



(12)发明专利

(10)授权公告号 CN 104978962 B

(45)授权公告日 2019.01.18

(21)申请号 201410149146.2

G10L 15/02(2006.01)

(22)申请日 2014.04.14

(56)对比文件

(65)同一申请的已公布的文献号
申请公布号 CN 104978962 A

CN 102375834 A, 2012.03.14, 说明书第0008-0032、0044-0076、0085-0102、0112-0123段以及附图1-8.

(43)申请公布日 2015.10.14

CN 1983253 A, 2007.06.20, 全文.

(73)专利权人 科大讯飞股份有限公司
地址 230088 安徽省合肥市高新开发区望江西路666号

CN 101471068 A, 2009.07.01, 全文.

CN 101916250 A, 2010.12.15, 全文.

CN 102841932 A, 2012.12.26, 全文.

(72)发明人 吴奎 蒋成林 刘丹 魏思
胡国平 胡郁 刘庆峰

US 2005/0086052 A1, 2005.04.21, 全文.

李珂 等. 基于音频检索的点歌系统. 《北京师范大学学报(自然科学版)》. 2006, 第42卷(第4期), 第383-386页.

(74)专利代理机构 北京维澳专利代理有限公司
11252
代理人 王立民 张应

审查员 李召卿

(51) Int. Cl.

G10L 15/08(2006.01)

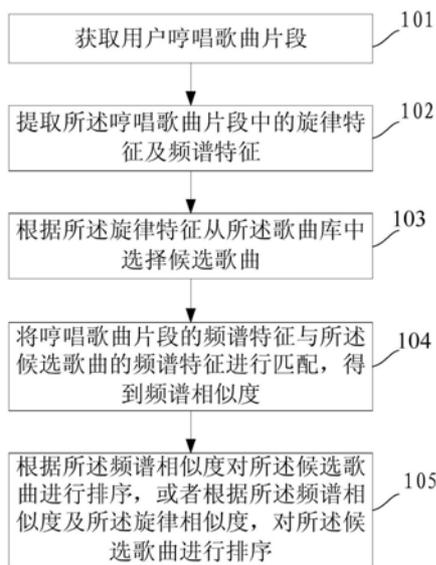
权利要求书3页 说明书9页 附图4页

(54)发明名称

哼唱检索方法及系统

(57)摘要

本发明涉及音乐检索技术领域,公开了一种哼唱检索方法及系统,该方法包括:获取用户哼唱歌曲片段;提取所述哼唱歌曲片段的旋律特征及频谱特征;根据所述旋律特征从所述歌曲库中选择候选歌曲;将所述频谱特征与所述候选歌曲的频谱特征进行匹配,得到频谱相似度;根据所述频谱相似度、或者根据所述频谱相似度及所述旋律相似度,对所述候选歌曲进行排序。本发明可以有效提高检索结果的准确率。



1. 一种哼唱检索方法,其特征在于,包括:
 - 获取用户哼唱歌曲片段;
 - 提取所述哼唱歌曲片段的旋律特征及频谱特征;
 - 根据所述旋律特征从歌曲库中选择候选歌曲,包括将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配,得到旋律相似度;根据所述旋律相似度从所述歌曲库中选择候选歌曲;
 - 将所述频谱特征与所述候选歌曲的频谱特征进行匹配,得到频谱相似度;
 - 根据所述频谱相似度、或者根据所述频谱相似度及所述旋律相似度,对所述候选歌曲进行排序。
2. 根据权利要求1所述的方法,其特征在于,所述旋律特征包括:音符序列表示的旋律特征或基频包络表示的旋律特征;
 - 所述将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配,得到旋律相似度包括:
 - 对于音符序列表示的旋律特征,利用序列匹配算法计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度;
 - 对于基频包络表示的旋律特征,利用动态时间规整算法计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度。
3. 根据权利要求1所述的方法,其特征在于,所述旋律特征包括:音符序列表示的旋律特征和基频包络表示的旋律特征;
 - 所述将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配,得到旋律相似度包括:
 - 计算所述哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列中各子音符序列的相似度;
 - 选择待匹配歌曲的音符序列中与所述哼唱歌曲片段的音符序列相似度最大的子音符序列,并确定所述子音符序列的起止时间点;
 - 计算所述哼唱歌曲片段的基频包络与所述待匹配歌曲在所述起止时间点内的基频包络的相似度;
 - 将所述基频包络的相似度作为旋律相似度;或者
 - 将所述基频包络的相似度与对应的最大子音符序列的相似度进行融合,并将融合结果作为旋律相似度。
4. 根据权利要求1所述的方法,其特征在于,所述根据所述旋律相似度从所述歌曲库中选择候选歌曲包括:
 - 选择旋律相似度大于设定阈值的歌曲作为候选歌曲;或者
 - 按照旋律相似度由大到小的顺序选择设定个数的歌曲作为候选歌曲。
5. 根据权利要求1所述的方法,其特征在于,所述旋律特征包括:音符序列表示的旋律特征和基频包络表示的旋律特征;
 - 所述根据所述旋律特征从所述歌曲库中选择候选歌曲包括:
 - 计算所述哼唱歌曲片段的音符序列与歌曲库中歌曲的音符序列的相似度;
 - 根据所述音符序列的相似度从所述歌曲库中选择初选歌曲;
 - 计算所述哼唱歌曲片段的基频包络与所述初选歌曲的基频包络的相似度;
 - 根据所述基频包络的相似度、或者根据所述基频包络的相似度及所述音符序列的相似度,对所述初选歌曲进行裁剪,得到候选歌曲。

6. 根据权利要求1至5任一项所述的方法,其特征在于,所述将所述频谱特征与所述候选歌曲的频谱特征进行匹配,得到频谱相似度包括:

采用动态时间规整算法计算所述频谱特征与所述候选歌曲的频谱特征的频谱相似度。

7. 一种哼唱检索系统,其特征在于,包括:

获取模块,用于获取用户哼唱歌曲片段;

特征提取模块,用于提取所述哼唱歌曲片段中的旋律特征及频谱特征;

选择模块,用于根据所述旋律特征从歌曲库中选择候选歌曲;所述选择模块包括:旋律特征匹配模块,用于将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配,得到旋律相似度;候选模块,用于根据所述旋律相似度从所述歌曲库中选择候选歌曲;

频谱特征匹配模块,用于将所述频谱特征与所述候选歌曲的频谱特征进行匹配,得到频谱相似度;

排序模块,用于根据所述频谱相似度、或者根据所述频谱相似度及所述旋律相似度,对所述候选歌曲进行排序。

8. 根据权利要求7所述的系统,其特征在于,所述旋律特征包括:音符序列表示的旋律特征或基频包络表示的旋律特征;

所述旋律特征匹配模块,具体用于对于音符序列表示的旋律特征,利用序列匹配算法计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度;或者对于基频包络表示的旋律特征,利用动态时间规整算法计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度。

9. 根据权利要求7所述的系统,其特征在于,所述旋律特征包括:音符序列表示的旋律特征和基频包络表示的旋律特征;

所述旋律特征匹配模块包括:

第一计算单元,用于计算所述哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列中各子音符序列的相似度;

选择单元,用于选择待匹配歌曲的音符序列中与所述哼唱歌曲片段的音符序列相似度最大的子音符序列;

确定单元,用于确定所述相似度最大的子音符序列的起止时间点;

第二计算单元,用于计算所述哼唱歌曲片段的基频包络与所述待匹配歌曲在所述起止时间点内的基频包络的相似度;

输出单元,用于将所述基频包络的相似度作为旋律相似度输出。

10. 根据权利要求9所述的系统,其特征在于,所述旋律特征匹配模块还包括:

融合单元,用于将所述基频包络的相似度与对应的最大子音符序列的相似度进行融合;

所述输出单元,用于将所述融合单元得到的融合结果作为旋律相似度输出。

11. 根据权利要求7所述的系统,其特征在于,

所述选择模块,具体选择旋律相似度大于设定阈值的歌曲作为候选歌曲;或者按照旋律相似度由大到小的顺序选择设定个数的歌曲作为候选歌曲。

12. 根据权利要求7所述的系统,其特征在于,所述旋律特征包括:音符序列表示的旋律特征和基频包络表示的旋律特征;

所述选择模块包括:

音符序列相似度计算单元,用于计算所述哼唱歌曲片段的音符序列与歌曲库中歌曲的音符序列的相似度;

初选单元,用于根据所述音符序列的相似度从所述歌曲库中选择初选歌曲;

基频包络相似度计算单元,用于计算所述哼唱歌曲片段的基频包络与所述初选歌曲的基频包络的相似度;

裁剪单元,用于根据所述基频包络的相似度、或者根据所述基频包络的相似度及所述音符序列的相似度,对所述候选歌曲进行裁剪,得到候选歌曲。

13. 根据权利要求7至12任一项所述的系统,其特征在于,

所述频谱特征匹配模块,具体用于采用动态时间规整算法计算所述频谱特征与所述候选歌曲的频谱特征的频谱相似度。

哼唱检索方法及系统

技术领域

[0001] 本发明涉及音乐检索技术领域,具体涉及一种哼唱检索方法及系统。

背景技术

[0002] 随着信息技术和网络技术的发展,音乐库规模呈几何级数增长,音乐歌曲的快速检索也成为新的研究热点。传统的音乐检索方法主要基于文本关键字检索,根据用户输入的歌曲名、歌手、流派、歌词等文本信息进行相关歌曲检索。而当用户只能哼唱出歌曲某一片段,但无法提供确切的歌曲名、歌手等文本信息时,传统的音乐检索系统显然无法满足这样的检索需求。为此,研究人员提出了音乐哼唱检索技术,即根据用户哼唱的歌曲片断,从音乐库中挑选相应的歌曲。

[0003] 现有的哼唱检索系统大都采用基于歌曲旋律特征匹配的检索方法,即比较用户输入歌曲片断和音乐库中的歌曲在旋律特征上的相似度。具体地,首先从用户输入歌曲片断中提取旋律特征,随后和系统预置的歌曲库中的各歌曲旋律特征模板一一匹配,并选择具有最大相似度的歌曲为期望歌曲。所述歌曲旋律是指连续的有音调(或称为音高)、有节奏的音乐音符序列,旋律特征可以用音符序列或音高曲线的形式表示。

[0004] 现有的基于歌曲旋律特征匹配的检索方法虽然解决了上述问题,但依然存在以下问题:

[0005] 1. 旋律特征音符序列提取中音符分割不准确或用户哼唱的歌曲存在跑调等问题时,提取的旋律信息无法正确表征歌曲特点,影响检索结果的准确率;

[0006] 2. 随着音乐库规模的扩大,歌曲之间的旋律特征区分性减弱,现有的基于歌曲旋律特征匹配的检索方法的检索结果准确率容易受到音乐库规模的影响。

发明内容

[0007] 本发明实施例提供一种哼唱检索方法及系统,以提高检索结果的准确率。

[0008] 为此,本发明提供如下技术方案:

[0009] 一种哼唱检索方法,包括:

[0010] 获取用户哼唱歌曲片段;

[0011] 提取所述哼唱歌曲片段的旋律特征及频谱特征;

[0012] 根据所述旋律特征从所述歌曲库中选择候选歌曲;

[0013] 将所述频谱特征与所述候选歌曲的频谱特征进行匹配,得到频谱相似度;

[0014] 根据所述频谱相似度、或者根据所述频谱相似度及所述旋律相似度,对所述候选歌曲进行排序。

[0015] 优选地,所述根据所述旋律特征从所述歌曲库中选择候选歌曲包括:

[0016] 将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配,得到旋律相似度;

[0017] 根据所述旋律相似度从所述歌曲库中选择候选歌曲。

[0018] 优选地,所述旋律特征包括:音符序列表示的旋律特征或基频包络表示的旋律特

征；

[0019] 所述将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配，得到旋律相似度包括：

[0020] 对于音符序列表示的旋律特征，利用序列匹配算法计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度；

[0021] 对于基频包络表示的旋律特征，利用动态时间规整算法计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度。

[0022] 优选地，所述旋律特征包括：音符序列表示的旋律特征和基频包络表示的旋律特征；

[0023] 所述将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配，得到旋律相似度包括：

[0024] 计算所述哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列中各子音符序列的相似度；

[0025] 选择待匹配歌曲的音符序列中与所述哼唱歌曲片段的音符序列相似度最大的子音符序列，并确定所述子音符序列的起止时间点；

[0026] 计算所述哼唱歌曲片段的基频包络与所述待匹配歌曲在所述起止时间点内的基频包络的相似度；

[0027] 将所述基频包络的相似度作为旋律相似度；或者

[0028] 将所述基频包络的相似度与对应的最大子音符序列的相似度进行融合，并将融合结果作为旋律相似度。

[0029] 优选地，所述根据所述旋律相似度从所述歌曲库中选择候选歌曲包括：

[0030] 选择旋律相似度大于设定阈值的歌曲作为候选歌曲；或者

[0031] 按照旋律相似度由大到小的顺序选择设定个数的歌曲作为候选歌曲。

[0032] 优选地，所述旋律特征包括：音符序列表示的旋律特征和基频包络表示的旋律特征；

[0033] 所述根据所述旋律特征从所述歌曲库中选择候选歌曲包括：

[0034] 计算所述哼唱歌曲片段的音符序列与歌曲库中歌曲的音符序列的相似度；

[0035] 根据所述音符序列的相似度从所述歌曲库中选择初选歌曲；

[0036] 计算所述哼唱歌曲片段的基频包络与所述初选歌曲的基频包络的相似度；

[0037] 根据所述基频包络的相似度、或者根据所述基频包络的相似度及所述音符序列的相似度，对所述初选歌曲进行裁剪，得到候选歌曲。

[0038] 优选地，所述将所述频谱特征与所述候选歌曲的频谱特征进行匹配，得到频谱相似度包括：

[0039] 采用动态时间规整算法计算所述频谱特征与所述候选歌曲的频谱特征的频谱相似度。

[0040] 一种哼唱检索系统，包括：

[0041] 获取模块，用于获取用户哼唱歌曲片段；

[0042] 特征提取模块，用于提取所述哼唱歌曲片段中的旋律特征及频谱特征；

[0043] 选择模块，用于根据所述旋律特征从所述歌曲库中选择候选歌曲；

- [0044] 频谱特征匹配模块,用于将所述频谱特征与所述候选歌曲的频谱特征进行匹配,得到频谱相似度;
- [0045] 排序模块,用于根据所述频谱相似度、或者根据所述频谱相似度及所述旋律相似度,对所述候选歌曲进行排序。
- [0046] 优选地,所述选择模块包括:
- [0047] 旋律特征匹配模块,用于将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配,得到旋律相似度;
- [0048] 候选模块,用于根据所述旋律相似度从所述歌曲库中选择候选歌曲。
- [0049] 优选地,所述旋律特征包括:音符序列表示的旋律特征或基频包络表示的旋律特征;
- [0050] 所述旋律特征匹配模块,具体用于对于音符序列表示的旋律特征,利用序列匹配算法计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度;或者对于基频包络表示的旋律特征,利用动态时间规整算法计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度。
- [0051] 优选地,所述旋律特征包括:音符序列表示的旋律特征和基频包络表示的旋律特征;
- [0052] 所述旋律特征匹配模块包括:
- [0053] 第一计算单元,用于计算所述哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列中各子音符序列的相似度;
- [0054] 选择单元,用于选择待匹配歌曲的音符序列中与所述哼唱歌曲片段的音符序列相似度最大的子音符序列;
- [0055] 确定单元,用于确定所述相似度最大的子音符序列的起止时间点;
- [0056] 第二计算单元,用于计算所述哼唱歌曲片段的基频包络与所述待匹配歌曲在所述起止时间点内的基频包络的相似度;
- [0057] 输出单元,用于将所述基频包络的相似度作为旋律相似度输出。
- [0058] 优选地,所述旋律特征匹配模块还包括:
- [0059] 融合单元,用于将所述基频包络的相似度与对应的最大子音符序列的相似度进行融合;
- [0060] 所述输出单元,用于将所述融合单元得到的融合结果作为旋律相似度输出。
- [0061] 优选地,所述选择模块,具体选择旋律相似度大于设定阈值的歌曲作为候选歌曲;或者按照旋律相似度由大到小的顺序选择设定个数的歌曲作为候选歌曲。
- [0062] 优选地,所述旋律特征包括:音符序列表示的旋律特征和基频包络表示的旋律特征;
- [0063] 所述选择模块包括:
- [0064] 音符序列相似度计算单元,用于计算所述哼唱歌曲片段的音符序列与歌曲库中歌曲的音符序列的相似度;
- [0065] 初选单元,用于根据所述音符序列的相似度从所述歌曲库中选择初选歌曲;
- [0066] 基频包络相似度计算单元,用于计算所述哼唱歌曲片段的基频包络与所述初选歌曲的基频包络的相似度;
- [0067] 裁剪单元,用于根据所述基频包络的相似度、或者根据所述基频包络的相似度及

所述音符序列的相似度,对所述候选歌曲进行裁剪,得到候选歌曲。

[0068] 优选地,所述频谱特征匹配模块,具体用于采用动态时间规整算法计算所述频谱特征与所述候选歌曲的频谱特征的频谱相似度。

[0069] 本发明实施例提供的哼唱检索方法及系统,不仅提取用户哼唱歌曲片段的旋律特征,而且还提取其频谱特征,对传统的哼唱检索方案中的旋律特征进行了有效的补充,从而使哼唱歌曲与歌曲库中歌曲不仅从旋律上而且从歌词内容上都进行相应的匹配,有效地提高了检索的准确度及本发明方案的应用范围,降低了音乐库规模大小对检索结果准确率的影响。

附图说明

[0070] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明中记载的一些实施例,对于本领域普通技术人员来讲,还可以根据这些附图获得其他的附图。

[0071] 图1是本发明实施例哼唱检索方法的流程图;

[0072] 图2是本发明实施例中进行旋律特征匹配的一种流程图;

[0073] 图3是本发明实施例中进行旋律特征匹配的另一种流程图;

[0074] 图4是本发明实施例中从歌曲库中选择候选歌曲的一种流程;

[0075] 图5是本发明实施例哼唱检索系统的一种结构示意图;

[0076] 图6是本发明实施例中选择模块的一种具体结构示意图;

[0077] 图7是本发明实施例中旋律特征匹配模块的一种优选结构示意图;

[0078] 图8是本发明实施例中选择模块的另一种具体结构示意图。

具体实施方式

[0079] 为了使本技术领域的人员更好地理解本发明实施例的方案,下面结合附图和实施方式对本发明实施例作进一步的详细说明。

[0080] 针对现有的基于歌曲旋律特征匹配的检索方法存在的问题,本发明实施例提供一种哼唱检索方法及系统,通过引入频谱特征对传统旋律特征进行有效补充,由于频谱特征是反映哼唱歌词内容,区分性更大,因此大大提高了哼唱检索方法及系统的应用范围及检索结果的准确度。

[0081] 如图1所示,是本发明实施例哼唱检索方法的流程图,包括以下步骤:

[0082] 步骤101,获取用户哼唱歌曲片段。

[0083] 步骤102,提取所述哼唱歌曲片段的旋律特征及频谱特征。

[0084] 旋律特征可用音符序列或音高曲线的形式表示,其特征提取具体过程如下:

[0085] (1)提取所述哼唱歌曲的基频特征序列

[0086] 基频特征反映说话人声带振动的频率,一般来说,音调越高则基频越大。基频特征提取有较成熟的算法,如自相关基频估计、平均幅度差函数法、谐波法等。如在自相关法基频估计时,首先对连续语音信号进行加窗分帧,通常取两个基频周期以上。然后计算每一帧的自相关函数,并获取该自相关函数的峰值点位置。最后根据该峰值点位置获得基频特征序列。

[0087] (2) 基频特征变换

[0088] 基频特征序列反映了哼唱时音调变化的走势,却不能被人耳直接感知。为了提取更符合人耳的感知和音乐原理的特征,在本发明实施例中,对提取的基频特征序列进行分段和量化,获得具有一定音高和时长的音符序列。具体地,可以对该基频特征序列进行音符分割,根据静音段位置及基频值变化幅度确定音符边界。然后使用相邻音符之间音高差作为当前音符的音高值。进一步按照十二平均律将所述音高值转化到以半音程为单位的音高值。

[0089] 另外,在本发明实施例中,还可以直接采用基频包络特征作为所述旋律特征,以减少音符分割错误对检索结果的影响。

[0090] 当然,在实际应用中,还可以综合利用音符序列及基频包络来确定用户哼唱歌曲片段与歌曲库中歌曲的旋律相似度,对此将在后面详细说明。

[0091] 步骤103,根据所述旋律特征从所述歌曲库中选择候选歌曲。

[0092] 具体地,可以将用户哼唱歌曲片段的旋律特征与歌曲库中歌曲的旋律特征进行匹配,并根据匹配结果从歌曲库中选出一些相似度高的歌曲作为候选歌曲,比如,可以选择旋律相似度大于设定阈值的歌曲作为候选歌曲;或者按照旋律相似度由大到小的顺序选择设定个数(比如20个)的歌曲作为候选歌曲。前面提到,在本发明实施例中,所述旋律特征可以由音符序列和/或基频包络来表征。相应地,在进行旋律特征匹配时,也可以有多种不同的匹配方法,比如:

[0093] 对于音符序列表示的旋律特征,可以利用序列匹配算法,比如ED(Edit Distance,编辑距离)算法、LCS(Longest Common Subsequence,最长公共子序列)算法等,计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度。

[0094] 对于基频包络表示的旋律特征,可以利用DTW(Dynamic Time Warping,动态时间规整)算法计算用户哼唱歌曲片段的旋律特征与歌曲库中歌曲的旋律特征的相似度。

[0095] 另外,为了进行更准确的匹配,还可以综合利用上述音符序列及基频包络来计算用户哼唱歌曲片段的旋律特征与歌曲库中歌曲的旋律特征的相似度,对此将在后面详细说明。

[0096] 步骤104,将哼唱歌曲片段的频谱特征与所述候选歌曲的频谱特征进行匹配,得到频谱相似度。

[0097] 具体地,可以根据哼唱歌曲片段在候选歌曲中匹配的起止时间位置(哼唱歌曲片段在候选歌曲中匹配的起止时间位置可以通过编辑距离算法来确定),获取候选歌曲的相应片段的频谱特征,然后计算哼唱歌曲片段的频谱特征与候选歌曲的相应片段的频谱特征的相似度,比如可以采用DTW算法等完成该计算。

[0098] 步骤105,根据所述频谱相似度对所述候选歌曲进行排序,或者根据所述频谱相似度及所述旋律相似度,对所述候选歌曲进行排序。

[0099] 本发明实施例的哼唱检索方法,通过引入频谱特征对传统旋律特征进行有效补充,由于频谱特征是反映哼唱歌词内容,区分性更大,因此大大提高了哼唱检索方法及系统的应用范围及检索结果的准确度。

[0100] 需要说明的是,在应用本发明实施例的环境中,歌曲库需要预先建立,而且,对应歌曲库中的每首歌曲,还需要建立与其对应的旋律特征及频谱特征。其中,旋律特征可以是

音符序列和/或基频包络,频谱特征可以是MFCC(Mel Frequency Cepstrum Coefficient,梅尔频谱倒谱系数)、PLP(Perceptual Linear Predictive,感知线性预测系数)等。

[0101] 另外,前面提到,为了得到更准确的匹配结果,还可以综合利用上述音符序列及基频包络来计算用户哼唱歌曲片段的旋律特征与歌曲库中歌曲的旋律特征的相似度。

[0102] 如图2所示,是本发明实施例中进行旋律特征匹配的一种流程图,包括以下步骤:

[0103] 步骤201,计算用户哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列的相似度。

[0104] 步骤202,计算用户哼唱歌曲片段的基频包络与歌曲库中歌曲的基频包络的相似度。

[0105] 步骤203,将对应待匹配歌曲的音符序列的相似度与基频包络的相似度进行融合,得到对应所述待匹配歌曲的旋律特征的相似度。

[0106] 具体的融合方式可以按以下公式进行:

$$[0107] \quad \text{Sim} = \alpha \cdot S_1 + \beta \cdot S_2 \quad (1)$$

[0108] 其中,Sim表示旋律特征相似度, S_1 表示音符序列相似度, S_2 表示基频包络相似度, α 和 β 为融合系数。

[0109] 如图3所示,是本发明实施例中进行旋律特征匹配的另一种流程图,包括以下步骤:

[0110] 步骤301,计算用户哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列中各子音符序列的相似度。

[0111] 步骤302,选择待匹配歌曲的音符序列中与所述哼唱歌曲片段的音符序列相似度最大的子音符序列,并确定所述子音符序列的起止时间点。

[0112] 基于音符序列特征,在进行动态匹配时,容许输入哼唱音符序列是从待匹配歌曲音符序列的任意一个音符开始。因此,在本发明实施例中,可以计算用户哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列中各子音符序列的相似度。然后选择相似度最大的子音符序列。

[0113] 步骤303,计算哼唱歌曲片段的基频包络与待匹配歌曲在所述起止时间点内的基频包络的相似度。

[0114] 步骤304,将所述基频包络的相似度作为旋律相似度,或者将所述基频包络的相似度与对应的最大子音符序列的相似度进行融合,并将融合结果作为旋律相似度。

[0115] 与图2所示流程不同的是,在图3所示的流程中,基于基频包络的相似度计算只针对待匹配歌曲的音符序列中的一个子音符序列,从而可以提高计算效率,避免不必要的运算资源浪费。

[0116] 另外,本发明实施例中还提供了一种综合利用音符序列表示的旋律特征和基频包络表示的旋律特征选择候选歌曲的方法,如图4所示,这种从歌曲库中选择候选歌曲的流程包括以下步骤:

[0117] 步骤401,计算用户哼唱歌曲片段的音符序列与歌曲库中歌曲的音符序列的相似度;

[0118] 步骤402,根据所述音符序列的相似度从所述歌曲库中选择初选歌曲;

[0119] 步骤403,计算所述用户哼唱歌曲片段的基频包络与所述初选歌曲的基频包络的

相似度；

[0120] 步骤404,根据所述基频包络的相似度、或者根据所述基频包络的相似度及所述音符序列的相似度,对所述初选歌曲进行裁剪,得到候选歌曲。

[0121] 比如,根据音符序列的相似度从歌曲库中选择了20首初选歌曲,当然,这20首初选歌曲是音符序列与用户哼唱歌曲片段的音符序列的相似度是最高的,而且还可以依音符序列的相似度从高到低排序。然后,再计算用户哼唱歌曲片段的基频包络与这20首初选歌曲的基频包络的相似度。最后,根据基频包络的相似度、或者根据基频包络的相似度及音符序列的相似度(比如,对基频包络的相似度及音符序列的相似度进行加权平均),对这20首歌曲进行裁剪,得到裁剪后剩余的歌曲作为候选歌曲,以筛选出匹配度更高的候选歌曲,并且提高计算效率。具体的裁剪方式可以有多种,比如选择基频包络相似度最低的10首歌曲直接裁剪掉,或者按照20首初选歌曲音符序列相似度由低到高的顺序,依次检查其基频包络相似度的大小,将基频包络相似度小于一定值的初选歌曲裁剪掉。当然,还可以有其它裁剪方式,在此不再一一列举。

[0122] 相应地,本发明实施例还提供一种哼唱检索系统,如图5所示,是该系统的一种结构示意图。

[0123] 在该实施例中,所述系统包括:

[0124] 获取模块501,用于获取用户哼唱歌曲片段;

[0125] 特征提取模块502,用于提取所述哼唱歌曲片段的旋律特征及频谱特征;

[0126] 选择模块503,用于根据所述旋律特征从所述歌曲库中选择候选歌曲;

[0127] 频谱特征匹配模块504,用于将所述频谱特征与所述候选歌曲的频谱特征进行匹配,得到频谱相似度;

[0128] 排序模块505,用于根据所述频谱相似度、或者根据所述频谱相似度及所述旋律相似度,对所述候选歌曲进行排序。

[0129] 上述选择模块503可以选择旋律相似度大于设定阈值的歌曲作为候选歌曲;或者按照旋律相似度由大到小的顺序选择设定个数(比如20个)的歌曲作为候选歌曲。

[0130] 需要说明的是,上述选择模块503具体可以将用户哼唱歌曲的旋律特征与歌曲库中歌曲的旋律特征进行匹配,并根据匹配结果从歌曲库中选择相似度最高的一定数量的歌曲作为候选歌曲。

[0131] 前面提到,在本发明实施例中,所述旋律特征可以由音符序列和/或基频包络来表征。相应地,上述选择模块503也可以有多种实现结构,

[0132] 比如,如图6所示,是上述选择模块的一种具体结构示意图。

[0133] 该选择模块包括:旋律特征匹配模块601和候选模块602,其中,旋律特征匹配模块601用于将所述旋律特征与歌曲库中歌曲的旋律特征进行匹配,得到旋律相似度;候选模块602用于根据所述旋律相似度从所述歌曲库中选择候选歌曲。

[0134] 所述旋律特征匹配模块601在进行旋律特征匹配时,也可以有多种不同的实现方式。比如,对于音符序列表示的旋律特征,旋律特征匹配模块601可以利用序列匹配算法,比如ED算法、LCS算法等,计算所述旋律特征与歌曲库中歌曲的旋律特征的相似度。对于基频包络表示的旋律特征,旋律特征匹配模块601可以利用DTW算法计算用户哼唱歌曲片段的旋律特征与歌曲库中歌曲的旋律特征的相似度。

[0135] 另外,为了进行更准确的匹配,旋律特征匹配模块601还可以综合利用上述音符序列及基频包络来计算用户哼唱歌曲片段的旋律特征与歌曲库中歌曲的旋律特征的相似度。

[0136] 比如,旋律特征匹配模块601可以分别计算用户哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列的相似度、以及用户哼唱歌曲片段的基频包络与歌曲库中歌曲的基频包络的相似度,然后将对应待匹配歌曲的音符序列的相似度与基频包络的相似度进行融合,得到对应所述待匹配歌曲的旋律特征的相似度。

[0137] 再比如图7所示,是本发明实施例中旋律特征匹配模块的一种优选结构示意图。

[0138] 在该实施例中,所述旋律特征匹配模块包括:

[0139] 第一计算单元701,用于计算哼唱歌曲片段的音符序列与歌曲库中待匹配歌曲的音符序列中各子音符序列的相似度;

[0140] 选择单元702,用于选择待匹配歌曲的音符序列中与所述哼唱歌曲片段的音符序列相似度最大的子音符序列;

[0141] 确定单元703,用于确定所述相似度最大的子音符序列的起止时间点;

[0142] 第二计算单元704,用于计算所述哼唱歌曲片段的基频包络与所述待匹配歌曲在所述起止时间点内的基频包络的相似度;

[0143] 输出单元705,用于将所述基频包络的相似度作为旋律相似度输出。

[0144] 在该实施例中,旋律特征匹配模块基于基频包络的相似度计算只针对待匹配歌曲的音符序列中的一个子音符序列,从而可以提高计算效率,避免不必要的运算资源浪费。

[0145] 如图8所示,是本发明实施例中选择模块的另一种结构示意图。

[0146] 在该实施例中,所述选择模块包括:

[0147] 音符序列相似度计算单元801,用于计算所述哼唱歌曲片段的音符序列与歌曲库中歌曲的音符序列的相似度;

[0148] 初选单元802,用于根据所述音符序列的相似度从所述歌曲库中选择初选歌曲;

[0149] 基频包络相似度计算单元803,用于计算所述哼唱歌曲片段的基频包络与所述初选歌曲的基频包络的相似度;

[0150] 裁剪单元804,用于根据所述基频包络的相似度、或者根据所述基频包络的相似度及所述音符序列的相似度,对所述候选歌曲进行裁剪,得到候选歌曲。具体的裁剪方式可以有多种,可以参照前面本发明方法实施例中的描述,在此不再一一列举。

[0151] 本发明实施例的哼唱检索系统,通过引入频谱特征对传统旋律特征进行有效补充,由于频谱特征是反映哼唱歌词内容,区分性更大,并且不受歌曲库影响的信息,因此大大提高了哼唱检索方法及系统的应用环境及检索结果的准确度。

[0152] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于系统实施例而言,由于其基本相似于方法实施例,所以描述得比较简单,相关之处参见方法实施例的部分说明即可。以上所描述的系统实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0153] 以上对本发明实施例进行了详细介绍,本文中应用了具体实施方式对本发明进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及设备;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

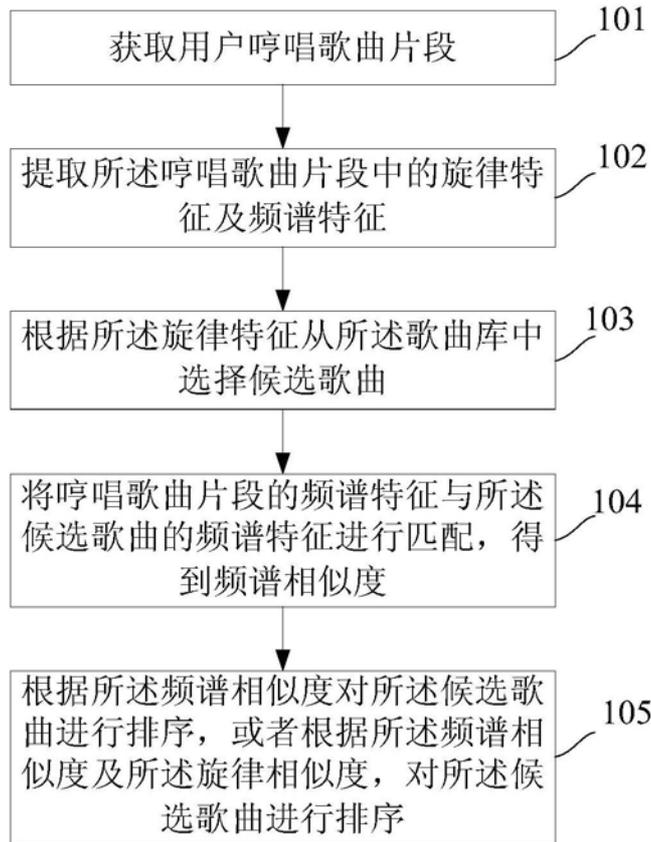


图1

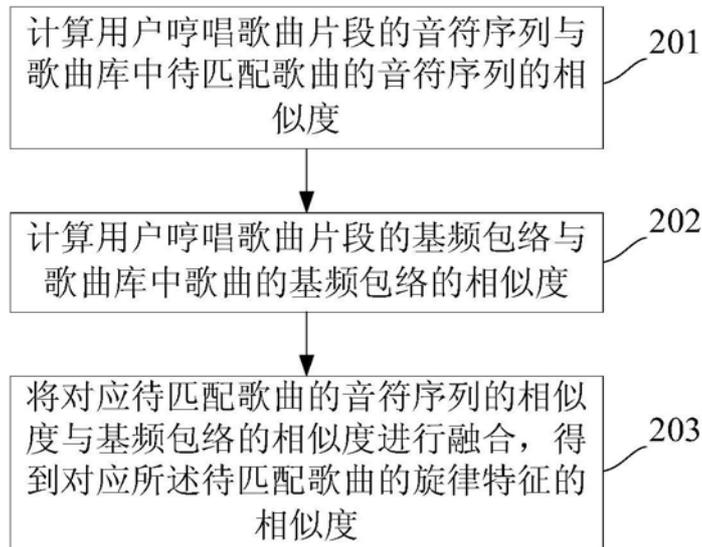


图2

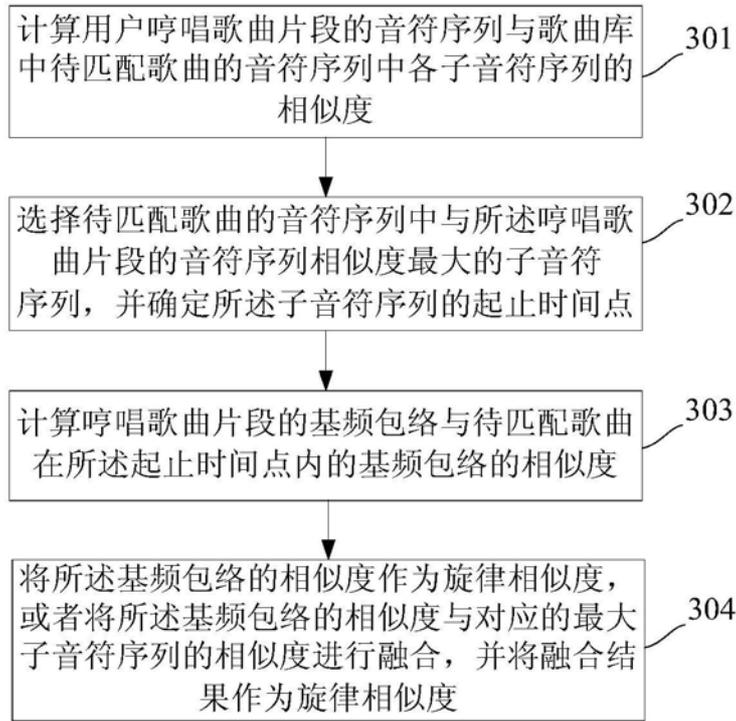


图3

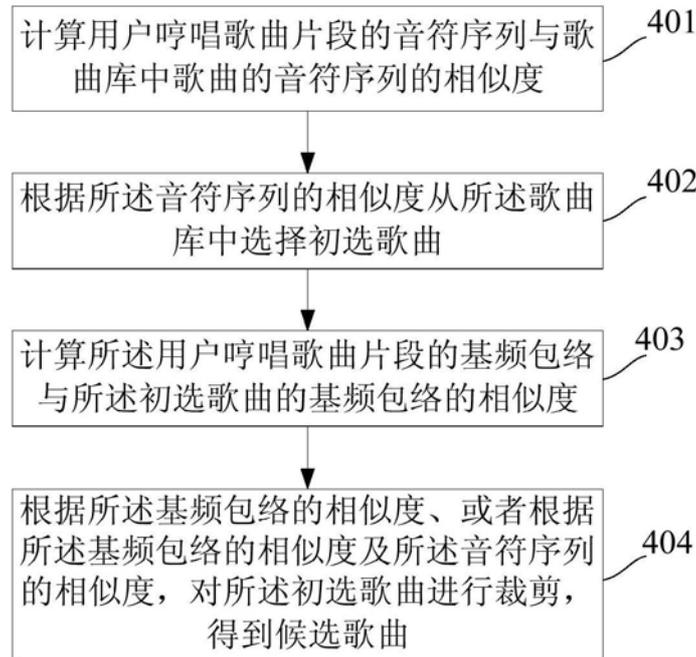


图4

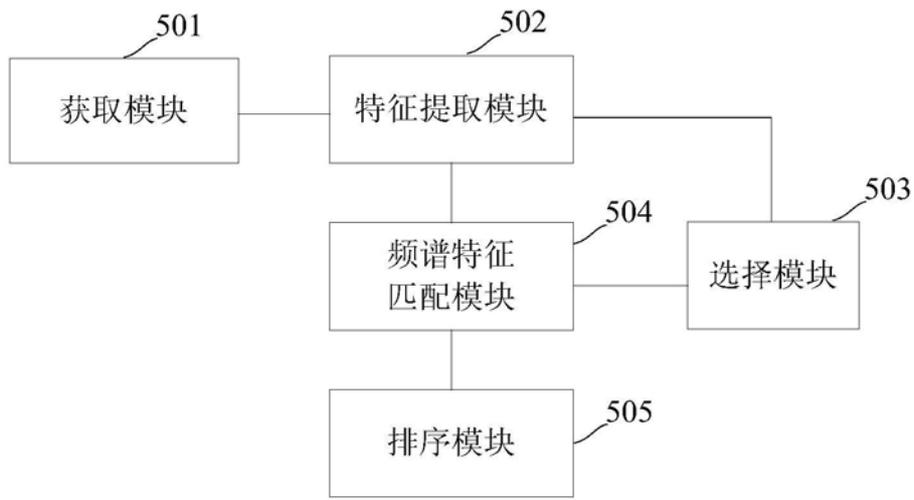


图5

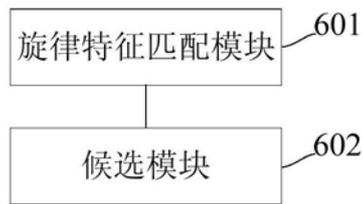


图6

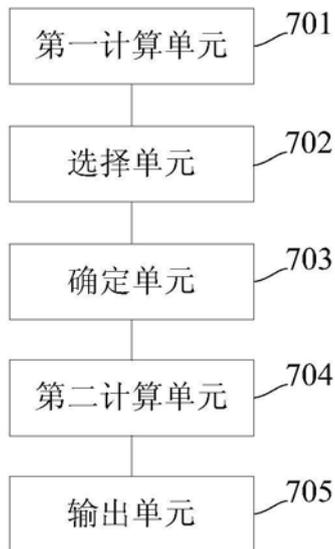


图7

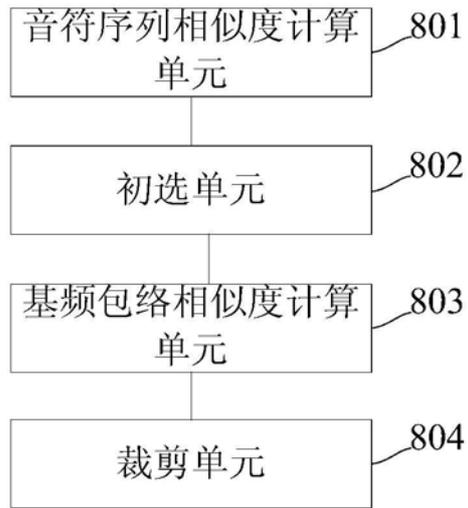


图8