

DOI: 10.7652/xjtuxb201404006

面向推荐系统的音乐内涵空间建模研究

秦继伟¹, 郑庆华²

(1. 新疆大学网络与信息技术中心, 830046, 乌鲁木齐; 2. 西安交通大学计算机科学与技术系, 710049, 西安)

摘要: 针对现有音乐推荐系统中音乐情感分析方法很难满足用户情感需求的问题,提出音乐内涵空间方法分析音乐情感。该方法选取音乐力度、速度、音强等表现要素,基于两极尺度的语义,构建音乐内涵空间。通过音乐内涵空间表示音乐情感,减少了音频特征客观水平与主观情感范围之间的差距,解决了情感标签标注音乐中由于用户情感经历不同造成的对音乐情感标注的主观差异性问题。在 10 672 条音乐评价数据集上,采用 Kendall's tau 距离进行有效性验证,与基于情感标签推荐相比,实验结果表明,基于音乐内涵空间推荐音乐能较好地满足用户的情感需求。

关键词: 内涵空间;音乐;推荐;Kendall's tau 距离

中图分类号: TP391 **文献标志码:** A **文章编号:** 0253-987X(2014)04-0031-04

Modeling of Music Connotative Space for Recommender System

QIN Jiwei¹, ZHENG Qinghua²

(1. Center of Network and Information Technology, Xinjiang University, Urumqi 830046, China;

2. Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Focusing on the music affective, a concept of music connotative space is introduced. The semantic polar scale is examined and the connotative space is constructed by musical strength, speed and intensity for music affective recommendation. The music connotative space bridges the gap between the objective levels of audio frequency characters and subject emotional rang inequality, and deals with the difference of music emotional label due to different listener experiences. According to 10 672 music records. the effectiveness of the proposed music connotative space is validated in terms of Kendall's tau distance. The music connotation space outperforms emotion label in music classification for music affective recommendation.

Keywords: connotative space; music; recommendation; Kendall's tau distance

音乐作为诱导情感的资源之一,不仅具有娱乐作用,同时蕴含着丰富的情感内涵,具有增强或者减弱听者情感的功能,常被用来推荐给听者满足其情感需求。由于情感受到上下文环境、以往经历、认知水平等影响,它的产生、变化是一个复杂动态的过程。因此,如何分析和确定音乐情感,成为音乐推荐研究的热点问题。

1 相关工作

目前,基于音乐情感的推荐系统中音乐情感分

析主要是通过对音乐内容分析添加情感标签实现。音乐内容的研究围绕音高、音强、音色、旋律、歌词等一系列给人们带来的听觉、感知特性的特征量进行分析。这里,音乐内容特征分为声学特征和语义特征^[1]。在声学特征研究方面,例如,采用 Thayer 的 AV (Arousal-Valence) 情感模型对声音数据的强度、音色和节奏等 3 个特征进行分析,建立基于高斯混合模型的层次化情感检测系统,实现对音乐片段进行分类^[2];文献[3]选取音乐的音频特征用支持向量机方法对音乐情感进行分类。在语义特征方面,

收稿日期: 2013-12-06。 作者简介: 秦继伟(1978—),女,工程师;郑庆华(通信作者),男,教授,博士生导师。 基金项目: 国家自然科学基金资助项目(41261087,61261036,61308120);中国博士后科学基金资助项目(2013M540783)。

网络出版时间: 2014-03-04

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20140304.0943.001.html>

研究者从歌词入手去分析音乐的情感,比较了高斯混合模型、K近邻、支持向量机3种分类方法,并结合歌词中的情感词辅助进行情感分析^[4];文献[5]以情感单元取代词汇,以情感单元的统计量作为情感特征,建立音乐情感向量空间模型,提高了情感的分类精度;在此基础上,文献[6]提出情感向量空间模型和“情感词-情感标签”共现的特征降维方法,与传统的文本特征向量模型相比,能够更好地提高歌曲情感分类准确率。

上述音乐情感分析的研究成果有效地推动了音乐推荐系统的发展,并被应用到基于音乐情感的推荐系统研究领域。通过深入分析,我们不难发现,在基于音频信号的声学特征研究方面,由于从音频信号中提取出来的特征众多,很难从中辨识与情感相关的特征;在歌词特征方面,歌词具有文本短小、不规范和存在重复段等特征,因此采用现有的文本分类方法,很难准确划分歌曲的情感类别,并且以歌词特征评价音乐情感类别的方法不适用于没有歌词的音乐,因而具有一定的局限性。因此,本文针对以上存在的问题,从音乐创作的角度出发,提出音乐内涵空间模型的方法来分析音乐的情感。

2 建立音乐内涵空间

音乐作为一种艺术表达形式,弹奏者或者演唱者都应反映创作者赋予音乐的内涵。本文从音乐创作者的意图出发,将内涵空间概念应用到音乐作品中,通过力度、速度和音强3个音乐表现要素的变化来反映音乐的情感,建立音乐的内涵空间,分析音乐的情感。

2.1 相关定义

音乐内涵空间以音乐的力度、速度和音强3个特征为纬度,反映音乐的情感。力度是音乐表现的重要手段。一般来说,力度越强,音乐越雄壮、紧张;力度越弱,音乐越缓和、委婉。在音乐内涵空间中,设力度轴为 H^L 。速度是一个非常重要的音乐元素,它影响着作品的情感。在音乐内涵空间中,设速度轴为 H^S 。音强是语音要素之一,指声音的客观物理强弱,对音乐情感表达起辅助作用。在音乐内涵空间中,设强度轴为 H^Y 。

2.2 音乐情感内涵的取值

基于两极尺度的语义将力度、速度和音强按照大小、快慢和强弱等程度划分为5个等级,分别用1、2、3、4、5表示。其中:1表示力度很小,速度很慢,音强很弱;5分别表示力度很大,速度很快,音强很强;2、3、4则表示音乐在3个特征上处于中间程度。

由此,音乐的力度、速度、音强3个值对应在内涵空间的点反映了该音乐作品的情感。例如,用户对音乐《寂寞在唱歌》在力度、速度和音强上的取值分别为2、3、4,则在音乐内涵空间中点 $\{2,3,4\}$ 反映了音乐《寂寞在唱歌》的情感内涵。

2.3 信度检验

为了评价用户在音乐内涵空间中标注值与音乐情感内涵实际值之间的一致程度,可采用内部相关系数 I_{CC} ^[9]来衡量和评价用户信度^[7]

$$I_{CC}(1, k) = (B_{MS} - W_{MS}) / B_{MS} \quad (1)$$

式中: B_{MS} 是所有用户对每首音乐的平均评分与所有音乐的平均评分之间的方差; W_{MS} 是某个用户对音乐的评分与所有用户对音乐评分之间的方差。假如, k 为用户数; M 为用户评价的音乐数量; x_{ij} 为用户 u_j 对音乐 m_i 的评分; $\bar{x}_i$ 表示所有用户对音乐 m_i 评分的平均值; $\bar{x}$ 表示所有用户对所有音乐评分的总平均值,则

$$B_{MS} = k \sum_{i=1}^M (\bar{x}_i - \bar{x})^2 / (M - 1) \quad (2)$$

$$W_{MS} = \sum_{i=1}^M \sum_{j=1}^k (x_{ij} - \bar{x}_i)^2 / M(k - 1) \quad (3)$$

显然地, I_{CC} 介于0~1之间,0表示不可信,1表示完全可信。一般认为,信度系数低于0.4表示信度较差,对于定量资料常常需要高的 I_{CC} 。例如,用户1、用户2、用户3对音乐作品《寂寞在唱歌》在内涵空间中的评价分别为 $\{2,3,4\}$ 、 $\{2,2,3\}$ 、 $\{2,4,4\}$,根据上述3个用户在《寂寞在唱歌》内涵空间上的取值,采用相关系数计算出评价《寂寞在唱歌》在力度、速度和音强上的用户间信度为0.703、0.826和0.722,则可信程度较高。

3 内涵空间模型有效性验证

将音乐情感内涵空间模型应用于音乐推荐,通过两组实验验证音乐情感内涵空间模型的有效性。其中,第一组实验在不同歌曲上,比较基于内涵空间推荐的音乐列表、基于情感标签的音乐列表与基于个人偏好形成的最佳音乐列表之间的差异性,验证本文所提的音乐内涵空间模型的准确性。第二组实验在不同用户数目情况下,通过最佳音乐列表与基于内涵空间推荐的音乐列表、基于情感标签推荐的音乐列表的比较,验证了基于音乐内涵空间推荐的音乐更能满足用户的情感需求。

3.1 数据集

情感的类型没有统一的标准,目前未有公共的

音乐数据集。本实验数据集来源于课题组建立的一个音乐推荐评价网站,此网站中102个用户对1548首音乐(来源于百度音乐和酷狗音乐网站)的10672条评分记录。

当多个用户标注一首音乐作品时,为了描述多个用户对这首音乐作品在内涵空间的取值,需分析所有用户在这首音乐作品力度、速度和音强上取值的概率分布。例如,上述3个用户对音乐作品《寂寞在唱歌》在内涵空间中的评价分别为 $\{2,3,4\}$ 、 $\{2,2,3\}$ 、 $\{2,4,4\}$,从以上取值分析《寂寞在唱歌》,在力度上3个用户取值都为2,则这首歌在力度值1、2、3、4、5上的概率分布为 $\{0,1,0,0,0\}$;在速度上3个用户取值为3、2、4,则这首歌在速度值1、2、3、4、5上概率分布为 $\{0,1/3,1/3,1/3,0\}$;在音强上3个用户取值为4、3、4,则这首歌在音强值1、2、3、4、5上概率分布为 $\{0,0,1/3,2/3,0\}$ 。同样地,多个用户对同一首音乐作品标注情感标签时,根据所有用户对这首音乐的情感标注的类别的概率分布进行描述。

3.2 内涵空间模型有效性验证过程

3.2.1 验证过程 在音乐推荐系统中,用户对音乐的偏好通过用户对音乐评分反映,用户对音乐评分的分值越高,表明此用户对这首音乐越喜欢。如果通过音乐数据库中用户 u_a 在悲伤情境下的对所有音乐的评分,根据评分进行排序,选择前10首音乐,形成用户 u_a 最佳音乐列表 $\{m_1^a, \dots, m_i^a, \dots, m_{10}^a\}$;从用户 u_a 最佳列表中选择音乐 m_i^a ,计算数据集中所有用户对音乐 m_i^a 标注的内涵空间的概率分布与用户 u_a 对音乐 m_i^a 标注的内涵空间的概率分布之间的距离,并按照距离由小到大重新排序,形成基于内涵空间推荐的音乐列表;同样的方式形成基于情感标签推荐的音乐列表,根据 m_i 的情感标签从音乐数据库中选择出与音乐 m_i 具有相同情感标签的音乐 m_j ,形成基于情感标签推荐的音乐列表;然后,整体上比较基于内涵空间、基于情感标签推荐的音乐列表与最佳列表的差异程度,差异越小则推荐得越准确。验证过程如图1所示。

3.2.2 评价指标 依据上述描述,首先采用欧氏距离计算数据集中所有用户对音乐 m_i^a 标注的(内涵空间或情感标签)概率分布与用户 u_a 对音乐 m_i^a 标注的概率分布之间的距离

$$d_i = \left(\sum_{k=1}^n (x_{ik} - x_{ik}^a)^2 \right)^{1/2} \quad (4)$$

式中: n 是概率分布向量的维数; x_{ik} 是所有用户对音乐 m_i^a 标注的概率分布向量中的第 k 分量; x_{ik}^a 是用

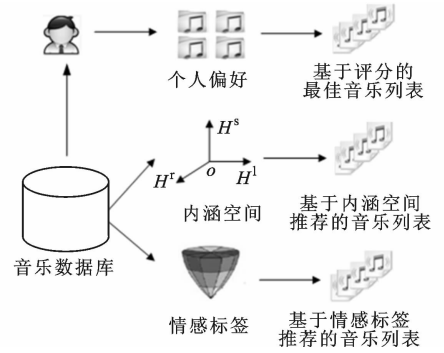


图1 音乐内涵空间验证过程

户 u_a 对音乐 m_i^a 标注的内涵空间的概率分布向量中的第 k 分量。

其次,采用 Kendall' tau 距离^[8]计算基于评分的最佳音乐列表与基于内涵空间推荐的音乐列表、基于情感标签推荐的音乐列表之间的差异程度。假如 x_i 是对应最佳音乐列表第 i 首音乐的推荐序, y_i 是对应在基于内涵空间(情感标签)推荐的音乐列表中第 i 首音乐的推荐序,则

$$\tau = \frac{\sum_{i < j} \text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j)}{(T_0 - T_1)^{1/2} (T_0 - T_2)^{1/2}} \quad (5)$$

$$\text{sgn}(\kappa) = \begin{cases} 1, & \kappa > 0 \\ 0, & \kappa = 0 \\ -1, & \kappa < 0 \end{cases} \quad (6)$$

$$T_0 = n(n-2)/2 \quad (7)$$

$$T_1 = \sum t_i(t_i - 1)/2 \quad (8)$$

$$T_2 = \sum u_i(u_i - 1)/2 \quad (9)$$

式中: t_i 是 x 的第 i 组结点 x 值的数目; u_i 是 y 的第 i 组结点 y 值的数目; n 为音乐列表中音乐的数目。

3.3 内涵空间准确性验证及评价

分析数据集,选取出在数据集上评价数目较多且评分值(评分值1~5,评分越高,用户越满意)分布合理的用户。从数据集中选出用户8,此用户评价了156首音乐,将用户8评分最高的10首音乐作为最佳音乐列表,如表1所示。

首先,按照图1所示验证过程,分别基于内涵空间、情感标签产生相应音乐列表,采用Kendall's tau 距离计算基于内涵空间推荐的音乐列表、基于情感标签推荐的音乐列表与基于评分推荐的最佳音乐列表的距离,结果如图2所示,基于内涵空间推荐的音乐列表与最佳音乐列表的距离小于基于情感标签推荐的音乐列表与最佳音乐列表的距离,本文所提的内涵空间模型准确度较高。

进一步,选定以上10首音乐作为最佳音乐列表,

表 1 最佳音乐列表中音乐内涵空间和情感标签

音乐 ID	音乐名称	内涵空间			情感标签
		力度	速度	音强	
M1	真的爱你	5	4	4	感激
M2	天路	5	3	5	感激
M3	上海滩	5	3	4	自豪
M4	朋友	4	3	4	感激
M5	男儿当自强	4	3	5	自豪
M6	美丽的神话	3	3	4	感激
M7	精忠报国	4	3	5	自豪
M8	红梅赞	2	2	2	自豪
M9	橄榄树	2	3	3	自豪
M10	沧海一声笑	4	3	4	自豪

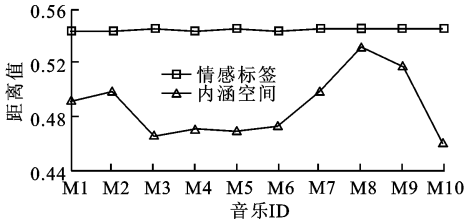


图 2 不同音乐下内涵空间的验证

在不同用户数目下,对用户的评分进行了平均处理,由此基于评分的最佳音乐列表同时被平均;分别通过概率分布量化每首音乐的情感内涵及情感标签,采用 Kendall's tau 距离计算基于内涵空间推荐的音乐列表、基于情感标签推荐的音乐列表与基于评分推荐的最佳音乐列表的距离,结果如图 3 所示,基于内涵空间推荐的音乐准确性高于基于情感标签推荐的音乐,并且当用户较少时,基于音乐内涵空间的推荐音乐更能满足用户的情感需求。

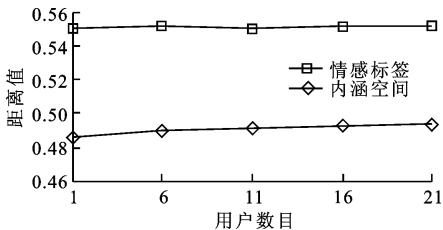


图 3 不同用户数目下音乐内涵空间的验证

通过以上两组实验验证,在相同数据集下,基于内涵空间推荐的音乐与基于情感标签推荐的音乐相比,基于内涵空间推荐的音乐准确度更高,更能满足用户的情感需求。

4 结 论

本文面向音乐推荐系统提出采用内涵空间模型分析音乐情感的方法。以推荐音乐满足用户的情感需求为背景,提出以音乐力度、速度、音强等表现要素为维度,构建音乐内涵空间实现对音乐作品的情感分析。实验结果表明,与基于情感标签推荐的音

乐列表相比,基于本文所提的内涵空间推荐的音乐准确度更高,更能满足用户的情感需求。下一步工作利用用户体验,对基于内涵空间的推荐系统继续展开深入研究。

参考文献:

[1] CASEY M A, VELTKAMP R, GOTO M, et al. Content-based music information retrieval: current directions and future challenges [J]. Proceedings of the IEEE, 2008, 96(4): 668-696.

[2] LIU Dan, LU Lie, ZHANG Hongjiang. Automatic mood detection from acoustic music data [C]// Proceedings of the International Symposium on Music Information Retrieval. Baltimore, MD, USA: The Johns Hopkins University Press, 2003: 81-87.

[3] LI Tao, OGIHARA M. Content-based music similarity search and emotion detection [C]// Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway, NJ, USA: IEEE, 2004: 705-708.

[4] YANG Dan, LEE W S. Music emotion identification from lyrics [C]// Proceedings of the 11th IEEE International Symposium on Multimedia. Piscataway, NJ, USA: IEEE, 2009: 624-629.

[5] 夏云庆, 杨莹, 张鹏洲, 等. 基于情感向量空间模型的歌词情感分析 [J]. 中文信息学报, 2010, 24(1): 99-103.

XIA Yunqing, YANG Ying, ZHANG Pengzhou, et al. Lyric-based song sentiment analysis by sentiment vector space model [J]. Journal of Chinese Information Processing, 2010, 24(1): 99-103.

[6] 李静, 林鸿飞, 李瑞敏. 基于情感向量空间模型的歌曲情感标签预测模型 [J]. 中文信息学报, 2012, 26(6): 45-50.

LI Jing, LIN Hongfei, LI Ruimin, et al. Sentiment vector space model based musical emotion tag prediction [J]. Journal of Chinese Information Processing, 2012, 26(6): 45-50.

[7] SU Xiaoyuan, KHOSHGOFTAAR T M. A survey of collaborative filtering techniques [J]. Advances in Artificial Intelligence, 2009, 2009: 1-19.

[8] GOVINDARAJULU Z. Rank correlation methods [J]. Technometrics, 1992, 34(1): 108.

[9] HUQ A, BELLO J P, ROWE R, et al. Automated music emotion recognition: a systematic evaluation [J]. Journal of New Music Research, 2010, 39(3): 227-244.

(编辑 武红江)