

基于外部信息源的 WWW 图像语义提取研究

张华¹ 张森¹ 孟祥增²

(哈尔滨工业大学(威海)计算机科学与技术学院 威海 264209)¹

(山东师范大学传播学院 济南 250014)²

摘要 HTML文档作为 WWW 图像的外部信息源和载体,蕴涵了丰富的描述图像内容的文本信息。为了实现基于语义的 WWW 图像检索,本文提出了一种 WWW 图像语义表征模型和图像语义词典的建设方法,给出了一种利用图像语义词典从嵌有 WWW 图像的 HTML 网页的相关外部文本信息中提取 WWW 图像语义信息的具体方法和实验结果。

关键词 图像语义,图像分类,信息提取,语义表征模型

Methods of Extracting WWW Image Semantics Based on External Information

ZHANG Hua¹ ZHANG Miao¹ MENG Xiang-Zeng²

(School of Computer Science and Technology, Harbin Institute of Technology at Weihai, Weihai 264209)¹

(School of Communications, Shandong Normal University, Jinan 250014)²

Abstract As external information sources and carriers of WWW image, HTML documents contained plentiful text information which related closely to image's semantics. A model of representing the semantics of WWW image and a method of constructing the dictionaries of the semantics of WWW image are presented to realize image retrieval based on semantic. A method of extracting the keywords describing the semantics of WWW image from the related texts in HTML documents based on the dictionaries is introduced and the results of primary experiment are shown in the paper.

Keywords Image semantics, Image classification, Information extracting, Model of representing semantics

1 引言

目前图像语义的提取主要有如下三种方法^[1]: 1) 基于知识的语义提取: 首先提取原始图像的视觉特征, 并进行图像分割, 然后利用对象模板、图像场景分类器等知识, 识别出用于描述图像的主题词、主体词及其属性词等语义信息。2) 人工交互语义提取: 先结合基于知识的方法, 自动学习关于原始图像的主题词、主体词及其属性词的描述, 再由人工进行修正。3) 基于外部信息源的语义提取: 根据图像来源处的其它信息提取与图像相关的相对高层的描述信息。

限于当前图像识别和理解的技术水平, 完全靠图像的视觉特征获取 WWW 图像的语义还相当困难。同时, 对于海量的 WWW 图像而言, 人工交互的语义提取又显得微不足道。WWW 图像的最大特点是嵌入在 HTML 文档中, HTML 文档中的文本内容作为 WWW 图像的外部信息源与其语义信息有着紧密的联系。因此, 采用第三种方法利用自然语言处理技术提取 WWW 图像语义信息是目前 WWW 图像检索采用最多的策略。

本文提出一种基于外部信息源的 WWW 图像语义提取方法, 可以为图像的语义描述和利用图像处理技术进一步理解图像提供重要的文字线索。

2 WWW 图像语义表征

网页中的图像包含三个方面的属性^[2,3]: 文件属性、视觉

属性和语义属性。图像的文件属性, 指图像文件名、类型、大小、建立时间、URL 等确切简单的数据; 图像的视觉属性, 指关于图像或图像中物体的颜色、形状、纹理、位置、大小、方向等直观的视觉特征; 图像的语义属性, 指图像所表现的主题、事件、场景以及图像中物体的名称、姿态、空间关系等语义信息^[4]; 视觉特征的中文描述, 也看作是图像的语义属性。

根据图像视觉属性与语义属性的含义, 按照从简单到复杂、从整体到局部、分层次逐渐级描述的原则, 建立了如图 1 所示的 WWW 图像语义表征模型。

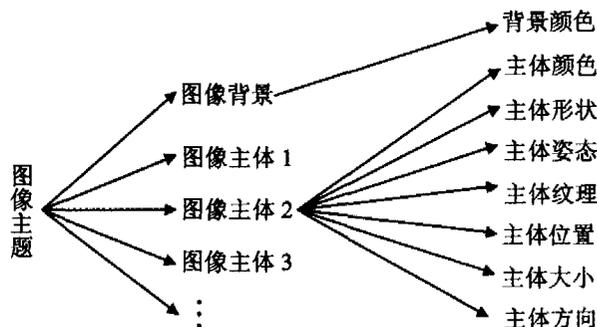


图 1 WWW 图像语义表征模型

图像语义表征模型首先描述的是图像主题, 然后描述图像的结构成分, 即图像的背景、主体, 进而再描述背景颜色及主体的颜色、形状、姿态、纹理、位置、大小、方向。其中, 图像

张华 硕士, 研究方向: 数据挖掘, ERP, 多媒体技术。张森 硕士, 研究方向: 人工智能, ERP。孟祥增 教授, 博士, 研究方向为: 多媒体检索、图像识别、CAI。

主题指概括图像内容和寓意的关键词、短语或句子,图像内容为图像突出表现的主体或场景,图像寓意为图像表现的深层意义,通常是抽象的概念;图像背景指对图像背景的描述,如天空、模糊或单一颜色等;背景颜色指背景的平均颜色;图像主体指图像突出表现而又明显区别于其背景物体,用名词表示;主体颜色指主体的主要颜色;主体形状指对主体的形状描述,如长、短、尖、方等;主体姿态指主体的形态,如站、卧、跑、飞等;主体纹理指主体表面的纹理特性,如条纹、网格、鱼鳞状等;主体位置指主体中心在图像画面中的位置;主体大小指主体在图像中的相对大小;主体方向指主体中心轴或头尾的方向,如水平、垂直、斜上等。

3 图像语义词典

根据图像语义表征模型,本文分别建立了图像主题词分类词典、图像主体词分类词典、图像主体属性词典和图像主题词对照词典,以便利用自然语言处理技术从 WWW 图像所在网页的相关外部文本信息中提取图像的主题词、主体词及其属性词等语义信息。

3.1 图像主题词分类词典

3.1.1 图像主题词的分类原则与分类结构 描述图像的主题词源于表现图像主题的词组或句子中的名词、动词和形容词,并按如下原则分类:

- 按图像的主题意义、所属领域,采用层次型分类结构。分类用路径形式表示,结构为:一级分类类名|二级分类类名|三级分类类名|……,父类包含子类,子类继承父类的属性,每个分类类名即为一个图像主题词。

- 分类不求太细,尽量全面,符合人们的一般认识,并非严格按学科分类。如植物分为:农作物、菜、花、树、藤、竹、草、叶、蕨、苔、藻、菌、病毒,而非严格按生物分类学的纲、目、科、属、种来划分的。

- 同义词用“主题词;主题同义词;主题同义词;…”表示,做同一类名处理。如“植物|树|红杉树;红杉”中的“红杉树”与“红杉”就是同义主题词。

- 主题分类体系是开放的,允许人为添加,也允许从网页中自动学习添加。

- 不同的分类路径可以包含相同的类名,如“物品/器材/运动器材/足球”类与“场景/事件/体育/球类/足球”类中都有“足球”,但属性完全不同。

根据以上原则,把 WWW 网页中描述图像的主题词分为人物、动物、植物、自然、建筑、物品、场景、创作共 8 类^[5,6],每类又分若干子类,按层次逐级展开,直至具体主题词,如图 2 所示。

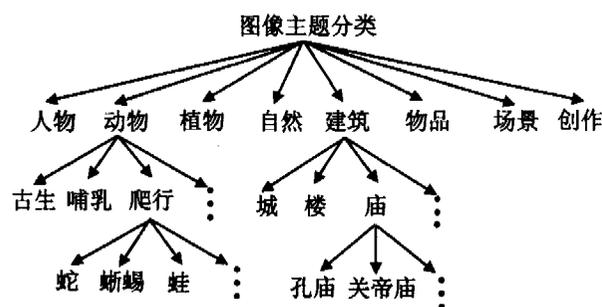


图 2 图像主题分类结构

3.1.2 图像主题词的在线学习 网页中的图像是海量

的,描述图像的主题词也难以穷举,所以,允许在线学习网页中的未登录主题词,并根据主题词典的分类结构,自动添加到相应的分类路径中^[7]。在线学习的语料来源即为与 WWW 图像语义密切相关的 HTML 文档中的外部文本信息。主题词在线学习过程如下:

1) WWW 图像外部信息提取。

2) 将图像的外部信息进行分词。

3) 自动组词以发现新的主题词,组词规则为:当名词前是名词或形容词时组合成新词,如“美洲(n)|虎(n)”合并为“美洲虎”,“红(a)|珊瑚(n)”合并为“红珊瑚”。如果新组合的词在主题词典中不存在,表示发现了一个新主题词,并按如下规则自动分类与添加:

4) 若词尾是分类结构中的某个类名,则该词属于该类,并作为子类自动添加到该类中,如“美洲虎”自动分类并添加的结果为“动物|哺乳动物|陆生|肉食|虎|美洲虎”。

5) 对于人名,将其添加到“人物/现代人物/普通人物”中。

6) 对于地名,若为地图,将其添加到“创作/绘图/地图”中,否则将其自动添加到“场景/风景、夜景、街景、鸟瞰”中。

7) 对于绘画、绘图名,将其添加到相应的“创作/绘画、绘图”类中。

8) 对于动词、形容词和非实物名词,将其添加到“创作/艺术摄影”类中。

图像主题是图像语义的重要成分,通过几个月的人工和自动添加,图像主题词典中已有 4 万多条主题词分类路径,而且随着在网络上的不断运行而不断增加。

3.2 图像主体词分类词典

图像主体词分类词典采用与图像主题词分类词典相同的分类原则和相似的分类结构,把图像中常见的主体分为人物、动物、植物、自然物、人造物、图形 6 类,每类又分若干子类。图像主体词分类词典同样也是主观开放的,但规模要小得多。

3.3 图像主体属性词典

图像主体属性词典列出对图像主体视觉特征的中文描述词,包括 262 个颜色词、82 个形状词、18 个纹理词、15 个姿态名词、11 个位置名词和 8 个方向名词。对于颜色词,大多数对应一组 R、G、B 值,个别颜色词对应几组 R、G、B 值,如肤色、蓝紫色等,而不同的颜色词也可能具有相同的 R、G、B 值,如紫红和洋红、棕和褐、蓝和兰等。

3.4 图像主题词对照词典

考虑到图像文件名、图像 URL、图像所在网页 URL、图像超链接 URL 等外部信息中多有汉语拼音、英文单词或缩写词,特建立了图像主题词对照词典,标出每个中文主题词的汉语拼音、英文单词或缩写词,如“风景(fengjing, scene)”。根据图像主题词对照词典,就可以把外部信息中的汉语拼音、英文单词或缩写词翻译成相应的主题。

4 WWW 图像语义信息提取

Heng Tao Shen 等人^[8,9]认为 HTML 文档中的图像名、图像注释、图像周围文本、图像所在网页标题等四项文本信息普遍和图像语义密切相关。我们还发现图像 URL、图像所在网页 URL、图像超链接 URL、图像超链接网页标题等四项文本信息亦和图像语义密切相关。因此可将上述八项文本信息作为表征 WWW 图像语义的外部信息来源来提取 WWW 图像的关键语义信息。将图像的文件名、图像的周围文本、图像的注释、图像所在网页标题、图像超链接网页标题以及 URL

(图像 URL、图像所在网页 URL、图像超链接 URL)的自动分词和词性标注的结果分别记为 T1~T6。WWW 图像的语义信息就来自对 T1~T6 的分析。

4.1 图像主题词的提取

描述图像主题的关键词至少要有有一个,不过不宜太多,以免不相关的关键词即噪声的干扰。从 T1~T6 中提取图像主题词时,采用了如下方法:

1)将 T1、T5、T2(不超过 20 个字时)、T3 按由高到低的优先级排列,然后逐一将各字符串中的词汇与图像主题词分类词典中的分类类名相匹配。如果匹配成功,就将符合条件的那个字符串中包含的所有分类类名提取出来作为图像的主题,不再考虑其它外部信息。

2)否则,如果从 T1 至 T6,有两个或两个以上外部信息中含有相同的图像主题词分类词典中的分类类名,则提取该分类类名作为图像主题;

3)否则,如果网页标题(T4)中含有图像主题词分类词典中的分类类名,则提取该分类类名作为图像主题;

4)否则,提取图像文件名(T1)作为图像主题。

4.2 图像主体词的提取

图像可以有多个主体,在 T1~T6 中提取下列内容作为图像主体名:

1)T1 中包含的图像主体词分类词典中的分类类名作为图像主体;

2)T1~T6 中两个或两个以上同时包含的图像主体词分

类词典中的分类类名作为图像主体;

3)T2 中出现频率最高的前 5 个图像主体词分类词典中的分类类名作为图像主体。

4.3 图像主体属性的提取

在提取了图像主体名以后,再对 T1~T6 中的相关语句进行语义分析,即可得到图像主体的相关属性词。限于自然语言理解的困难性,仅对带“是”和“的”的语法结构进行简单的语义分析来提取图像主体的属性词;对其它复杂的句子结构则采用统计的方法来分析图像主体的属性。步骤如下:

1)在主体名中有属性词或主体名前有“的”字,而“的”字前是属性词,则将属性词提取出来;

2)在包含图像主体名的句子中,若有属性名,而且属性名后有“是”字,“是”字后是属性词,则将属性词提取出来;

3)提取外部信息中出现频率最高的属性词。

5 WWW 图像语义提取系统介绍

根据上述研究,我们开发了一个 WWW 图像语义提取系统,结构如图 3 所示。该系统由元搜索和预处理、语义提取及主题词在线学习三个模块构成。图像元搜索和预处理模块负责从 Google 与 Baidu 上搜索图像与相关网页,并提取与图像相关的八项外部信息。图像语义信息提取模块负责从外部信息中提取图像主题词、主体词及其属性词。主题词在线学习模块负责主题词分类词典的在线学习与自动添加。

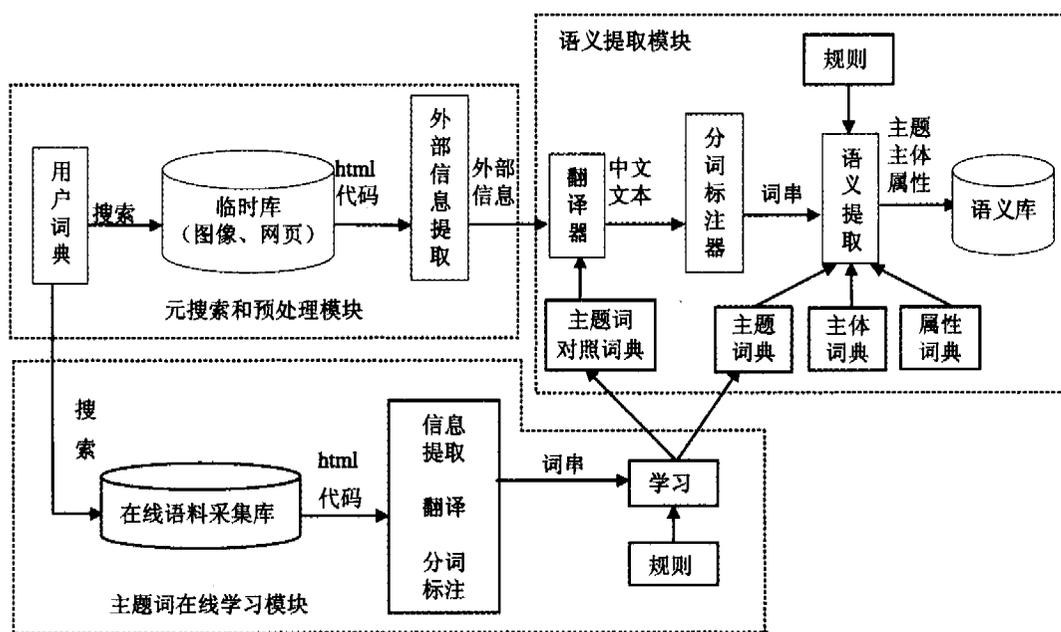


图 3 WWW 图像语义提取系统结构图

6 实验结果

利用 Google 的图像搜索功能,以“风光”、“森林”、“草原”、“大海”、“黄河”、“长江”等 25 个词为关键词,搜索了 1339 幅相关的图像和相应的中文简体网页。采用本文所述方法,提取每幅图像的主题、主体名和主体属性词。然后,根据图像所在网页的说明文字,分析图像的语义属性,计算图像语义信息自动提取的覆盖率(自动提取的正确数与人为提取的总数之比)和正确率(自动提取的正确数与自动提取的总数之比),实验结果如表 1 所示。

表 1 实验结果

	图像主题	图像主体	主体颜色	主体形状	主体纹理
覆盖率	83%	77%	60%	58%	56%
正确率	72%	51%	71%	67%	45%

实验发现,基于简单的关键词搜索,其搜索结果与用户期望结果相差甚远。利用本文介绍的图像语义提取方法标注图像,可以有效地缩小目标图像范围,提高图像搜索的查准率。

结束语 实验结果表明,本文中的 WWW 图像语义提取方法,能够实现 WWW 图像语义的自动标注,对基于内容和

语义的图像检索具有良好的可行性。我们今后的工作是提高图像外部信息的提取精度、扩大图像语义词典的规模、改进语义提取算法并结合图像的底层视觉特征进一步确认图像的语义属性。

参考文献

- 王惠锋,孙正兴,王箭. 语义图像检索研究进展. 计算机研究与发展, 2002(5)
- 孟祥增,张华,王翔英,钟义信. WWW 中图像的语义信息提取. 中国人工智能进展; 2003. 北京: 邮电出版社, 2003
- 孟祥增,钟义信,白成杰. 基于属性表征和信息测度的图像检索. 山东师范大学学报(自然科学版), 2004, 3
- Al-Khatib W. Semantic Modeling and Knowledge Representation

(上接第 207 页)

表 2 $\tau_{ij}=\sigma_{ij}=5, w=0.001$ 系统(8)收敛迭代速度统计表(次数)

实验次数	1	2	3	4	5	6	7	8	9	10	平均
$u_1(t)$	859	1164	550	1162	1263	1372	1267	1266	1060	1166	1112.9
$u_2(t)$	1054	1266	652	1265	1367	1568	1369	1369	1163	1268	1234.1
$v_1(t)$	697	907	294	907	1009	1210	1010	1010	804	1004	885.2
$v_2(t)$	1011	1315	702	1314	1414	1523	1418	1418	1211	1318	1264.4

表 3 $\tau_{ij}=\sigma_{ij}=10, w=0.001$ 系统(8)收敛迭代速度统计表(次数)

实验次数	1	2	3	4	5	6	7	8	9	10	平均
$u_1(t)$	1910	1507	901	1872	2115	1912	1873	1506	699	1710	1600.5
$u_2(t)$	2112	1708	1298	2109	2510	2305	2109	1708	1097	2105	1906.1
$v_1(t)$	1404	1000	591	1402	1804	1601	1402	999	389	1399	1199.1
$v_2(t)$	2211	1808	1202	2206	2415	2212	2206	1807	1000	2011	1907.8

表 4 $\tau_{ij}=\sigma_{ij}=10, w=0.001$ 系统(8)收敛迭代速度统计表(次数)

实验次数	1	2	3	4	5	6	7	8	9	10	平均
$u_1(t)$	1913	2043	2316	2648	1912	2107	2313	2310	2512	2114	2218.8
$u_2(t)$	2309	3210	2710	2915	2242	2311	2515	2513	2715	2508	2504.8
$v_1(t)$	1603	1603	2005	2209	1599	1604	1808	1807	2009	1803	1805.0
$v_2(t)$	2341	2407	2617	2979	2212	2409	2614	2611	2813	2415	2541.8

参考文献

- Kosko B. Bidirectional associative memories. IEEE Trans. Syst. Man and Cybernetics, 1987, 18, 49~60
- Kosko B. Adaptive bidirectional associative memories. Appl. Optics, 1987, 26, 4947~4960
- Gopalsamy K, He X Z. Delay-independent stability in bidirectional associative memory networks. IEEE Trans. on Neural Networks, 1994, 5, 998~1002
- Harirao V S, Phaneendra Bh R M, Prameela V. Global dynamics of bidirectional associative memory networks with transmission delays. Differential Equations and Dynamical Syst., 1996, 4, 453~471

- in Multimedia Database[J]. IEEE Trans. On Knowledge and Data Engineering, 1999, 11(1), 64~80
- 刘怡,于沛. 基于“知网”的新闻图像检索方法. 河北师范大学学报(自然科学版), 2003(5)
- 曹莉华,胡晓峰. 基于 WWW 的多媒体信息检索. 微型电脑应用, 1998(4)
- 孔桃,赖茂生. 基于语义联想的中文图像搜索引擎——构想与实验. 现代图书情报技术, 2002(3)
- Shen H T, Ooi B C, Tan K L. Giving Meanings to WWW Images. In ACM MM, 2000, 39~47
- Ooi B C, Shen H T, Tan K L. ICICLE: A Semantic-based Retrieval System for WWW Images
- Harirao V S, Phaneendra Bh R M. Global dynamics of bidirectional associative memory networks involving transmission delays and dead zones. Neural Networks, 1999, 12, 455~465
- 刘妹琴. 连续 BAM 神经网络的稳定性分析—LMI/BMI 方法. 电路与系统学报, 2005, 10(3), 52~57
- 周进. 具有时滞的双向联想记忆(BAM)的神经网络的全局动力学行为. 应用数学和力学, 2005, 26(3), 300~308
- Liao X, Yu J. Qualitative analysis of bidirectional associative memory with time delays. Int. J. Circuit Theory Applicat., 1998, 26, 219~229
- Mohamad S. Global exponential stability in continuous-time and discrete time delayed bidirectional neural network. Physica D, 2001, 159, 233~251
- 张强. 时滞双向联想记忆神经网络的全局稳定性. 物理学报, 2003, 52(7), 1600~1605
- 周小平. 含时延的双向联想记忆神经网络的指数吸引性分析. 四川师范大学学报(自然科学版), 2005, 28(4), 386~390
- Forti M, Tesi A. New conditions for global stability of neural networks with application to linear and quadratic programming problems. IEEE Trans. Circuits Syst., 1995, 42, 354~365
- Sudharsanan S I, Sundareshan M K. Exponential stability and a systematic synthesis of a neural networks for quadratic minimization. Neural Networks, 1991, 599~613
- Kosko B. Neural networks and fuzzy systems—a dynamical systems approach to machine intelligence. New Delhi; Prentice-Hall of India, 1994
- Mitrinovic D, Vasic P. Analytic Inequalities. New York; Springer-Verlag, 1970
- Zhao H. Existence of periodic oscillatory solution of reaction-diffusion neural network with delays. Physics Letters A, 2005, 343, 372~383
- Zhao H. Global exponential stability and periodicity of cellular neural networks with variable delays. Physics Letters A, 2005, 336, 331~331
- 陈光俭. 变时滞 Hopfield 神经网络模型的全局指数稳定性和全局吸引性. 工程数学学报, 2005, 22(5), 821~826
- 余昭旭. 一类时滞神经网络系统的指数稳定性. 控制理论与应用, 2005, 22(2), 321~324
- Liu X, Dickson R. Stability analysis of Hopfield neural networks with uncertainty. Mathematical and Computer Modeling, 2001, 34, 353~363