

文章编号:1001-9081(2010)01-0090-04

基于个性化本体的图像语义标注和检索

史婷婷, 闫大顺, 沈玉利

(仲恺农业工程学院 计算机科学与工程学院, 广州 510225)

(to_shitingting@126.com)

摘要:针对目前图像检索系统较难实现语义检索的问题,提出了一种新的以本体为核心的图像语义标注和检索模型。构建个性化本体描述图像语义,继而提取基于概念集的图像语义特征并利用本体中“Is-A”关系设计相似性度量方法最终实现语义扩展检索。其难点在于顶级本体向个性化本体进化,以及基于概念集和“Is-A”关系实现语义相似度量的方法。通过系统的初步实现与相关实验的验证,该模型的检索准确度可达88.6%,明显高于传统的基于关键字和基于通用本体的图像检索,实现了图像智能检索功能。

关键词:基于语义的图像检索;本体;个性化本体;图像语义标注;检索模型

中图分类号: TP391 **文献标志码:** A

Personalized ontology-based semantic annotation and retrieval of image model

SHI Ting-ting, YAN Da-shun, SHEN Yu-li

(School of Computer Science and Engineering, Zhongkai University of Agriculture and Engineering, Guangzhou Guangdong 510225, China)

Abstract: In order to solve the problem that it is difficult to implement semantic retrieval of image, a novel ontology-cored semantic annotation and retrieval model of image was proposed, which constructed a personalized ontology to describe image, extracted the image semantic features based on the concept set and used inter-relationship "Is-A" of ontology to measure the similarity between the images and query. The difficulty lies in the personalized evolution of the top ontology, as well as the semantic similarity measure based on concept set and inter-relationship "Is-A". This model was implemented and some related experiments were carried out. The experimental results validate the retrieval accuracy as high as 88.6%, significantly higher than the traditional keyword-based and generic ontology-based image retrieval models, which realized intelligent retrieval of image.

Key words: semantic-based image retrieval; ontology; personalized ontology; semantic annotation of image; retrieval model

0 引言

基于内容的图像检索依靠计算机自动抽取图像特征,如颜色、形状、纹理及对象的空间关系等信息,建立图像的特征索引,检索时从图像数据库中提取相似图像^[1]。然而,计算机自动提取的图像视觉特征和人所理解的图像语义间存在很大差距,即“语义鸿沟”。有效的图像检索必须充分利用图像的语义信息。鉴于此,有些学者提出基于语义的图像检索(Semantic-based Image Retrieval, SBIR)^[2],试图将可视特征识别为语义概念,使之模拟人类识别事物的方式。

图像分析和本体技术相结合进行图像语义标注是对视觉信息语义理解的一个有效途径。具有如下优点:1)本体可为图像语义标准化标注提供形式化基础,2)标注后的图像语义信息与领域本体相对应,使基于关系和公理的图像智能检索成为可能,3)将原来孤立的图像文件联系起来,加大了不同图像之间的耦合度^[3]。研究人员尝试利用本体表示图像特征,取得了一些成果^[3-9]。文献[4-6]将图像分割成不同的区域,然后提取每个区域的低层索引特征,通过本体将空间区域的低层索引特征和人类更熟悉的高层语义关联起来。文献[7]用领域本体的概念来描述 DICOM 医学图像的高级语义。

文献[9]将本体技术和基于文本的检索相结合,使用 RDF 三元组标注文件,并将基于本体的查询简化为布尔串检索,但是大大降低了查询能力。然而,上述基于本体对图像进行语义检索仍未取得很好的检索结果。原因有三:本体只是用来解决文本间概念异构;构建领域本体往往需要领域专家指导,花费大量额外的时间;通用本体太庞大,使结果产生冗余。

综上所述,构建合适的本体以及合理的相似性度量方法是图像语义检索的关键。由此,本文提出基于个性化本体的图像语义标注和检索方法,即:利用个性化本体描述图像语义,继而抽象基于概念集的图像语义特征并利用本体中 Is-A 关系设计相似性度量方法实现新的以本体为核心的图像语义标注和检索模型,并利用实验验证其性能。

1 图像语义标注

1.1 本体与图像语义标注

本体作为一种知识表示模型,包含了某个领域内的专家或学者所共同认可的词汇,是共享概念模型的形式化的规范说明。一幅特定的图像由一组用户感兴趣的对象及对象间的关系组成。其中:对象包含自身的属性信息;图像本体是对图像的一种描述模式,任何一幅图像都是图像本体的一个实例,

收稿日期:2009-06-24;修回日期:2009-09-05。

基金项目:仲恺农业工程学院校级科研基金资助项目(G3081807);广东省科技厅项目(2008B021300002)。

作者简介:史婷婷(1982-),女,安徽芜湖人,讲师,硕士,主要研究方向:图形图像处理; 闫大顺(1974-),男,河北大名,副教授,硕士,主要研究方向:程序原理、程序测试、智能控制; 沈玉利(1955-),男,山东费县人,教授,博士,主要研究方向:模式识别、智能系统、农业信息化工程。

并可用本体语言形式化表示,即:图像语义标注。

图1用一个有向无环图(Directed Acyclic Graph, DAG)表示学校组织本体, DAG中的节点代表概念 concept, 每一个概念包含三个基本元素:名称 term、属性 attribute 和关系 relation。如图1中概念 Teacher 包含一系列属性: name、degree 等。两个概念之间有连线表明这两个概念有关系。基本关系有3种: Is-A、Instance-Of 和 Part-Of。Instance-Of 表达概念的实例和概念之间的关系, 类似于面向对象中的对象和类的关系, 表示为概念 C_j 是概念 C_i 的一个例子。例如李教授是概念“Teacher”的一个实例; Part-Of 表达概念之间部分与整体的关系。概念 C_i 包含概念 C_j 也可以说 C_j 是 C_i 的一部分。Is-A 表达概念之间的继承关系, 类似于面向对象中的父类和子类之间的关系, 如概念 Faculty 和 Teacher, 其中 Faculty 为父概念, 而 Teacher 为子概念。Is-A 关系对于建立具有语义扩展能力的图像检索系统尤为重要。例如用户需要查询一副包含 Faculty 的图像, 则检索结果应包含与其子概念 Teacher 有关的图像, 这需要合理的相似性度量算法支持。

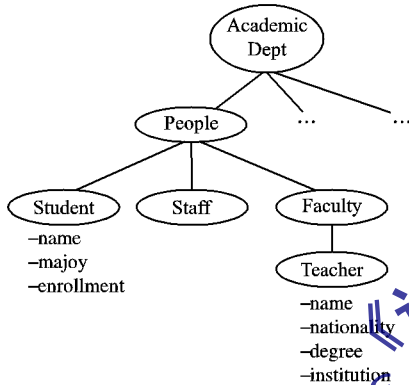


图1 学校组织本体示例

1.2 个性化本体构建

个性化本体, 顾名思义, 指用户利用其专业知识对现有本体中的概念、属性、关系进行修改、添加或删除。文献[3]提出对特定领域, 模型只提供顶级本体, 用户在使用过程中随时对本体进行修改、添加或者删除, 实现领域本体个人化。基于此, 模型在最初仅提供顶级本体, 描述最普遍的概念之间的关系。采用7分类^[10]: 自然现象、食物、动物、植物、人、自然对象、人造品。通过输入操作, 用户所描述的概念会自动加入现有本体, 由此建立个性化的理想本体。

假设: 用户需要存储并检索有关“王励勤”的图像。首先, 通过互联网或其他媒体找到“王励勤”的图像, 并根据自己的理解以自然语言描述图像信息, 如图2所示。该信息被分解为三个语义层次。

- 对象、场景描述: object: player, pingpang; background: stadium;
- 对象及背景属性描述: player→profession: table tennis; player→nationality: China; player→full name: Liqin Wang;
- 对象间的语义关系, 即 player→playing→pingpang。

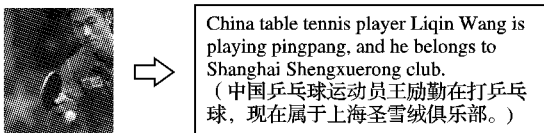


图2 例图及自然语言描述

接着, 个性化本体构建模块通过自然语言解析工具向顶级本体中添加相关概念如“Liqin Wang”等, 如图3所示。

重复上述步骤, 用户继续描述图像信息, 顶级本体越来越大并逐步进化为个性化的乒乓球运动员本体, 而与该本体有关的图像也同步存入图像语义库中, 如图4所示。

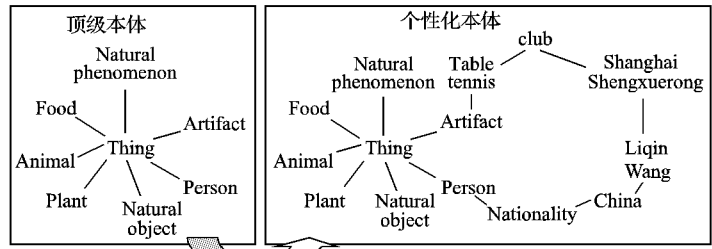


图3 顶级本体向个性化本体进化

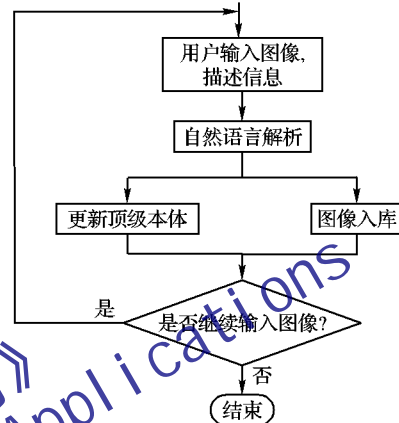


图4 个性化本体构建流程

1.3 标注文档生成

目前, 本体主要以 RDFS 文件或者 OWL 文件的形式存储, 这些文件可由相关专家根据标准的格式以 XML 基本语法手工编辑, 也可由本体构建工具自动导出生成。本文利用斯坦福大学开发的开源的 Protégé3.3 构建本体, 结合 Jena 开发包实现了图像语义标注文档的自动生成。对图2中的图像进行语义标注文档片段如下:

```
<owl: Class rdf: ID = "artifact" >
  <rdfs: subClassOf rdf: resource = "thing"/>
</owl: Class>
<owl: Class rdf: ID = "nationality" >
  <rdfs: subClassOf rdf: resource = "artifact"/>
</owl: Class>
<owl: Class rdf: ID = "China" >
  <rdfs: subClassOf rdf: resource = "nationality"/>
</owl: Class>
...
<owl: Class rdf: ID = "Liqin Wang" >
  <owl: intersectionOf rdf: parseType = "Collection" >
    <owl: Class rdf: about = "Shanghai Shengxuerong" />
    <owl: Class rdf: about = "China" />
  </owl: intersectionOf>
</owl: Class>
```

由此可见, 图像语义标注依据个性化本体概念的层次结构, 自上而下进行, 实现了图像语义对象与本体中概念的准确对应, 并标准描述了图像所传达的语义信息。

2 图像语义检索

通过上述语义标注, 图像语义库中的图像由一系列概念描述, 且与 OWL 文档一一对应。基于此, 本文利用概念集构造图像语义特征。同时, 图像检索的目的在于检索与用户需

求尽可能相似的图像,所以选取合适的相似性度量方法也是关键。因而,图像检索包含三个子模块,如图5所示。

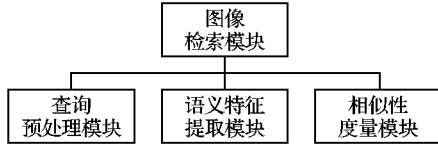


图5 图像检索模型框架

2.1 查询预处理

本体不仅能用于语义标注,还可帮助用户明确查询条件。当用户对所查询的领域较陌生时,通过语义浏览可只使用本体中已定义的概念描述查询对象;同时,允许用户对相应的概念设置属性以此限定查询范围;当用户对该领域知识较熟悉时,可基于自然语言检索,简单如“Liqin Wang”或者复杂如“The table tennis player who is a Chinese, and is member of Shanghai Shengxuerong team”。选择上述两者中任一检索方式并提交查询条件后,系统自动建立一个OWL文件规范查询条件,从而与图像语义库中现有的图像所对应的OWL文件作语义相似性度量。

2.2 图像语义特征提取

OWL文件是结构化文档,很容易从元素、标记及属性中抽取有用的关键词。本模型采用DOM解析文档并在内存里创建对文档内容的树状表达式,程序从树的顶部开始遍历逐一读取相应图像所涉及的概念。

从图2中提取概念“Linqin Wang”以及一系列相关属性“table tennis”、“Shanghai Shengxuerong”、“China”,由此抽象出图像涉及的语义概念及其属性和概念间的空间关系形成一个一维语义特征向量 SF ,如式(1)所示:

$$SF = (C_1, C_2, \dots, C_n)$$

其中 $C_i (i = 1, \dots, n)$ 代表图像所对应的概念。

每个概念都有一个核心名称term和相关属性attribute,故 C_i 进一步细化为式(2):

$$C = (T(A_1, A_2, \dots, A_m)) \quad (2)$$

其中: $A_i (j = 1, \dots, m)$ 代表图像所对应的属性, T 代表概念名称。

图2中的图像语义特征可表达为:(Linqin Wang(table tennis, Shanghai Shengxuerong, China))

2.3 相似性度量

经过上述步骤,图像语义库中所有图像和查询条件都被转换为概念集,接下来便比较图像和查询条件的概念集间的语义相似度量,从而得到符合条件的图像结果集。

假设 $SF_{image} = (C_{image,1}, C_{image,2}, \dots, C_{image,m})$ 和 $SF_{query} = (C_{query,1}, C_{query,2}, \dots, C_{query,n})$

分别描述图像和检索查询的概念集,二者间的相似性概率 $SP(SF_{image}, SF_{query})$ 如式(3)所示:

$$SP(SF_{image}, SF_{query}) = \sum_{i=1}^n MP(C_{query,i}) \quad (3)$$

其中满意概率 $MP(C_{query,i}) (i = 1, \dots, n)$ 代表查询条件中每个概念的满意概率。

由此,计算图像和查询条件之间相似性就转化为计算满意概率。对于查询条件概念集中的每一个概念 $C_{query,i}$,将其与图像概念集中所有概念进行对比,找出最大的相似概率即为 $MP(C_{query,i})$,定义为式(4):

$$MP(C_{query,i}) = \max(PS(C_{image,k}, C_{query,i})) ; k = 1, \dots, m \quad (4)$$

相似性度量进一步转化为计算两个概念之间的相似概率 $PS(C_{image,k}, C_{query,i}) ; k = 1, \dots, m ; i = 1, \dots, n$ 。

在本体中,概念之间具有类层次的上下位关系、同义关系等,概念和概念之间通过这些关系实现连接组成网状结构。下面分两种情况讨论概念之间的相似概率。

假设:两个概念 $C_1 = (T(A_{1,1}, A_{1,2}, \dots, A_{1,p}))$ 和 $C_2 = (T(A_{2,1}, A_{2,2}, \dots, A_{2,q}))$

1) 对于 $\forall A_{1,i} (i = 1, \dots, p)$,都 $\exists A_{2,j} (j \leq q)$ 与之属于同一概念,只是属性值不同,则 $PS(C_1, C_2) = 0$ 。

2) 当1)不成立的时候, $PS(C_1, C_2) = R * (1/q) * \sum_{i=1}^q PF(A_{2,i})$ 。

$$R = \begin{cases} 1, & C_1 \text{ 是 } C_2 \text{ 的子概念} \\ \frac{NUM_2}{MUM_1}, & C_1 \text{ 是 } C_2 \text{ 的父概念} \end{cases}$$

其中: $PF(A_{2,i})$ 是概率函数,当 $\exists A_{1,j} (j \leq p)$ 满足条件 $A_{2,i}$ 和 $A_{1,j}$ 是同一概念并且值相同,则 $PF(A_{2,i}) = 1$;否则,需要根据概念 $A_{2,i}$ 的值域计算 $PF(A_{2,i})$ 。因为不同概念的值域不同,所以计算需对应于某一具体的属相概念。 NUM_1 和 NUM_2 分别代表 C_1 和 C_2 的所有子概念数目。

3 基于个性化本体图像语义标注和检索模型

3.1 模型结构框架

采用前述原理,基于个性化本体的图像语义标注和检索模型如图6所示,包含两大模块:图像语义标注、图像语义检索。具体的工作流程及基本思想概括如下。

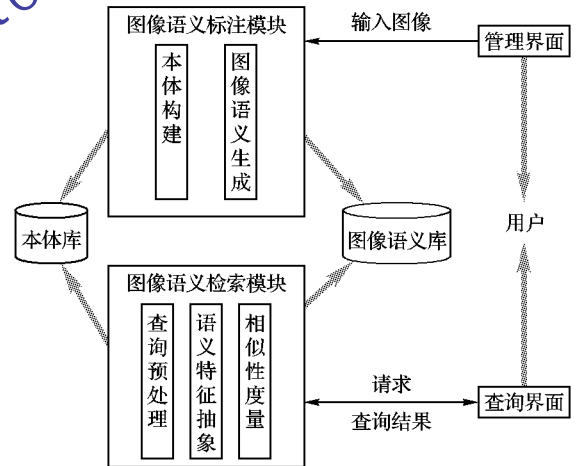


图6 基于个性化本体的图像语义标注和检索模型

1) 构建本体。本体是信息组织的核心,定义了知识的概念层次、概念与概念之间的关系。故构建检索模型首先要建立理想的本体,本文利用个性化的本体构造本体库。

2) 描述图像语义。本模型的图像语义采用W3C推荐的OWL描述。

3) 查询预处理。接受用户的查询请求,参照已经构建的本体库,将用户的查询条件映射为本体中的概念,并为其建立一个OWL文件规范查询条件。

4) 语义特征提取。采用DOM解析OWL文件,抽象出其所涉及的语义概念形成一个一维语义特征向量表征图像和查询条件。

5) 相似性度量。概念之间具有类层次的上下位关系、同义关系等,并通过这些关系实现连接组成网状结构。基于本体中Is-A关系设计相似性度量算法,实现语义扩展。

3.2 实验和结果分析

3.2.1 检索结果

为检验模型性能,本文基于 eclipse 开发平台,选用 Protégé3.3 作为本体建模工具,并采用了惠普实验室开发提供的 Jena2.1 API 接口方法构建检索实验系统。

系统中存入 1000 幅有关乒乓球运动员的图像继而建立一个关于乒乓球运动员的个性化本体库。对于若干种查询均取得了预期结果,图 7(a)是以“China table tennis player”为检索条件的结果集,图 7(b)是以“China table tennis player and a member of Shanghai Shengxuerong club”为检索条件的结果集。



图7 结果集

3.2.2 模型的性能评价

为了进一步说明本文所提出模型的有效性,使用三个测试系统:基于关键字的图像检索:Baidu 图片搜索;基于通用本体 WordNet 的图像检索以及本文基于个性化本体的图像检索。针对下列四类查询条件做检索。利用 Baidu 搜索图片时,将四类查询翻译为中文进行检索。

- 1) Linqin Wang;
- 2) China table tennis player;
- 3) China table tennis player is playing pingpang;
- 4) China table tennis player and a member of Shanghai Shengxuerong club.

每个系统的图片集不同,无法比较查全率,故本文使用查准率衡量每个系统检索结果优劣,公式如下:

$$PR = n/T \quad (5)$$

其中: N 为图像检索系统自动检索输出的包含在 N 中的图像数目, N 为人眼主观从图像库中找出的与图像 Q 相似的图像数目, T 为图像检索系统自动检索输出的总的图像数目。

从检索结果图像集来看,本文所提之模型可以检索出语义上相关联的图像。同时,从图 8 可知,根据检索模型实现的系统较 Baidu 图片搜索和基于 WordNet 的图像检索系统而言,在简单查询中优势并不明显,如查询条件 1。查询条件越复杂,检索准确性能越好,且平均查准率高达 88.6%。因此,个性化的本体以及基于概念集的图像语义特征表达有助于寻找更多在语义上相关的图像,是解决目前图像语义检索的有效方式。

4 结语

在基于语义的图像检索中,图像语义特征描述和相似性度量是检索系统的中心环节。现有的低层特征模型对图像的描述能力仍然远远无法涵盖基本的图像内容,使得语义信息

的标注和检索成为图像检索领域近年来的一个热点问题。在各种方法的启发下,本文采用一种全新的处理流程实现基于本体的图像语义标注和检索,其中:

- 1) 针对通用本体和领域本体在构建和表达图像语义方面的缺陷,提出使用个性化本体描述图像语义,构建本体的同时,完成对图像语义的标注;
- 2) 提出一种新的图像语义特征表达,即利用图像本体中定义的概念描述图像语义,形成基于概念集的图像语义特征,且完全独立于具体本体;
- 3) 提出基于“Is-A”关系的相似性度量法则,形成具有语义扩展功能的图像检索模型。

通过与基于关键字的图像检索系统(Baidu 图片搜索)和基于通用本体 WordNet 的图像检索系统进行对比分析,本文的检索模型取得了很好的检索效果,提高了图像检索的效率、精确度、智能性。

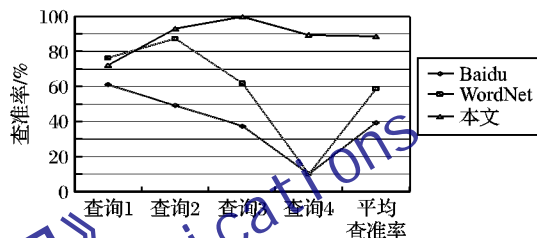


图8 四类查询条件的查准率比较和三个系统的平均查准率

参考文献:

- [1] KRODLEY C E, KAK A C, DY J G, *et al.* Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision [C]// Proceedings of the 16th National Conference on Artificial Intelligence. Washington, DC: IEEE Computer Society, 1999: 760 - 767.
- [2] WU YI, ZHUANG YUE-TING, PAN YUN-HE. Web image retrieval system combining with semantics [J]. Pattern Recognition and Artificial Intelligence, 2001, 14(2): 156 - 160.
- [3] 邓涛, 郭雷, 杨卫莉. 基于本体的图像语义标注与检索模型[J]. 计算机工程, 2008, 34(17): 188 - 190.
- [4] 钟洪, 夏利民. 基于本体的图像检索[J]. 计算机工程与应用, 2007, 43(17): 37 - 40.
- [5] HU B, DASMAHAPATRA S, LEWIS P, *et al.* Ontology-based medical image annotation with description logics [C]// Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence. Sacramento: IEEE Computer Society Press, 2003: 77 - 82.
- [6] GRUBER T R. A translation approach to portable ontology specifications[J]. Knowledge Acquisition, 1993, 5(2): 199 - 211.
- [7] 刘亮亮, 倪天权. 基于本体的 DICOM 医学图像检索[J]. 微计算机信息, 2007, 23(27): 296 - 297, 235.
- [8] 吴静. 基于本体的画藏检索系统设计与实现[D]. 北京: 北京交通大学, 2006.
- [9] MAYFIELD J, FININ T. Information retrieval on the semantic Web: Integrating inference and retrieval [C]// Proceedings of the SIGIR 2003 Semantic Web Workshop. Toronto: ACM press, 2003: 84 - 92.
- [10] KONG H, HWANG M, KIM P. The study on the semantic image retrieval based on the personalized ontology [J]. International Journal of Information Technology, 2006, 12(2): 35 - 45.