

图像数据的结构化协同标注与检索

朱 麟, 高丽萍, 卢 瞰

(复旦大学计算机科学技术学院协同信息与系统实验室, 上海 200433)

摘 要:通过分析图像数据对语义信息的检索需求, 提出一种表达图像信息的结构化方法, 并定义一种根据图像中隐含语义信息来判定相似度的算法。该方法在一定程度上克服了基于内容的图像检索和基于文本的元数据图像检索这 2 种常用图像检索方法缺乏语义信息匹配上的不足。原型系统根据结构化协同标注中语义信息的丰富程度与图像检索的精确程度, 证明了该方法的有效性。

关键词:协同标注; 图像检索; 结构化; 相似度

Structural Collaborative Annotation and Retrieval of Image Data

ZHU Lin, GAO Li-ping, LU Tun

(Cooperative Information and Systems Lab, School of Computer Science, Fudan University, Shanghai 200433)

【Abstract】 By analyzing image data's demands for semantic information, a structural method is adopted to express information in images, and a similarity algorithm which contains semantic information to help retrieve images is defined. To a certain extent, this method overcomes limitations of two common image retrieval methods, i.e., content based image retrieval, and text based metadata image retrieval, both of which lack semantic information for match. The prototype system proves its effectiveness through fruitful semantic information of collaborative annotation and accuracy of image retrieval.

【Key words】 collaborative annotation; image retrieval; structural; similarity

1 概述

随着互联网的发展, 图像逐渐成为一种非常重要的信息来源, 如何在大量图像中检索自己所需的图像, 成为信息检索的一个热点问题。传统的图像检索主要有 2 种方法: 基于文本的元数据图像检索及基于内容的图像检索(CBIR)。本文关注的是在一种专业领域的社区中, 利用大量专家的协同标注内容提供领域相关图像的检索。文献[1]构建了一种对图像进行语义标注的工具, 在标注的同时提供用户一种本体发育的方法。画册标注^[2]设计了一种基于专业本体的特定领域图像标注和检索的工具。根据图像某一部分对象来检索含有类似对象的其他图像的功能, 而这种检索的基础是用户的标注。基于本体的图像检索^[3]提出了标注图像中蕴含对象的方法, 并以此作为检索的依据。这些方法的提出能够从一定程度上满足用户图像检索的需求, 但对于那些信息量大的图像, 单纯提取特征或者使用关键字查询的做法并不能满足用户语义检索的需求。实际上计算机并不能理解图像内在的意义, 不同人对同一张图像往往有不同的认识, 而这种不同很大程度上源自人们价值观的不同。单纯用提取特征的 CBIR 并不能反映这种差别。一个图像除了各种对象外, 还潜在地包含着对象间的结构化信息; 而以整个图像为单位的标注往往不能表示这些信息。此外, 不同领域的图像千差万别, 它们所对应的语义信息也有很大的不同。所以本文的观点是将特定领域的价值观反映到图像中。

本文提出了一种让用户自由选定局部范围对象的图像标注方法, 并提供对象间关系的标注, 目的是抽取图像的概念及结构化信息。标注的手段是让互联网上大量用户协同地标注图像中对象的概念及概念间的关系, 并用 RDF 格式保存元数据, 最后利用这些元数据来进行检索。

2 结构化协同图像标注

图像标注的目的在于利用标注结果来反映图像的内容, 以达到方便图像检索的目的。而与普遍采用的以整张图像作为单位的加 tag 方式不同, 这里提供了一种用户自定义的标注方式: 利用矩形框选定图像的某个区域, 标注上“概念”。用户还可以标注选定对象所代表概念之间的关系。标注结果都被获取并作为元数据保存在 RDF 文件中。在专业领域社区的图像标注中, 概念及其关系通过领域本体库来限定。下面用一个简单的例子来说明。

2.1 图像中自定义对象的标注

在给定的图像上, 用户可以用矩形框任意选定范围, 并做出自己的标注。

图 1 是一个自定义对象标注的例子。标注的内容包括对象的关键字及该关键字的属性。所以图中标注“Jason Schiffman”和“Siobhan”的对象内容为

```
Father
  -name: Jason Schiffman
  -occupation: soldier
Mother
  -name: Siobhan
  -occupation: housewife
Son
  -name: Darren
```

概念项是对该对象的描述性词汇。这里, “father”和“mother”这 2 个概念均出自自己已经给定的亲属概念本体库

作者简介:朱 麟(1984 -), 男, 硕士研究生, 主研方向: 计算机支持协同工作; 高丽萍, 博士研究生; 卢 瞰, 博士

收稿日期:2009-03-10 **E-mail:** zhulin1984@fudan.edu.cn

Family。



图 1 标注实例

2.2 对象间关系的标注

在标注自定义对象的基础上，本方法还提供给用户自由标注对象间关系的功能。比如上面那对 couple 之间的关系：

```
Relation
-object1: Father
-object2: Mother
-relation: Couple
```

提供给用户的工具是在“关系”中输入 2 个对象的标识符及关系的名称。关系概念的选择来自已有的亲属关系本体库 Family-Relation。这里的“couple”关系来自这个本体库。

2.3 协同标注以及冲突的解决

由于标注是在互联网上进行，因此所有用户标注的内容是共享的。相比单人做标注，多人标注的语义信息更加丰富。但在标注的过程中，不同人对同一图像的理解有偏差。冲突的来源有以下几点：

- (1) 选定区域的冲突；
- (2) 概念上的冲突。

如果选定范围重叠率高达 90%，认为他们选定的是相同的对象。这样的结果是一个对象可能具有多个概念。这些不同的概念都认为是不同的标注。

2.4 标注结果和元数据的存储

标注的结果以 RDF 的格式存储，作为检索的元数据。自定义对象的标注结果存储于 concept.xml，而对对象间关系的标注结果存于 relation.xml。

对图 1 例子的对象标注结果 concept.xml 代码如下(只列出了一个对象)：

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="rdf-syntax-ns#" xmlns:family="family#">
<rdf:Description rdf:about="family/father">
<family:name>Jason Schiffman</family:name>
<family:occupation>
Soldier
</family:occupation>
</rdf:Description>
...
</rdf:RDF>
```

而对于关系的标注，Relation.xml 保存对象间的关系，他们以三元组的形式：概念-关系-概念的形式存储。标注的最后结果是一个三角状逻辑结构：3 个顶点表示 3 个对象，而 3 条边分别表示对象间的关系。

3 基于标注内容的图像检索

本文中图像的含义都是通过对象的概念以及概念间的关系获取，因此，通过以上的标注，结构化立体的图像内容就

被采集了出来。图像检索是通过比较用户输入的查询语句(query)和目标图像库中的所有图像，获取相似度符合要求的那些图像。用户可以输入一个含有关键字及其关系的查询，也可以通过某个已经被标注过的图像来检索具有类似内容的图像。

3.1 元数据中特征的抽取

因为语义标注的基础是概念和关系，而概念和关系的基础是 2 个本体库，比如上面例子中的 Family 和 Family-Relation 库，所以需要根据这些概念与本体库的关系抽取信息。

对于形式化的 RDF，可以通过 RDF parser 提取出概念。图 2 对应的 concept.rdf 中具有概念“Mother”和“Father”及其相关属性“name”和“occupation”。这样，用 $CS = (c_1, c_2, \dots, c_n)$ 表示标注所使用到的概念集合。 c_i 表示某个特定的概念项(term)。此外，因为每个概念都有属性，所以必须将这些属性包含进来，于是用 $CAS = (ca_1, ca_2, \dots, ca_m)$ 表示包含属性的概念集。其中， $ca_i = (c_i, (a_1, a_2, \dots, a_k))$ ， a_m 表示 c_i 的第 m 个属性。同理，关系集合 $RS = (r_1, r_2, \dots, r_n)$ ，其中， r_i 表示关系。

3.2 Query 和目标图像集的相似度比较

用户经常碰到的图像检索问题之一就是已有一张图片但不能简洁地概括它，而他却需要搜索和此相似的图片。这个问题实际上就是由一个形式化 query(具有 rdf 元数据)，去逐一匹配目标图像的元数据。对 2 个结构化标注的比较，目标是抽取结构信息，最后转化为概念之间的比较。

整个图像元数据的内容可以用多个三元组来表示，这些三元组的形式诸如：概念-关系-概念。于是可以将整个 rdf 元数据通过 RDF parser 提取概念后拆分成多个 Triple 的形式。这里用 $Triple(a-R-b)$ 表示三元组。

定义 2 个图像 query 与 image 的相似度：

$$Sim(query, image) = \sum_i \sum_j sim(Triple(a_i - R_i - b_i), Triple(a_j - R_j - b_j))$$

Triple 相似度计算方法见下节。

3.3 三元组 Triple 之间的相似度计算

对于 2 个三元组 $Triple(a-X-b)$ 和 $Triple(c-Y-d)$ ，有 $sim(Triple(a-X-b), Triple(c-Y-d)) = sim(concept(a-b), concept(c-d)) \times 2^{sim(X,Y)}$

当关系相似度 $sim(X, Y)$ 为 100% 时，整个相似度就在概念相似的基础上翻倍；若 $sim(X, Y) = 0$ ，那么只需要考虑概念相似度。

而 $sim(concept(a-b), concept(c-d))$ 需要考虑 $a-c$ 、 $b-d$ 和 $a-d$ 、 $b-c$ 的搭配，那么有

$$sim(concept(a-b), concept(c-d)) = \max \{ w_a w_c sim(a, c) + w_b w_d sim(b, d), w_a w_d sim(a, d) + w_b w_c sim(b, c) \}$$

这里考虑对象区域面积的影响：

$$w = \frac{S(selected_region)}{S(image_region)}$$

3.4 2 个概念间的相似度计算

上文已经提到，标注后的概念可以表示为

$$ca = (c, (a_1, a_2, \dots, a_k))$$

其中， c 为概念项； a_i 为附带属性。现在考虑 2 个概念： $ca_a = (c_a, (a_{a,1}, a_{a,2}, \dots, a_{a,p}))$ 和 $ca_b = (c_b, (a_{b,1}, a_{b,2}, \dots, a_{b,q}))$ 之间的相似度。

对于一组概念集 c_1, c_2, \dots, c_s ，如果每个 c_i 是 c_{i+1} 的子概念，

也就是本体树中具有父子关系的节点。这样， c_1 是 c_s 的派生概念，而 c_s 是 c_1 的祖先概念。根据这个定义，考虑概念项及其属性：

$$\text{sim}(ca_a, ca_b) = \text{Rate} \times \frac{1}{q} \times \sum_{i=1}^q \text{Pr}(a_{b,i})$$

在 c_a 和 c_b 是相同的概念或者前者是后者的派生概念时， Rate 值为 1； c_a 是 c_b 的祖先概念时， $\text{Rate} = \frac{\text{Num}_a}{\text{Num}_b}$ 。 Num_a 和 Num_b 分别是 c_a 和 c_b 的派生概念总数。此外， $\text{Pr}(a_{b,i})$ 是属性 $a_{b,i}$ (属性本质上也是一个概念词条) 被满足的概率。如果 2 个属性概念 $a_{b,i}$ 和 $a_{a,i}$ 是相同的概念，并且具有相同的值，则 $\text{Pr}(a_{b,i}) = 1$ ；如果没有属性概念 $a_{a,i}$ ，使得 $a_{b,i}$ 和 $a_{a,i}$ 是相同的概念，就可以将它们放入特定领域后的特殊关系来计算 $\text{Pr}(a_{b,i})$ 。总之，如果是上面这种情况，而且不是在特定领域的情况下，就认为 $\text{Pr}(a_{b,i}) = 0$ 。

4 原型系统及实验评估

本部分根据前面的方法建立了一个足球领域的社区，提供用户对球星和俱乐部相关的图像进行标注和检索。系统需要 2 个本体库：球星及足球术语本体库和描述足球相关关系的本体库。

4.1 系统架构和实验方法

标注系统采用 JS 的技术创建，浏览器是 Firefox，系统包括标注工具和检索工具。标注工具可以让用户对选定图像进行概念和关系的标注；而检索工具提供用户根据已有图像搜索具有相似结构和内容的图像；标注界面如图 2 所示。其中，左边的图像中的小框是选定对象；中间工具栏包括 2 个选项“加标签”和“加关系”；这里 2 个小框以及关系表示的语义是“卡纳瓦罗-举起-世界杯”。



图 2 标注系统界面

本系统将 1 000 张足球相关图像放到局域网，让不同的足球爱好者进行标注。

实验分成 2 个部分：(1)将结构化标注结果和以往普通加文本 tag 结果相比较来检测其对图像信息精确度的影响。(2)将多人协同标注的结果和单人标注结果相比较来检测其对图像信息精确度的影响。查询的精确度使用查准率来衡量。

4.2 实验结果评估

实验(1)和实验(2)各使用 5 次查询，图 3、图 4 是查询结

果比较。由图 3 可以看出，后者效果明显好于前者，其中 Query2 甚至将结果提升了 30 个百分点。图 4 表示协同标注和单人标注对精确度的影响，协同标注所含的语义信息略多于单人标注。

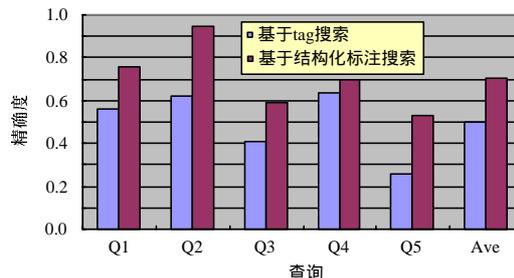


图 3 基于结构化标注与基于 tag 检索精确度的比较

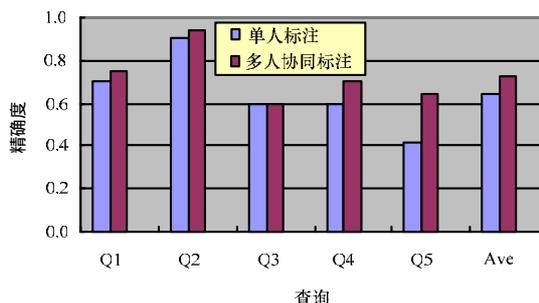


图 4 多人协同标注和单人标注对精确度影响的比较

5 结束语

本文介绍了一种提取图像内部结构化信息的协同标注方法，以及基于该标注的相似度算法。通过 2 个比较实验：(1)本方法与普通增加 tag 标注方法对检索精确度的影响；(2)多人协同标注与单人标注对检索精确度的影响，从一定程度上展示了本方法的优越性。图像的结构化信息能很好地将人的认识和图像内在关系表达出来，因此，提高了检索效率。

图像数据的结构化协同标注的是一种新的尝试，主要面向专业领域。如何让这种方法在通用的情况下有更好的效果，以及设计一种怎样的游戏，让用户去标注图像，也是值得研究的课题。此外，如何改进概念的精确度以及概念间关系的精确度，如何改进检索效率等，也是非常重要的课题。这些将在未来的工作中继续研究。

参考文献

- [1] Braun S, Schmidt A, Walter A. Ontology Maturing: A Collaborative Web 2.0 Approach to Ontology Engineering[C]//Proc. of the Workshop on Social and Collaborative Construction of Structured Knowledge. Banff, Canada: [s. n.], 2007.
- [2] Hollink L. Semantic Annotation of Image Collections[C]//Proc. of the Workshop on Knowledge Capture and Semantic Annotation. Florida, USA: [s. n.], 2003.
- [3] Zheng Wei, Ouyang Yi, Ford J. Ontology-based Image Retrieval[EB/OL]. (2008-05-10). <http://devlab.dartmouth.edu/labpapers/accepted/WSEAS03/ontologyImage.pdf>.

编辑 顾逸斐