

面向中文歌词的音乐情感分类方法^①



王 洁, 朱贝贝

(北京工业大学 信息学部, 北京 100124)

通讯作者: 朱贝贝, E-mail: zbb@emails.bjut.edu.cn

摘 要: 情感是音乐最重要的语义信息, 音乐情感分类广泛应用于音乐检索, 音乐推荐和音乐治疗等领域. 传统的音乐情感分类大都是基于音频的, 但基于现在的技术水平, 很难从音频中提取出语义相关的音频特征. 歌词文本中蕴含着一些情感信息, 结合歌词进行音乐情感分类可以进一步提高分类性能. 本文将面向中文歌词进行研究, 构建一部合理的音乐情感词典是歌词情感分析的前提和基础, 因此基于 Word2Vec 构建音乐领域的中文情感词典, 并基于情感词加权和词性进行中文音乐情感分析. 本文首先以 VA 情感模型为基础构建情感词表, 采用 Word2Vec 中词语相似度计算的思想扩展情感词表, 构建中文音乐情感词典, 词典中包含每个词的情感类别和情感权值. 然后, 依照该词典获取情感词权值, 构建基于 TF-IDF (Term Frequency-Inverse Document Frequency) 和词性的歌词文本的特征向量, 最终实现音乐情感分类. 实验结果表明所构建的音乐情感词典更适用于音乐领域, 同时在构造特征向量时考虑词性的影响也可以提高准确率.

关键词: 音乐情感分类; 中文情感词典; 情感词加权; TF-IDF; 词性

引用格式: 王洁, 朱贝贝. 面向中文歌词的音乐情感分类方法. 计算机系统应用, 2019, 28(8): 24-29. <http://www.c-s-a.org.cn/1003-3254/6959.html>

Classification of Musical Emotions Oriented to Chinese Lyrics

WANG Jie, ZHU Bei-Bei

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: Emotion is the most important semantic information of music. Music emotional classification is widely used in music retrieval, music recommendation, and music therapy. Traditional music emotional classification is mostly based on audio. Nevertheless, based on the current technology level, it is difficult to extract semantic-related audio features from audio. There is some emotional information in the Lyric text, and music emotional classification is carried out with lyrics. This study focuses on Chinese lyrics and constructs a reasonable music emotion dictionary, which is the premise and foundation of lyric emotion analysis. Therefore, a Chinese emotion dictionary in music field is constructed based on Word2Vec, and Chinese music emotion analysis is carried out based on the weighting of emotional words and the part of speech. Firstly, this study constructs the emotional lyrics table based on the VA emotional model and adopts Word2Vec. The idea of word similarity calculation in Word2Vec extends the emotional vocabulary and constructs a Chinese music emotional dictionary, which contains the emotional categories and emotional weights of each word. Then, according to the dictionary, emotional words weights are obtained, and feature vectors of lyric texts based on TF-IDF (Term Frequency-Inverse Document Frequency) and lexical features are constructed. Finally, music emotional classification is realized. The constructed music emotion dictionary is more suitable for music field, and the accuracy can be improved by considering the influence of part of speech when constructing feature vectors.

Key words: music emotional classification; Chinese emotional dictionary; emotional word weighting; TF-IDF; part of speech

① 基金项目: 国家自然科学基金 (61876010)

Foundation item: National Natural Science Foundation of China (61876010)

收稿时间: 2018-12-05; 修改时间: 2018-12-25, 2019-01-11; 采用时间: 2019-01-28; csa 在线出版时间: 2019-08-08

1 引言

随着计算机网络与多媒体技术的快速发展,互联网上涌现了越来越多的文本、图像、音频和视频等多媒体数据.音乐是多媒体数据的重要组成部分.面对音乐作品数量的爆炸式增长,音乐种类的不断增多,如何有效的管理音乐成为一个值得关注的问题,对音乐资源进行合理的分类是一个有效的解决办法.音乐是情感的载体,情感是音乐最重要的语义信息,音乐的情感分析广泛应用于音乐检索、音乐推荐和音乐治疗等领域^[1].音乐情感的自动识别属于音乐心理学和计算机科学的交叉领域^[2],至今已有了十几年的历史,国内外众多学者都对该领域做了深入的研究,并取得了一定成果.2010年, Kim 等曾对当时音乐情感识别研究的进展做了全面综述^[3].2012年, Yang 等对音频音乐情感识别研究做了详尽的总结^[4].2017年,陈晓鸥等人回顾了最近几年音频音乐情感识别的研究进展情况,提出了若干问题及可能的解决方案^[5].

音乐主要由音频信号与歌词文本这两种模态的数据组成.目前大多数研究者致力于使用机器学习方法研究音频信息对音乐情感分类的作用^[6-8].2003年, Li 等提取音色、节奏和音阶等音频特征,并最早使用 SVM (Support Vector Machine) 对音乐情感分类^[9].2014年, Weninger 等提取 MFCC (Mel Frequency Cestrum Coefficient) 等底层音频特征后,使用 LSTM (Long Short-Term Memory) 对音乐情感分类^[10].2017年, Jakubik 等使用 GRU (Gated Recurrent Unit) 对音频进行特征学习,并引入语义嵌入的思想^[11].同年,邓永莉等提出基于中高层特征的音乐情感识别模型^[12].

采用现有频域和时域特征的机器学习方法,很难使音乐情感识别的性能再提高^[5].一些心理学研究表明歌词文本中确实蕴含着一些特有的语义信息,包括情感信息^[13].结合歌词进行音乐情感分类可以进一步提高分类性能,基于歌词的情感分类也逐渐成为热门的研究方向.歌词本质上属于文本的范畴,词语的情感判别是歌词情感分析的基础.构建一部合理的音乐情感词典,实现词汇的情感分析是歌词情感分析的前提和基础.国内情感词典构建起步较晚,情感词典的领域特性也日趋明显.已有研究构建的情感词典主要集中在中文评论领域,如微博评论和商品评论,评论情感分析一般只判别情感极性,即褒贬性,而歌词具有更加丰富的情感,目前还没有被广泛认可的音乐领域的中文情

感词典.2010年,夏云庆等基于文本向量空间模型提出了情感向量空间模型,并利用实验验证了情感向量模型在解决稀疏性、特征表示能力、表示效率和特征歧义消除等方面具有较明显的优势^[14].2014年,蒋盛益等利用 HowNet 中语义相似度计算的思想,构建音乐领域的中文情感词典,再进行音乐情感分类^[15].2015年, Furuya 等通过对非情感词加权构造情感向量,使用聚类方法进行音乐情感分类^[16].2017年,黄仁等基于 Word2Vec 建立情感词典,用构建的情感词典对互联网商品评论进行情感分类^[17],该词典仅包含词的情感极性,不包含词的情感强度,没有考虑情感词本身的分类影响程度.在基于歌词文本的情感向量的构建过程中,夏云庆等仅统计每个情感类别的情感词个数,忽略了情感词的情感强度是不同的.蒋盛益和 Furuya 等均未考虑实词词频和词性的影响作用.

针对以上问题,本文基于 Word2Vec 构建音乐领域的中文情感词典,并基于情感词加权和词性进行中文音乐情感分析.本文首先以 VA 情感模型为基础构建情感词表,采用 Word2Vec 中词语相似度计算的思想扩展情感词表,构建中文音乐情感词典,词典中包含每个词的情感类别和情感权值.然后,依照该词典获取情感词权值,基于 TF-IDF 构造特征向量,并进一步考虑词性对情感分类的影响,最终实现音乐情感分类.

2 基于歌词的音乐情感分类

2.1 构建情感词表

音乐是情感的载体,情感是音乐最重要的语义信息,不同的音乐必然与不同的情感相联系.为了更准确的描述音乐情感,与人的情感体验一致,需要选择合适的音乐情感模型.

选择 VA 模型作为音乐情感分类的依据.VA 模型(也称为环形情感模型)是由 Russel 提出的,是广泛采用的通用连续维度模型^[18,19].该模型认为情感状态是分布在一个包含效价度(valence)和激活度(arousal)的二维环形空间上的点,如图1所示.其中横轴表示效价度,纵轴表示激活度,圆心代表中性的效价度和中等水平的激活度.

为了适应中文情感分类,结合 VA 模型的情感类别以及音乐表达情感的特点构建情感词表.情感词表如表1所示,第1列为情感类别,即本文用于情感分类的标签.其它列为对应每一个情感类别的情感词汇.

2.2 中文音乐情感词典

情感词表中的词是经过提炼而成的情感词,而歌词中包含的情感词是有限的,因此,需要对情感词表进行扩展以满足歌词情感分类的需求.借助 Word2Vec 计算词语相似度,使用《哈工大同义词林》和歌词语料库对情感词表的 38 个情感词汇进行扩展,从而构建出一部适用于中文歌词情感分类的词典,用于歌词文本分析.

2.2.1 计算词语相似度

情感词一般是指能表达情感的形容词,如“喜悦”、“悲伤”、“愤怒”.歌词中包含的情感词较少,而歌词中的某些非情感词也能表达情感.比如,“夜”可以表达“静”的情感,“告别”可以表达“悲伤”的情感.因此,可以通过计算有情感倾向的实词(包括名词、动词、形

容词和副词)与情感词表中的情感词汇的词语相似度,对实词进行情感分类,从而实现情感词表的扩展.

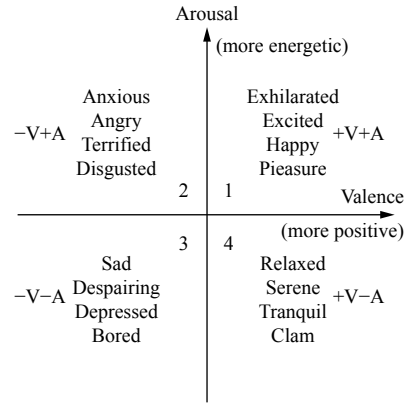


图1 VA 情感模型

表1 情感词表

情感类别	词 1	词 2	词 3	词 4	词 5	词 6	词 7	词 8	词 9	词 10
+V+A	快乐	明亮	激动	热情	胜利	生机	神圣	轻盈	活泼	严肃
-V+A	愤怒	生气	恼火	恼怒	悲愤	愤恨	痛恨	仇恨	敌视	
-V-A	悲伤	哀愁	困惑	黯淡	沉重	压抑	忧郁	忧伤	屈从	
+V-A	抒情	从容	平静	满足	安谧	抚慰	向往	多情	柔和	梦幻

采用 Gensim 的 Word2Vec 进行相似度计算,并用搜狗实验室的语料数据训练模型.因为该语料库规模较大,训练好的模型可以涵盖尽可能多的词语.因此,可以利用训练好的模型计算任意两个词语的相似度,相似度值在 0-1 之间.

2.2.2 构建音乐情感词典

首先使用《哈工大同义词林》对情感词表进行扩展,构建基础情感词典.并结合歌词语料库对基础情感词典做进一步扩展,构建音乐情感词典.

使用《哈工大同义词林》对情感词表进行扩展,扩展后的词典仅包含情感词,称为基础情感词典.首先,从《哈工大同义词林》获取情感模型中的情感词对应的同义词列表.然后,计算同义词列表中的实词与情感词的词语相似度,选择相似度值最大的情感词与实词对应.最后,提取相似度高于 0 的实词构成基础情感词典.实现过程如图 2 所示.情感词典的存储结构为四元组(实词,情感词,情感类别,情感权值),比如(伤感,忧伤, -V-A, 0.91).其中,“伤感”是选自《哈工大同义词林》的实词;“忧伤”是选自情感词表的情感词;“-V-A”是“忧伤”对应的情感类别;0.91 对应于情感权值,即“忧伤”与“伤感”的词语相似度.基础情感词典共包含

1527 个词汇,其中 4 类情感的词汇量分布情况如表 2 所示.

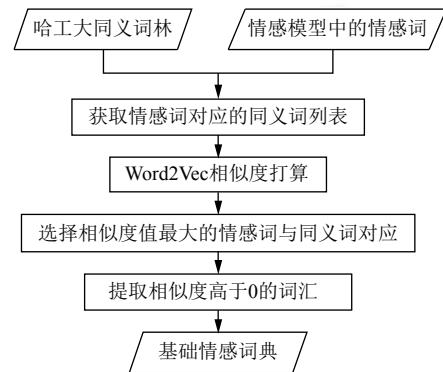


图2 基础情感词典构建流程

表2 基础情感词典词汇量分布

情感类别	+V+A	-V+A	-V-A	+V-A
词汇数量	807	269	183	268

基础情感词典仅包含情感词,而歌词中所包含的情感词较少,但歌词中的某些非情感词也有情感倾向.因此可以利用包含一万首中文歌曲的歌词语料库对基础情感词典做进一步扩展,以更适用于中文歌词情感

分类, 扩展后的词典称为音乐情感词典. 首先, 将歌词语料库中的歌词进行处理, 并提取实词构建歌词词汇语料库. 然后, 计算歌词词汇语料库中的实词与情感词的词语相似度, 选择相似度值最大的情感词与实词对应. 最后, 提取相似度高于 0 的实词与基础情感词典合并, 构成音乐情感词典. 实现过程如图 3 所示. 音乐情感词典的存储结构与基础情感词典相同. 音乐情感词典共包含 45 374 个词汇, 其中 4 类情感的词汇量分布情况如表 3 所示.

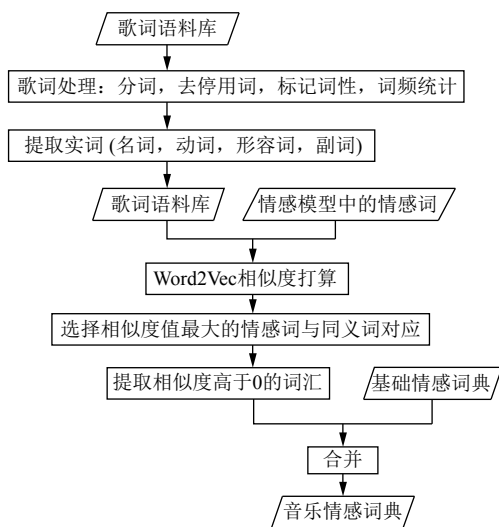


图3 音乐情感词典构建流程

表3 音乐情感词典词汇量分布

情感类别	+V+A	-V+A	-V-A	+V-A
词汇数量	22 607	4125	8342	10 300

2.3 构建基于歌词的特征向量

本文依照所构建的情感词典获取情感词权值, 基于 TF-IDF 构造特征向量, 并进一步考虑词性对情感分类的影响, 将特征向量扩展到 16 个维度.

2.3.1 TF-IDF

基于情感向量模型, 采用情感词表的 4 类情感类别作为音乐的情感特征向量, 共 4 个维度. 该特征向量的每个维度表示音乐与每类情感的相似关系. 将歌词中的实词与情感词典进行匹配, 可以得到每个实词的情感类别和情感权值, 从而计算出每个情感类别的统计值.

采用 TF-IDF 规则计算情感特征. TF-IDF 是一种统计方法, TF 表示词频, 可以评估歌词中某个实词在

特定歌词文件中的情感重要程度. IDF 表示逆向文件频率, 可以评估某个实词对于区分特定歌词文件和其他歌词文件的情感重要性.

$$tf(t_i) = \frac{s_i \times n_{i,j}}{\sum_{k=1}^N (s_k \times n_{k,j})} \quad (1)$$

对于歌词文本中的实词 t_i , $tf(t_i)$ 表示 t_i 对情感类别的重要程度, 反映出一个实词在特定歌词文件中的局部统计特征. s_i 表示 t_i 在情感词典中对应实词的情感权值, 如果 t_i 匹配失败, s_i 值为 0. $n_{i,j}$ 表示 t_i 在该歌词文本中的出现次数. N 表示该篇歌词经过分词后得到的词语总数.

IDF 由式 (2) 计算:

$$idf(t_i) = \frac{|D|}{|\{doc_j : t_i \in doc_j\}|} \quad (2)$$

对于歌词文本中的实词 t_i , $idf(t_i)$ 表示 t_i 对于区分特定歌词文件和其他歌词文件的情感重要性. $|D|$ 表示歌词文档总数, $|\{doc_j : t_i \in doc_j\}|$ 表示歌词中出现 t_i 的歌词文档数.

歌词文本的特征向量表示为:

$$f_{lyric} = (f_1, \dots, f_c)^T \quad (3)$$

其中, c 表示情感类别数, 特征向量的每个维度由式 (4) 计算:

$$f_k = \sum_{\{i|t_i \in lyric, t_i \in e_k\}} tf(t_i) \times idf(t_i) \quad (4)$$

其中, e_k 表示特征向量第 k 维度对应的情感类别, $lyric$ 表示歌词经过分词后得到的词语.

2.3.2 词性

不同的词性表达情感的能力是不同的, 比如, 形容词比名词表达的情感更丰富. 因此, 本文将结合词性信息构造特征向量. 我们将每篇歌词文本经过分词后得到的词语根据词性分为 4 类, 每类词性对应的特征向量如式 (5) 所示. 将 4 类词性的特征向量结合得到最终的特征向量, 如式 (6) 所示.

$$f_{lyric, POS} = (f_1, POS, \dots, f_c, POS)^T \quad (5)$$

$$f_{lyric} = (f_1, noun, \dots, f_c, noun, \dots, f_1, adv, \dots, f_c, adv)^T \quad (6)$$

2.4 歌词情感分类

音乐情感分类通常采用标准的监督学习模型, 如 K 近邻 (KNN)、高斯混合模型 (GMM) 和支持向量机

(SVM). 本文采用 KNN 分类算法用于歌词情感分类.

KNN 的核心思想是如果一个样本在特征空间中的 k 个最相邻的样本中的大多数属于某一个类别, 则该样本也属于这个类别, 其中 k 通常是不大于 20 的整数.

选择中文歌词数据集中的 80% 作为训练集, 20% 作为测试集. 利用 KNN 分类的思想, 对每个测试样例, 计算它与所有训练样例之间的情感特征向量的欧几里德距离, 以确定其最近邻列表, 选择最近邻中出现次数最多一个类别作为测试样例的类别. 在本文实验中, 当 $k=5$ 时分类效果最佳.

3 实验结果及分析

3.1 数据集

数据集是 200 篇情感类别鲜明的中文歌词文档, 每个情感类别约 50 篇. 每篇歌词的情感标签由多个人进行标记, 选择标记人数最多的类别作为该歌词的标签.

3.2 评价指标

本文使用准确率、精确率、召回率和 F 值来评价分类效果. 以情感类别“-V-A”为例, 正类 (Positive) 表示该歌词的预测结果是“-V-A”, 负类 (Negative) 表示该歌词的预测结果不是“-V-A”. TP , FP , TN 和 FN 的具体含义如表 4 所示. 则准确率 A 、精确率 P 、召回率 R 、 F 值的计算公式如下:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

$$F = \frac{2 \times P \times R}{P + R} \quad (10)$$

表 4 TP , FP , TN 和 FN 含义表

	正类	负类
真	真正 (TP)	真负 (TN)
假	假正 (FP)	假负 (FN)

3.3 实验结果与评价

本文做了两组对比实验, 所有实验均使用相同的数据集. 第一组对比实验中, 情感词典的选择不同, 分别使用 2.2.2 节构造的基础情感词典和音乐情感词典,

构造特征向量均考虑词性的影响. 实验结果如表 5 所列. 可以看出, 基于音乐情感词典的中文歌词情感分类能够达到更好的效果, 所有情感类别的准确率以及 4 种情感类别的精确率、召回率和 F 值均高于基础情感词典. 同时, 实验结果表明, 基于歌词的音乐情感分类更适用于“+V+A”和“-V+A”这两类情感的识别.

第二组对比实验中, 特征向量的构造方式不同, 均使用 2.2.2 节构造的音乐情感词典. 实验结果如表 6 所列. 可以看出, 在构造特征向量时考虑词性的影响可以提高所有情感类别的准确率, “-V+A”和“-V-A”的评价指标值都有一定提高.

表 5 情感词典对比实验结果 (%)

情感类别	评价指标	基础情感词典	音乐情感词典
+V+A	P	56.0	69.6
	R	82.4	94.1
	F	66.7	80.0
-V+A	P	51.3	93.0
	R	45.7	90.4
	F	49.5	92.7
-V-A	P	50.0	75.0
	R	33.3	50.0
	F	40.0	60.0
+V-A	P	42.7	50.0
	R	13.9	16.7
	F	20.2	25.0
全部	A	53.3	70.0

表 6 特征向量对比实验结果 (%)

情感类别	评价指标	TF-IDF	TF-IDF + 词性
+V+A	P	58.3	69.6
	R	85.2	94.1
	F	70.4	80.0
-V+A	P	50.0	93.0
	R	48.3	90.4
	F	66.7	92.7
-V-A	P	66.7	75.0
	R	33.3	50.0
	F	44.4	60.0
+V-A	P	48.0	50.0
	R	13.2	16.7
	F	20.5	25.0
全部	A	66.7	70.0

4 结论与展望

本文基于 Word2Vec 构建音乐领域的中文情感词典, 并基于情感词加权和词性进行中文音乐情感分析. 首先以 VA 情感模型为基础构建情感词表, 采用 Word2Vec 中词语相似度计算的思想扩展情感词表, 构

建中文音乐情感词典,词典中包含每个词的情感类别和情感权值。然后,依照该词典获取情感词权值,基于TF-IDF构造特征向量,并进一步考虑词性对情感分类的影响,最终实现音乐情感分类。实验结果表明基于所构建的音乐情感词典进行中文歌词情感分类能够达到更好的效果,同时在构造特征向量时考虑词性的影响也可以提高准确率。但基于歌词的音乐情感分类更适用于“+V+A”和“-V+A”这两类情感的识别。今后将研究结合歌词与音频的多模态融合的音乐情感分类,以实现多个信息源互补,从而提高所有情感类别的分类精度。

参考文献

- 1 Furuya M, Huang HH, Kawagoe K. Evaluation of music classification method based on lyrics of English songs. Proceedings of 2015 International Multi Conference of Engineers and Computer Scientists. Hong Kong, China. 2015. 134–137.
- 2 Yang YH, Chen HH. Music Emotion Recognition. Boca Raton: CRC Press, 2011. 1–25.
- 3 Kim YE, Schmidt EM, Migneco R, *et al.* Music emotion recognition: A state of the art review. Proceedings of the 11th International Society for Music Information Retrieval Conference. Utrecht, Netherlands. 2010. 255–266.
- 4 Yang YH, Chen HH. Machine recognition of music emotion: A review. ACM Transactions on Intelligent Systems and Technology, 2012, 3(3): 40.
- 5 陈晓鸥, 杨德顺. 音乐情感识别研究进展. 复旦学报(自然科学版), 2017, 56(2): 135–148.
- 6 Lu L, Liu D, Zhang HJ. Automatic mood detection and tracking of music audio signals. IEEE Transactions on Audio, Speech, and Language Processing, 2006, 14(1): 5–18. [doi: 10.1109/TSA.2005.860344]
- 7 Fu ZY, Lu GJ, Ting KM, *et al.* A survey of audio-based music classification and annotation. IEEE Transactions on Multimedia, 2011, 13(2): 303–319. [doi: 10.1109/TMM.2010.2098858]
- 8 Ren JM, Wu MJ, Jang JSR. Automatic music mood classification based on timbre and modulation features. IEEE Transactions on Affective Computing, 2015, 6(3): 236–246. [doi: 10.1109/TAFFC.2015.2427836]
- 9 Li T, Ogihara M. Detecting emotion in music. Proceedings of the 4th International Conference on Music Information Retrieval. Baltimore, MD, USA. 2003. 239–240.
- 10 Weninger F, Eyben F, Schuller B. On-line continuous-time music mood regression with deep recurrent neural networks. Proceedings of 2014 IEEE International Conference on Acoustics, Speech and Signal Processing. Florence, Italy. 2014. 5412–5416.
- 11 Jakubik J, Kwaśnicka H. Music emotion analysis using semantic embedding recurrent neural networks. Proceedings of 2017 IEEE International Conference on Innovations in Intelligent Systems and Applications. Gdynia, Poland. 2017. 271–276.
- 12 邓永莉, 吕愿愿, 刘明亮, 等. 基于中高层特征的音乐情感识别模型. 计算机工程与设计, 2017, 38(4): 1029–1034.
- 13 Besson M, Faïta F, Peretz I, *et al.* Singing in the brain: Independence of lyrics and tunes. Psychological Science, 1998, 9(6): 494–498. [doi: 10.1111/1467-9280.00091]
- 14 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. 中文信息学报, 2010, 24(1): 99–103. [doi: 10.3969/j.issn.1003-0077.2010.01.018]
- 15 蒋盛益, 阳垚, 廖静欣. 中文音乐情感词典构建及情感分类方法研究. 计算机工程与应用, 2014, 50(24): 118–121. [doi: 10.3778/j.issn.1002-8331.1305-0147]
- 16 Furuya M, Oku K, Kawagoe K. Music feeling classification based on lyrics using weighting of non-emotional words. Proceedings of the 13th International Conference on Advances in Mobile Computing and Multimedia. Brussels, Belgium. 2015. 380–383.
- 17 黄仁, 张卫. 基于 Word2Vec 的互联网商品评论情感倾向研究. 计算机科学, 2016, 43(S1): 387–389.
- 18 Russell JA. A circumplex model of affect. Journal of Personality and Social Psychology, 1980, 39(6): 1161–1178. [doi: 10.1037/h0077714]
- 19 Posner J, Russell JA, Peterson BS. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and Psychopathology, 2005, 17(3): 715–734.