据库的研究与开发发挥着越来越大的作用,图像数据库技术的应用范围已涉及到各行各业^[1].

如何从大规模图像数据库中快速准确地检索到所需要的图像是当前的研究热点.基于文本的图像检索使用关键词对图像进行刻画,实现基于关系数据库的检索,很难反映图像的多种特征,并且带有强烈的主观性.传统的图像检索通常基于内容,例如利用图像的形状、颜色和纹理等特征来检索图像.该方法存在低层语义和高层语义之间的语义鸿沟.此外,传统的数据库技术通常只处理静态的数据集,而对于动态变化的数据流很难进行分析和处理.主要原因是数据流具有数据规模大、到达速度快和数据维数高等特点.高维数据存在冗余度高,噪声干扰多以及由于维数高所导致的存储和计算资源消耗大等问题.多个维度表示多个特征,然而并不是每个特征都能够对解决所面临的问题有所帮助.因此,

如何有效地合理地降低数据维数是高维数据得到广泛应用的关键.如何设计合理高效的数据集处理算法,降低数据维数、缩小数据规模的内存空间和减少计算量是解决上述问题的重点.

针对上述问题,本文引入本体作为表达图像数据潜在语义的基础,构建基于本体的潜在语义索引(Latent Semantic Indexing, LSI)模型.该模型基于领域本体对图像数据内容进行潜在语义标引和构建索引词-标注矩阵.对矩阵进行切片,得到针对图像数据的某一显著维或希望研究的维度,针对切片后的矩阵进行奇异值分解后,得到切片矩阵的数据分布特征,即奇异值特征.从而降低数据维数和缩小数据规模的内存空间和计算量,降低图像数据库中的图像特征索引的复杂度.提出基于本体和奇异值分解的图像数据流索引查询算法,针对用户语义信息的潜在语义索引进行扩展,实现用户查询和文档之间深层次的语义匹配.

实验结果表明,本文算法能有效地对高维图像数据特征进行提取和检测,并具有图像进行转置、旋转、位移、镜像变换后的不变性等特点,在检测的准确率、召回率、错报率等方面都有较大改善,还能有效提高索引效率,可大幅度提高图像检索的查全率和查准率.

2 相关工作

当前图像检索的主流方法分为基于内容的和基于语义的两大类. 基于内容的图像检索^[2]是根据图像的颜色、纹理和形状等特征进行检索, 然后通过这些特征计算相似度, 从而找到具有指定特征或含有特定内容的图像. 该方法可降低图像资源的各种初级特征与检索时检索词的丰富语义之间的差异性, 但该方法检索效率不理想. 基于语义的图像检索方法融合了图像分析和模式识别等技术^[3], 首先利用各种多媒体处理技术对图像资源进行表层特征提取, 然后采用统计学习方法等获得其对应的语义特征, 并使用分层抽取机制来支持图像的语义处理, 优点是特征抽取和模式匹配算法的复杂度较低, 缺点是检索结果中有很大比例的不相关图像.

就结合语义描述、语义提取以及检索方法三个方面而言, 目前国内做的比较好的图像检索系统有以下几种. 由朱兴全等人开发的 iFind^[4]系统, 结合语义与视觉特征信息的综合内容, 并将语义和底层特征都结合到相关反馈过程中, 将图片和数据库中的语义注释在语义网络中关联起来, 能在图像数据内容的表现中不但体现图像的底层可视特征, 而且更重要的是能体现图像的高层语义信息. Visengine^[5]图像检索原型系统基于主要区域语义, 采用图像分割技术和 XML 机制, 将主要内容的对象语义作为主要区域的类别, 主要内容的视觉特征即等价于主要区域的视觉特征. SemView^[6]作为分布式的图像检索系统, 支持视觉特征和语义查询, 智能化地将分布式图像数据库中分散存放的各类图像数据进行排列.

随着本体研究的广泛和深入, 基于本体的检索技术得到较快的发展. 文献[7]提出一种基于本体的信息检索模型, 采用本体作为表达潜在语义的基础, 能有效反映用户信息需求和文档的潜在语义. 文献[8]给出了本体的形式化模型, 并提出六种基于本体模型的查询模式. 本体具有清晰的概念层次结构, 能够表达隐含的概念模型, 并且可以在模型的基础上进行逻辑推理. 基于本体的检索方便用户和系统之间的交互, 使得检索系统某些方面获得相当于人的理解水平.

3 基于本体和奇异值分解的图像数据索引查询算法

采用本体语言描述图像数据库中的图像数据并得到本体信息. 本体信息包含图像数据特征之间的语义关联. 基于语义关联建立图像数据本体库. 通过有效的推理, 生成相关的索引项集合. 根据查询的索引项集合构建索引词-标注向量, 然后依据潜在语义索引模型理

论, 对向量进行奇异值分解, 将索引词和标注矩阵转换到图像数据本体的语义空间.

同时, 将图像检索请求转化为相应的图像查询向量. 基于图像数据本体库设计查询词权重赋值策略. 通过对图像查询向量和权重向量进行语义扩展, 并把它们映射到潜在语义空间. 采用相似度计算方法得到图像查询向量和图像语义标注之间的相似度, 作为阈值将查询结果排序, 最后将排序结果返回给查询用户. 上述过程即为本文提出的基于本体的潜在语义索引查询扩展算法.

在原始的“索引词-标注矩阵”中, 存在冗余、词语多义和异常等问题. 如果能建立一个比原始矩阵小得多, 并只包含有效潜在语义的子空间, 则检索效果将得到极大改善. 有效的方法是维数约减. 即将冗余的信息合并, 而词语多义问题可通过分析上下文相关信息来改进. 把相对影响不大的一些特征去掉则可以部分解决异常问题. LSI 就是这样一种维数约减方法, 通过对“索引词-标注矩阵”进行奇异值分解(SVD: Singular Value Decomposition)运算, 自动计算得到一个比原始空间小得多的有效潜在语义空间.

在原始的“索引词-标注矩阵”结构中, 针对本体词汇集合 B 构建出本体词汇-文档矩阵 $(b_{ij})_{m \times n}$, 该矩阵中的 m 维表达本体词汇的个数, n 表达文档数, 元素 b_{ij} 为正整数, 代表第 i 个本体词汇在第 j 个图像数据文档中出现的频数. 根据潜在语义索引项集合 I' ($I' = \{[X_1], [X_2], \dots, [X_t]\}$), 将矩阵 $(b_{ij})_{m \times n}$ 中的一些行向量合并, 得到索引词-文档矩阵 A . A 中元素 a_{ij} 为非负整数, 代表与索引词 X_i 意义相同的本体词汇在第 j 个图像数据文档中出现的频数代数. 本文 a_{ij} 采用如下公式计算: $a_{ij} = L(i, j) \times C(i)$. $L(i, j)$ 表示局部权重, 代表索引词在第 j 个图像数据文档中的重要程度; $C(i)$ 表示全局权重, 表示第 i 个索引词在整个图像库中的重要程度. 权重计算采用 TF-IDF (Term Frequency-Inverse Document Frequency) 方法, 用统计方式评估一个字词对于数据库中一个文档的重要程度.

根据建立的索引词-文档矩阵 A , 对 A 进行 SVD 奇异值分解, 分解结果得到三个矩阵: 正交矩阵 U (包含 A 的左右奇异向量)、对角矩阵 S (由 A 的奇异值按递减顺序排列构成) 以及 V . 将矩阵 A 变换为下式: $A = USV^T$. 根据算法能得到 S 中最大的前 k 个奇异值, 并根据 U 和 V 中对应的行和列, 得到 A 的 k 维近似矩阵 $A_k = U_k S_k V_k^T$. 将 $U_k S_k$ 的行向量近似成索引词在 k 维潜在语义空间中的向量, $S_k V_k^T$ 的列向量近似等于图像数据特征文档在潜在语义空间中的向量. 根据对角矩阵 S_k 的特性, 对 k 维空间中的向量通过缩放成 $U_k S_k$, 所以能

将 U_k 中行向量视作索引词向量. 与此同时将 V_k^T 中列向量视作文档向量.

在潜在语义索引空间中通过形成词汇—文本矩阵, 并对其进行一系列计算, 最后, 根据相似度得到排序结果. 在检索时, 首先将用户查询向量映射到 k 维潜在语义空间, 然后比较已被映射到潜在语义空间的文档, 进行潜在语义相似度计算. 通过公式 $\hat{q} = q^T U_k S_k^{-1}$ 计算查询词向量 Q 在潜在语义空间中的向量. 其中, q 等于两个向量 Q 与 $Q^T W$ 中权重的乘积, 而 $q^T U_k$ 等于将 q 映射到潜在语义空间向量中, S_k^{-1} 代表潜在语义空间向量每一维不同的权重. 根据向量夹角的余弦值计算得出 \hat{q} 和文档向量之间的相似度, 公式如下:

$$\text{sim}(\hat{q}, s_j) = \frac{\hat{q} \cdot s_j}{|\hat{q}| |s_j|} \quad (1)$$

其中, s_j 代表特征文档 d_j 在 $S_k V_k^T$ 的第 j 个列向量, 等于文档 d_j 在潜在语义空间中的向量表示, $\hat{q} \cdot s_j = \sum_{i=1}^k \hat{q}_i s_{ji}$, $|\hat{q}| |s_j| = \sqrt{\sum_{i=1}^k \hat{q}_i^2} \cdot \sqrt{\sum_{i=1}^k s_{ji}^2}$. 计算得出相似度大小后将结果降序排列, 同时如果相似度大于阈值 S_{\min} 的图像特征文档将作为检索结果 D_{Result} 返回:

$$D_{\text{Result}} = \{d_j | \text{sim}(\hat{q}, s_j) \geq S_{\min}\} \quad (2)$$

算法描述:

步骤 1 对“索引词-标注矩阵”进行奇异值分解, 得到一个比原始空间小得多的有效潜在语义空间.

步骤 2 初始化最小潜在语义相似度向量 $S_k = [0, \dots, 0]$. 向量 S_k 用来保存查询词 q 的 k 个最邻近向量和它们到查询词 q 的最小相似度.

步骤 3 对查询向量中的每个查询词进行语义扩展, 并使用 TF-IDF 方法进行权重分配.

步骤 3.1 如果查询词为本体中的个体

步骤 3.1.1 对于查询词等价类中的每个个体

步骤 3.1.1.1 以个体为概念进行语义扩展和概念扩展

步骤 3.1.1.2 使用 TF-IDF 方法对扩展后的查询词进行权重分配.

步骤 4 根据式(1)计算扩展后的查询词和向量 s_j 之间的相似度. 向量 s_j 表示在潜在语义空间的第 j 个列向量.

步骤 5 计算得出相似度大小后将结果降序排列.

步骤 6 如果相似度大于阈值 S_{\min} 的图像特征文档, 按照式(2)将最小潜在语义相似度向量 S_k 的目标结果更新为 $\text{sim}_{\min}(\hat{q}, s_j)$, 并把这些文档作为检索结果返回.

4 实验分析

实验所采用的数据集为卡通动画图像数据, 由 3000 多幅卡通动画素材图片构建图像库. 该图像库包括人物、动物、植物和室内物品等图片. 每类有 700 多幅

图像, 用 200 幅图像对卡通动画图像库进行检索. 目前, 存在多种图像检索算法^[9~11]. 文献[10]采用小波模型获得图像的纹理特征, 然后利用欧式距离计算图像文件之间的特征相似度. 文献[11]通过划分图像的语义层次, 采用较高层次的语义来描述图像, 并基于语义检索图像. 实验针对提高检索性能的两个目标: 减少系统的处理时间, 在达到某一查全率下追求更高的查准率, 对这些算法进行分析比较. 检索结果得到的图像和图像素材库中的总图像的比值为查全率, 表示检索系统按要求查询图像的能力. 检索结果得到的图像和返回给用户的图像个数比值为查准率. 实验结果如表 1 所示.

表 1 不同图像检索算法的性能比较

	准确率	查全率	检索时间(ms)
文献[10]	0.543	0.221	89
文献[11]	0.576	0.196	78
本文算法	0.821	0.087	57

由表 1 可知, 文献[10]和文献[11]的查准率和查全率比较低, 且平均检索时间较长. 原因是这两种方法的图像特征维数较大, 且采用低层语义特征进行图像检索. 为进一步比较算法之间在不同条件下(例如, 查准率和查全率等)图像检索的性能. 三种算法在不同查全率要求下的查准率如图 1 所示.

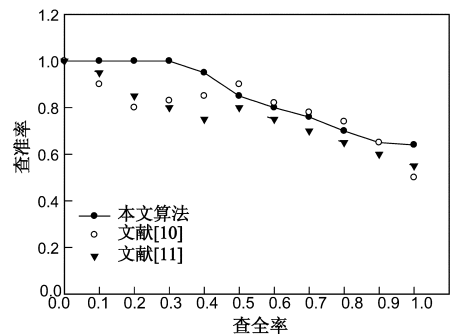


图 1 三种不同算法的查准率曲线图

在相同的查全率情况下, 三种算法的查准率都比较高, 虽然在个别查全率下(例如, 0.5 和 0.6), 本文算法的查准率低于文献[10], 但它的整体查准率稳定, 且符合查全率提高其查准率相应降低的规律.

本文算法通过潜在语义的构建, 生成的图像特征向量在语义概念上更加清晰和明确, 计算机和用户请求之间的语义差距较小.

三种算法在不同图像数据维度下的查准率如图 2 所示. 文献[10]和文献[11]的查准率随着图像数据维度 k 的值增大迅速增加, 在维度 32 左右时达到最高, 接着迅速下降. 本文算法在较小的维度 k 值时就具有较优的查准率.

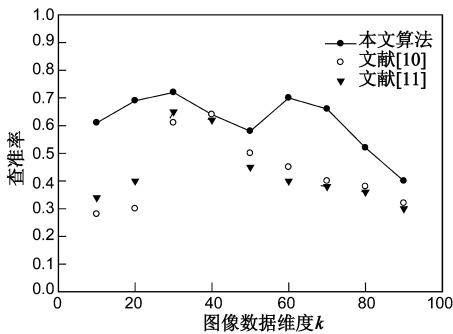


图2 三种不同算法的维度扩展性曲线图

三种算法在不同查全率下的查询时间如图 3 所示。本文算法采用权重赋值策略使得对图像特征矩阵进行奇异值分解时产生的对角矩阵在查全率较高时都能较好地近似原始图像特征矩阵,极大地降低了图像语义索引的存储空间和计算时间,提高了查询响应速度。

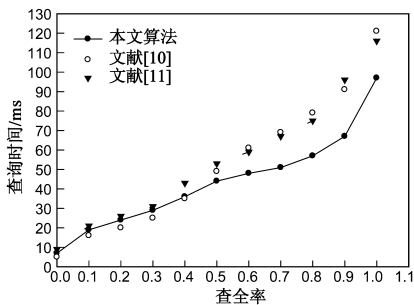


图3 三种不同算法的查询时间曲线图

此外,本文算法在检索时,将用户查询向量映射到潜在语义空间,通过比较已被映射到潜在语义空间中的文档进行潜在语义相似度计算。由于潜在语义索引有效地解决了冗余、词语多义和异常等问题,进一步降低了索引时间。

5 总结

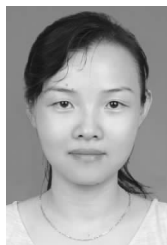
本文算法基于本体查询词权重赋值策略和关键词查询扩展,通过潜在语义扩展和奇异值分解的查询扩展向量,将查询扩展向量映射到先前建立的潜在语义空间,再根据相似度的计算方法计算用户查询向量与图像特征数据文档之间的相似度,并将相似度大于阈值的图像数据作为检索结果并降序排列返回给用户。实验结果表明,本文算法与传统空间向量模型相比,能更有效地提高图像检索的查准率和查全率。进一步的工作将基于潜在语义的查询优化策略,与已有的图像数据库系统进行整合,构建准确高效的语义索引结构以加速图像语义数据的标识和查询。

参考文献

- [1] 谢从华,宋余庆,等. 图像数据库管理系统的关键技术研究综述[J]. 计算机应用研究, 2011, 28(9): 3206 - 3225.
Xie Cong-hua, Song Yu-qing, et al. Key technologies for image database management system: a survey [J]. Application Re-

- search of Computers, 2011, 28(9): 3206 - 3225. (in Chinese)
- [2] 李向阳,等. 基于内容的图像检索技术与系统[J]. 计算机研究与发展, 2001, 38(3): 344 - 354.
Li Xiangyang, et al. The technique and systems of content-based image retrieval[J]. Journal of Computer Research & Development, 2001, 38(3): 344 - 354. (in Chinese)
- [3] J Golbeck, et al. The national cancer institute's thesaurus and ontology[J]. Journal of Web Semantics, 2003, 1(1): 75 - 80.
- [4] 朱兴全,等. iFind: 一个结合语义和视觉特征的图像相关反馈检索系统[J]. 计算机学报, 2002, 25(7): 681 - 688.
Zhu Xingquan, et al. iFind: An image retrieval system with relevance feedback based on the combination of semantics and visual features[J]. Chinese Journal of Computers, 2002, 25(7): 681 - 688. (in Chinese)
- [5] 徐晓刚,彭彬彬,孙正兴. 一个基于语义的图像检索系统: VisEngine[J]. 计算机工程, 2004, 30(4): 51 - 53.
Xu Xiaogang, Peng Binbin, Sun Zhengxing. A semantic-based image retrieval system: VisEngine [J]. Computer Engineering, 2004, 30(4): 51 - 53. (in Chinese)
- [6] 刘欣. 一种基于潜语义标引的全文检索系统的研究与实现[D]. 天津: 南开大学, 2003.
- [7] 王洪伟,吴家春,蒋馥. 基于本体模型的信息检索机制研究[J]. 情报学报, 2004, 23(1): 57 - 62.
- [8] I Horrocks, P F Patel-Schneider. Reducing OWL entailment to description logic satisfiability [J]. Journal of Web Semantics, 2004, 1(4): 345 - 357.
- [9] 程刚. 一种基于形状的图像相似性检索方法[J]. 合肥工业大学学报, 2007, 30(2): 148 - 150.
Cheng Gang. A new image similarity retrieval method based on shape [J]. Journal of Hefei University of Technology, 2007, 30(2): 148 - 150. (in Chinese)
- [10] Haley GM, Manjunath B S. Rotation-invariant texture classification using a complete space-frequency mode [J]. IEEE Trans Image Processing, 1999, 8(2): 255 - 269.
- [11] 吴楠,宋方敏. 一种基于图像高层语义信息的图像检索方法[J]. 中国图像图形学报, 2006, 11(12): 1174 - 1180.
Wu Nan, Song Fangmin. An image retrieval method based on high-level Image semantic Information [J]. Journal of Image and Graphics, 2006, 11(12): 1174 - 1180. (in Chinese)

作者简介



黄 霞 女, 1975 年 4 月出生, 湖南岳阳人。1998 年毕业于中南大学信息工程学院, 2012 获中南大学计算机应用专业硕士学位。一直在高校从事计算机教学工作, 研究方向为数据库。

E-mail: 4947249@qq.com