高技术通讯 2013 年 第23 卷 第1 期:49 ~ 54

doi:10.3772/j.issn.1002-0470.2013.01.008

基于网络信息挖掘的视频博客自动语义标注①

张晓宇②

(中国科学技术信息研究所 北京 100038)

摘 要 为获得高质量的视频博客语义标注,针对视频博客的特点,提出了一种基于网络信息挖掘的自动语义标注算法,该算法首先从分析视频博客自身所包含的信息入手,从中提取基本标注;然后借助丰富而便捷的网络资源,通过深入挖掘网络信息获取在底层特征和高层语义上都相关的信息,对基本标注进行改进和完善,最终实现基于上下文的标注扩展。为了更加全面客观地评价语义标注结果,提出了一种基于分值的评价标准,该标准有效兼顾了标注的正确性和完整性这两大重要指标,从而能更加准确地反映标注质量。实验结果表明,这种基于网络信息挖掘的自动语义标注算法能够显著提高语义标注质量,对于海量视频博客的有效获取与管理具有重要意义。

关键词 信息挖掘,语义标注,标注扩展,评价标准,视频博客

0 引 言

语义标注是多媒体信息管理的一个重要手段,基于语义标注可以实现信息的高效检索。近年来,随着互联网的发展与推广,博客(weblog,简称 blog)成为一种新的信息发布和交流媒介[1]。最早出现的博客大都采用文字形式,随着网络数字媒体技术的飞速发展,出现了一系列新型博客,其中视频博客(video blog,简称 vlog)[2]因采用了以视频为主、以文学为辅的生动形象的表现形式,在广大博客用户中大受欢迎。随着视频博客的不断涌现,如何对其进行自动语义标注,实现有效管理,则成为一个极具挑战性的课题。本文关注了这方面的研究,在借鉴已有自动语义标注方法的基础上,根据视频博客的特点,提出了一种基于网络信息挖掘的视频博客自动语义标注算法,该算法的有效性已通过一系列实验得到了验证。

1 相关知识

从语义标注的角度看,博客视频相比单纯的视频而言有其自身的优势:以视频为中心的,其中的文字信息往往都为描述视频内容服务。因此,如果利用视频博客自身的文本对视频内容进行标注,其结果往往比从一般视频的环绕文字中获取的信息更加可靠。但是视频博客的标注也有其自身的难点:由于视频博客是由不同用户制作和发布的,因而其用语往往不可避免地存在着不统一、不规范的问题。所以,直接从视频博客的文字内容中提取关键词作为其语义标注的效果往往不佳,使用这样的语义标注最终也会影响到视频博客的检索性能。

对视频博客的语义标注是一个多标签的标注过程[3,4],因为一个视频博客可以同时用多个关键词进行描述。在多标签的图像、视频语义标注方面已经有一些相关的工作,目前的一个趋势是:不仅仅从待标注对象本身提取语义标注,而且还从相关的其它对象中获取有用信息。文献[5-7]提出了一系列图像标注扩展的方法,其主要思想是:对于一个待标注图像,在已标注样本集上通过基于文本和内容的检索方式获取在语义和视觉上相似的其它图像,然后利用这些相关图像标注去描述待标注图像。这种方法充分利用现有资源对标注信息进行扩充和完善,因而可以获得比仅仅基于自身信息更好的标注

① 中央级公益性科研院所基本科研业务费专项资金(XK2012-2, ZD2012-7-2)和中国科学技术信息研究所科研项目预研资金(YY201208)资助项目。

② 男,1983 年生,博士;研究方向:模式识别与智能系统,科技信息分析与挖掘;联系人,E-mail: zhangxy@istic.ac.cn

(收稿日期:2012-03-10)

效果,然而,无关信息的引入影响标注的准确性是其需要着力解决的问题。

2 自动语义标注

本文针对视频博客的特点提出的基于网络信息挖掘的视频博客自动语义标注算法,一方面从视频博客自身所包含的文本信息中获取基本标注,另一方面深入挖掘网络信息,充分利用相关网络资源通过标注扩展的方法对现有标注进行改进与完善。

2.1 词的语义相关性度量

在语义标注中,词的语义相关性度量是一个非常关键的问题。现有方法主要分为两大类,即基于词典的方法和基于统计的方法。

基于词典的方法主要借助结构化的词典来计算词在语义上的相互关系。例如,WordNet[8]是在自然语言处理中广泛应用的一种结构化词典,基于词典所构造的语义关系可以衡量词的语义相关性。但是,正如文献[9]所言,基于词典的方法容易受到词的多义性影响,导致语义相关性度量产生偏差。此外,词的数量庞大并且不断增长,而词典的规模则毕竟有限,因而基于词典的方法无法对任意词的语义相关性进行度量。

基于统计的方法是数据驱动(data-driven)的,其主要思想是从词的共生关系中挖掘词汇相关性。实验表明,具有较高共生频率的词之间往往具有较高的语义相关性,因此词的共生关系可以有效地反映其语义相关性[10]。目前,归一化 Google 距离(normalized Google distance,NGD)[11]是利用共生性度量词的语义相关性的一种常用方法,它借助 Google 搜索引擎获取词在网络文档中同时出现的相关信息,并以此计算其语义相关性。基于的统计的方法不依赖固定词库,相比基于词典的方法具有更大的可扩展性和灵活性,因此本文采用归一化 Google 距离作为词的语义相关性度量。

2.2 基本标注生成

利用视频博客自身文字信息,可以直接从中提取关键词作为基本标注。

一个视频博客中的文字信息主要包括标题、正文以及浏览评价,其中标题和正文是与视频博客的语义直接相关的,因此可以从中获取语义标注的有用信息。

由于标题点明了整个视频博客的主题,因而对于理解其语义是至关重要的。算法首先从标题中提取关键词集合 W_{title}。

视频博客的标题的长短有一定限制,一般非常简短,无法涵盖其所有重要内容,此外很多视频博客作者比较倾向于使用引人注目的标题,而并不关注于其对语义内容的描述是否准确,因此,有必要从正文中获取视频博客的语义标注信息。基于标准的文字处理手段,如本文中所采用的词频-反文档频率(TF-IDF),本文算法进一步获取正文的关键词集合 $W_{\text{description}}$。从标题和正文分别得到关键词集合之后,算法将两者融合,得到基本标注集 $W_{\text{intrinsic}}$:

$$W_{\text{intrinsic}} = W_{\text{title}} \cup W_{\text{description}} \tag{1}$$

2.3 自动语义分类

常用视频博客网站,均按照不同的语义类别进行内容组织与管理。有了语义类别信息,视频博客的语义内容可以被更加准确地界定与刻画。例如,一个有关“Houston Rockets”的 NBA 篮球比赛精彩镜头视频博客,其标注内容中将包含关键词“rockets(火箭)”,但是,仅仅从“rockets”一词出发,用户将会误认为这是一个有关军事题材的视频博客。一旦将其语义类别限定为“sports(体育)”,则此处“rockets”将明确无误地表示一个篮球队的名字。可见,语义分类是一个语义消歧(disambiguation)的有效手段。

本文主要考虑 7 个语义类别的视频博客,其基本囊括了目前视频博客的大部分热门主题,分别是:Autos & Vehicles,Film & Animation,Music,News & Politics,Pets & Animals,Sports 和 Travel & Events。这些类别可以方便地进行扩展。

基本标注集 $W_{\text{intrinsic}}$ 中的所有关键词构成了当前视频博客的一个全局上下文语境(global context),由此可以确定视频博客所属的语义类别。在上例中,尽管“rockets”本身具有一定的歧义性,但是综合考虑其它关键词:“NBA”、“points”、“rebounds”等,可以断定该视频博客应该属于“sports”类别。

给定一个视频博客的基本标注集 $W_{\text{intrinsic}}$,本文采用投票的方式决定其语义类别的归属。

首先,确定标注集中每一个关键词的类别信息。由于类别的名称本身就揭示了其语义,因而可以通过计算关键词与类别名称的相关性来度量该关键词属于该类别的程度。与一个关键词在语义上最相关的类别便作为该关键词所属的类别:

$$Cat_{\text{word}}(w) = \arg \max_{c \in C} Sim_{\text{word}}(w, c) \tag{2}$$

其中,w 是关键词,c 是语义类别,C 是所有语义类别的集合。

在获取每一个关键词的类别信息之后,便可确定整个视频博客的类别。由于一个视频博客的类别取决于其所在的全局上下文,因此将具有最多关键词的类别作为该视频博客的类别:

$$Cat_{\text{vlog}}(v) = \underset{w \in W_{\text{intrinsic}}(v)}{\text{mod}} Cat_{\text{word}}(w) \quad (3)$$

其中,函数 mod 返回具有最多出现频率的语义类别。

2.4 基于上下文的标注扩展

基本标注往往并不足以全面地描述视频博客的语义,为了进一步提高标注的质量,本文深入挖掘网络信息,进行基于上下文信息的标注扩展。

2.4.1 标注扩展模式

"基于搜索的标注"(search-based annotation)[5-7]是进行标注扩展中最为常用的一种手段,其主要思想是:搜索与待标注对象在底层特征和高层语义上相似的已标注对象,利用返回结果的标注信息对待标注对象进行标注。对基于搜索的标注方法来说,一个已标注的样本集和一个搜索引擎是必不可少的。众所周知,YouTube 是目前最大的视频共享网站之一,其中的每一个视频都是已标注的,因而我们使用 YouTube 作为已标注视频集。在给定一个查询之后,YouTube 中基于 Google 的搜索引擎可以返回相当好的搜索结果,因此我们可以使用 YouTube 去获取语义上相关的视频。利用现有的 YouTube 来进行标注扩展与其它的基于搜索的标注方法相比,具有突出优点:首先,基于搜索的标注方法需要事先构建一个大规模、高质量的已标注样本集,该过程相当费时费力,而利用 YouTube 中具有可靠标注信息的视频集,则省去了上述繁琐的样本集构建与标注工作;其次,基于搜索的标注方法需要一个可靠的搜索引擎,而 YouTube 中所采用的基于 Google 的搜索引擎是在实际应用中公认的高质量搜索引擎,因而保证了标注扩展的准确性。

下面对标注扩展的具体模式进行深入讨论。最直接的基于搜索的标注扩展模式有以下 3 种:

模式 1: $w_o \rightarrow w_e$,其中 w_o 和 w_e 分别为初始关键词和扩展后的关键词。

模式 1 等价于向搜索引擎提交单一关键词作为查询以获取扩展信息,这种扩展模式常常会受到歧义性的影响。例如,"rockets→missile"与"rockets→basketball"都是对关键词"rockets"的合理扩展。但是,对于有关篮球比赛的视频博客,前一个扩展显然并不适用。为了解决模式 1 存在的问题,需要加入更多的上下文信息,以更严格地对扩展进行约束[12],于是提出如下扩展模式:

模式 2: $\{w_o, w_c\} \rightarrow w_e$,其中 w_c 是 w_o 的局部上下文语境(local context)关键词。

模式 2 等价于向搜索引擎同时提交两个关键词作为查询以获取更加准确的扩展信息。由于 w_c 的使用为扩展提供了额外的信息,因而扩展被很自然地限定在了更准确的范围内。在上例中,如果使用"NBA"作为局部上下文语境关键词,则很显然地只有"{rockets, NBA}→basketball"是合适的扩展。基于前文已经获取的视频博客类别信息,可以对其充分利用来进一步改进现有扩展模式。本文提出如下扩展模式:

模式 3: $\{w_o, w_c\} \mid c \rightarrow w_e$,其中 c 是视频博客的类别信息。

与模式 1 和模式 2 相比,模式 3 类似于利用类别信息进行高级搜索的过程,因而进一步对语义扩展进行了限定。结合上例,在模式 2 中,"{rockets, shot}→missile"依然可以被认为是一个正确的扩展,因为"shot"一词具有多义性(发射/投篮),但是,如果采用模式 3,以类别"sports"作为扩展条件,则"{rockets, shot} | sports→missile"显然是一个错误的扩展。

对于每一个关键词 w_o,其局部上下文语境关键词 w_c 的选择取决于在基本标注集 $W_{\text{intrinsic}}$ 中关键词之间的关系。给定一个 w_o,其它关键词 $w_t(W_{\text{intrinsic}} - \{w_o\})$ 均为其提供了语义上的局部上下文信息,但是选取不同的关键词作为上下文限定在语义扩展中的效果不尽相同。实验表明,与 w_o 的相关性越小的 w_t 对于上下文语义的限定就越精确。因此,本文选取与 w_o 相关性最小的关键词作为标注扩展中的上下文限定词:

$$w_c = \underset{w_t \in W_{\text{intrinsic}} - \{w_o\}}{\arg\min} Sim_{\text{word}}(w_o, w_t) \quad (4)$$

2.4.2 标注扩展算法

基于上下文的标注扩展算法可作如下描述:

对一个待标注视频博客,将关键词 w_o 及其上下文关键词 w_c 提交给 YouTube 搜索引擎,并以相应的语义类别作为扩展条件,从而获得一系列在语义上相关的搜索结果(为简单起见,只采用返回的前 20 个结果);对于每一个返回结果 r,提取视频的关键帧 $Frame(r)$ 以及相应的标注 $Tag(r)$。对于从 YouTube 返回的语义上相关的视频,进一步将其与待标注的视频博客进行比较,以筛选出底层特征上相似的视频。我们将视频博客 v 与返回的视频 r 在底层特征上的相似性定义如下:

$$Sim_{\text{video}}(r,v) = \frac{1}{|Frame(v)|}\sum_{f_v \in Frame(v)} \max_{f_r \in Frame(r)} Sim_{\text{image}}(f_r, f_v) \quad (5)$$

其中，Sim_{image} 表示关键帧之间的图像相似度。

至此，我们从关键词 w_o 出发，获取了在高层语义和底层特征上均与待标注视频博客相似的视频，我们将这些视频所对应的标注信息作为候选的扩展标注放入集合 $ExTag(w_o)$ 中。

对每一个 $w_o \in W_{\text{intrinsic}}$ 重复上述过程，最终可以得到如下形式的候选扩展标注集：

$$W_{\text{external}} = \bigcup_{w_o \in W_{\text{intrinsic}}} ExTag(w_o) \quad (6)$$

最后，将基本标注集 $W_{\text{intrinsic}}$ 与扩展标注集 W_{external} 进行融合，获得扩展后的语义标注：

$$W = W_{\text{intrinsic}} \cup W_{\text{external}} \quad (7)$$

3 评价标准

对于语义标注的评价主要分为两类：

基于检索的评价(search-based criteria)，即通过检索结果的好坏评价语义标注的质量。

基于分值的评价(score-based criteria)，即直接由用户对语义标注给出评价分值。在文献[5]中，定义了三类评价等级："perfect(优)"、"correct(可)"和"wrong(差)"，分别对应于权值 1、0.5 和 -1。最终，一个标注集的评价分值由下式计算：

$$E = (p + 0.5r - w)/N \quad (8)$$

其中，N 表示标注集中关键词的个数，p、r、w 分别表示"优等"、"可等"和"差等"这三类关键词的个数。这种评分标准是文献[13]所提出的规范化评分(normalized score)方法的扩展，在该方法中仅仅定义了"right(正确，权值为 1)"和"wrong(错误，权值为 -1)"两种关键词评价。

实践表明，一个具有高质量的标注必须同时满足两方面的要求，即准确性(accuracy)和完整性(completeness)。准确性要求标注信息能正确无误地描述对象的语义内容，而完整性则要求标注信息能全面地涵盖对象中所涉及的语义内容。例如，对于电影《Titanic》中"Titanic 号"沉没的视频片段，有两个标注：$W_1 = \{\text{Titanic}\}$，$W_2 = \{\text{Titanic, crash, sink, iceberg, USA}\}$。显然，$W_1$ 非常准确但却过于简略(不完整)，而 W_2 则明显优于 W_1。

式(8)所示的评分标准偏向于准确性，而忽略了完整性方面的要求。在上例中，依据式(8)，$W_1(p=1, r=0, w=0, N=1)$ 与 $W_2(p=4, r=1, w=0, N=5)$ 的评分分别为 1 和 0.9，$E(W_1) > E(W_2)$，与上述分析相矛盾。

本文提出了一种新的评分标准，同时兼顾标注的准确性和完整性。主要思想为：一个标注中若存在(present)一个"优等"关键词，则奖励分值 1；若缺失(absent)了一个本应存在的"优等"关键词，则惩罚分值 -1。类似地，若存在一个"差等"关键词，则惩罚分值 -1；若不存在一个在别的标注中存在的"差等"关键词，则奖励分值 1。对于"可等"关键词，似乎可有可无，但从简洁性考虑宁愿其不存在，因此对存在的"可等"关键词赋以 0.5 分，而对不存在的"可等"关键词赋以 1 分。新的评分标准归纳在表 1 中。

表 1 标注关键词评分标准

	优等	可等	差等
存在	1	0.5	-1
不存在	-1	1	1

基于新的评分标准，可以计算相应的标注评价分值。假设有 K 个不同的标注：W_1，W_2，…，W_K。为了对其进行比较，首先将所有 K 个标注融合成为一个集合 W_{all}：

$$W_{\text{all}} = W_1 \cup W_2 \cup \cdots \cup W_K \quad (9)$$

因此，W_{all} 中含有所有标注中的关键词。然后，我们对 W_{all} 中的每一个关键词进行判定，评出优、可、差三等。对于一个标注 W_i，我们根据下式计算其评价分值：

$$E_{a\&c} = \frac{p_{\text{present}} - p_{\text{absent}} + 0.5r_{\text{present}} + r_{\text{absent}} - w_{\text{present}} + w_{\text{absent}}}{|W_{\text{all}}|} \quad (10)$$

其中，p_{present}、r_{present}、w_{present} 分别表示 W_i 中存在的优等、可等、差等关键词数，p_{absent}、r_{absent}、w_{absent} 分别表示 W_{all} 中存在而 W_i 中不存在的优等、可等、差等关键词数。

利用式(10)，我们对上述 Titanic 视频片段的标注进行评分，$W_1(p_{\text{present}}=1, p_{\text{absent}}=3, r_{\text{absent}}=1, r_{\text{present}}=w_{\text{present}}=w_{\text{absent}}=0, |W_{\text{all}}|=5)$ 得分：-0.2，$W_2(p_{\text{present}}=4, r_{\text{present}}=1, p_{\text{absent}}=r_{\text{absent}}=w_{\text{present}}=w_{\text{absent}}=0, |W_{\text{all}}|=5)$ 得分：0.9。显然，这种评分方式更加符合用户体验。

与以前的评分方法相比，新的评价标准更好地

平衡了准确性和完整性,因而是一种更加合理的方法。该评分方法不仅运用于视频博客的标注评价,而且可以作为任何一种多标签标注问题的评价标准。

4 实验结果

实验视频博客库由前文介绍的7个语义类别构成,每类包含200个视频博客,共1400个。每个视频博客,包含视频文件及相应的文字内容(标题和正文),系统将自动对其进行语义标注。

4.1 基于分值的评价结果

自动语义标注完成之后,标注中每一个关键词通过人工判定的方式将其划归到优、可、差3个等级中。之后,根据表1所示的评价准则,利用式(10)计算出相应的标注评价分值。实验对语义标注每一个阶段所获得的标注信息进行评分,并比较了2.4.1中所介绍的3种标注扩展模式,比较结果列在表2中。

表2 标注评价分值

	$W_{\text{intrinsic}}$	W_{external}	
$E_{a\&b}$	0.47	0.63	$w_o \rightarrow w_e$
		0.70	$\{w_o, w_c\} \rightarrow w_e$
		0.74	$\{w_o, w_c\} \mid c \rightarrow w_e$

由表2可见,基本标注集质量不高,相应的评价分值也较低。在进行标注扩展之后,标注质量随着所采用扩展模式的不同而各异:利用模式1,标注质量提升有限,这是因为尽管标注扩展提高了标注的完整性,但是也不可避免地混入了一些可等甚至是差等的关键词,从而在评分中受到惩罚;利用模式2和模式3,改善效果则明显得多,这主要归功于一些约束条件的引入有效地遏制了扩展过程中歧义性的出现。在所有参与比较的方法中,本文提出的标注扩展模式3是最优的。

4.2 基于检索的评价结果

实验采用基于文本的检索,利用查询与视频博客的标注信息的匹配返回检索结果。实验对随机选择的50个查询关键词分别计算P@m值[14]并取平均,不同标注情况下的结果对比显示在图1和图2中。

图1比较了$W_{\text{intrinsic}}$和依据扩展模式3所得的W_{external}。由图可见,基于基本标注的检索结果并不十分令人满意;而在引入扩展信息之后,P@m值普遍得到显著提高。图2比较了三种扩展模式下的W_{external}。由图可见,采用扩展模式2和模式3的标注所获得的检索结果优于模式1,而本文提出的模式3效果最优。图1和图2的结果与表2一致,这从一个侧面也证明了本文提出的标注评分方法的正确性。

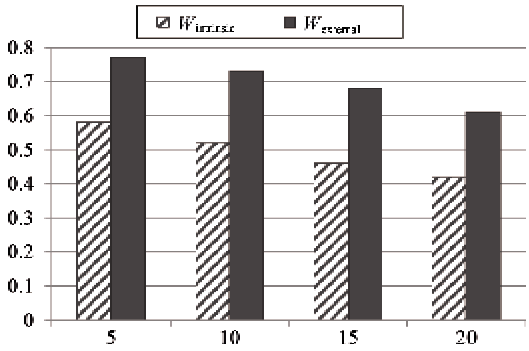

图1 平均P@m值(m=5, 10, 15, 20)

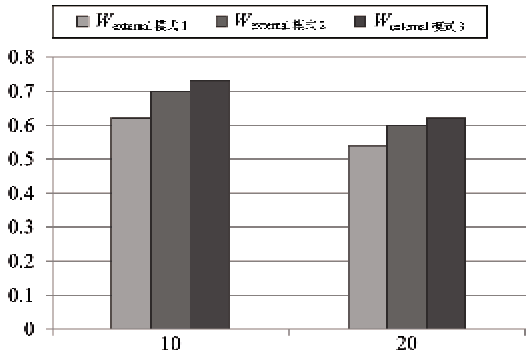

图2 平均P@m值(m=10, 20)

5 结 论

本文提出了一种基于网络信息挖掘的视频博客自动语义标注算法,实现了视频博客的语义内容分析。该算法从视频博客自身所包含的信息中提取基本标注,并通过深入挖掘网络信息实现基于上下文的标注扩展,从而获得高质量的语义标注。本文的贡献在于:借助丰富而便捷的网络资源,对基本标注进行改进与完善,提高标注质量;充分利用全局和局部上下文语境,获取重要类别信息,并为标注扩展提供有效约束;针对多标签标注问题,定义了基于分值的评价标准,更好地兼顾了准确性和完整性。实验证明,本文提出的基于分值的评价标准与实际检索结果高度一致,并且两者共同验证了基于网络信息挖掘的自动语义标注算法的有效性,为海量视频博

客的有效获取与管理提供了保证。

参考文献

[1] Rosenbloom A. The blogosphere: introduction. *Communications of the ACM*, 2004, 47(12): 31-33

[2] Parker C, Pfeiffer S. Video blogging: content to the max. *IEEE Multimedia*, 2005, 12(2): 4-8

[3] Qi G, Hua X, Rui Y, et al. Correlative multi-label video annotation. In: Proceedings of the 15th International Conference on Multimedia, Augsburg, Germany, 2007. 17-26

[4] 张晓宇. 基于多视角二维主动学习的多标签分类. 高技术通讯, 2011, 21(12):1312-1317

[5] Wang X, Zhang L, Jing F, et al. AnnoSearch: image auto-annotation by search. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, USA, 2006. 1483-1490

[6] Wang C, Jing F, Zhang L, et al. Scalable search-based image annotation of personal images. In: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval, Santa Barbara, USA, 2006. 269-278

[7] Rui X, Li M, Li Z, et al. Bipartite graph reinforcement model for web image annotation. In: Proceedings of the 15th International Conference on Multimedia, Augsburg, Germany, 2007. 585-594

[8] Miller G. WordNet: a lexical database for English. *Communications of the ACM*, 1995, 38(11): 39-41

[9] Natsev A, Haubold A, Tesic J, et al. Semantic concept-based query expansion and re-ranking for multimedia retrieval. In: Proceedings of the 15th International Conference on Multimedia, Augsburg, Germany, 2007. 991-1000

[10] Schutze H, Pedersen J. A cooccurrence-based thesaurus and two applications to information retrieval. *Information Processing & Management*,1997, 33(3): 307-318

[11] Cilibrasi R, Vitanyi P. Automatic extraction of meaning from the web. In: Proceedings of IEEE International Symposium on Information Theory, Seoul, Korea, 2006. 2309-2313

[12] Bai J, Nie J, Cao G, et al. Using query contexts in information retrieval. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, Netherlands, 2007. 15-22

[13] Barnard K, Duygulu P, Forsyth D, et al. Matching words and pictures. *Journal of Machine Learning Research*, 2003, 3: 1107-1135

[14] Liu J, Wang B, Li M, et al. Dual cross-media relevance model for image annotation. In: Proceedings of the 15th International Conference on Multimedia, Augsburg, Germany, 2007. 605-614

Automatic semantic annotation for video blogs based on web information mining

Zhang Xiaoyu

(Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract

This paper presents a new automatic algorithm for semantic annotation of video blogs based on web information mining with the aim to acquire high-quality annotation. The algorithm first extracts intrinsic annotations from the original content of the target video blogs, and then uses external resources to obtain informative contents that are relevant to the video blogs in both high level semantics and low level features and improve intrinsic annotations based on web information mining. Finally, the context-based annotation expansion is achieved. The paper also defines a new score-based evaluation criterion for multi-label annotation problems, which takes both the accuracy and the completeness in annotation into account. The experimental results demonstrate the effectiveness of the proposed annotation algorithm and evaluation criterion, both of which are significant for acquisition and management of a large amount of video blogs.

Key words: information mining, semantic annotation, annotation expansion, evaluation criterion, video blog