

## 一个使用歌谱信息进行哼唱检索的系统

王小凤<sup>1)</sup> 周明全<sup>1,2)</sup> 耿国华<sup>1)</sup> 郭红波<sup>1)</sup>

<sup>1)</sup>(西北大学信息科学与技术学院 西安 710069)

<sup>2)</sup>(北京师范大学信息科学与技术学院 北京 100875)

(flysuya@sina.com)

**摘 要** 提出一种直接使用歌谱信息进行哼唱检索的方法,并在此基础上实现了一个基于内容的哼唱检索系统——Music Angel。该系统采用歌谱轮廓作为检索目标,首先提出歌谱轮廓特征提取算法并构造标准音调差值图,然后提取哼唱片段的基音轮廓特征并将其转换为歌谱轮廓特征,最后使用动态规划算法进行字符串相似度计算,得出检索结果。结果表明,该系统对环境噪声有较好的鲁棒性;在含有 405 首歌曲的搜索空间中,检索前 5 位成功率超过 90%。

**关键词** 哼唱检索;音调差值表;歌谱轮廓;基音轮廓  
**中图法分类号** TP391.4

## A Query by Humming System Based on Score

Wang Xiaofeng<sup>1)</sup> Zhou Mingquan<sup>1,2)</sup> Geng Guohua<sup>1)</sup> Guo Hongbo<sup>1)</sup>

<sup>1)</sup>(School of Information Science and Technology, Northwest University, Xi'an 710069)

<sup>2)</sup>(School of Information Science and Technology, Beijing Normal University, Beijing 100875)

**Abstract** In the paper, a new method based on score is presented and a query by humming system—"Music Angle" is implemented. The score contour and the pitch difference table are firstly constructed. Then the pitch contour characters are picked up and converted to score contour characters. In the end the querying results are obtained by dynamic programming algorithm. The results show that the system are very robust against surroundings noise, and the top-5 success rate is over 90% in a music database comprised 405 songs.

**Key words** query by humming; pitch difference table; score contour; pitch contour

音乐与人的听觉感知紧密相关,它更多地传达了一种感情,一种很难量化的情绪,音乐的这种特性决定了在音频的分类检索技术中所用到的歌名、演唱者等外在信息对音乐分析并不适用。当前绝大多数的音乐搜索引擎,如 Yahoo.com, Google.com 等,都是使用歌曲的名称、演唱者、作者或者歌词等来建立索引架构供用户使用。但用户经常会遇到这种情况,他能记得歌曲的某些片断,却无法想起歌名或演唱者等信息,利用上面搜索引擎无法找到目标歌曲。

使用哼唱检索技术(query by humming, QBH)可有效地解决上述问题。哼唱检索是一种基于内容的音乐检索方式,它通过用户哼唱来进行检索;相对于传统的使用歌名、演唱者等外在信息的检索方式,它是根据音乐的旋律、节奏等内在特征来进行检索。因此相对于传统的关键字形式的用户接口,哼唱检索使用户能得到更佳搜索体验。

哼唱检索以其独特的以用户体验为核心的检索方式及自身所拥有的商业发展潜力,已经被越来越多

收稿日期:2007-05-18。基金项目:国家自然科学基金(60673100)。王小凤,女,1979年生,博士研究生,主要研究方向为音频处理、数据挖掘、模式识别、图形图像处理等。周明全,男,1954年生,硕士,教授,博士生导师,主要研究方向为可视化技术、图形图像处理、数据挖掘等。耿国华,女,1955年生,博士,教授,博士生导师,CCF高级会员,主要研究方向为智能信息处理、数据库、数据挖掘等。郭红波,男,1978年生,硕士研究生,主要研究方向为音频处理、智能挖掘等。

的人所关注. 1995 年, Ghias 等<sup>[1]</sup>展示了首个 QBH 系统, 该系统将歌曲转换为基音轮廓信息进行匹配, 使用三步轮廓法——上升(up, U)、下降(down, D)和不变(same, S)来描述音符间的音程, 并用记录音程信息的字符串来表示音乐的旋律轮廓. 一段旋律中的字符表示当前音符与其前面音符的比较, S 表示音调的重复, U 表示比其前面音符的音调高, D 表示比其前面音符的音调低. McNab 等<sup>[2-3]</sup>增加了对音乐节奏信息的提取, 以提高检索成功率. 上述检索系统都使用字符串匹配的相似度方法进行结果排序. Blackburn 等<sup>[4]</sup>, Rolland 等<sup>[5]</sup>发展了 McNab 的方法, 使用基于树的数据库搜索技术提高检索精度和速度. Kosugi 等<sup>[6-7]</sup>使用欧氏距离进行搜索, 且其中用户输入及数据库内容都被分割为固定长度窗口. Hu 等<sup>[8]</sup>在旋律提取中也使用了固定长度窗口技术. Shih 等<sup>[9]</sup>在其 QBH 系统中使用了隐马尔科夫模型(hide Markov model, HMM), 这项技术已经被成功地应用到语音识别等领域. Pardo 等<sup>[10]</sup>将 2 种不同的相似度计算方法用于哼唱检索, 一是使用距离估计目标和数据库中数据的差异; 二是将数据库中的旋律序列看成 HMM 形式, 而输入数据看作观察序列, 只有当某个目标旋律的 HMM 看起来能够产生此观察序列时, 才认为此目标旋律和观察序列相匹配. Lu 等<sup>[11]</sup>提出一种新的旋律字符串——使用基音轮廓、基音间隔和基音长度的组合, 该系统将字符串匹配和动态规划相结合, 用层次化的搜索方法来提高检索结果. Zhu 等<sup>[12]</sup>动态时间规整(dynamic time warping, DTW)索引技术将演唱歌曲直接与数据库中的歌曲进行比较. 各系统的解决方案如表 1 所示.

表 1 哼唱检索技术发展总结

文献	特征提取技术	搜索技术	时间
[1]	基音轮廓(U, D, S)	字符串匹配	1995
[2-3]	基音轮廓(U, D, S)、音长	字符串匹配	1996, 2000
[4-5]	基音轮廓(U, D, S)	基于树的搜索	1998, 1999
[11]	基音轮廓、基音间隔和基音长度的组合	字符串匹配和动态规划相结合	2001
[6-8, 12]	固定窗口长度基音信息	动态规划	1999—2003
[10]	Midi, HMM	相似度	2004

传统的哼唱检索系统使用哼(鼻子)或者唱(口)固定的符号如: Ba, La, Da 等; 更早的检索系统还需要在用户哼唱过程中加入辅助手段(如节拍器)以便分割音符. 本文提出的哼唱检索系统——Music Angel 不用哼唱固定符号, 更不用加入辅助手段, 用户直接哼一段音乐或唱一段歌谱, 系统就会返回查找结果. 实验证明, 即便在含有一定噪声的环境下, 系统仍然能够保持较好的成功率.

## 1 系统结构

一个鲁棒性好、搜索结果准确的哼唱检索系统, 需要有乐理、信号处理、概率论及搜索方法等多方面技术的支持. 图 1 所示为 Music Angel 的系统结构图.

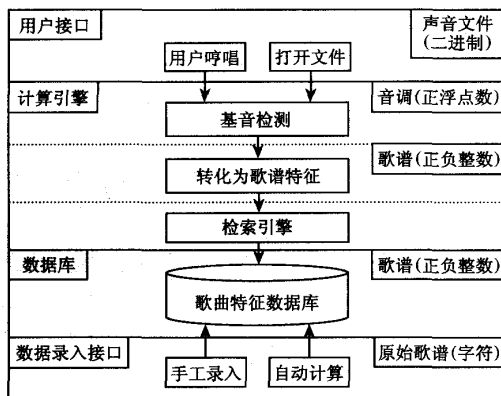


图 1 Music Angel 系统结构

图 1 分为左、中、右 3 个部分. 左侧是系统组织结构划分, 中间为组成系统的各模块, 右侧描述音乐数据在系统各模块下的状态和类型.

歌曲特征数据库是 Music Angel 系统的基础, 它和计算引擎的设计是密切相关的. 数据库中存储的音乐特征将决定计算引擎的设计; 相反, 计算引擎也可以决定数据库中需要存储的音乐特征. 计算引擎是系统的核心部分, 它首先将待检索的声音文件(现场哼唱得到的文件或已有哼唱文件)转化为整数形式的歌谱特征, 然后检索引擎计算该特征与数据库中每条记录的相似度, 最后返回相似度较高的几首歌曲. 图 1 中数据录入接口可采用手工方式或编写采集程序, 将数据录入到数据库中. 用户接口为用户哼唱或打开声音文件, 用户哼唱是使用麦克风哼唱得到数据, 打开文件是打开已有歌曲或哼唱片段获取数据.

## 2 关键技术

Music Angel 系统的关键技术是歌谱轮廓提取算法和标准音调差值图的生成。其中,歌谱轮廓提取算法是系统的基础和核心,用于将二进制声音文件转换为歌谱特征,以便检索;标准音调差值图是对标准音调进行计算后得出的规律图。

### 2.1 歌谱轮廓提取算法

歌谱信息是本文系统检索目标,歌谱轮廓特征从基音轮廓特征转换而来,而基音是哼唱检索的基础。哼唱过程中的音调和每一时刻的声音振动频率相关联:频率越大,音调越高;频率越小,音调越低。这种振动频率也称作声音信号中的基音频率。基音提取的方法大致可分为时域法、频域法和混合法 3 种。时域法包括过零率法、短时自相关函数法、平均幅度差函数法和并行处理技术法<sup>[13-15]</sup>等;频域法包括频率极值点检测法、简单逆滤波器跟踪法、小波函数法和 FIR 滤波器法<sup>[16-17]</sup>等;此外还有倒谱法,即时频率法、Comb 变换法<sup>[18-20]</sup>和基于 Hilbert-Huang 变换的基音周期检测方法<sup>[21]</sup>等多种方法。各种算法在运算量、精确度和抗噪性能等方面不尽相同,每种算法都有其固有的局限性<sup>[15]</sup>。通常,时域相关法易产生倍基音误差,而频域相关法易产生次基音误差,因此人们提出了各种改进和综合算法,以提高算法精确度和对各种环境下基音提取的健壮性。

本文使用文献<sup>[19]</sup>中的方法提取基因序列,首先找到声音频谱上的各个峰值点,再估计基音频率。对一哼唱片段,加窗时间长度为 20 ms,窗口偏移设为 10 ms。若一片断采样率用 Rate 表示,单位为 Hz,则每帧数据长度可设为  $\text{Rate} \times 0.02$ <sup>[22]</sup>;然后对此数据作快速傅里叶变换。但即使在高采样率(如 44 100 kHz 下,每个窗口的数据长度也仅为 4 096,对 4 096 个数据作快速傅里叶变换,其精度显然是不足的。本文对每帧数据扩展 3 倍长度,如 4 096 个点扩展为 16 384 个点,扩展数据全部补 0,然后再作快速傅里叶变换。

对于周期为  $T$  的信号  $s(t)$ ,可分解为正弦级数

$$s(t) = \sum_{k=1}^{\infty} a_k \sin(2\pi kft + \beta_k) \quad (1)$$

其中  $f=1/T$ ,为  $s(t)$  的频率。那么对  $s(t)$  进行快速傅里叶变换再取模后,在  $f, 2f, 3f, \dots$  处将会出现峰值。取所有峰值的最大公约数作为最后的基音检测结果,即所有的峰值都是它的倍频。

**算法 1. 歌谱轮廓提取算法 (score contour extract algorithm, SCEA)**

Step1. 基音估计。求得一个音乐的基音序列  $S_n$ ,  $n$  为序列长度。

Step2. 初步分段。对  $S_n$  按大小进行分段,相同的值分成一段,各段分别记为  $L_m$ ,其值大小记为  $L_m$ ,  $m$  为基音段总个数 ( $1 \leq m \leq n$ )。

Step3. 基音孤立段(点)剔除。

实验中发现,经初步分段计算后得到的基音段序列  $L_m$  中常常含有一些“野点”或“孤立段”,即该段长度很短,一般为 1 或 2,但基音值与前后基音段值相差明显。文献<sup>[23]</sup>使用差值平滑处理算法对基音检测结果进行后处理。本文处理过程如下:

设阈值  $K$ ,某“孤立段”为  $L_i$  ( $1 < i < m$ ),记  $\Delta\text{Left} = |L_i - L_{i-1}|$ ,  $\Delta\text{Right} = |L_i - L_{i+1}|$ ,则

若  $\Delta\text{Left} > K$  and  $\Delta\text{Right} > K$ ,该段(点)剔出。

若  $\Delta\text{Left} > K$  and  $\Delta\text{Right} < K$ ,该段(点)并入  $L_{i+1}$ ;反之,并入  $L_{i-1}$ 。

若  $\Delta\text{Left} < K$  and  $\Delta\text{Right} < K$ ,但  $\Delta\text{Left} > \Delta\text{Right}$ ,则该段(点)并入  $L_{i+1}$ ;反之,并入  $L_{i-1}$ 。

$L_m$  经上述操作后记为  $L_v$ ,各段值为  $L_v$ ,  $v$  为当前基音段总个数 ( $1 \leq v \leq m$ )。

Step4. 基音序列合并。对  $L_v$  再次进行合并,步骤如下:

设阈值  $P$ ,对任意段  $L_i$  ( $1 < i < v$ ),记  $\Delta\text{Left} = |L_i - L_{i-1}|$ ,  $\Delta\text{Right} = |L_i - L_{i+1}|$ ,令函数  $f(L_i)$  表示段  $L_i$  的长度,则

若  $\Delta\text{Left} > P$  and  $\Delta\text{Right} < P$ ,则  $L_i$  和  $L_{i+1}$  合并;反之,  $L_i$  和  $L_{i-1}$  合并。

若  $\Delta\text{Left} < P$  and  $\Delta\text{Right} < P$ ,并且  $\Delta\text{Left} > \Delta\text{Right}$ ,则  $L_i$  和  $L_{i+1}$  合并;反之,  $L_i$  和  $L_{i-1}$  合并。

合并后产生的新基音段基音值按以下方法计算:记合并的 2 段分别为  $L_i$  和  $L_j$ ,合并后新段为  $L_t$ ,则

若  $f(L_i) = f(L_j)$ ,  $L_t = (L_i + L_j)/2$ 。

若  $f(L_i) > f(L_j)$ ,  $L_t = L_i$ ;反之,  $L_t = L_j$ 。

$L_v$  经上述操作后记为  $L_w$ ,各段值为  $L_w$ ,  $w$  为当前基音段总个数 ( $1 \leq w \leq v$ )。

图 2 所示为用户哼唱 7 个基本音级 do, re, mi, fa, sol, la, si(或 ti)的时域波形图和提取的基音序列图。图 2b 所示为用户在哼唱基本音级后的基音计算结果,观察此图不难发现,基本音级的音调值呈现出阶梯状上升趋势,正好与人哼唱这 7 个基本音级的频率逐步上升相对应。由此可证明,本文系统中的音调提取算法是非常有效的。

Step5. 计算基音轮廓。每个人的发声器官都是独一无二的,因而同一首歌不同人演唱其频率很难相同,但是演唱者都会在演唱过程中尽量使发音的高低起伏与歌谱一致,从而可采用基音变化,即基音轮廓作为搜索准则。也就是说,音调之间的变化比音调本身的值更重要,一首歌曲排除节奏

后,歌谱之间的变化才是这首歌的本质.

对  $L_w$  中任意连续段  $L_i, L_{i+1}(1 \leq i < w)$ , 定义其连续递增序列为  $\Delta L_z(1 < z < w)$ , 并记  $\Delta L_z = L_{i+1} - L_i, \Delta L_z$  就是音调的轮廓序列.

对于 Step4 中所演唱的 7 个基本音级而言,得到的基音轮廓为 1.690 350 43, 1.787 060 10, 1.071 208 06, 1.650 042 28,

1.506 370 59 和 1.166 866 95.

Step6. 将基音轮廓转换为歌谱轮廓. 歌谱轮廓就是歌谱中两两音符之间的差异值. 有了  $\Delta L_z$  的基音轮廓, 再根据第 2.2 节介绍的标准音调差值图或表, 就可以将其转换为歌谱轮廓. 然后将此轮廓序列与数据库中存储的歌曲歌谱轮廓特征进行比较, 选取相似度较高的歌曲作为检索结果.

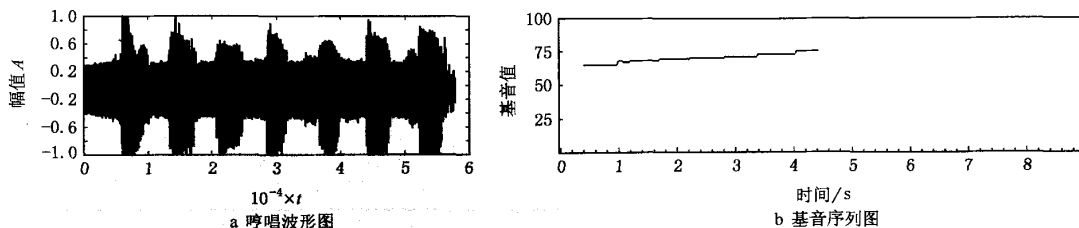


图2 哼唱音乐片段的波形图和基音序列图

## 2.2 标准音调差值图的生成

GUITAR PRO<sup>①</sup>是 AROBAS 公司开发的一个音乐软件, 它可模拟吉他或钢琴键盘输入音符, 生成 Midi 序列, 并可导出为 wav 格式文件. 本文用 GUITAR PRO 生成标准的 7 个基本音级的音调旋律, 再对此旋律使用 SCEA 算法计算音调. 实验过程中发现了标准音调的差值规律. 计算步骤如下:

Step1. 使用 Guitar Pro 4.0 生成基本音级音调旋律(do,

re, mi, fa, sol, la, si) 的 wav 文件. 发音器选用钢琴, 图 3a 所示为时域波形图.

Step2. 使用 SCEA 算法对基本音级音调旋律进行计算, 得到标准基音序列, 如图 3b 所示.

Step3. 计算标准基音序列两两之间的增量序列, 得到基音轮廓. 此轮廓序列将作为其他哼唱音乐片段进行基音轮廓计算的量化标准. 经过多次实验验证后, 得到标准音调差值图, 如图 3c 所示.

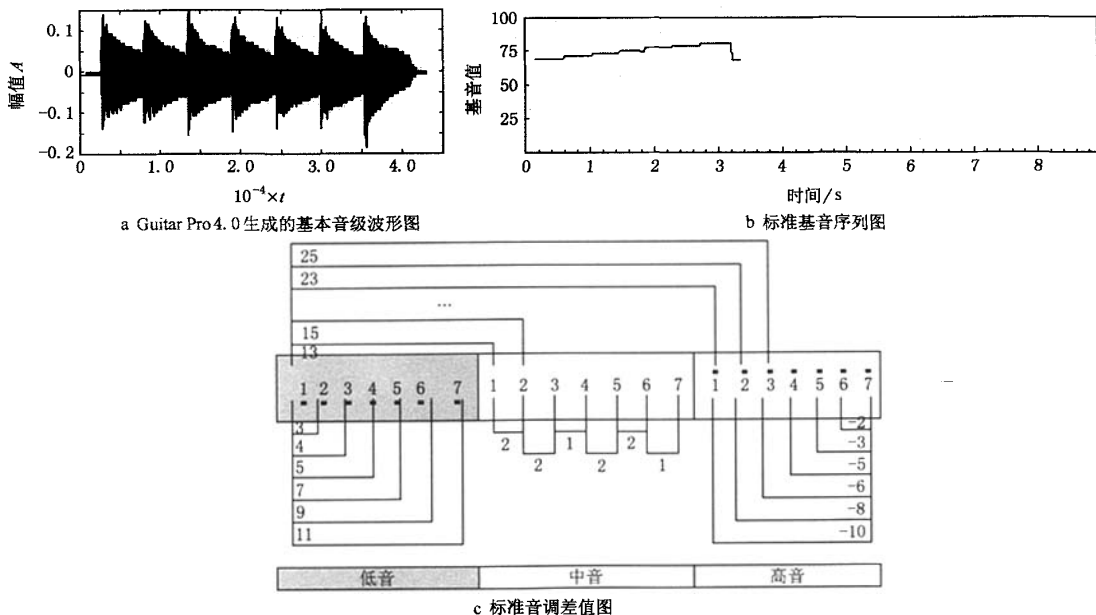


图3 标准音调差值图的生成

图 3c 分为左、中、右 3 个部分. 左侧是低音, 中间是中音, 右侧是高音. 图中所有运算结果均为经 SCEA 算法计算后取整所得; 按从左向右的先后顺序, 正值代表线段两端排列在后的音调值减去排列

在前的音调值所得, 负值相反; 对中音部分, 图中标出了相邻 2 个音调之间的差值. 由图 3c 可得到如表

① <http://www.guitar-pro.com/>

2 所示的标准音调音符差值表。

表 2 含义为:当基音差值为  $\Delta n$  时,这 2 个音调间相差音符距离为  $m$ 。由基音差值序列,然后查表 2 可迅速找到相应音符。对算法 1 Step5 出现的基音轮廓数据,查表后得到歌谱轮廓为 111111。

表 2 标准音调音符差值表

基音差值 $\Delta n$	相差音符个数 $m$
1,2	1
3,4	2
5,6	3
6,7,8	4
8,9	5
10	6
...	...

经过第 2.1,2.2 节后,实现了从哼唱音乐片段到歌谱轮廓的转换,因而数据库中存储歌曲的歌谱轮廓即可,然后使用字符串匹配搜索引擎得出相似度最好的若干首歌曲,完成检索过程。

3 系统实现及性能测试

3.1 系统实现

Step1. 创建音乐特征数据库。编写数据录入程序,将音乐数据输入后,该程序将歌曲基本信息和歌谱轮廓特征写入数据库。

Step2. 创建标准音调差值表,供 SCEA 算法使用。

Step3. 实现 Music Angel 系统。编写基于 SCEA 算法的 GUI 界面应用程序,其中用户接口采用文件输入或现场哼唱 2 种方式;根据本文提取的哼唱特征的特点,搜索算法使用基于动态规划的字符串匹配算法。图 4 所示 Music Angle 系统的开发平台为:Windows OS;Visual Studio.net 2003 开发环境;SQL Server2000 数据库。



图 4 系统哼唱检索界面

3.2 性能评测

系统性能测试中,我们邀请了 8 位测试人员(4 男 4 女)进行实验,每人随机演唱数据库中的 6 首歌曲,演唱方式为演唱歌谱、使用鼻子哼唱或使用“LaLa...”哼唱,分别统计了总成功率(如表 3 所示)和 3 种演唱方式的不同成功率(如表 4 所示)。

表 3 总实验结果

搜索结果位置	精度/%
前 3 位	86
前 5 位	90
前 10 位	94

表 4 “Music Angel”对不同演唱方式的前 3 位成功率统计结果

演唱方式	总次数	平均成功率/%
演唱歌谱	16	93
使用“LaLa...”声	14	86
鼻子哼唱	18	80

实验结果表明:演唱歌谱的效果最好,其次是使用“LaLa...”声,最差的是鼻子哼,这与我们的预测完全相同:只要能将旋律的高低差异表达出来,就能取得良好的效果。显然,大多数人演唱歌谱都能将歌谱差异表达得比较准确,而哼唱时往往不能准确地表达歌谱的旋律。

实验还发现,检索结果相似度相同的很多;要求返回前 3 位时,返回结果往往多于 3 首。主要是由于本文仅使用歌谱轮廓进行检索,并没有考虑歌曲节奏信息。由于数据库很大,存在大量重复或相似度很高的片断,就会出现上述结果,因此进一步的工作就是添加节奏等信息作为特征,以期提高检索精度。

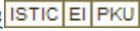
4 结 论

本文提出了一种直接使用歌谱信息进行匹配的哼唱检索方法,并在此基础上实现了 Music Angel 系统。该系统对用户哼唱的一段旋律可检索出目标歌曲。与已有方法相比,Music Angel 系统在以下方面做出了创新或改进:

- 1) 直接使用歌谱信息作为搜索目标,在大型音乐数据库的创建及搜索速度上将会获得优势。
- 2) 给出了标准音调的差值图,提出了使用标准音调对用户演唱音调进行量化的方法。实验表明,该方法具有良好效果。

## 参 考 文 献

- [1] Ghias A, Logan J, Chamberlin D, *et al.* Query by humming: musical information retrieval in an audio database [C] //Proceedings of ACM International Conference on Multimedia, San Francisco, 1995: 231-236
- [2] McNab R J, Smith L A, Witten I H, *et al.* Towards the digital music library: tune retrieval from acoustic input [C] //Proceedings of ACM International Conference on Digital Libraries, Bethesda, 1996: 11-18
- [3] McNab R J, Smith L A, Witten I H, *et al.* Tune retrieval in multimedia library [J]. *Multimedia Tools And Applications*, 2000, 10(4): 113-132
- [4] Blackburn S, Roure D D. A tool for content based navigation of music [C] //Proceedings of ACM International Conference on Multimedia, Bristol, 1998: 361-368
- [5] Rolland P Y, Raskinis G, Ganascia J G. Music content-based retrieval: an overview of Melodiscov approach and systems [C] //Proceedings of ACM International Conference on Multimedia, Orlando, 1999: 81-84
- [6] Kosugi N, Nishihara Y, Sakata T, *et al.* Music retrieval by humming [C] //Proceedings of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Victoria, 1999: 404-407
- [7] Kosugi N, Nishihara Y, Sakata T, *et al.* A practical query-by-humming system for a large music database [C] //Proceedings of ACM International Conference on Multimedia, Marina Del Rey, California, 2000: 333-342
- [8] Hu N, Dannenberg R B. A comparison of melodic database retrieval techniques [C] //Proceedings of ACM International Conference on Digital Libraries, Portland, Oregon, 2002: 301-307
- [9] Shih H-H, Narayanan S, Kuo C-C J. An hmm-based approach to humming transcription [C] //Proceedings of IEEE International Conference on Multimedia and Expo, Lausanne, 2002: 337-340
- [10] Pardo B, Shifrin Jonah, Birmingham William. Name that tune: a pilot study in finding a melody from a sung query [J]. *Journal of the American Society for Information Science and Technology*, 2004, 55(4): 283-300
- [11] Lu L, You H, Zhang H-J. A new approach to query by humming in music retrieval [C] //Proceedings of IEEE International Conference on Multimedia and Expo, Tokyo, 2001: 776-779
- [12] Zhu Y, Shasha D. Warping indexes with envelope transforms for query-by-humming [C] //Proceedings of SIGMOD, San Diego, CA, 2003: 181-192
- [13] Cheng Y M, O' Shaughnessy D. Automatic and reliable estimation of glottal closure instant and period [J]. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1989, 37(12): 1805-1815
- [14] Ross M J, Shafer H L, Cohen A, *et al.* Average magnitude difference function pitch extractor [J]. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1974, 22(5): 353-362
- [15] Rabiner L R, Cheng M J, Rosenberg A E, *et al.* A comparative performance study of several pitch detection algorithms [J]. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1976, 24(5): 399-418
- [16] Hu Jianling, Xu Sheng, Chen Jian. A modified pitch detection algorithm [J]. *IEEE Communications Letters*, 2001, 5(2): 64-66
- [17] Ghaemmaghami S, Deriche M, Boashash B. A new approach to pitch and voicing detection through spectrum periodicity measurement [C] //Proceedings of TENCON'97, IEEE Region 10 Annual Conference, Speech and Image Technologies for Computing and Telecommunications, Brisbane, 1997, 2: 743-746
- [18] Huici Hernandez-Diaz, Ginori Lorenzo. Combined algorithm for pitch detection of speech signals [J]. *Electronics Letters*, 1995, 31(1): 15-16
- [19] Abe T, Kobayashi T, Imai S. Harmonics tracking and pitch extraction based on instantaneous frequency [C] //Proceedings of International Conference on Acoustics, Speech, and Processing, Detroit, 1995, 1: 756-759
- [20] Janer L. New pitch detection algorithm based on wavelet transform [C] //Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, 1998: 165-168
- [21] Yang Zhihua, Qi Dongxu, Yang Lihua. Detecting pitch period based on Hilbert-Huang transform [J]. *Chinese Journal of Computers*, 2006, 29(1): 106-115 (in Chinese)  
(杨志华, 齐东旭, 杨力华. 一种基于 Hilbert-Huang 变换的基音周期检测新方法[J]. *计算机学报*, 2006, 29(1): 106-115)
- [22] Unal E, Narayanan S, Chew E, *et al.* A dictionary based approach for robust and syllable-independent audio input transcription for query by humming systems [C] //Proceedings of Audio and Music Computing Multimedia, Santa Barbara, 2006: 37-43
- [23] Zhao Li. The processing of voice signal [M]. Beijing: China Machine Press, 2003: 65-74 (in Chinese)  
(赵 力. 语音信号处理[M]. 北京: 机械工业出版社, 2003: 65-74)

作者: 王小凤, 周明全, 耿国华, 郭红波, Wang Xiaofeng, Zhou Mingquan, Geng Guohua, Guo Hongbo  
作者单位: 王小凤, 耿国华, 郭红波, Wang Xiaofeng, Geng Guohua, Guo Hongbo(西北大学信息科学与技术学院, 西安, 710069), 周明全, Zhou Mingquan(西北大学信息科学与技术学院, 西安, 710069; 北京师范大学信息科学与技术学院, 北京, 100875)  
刊名: 计算机辅助设计与图形学学报   
英文刊名: JOURNAL OF COMPUTER-AIDED DESIGN & COMPUTER GRAPHICS  
年, 卷(期): 2007, 19(7)  
被引用次数: 7次

参考文献(24条)

1. Ghias A; Logan J; Chamberlin D [Query by humming: musical information retrieval in an audio database](#) 1995
2. McNab R J; Smith L A; Witten I H [Towards the digital music library: tune retrieval from acoustic input](#) 1996
3. McNab R J; Smith L A; Witten I H [Tune retrieval in multimedia library](#)[外文期刊] 2000(04)
4. Blackburn S; Roure D D [A tool for content based navigation of music](#) 1998
5. Rolland P Y; Raskinis G; Ganascia J G [Music content-based retrieval: an overview of Melodiscov approach and systems](#) 1999
6. Kosugi N; Nishihara Y; Sakata T [Music retrieval by humming](#) 1999
7. Kosugi N; Nishihara Y; Sakata T [A practical query-by-humming system for a large music database](#) 2000
8. Hu N; Dannenberg R B [A comparison of melodic database retrieval techniques](#) 2002
9. Shih H -H; Narayanan S; Kuo C -C J [An hmm-based approach to humming transcription](#)[外文会议] 2002
10. Pardo B; Shifrin Jonah; Birmingham William [Name that tune: a pilot study in finding a melody from a sung query](#)[外文期刊] 2004(04)
11. Lu L; You H; Zhang H -J [A new approach to query by humming in music retrieval](#) 2001
12. Zhu Y; Shasha D [Warping indexes with envelope transforms for query-by-humming](#) 2003
13. Cheng Y M; O' Shaghnesy D [Automatic and reliable estimation of glottal closure instant and period](#) [外文期刊] 1989(12)
14. Ross M J; Shafer H L; Cohen A [Average magnitude difference function pitch extractor](#) 1974(05)
15. Rabiner L R; Cheng M J; Rosenberg A E [A comparative performance study of several pitch detection algorithms](#) 1976(05)
16. Hu Jianling; Xu Sheng; Chen Jian [A modified pitch detection algorithm](#) 2001(02)
17. Ghaemmaghami S; Deriche M; Boashash B [A new approach to pitch and voicing detection through spectrum periodicity measurement](#)[外文会议] 1997
18. Huici Hernandez-Diaz; Ginori Lorenzo [Combined algorithm for pitch detection of speech signals](#)[外文期刊] 1995(01)
19. Abe T; Kobayashi T; Imai S [Harmonics tracking and pitch extraction based on instantaneous frequency](#) [外文会议] 1995
20. Janer L [New pitch detection algorithm based on wavelet transform](#)[外文会议] 1998