

文章编号:2096-1383(2017)01-0076-04

# 基于双向情感分析的实时性 音乐推荐系统设计

毋亚男,刘德然,许小可

(大连民族大学,信息与通信工程学院,辽宁 大连 116605)

**摘要:**目前的音乐推荐系统,一般采用基于个体兴趣的推荐方法,这种方式虽然能满足大部分情景下的用户需求,但无法感知到用户实时性的心情变化。考虑到不同情绪状态下用户对于音乐的需求往往也会发生改变,提出一种基于双向情感分析的算法并构建了实际系统,实时分析用户的情感需求来进行音乐推荐。一方面基于音乐在频域的梅尔倒谱系数构建特征分类器完成歌曲的情感分类;另一方面通过获取用户在社交网络中的实时文本信息,基于自然语言理解分析出用户当前的情感需求,最终为用户产生音乐推荐列表,实现基于情境感知的实时音乐推荐。实验表明,使用该个性化推荐算法具有更高的准确性,用户群体可以获得更为满意的用户体验。

**关键词:**情感分析;推荐系统;贝叶斯分类器

**中图分类号:**TP399      **文献标志码:**A

## Designing of Real-time Music Recommendation System Based on Two-way Emotion Analysis

WU Ya-nan, LIU De-ran, XU Xiao-ke

(School of Information and Communication Engineering, Dalian Minzu University,  
Dalian Liaoning 116605, China)

**Abstract:** The current music recommendation system generally adopts the recommendation method on the basis of individual interest. Such method can meet users' needs under most situations, but it fails to perceive the real-time mood changes of users. In view of the users' demands for music under different emotional state always changing, the paper puts forward a two-way emotion analysis algorithm and constructs a real system to recommend music through real-time analysis of users' emotion needs. On the one hand, based on MEL cepstrum coefficient of the music in the frequency domain, a feature classifier is built to complete the music emotion classification. On the other hand, by capturing the users' real-time text social information in social networks, the emotional needs of the users are analyzed all the time based on natural language understanding. Finally, a music recommendation list is produced to the user for a specific situation, achieving the real-time music recommendation based on situational awareness. The experiment results show that the user groups who adopt such personalized recommendation algorithm can obtain more satisfying user experience.

**Key words:** sentiment analysis; recommendation system; naive Bayes classification

随着移动互联网的高速发展,在线数字音乐已 是人们上网消遣、日常休闲等场景中非常流行的娱

收稿日期:2016-09-14;最后修回日期:2016-12-13

基金项目:国家自然科学基金项目(61374170);中央高校基本科研业务费专项资金资助项目(DC201502060201)。

作者简介:毋亚男(1990-),女,辽宁大连人,大连民族大学信息与通信工程学院硕士研究生,主要从事数据挖掘和社交网络分析研究。

乐方式,但海量的音乐资源往往使得用户受信息过载的困扰,很难在众多曲目中找到要听的旋律。针对大规模的音乐数据,个性化推荐系统应运而生<sup>[1-3]</sup>。目前常用的推荐方式有基于歌曲相似性的推荐,基于用户相似性的推荐,人工推荐,朋友推荐等诸多方式,这些推荐方式虽然在一定程度上解决了音乐和用户需求相匹配的问题,但没有考虑到用户时刻变化的心情以及所处的情景。

本文以用户体验满意度为主要指标,一方面使用基于梅尔频率倒谱系数对歌曲情感的倾向性<sup>[4-5]</sup>进行分类;另一方面实时爬取用户在社交网络和社交媒体上的留言信息,通过自然语言理解<sup>[6]</sup>的方法来感知用户心情的实时变化;然后采用贝叶斯分类器将两部分情感分析的结果进行整合和匹配,完成时刻契合用户心情的音乐推荐。使用该个性化推荐算法的群体可以获得更为满意的用户体验。

### 1 实时音乐推荐系统总体框架

实时音乐系统可以分为用户音乐情感分类、用户社交网络情感分析与推荐系统界面三个模块,如图 1。其中,音乐情感和用户情感均分为四个维度,分别是喜悦、励志、躁动和悲伤。具体过程:首先获得用户的音乐曲目列表,计算所有歌曲在频域的梅尔倒谱系数,完成曲目的情感分类。这部分的计算是离线的,非实时感知的。此外,还需要使用网络爬虫实时抓取用户在社交网络或社交媒体上的文本信息,使用自然语言处理模块分析出用户每时每刻的情感,要求此部分必须是在线的,实时感知的。将用户的实时心情和对应的曲目相匹配,推荐出满足当前心情的歌曲,再整合到推荐系统界面,最终进行可视化展示和音乐播放<sup>[7-8]</sup>。

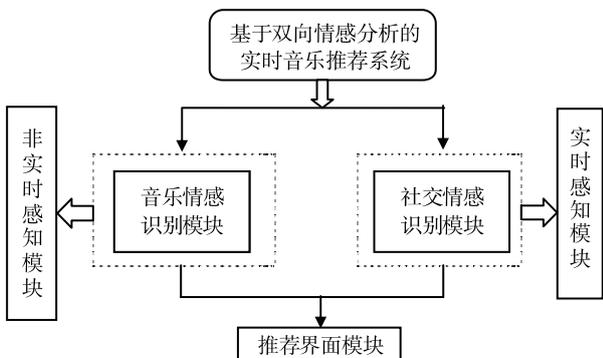


图 1 总体设计图

### 2 音乐情感分类

首先对每首歌曲的音乐数据进行预处理,分别使用歌曲的频域特征和梅尔倒谱特征来设计分

类器,测试和评估分类效果。在对基于不同音频特征构建的分类器结果进行比较后,选取梅尔频率倒谱系数这一准确率较高的特征作为最终分类的依据,完成音乐情感分类。

#### 2.1 实现过程

(1)音乐数据预处理。由于目前在音乐情感识别领域未有公开的音乐数据库,故本文用于测试和训练的音乐数据库为项目成员收集并通过多次筛选的 1 600 首中文歌曲,采用 SoX 音频处理软件将所有 MP3 类型歌曲统一转化为单声道,音轨为 22 050 Hz 的 WAV 文件。

(2)人工标定音乐情感分类。选取 20 位本科在读学生作为测评人员,利用人工判读的方式,针对处理后的音频文件按照励志、喜悦、悲伤、躁动四个情感维度对歌曲进行分类。由于音乐情感划分受主观性影响较大,故采用多人对同一音乐片段进行标定,取人数最多的方式完成歌曲的情感分类。

(3)划分训练集和测试集。将预处理后的 WAV 文件按照 7:3 分成两个子集,前者作为训练集,用于构建分类器,后者作为测试集,用来评估分类器的性能。

(4)音乐特征提取。根据分析的角度不同,可以对音乐信号进行时域分析、频域分析等。时域特征提取方法计算量较小,提取相对较容易,但不足之处是鲁棒性差。时域特征参数包括音乐的短时能量,短时自相关函数,短时过零率等。音乐信号的功率谱可从多方面反映音乐的情感特性,故本文主要从频域提取音乐特征,分别提取出歌曲的频谱特征和梅尔倒谱特征。

(5)训练分类器。考虑到本项目是个多分类问题,采用逻辑回归分类器进行音乐分类。

(6)测试并评估分类结果。将测试集作为输入,运行分类器,得到最终的分类结果并进行评价。

#### 2.2 音频信号特征提取过程

从频域特征参数的分析角度,分别提取每首歌曲的频谱特征和梅尔倒谱特征<sup>[9]</sup>。对于音频信号等数字信号的处理过程,离散傅里叶变换是常用的变换方法。使用快速傅里叶变换可分析信号的频谱、计数滤波器的频率响应,以及实现信号通过线性系统的卷积运算等,因而在信号的频谱分析方面有广泛的应用。本文采用快速傅里叶变换来计算每首歌曲的频谱特征。

梅尔频率倒谱系数 (Mel - scale Frequency Cepstral Coefficients,简称 MFCC)是标度频率域提

取出的倒谱参数,描述了音频信号在频域上的非线性特性。原始的频谱由两部分组成:包络和频谱的细节。如果从低频到高频按临界带宽的大小依次设置一组带通滤波器,对输入信号进行滤波,再将每个带通滤波器输出的信号能量作为信号的基本特征,就可以提取出梅尔频率倒谱系数。由于该特征不依赖于信号的性质,因此具有良好的识别性和抗噪能力,更符合人耳的听觉特性。MFCC 参数提取基本流程如下:

(1)将连续音乐信号  $s(n)$  经过预加重,分帧及加窗得到一帧音乐信号  $x(n)$ 。

(2)将  $x(n)$  信号进行快速傅里叶变换(FFT)得到频域信号,即线性频谱  $X(k)$ ,

$$X(k) = \sum_{n=0}^{N-1} x(n) W^{nk}, k = 0, 1, 2, \dots, N-1. \quad (1)$$

式中,  $x(n)$  为输入语音信号,  $N$  为傅里叶变换的点数。

(3)将线性频谱  $X(k)$  输入三角滤波器组  $H_m(k)$ ,  $0 \leq m \leq M$ ,  $M$  为滤波器组个数。每个滤波器的中心频率为  $f(m)$ 。三角滤波器组的中心频率在梅尔频率域内的间隔是相等的。由此,可以计算两个相邻的三角带通滤波器的中心频率距离  $\Delta Mel = \frac{f_{\max}}{k+1}$ 。每个带通滤波器组传递参数为

$$H_m(k) = \begin{cases} 0, & k < f(m-1) \\ \frac{2(k-f(m-1))}{(f(m+1)-f(m-1))(f(m)-f(m-1))}, & f(m-1) \leq k \leq f(m) \\ \frac{2(f(m+1)-k)}{(f(m+1)-f(m-1))(f(m)-f(m-1))}, & f(m) \leq k \leq f(m+1) \\ 0, & k \geq f(m+1) \end{cases}. \quad (2)$$

(4)Mel 坐标上的能量谱  $s(m)$  经滤波器组输出为:

$$s(m) = \ln\left(\sum_{k=0}^{N-1} |X_u(k)|^2 |H_m(k)|^2\right), 0 \leq m \leq M. \quad (3)$$

(5)经离散余弦反变换,得到频率域上的 Mel 频率倒谱  $C(n)$ :

$$C(n) = \sum_{k=0}^{N-1} s(m) \cos\left(\frac{\pi n(m-0.5)}{M}\right), n = 1, 2, \dots, L. \quad (4)$$

式中,  $L$  指 MFCC 系数的阶数。由于前若干维及后若干维的 MFCC 系数对语音的区分性能较大,所以本文采用前 12 维 MFCC 系数和前 12 维一阶差分系数作为特征参数。

### 2.3 音乐情感分类结果的评估与比较

在得到每首歌曲的频谱特征和梅尔倒谱特征以后,就可以分别使用这两种特征来训练分类器,

使用测试集数据评估分类结果。为更直观比较这两种特征分类效果的优劣,本文通过混淆矩阵对分类结果进行可视化。混淆矩阵是用于比较分类结果准确性的一种可视化工具,通过比较每个实测单元与预测分类图像的相应位置,对应单元模块的颜色浓度直观的显示分类结果。本文中混淆矩阵中横轴方向代表预测分类结果,纵轴方向代表实际类别结果。对于完美的分类器,从左上角到右下角都是深色的方格,而剩下的区域都是浅色,甚至无色。因此,可以通过观察混淆矩阵中主对角线区域内方格颜色的深浅,直观快速地评估分类效果。

基于频谱特征和梅尔倒谱特征构建分类器的分类效果比较如图 2。图中横轴表示歌曲预测分类,纵轴代表经主观评定后歌曲的真实情感分类。左图为基于频谱特征构建的分类器,从它的混淆矩阵中可看出关于喜悦、躁动的格子颜色较深,但励志、悲伤的方格颜色较浅,而且非对角区域颜色并不都是如期望的均为白色,说明错误分类的数量较多,整体分类效果一般。相对于左图,右图中基于梅尔频率倒谱系数构建的混淆矩阵中对角线区域格子的颜色更深,其中关于躁动这一类别的分类优势更为明显,非对角区域颜色较左图更浅,表明误分类情况出现较少,整体分类效果较左图获得了明显提升。综上所述,基于梅尔倒谱特征构建的分类器在音乐情感分类中取得了更好的分类效果,本系统采取该方法对每首歌曲进行情感分类。

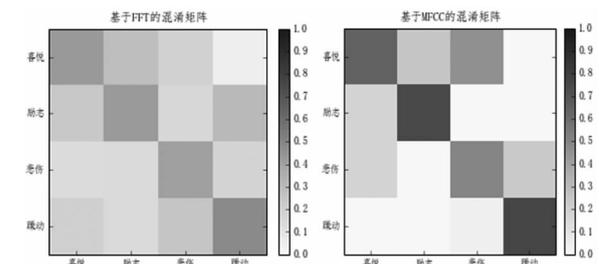


图2 基于频谱特征和梅尔倒谱特征构建分类器的分类效果比较

## 3 用户实时心情检测

在对歌曲进行分类后,下一步就是要实时检测出用户的心情变化。本文首先对实现过程进行阐述,然后使用朴素贝叶斯分类器<sup>[10]</sup>对文本情感分类结果进行评估。

### 3.1 情感分析实现过程

(1)用户情感文本数据的收集和预处理。本文利用新浪微博公开的 API,采集热门话题下的微博数据共计 7 500 条,筛选出具有情感色彩的

2 500条微博。

(2)人工标定文本情感分类。选取25名本科在读学生作为测评人员,与上述音乐分类情况相同,通过人工判读的方式,对所收集的微博文本按照励志、喜悦、悲伤、躁动四个情感维度的特征属性对文本进行分类。

(3)构建训练和测试语料库。首先,从2 500条微博文本信息中去除表情符号,保留中文部分并利用中文分词工具(“结巴”分词)进行分词处理,构成文本情感集。在基本情感字典基础上加入“杯具”、“给力”等带有感情色彩的网络词汇构建适用于微博的情感词典。其次,利用已构建的情感字典,对分词所得的稀疏矩阵进行降维,去掉停用词,提取能显著表征文本情感类别的词汇,进行情感特征选择。最后,进行情感特征权值的计算。情感文本集经过处理后得到的向量矩阵,按照7:3划分为训练语料库和测试语料库,利用训练语料库进行分类器训练。

(4)训练分类器。将训练语料库作为输入,对基于朴素贝叶斯构建的分类器进行训练,通过计算  $P(C_i) = \frac{s_{ik}}{s}$  (预先设定的喜悦、躁动、励志、悲伤四个分类属性在各训练样本中的出现概率)和  $P(X|C_i)$  (每个特征属性对不同分类的条件概率)生成分类器。

(5)测试并评估分类结果。利用经过训练阶段的分类器对测试语料库进行分类测试。输入是待分类的测试语料库,输出是待分类项与类别的映射关系。具体的分类过程如下:获取训练语料库各条文本数据的向量矩阵  $d = \{w_1, w_2, \dots, w_n\}$ , 用一个  $n$  维特征向量  $X = \{x_1, x_2, \dots, x_n\}$  表示,其中  $x_k$  为特征属性  $A_k$  的值,所有的样本分为  $m$  类:  $C = \{C_1, C_2, \dots, C_m\}$ 。本文中  $m$  值取为4,  $C_1, C_2, C_3, C_4$  分别代表喜悦、悲伤、励志和躁动四个类别。将一个未分类的文本数据分配给类别,当且仅当  $P(C_i|X) > P(C_j|X)$  时,将其归为  $C_i$  类。

### 3.2 文本情感分类结果的评估

通过对文本分析分类性能进行评估,常用的评价方法包括准确率、召回率、F1值等。本文采用前两种方法进行评估。精确率(Precision)是指在所有被判断为正确的文档中,正确数量所占比例。而召回率(Recall Rate)是指在所有事实上正确的文档中,有多大比例被判为正确。将测试语料进行分类测试,实验结果见表1。从分类结果

可看出,贝叶斯分类器在文本情感分析中取得了较好的分类效果。最后,将歌曲情感分类和文本情感分类进行匹配,完成音乐的实时性推荐。

表1 基于朴素贝叶斯构建的分类器分类结果

评测指标	喜悦	励志	悲伤	躁动
准确率	85.03	78.69	73.23	86.37
召回率	73.22	70.24	69.37	75.03

## 4 结语

本文提出了一种双向的情感分析算法并构建了实际系统,重点在于对音乐情感属性划分以及感知用户的社交网络情境。实验结果表明,使用该个性化推荐算法的用户群体可以获得更为满意的用户体验,推荐效果较好。今后将融合曲目的歌词信息用于提高歌曲情感分析的准确性,采用Adaboost等集成分类器来提高实时情感分析的分类准确性,从整体上进一步提升推荐质量。

### 参考文献:

- [1] 陈雅茜. 音乐推荐系统及相关技术研究[J]. 计算机工程与应用,2012(18):9-16.
- [2] 李瑞敏,闫俊,林鸿飞. 基于音乐基因组的个性化移动音乐推荐系统[J]. 计算机应用与软件,2012(9):27-30.
- [3] 谭学清,何珊. 音乐个性化推荐系统研究综述[J]. 现代图书情报技术,2014(9):22-32.
- [4] 李瑞敏,林鸿飞,闫俊. 基于用户-标签-项目语义挖掘的个性化音乐推荐[J]. 计算机研究与发展,2014(10):2270-2276.
- [5] 牛滨,孔令志,罗森林,等. 基于MFCC和GMM的个性音乐推荐模型[J]. 北京理工大学学报,2009(4):351-355.
- [6] 孙艳,周学广,付伟. 基于主题情感混合模型的无监督文本情感分析[J]. 北京大学学报(自然科学版),2013(1):102-108.
- [7] 伯德,克莱恩,洛佩尔. Python自然语言处理[M]. 北京:人民邮电出版社,2014:266-271.
- [8] 查美丽. 基于情感的音乐分类系统的研究与实现[D]. 南京:南京邮电大学,2014.
- [9] 里彻特,柯埃略. 机器学习系统设计[M]. 刘峰,译. 北京:人民邮电出版社,2014:139-153.
- [10] 杨鼎,阳爱民. 一种基于情感词典和朴素贝叶斯的中文文本情感分类方法[J]. 计算机应用研究,2010(10):3737-3739.

(责任编辑 王楠楠)