

硕士学位论文

基于情感标签的音乐检索算法研究

Research of Emotion Tag Based Music Retrieval Algorithm

作者姓名：_____李静_____

学科、专业：_____计算机应用技术_____

学号：_____20809400_____

指导教师：_____林鸿飞教授_____

完成日期：_____2010.11_____

大连理工大学

Dalian University of Technology

大连理工大学学位论文独创性声明

作者郑重声明：所呈交的学位论文，是本人在导师的指导下进行研究工作所取得的成果。尽我所知，除文中已经注明引用内容和致谢的地方外，本论文不包含其他个人或集体已经发表的研究成果，也不包含其他已申请学位或其他用途使用过的成果。与我一同工作的同志对本研究所做的贡献均已在论文中做了明确的说明并表示了谢意。

若有不实之处，本人愿意承担相关法律责任。

学位论文题目：_____

作者签名：_____ 日期：_____年____月____日

摘 要

随着数字音乐的爆炸式增长，越来越多的人通过互联网去检索音乐，由此，音乐检索研究越来越受到人们的重视。传统的音乐检索基本上是基于歌名、歌手、歌词中的关键词等从曲库中检索出符合条件的音乐作品。这种检索方式要求用户必须记得相关信息才能进行检索，已经不能够满足人们日益增强的检索需求。在基于内容的音乐检索技术的研究中，通常考虑音乐的节奏、音色、音高、强度等物理特征作为内容进行匹配检索，由于这些特征过于专业化，很难得到用户的接受。音乐作为情感的载体，聆听音乐的时候，就是体会人类情感、抒发人类情怀的时候。基于情感的音乐检索能够满足用户通过输入自己的心情来检索音乐，更加个性化。

Web2.0的发展提供了用户自由的行为，用户可以通过自己对音乐的理解定义标签，不同的标签从不同角度诠释着用户对音乐的理解。其中，大量的情感类标签从多个角度描述同一首音乐所表达的情感信息，同一个情感标签也可以去描述不同的音乐。本文提出了一种基于用户情感标签的音乐检索算法，让用户通过输入情感标签检索到满足心情的音乐，满足其个性化的检索需求。

本文提出的方法对返回的音乐主要有两个要求：一是相似度高，例如：用户检索“悲伤”的歌曲时，要求返回的音乐能够最大程度地表达出用户此时悲伤难过的心情，本文通过构建以情感标签和音乐为节点的二部图，利用 T_SimRank 算法计算标签和音乐之间的语义相似度，然后，利用短文本的语义相似度计算来得到查询标签串和音乐的相似度；另一个是热度，即返回歌曲的“流行度”。通过对大量音乐网站的调研发现，用户更加倾向于选择比较热门的歌曲，因此，本文基于音乐为节点的图模型提出了音乐热度排序 T_PageRank 算法。实验中，将二者进行融合，同时考虑相似度和热度两个因素，通过排序学习的方法得出音乐的最终排序，从而更好的满足用户体验。

另外，针对新音乐没有标签出现的“冷启动”问题。本文通过分析歌曲的歌词内容、结构、唱速等信息来挖掘出新音乐的“隐性”情感基因，即给一首新的音乐去标注一些“虚拟”的情感标签，从而让它在基于情感标签的音乐检索中适时的返回给用户。

关键词：音乐检索；音乐情感相似度；音乐热度；排序学习；标签推荐

Research of Emotion Tag Based Music Retrieval Algorithm

Abstract

With the booming of digital music on the Internet, more and more people search music on internet, so people pay more and more attention on the research of music information retrieval. Traditional MIR (music information retrieval) system which is based on the query of song's name, singer and lyric cannot meet people's various needs. Content based music retrieval technology usually consider music physical characteristics such as rhythm, tone, pitch and intensity, it is too specialized to be accepted. As there is a close connection between music and emotion, the idea of emotion based music retrieval is proposed. Emotion based MIR system allows users to search music via imputing emotional words, which satisfies the individual needs.

As the prevalence of Web 2.0, web users are bestowed more freedom of motion to define music tags according to their own understanding of music, which denote users' distinguishing comprehension of music. Music is to express emotions and interpreted by tags, so the tags annotated by users inevitably involve emotion tags that reflect joviality or sorrow. In this paper, we study the novel problem of integrating emotion tags into music retrieval. Two novel iterative algorithms are proposed to improve the performance on MIR: (1) T_SimRank algorithm, which is to compute latent semantic similarity between tags and music based on a Bipartite graph of tags and music. (2) T_PageRank algorithm, which is to measure the popularity of music using tags. In the experimentation, taking the advantage of learning to rank, we combine many methods to get the final ranking results.

In addition, as the new music is't labeled by any tags, it could cause the emergence of "cold start" problems. This paper analyzes the lyrics, structure, and sing speed to mine the "hidden" emotional genes, after adding some "virtual" emotions tags, the new music could be returned to user by emotion tag based MIR system.

Key Words: Music Retrieval; Music Emotion Similarity; Music Popularity; Learning to rank; Tag Recommendation

目 录

摘 要.....	I
Abstract.....	II
1 绪论.....	1
1.1 研究背景.....	1
1.2 研究现状.....	3
1.3 本文的工作.....	4
1.4 本文的结构.....	5
2 相关知识及评价方法.....	6
2.1 音乐的情感分析.....	6
2.2 音乐情感标签及语义相似度计算.....	7
2.3 排序学习方法.....	10
2.4 评测方法.....	12
2.4.1 语料.....	12
2.4.2 评测指标.....	13
2.5 本章小结.....	14
3 情感标签语义相似度算法.....	15
3.1 基于标签共现的语义相似度.....	15
3.2 SimRank 算法及改进.....	15
3.2.1 SimRank 基本思想.....	15
3.2.2 T_SimRank 算法.....	18
3.2.3 T_SimRank 算法复杂度分析及改进.....	20
3.3 实验设计及结果分析.....	23
3.3.1 实验语料及评价方法.....	23
3.3.2 对比实验设计.....	23
3.3.3 实验结果及分析.....	24
3.4 本章小结.....	25
4 基于标签的音乐热度算法.....	27
4.1 音乐热度计算.....	27
4.1.1 PageRank 算法.....	27
4.1.2 T_PageRank 算法.....	28
4.1.3 T_PageRank 算法复杂度分析及改进.....	30

4.2	实验设计与结果分析	30
4.2.1	单种方法融合热度信息结果	31
4.2.2	多种方法融合结果	32
4.3	本章小结	33
	下一步计划	34
5	新音乐的情感标签推荐	35
5.1	标签推荐技术	35
5.2	基于歌词情感相似度的标签推荐	36
5.3	本章小结	41
结 论	42
参 考 文 献	43
致 谢	48
	大连理工大学学位论文版权使用授权书	49

1 绪论

1.1 研究背景

音乐作为一种表达内心情感的艺术，一直以来在人们的生活中占据不可缺失的重要地位。随着互联网技术的不断发展，更多的人开始在互联网上听歌、看电影、浏览电子小说，去丰富他们的业余生活。艾瑞咨询 (<http://www.iresearch.com.cn/>) 最新调查显示，2009年97.6%的中国网民选择从互联网上获取音乐，这一比例远远高于传统的CD、磁带以及新兴的移动互联网渠道。互联网已成为音乐内容传播的重要渠道。

传统的音乐检索基本上是基于音乐的外部基因^[1]（比如标题、词曲作者、演唱/演奏者、歌词中的关键词等），从音乐库中检索出符合条件的音乐作品，本质上只是一种基于关键词的文本搜索，用户必须记得对应的相关信息才能够去检索音乐，比如Google的音乐搜索功能或百度mp3搜索等。当一个用户在商场或者路边听到一首好听的歌曲时，由于他无法获知这首歌曲的歌手或者歌名等信息，因此，传统的音乐检索方法不能够满足他的需求。

从20世纪90年代开始，人们开始研究一种基于内容的音乐检索CBR（Content Based Retrieval）技术。基于内容的音乐检索通过对音频的幅度、频道等物理特征和响度、音高、音色等听觉特征进行分析，对不同音频数据赋予不同的语义，保证相同语义的音频在听觉上保持相似性。基于内容的音乐检索可以有两种输入方式，音频信息或用户哼唱信息。在实际应用中，通过输入一小段音频去检索音乐的方式推广起来不太容易，因此，一般推荐用户通过哼唱一段旋律去检索相应的音乐。哼唱检索方式不仅要求有相应的语音输入设备，对于唱歌走调的音乐爱好者苦不堪言，往往检索不到用户所期望的音乐。

情感在音乐艺术中，占有极为重要的地位。音乐的情感分析一直以来是音乐学家和心理学家的重要研究课题，人们通过音乐去表达他内心的情感。法国作曲家柏辽兹曾说过：“音乐是心灵的迸发。它不象化学那样能进行实验分析。对伟大的音乐来说只有一种真正的特性，那就是感情。”。匈牙利著名的钢琴家李斯特也指出：“音乐是不假任何外力，直接沁人心脾的最纯的感情的火焰；它是从口吸入的空气，它是生命的血管中流通着的血液。感情在音乐中独立存在，放射光芒。”。我国汉代的《毛诗序》也指出：“情动于中而形于言，言之不足故嗟叹之，嗟叹之不足故咏歌之，咏歌之不足，不知手之舞之，足之蹈之也。”。歌词、节奏、旋律、曲风等从不同的角度反应出音乐所携带的情感信息，随着互联网的发展，基于情感的音乐检索已经开始起步，人们希望通过描

述自己的心情来检索到合适的音乐，例如：在一次聚会上，大家都非常高兴，希望能够播放一些欢快的音乐来配合当时活跃的气氛，用户就可以通过检索“欢快”、“高兴”等情感词语来达到目的，返回的这些歌曲跟歌手、歌词、歌名等毫无关系，只需能够烘托当时的气氛即可。Google 音乐挑歌功能允许用户自定义节奏、声调、音色、年代、歌手、流派和语言等信息来选择歌曲，但是，一般用户对音色、节奏、声调等比较专业的音乐知识了解相对较少，所以，让用户自己设置音乐音频特征的方式不能满足大多数人的需求。图 1.1 为 google 音乐挑歌功能的检索界面。



图 1.1 google 音乐挑歌界面

Fig. 1.1 Frame of google music

Web2.0 的发展提供了用户自由的行为，用户可以通过自己对音乐的理解定义标签，不同的标签从不同角度诠释着用户对音乐的理解。由于音乐跟情感密不可分的关系，用户的音乐标签中存在大量的情感类标签，这些标签或者表达欣喜欢快的心情，或者表达悲伤失落的情愫。基于情感类标签的音乐检索，允许用户通过描述自己的心情来检索适合的歌曲，更能够满足用户个性化的检索需求。现有的音乐搜索引擎中，基于标签的音乐检索主要是通过标签云或分类标签中的链接来实现的，这种方法只允许用户通过单个标签进行检索，当用户想听“悲伤”的“爱情”歌曲时，只能通过检索“悲伤”或者“爱情”来查找，而无法查找同时满足“悲伤”和“爱情”歌曲；另外，传统的音乐搜索引擎一般是通过布尔检索返回所有包含该标签的音乐，然后按照一定的顺序排序，如：音乐的播放次数，或者出版时间等，很少考虑音乐跟标签的相似度信息。百度 MP3 通过搜集用户在“音乐掌门人”中对歌曲的描述，计算标签跟歌曲的相关性，最终按照标签

与歌曲的相关性进行排序。有道音乐“随心听”通过爬取互联网上大量的用户对歌曲的评价信息，将用户听歌的心情分为伤感，爱情，温馨，安静等 10 类，只要用户选择自己的心情，就可以检索到适合这种心情的音乐。

1.2 研究现状

目前，基于用户情感的音乐检索和音乐推荐，主要是从音乐自身的曲风、旋律、节奏、音色、强度、频谱质心、能量等来获取音乐的情感信息，没有利用用户标签中大量的情感标签信息^[2-3]。Dan Liu 等^[4]提出了基于 GMM 的层次化情感检测系统，通过对声音数据的强度、音色和节奏三个特征使用 Thayer 情感检测模型对音乐片段进行分类。浙江大学的 Yazhong Feng 等^[5]使用音乐情感分析实现了一套可视化的音乐信息检索系统，在情感描述方面，使用了 Juslin 的理论，将音乐的情感简单分为四类：快乐、悲伤、气愤和恐惧。在分类之前，通过对节拍进行检测，对清晰度进行建模，最后使用 BP 神经网络构建分类器。文献[6]也是通过选取音乐的音频特征用 SVM 进行情感分类。但是，音频信号的研究已有近 20 年的历史，至今无法获得准确反映情感的音频特征，所取得的成效非常有限，无法达到满意的水平^[7]。

考虑到目前音频信号在音乐情感分析中的局限性，越来越多的研究者开始从文本的角度去分析音乐的情感，主要是从歌曲的歌词内容、结构、节奏等角度去分析，文献[8]从歌词文本中提取了 182 个能够描述心理的特征词，通过机器学习的方法将音乐划分为 23 个类别。陈若涵等^[9]用 KNNR、GMM、SVM 三类分类方法，通过解析歌词中的情感词汇辅助进行情感分析，将古典音乐和流行音乐分为四种情感。文献[10-12]同时考虑音乐的音频和歌词文本两种信息进行情感分类，但是，由于所选的语料集规模较小，同时，在分析歌词文本信息时，只考虑了简单的基于“词袋”的方法，用 TF-IDF 计算特征词权重，效果不是特别明显。Xiao Hu 等^[13]比较了音频跟歌词在情感分类中的作用，实验发现，在音乐样本比较稀疏或者歌词的语义信息比较明确时，基于歌词文本的情感分类效果更好。结合音频和歌词的情感分类对于大部分的音乐情感类别效果都比较好大部分类别，文本特征数太多容易引入较多的噪音信息，从而不利于音乐的情感分类。在歌词特征提取方面，比较 BOW、POS、FW 三种词特征选取方法，证明基于“词袋”（BOW）的方法效果最好。清华大学的夏云庆等^[7]在利用歌词分析音乐情感时，提出了情感向量空间模型的初步设想，在特征定义中以情感单元取代词汇，以情感单元的统计量作为情感特征，使得音乐情感分析取得较明显地提高。

无论是从音频的角度去分析音乐情感，还是从文本的角度去挖掘音乐情感，研究者对音乐的情感判断主要有两个来源：（1）从研究者自己的角度去进行标注，这个容易

导致情感判断的主观性，以及没有公开语料对研究的扩展性和可信度带来的挑战，文献[2,4,7,9,10,11,12,14,15]均采用这种标注形式；（2）专家标注，从2005年开始的MIREX（Music Information Retrieval Evaluation eXchange）比赛中，有一个AMC（Audio Mood Classification）子任务，提供部分专家对音乐的情感标注信息，为研究者提供了比较公开的研究语料。但是，专家的标注是建立在他们严格的专业水平之上，跟普通的用户之间会产生一个明显的“代沟”。文献[16]中提到，2007年ACM任务中提供的专家情感标注跟大众情感判断存在30%的差异。作为Web2.0技术的核心，标签体现了用户从自己的角度概括网络资源内容的着重点或者兴趣点，往往能够比较真实的反映大众的观感^[17]。文献[17-20]研究表明，利用标签信息可以优化搜索引擎，提高搜索引擎的准确率。文献[21]通过对基于标签和视频元数据的检索效果进行分析，表明标签信息对视频检索的重要性。文献[22-23]研究表明，标签可以有效提高图片检索的效率。文献[24-25]分别对音乐网站last.fm和“百度音乐掌门人”的标签进行统计分析，文献[24]将标签分为音乐情感、曲风、适合状况、类别、歌手、地域语言6大类，其中，音乐的情感类标签所占的比重最大。文献[25]将用户标签分为主题、时代、地域、类型、歌手、情感、适用场合及用户个人信息8大类，经过统计发现，用户情感类标签是除了曲风类标签外数量最多的，并且对于同一首音乐，用户的情感认知基本是相似的。Xiao Hu等^[13]在将音乐情感分类时，没有直接利用音乐情感心理模型，而是考虑到用户行为，通过用户标签将音乐情感分为18大类，与传统的专家情感分类相比较，更贴近一般用户的情感。因此，综合考虑标签在信息检索中的贡献以及情感与音乐密不可分的关系，研究情感类标签对音乐检索的价值十分有意义，基于情感标签的音乐检索能够满足用户通过描述自己的心情来检索音乐的个性化需求。

1.3 本文的工作

本文考虑到目前基于音频的音乐情感分析的局限性，通过分析音乐与情感的内在联系，实现了基于多个用户情感标签的音乐检索，返回的音乐主要有两个特点：一是相似度高，例如：用户检索“悲伤”的歌曲时，要求返回的音乐能够最大程度地表达出用户此时悲伤难过的心情，本文通过构建以情感标签和音乐为节点的二部图，利用相似度计算方法T_SimRank去计算标签和音乐之间的语义相似度，然后，利用短文本的语义相似度计算来得到查询标签串和音乐的相似度；另一个是热度，即返回歌曲的“流行度”。通过对大量音乐网站的调研发现，用户更加倾向于选择比较热门的歌曲，文献[15]将音乐的热度引入了音乐检索中，在很大程度上提高了音乐的检索效率。本文考虑到标签本身的相似度信息，对文献[15]中的算法进行改进，得到音乐热度排序T_PageRank算法。

本文将二者进行融合，同时考虑相似度和热度两个因素，通过排序学习的方法得出音乐的最终排序，从而更好的满足用户体验。

基于情感标签的音乐检索依赖于用户的标注，是利用歌曲在被用户所关注的前提下所显示出来的“显性”基因，但是，对于一首新的音乐，在它还未被大众所熟识的情况下是没有情感标签的，因此，会出现“冷启动”问题。本文通过分析歌曲的歌词内容、结构、唱速等信息来挖掘出新音乐的“隐性”情感基因，即给一首新的音乐去标注一些“虚拟”的情感标签，从而让它在基于情感标签的音乐检索中适时的返回给用户。

1.4 本文的结构

论文共分为五章，详细阐述了基于情感标签的音乐检索算法、实验设计、性能评估以及针对新歌曲采用的情感标签预测方法。具体章节安排如下：

第一章，绪论，综述了论文所研究内容的背景以及研究现状，介绍了本文研究的主要方法和论文的结构安排。

第二章，主要介绍基于情感的音乐检索算法的相关技术以及本文所用的实验语料和评价方法。

第三章，详细介绍音乐检索中的标签语义相似度计算方法，以及实验设计与结果分析。

第四章，详细介绍音乐检索中基于标签的音乐热度计算方法，以及实验设计与结果分析。

第五章，详细介绍针对新歌曲的“冷启动”问题，介绍一种基于内容推荐的新音乐情感标签预测方法。

论文的总结，介绍了本文研究的内容、主要工作及下一步的工作。

2 相关知识及评价方法

2.1 音乐的情感分析

情感是音乐的本质特征，音乐通过声波振动的非语义性组织结构与人类的情感、意志活动具有直接同构的对应关系^[26]。一直以来，音乐的情感分析是音乐学家和心理学家的重要研究课题，随着计算机科学技术与互联网技术的飞速发展，互联网上大量的数字音乐信息让人们目不暇接，利用计算机去分析音乐的情感信息也越来越得到人们的重视，通过分析歌词、节奏、旋律、曲风等各方面所携带的情感信息，可以应用到基于情感的音乐检索领域。目前，在研究领域中应用比较广泛的音乐情感认知模型有离散形式的 Hevner 模型^[27]和维度形式的 Thayer 模型^[28]等。

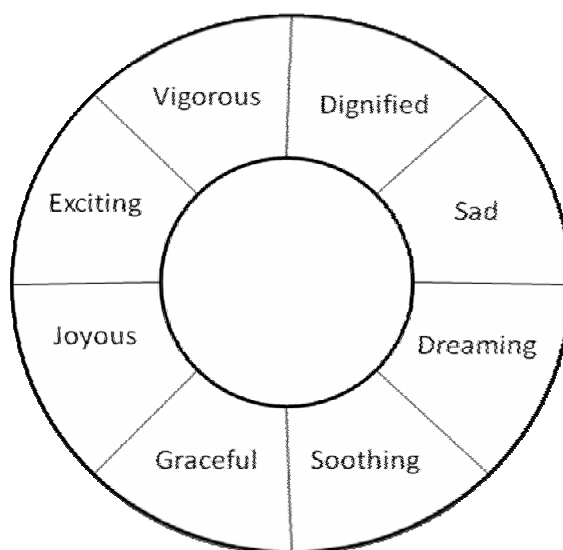


图 2.1 Hevner 情感环模型

Fig. 2.1 Hevner Emotion Model

Hevner 模型采用一个如图 2.1 所示的环形结构，用来象征心境、情绪或者情感反应，它运用 67 个情感形容词描述艺术领域的不同情感属性，并按照类别归为 8 个情感相近的类型。在 Hevner 情感环中，其中每一个环节都与它前后相连的环节在情感逻辑上存在一定程度的递进关系，象征着不同情感在音乐描述中的有规律的变化。音乐的 Hevner 情感环首次提出是在西方，研究对象也是西方音乐，研究受体是西方群体，由于由于文化、思维、习惯、历史背景等不一样，中外在表达情感时也存在差异，针对这个问题，

文献[2]通过调研比较，将 Hevner 情感环中的词表数从 67 减到 47，所选择的情感形容词更符合中国人的表达习惯。

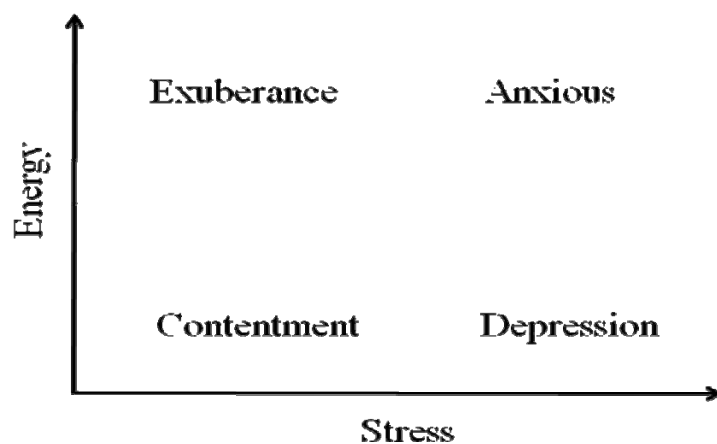


图 2.2 Thayer 二维模型
Fig. 2.2 Thayer Emotion Model

Thayer 模型采用如图 2.2 所示的能量 (Energy) 与压力 (Stress) 两个维度描述情感。能量维度，是指从“充满活力的”到“平静的”不同水平的反应，压力是从“快乐”到“焦虑”的反应。一般研究在用到 Thayer 模型时，将其分为满意 (Contentment)、消沉 (Depression)、充沛 (Exuberance) 以及忧伤或狂乱 (Anxious) 四大类。

情感模型的选取依赖于具体的应用和情感识别方法，其中，Hevner 模型从音乐学角度提出，主要考虑歌曲演唱者和听众的心理感受，符合音乐情感的心理交互实际情况，文献[2,29]在研究中使用了 Hevner 情感模型。Thayer 模型是心理学角度提出的，采用维度的思想描述音乐情感，比较与音乐特征中的能量、幅值等进行关联，因此适用于音频的情感识别中，文献[30-32]使用了 Thayer 模型。

2.2 音乐情感标签及语义相似度计算

潘宇等^[1]提出了“音乐基因组”概念，将音乐分为“内部基因”和“外部基因”两部分。图 2.3 为音乐的基因组结构，主要分为内部基因和外部基因两大组。内部基因是用户在欣赏音乐时可以直接感触到的，如音乐的歌词和旋律 (音频)。这些特征对于一首音乐来说是唯一的且无法改变的。外部基因是指用户欣赏音乐时，不会被用户直接感知到，但是却从不同的方面反映出音乐的特征，如歌手性别、歌手名、语种、地域、音乐所携带的情感信息和曲风信息等。通过观察可以发现，在音乐的外部基因中，有些基因是确定唯一不变的，例如音乐名，歌手名，歌手性别等；有些基因，不同的用户则可能会有不同的认知，例如音乐的情感，曲风等。针对这一特征，将音乐的外部基因组划

分为固定基因和自由基因两种类型。固定基因是指用户无法改变的音乐固有的特征。自由基因是由用户定义的，能够完全反映用户认知的特征。

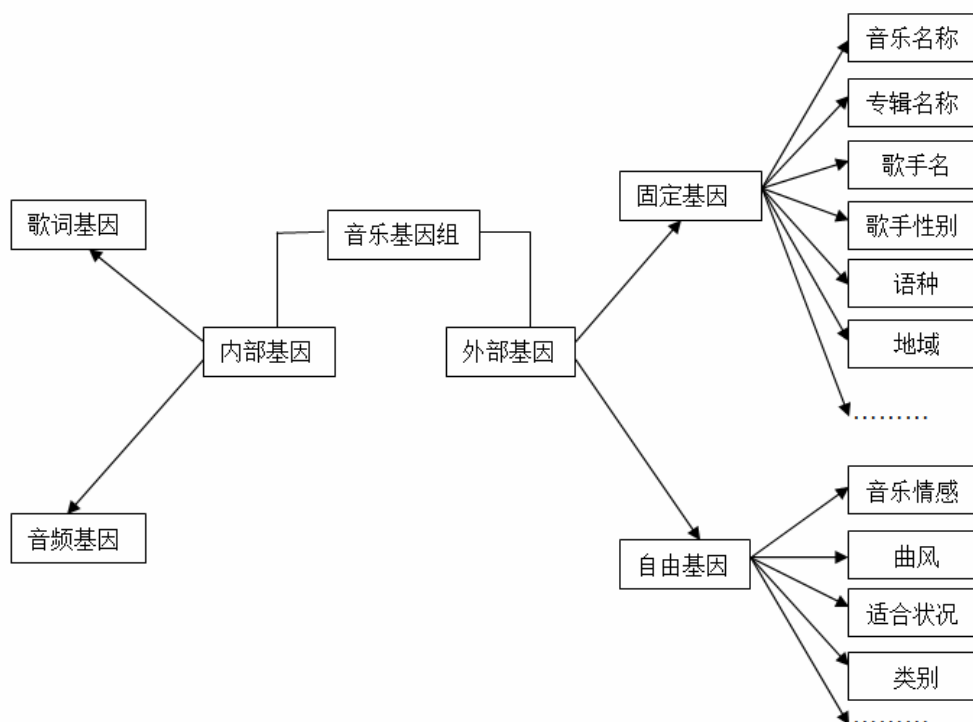


图 2.3 音乐基因组结构

Fig.2.3 Music genome structure

标签作为用户自由标注的产物，包含了用户对所分享内容理解，是对所分享内容语义的附加信息。标签通常是一些非常精炼的短语，可以帮助用户简捷的获取所需信息，节省用户的浏览时间。目前的 Web2.0 网站，如 last.fm、Flickr、Delicious、YouTube、Youku、豆瓣等都支持用户的标注行为。在音乐网站中，标签是用户对音乐的理解和诠释，在音乐基因组中属于“自由基因”。对标签的研究有助于统计和分析用户的行为、兴趣和习惯等。作为 Web2.0 技术的核心，标签体现了用户从自己的角度概括网络资源内容的着重点或者兴趣点，往往能够比较真实的反映大众的观感。文献[17-20]研究表明，利用标签信息可以优化搜索引擎，提高搜索引擎的准确率。文献[21]通过对基于标签和视频元数据的检索效果进行分析，表明标签信息对视频检索的重要性。文献[22-23]研究表明，标签可以有效提高图片检索的效率。标签不仅支持用户的浏览行为，还支持基于标签的用户检索行为。百度音乐掌门人允许用户以“Tag:+标签”的方式来检索音乐。

同时,还允许用户通过“标签云”(Tag Cloud)^[32]的方式来进行检索。标签云是标签的视觉化描述,它把标签的字体大小与它的热门度关联起来,如果一个标签被越多的用户标记,那么它的字体就越大。图 2.4 为国内某音乐网站的标签云分布。标签的排列顺序一般依照字典排序或者随机排序,按照热门程度确定字体的大小和颜色。

johnny's live mika muse music neo-folk newage piano pop pop-punk post-rock postrock punk rap rock shoegaze single trip-hop vr yui 嵐 ドラマCD 爱尔兰 澳大利亚 冰岛 蔡健雅 蔡依林 曹方 陈珊妮 陈升 陈奕迅 陈绮贞 纯音乐 大塚爱 大陆 戴佩妮 单曲 德国 电影原声 电子 东方神起 动漫 独立 独立音乐 法国 翻唱 范晓萱 范玮琪 方大同 芬兰 风潮 钢琴 港台 古典 古典音乐 国内 国语 韩国 合辑 后摇 华语 华语音乐 加拿大 精选 经典 久石让 爵士 雷光夏 李志 梁静茹 梁咏琪 林俊杰 林一峰 林宥嘉 刘若英 流行 卢广仲 马来西亚 美国 迷幻 民谣 摩登天空 莫文蔚 男歌手 男声 内地 女歌手 女声 挪威 欧美 清新 日本 日本音乐 瑞典 手寫藝 苏打绿 孙燕姿 蘇打綠 台灣 台湾 陶喆 同人音乐 王菲 王力宏 王若琳 五月天 香港 萧亚轩 小清新 小野丽莎 新加坡 熊木杏里 许巍 许哲佩 演唱会 摇滚 音乐 英国 英伦 英伦摇滚 宇多田ヒカル 宇多田光 原声 粤语 张国荣 张惠妹 张信哲 张悬 张学友 张震岳 张靓颖 治愈系 中岛美嘉 中国 中国摇滚 周杰伦 椎名林檎 自然卷 组合 左小祖咒 菅野よう子 浜崎あゆみ 陳綺貞 窦唯

图 2.4 标签云示例

Fig.2.4 The example of tag cloud

文献[24-25]分别对音乐网站 last.fm 和进行统计分析,文献[24]将“百度音乐掌门人”的标签分为音乐情感、曲风、适合状况、类别、歌手、地域语言 6 大类,其中,音乐的情感类标签所占的比重最大。文献[25]将音乐网站 last.fm 的标签划为主题、时代、地域、类型、歌手、情感、适用场合及用户个人信息 8 大类,经过统计发现,用户情感类标签是除了曲风类标签外数量最多的,并且对于同一首音乐,用户的情感认知基本是相似的,因此,情感类标签对基于标签的音乐检索有非常大的贡献。

同一首音乐所蕴含的情感不是唯一的,从开始到高潮会有一个情感的转变过程,如,德彪西的《水中倒影》的情感变化从开头的静谧、梦幻、轻快、略有活泼到中部的不安、兴奋、激情、欢快等,这种情感的变化也符合 Hevner 情感环。另外,不同的用户由于年龄、性别、职业、性格、习惯、文化背景、所处环境不同对音乐的理解也是不同的,比如,对于周杰伦的《稻香》,温馨的旋律、温暖的歌词,有的人从节奏的角度去评价,认为是一首比较欢快的歌曲;有的人从歌词的词义分析,里面含有“不要这么容易就想

放弃 就像我说的 追不到的梦想 换个梦不就得了 为自己的人生鲜艳上色 先把爱涂上喜欢的颜色 笑一个吧 功成名就不是目的 让自己快乐快乐这才叫做意义”认为是一首积极的歌曲，也有的人认为多次提到“童年的纸飞机”、“乡村的歌谣”、“萤火虫”、“赤脚在田里抓蜻蜓”等回忆童年，回忆家乡的信息，这首歌曲表达的是一种思念跟回忆的情愫等。综上所述，同一个歌曲本身就有多种情感，不同的人又会有不同的情感认知，因此，同一首歌曲可以被标注为多个情感类标签，这些标签从不同的角度描述了音乐所要表达的情感信息。以周杰伦的《简单爱》为例，由于不同的人感受不同，这首歌被标注为“优美”、“休闲”、“初恋”、“夏天”、“安静”、“开心”、“柔情”等多个情感标签。

本文中的情感类标签不仅仅包括“欢快”、“悲伤”、“开心”、“忧郁”等传统意义的情感词，还包括能够表达情感信息的“失恋”、“夏天”、“阳光”、“爱情”等词语。对于词性也不要求必须是描述心情的形容词，还可以是相应的名词、动词、副词等类型的词。

不同的音乐也可以表达同一种情感，换句话说，同一种情感也可以通过不同的音乐来表达，尤其是对于大众音乐爱好者而言，表达情感的词汇相对较集中，通常，同一个情感标签也可以描述不同的歌曲，例如：“悲伤”同时描述了梁静茹的《会呼吸的痛》和《可惜不是你》、南拳妈妈的《下雨天》、阿桑的《一直很安静》等多首歌曲。标签和音乐构成“多对多”的关系。

基于标签语义相似度的音乐检索避免了传统基于标签的音乐检索系统中布尔匹配的局限性和返回结果对标签的完全依赖性，能够从更高的语义层次去分析标签跟音乐之间的相似度，从而更加客观的返回结果。例如，基于标签语义相似度的音乐检索系统能够检索到不包含该标签，但是却与它相似程度很高的音乐。例如：以标签“难过”进行检索时，音乐“会呼吸的痛-梁静茹”虽然没有被标注为“难过”，但是它所有的标签与“难过”的相似度都很高，所以，它可能以较高的相似度被排在其他被标注为“难过”的音乐前面。

2.3 排序学习方法

排序学习（learning to rank）是现在非常热门的一个研究领域，它是将传统的信息检索与机器学习相结合的研究课题，其核心问题是如何构造一个函数或模型反映文档对于查询的相关度^[33]，换句话说，就是将传统的排序方法作为多个特征融合成一个排序函数的方法，其目的是优化损失函数。用机器学习的方法来解决排序学习问题大致可以分