

基于情感音乐模板的音乐检索系统研究

马希荣^{1,2} 梁景莲¹

(天津师范大学计算机与信息工程学院 天津 300387)¹ (北京科技大学信息工程学院 北京 100083)²

摘 要 传统的基于文本信息描述的音乐检索技术已经无法满足人们对检索智能化的需求,于是产生了基于内容的音乐检索方法。在此基础上将情感需求引入到检索中,对基于情感的音乐检索方法及模型进行了相关研究。首先构建了音乐情感空间来获得用户的情感描述;然后通过对情感音乐模型进行定义提出了情感音乐模板库,以得到满足用户情感需求的匹配模板;最后,在此基础上提出了基于情感音乐模板的音乐检索系统模型,力求探讨出一种基于情感的有效检索方法。

关键词 基于内容的音乐检索,音乐情感空间,情感音乐模型,情感音乐模板

Research on Music Retrieval System Based on Affective Music Pattern

MA Xi-rong^{1,2} LIANG Jing-lian¹

(College of Computer and Information Engineering, Tianjin Normal University, Tianjin 300387, China)¹

(Information Engineering School, University of Science and Technology Beijing, Beijing 100083, China)²

Abstract The traditional information retrieval technology which is based on text has been unable to satisfy the need for more and more music information. So content-based music retrieval technology was developed, which applies audio characteristics. This paper researched the music retrieval method and model based on affection by importing affective requirements from users. We constructed a music affective space in order to acquire users' affective description; introduced an affective music pattern database through defining PDA model, so as to get a music pattern satisfied by users' affective requirements; we put forward a music retrieval system model based on affective music pattern, aiming to think out an efficient affective-based retrieval method.

Keywords Content-based music retrieval, Affection space of music, PDA, Affective music pattern

1 引言

随着信息数字化与网络化的飞速发展,信息资源检索技术受到了越来越多人的关注。对于音频信息的检索来说,传统的基于常规属性文本信息(名称、作者或格式等标注信息)的检索方法已经越来越不能满足人们对信息检索形式的需求,基于内容的音频信息检索技术已经成为多媒体信息检索技术的研究重点之一。在基于内容的音乐检索技术的研究中,通常将音乐的时域、频域等声学特征作为内容进行匹配检索,然而这些特征由于过于专业化,很难得到用户的接受。音乐是情感的载体,通过各种基本要素的巧妙组合展现出一个丰富的情感世界^[1]。换句话说,情感才是音乐的主要内容,而目前的音乐检索方法却忽略了用户对音乐的情感需求。

人类对于信息的检索,是一种智能检索,更多地关注的是信息的内容本质;而对于音乐来说则是听觉上的感官以及情感上的共鸣。那么如何让机器检索也具有智能,则是本文主要的研究内容。情感作为一种语义特征是不同于音高、音调等声学特征的,它在人类智能活动中扮演着重要的角色,在认知和人机交互中起着关键性的作用。人工智能是使机器具有类似人一样的智能,而要实现真正的人工智能,关键是使机器

具有情感能力。美国 MIT 大学 Minsky 教授在她 1985 年的专著“*The Society of Mind*”^[3]中首先提出了让计算机具有情感能力,并指出问题不在于智能机器能否有任何情感,而在于机器实现智能时怎么能够没有情感。情感计算是关于情感、情感产生以及影响情感方面的计算,其目的是赋予计算机识别、理解、表达和适应人情感的能力,从而建立和谐的人机环境。

本文在基于内容的音频检索方法的基础上,通过对音乐的情感特征与声学特征的研究,根据音乐的情感状态特点,构建了音乐的情感状态空间;然后基于该状态空间,并结合音乐的声学特征,提出基于情感特征的音乐模板库;最后利用情感音乐模板库,对传统的基于内容的音乐检索方法进行改进,构建出基于情感特征的音乐检索系统模型,旨在为基于内容的音乐检索的智能化,特别是情感的智能化,探讨一种行之有效的方法。

2 基于内容的音频检索

不同音频类型具有不同的内在特征,但从整体上看,这些内在特征可划分为 3 级^[2]:最低层的物理样本级、中间层的声学特征级和最高层的语义级,如图 1 所示。

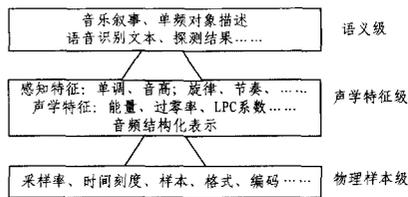


图1 音频内容分层描述模型

从低级到高级,内容逐级抽象,内容的表示逐级概括。在物理样本级,音频内容呈现的是流媒体形式。声学特征级中的一些特征,如音调、音高等可用于语音的识别或检测,并且还能够支持更高层的内容表示。处于最高层的语义级,则是音频内容、音频对象的概念级描述。在图1所示的3层描述模型中,可以看出每一层所提供的内容形式都是不同的,因此就决定了其不同的应用技术。其中,声学特征级和语义级得到了基于内容的音频检索技术的关注。在这两个层次上,用户可以通过提交概念进行查询或按照听觉感知来实现检索。

音频是声音信号形式,作为一种信息载体,可分为波形声音、语音声音和音乐声音3种类型。不同的音频类型,其音频特性有所不同,从而检索的方法也就不同。音乐是我们经常接触的媒体,其检索方法主要利用音乐的节奏、音符和乐器特征等内容进行检索。音频的听觉特性决定其查询方式不同于常规的信息检索。基于内容的查询是一种相似查询,查询中可以指定相似度的大小,还可以强调或忽略某些特征成分。用户则通过提供音乐示例,或哼唱出要查找的曲调来表达其查询的要求,从而查找出在某些特征方面相似的所有声音。

本文基于音乐的情感语义特征,构建了一个基于情感音乐模板的音乐检索系统模型,旨在探讨面向情感语义层的音乐检索途径与方法。

3 情感音乐模板

目前,针对音乐的检索大多是利用提取到的能量、频率等物理特征,或节奏、旋律等声学特征,通过特征聚类方法获得不同类别的特征模板,然后进行音乐样本的匹配判断。然而这种分类方法所得到的特征模板,仅仅是对大量音乐物理特征相似性的一种表现,它忽略了音乐中最本质的情感内涵。就目前的音乐分类方法来看,主要是基于音频的物理特征层与声学特征层实现的,那么如何设计出一种面向情感语义层的有效分类方法,就成为了本文下面将要讨论的问题。

本文所提出的基于情感音乐模板的音乐检索方法,是通过利用“情感音乐模板”来对音乐数据库中的乐曲进行匹配检索,从而获得具有某类情感的乐曲。“情感音乐模板”的定义和构建过程如图2所示,通过如下4个步骤完成:

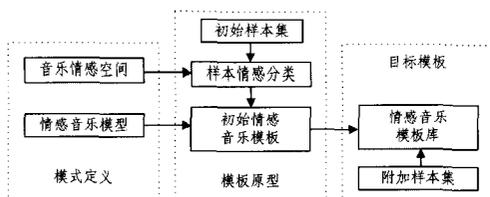


图2 情感音乐模板构建过程

- (1) 构建音乐情感空间,形成初步的音乐情感分类;
- (2) 定义情感音乐模型,从而为情感音乐模板的构建提供

量化标准;

(3) 依据情感音乐模型的定义,利用情感分类得到的样本子集构造各个情感类的音乐模板原型;

(4) 利用附加样本集,形成最终的目标情感音乐模板库。

3.1 构建音乐情感空间

为了实现以情感为需求进行的音乐检索,首先需要确定情感的描述方式和分类方法。Thayer提出的二维情感模型^[13]从影响情感的本质因素入手,认为音乐情感受限于压力和能量这两个维度因素,于是将情感分为渴望、满足、消沉和生气勃勃4类。本文从音乐检索的应用角度对该模型进行分析,认为这种情感描述方法过于抽象,对于能量与压力的变化用户很难准确地表达出来。也有人利用18组形容词对,从3个方面对音乐的情感色彩进行分析描述^[14];然而该方法提出的一些形容词对存在着交叉的现象,不利于情感的分类。基于以上两种方法在音乐检索应用中存在的实际问题,本文提出基于等级度量的音乐情感空间。

本文首先根据能量这一影响情感的本质因素,结合人类对音乐产生的情感状态,设计出12种情感状态描述符,如表1所示。并以此12种情感状态构建得到12维度的情感状态空间。对表1中的12种情感描述进行等级量化,分别为“非常”、“比较”、“一般”,其量化系数对应为2,1.5,1,以此作为被调查者对音乐的评价标准。然后,初步选取400首乐曲作为初始样本集,包括轻音乐、钢琴曲、摇滚、二胡等各种风格的乐曲,由被调查者进行评测,并根据评价结果中所选择的情感类别进行音乐归类,从而将200首音乐划分为12类。对于该分类过程,本文没有采用分类算法,主要考虑到情感本身就是人的一种感知,直接根据被调查者对音乐的情感感知获得初步分类结果,更具真实性。当然,由于人对音乐情感体验的个性化特点,使得这种分类不够精确。因此,本文在此基础上作进一步的改进,来获得最终情感音乐分类的特征模板。具体方法将在后面的内容中进行描述。

表1 音乐情感描述符

1. 振奋——消沉
2. 高兴——悲伤
3. 欢快——沉重
4. 活泼——严肃
5. 平静——烦躁
6. 放松——紧张

3.2 PDA情感音乐模型

根据音乐情感空间,对初始样本集进行了划分,从而得到12个音乐样本子集。在此基础上,本文提出利用PDA情感音乐模型来描述每一个样本子类。该模型从多角度定义了情感音乐模板的基本模式,其中包含音乐的声学特征、模板所对应的情感状态以及情感度量,涵盖了音乐内容模型中的多个层次。将PDA情感音乐模型用三元组 (P, D, A) 描述,定义如下:

P 为音乐的声学特征向量,我们选取了6个音乐声学特征作为其分量,即最大频率、两相邻测量点频率相差 >500 的次数、两相邻测量点频率相差 >1000 的次数、两相邻测量点频率不变的次数、每1000点频率跳跃次数和平均振幅,这6个分量可以较好地反映音乐的基音、节奏快慢、音高及响度。

在情感音乐模板中,我们将采用每个情感类中样本的平均声学特征向量 \bar{P} 来表示。

D 为每个情感类中的音乐样本距离向量,该向量包含3个分量,即对应情感类中样本间的平均距离 \bar{d} 、最大距离 d_{\max} 和最小距离 d_{\min} 。这3个距离用来对音乐的情感度(强、中、弱)进行分级判定。

A 为每个情感类模板所对应的情感状态值,区间范围为1~12,分别对应表1中的12种状态,用来对情感音乐模板进行索引。

情感音乐模板将在该模型的基础上完成最终的构建。

3.3 构建情感音乐模板库

为了构建情感音乐模板库,本文采用了基于样本和核的相似性度量的动态聚类算法^[14],来获得每个音乐样本分类模板的声学特征向量 P 。具体方法如下:

(1)根据音乐情感空间所获得的初始样本集(400首音乐)分类(12类),可以确定每一类的初始核 $K_j, j=1, 2, \dots, 12$,这里我们取类中全部样本声学特征向量 P 的平均值,即 $K_j=(\bar{P}_j, D_j, A_j)$ 。

(2)再次选取200首不同风格的乐曲作为附加样本集,按照算法1将每个样本分到相应的情感类中去。

算法1

若 $\Delta(p, P_j) = \min_k \Delta(p, P_k) \quad k=1, 2, \dots, 12$

则 $p \in \Gamma_j$

其中, p 为附加样本集中的每一个音乐样本, Γ_j 为每个情感类。

(3)重新修正核 K_j 。若核 K_j 保持不变,则继续执行步骤

(4),否则转步骤(2),并且重新获得每个情感类的核向量 K_j' 。

(4)将该模板存入情感音乐模板库,该数据库以情感状态值 A 作为索引值,用于情感模板的索引。

情感音乐模板构建完成,我们可以实现情感音乐的检索。过程如下:

(1)情感信息输入。该信息包括情感状态的描述、情感度(情感表达程度的强弱)等。

(2)情感音乐模板索引。利用情感信息量化值对情感音乐模板库进行索引查找。

(3)模板匹配检索音乐。利用获得的情感音乐模板对音乐特征库和原始音乐库进行检索。对于模板的匹配过程将由两部分组成,首先利用模板的声学特征向量 P_j 进行检索比较,当 $\Delta(p, P_j) > \mu$ (μ 为匹配相似度阈值)时,则该音乐不属于此情感类;否则,可判断属于该模板类,同时继续对模板的距离向量进行比较,判断 $\Delta(\Delta(p, P_j), d_{j, \max}) < \theta, \Delta(\Delta(p, P_j), \bar{d}_j) < \theta, \Delta(\Delta(p, P_j), d_{j, \min}) < \theta$, 3个相似度比较中哪一个能够满足,那么就可以确定该音乐具体属于哪一程度的情感状态(划分为强、中、弱3个级别程度)。

情感音乐模板库构建完成后,我们将通过选取模板来进行最终的基于情感模板匹配的音乐检索。

4 基于情感模板的音乐检索系统模型

在日常生活中,音乐带给人们不仅仅是听觉上的享受,更多的则是情感上的表达、宣泄或共鸣。语言、表情可以被人们用来传递情感,音乐也是如此。以往的音乐检索方法,不论是

采用音乐注释信息作为查询依据,还是通过由用户提供音乐示例来进行声学特征的匹配检索,都没有包含音乐情感的参与。从这个角度出发,本文提出了基于情感音乐模板的音乐检索系统模型,该模型对传统音乐检索方法进行了改进,将人的情感作为主要的检索需求,充分考虑了音乐的情感内涵,旨在探讨在音乐检索过程中情感的参与模式,从而实现情感驱动的人机和谐交互,如图3所示。

基于情感音乐模板的音乐检索系统模型主要由3个模块构成,分别为音乐库产生模块、情感音乐模板构建模块和用户查询模块。音乐库产生模块主要根据传统的基于内容的音频检索中数据库构建方法^[15]实现,首先由原始音乐生成原始音乐库,然后通过对音乐样本进行预处理、声学特征提取,从而得到音乐特征库,最后利用聚类方法形成聚类参数集。该模块的形成在系统没有涉及情感因素,仅是独立地完成音乐库及音乐声学特征分类的环节,为用户进行检索匹配时提供资源。这种基于声学特征的音乐库构建过程已较为成熟,本文在这里将不做更多的描述。

情感音乐模板构建模块通过生成情感音乐模板库来为用户提供情感的匹配模板。在检索过程中,用户为了表达情感只能通过一些特定的语言、文字描述出来,而这些情感信息在音乐原始库中很难被理解、被识别,换句话说就需要一种中间模式来表现用户的情感,基于此目的本文提出了“情感音乐模板”,以此作为情感与音乐的中间模式,架起情感与音乐实现通信的桥梁。对于该模块的设计思想及工作原理在第3节中已做了详细的介绍,下面将从系统模型的整体,对用户查询提交、处理、检索过程等进行描述,过程如下。

如图3所示,首先用户通过对情感的描述向系统提出情感需求。这里的情感描述主要以情感状态选择的方式来提供,即系统为用户列出系统构造的情感状态空间中的情感选项,由用户进行有范围限定、形式化描述的选项,同时对情感表现程度进行高、中、低的确定。用户除了需要提交情感需求外,还要对音乐的格式、大小等属性特征做出选择,一并提交给系统。系统根据用户的情感描述选择出相似度范围内的情感音乐模板,该模板与音乐的基本属性特征要求一起形成用户的查询矢量,提交给音乐数据库进行匹配检索,最后将检索结果反馈给用户,从而完成基于情感的音乐检索。

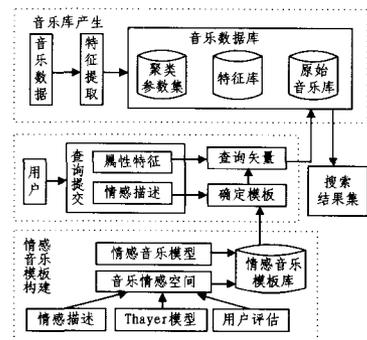


图3 基于情感语义的音频检索系统模型

结束语 随着计算机技术的发展,人们开始让计算机具有“智能”,这种智能不仅仅是知识上的,更是情感上的。基于这一目的,信息检索也开始向着智能的方向发展,产生了基于

(下转第278页)

SS	0.93024	0.92947	0.87661	0.94444	0.97720	0.98009	0.99276	7>6>5>4>1>2>3
结论	7 最佳,6,5 很好							
注	4. 曲线拟合 $R(t) = \exp(-\lambda t)$ 5. 改进指数分布曲线拟合一 $R(t) = a * \exp(-\lambda t)$ 6. 改进指数分布曲线拟合二 $R(t) = b + a * \exp(-\lambda t)$ 7. 二参数威布尔分布曲线拟合方法 $R(t) = e^{-\left(\frac{t}{\gamma}\right)^m}$							

结束语 系统可靠性是系统质量的重要指标,准确评估系统的可靠性对于系统质量的评定有着非常重要的作用。根据系统可靠性计算系统的平均失效时间 MTTF,从而定制系统巡检维修的周期对于现实系统投入运行后的巡检维修工作有重要的指导意义。本文从收集大量实际数据出发,优化可靠性试验方案,采用多种系统可靠性评估方法对目标系统可靠性进行评估并比较其评估结果的准确程度。最后选择根据实际最为准确的系统可靠性评估方法,对目标系统进行可靠性指标计算,根据现实需要定制目标系统的巡检维修周期用于指导现实的实践工作,收到了良好效果。

参 考 文 献

[1] 金星,洪延姬. 系统可靠性评定方法. 国防工业出版社,2005
 [2] Pham H. Software Reliability. Springer,2000
 [3] 黄锡滋. 软件可靠性、安全性与质量保证. 电子工业出版社,2002
 [4] 王学仁,等. 应用回归分析. 重庆大学出版社,1989
 [5] Li Zehui,Zhao Peng. Reliability Analysis on the δ -Shock Model of Complex Systems. Reliability, IEEE Transactions on, 2007, 56:340-348
 [6] Sarhan A M, Guess F M, Usher J S. Estimators for Reliability Measures in Geometric System Life Test Data Distribution

Model Using Dependent Masked. Reliability, IEEE Transactions on, 2007, 56:312-320

[7] Reliability testing - Compliance tests for constant failure rate and constant failure intensity. IEC61124-International Electro-technical Commission, 2006
 [8] Chen Jiading, Wei Xuelian. Improvement for the Sequential Scheme in MIL-HDBK-781 and IEC61124 (2002) Draft // The Joint Meeting of the Chinese Society of Probability and Statistics (CSPS) and the Institute of Mathematical Statistics (IMS). Beijing, 2005
 [9] 隗雪莲,陈家鼎. 对 MIL-HDBK-781 和 IEC61124(2002)草案中定时试验方案的改进. 应用概率统计, 2004(04):414
 [10] Gokhale S S, Lyu M R, Trivedi K S. Reliability Simulation of Component-Based Software Systems. Software Reliability Engineering, Nov. 1998:192-201
 [11] Stringfellow C, Andrews A A. An Empirical Method for Selecting Software Reliability Growth Models. Empirical Software Engineering, 2002, 7:319-343

(上接第 241 页)

内容的检索,基于多信息融合的检索等方法。而对于音乐这种特殊的媒介,情感是其本质的内涵,人们获取音乐的目的不只是为了听觉上的满足,更是情感上的交流。本文提出的基于情感音乐模板的音乐检索方法,正是在传统音频检索的基础上注入了情感因素,使查询更加贴近人性化,更加接近情感上的沟通,从而实现人机间情感的和谐交互。

参 考 文 献

[1] 王次昭. 音乐美学[M]. 北京:高等教育出版社,1994
 [2] 范宝元,韩秀苓. 音频工程基础[M]. 北京工业大学出版社,2002
 [3] 李国辉,李恒峰. 基于内容的音频检索概念和方法[J]. 小型微型计算机系统, 2000;1000-1220,11-1173-05
 [4] Picard R W. Affective computing[M]. London: MIT Press, 1997
 [5] 胡晓峰,李国辉. 多媒体系统[M]. 北京:人民邮电出版社,1997
 [6] 焦玉英,周华敏. 基于音频内容的交互渐进式音乐检索系统的设计[J]. 情报科学, 2004; 1007-7634, 12-1458-03

[7] 程凯,李应,黄樟钦. 音频数据的一种空间特征模型[J]. 计算机应用, 2004; 1001-9081, 01-0143-03
 [8] 赵雪雁,吴飞,庄越挺,等. 基于模糊聚类表征的音频例子检索及相关反馈[J]. 浙江大学学报, 2003; 10082973X, 0320264205
 [9] Bertini M, Bimbo A D, Pala P. Content - based indexing and retrieval of TV news. ELSEVIER Journal, 2001, 22:503-516
 [10] Li Dongge, Sethi I K, et al. Classification of general audio data for content-based retrieval[J]. ELSEVIER Journal, 2001, 22: 533-544
 [11] 陶建华,谭铁牛. 让计算机更善解人意[J]. 计算机世界, 2005(1)
 [12] 傅小兰. 人机交互中的情感计算[J]. 计算机世界, 2004(5)
 [13] Thayer R E. The biopsychology of mood and arousal[M]. Oxford University Press, 1989
 [14] 边肇祺,张学工. 模式识别[M]. 北京:清华大学出版社, 2000
 [15] 李恒峰,李国辉. 基于内容的音频检索与分类[J]. 计算机工程与应用, 2000(7)