

# 基于模糊数学的音乐情感检索技术

高友平<sup>1</sup> 童名文<sup>1</sup> 张 凯<sup>2</sup> 野菊苹<sup>1</sup> 陈林林<sup>1</sup>

(华中师范大学信息技术系 武汉 430079)<sup>1</sup> (华中师范大学音乐学院 武汉 430079)<sup>2</sup>

**摘 要** 如何以音乐的情感为检索关键词来准确地匹配检索算法是目前研究的热点问题之一。在此,以音乐情感的模糊性特征为基础,采用模糊数学理论和 AV 情感向量空间模型,提出了一种音乐情感模糊检索技术:将音乐情感进行量化处理的同时,也体现一首音乐所包含的不同种类的情感;另外,同一情感类型的音乐能够按其隶属度值的高低进行智能排名,使人们能更准确地找到所需的音乐资源。通过对比实验,计算此检索方法的查全率、查准率和 F 值,相关度, C 值、V 值覆盖率和出新率等评价指标,从而证明了该检索算法的有效性。

**关键词** 模糊数学,音乐情感,AV 情感向量空间模型,查全率,查准率

**中图分类号** TP391.3 **文献标识码** A

## Research on Music Emotion Retrieval Technology Employing Fuzzy Mathematics

GAO You-ping<sup>1</sup> TONG Ming-wen<sup>1</sup> ZHANG Kai<sup>2</sup> YE Ju-ping<sup>1</sup> CHEN Lin-lin<sup>1</sup>

(Department of Information Technology, Central China Normal University, Wuhan 430079, China)<sup>1</sup>

(School of Music, Central China Normal University, Wuhan 430079, China)<sup>2</sup>

**Abstract** One of the hot issues of the present study is how to match retrieval algorithm accurately while regarding the music emotion as retrieval keywords. Based on the fuzzy characteristics of emotion of music, the theory of fuzzy mathematics and AV emotion of vector space model were used to propose a novel music emotional fuzzy retrieval technology (MEFRT). The most significant feature of this music emotion retrieval technology is music emotion to quantitative treatment. We can see a music contains different types of emotion from this method of retrieval. According to the membership value, the same emotional type of music can be smart ranking. People can more accurately find the music resources by this way. In order to prove the effectiveness of the MEFRT, a series of experiments were designed, and the metrics were calculated, such as recall ratio, precision, F value, relevance, C value, V value coverage and new rate, etc.

**Keywords** Fuzzy mathematics, Music emotion, AV emotion of vector space model, Recall ratio, Precision

近年来,数字化音乐资源迅速增长并在教学中得到了广泛应用。教师开始使用数字化音乐资源辅助教学,并取得了良好的学习效果。但是,随着数字化音乐资源的日益增加,如何有效地管理这些资源成为音乐辅助教学中亟需解决的问题。数字化音乐资源管理的一个重要功能是资源检索。国内外学者对音乐资源检索方式做了广泛研究,提出了基于内容的音乐检索——通过记录音乐的声学信号、节奏或旋律来实现基于内容的音乐检索<sup>[1]</sup>、基于哼唱的音乐检索和基于旋律的音乐检索等检索方式。在文献[2]中提出利用旋律识别技术来对比用户自己哼唱的歌曲的旋律和音乐资源系统里的相关资源的旋律,从而实现哼唱检索。张宝华和张品在文献[3]中提出使用基于时域自相关测量的音调追踪算法进行音乐片段的旋律提取,来实现给予旋律的音乐检索。近几年,研究人员提出基于情感的音乐检索。文献[4]中提出了以音乐的音频特征来区别音乐的不同情感,从而对音乐进行情感分类来实现给予情感的音乐检索。而基于音乐情感的检索方式的实

现大都是将音乐准确地定义在某一类情感基调中,如悲伤、欢喜、兴奋,如文献[5,6]提出的音乐情感分类方法都是以音乐的自身属性,如频率、音调等作为音乐分类的基本依据,同时将音乐片段分类到某一类情感领域中。音乐本身具有很强的感情色彩,因此基于情感的数字音乐检索方式符合音乐的本质特征,也是用户的需求所在。然而,情感不能精确测量,且具有很强的主观性。所以,如何量化情感是音乐情感检索技术的核心问题。模糊数学是人们日常生活中所接触到的一些模糊概念,如:高矮胖瘦、能量化、数学化<sup>[7]</sup>。目前,模糊数学的原理在事物分类和个体评价这两方面的应用是最广泛的。同时,模糊数学在财务管理、人工智能、医学诊疗、环境监测等领域都有应用。模糊数学之所以能够应用到与分类和程度判断有关的领域中去,是因为隶属度函数的结果能够表示出某个体属于某种类型的相对程度。这也体现了个体的多面性。如果在音乐情感检索中引入模糊数学这个概念,人们就能很客观地描述某首乐曲属于悲伤的程度或者属于高兴的程

到稿日期:2012-09-05 返修日期:2012-12-26 本文受基金教育部人文社科基金项目(10YJC88011),湖北省自然科学基金项目(2011CDB163),中央高校自主科研专项资金资助项目(ccnu10b01003)资助。

高友平(1987-),女,硕士生,主要研究方向为多媒体内容适配、教育资源管理与服务,E-mail:410677538@qq.com;童名文(1975-),男,博士,副教授;张 凯(1972-),女,副教授;野菊苹(1986-),女,硕士生;陈林林(1985-),男,硕士生。

度到底有多少。在此基础上,从音乐情感分类的分类模型中,找到了与模糊数学相契合的 AV 情感空间。进一步提出了相关的隶属度函数,并利用隶属函数将选取的音乐片段归到不同的情感类型中,来实现基于模糊数学的音乐情感检索。最终通过实验验证了这种检索方式的有效性。

## 1 音乐的情感特征和情感分类

### 1.1 音乐的情感特征

音乐除包含节奏、音调、拍子、歌词等与自身组成相关的因素外,还包括情感、视觉形象、想象空间等外延因素。从某种意义上来说,音乐的情感是客观的。因为音乐的情感所体现的主要是创作者创作时的一种主观的情感体验,所以对听众来说,这种情感体验已经被创作者定位了。不同的听众在不同的心理状态下,对同一首音乐的情感体验是不尽相同的,但这种不同不会与原创者的情感状态出现很大的差距。总之,音乐的情感基调具有客观性是我们实现模糊检索的一个大前提。除此之外,甘露认为音乐的情感还带有典型性、概括性和易感性<sup>[8]</sup>。

### 1.2 音乐的情感分类

在音乐情感的分类中,目前比较认可的情感分类模型是 Ressel 的 AV 情感空间<sup>[9]</sup>和 Thayer 改进的 AV 情感空间<sup>[10]</sup>。Ressel 的 AV 情感空间里的“A”是指 activation,心理学上叫做“活跃度”,用它来度量人们情绪是否高涨;“V”是指 valence,心理学上称为“诱发力”,用它来度量人们是否感到愉悦。Ressel 发现人们的情感偏好和 A 呈“倒 U 形”关系,V 和人们的情感偏好呈线性关系;同时,很多感情还可以被 Pleasantness 和 Activation 两维值组合而成<sup>[5]</sup>。Thayer 将音乐的情感分为 4 大类:joy(高兴)、sadness(伤感)、clam(平静)和 anxiety(焦虑)。在此使用 Thayer 的 AV 情感向量空间模型,并保留 Ressel 所描述的 A 和 V 与情感的函数关系及其组合特性。

## 2 情感模糊检索

### 2.1 情感模糊检索的基本思想

在此检索方法中,提供给用户的检索词是高兴、焦虑、悲伤和宁静。在模糊检索的过程中主要涉及模糊集表示、隶属度函数的确定、检索结果的排序和检索关键词的提取。设论域  $U = \{x_1, x_2, \dots, x_i\}$ ,  $x_i$  表示具体的某一首音乐。同时,将这些音乐分成 4 种不同的情感类型:高兴、焦虑、悲伤和宁静。这些类型在  $U$  上表现为 4 个模糊子集  $A_i \in (U) (i=1, 2, 3, 4)$ 。这 4 个模糊子集分别由 2 个主要的指标—A、V 来描述。由于音乐的情感本身就是一个比较复杂的模糊集合体,同时 A 的大小和 V 的大小也不是精确的,因此对每个模糊集合  $A_i$  而言,A 和 V 这两个指标也是模糊集,即  $A_i = (A_{i1}, A_{i2}) (i=1, 2)$ 。隶属度函数的设计则是根据 A 和 V 与情感 P 的关系来实现的。在模糊检索中,检索结果的排序是根据个体资源对每个模糊集的隶属度值的大小来确定的。对于同一类的模糊集,其资源的排序应当是按照隶属度值的降序排列。

### 2.2 情感模糊检索的流程

(1)确定检索的检索关键词为高兴、焦虑、悲伤和宁静。

(2)确定每个音乐资源对这 4 类情感的隶属度值。根据隶属度函数来求得个体对与每个模糊集的隶属度值。(3)确定构造检索系统的平台,包括所使用的页面语言、数据库语言。同时搭建检索系统。系统的搭建主要采用的是 MySQL 数据库语言和 jsp 及 java 语言。(4)测试检索系统的功能,整理检索结果,并对检索系统的性能进行分析。检索性能的分析主要从检索方法本身的角度和用户的角度这两个方面来进行评价。

### 2.3 隶属度函数设计

为了与模糊数学的处理方法更契合,可以将 AV 模型看成是一个二维平面坐标轴。从 David 和 Meyer 在 2009 年所做的一系列实验数据可以看出:横坐标 Valence 简称 V,变量名记为  $v$ ,从左到右的范围为  $[-0.33, 0.33]$ ;纵坐标 Activation 简称 A,变量名记为  $a$ ,从下到上的取值范围依次为  $[-0.26, 0.26]$ <sup>[12]</sup>。

#### 2.3.1 A 变量的隶属函数

当情感为“兴奋”等相当振奋的状态时,A 的取值接近 0.26。而当情感处于一般愉悦状态时,A 的值在 0.1 附近。当人们的情感偏好处于害怕、愤怒、紧张这些情绪中时,A 的取值在 0.23~0.16,或者在 0.03 附近。当处于悲伤时,A 的取值主要在  $-0.03 \sim -0.20$  这一区域内。当 A 值在  $-0.23$  至  $-0.16$  时,则可断定情感状态为“平静”。

首先,确定情感为“开心”时所对应的隶属函数,并将其隶属函数记为  $A_{11}$ 。从上面所述可以确定,“开心”这一模糊集所对应的论域为  $[-0.26, 0.26]$ 。从文献[12]的 table2 中发现,与“happy”意思相近的词汇几乎集中在某几段范围内,因此认定隶属函数  $A_{11}$  的整体结构应该是中间型。其次,确定隶属度为 1 和 0 的区域。隶属度为 1 的区域代表当某首音乐的 A 值落在这一区域时,这首音乐的情感属于“高兴”的隶属度是 1。同样,隶属度为 0 的区域表示当某首音乐的 A 值落在这一区域时,这首音乐隶属于“高兴”的程度为 0。因此,确定隶属度为 1 的论域为  $(0.13, 0.07)$ ;隶属度为 0 的论域为  $[-0.26, 0.02]$ 。然后确定过渡带。所谓过渡带,即指 A 值落到这一区域时,很难说这首音乐的情感到底是不是属于“高兴”。从文献[12]的 table2 中也可以看出, $A_{11}$  的过渡带论域为  $(0.02, 0.07) \cup (0.13, 0.26)$ 。由于 A 和人们的情感偏好是呈“倒 U 型”关系,因此过渡带的形状应该是非线性的上凸型过渡。最后,可以得出  $A_{11}$  的函数形式,即式(1)。

$$A_{11} = \begin{cases} 0, & -0.26 \leq a < 0.07 \\ 2[(a-0.07)/0.02]^2, & 0.07 \leq a < 0.08 \\ 1-2[(a-0.09)/0.02]^2, & 0.08 \leq a < 0.09 \\ 1, & 0.09 \leq a < 0.13 \\ 1-2[(a-0.13)/0.04]^2, & 0.13 \leq a < 0.15 \\ 2[(a-0.17)/0.04]^2, & 0.15 \leq a < 0.17 \\ 0, & 0.07 \leq a < 0.24 \\ 1, & 0.24 \leq a \leq 0.26 \end{cases} \quad (1)$$

类似的方法可以得出情感为“焦虑”的隶属度函数  $A_{12}$ 、情感为“悲伤”的隶属度函数  $A_{13}$ ,以及情感为“平静”的隶属度函数  $A_{14}$ 。

$$A_{12} = \begin{cases} 0, & -0.26 \leq a < -0.06 \\ 2[(a+0.06)/0.09]^2, & -0.06 \leq a < -0.02 \\ 1-2[(a-0.03)/0.09]^2, & -0.02 \leq a < -0.03 \\ 1, & 0.03 \leq a < -0.08 \\ 0, & 0.08 \leq a < 0.13 \\ 2[(a-0.13)/0.04]^2, & 0.13 \leq a < 0.15 \\ 1-2[(a-0.17)/0.04]^2, & 0.15 \leq a < 0.17 \\ 1, & 0.17 \leq a < 0.24 \\ 0, & 0.24 \leq a \leq 0.26 \end{cases} \quad (2)$$

$$A_{13} = \begin{cases} 0, & -0.26 \leq a < -0.16 \\ 2[(a+0.16)/0.06]^2, & -0.16 \leq a < -0.13 \\ 1-2[(a+0.1)/0.06]^2, & -0.13 \leq a < -0.1 \\ 1, & -0.1 \leq a < -0.06 \\ 1-2[(a+0.06)/0.09]^2, & -0.06 \leq a < -0.02 \\ 2[(a-0.03)/0.09]^2, & -0.02 \leq a < -0.03 \\ 0, & 0.03 \leq a \leq 0.26 \end{cases} \quad (3)$$

$$A_{14} = \begin{cases} 0, & -0.26 \leq a < -0.17 \\ 1-2[(a+0.17)/0.07]^2, & -0.17 \leq a < -0.14 \\ 2[(a+0.1)/0.07]^2, & -0.14 \leq a < -0.1 \\ 0, & -0.1 \leq a \leq 0.26 \end{cases} \quad (4)$$

### 2.3.2 V变量的隶属函数

V变量和人们的情感偏好呈线性关系。当V值从小到小时,情感状态的排序依次为:悲伤、愤怒、平静、高兴。V变量也有4个模糊集:高兴、焦虑、悲伤、平静,将其隶属函数依次记为 $A_{21}$ 、 $A_{22}$ 、 $A_{23}$ 和 $A_{24}$ 。V的取值范围为 $[-0.33, 0.33]$ 。与A值一样,从文献[12]中可以看出,当V的取值在0.09以上时,“高兴”这一情绪会随着V值的变大表现得更明显。而当V值处在0.1附近时,情感偏向于平静。当V处于-0.16到-0.05之间时,“紧张、愤怒、焦虑”这一类情绪出现的几率很大。当V值小于-0.16时,伤感这一类的情绪会随着V值的变小而变得更明显。

确定这4个函数的具体形式。与上面确定A变量函数形式的步骤一样,先确定函数的整体结构。从文献[12]的table2中给出的数据可以大致确定 $A_{21}$ 是属于偏大型的函数,即在0.17这个值以上,会有“高兴”或与高兴相近的情感出现。然后确定隶属度为0和1的论域。隶属度为0的论域为 $[-0.33, 0.09]$ ,隶属度为1的论域为 $(0.19, 0.33]$ 。过渡带的范围为 $[0.09, 0.19]$ ,而且过渡带的形状是线形的。因此,最后选择梯形分布的隶属函数,且得出函数 $A_{21}$ 。

$$A_{21} = \begin{cases} 0, & -0.33 \leq v < 0.09 \\ (v-0.09)/0.1, & 0.09 \leq v < 0.19 \\ 0, & 0.19 \leq v < 0.33 \end{cases} \quad (5)$$

同样,根据以上方法,可以确定 $A_{22}$ 、 $A_{23}$ 和 $A_{24}$ 的函数表达式。

$$A_{22} = \begin{cases} 0, & -0.33 \leq v < -0.2 \\ (v+0.2)/0.44, & -0.2 \leq v < -0.16 \\ 1, & -0.16 \leq v < -0.06 \\ -v/0.06, & 0.16 \leq v < 0 \\ 0, & 0 \leq v \leq 0.33 \end{cases} \quad (6)$$

$$A_{23} = \begin{cases} 1, & -0.33 \leq v < -0.2 \\ (-0.16-v)/0.04, & -0.2 \leq v < -0.16 \\ 0, & -0.16 \leq v \leq 0.33 \end{cases} \quad (7)$$

$$A_{24} = \begin{cases} 1, & -0.33 \leq v < -0.04 \\ (v+0.04)/0.04, & -0.04 \leq v < 0 \\ 0, & 0 \leq v < 0.11 \\ (0.19-v)/0.08, & 0.11 \leq v < 0.19 \\ 0, & 0.19 \leq v \leq 0.33 \end{cases} \quad (8)$$

最后,根据多因素的综合隶属函数的实现方法,得出 $A_i = (A_{1i} + A_{2i})/2, (i=1, 2, 3, 4)$ 。

### 2.3.3 A值和V值的获取

由于A值和V值没有特定的度量单位,因此不可能像获取身高那样精确、客观地取得某首音乐的AV值。但是,根据这4种感情A值和V值对的分布情况,音乐的节奏、旋律以及音乐的一些背景知识,可以尽可能准确地确定AV值。比如,欢快的音乐大都音调比较高,起伏变化不是很大<sup>[13]</sup>。

## 3 情感模糊检索性能分析

### 3.1 实验设计与结果

首先,选取实验样本。为探讨此隶属函数的有效性和其在计算机检索方面的实用性,随机选取了90首音乐样本。这些音乐样本包括40首欧美经典音乐和50首中国经典音乐。为扩大其包容性,还包含了民族歌曲、乡村歌曲、流行乐曲、摇滚乐曲、古典乐曲等多种类型。

其次,确定每首乐曲的隶属值。对每一首样本音乐,先确定其A值和V值,再利用隶属函数得出它分别属于以上4种感情的隶属度。如样本音乐《彩云追月》,其曲调优美,节奏欢快,旋律起伏变化不大,整体上给人一种轻快、追逐的感觉,情绪主动而不亢奋。根据这样一个基本判断,取A为0.03, V为0.17。得出它的 $A_{11} = 1, A_{12} = 0, A_{13} = 0, A_{14} = 0, A_{21} = 0.8, A_{22} = 0, A_{23} = 0, A_{24} = 0.25$ 。进一步得到 $A_1 = 0.9, A_2 = 0, A_3 = 0, A_4 = 0.13$ 。

再次,建立简单的检索系统。检索系统的建立分两部分:数据表的建立和检索页面的创建。模糊检索的数据表为music1,包含5个属性字段:title, happy, anxiety, sad, calm。将每首乐曲的曲名和相应隶属度的值填入相应的字段。例如:将《彩云追月》填入数据表music1中,则在title字段中输入“彩云追月”,在“happy”字段中输入“0.9”,在后3个字段中依次输入“0”、“0”、“0.13”。

同时,对这90首音乐按照传统的方法来估量其情感类型,并将其对应的数据表记为music2。同music1一样, music2也有5个属性字段:title, happy, anxiety, sad, calm。如果某一首音乐属于某种情感类型,则将相应的情感值记为“1”;反之则记为“0”。例如,歌曲《北京的金山上》其感情基调明显是欢快型的。因此在music2数据表的happy字段中填入“1”,在另外3个属性字段中填入“0”。

在制作检索页面时,需要与数据库连接,其关键就是检索语言的设置。对于传统的检索方式,其sql="select \* from music2 where happy=1"。而对于模糊检索,则不能让其sql="select \* from music1 where happy>0",因为并不是所有隶属于“高兴”这一类情感的隶属值大于0的乐曲一定是愉悦的感情。因此,需要确定隶属度在什么范围才能让人感受到

比较明显的愉悦情感体验。通常采用“专家评估”的方法来确定范围。同样以“happy”这一类的情感为例,先将所有的隶属于“happy”这类感情的隶属度大于0的乐曲选出来,共有41首。其中有一些乐曲的隶属度值为0.13,有的为0.15,有的为0.25。请一位具有较丰厚的音乐知识的专家来对这些音乐进行判断,看看这些隶属度值比较小的乐曲能否让听众产生比较明显的愉悦情感体验。如果在某一确定的隶属度值之上的乐曲都能让人有比较愉悦的情感体验,而在此值以下的乐曲并不能让人有比较明显的愉悦体验,则将这个值确定为临界值。最后,确定“happy”这一类情感的临界值为0.3,故它的sql="select \* from music1 where happy >= 0.3 order by happy1 desc"。其他情感类型的检索语句按同样的方法做类似的处理。

最后,搜集和整理数据。首先,在数据表 music1 中检索出情感为“happy”的所有音乐曲目,并记下总的检索记录的条数 n1,将其他3种情感也以同样的方式记录下来。数据表 music2 也做类似处理,记为 n2。检索记录数的统计结果如表1所列。

表1 music1 和 music2 的检索结果

统计项	n1	n2
happy	33	30
anxiety	14	8
sad	33	30
calm	26	22
总记录条数	106	90

另外,此种检索方法的检索能力和检索准确度从查全率、查准率和 F 值这3个方面来考察。查全率是指被检出的相关文献占总文献内所有相关文献的百分比,记为 R。查准率是指被检出的相关文献占被检出文献总数的百分比,记为 P。例如,对表1中的“happy”一类的情感类型,其 R=89.2%。“33”表示查出的与 happy 相关的音乐资源的记录条数;“37”表示那些“happy”隶属度值大于零的音乐资源中,能够稍微让人体会出“happy”这一类情感的所有音乐资源的总的记录条数。对表1中“happy”一类的情感类型,其 P=100%。图1和图2是用以上两种检索方式得出的4种不同类型情感的查全率和查准率。在一般情况下,某一提问的全部相关文献数量不能精确获知,因此 R 的计算结果一般都是近似值。但若事先已经明确了所有被测资源的具体信息和样本数,则可以在此获知每种情感类型的 R 值。查全率和查准率标示了检索方式的某些方面,但究竟哪种检索方式更好,则需要进一步对查全率和查准率进行融合。因此,在此引入 P 和 R 的调和平均值 F 来考察检索能力。其中  $F=2/(1/R+1/P)$ ,  $R \neq 0$ ,  $P \neq 0$ 。由此可以计算出两中检索方式不同情感类型的 F 值,并得到图3。

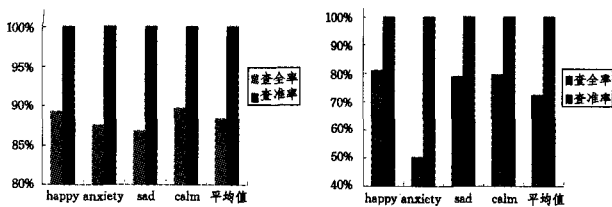


图1 music1 的查全率和查准率

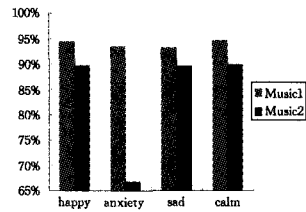


图3 两种检索方式的 F 值

最后,从面向用户的角度对这两种检索方式的性能进行评价。主要考察覆盖率 C 和出新率 N。其中  $C=Ru/U$ , U 表示用户已知的相关文档的总数, Ru 表示检索结果和 U 的交集的结果数。C 表示系统找到的用户已知的相关文档的比例。对于集合 U 的来源,随机采访了 10 位用户,让他们按自己想法在这 90 首乐曲中选出属于每种情感的乐曲。如对于“happy”这一类情感,让 10 位用户说出自己认为属于此类情感的音乐集合,并分别将其记为  $u_1, u_2, u_3, \dots, u_{10}$ 。最后“happy”这一类情感的集合  $U(h)=u_1 \cap u_2 \cap u_3 \cap \dots \cap u_{10}$ 。其他类型情感的 U 集合依此类推。出新率  $N=R_k/(R_u+R_k)$ ,  $R_k$  表示检索结果中返回的以前用户未知的相关文档集合。N 表示系统返回的新相关文档的比例。最后,经过数据收集和整理,得出图4和图5。

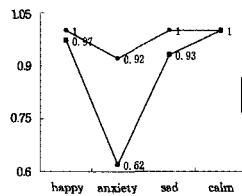


图4 两种检索方式的 C 值

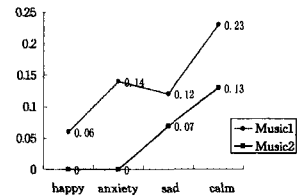


图5 两种检索方式的 N 值

### 3.2 实验数据分析

#### 3.2.1 检索数量

表1中 music1 同种情感类型的音乐记录数都要比 music2 中相同情感类型的音乐记录数多,且其检索总条数更是比 music2 多出16。这说明,传统情感检索的方法忽略了音乐情感的多样性,而模糊检索更能体现音乐情感的多样性,例如:赵咏华演唱的《最浪漫的事》不仅仅有“高兴”这一类的情感,同时也有“平静”这一类的情感。但总体上来说,它还是更偏向于“高兴”。因此,我们最后得出它隶属于“happy”和“calm”的值分别是0.7和0.41。但传统的检索方式则没有这种量化的结果。

#### 3.2.2 检索结果排序

从检索结果的排序上,模糊检索得出的结果是按照隶属程度从高到低按降序依次排列。以“happy”这一类情感为例,如果用户需要找出情感为“happy”的所有乐曲,那么最终呈现给用户的结果是隶属度为1的乐曲排在最前面,比较高兴的乐曲排在其次,排在最后的是隶属度为0.3的乐曲。若是传统的检索模式,用户检索出来的结果将不具有这种规律,只是按照填入数据表 music2 的先后顺序来进行排列。

#### 3.2.3 查全率、查准率和 F 值

从图1可得,就查准率而言,模糊检索的检索能力是毋庸置疑的。查全率接近90%,虽然没有达到100%,但对于那些没有被查询出来的资源,它们与用户想要的资源的相关性都比较小。根据我国学者对查全率与查准率这两者之间的联系探讨,不少人认为,对检索系统而言,查准率是第一位的。查全率依赖于查准率,查准率的提高有利于查全率的提

高<sup>[14]</sup>。因此,总的来说,此种检索方式的检索能力是比较优秀的。

从图2中看到,传统检索方式的查全率的平均值只有72.3%,与模糊查询相比,其查全率要低于16%。而且anxiety这一类的情感和sad这一类的情感具有一定的类似性,所以在非量化判断某首乐曲到底归属于哪一类情感类型时,出现判断失误的概率较其他类型的情感也要高一些。同时,对这两类情感的查准率和查全率也会有影响。通过对比,发现模糊检索的检索能力比普通的非量化查询的查询能力要高很多。

从图3中可得,这两种检索方式的F值都在65%以上。对于模糊检索来说,每种情感的F值都在93%以上,显然比传统检索方式的F值高。尤其是“anxiety”这一类情感与“sad”一类情感具有较大的相似性,若按照传统的检索方式,就很容易将那些带有伤感又带有焦虑情绪的音乐直接归到“sad”这一类情感里,从而导致其查全率大幅度降低。不妨设想一下,如果是对于大量的音乐资源,按照传统的这种检索方式,其检索能力必然会大大下降,尤其是查全率。若是用模糊检索这种方式,其查全率也会保持良好的势头。

### 3.2.4 相关度

音乐资源与这些情感类型不只有相关和不相关这两种描述,同时还有其相关度的级别。对于返回结果来说,其相关度级别越高,排名越前。由于对相关度的描述还没有统一的函数模型,在此,将隶属度函数的值直接作为其相关度的级别值。所以对模糊检索来说,能够清楚地知道每个音乐资源与每种情感类型的相关度。如果相关度值为0,则说明两者完全不相关,反之亦然。但是,如果按照隶属度来计量相关度,则不能肯定地说,“happy”隶属度值为0.05的音乐资源与“happy”这一类情感相关。由于相不相关是由用户来判断的,因此,对于相关这一概念的值的界定,需要若干用户或专家来确定到底这个隶属度值在什么范围才能说明相关。然而,对于传统的检索方法,则不能看出音乐资源与情感类型的相关度。

### 3.2.5 覆盖率和出新率

从图4可得,模糊检索能更多地覆盖用户已经知晓的相关资源,尤其是对于那些情感界限很明显的音乐资源,如“happy”和“sad”这两类情感类型的音乐资源,它们的覆盖率都达到了100%。显然,对于中间类型的情感“anxiety”,模糊检索的覆盖率要比传统检索的覆盖率高出很多。这也说明模糊检索更能体现出音乐情感的多样性。

图5中模糊检索的N值曲线位于传统检索的N值曲线的上方,这表示模糊检索的出新率要高于传统检索的出新率。从用户的角度上来说,模糊检索一方面更能包含用户的检索需求,也能为用户提供认知以外的一些新的相关性比较大的音乐资源。

**结束语** 音乐模糊检索方法从技术上满足了音乐情感的多样性,使音乐的数字化工作更贴近音乐的本质特征。实验结果证明,基于模糊数学来实现对音乐情感的检索是可行的、有效的。当然,由于实验设计中音乐样本数目不是很大,也没有考虑不同环境和不同心理状态下用户对音乐情感的鉴定会有所变化,因此导致A值和V值有一定程度的偏差。下一步的研究中,将针对用户情感波动对A值和V值的影响进行深入分析,使其根据情感变化自动调整,以实现更加可靠的音乐情感检索。

## 参 考 文 献

- [1] 康庆阳,聂自非. 基于内容的音乐检索及其在媒资管理系统中的应用[J]. 现代电视技术, 2007, 5: 90-95
- [2] 鞠源. 基于哼唱的音乐检索系统的研究与探索[J]. 情报杂志, 2010, 29: 160-163
- [3] 张宝华, 张品. 基于旋律的音乐检索系统[J]. 电声基础, 2005, 12: 4-8
- [4] 徐欣, 周运, 邵曦. 基于音乐情感特征提取的音乐检索分析[J]. 信息通信, 2011, 115(5): 9-12
- [5] 孙晓晔. 自动化音乐情感分类问题的研究[D]. 杭州: 浙江大学计算机科学与技术学院, 2010
- [6] 马希荣, 梁景莲. 给予情感音乐模板的音乐检索系统研究[J]. 计算机科学, 2009, 36(1): 239-242
- [7] 谢季坚, 刘承平. 模糊数学方法及其应用[M]. 武汉: 华中科技大学出版社, 2007
- [8] 甘露. 论音乐的情感特征[J]. 吉林艺术学院报, 2006, 67(4): 10-13
- [9] Russell J S. A Circumflex Model of Affect [J]. J. Personality Social Psychology, 1980, 39(6): 1161-1178
- [10] Thayer R E. The biopsychology of mood and arousal [M]. Oxford University Press, 1989
- [11] 王浩, 庄钊文. 模糊性可靠分析中的隶属函数确定[J]. 电子产品可靠性与环境试验, 2000, 8(4): 2-7
- [12] Stanley D J, Meyer J P. Two-Dimensional Affective Space: A New Approach To Orienting the Axes [J]. American Psychological Association, 2009, 9(1): 214-237
- [13] 王胜, 张石清. 基于二维情感空间的语音情感识别[J]. 台州学院学报, 2007, 29(6): 49-51
- [14] 沈建人. 查准率与查全率之间的关系[J]. 情报探索, 2006, 102(4): 32-34
- [15] Yang Y-H, Lin Y-C, Su Ya-fan, et al. A Regression Approach to Music Emotion Recognition[J]. IEEE Trans. Audio, Speech and Language Processing, 2008, 16(2): 448-457
- [16] 李名标. 模糊分类与模糊匹配相结合的模糊检索[J]. 计算机科学, 2000, 28(8): 37-39
- [17] Skyline 查询处理[J]. 计算机学报, 2011, 34(10): 1875-1796
- [18] 覃雄派, 王会举, 杜小勇, 等. 大数据分析—RDBMS 与 MapReduce 的竞争与共生[J]. 软件学报, 2012, 23(1): 32-45
- [19] 吴广君, 王树鹏, 陈明, 等. 海量结构化数据存储检索系统[J]. 计算机研究与发展, 2012, 49(Suppl. ): 1-5
- [20] 曾志勇, 杨辉, 余建坤. 基于 HMT 和哈希树的 Apriori 并行算法研究[J]. 计算机工程与设计, 2012, 33(1): 214-248
- [21] 关晓蕾, 钱宇华. 基于不完备信息系统的决策树生成算法[J]. 计算机科学, 2012, 30(1): 156-158
- [22] 王柯柯, 崔贯勋, 倪伟, 等. 基于单元的快速的大数据集离群数据挖掘算法[J]. 重庆邮电大学学报: 自然科学版, 2010, 22(5): 673-677

(上接第 195 页)

- [6] Allcock B, Chervenak A. Data Grid tools: enabling science on big distributed data [C] // Journal of Physics: Conference Series, 2005: 1-5
- [7] Costa P, Donnelly A. Camdoop: Exploiting In-network Aggregation for Big Data Application [C] // 9th USENIX Symposium on Networked Systems Design and Implementation, April 2012: 1-14
- [8] 王珊, 王会举, 覃雄派, 等. 架构大数据: 挑战、现状与展望[J]. 计算机学报, 2011, 34(10): 1741-1752
- [9] 张延松, 焦敏, 王占伟, 等. 海量数据分析的 One-size-fits-all OLAP 技术[J]. 计算机学报, 2011, 34(10): 1936-1946
- [10] 琳琳, 信俊昌, 王国仁, 等. 基于 Map-Reduce 的海量数据高效