

《智能信息处理》课程考试

基于本体的音乐语义信息检索研究

亓泽东

考核	到课[10]	作业[20]	考试[70]	课程成绩[100]
得分				

2020 年 12 月 09 日

基于本体的音乐语义信息检索研究

元泽东

(大连海事大学 计算机科学与技术 辽宁省大连市 中国 116026)

摘 要 针对目前互联网上使用的音乐检索系统是以文本信息为主要搜索依据,难以满足智能化检索需求的问题,提出了一种基于内容的音乐检索方法。在保留传统检索方法的基础上,将情感和场景需求引入到检索中,对基于语义的音乐检索方法及模型进行了相关研究。首先根据音乐所蕴含的意境及适合播放的场景进行分析,构建音乐领域本体;然后对用户提出的查询请求进行分词分析和词性标注,并进行语义分析;最后通过推理机对本体进行推理,检索出满足用户需求的音乐集。

关 键 词: 语义检索; 音乐; 本体; 自然语言

中图法分类号 TP311.20 DOI号 10.3969/j.issn.1001-3695.2014.01.030

Research on Music Semantic Information Retrieval Based on Ontology

Zedong Qi

(Computer science and technology, Dalian maritime university, Liaoning Dalian, 116026, China)

Abstract Aiming at the problem that the current music retrieval system used on the Internet is based on text information as the main search basis, it is difficult to meet the needs of intelligent retrieval and a content-based music retrieval method is proposed. On the basis of retaining the traditional retrieval methods, the emotion and scene requirements are introduced into retrieval and related researches on semantic-based music retrieval methods and models are carried out. First, analyze the artistic conception contained in the music and the scenes suitable for playback to construct the music domain ontology; then perform word segmentation analysis and part-of-speech tagging on the query request submitted by the user and perform semantic analysis; finally use the inference engine to reason about the ontology and retrieve it A music collection that meets the needs of users.

Key words: Semantic retrieval; music; ontology; natural language

1 引言

近几年来,随着全球信息数字化和网络一体化进程的飞速发展,人们对日常所需信息资源的索取越来越多地依赖于网络检索技术。传统对于音乐文件的检索采用基于文件常规属性(音乐名、歌手/词(曲)作者、音频格式或者歌词等标注信息)的检索方法,已经不能满足人们对检索形式的需求,基于内容的音乐检索技术也因此应运而生。在基于内容的音乐检索技术的研究中,通常将音乐的时域、频域等声学特征作为内容进行匹配检索,然而这些特征由于过于专业化,很难受到用户的接受,因此在对音乐的检索上,用户希望通过简单的自然语

言就可以查询到所需的音乐。而现有的音乐搜索引擎还没有采用语义技术实现用户的自然语言查询。系统借助本体建立音乐间的语义关系,通过对用户查询语句的分词处理将关键词映射到本体概念,最后通过推理机推理得到相应的检索结果。

2 语义网概述

语义网(Semantic Web)是一种包含语义关系的网络,是万维网的发展趋势。它可以将不同网站内使用不同语言语言、面向不同应用的数据联系起来,并帮助用户从这些联系中获取知识。它的提出为万维网绘制了美好蓝图,在全世界范围内被广泛关注。

2.1 语义网的概念

“语义网”(Semantic Web)这一概念出现在人们的生活中,是受 Web 的创始人 Tim Berners-Lee 的影响,他曾于 1998 年首次提出语义网的基本思想,并对其做了如下定义:语义网首先是一个网,在这个网中有许多描述事物之间关系文档,同时也包含了相关语义信息,使机器可以自动处理。所谓“语义”就是文本的含义,主要用来对网络中文本的内容与结构进行说明,与文本的格式无关。有了语义,网络就可以进行“判断”,因此我们说语义网是一种具有智能的网络,它能理解自然语言并且可以实现人与计算机在思想上的交互。语义网在当前万维网的基础上进行扩展,为网络中的信息加入对应的语义,使计算机可以理解并自动处理。它是如何让计算机理解这些信息的呢?这就是语义网研究的关键,也就是说,我们需要向网络中添加一些知识并使机器理解,这些知识就是我们所说的概念。

2.2 模糊形式概念

语义网和其他的网络一样,也具有一定的体系结构,目前公认的语义网体系结构是万维网的创始者 Tim Berners-Lee 提出的。

Unicode 和 URI 层:该层是整个语义 Web 体系结构的基础。Unicode 是一种对系统中的字符进行编码方法,在语义网中主要用于对资源编码,不用考虑资源描述使用的是哪种语言。URI 是负责标识资源,是 URL 的超集。它可以对语义网上的概念或资源进行标识,便于对资源快速而准确的检索。

XML + NS + XML Schema:该层主要是在语义层面对数据进行处理。XML 是一种资源描述语言,不像 HTML,XML 可以表达数据中包含的内容。它在处理文档时有很大的优势,可以独立出文档的三要素,然后分别对他们进行处理,这样处理后的数据就可以以任何格式存储。虽然 XML 能够描述各类信息,但是从方便信息搜索的角度来看,仅有 XML 还是不够的。NS 即命名空间,是由索引来确定的。命名空间提供了一种可以将 XML 文档中的元素与属性的字符描述和用 URI 确定的命名空间联系起来的方法。XML Schema 拥有比 DTD 更丰富的功能。它直接参考了 XML 的语法规则,使 XML 外达到了完美的统一。

RDF + RDF Schema:该层主要是用一种标准的格式对资源进行统一处理。RDF 并不能理解语法,

它仅仅可以提供一种用来描述元数据的机制,这种机制领域无关。RDF Schema 即 RDF 词汇描述语言,是 RDF 的类型系统,与 RDF 相比,它可以更好的描述资源。

Ontology Vocabulary:在语义 Web 中,XML 和 RDF 都无法处理“一个概念可以由多个词语来表示,每一个词语有多重含义”这一类问题,所以很自然的就要引入 Ontology。在第二章本体的定义中,我们已经知道,本体概念的形式化说明,它的最终目标是“精确地表示那些隐含或不明确的信息”^[1]。该层的主要作用是定义一中参照,通过该参照,系统间可以进行语义上的交互。

此外,Logic 层的主要作用是进行本体的逻辑推理,生成一些待验证的逻辑规则;Proof 层主要来执行 Logic 层产生的逻辑规则并进行验证;Trust 层主要用于保证用户与计算机之间传递信息的可靠性。

3 本体

3.1 本体的定义

本体(Ontology)的概念虽然很早就出现了,也陆陆续续的有学者进行研究,但是对本体的定义一直没有统一的定论。目前普遍认可的说法是 1998 年 Studer 等人提出的“本体是对可共享的概念模型的一种规范、明确、形式化的说明”。这个定义包含 4 层含义:概念模型、明确、形式化和共享。具体描述如下:“概念模型”意思是本体一个概念集,这些概念用来对领域内的术语作出描述。依据这些描述体现出的模型具有很好的适应性,不会随环境的变化而变化。“明确”是指本体中的概念以及对概念的约束都经过明确的定义,不会产歧义。

“形式化”是指本体可以本体语言描述,这些描述可以被计算机理解和处理。“共享”是指本体所描述的是某一领域中公认的知识,因此可以将该本体应用于相同领域中。概括地说,本体是一种领域内的概念集,这个概念集中包含了领域内公认的词汇,这些词汇分别处于概念模型中的不同层次,并且它们之间的关系也都有明确的说明,因此可以实现对领域内知识的推理。

3.2 建立音乐领域本体

本文所建立的音乐领域本体是参照现有的本体构建方法,结合音乐的特殊性与人类情感的多样

性,使用由美国 Stanford 大学开发的本体编辑器 Prot6964.0 来构建音乐领域本体。

本文针对人类情感的复杂性,多样性和普遍性,以及人类在不同场景对音乐的需求进行了深入浅出的分析。抽取其中具有代表性的情绪特征和场景进行分类,对其进行规范的定义和描述,形成领域本体的概念。本系统针对音乐所蕴含的情感和适合播放的场景将试验用的音乐分为表达情感和播放场景 2 类。图 1 以表达悲伤情绪为例,描述了领域本体中概念的层次结构及相关属性。

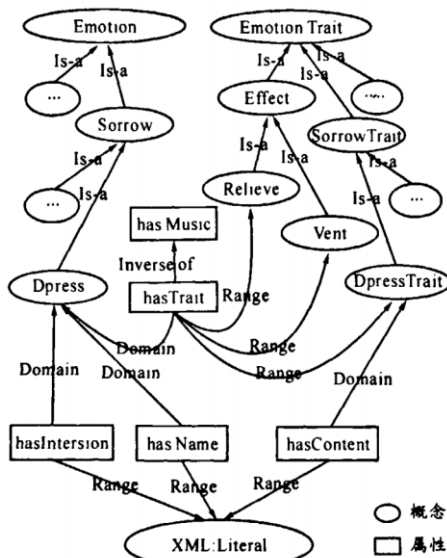


图 1 音乐领域本体构造

音乐领域本体中的概念描述如表 1 所示。本体中的概念层次描述的是概念间的继承关系,而在继承关系之外的概念间的关系则是由表 2 中的相关属性表示,同样还可以通过对类属性建立约束条件来体现各个概念以及概念间的语义关系。

表 1 音乐领域本体部分概念

类名	中文描述	类名	中文描述
Emotion	人的情绪	Vent	发泄情绪
Occasion	人所处于的场景	Sorrow	悲伤
EmotionTrait	情绪特征	Dpress	忧郁
OccasionTrait	场景特征	Lose	失落
Effect	歌曲的作用	Fear	恐惧
Relieve	缓解情绪		

表 2 音乐领域本体部分属性

属性名	Domain	Range	中文描述
HasSimilarMusic	Music	Music	类似的歌曲
hasMusic	MusicTrait	Music	音乐所描述的情绪或场景特征所关联的音乐
hasTrait	Music	MusicTrait	音乐所表达的对象特征
HasIntension	Music	XML;Literal	音乐所具有的内涵
hasName	Music	XML;Literal	音乐名称
hasContent	Music	XML;Literal	特征内容

3.3 汉语分词技术

分词的主要目的是将用户以自然语言提出的查询请求分割成多个词语,并对每个词语进行词性标注。本文使用中国科学院计算技术研究所的汉语词法分析系统 ICTCLAS 进行分词处理。它的功能包括中文分词、词性标注、未定义词识别等。例如“缓解失恋情绪的 歌”这一句子通过 ICTCLAS 系统处理后便得到一个带有词性标注的关键词序列:“缓解 / v 失恋 / v 情绪 / n 的 / u 歌 / n”。

中文语句的关键组成语句是主语、谓语和宾语,而定语,状语和补语则是对主、谓、宾的修饰,因此,在对中文语句的分词处理中,只需要考虑语句中的主、谓、宾 3 个部分。结合中文词语的词性,只需要在提取关键词的时候提取句子中的名词、动词和形容词形成关键词序列。例如对“缓 解 / v 失 恋 / v 情 绪 / n 的 / u 歌 / n”进行关键词提取,就得到了一个相应的关键词序列:“缓解 / v 失恋 / v 情绪 / n 歌 / n”。

3.4 同义词处理技术

人类的语言非常丰富,同一事物往往会有不同称呼,比如“高兴”和“开心”是同一种概念。为了防止漏检,必须用到同义词处理技术。本系统所采用的是哈尔滨工业大学信息检索研究室的<同义词词林: >扩展板。<同义词词林>按照树状的层次结构把所有收录的词条组织到一起,把词汇分成大,中,小 3 类。大类有 12 个,中类有 97 个,小类有 1 428 个。图 2 展示了<同义词词林>词义结构的部分示例,每一个节点为一个语义类,所有语义类都表示一个概念,不对应汉语中的词,只有和叶节点对应的语义类是一个具有相似意义的词群。

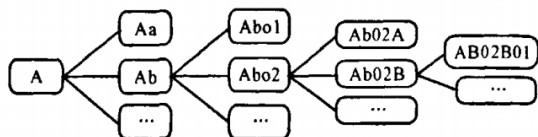


图2 《同义词词林》结构示例

3.4 句子相似度计算

句子相似度计算是满足用户使用自然语言查询,并返回准确检索结果的重要理论基础和关键实施技术。现有的句子相似度计算方法主要有向量空间模型法,语义词典法,词形词序综合法和依存树法。本文采用的句子相似度计算方法是语义词典法。语义词典法主要是利用同义词词林,知网等现有的,较成熟的语义资源,通过计算句子之间的词语相似度,进而计算句子相似度。要研究概念、概念的属性相关的一系列关系。

首先对句子进行分词和词性标注,然后将词性相同的词放入一组,只考虑具有实际意义的实词,则相应的词组为名词词组(N),动词词组(V),形容词词组(A)。设句子S经过分词后包括k个词: $S=\{w_1, w_2, \dots, w_k\}$,这k个词可以根据实词的种类分为3个集合: $S=\{N, V, A\}$,根据分词后的词语属性将实词取出并归入相应的集合中:

$$S = \{N(w_1, w_2, \dots, w_i), V(w_1, w_2, \dots, w_j), A(w_1, w_2, \dots, w_k)\}$$

然后分别计算2个句子对应的词性相同的集合的相似度。设待计算的2个句子S1和S2的实词集合分别为 $\{N1, V1, A1\}$ 和 $\{N2, V2, A2\}$,以名词集合N为例,设名词集合为:

$$N_1 = \{w_1, w_2, \dots, w_m\}$$

$$N_2 = \{w_1, w_2, \dots, w_n\}$$

对2句的名词进行两两相互词语相似度计算,取相似度最大的一对词,从集合里消去,然后再在剩下的词语里找最大相似度的一对词,如此直至全部或者其中一个句集中没有词语,最后把全部配对的词语的相似度相加求均值即为句子的名次的相似度 $SimN(S1, S2)$ 。同理可以分别求出 $SimV(S1, S2)$ 和 $SimA(S1, S2)$ 。整个句子的相似度 $Sim(S1, S2)$ 由各类词的相似度加权平均得到:

$$Sim(S_1, S_2) = \sum_{i=1}^3 \beta_i SimS_i(S_1, S_2)$$

式中 $\beta_i (1 \leq i \leq 3)$ 为部分义原计算权值,其取值根据语言学知识及试验得到。

4 基于本体的音乐语义系统模型研究

本系统采用B/S架构,以Java EE技术平台为基础,其总体的软件体系结构如图3所示。

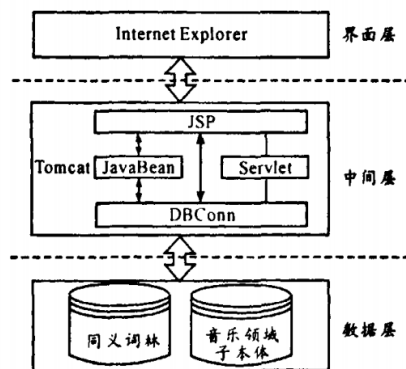


图3 基于语义的音乐检索系统体系结构

系统共分3层,主要工作流程是:用户通过Web浏览器访问系统,在界面层提出检索请求并发送到Tomcat服务器。通过中间层的Java组件对用户输入的自然语句进行分词,关键词提取等操作,将分析后的关键词集合映射到数据层,再通过句子或词语的相似度计算进行概念匹配。如果达到预定匹配度,则进入领域本体的描述逻辑推理机推理并向上输出检索结果;如果概念达不到预定匹配度,则输出显示“无满足要求的结果”。

下面是一个具体实例。用户向系统输入检索请求:“缓解忧伤的歌曲”。系统首先对这一请求进行语法分析和分词处理,得到关键词序列:“缓解 / v 忧伤 / n 歌 / n”,然后进行相似度分析与本体概念进行匹配,匹配结果为:“缓解 / 忧伤 / 音乐”。系统通过调用本体的描述逻辑推理机对概念间的语义关系进行推理就得到相应的音乐个体集合。

结束语

本文针对传统音乐检索系统的不足,构建了音乐领域本体,使用户可以通过自然语言检索音乐。当然,本系统也存在很多的不足之处:一是对本体的分类不够细致不够全面;二是本体的添加和分类还没有实现自动化;三是对本体的利用还不够完全。下一步的工作主要有:①进一步完善音乐领域本体,将现有的音乐数据库添加到本体中来,即实现用户可以进行传统模式与语义模式的双项查询;②与现有的正在研究的基于内容的音频检索系统相结合,实现文字和多媒体的双模查询。

参 考 文 献

- [1] Guarino N , Masolo C , Vetere G 。 Onto Seek:Content-based Access to the Web[J]。 IEEE Intelligent System, 1999。 14(3):70-80。
- [2] 高志强, 潘越, 马力。 语义 Web 原理及应用[M]。 北京: 机械工业出版社, 2009:8。
- [3] 陆建江, 张亚飞, 苗壮等。 语义网原理与技术[M]。 北京: 科学出版社, 2007:122-124。
- [4] 刘建飞。 领域本体的构建及语义检索模型研究[D]。 长春: 吉林大学硕士学位论文, 2011。
- [5] 李朝印。 基于本体的语义检索技术研究与实践[D]。 西安: 西安电子科技大学硕士学位论文, 2015。
- [6] 冉婕, 漆丽娟。 一种基于本体的语义检索设计与实现[J]。 电子设计工程, 2015, 1(5):12-14, 17。