

音乐检索特征库构建方法^{*}

李 鹏¹, 周明全²⁺, 李 娟², 黎南杉²

¹(北京师范大学 虚拟现实与可视化技术研究所,北京 100875)

²(北京师范大学 信息科学与技术学院,北京 100875)

Music Retrieval Feature Database Construction Methods

LI Peng¹, ZHOU Ming-Quan²⁺, LI Juan², LI Nan-Shan²

¹(Institute of Virtual Reality and Visualization Technology, Beijing Normal University, Beijing 100875, China)

²(College of Information Science and Technology, Beijing Normal University, Beijing 100875, China)

+ Corresponding author: E-mail: mqzhou@bnu.edu.cn, http://www.bnu.edu.cn

Li P, Zhou MQ, Li J, Li NS. Music retrieval feature database construction methods. *Journal of Software*, 2009,20(Suppl.):213-220. <http://www.jos.org.cn/1000-9825/09025.htm>

Abstract: Construct music feature database properly is significant to accurate music retrieval. On the foundation of content-based music retrieval framework, this paper divides the database construct methods into pitch extraction, score information and MIDI analysis method. It proposes a pitch extraction, post-processing methods and MIDI analysis methods to implement the database construction. Experimental results show that the two methods are correct and effective, could construct a music database accurately and rapidly.

Key words: music retrieval; music feature database; pitch extraction; score information; MIDI analysis

摘 要: 合理的音乐特征库构建方法对于实现准确音乐检索具有重要意义.以基于内容音乐检索框架为基础,将音乐特征库的构建分为采用基音提取方法建库、采用乐谱信息建库和采用 MIDI 分析方法建库 3 种方式,并提出一种基音提取及后处理方法和 MIDI 分析方法实现音乐特征库的构建.实验结果表明,这两种建库方法是正确有效的,可以准确、快速地实现音乐特征库的构建.

关键词: 音乐检索;音乐特征库;基音提取;乐谱信息;MIDI 分析

音频识别检索技术是在网络环境下处理多媒体海量数据的一项重要技术,与图像检索、视频检索并列已成为当今基于内容检索研究的热点.将音频的识别和检索技术与传统的文本检索相结合可以大幅度提高音频数据检索的效率和准确率,降低检索成本.基于内容的音乐检索是一个新兴的研究领域,随着多媒体数据库的发展,可用的音频数据集迅速增长,实现基于内容的音乐检索的需求得到凸显.

国外研究机构对音乐检索进行了多方面的研究.1995 年,Ghias^[1]对单音轨的 MIDI 音乐的哼唱检索进行了开创性研究,采用时域自相关算法提取音高信息,然后使用字符串近似匹配的方式实现对单声部音乐的检索.新西兰 Waikato 大学的 McNab^[2-5]与新西兰数字音乐图书馆合作开发了一套名为 MT(melody transcript)和

* Supported by the National Natural Science Foundation of China under Grant Nos.60673100, 60736008 (国家自然科学基金); the Beijing Municipal Natural Science Foundation of China under Grant No.4081002 (北京市自然科学基金)

Received 2009-05-15; Accepted 2009-07-23

MELDEX系统,但是MELDEX系统无法正确切割音符,用户在哼唱时,必须在音符与音符之间自行留下短暂的停顿或插入特定断句声音,用户使用起来不方便,也不自然.Tomonari等人^[6]提出同时使用音长和音高作为旋律特征,然后采用基于“动态闭值调整”的匹配方法,采用了先粗选后细选的匹配策略.Kosugi^[7]提出同时使用音高变化和音高分布来提高系统的性能.他们开发出了一套名为SoundComPass的系统,系统中有 10 086 首歌曲,在使用时需配合乐器的节拍哼唱.韩国的Seungmin等人^[8,9]改进了基音提取算法,在传统检索算法的基础上增加了访问频度索引功能,采用基于遗传算法的相关反馈机制,使算法的检索正确率有了显著提升.

国内对音乐信息检索的研究正在逐步深入,并且取得了一定进展.台湾清华大学的张智星等人^[10-12]在哼唱式音乐检索方面起步较早.他们在提取基音值之后,采用分级过滤算法对音高进行处理,首先排除掉一部分候选音乐,然后采用DTW算法进行精确匹配,并开发出了一套名为“Super MBox”的点歌系统.微软亚洲研究院的卢烈^[13]、中国科学院声学研究所的李明^[14]、上海交通大学的李扬等人^[15]、浙江大学的冯雅中等人^[16]、西北大学的王小凤等人^[17,18]在基于内容的音乐检索领域提出过不少解决方案和实验研究.北京师范大学在该领域也展开了研究,先后提出了一种快速生成MIDI文件的方法^[19]、旋律轮廓表示方法和哼唱检索方案,设计实现了多个依靠哼唱/歌唱实现检索的原型系统^[20-23].

1 音乐检索框架体系

通过对国内外研究机构实现基于内容音乐检索方案的考察,可以得到如图 1 所示的音乐检索框架体系.虽然每个机构的具体解决方案各有不同,但都包含了框架中所述的基本功能单元.在解决具体问题中,注意力大多集中在基音提取、旋律轮廓(音乐特征)表示方法和匹配算法上边,在这 3 个方面每年都有大量论文进行多方面的研究,但在如何组建乐曲旋律特征数据库方面,却没有引起广泛重视.音乐特征库是实现音乐检索的基础,具有完善的特征定义和正确的特征值才能使用户输入较为准确时得到正确的曲目.如果特征库定义本身不能完整地表达乐曲旋律变化的特征,或者在创建特征库时因为种种原因未能生成可靠的特征值,就会导致即使歌曲特征输入非常准确也无法得到相应的歌曲列表输出.可见,完备、准确的音乐特征数据库对于实现准确检索、基音提取算法和匹配算法效果测试都具有非常重要的作用.

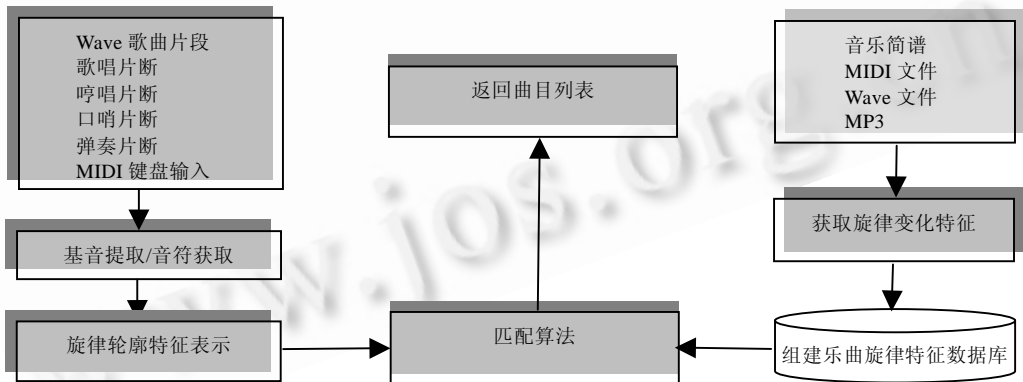


Fig.1 Framework of music retrieval

图 1 音乐检索框架

目前实现音乐检索最流行的形式是使用人声输入进行查询,无论是歌唱还是哼唱,在表现音乐旋律变化特征方面只可能获取到 3 个方面的信息,即音高信息、音长信息和节拍信息.音高信息反映的是所唱歌曲中的音调变化,音长信息表现的是某个音调持续的时间长短,节拍信息反映了某段歌曲哼唱/歌唱的快慢程度.基于这种分析,音乐特征库的定义中一般包括这 3 种信息,或者仅包括音高、时长的变化信息,将节拍信息设为默认速率,然后采用弹性匹配算法实现在一定节拍变化限度内的旋律特征匹配.

本文以构建音乐特征数据库作为研究对象,在广泛调查了国内外研究机构在实验音乐检索中各自不同的

建库方式之后,对使用比较多的建库技术做了论述,同时,在研究采用基音提取方法建库时提出了一种新的基音分析方法,使得音乐库特征值能够更准确地反映乐曲旋律变化,在研究使用 MIDI 文件建库时,提出了一种 MIDI 分析方法,对出现的一些问题提出了解决办法,实现了音乐旋律的自动抽取,达到高效建库的目的.最后通过原型系统的实验证明了该建库方法的正确性和有效性.

目前由文献调研正式地得到以及通过对其系统分析进行推论得到的建库方式主要有 3 种体系:通过基音提取技术建库、通过简谱建库和通过 MIDI 文件建库.下面分别加以论述.

2 采用基音提取方法建库

在音乐特征库的构建过程中,可以借助音频信号分析中的基音检测技术实现音乐旋律获取,实现建库.具体包括 3 个步骤:单声部 wave 文件获取、旋律自动分割、特征存储.

2.1 单声部 wave 文件获取

基音提取方法处理的是时域信号,对于待处理的声音信号需要 wave 格式文件,或者将其他压缩格式转换为 wave 格式.单声部是对 wave 格式文件的另一个要求,即用于建库的 wave 文件中应该只包含一种乐器对主旋律的演奏.歌曲中通常包括人声演唱部分和背景音乐伴奏部分,歌曲的主旋律应该被定义为原歌曲的人声歌唱部分,因为实现检索的人声输入一般都是针对具有歌词的部分旋律,所以最佳的 wave 文件是某个乐器对主旋律的演奏录音.但要搜集这类具有特殊要求的数字音乐并不是一件容易的事情.一个好的解决办法就是首先获取待入库歌曲的单声部 midi 文件,然后使用“MIDI2WAVE”软件或者内录实现目标 wave 文件的生成.

2.2 音乐特征获取

2.2.1 基音检测方法

得到合适的 wave 文件之后需要对其进行基音提取,以获取旋律变化特征.基音提取方法是实现音乐检索,尤其是哼唱式音乐检索必不可少的步骤,作为歌曲旋律变化特征获取的基础具有至关重要的地位.其主要功能是对用户的查询信号进行分析得到其频域变化信息,该信息直接反映了查询片段的音高旋律变化特征.基音提取的常规处理办法是对一查询片段进行加窗分帧,并且重叠移动形成多帧,然后对每帧数据作基音提取操作,最终得到代表音调变化的基音序列.目前国际上出现的基音检测方法有自相关函数法、平均幅度差函数法、小波变换法、倒谱法、并行处理技术方法、简化逆滤波法、基于 Hilbert-Huang 变换的基音周期检测等方法.其中,自相关函数(autocorrelation function,简称 ACF)法因其实现简单,检测效果好而被广泛使用.

自相关函数是对信号进行短时相关分析时最常用到的特征函数.音乐信号 $s(m)$ 经窗长为 N 的窗口截取为加窗帧信号 $S_n(m)$,定义每帧的自相关函数 $R_n(k)$ 为

$$R_n(k) = \sum_{m=0}^{N-k-1} S_n(m)S_n(m+k) \quad (1)$$

其中, $k = (-N+1) \sim (N-1)$. 由于信号的自相关函数在基音周期的整数倍位置上会出现峰值,因此可通过检测峰值的位置来提取基音周期值,进而得到该帧的频率,即基音值,反映的是音调高低信息.

因为环境、电流等噪声的干扰,得到的基音序列中存在着大量的偏离点,传统的平滑滤波算法在去除粒偏离点的同时会对原波形固有的高低突变造成影响,因此,为了解决这个问题,提出了一种“小阁”方法,可以很好地达到滤除偏离点、保留固有突变的目的.具体步骤如下:

将基音序列表示为 $X(m)$,滤波宽度定义为 L ,基音值波动范围是 $Range$,定义一个二维数组 $A[Range,L]$,并定义 $Y(m)$ 保存滤波之后的数据.实施步骤如下:

Step 1. 创建一个二维数组 $A[Range,L]$,以滤波宽度作为其横坐标,基音值波动范围作为其纵坐标.

Step 2. 申请一个与基音序列数组等大的数组空间,作为滤波之后数值的保存空间.

Step 3. 截取基音序列的前 L 个基音值,并将其按照值的大小分别放入到数组 A 中, A 中的每一列代表一个小阁.

Step 4. 统计数组 A 中每个小阁中存入数据的个数,找到数据数目最多的那一列,并计算其均值.

Step 5. 将该均值作为滤波之后的基音值写入到新的基音序列空间对应位置.

Step 6. 清空二维数组 A .

Step 7. 将滤波指针向后移动一个位置.

Step 8. 对基音次序在 $\left(\left[\frac{L}{2}\right], \left[m - \frac{L}{2}\right]\right)$ 之间的数据反复执行 Step 3~Step 7 的操作,直到数组数据取完为止.

以上的方法是完成一次滤波的情形.实验结果表明,多次滤波可使波形变化保持稳定,达到最佳滤波效果.完成该步骤之后即可得到能够真实反映查询片段音调变化的规整信息.

2.2.2 旋律自动分割

自动断句(基音序列自动分割)是另外一个必不可少的环节.实现人声输入检索具有一个重要特征,人们在使用声音(歌唱/哼唱/口哨...)进行查询时,都是针对歌曲中某几个连续的完整短句,不会出现部分短句或者部分短句拼接的情况.而在音乐特征库的构建中,如果忽略自动断句环节,就会导致将用户查询特征与整个歌曲进行步进式匹配,从而大量出现查询特征序列与不完整短句组合匹配的情况,这将使匹配算法的准确度大为降低,同时急剧增加系统运行时间.

歌曲在演唱过程中会出现多处停顿,停顿各有长短,选择一个合适的断句标准至关重要,应该遵从两个原则:(1) 不要断开用户在检索时以最小单元演唱的句子;(2) 尽量使断开后短句的数量减少.鉴于两个连续的句子未能正确断开会使用户唱歌检索时的准确率受到严重影响,而短句数量的增加只是对检索速率略有影响,因此应该适当将断句时长标准调低.

实现基音序列自动分割的传统方法是按照其时域信号的幅差音符进行分割,但该方法会因计算音符占帧不准而影响基频平均值,或者由于错误的划分而影响音符的数量进而影响匹配准确性.北京师范大学的秦静提出了一种基于幅度差函数的动态阈值音符分割算法^[20,21],较好地解决了断句问题.

2.3 特征存储

得到乐曲旋律变化信息和句子分段信息之后,可以采用以下 3 种形式实现存储:

1. 直接存储.将带有断句标志的基音序列直接存储在二进制文件中,每个乐曲建立一个文件,在实现匹配时依次读取实现检索.该方法实现简单,但是当乐曲数目较多时,检索需要依次载入全部文件,耗时较多.
2. 建立特定文件结构存储.上海交通大学在这方面的研究中建立了一种自定义的ref文件^[24],以相对音高差和时长的形式实现存储.
3. 按照向量格式存储.即以特征向量(音高,音长,是否断点)的形式对一系列基音序列进行存储.

后两种方法的优点是节约存储空间,检索时载入较快.缺点是对基音序列的后处理有严格要求,在每个音符的平稳发声段,基音值必须严格相等.

3 采用乐谱信息建库

获取音乐旋律变化信息最直接的方法之一就是通过对该乐曲的简谱或者五线谱,目前也容易在短时间内搜集到大量歌曲的乐谱.有两种以乐谱信息建库的方法:

1. 直接手动输入乐谱信息.输入乐曲的音符信息、节拍信息和断句信息,直接保存在数据库中.该方法因有人工参与而能很好地实现断句,但效率很低,只能在小规模建库时使用,同样也因为人工干预导致容易出现疏漏,而且不易察觉.
2. 使用光学乐谱识别(optical music recognition,简称 OMR)技术.计算机光学乐谱识别技术利用扫描仪等数码设备对已有的纸质乐谱进行扫描,形成数字乐谱图像,再利用专用的软件系统进行处理与识别,提取出音符、音长、节拍及断句信息.目前,国际、国内也有多家机构在这方从事研究,最高的识别率可达 90% 以上.通过 OMR 技术可以实现将乐谱信息转换为 MIDI 文件.这也为通过 MIDI 文件建库提供了便利.该技术属于图像处理与识别领域,本文不加详述.

4 采用 MIDI 分析方法建库

MIDI 是 Musical Instrument Digital Interface 的缩写,直接翻译过来的意思就是乐器数字化接口.可以把 MIDI 理解成是一种协议、一种标准,或是一种技术.MIDI 也是一种数字音乐文件的格式,文件由很多信息和命令构成,用来实现使用不同的乐器演奏不同的音调,用不同方法调整的音色参数等.MIDI 就是一组命令,像音乐盒里拨动簧片的滚轮,MIDI 本身并不存储声音信号本身,而是告诉声卡在什么时间、什么位置(音高)、什么强度下发一个什么音,发多长时间,是否颤音,是否力度变化等.通过网络可以很容易地获取到大量歌曲的 MIDI 格式文件,MIDI 键盘也是一个快速生成 MIDI 文件的设备,但需要专业人员操作.目前,使用 MIDI 文件实现音乐特征库构建在国内外的研究中比较常见,但使用中仍然存在一些问题,相关论述并不多.具体实现过程包括 MIDI 预处理、MIDI 文件分析和旋律切分.

4.1 MIDI预处理

MIDI 文件格式有多种,首先要将其转换为标准 MIDI 格式.MIDI 文件中通常包含多个音轨,每个音轨代表一个特定的乐器,主旋律归属某个音轨,该音轨还存在着丰富的附加信息,但在建库时,这些附加信息及非主旋律音轨都成了干扰,它们会给建库和实现检索带来严重影响.根据不同情况,有 3 种不同的处理方案.

1. 如果 MIDI 文件只有 1 个主旋律音轨,则直接进行 MIDI 文件分析,如果 MIDI 文件的作者用类似“MELODIES”,“VOCAL”,“SING”,“SOLO”,“LEAD”,“VOICE”等关键词标识了主旋律所在音轨,则保留该音轨,删掉所有的其他音轨.
2. 如果文件有多个音轨,并且主旋律音轨未知,这时需要考察每个音轨的左、右声道平衡度、主音量、发音时间等因素识别主旋律音轨,主旋律所在音轨的声道发音比较平衡,音量较大,而且贯穿整个文件始末.发音时间较短的音轨可以直接删除.
3. 对于主旋律出现在两个及以上音轨的 MIDI 文件则需要人工干预,将两个音轨合并,删除多余音轨.对于无法准确判断主旋律的 MIDI 文件也应该用 Akoff music composer 等软件辅助加以鉴别.

4.2 MIDI文件分析

确定了 MIDI 文件主旋律所在音轨之后,读取文件的二进制内容,按照 MIDI 编码规则获取音高和时长.读取 MIDI 头文件后,可以获得调号、节拍、每分钟节拍数、音轨个数等信息.调号表示音乐基准音的定位,占用 1 个字符,必须为 A,B,C,D,E,F,G,通常为 C 调;节拍取值如 2/4,3/4,4/4,3/8,6/8 等;每分钟节拍数表示每分钟演奏的节拍总数,取值在 40~200 之间.音轨个数表示此歌曲的声部数.

按照主旋律音轨号码,读取数据块中该音轨的信息.MIDI 以事件和消息的方式控制乐器的演奏方式,包括 MIDI 事件、Sysex 事件和 Meta 事件.MIDI 消息的种类很丰富,而且没有统一的格式,这意味着只能把每一种消息单独处理.但需要获取的音高时长信息通过对 MIDI 信息的分析即可获得.消息中最常用的是音符打开消息(note-on)和音符关闭消息(note-off).MIDI 中没有音符的概念,因此要通过将对应的音符开启和关闭事件配对形成一个音符,称为原始音符,之后还需要将音符开始时间戳和结束时间戳转换成音符开始时间戳和音符持续长度.

为了完成上述两个任务,使用一个大数组缓存 16 个通道里的 128 个音的状态.在接收到音符打开与关闭消息时进行记录,并同时计算开始时间与持续时间.最终得到如图 2 所示的音高、音长序列.

4.3 旋律切分

按照时长的分布状况,设定合适的阈值对音高序列进行自动切分,一般的方法是绘制时长信息直方图,将第一个波谷所在值作为断句阈值,通过统计将本例阈值设为 500 可得到如图 2 所示的断句效果.之后可以按照第 2.3 节所述方法进行存储.

Track #3:				720	2 400	82	c
时长	起始时间	音高	断句	120	3 120	82.	
240	0	82.		120	3 240	83.	
240	240	80.		240	3 360	80.	
720	480	78	c	120	3 600	78.	
240	1 200	78.		360	3 720	80.	
240	1 440	78.		120	4 080	77.	
240	1 680	80.		120	4 200	77.	
240	1 920	82.		600	4 320	78.	c
240	2 160	83.		120	4 920	75.	(未完)

Fig.2 Results of MIDI analysis

图 2 MIDI 分析结果

以上论述的内容主要是从各种形式的音乐信息载体中获得描述该音乐音调变化特征信息及组织方式,具体的实现中需要在数据库中创建至少 1 个表,保存包括歌名、歌词、演唱、曲作者、词作者、流派分类、音乐文件地址、特征文件地址、特征字符串等众多信息,并且开发一个音乐特征库维护界面,实现对乐曲旋律特征的查阅、入库、修正、删除等操作。

完成了音乐特征库的创建之后就可以按照合适的查询方式、基音提取算法和旋律特征表示方法,形成查询特征序列,采用模糊匹配算法与音乐特征库中每首歌曲进行匹配检索,按照相似度由高到低返回歌曲列表。

5 系统及实验

5.1 验证实验

为确保建库方法有效,开展如下实验进行验证:首先得到某歌曲的单声部 MIDI 音乐,直接对其分析提取旋律特征,并绘制音高图,如图 3(a)所示绘制的是歌曲《世上只有妈妈好》前两句的音高曲线;然后将该 MIDI 转录为 wave 文件,用基因提取方法获得其音高曲线,如图 3(b)所示;最后使用著名音频分析软件 Praat 对 wave 文件作分析,绘制音高曲线,如图 3(c)所示。三者具有很高相似性,证明了使用 MIDI 分析方法和基音提取方法得到的旋律特征是正确的。MIDI 分析图 3(a)中的 C 表示旋律切分位置。

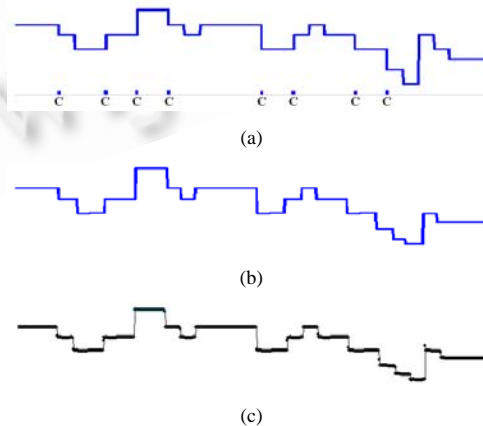


Fig.3 Validate experiments figure

图 3 验证实验图

5.2 系统实现

按照基于内容音乐检索的框架,设计实现了一个可以通过哼唱/歌唱实现歌曲检索的系统.系统所使用的数据库分别为按照基音提取方法和 MIDI 分析构建的音乐特征库,库中包含了 121 首中文流行歌曲.系统设计了一种简易的交互界面(如图 4 所示),用户可以在单击按钮之后现场哼唱/歌唱,系统实时记录人声输入,并在录音结束之后对查询文件进行预处理、提取基音、后处理、匹配检索等操作,最终按相似度返回若干首歌曲名称.单击歌曲即可实现播放,并显示该歌曲歌词.邀请 4 位未参加过特别音乐培训的用户(2 男 2 女)进行哼唱/歌唱 8s 左右或者两句歌词,共生成声音文件 48 个.评判标准为前 3 位即算作检出.共开展 3 组实验:第 1 组采用基音提取方法,不带自动断句功能,wave 建库文件来自单声部 MIDI 文件的转录;第 2 组采用基音提取方法并且实现自动断句;第 3 组使用 MIDI 分析并且已完成自动断句.实验结果见表 1.实验结果表明,使用自动断句功能可以大幅度提高检索准确率,并且节约检索时长.MIDI 建库的效果也很好,但因为 MIDI 文件本身的复杂性,需要在对其进行分析的过程中再做更细致的工作.



Fig.4 System interface

图 4 系统界面

Table 1 Experimental results

表 1 实验结果

第 1 组	第 2 组	第 3 组
5.4%	93.8%	89.6%

6 结 论

本文以基于内容音乐检索框架为基础,以构建音乐特征数据库为研究对象,将音乐特征库的构建方式分为 3 种,采用基音提取方法建库、采用乐谱信息建库和采用 MIDI 分析方法建库.分别按照不同的建库方法进行阐述,提出了一种基音提取及后处理的方法和对 MIDI 进行分析提取特征的办法.基于这种建库方法,进行了原型系统设计,开展实验.实验结果表明,两种建库方法对于实现基于内容音乐检索是正确、有效的.

References:

- [1] Ghias A, Logan J, Chamberlin D, Smith BC. Query by humming: musical information retrieval in an audio database. In: Zellweger P, ed. Proc. of the 3rd ACM Int'l Conf. on Multimedia. New York: ACM Press, 1995. 231-236.
- [2] McNab RJ, Smith LA, Witten IH. Towards the digital music library: Tune retrieval from acoustic input. In: Hall W, Little TDC, eds. Proc. of the 4th ACM Int'l Multimedia Conf. New York: ACM Press, 1996. 11-18.
- [3] McNab RJ, Smith LA, Witten IH. Signal processing for melody transcription. In: Hamilton. Proc. of the 19th Australasian Computer Science Conf. Department of Computer Science, University of Waikato, 1995. 301-307.
- [4] McNab RJ, Smith LA. Melody transcription for interactive applications. Department of Computer Science, University of Waikato, 1996. 1-22.
- [5] McNab RJ, Smith LA, Witten IH, Henderson CL. Tune retrieval in the multimedia library. Multimedia Tools and Applications, 1996,10(2-3):113-132.

- [6] Sonoda T, Goto M, Muraoka Y. A WWW based melody retrieval system. *Electronics and Communications in Japan*, 2002,85(9): 721–731.
- [7] Kosugi N, Nishihara Y, Sakata T, Yamamuro M, Kushima K. A practical query-by-humming system for a large music database. In: Ghandeharizadeh S, ed. *Proc. of the 8th ACM Int'l Conf. on Multimedia*. New York: ACM Press, 2000. 333–342.
- [8] Rho SM, Hwang E. FMF: Query adaptive melody retrieval system. *The Journal of Systems and Software*, 2006,79(1):43–56.
- [9] Rho SM, Han BJ, Hwang EJ, Kim MK. MUSEMBLE: A novel music retrieval system with automatic voice query transcription and reformulation. *The Journal of Systems and Software*, 2008,81(7):1065–1080.
- [10] Jang JSR, Lee HR. Hierarchical filtering method for content-based music retrieval via acoustic in put. In: Georganas ND, ed. *Proc. of the 9th ACM Multimedia Conf*. New York: ACM Press, 2001. 401–410.
- [11] Jang JSR, Lee HR, Chen JC. Super MBox: An efficient/effective content-based music retrieval system. In: Georganas ND, ed. *Proc. of the 9th ACM Multimedia Conf*. New York: ACM Press, 2001. 636–637.
- [12] Jang JSR, Gao MY. A query-by-singing system based on dynamic programming. In: *Proc. of the Int'l Workshop on Intelligent Systems Resolutions*. Cweb Inc., 2000. 85–89.
- [13] Lu L, You H, Zhang HJ. A new approach to query by humming in music retrieval. In: Alghoniemy M, ed. *IEEE Int'l Conf. on Multimedia and Expo*. Los Alamitos: IEEE Computer Society, 2001. 22–25.
- [14] 李明,颜永红.一种基于哼唱的音乐检索方法.见:第8届全国人机语音通讯学术会议论文集.2005.433–437.
- [15] 李扬,吴亚栋,刘宝龙.一种新的近似旋律匹配方法及其在哼唱检索系统中的应用.计算机研究与发展,2003,40(11):1554–1560.
- [16] 冯雅中,庄越挺,潘云鹤.一种启发式的用哼唱检索音乐的层次化方法.计算机研究与发展,2004,41(2):333–339.
- [17] 王小凤,周明全,耿国华,郭红波.一个使用歌谱信息进行哼唱检索的系统.计算机辅助设计与图形学学报,2007,19(17):941–946.
- [18] 郭红波.音乐哼唱检索关键技术研究.西安:西北大学,2007.
- [19] 周荣.基于MIDI乐音的音符编码和自动乐谱编制系统的研究.北京:北京师范大学,2007.
- [20] 秦静,周明全,王醒策,沈复兴.基于动态分割和加权综合匹配的音乐检索算法.计算机工程,33(13):194–199.
- [21] 秦静.基于旋律的音乐检索关键技术研究.北京:北京师范大学,2007.
- [22] Li NS, Zhou MQ. Audio denoising algorithm based on adaptive wavelet soft-threshold of gain factor and teager energy operator. In: *Int'l Conf. on Computer Science and Software Engineering*. Los Alamitos: IEEE Computer Society, 2008. 787–790.
- [23] Li P, Zhou MQ, Wang XS, Li NS. A novel MIR system based on improved melody contour definition. In: Tian XZ, ed. *Proc. of the Int'l Conf. on MultiMedia and Information Technology*. Los Alamitos: IEEE Computer Society, 2008. 409–412.
- [24] 张俊杰.基于和谐泛音检测的主旋律提取技术.上海:上海交通大学,2007.



李鹏(1981—),男,陕西扶风人,博士生,主要研究领域为音频检索,模式识别,图像检索.



李娟(1985—),女,硕士生,主要研究领域为音乐信息检索.



周明全(1954—),男,教授,博士生导师,主要研究领域为虚拟现实与可视化技术,多媒体处理技术,数据挖掘.



黎南杉(1983—),女,硕士生,主要研究领域为音乐数据库构建,检索技术.