(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2020 年 7 月 9 日 (09.07.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/140390 A1

(51) 国际专利分类号:
G10H 1/00 (2006.01)

(21) 国际申请号: PCT/CN2019/091093

(22) 国际申请日: 2019 年 6 月 13 日 (13.06.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910008576.5 2019年1月4日 (04.01.2019) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(72) 发明人: 朱清影 (ZHU, Qingying); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 程宁 (CHENG, Ning); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 王健宗 (WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市精英专利事务所 (SHENZHEN TALENT PATENT SERVICE); 中国广东省深圳市福田区深南中路6009号绿景广场B栋20层B, Guangdong 518000 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,

(54) **Title:** VIBRATO MODELING METHOD, DEVICE, COMPUTER APPARATUS AND STORAGE MEDIUM

(54) **发明名称:** 颤音建模方法、装置、计算机设备及存储介质

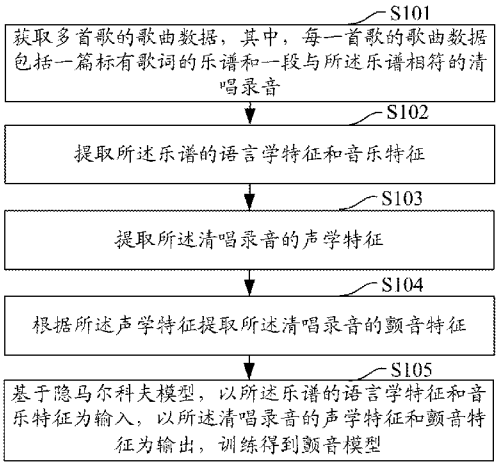

图 1

(57) **Abstract:** Provided are a vibrato modeling method, a device, a computer apparatus and storage medium. The method comprises: acquiring song data of multiple songs, wherein the song data of each song comprises a music score marked with lyrics and an accapella record matched with the music score (S101); extracting linguistic feature and music feature of the music score (S102); extracting acoustic feature of the accapella record (S103); extracting vibrato feature of the accapella record according to the acoustic feature (S104); based on the hidden Markov model, taking the linguistic feature and music feature of the music score as input, and taking the acoustic feature and the vibrato feature of the accapella record as output, performing training to obtain a vibrato model (S105).

WO 2020/140390 A1

[见续页]

MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要：一种颤音建模方法、装置、计算机设备及存储介质。该方法包括：获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与乐谱相符的清唱录音（S101）；提取乐谱的语言学特征和音乐特征(S102)；提取清唱录音的声学特征(S103)；根据声学特征提取清唱录音的颤音特征(S104)；基于隐马尔可夫模型，以乐谱的语言学特征和音乐特征为输入，以清唱录音的声学特征和颤音特征为输出，训练得到颤音模型(S105)。

颤音建模方法、装置、计算机设备及存储介质

本申请要求于 2019 年 1 月 4 日提交中国专利局、申请号为 201910008576.5、发明名称为“颤音建模方法、装置、计算机设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及人工智能技术领域，尤其涉及一种颤音建模方法、装置、计算机设备及存储介质。

背景技术

近年来，基于隐马尔可夫模型进行参数合成的歌曲合成方法在业界非常受关注。使用隐马尔可夫模型合成歌曲的最大优点是，在不需要庞大的歌唱数据库的情况下，可以有效的模拟不同的声音特征，歌唱风格，甚至是情绪。而颤音，作为一种重要的歌唱技巧，对合成歌曲的自然度有很大的影响。颤音在声学特征上的具体体现为基频上的小幅震动，颤音的具体时间点和强度因歌手而异。然而普通的隐马尔可夫模型会在训练和合成时平滑基频上的小幅度的起伏，如此会平滑掉颤音，导致合成的歌唱中并没有颤音的效果。

发明内容

本申请实施例提供一种颤音建模方法、装置、计算机设备及存储介质，可保留颤音特征，以提高合成歌曲的自然度。

第一方面，本申请实施例提供了一种颤音建模方法，该方法包括:

获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与所述乐谱相符的清唱录音；提取所述乐谱的语言学特征和音乐特征；提取所述清唱录音的声学特征；根据所述声学特征提取所述清唱录音的颤音特征；基于隐马尔可夫模型，以所述乐谱的语言学特征和音乐特征为输入，以所述清唱录音的声学特征和颤音特征为输出，训练得到颤音模型。

第二方面，本发明实施例提供了一种颤音建模装置，该颤音建模装置包括用于执行上述第一方面所述的方法对应的单元。

第三方面，本发明实施例提供了一种计算机设备，所述计算机设备包括存储器，以及与所述存储器相连的处理器；

所述存储器用于存储计算机程序，所述处理器用于运行所述存储器中存储的计算机程序，以执行上述第一方面所述的方法。

第四方面，本发明实施例提供了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时，实现上述第一方面所述的方法。

附图说明

为了更清楚地说明本发明实施例技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 是本申请实施例提供的颤音建模方法的流程示意图；

图 2 是本申请实施例提供的一个音符所对应的标签对数据；

图 3 是本申请实施例提供的一段清唱录音所对应的基频的示意图；

图 4 是本申请实施例提供的颤音建模方法的子流程示意图；

图 5 是本申请另一实施例提供的颤音建模方法的子流程示意图；

图 6 是本申请另一实施例提供的颤音建模方法的流程示意图；

图 7 是本申请实施例提供的颤音建模装置的示意性框图；

图 8 是本申请实施例提供的颤音定位单元的示意性框图；

图 9 是本申请另一实施例提供的颤音定位单元的示意性框图；

图 10 是本申请另一实施例提供的颤音建模装置的示意性框图；

图 11 是本申请实施例提供的计算机设备的示意性框图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清

楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

图 1 是本申请实施例提供的颤音建模方法的流程示意图。如图 1 所示，该方法包括 S101-S105。

S101，获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与所述乐谱相符的清唱录音。

其中，每一首歌的清唱录音是与该乐谱对应的清唱录音，即乐谱和清唱录音是对应的。其中，乐谱是以文本文件的形式存在。通常情况下，用一些软件如 MuseScore 制作电子乐谱时，可以直接将乐谱文件保存为一些预设的格式文件如 musicxml 格式。这类的格式文件其本质是有特殊格式的文本文件。

S102，提取所述乐谱的语言学特征和音乐特征。

从整体来看，一个乐谱文件的结构如下:

- 音调
- 谱号
- 拍号
- 速度
- （一些显示信息）
- 第一小节:
  - 音符
  - 音符
  - 音符
  ……
- 第二小节:
  - 音符
  ……

……

一个乐谱文件中通常包括音调、谱号、拍号、速度、显示信息（显示信息与音符在乐谱编辑软件中的视觉显示有关，与模型的训练无关，可忽略）、多个小节等信息。其中，每个小节中包括多个音符。每个音符中包括音高、时长、声部、音符类型、歌词等信息。而音高包括音阶和八度等信息，歌词包括音节、文本等信息。

乐谱文件中以标签对的形式来标识乐谱文件的各种信息，如<></>即为一个

标签对。图 2 是本申请实施例提供的一个音符所对应的标签对数据。需要注意的是，图 2 中的汉字所对应的是注解，以方便理解该音符所对应的标签对数据的含义。如图 2 所示，<note>和</note>这个标签之间包含了关于这个音符所有信息。而其中的<pitch>和</pitch>之间记录了该音符的音高信息，<duration>和</duration>之间记录了该音符的时长信息，<voice>和</voice>之间记录了该音符的声部信息，<type>和</type>之间记录了该音符的音符类型，<stem>和</stem>之间记录了该音符的显示信息（显示信息与音符在乐谱编辑软件中的视觉显示有关，与模型的训练无关，可忽略），<lyric>和</lyric>之间记录了该音符的歌词信息。其中，音高信息<pitch>和</pitch>中的<step>和</step>之间记录了音高信息中的音阶，<octave>和</octave>之间记录了音高信息中的八度；<lyric>和</lyric>中的<syllabic>和</syllabic>之间记录了歌词的音节，<text>和</text>之间记录了歌词文本。通过寻找标签对，可以读取到每个标签对中具体的关键字（如 pitch、step、lyric 等关键字）和关键字对应值的信息。

在一个乐谱文件中，“音调”、“谱号”、“拍号”、“速度”、每个音符的“音高”、“时长”、“声部”和“音符类型”都是音乐特征，“歌词”则是文本特征，对应语言学特征。具体地，步骤 S102，即提取所述乐谱的语言学特征和音乐特征，包括：获取乐谱文件中的便签对；解析标签对，以提取标签对中所对应的音乐特征以及语言学特征的值。其中，语言学特征包括歌词的发音，以及上下前后文的关系等。如此通过获取并解析乐谱文件中的标签对，提取乐谱中的语言学特征和音乐特征，达到提取特征的目的。

S103，提取所述清唱录音的声学特征。

其中，声学特征包括基频和梅尔频谱系数等。每一帧对应一组特征。

基音的频率称为基频，而基音就是指发浊音时声带振动的周期性，基频是语音中频率最低的纯音，但振幅确是最大的，决定整个音的升高。基频作为语音的主要特征，广泛应用于语音编码、语音识别、语音合成等领域。

清唱录音的基频特征的提取方法有很多种，大致可以分为三类：一，时域分析算法，如自相关法（ACF）、短时平均幅度差法（AMDF）等；二，频域分析算法，如倒谱法（CEP）等；三，时频结合的分析算法，如小波分析算法等。还可以使用其他的基频提取方法来提取清唱录音的基频。其中，提取出的

清唱录音的基频为基频序列。

梅尔倒谱系数是在 Mel 标度频率域提取出来的倒谱参数，Mel 标度描述了人耳频率的非线性特性。提取梅尔倒谱系数的方法主要包括:

对清唱录音的语音信号进行采样处理以得到清唱录音的数字语音信号；将清唱录音的数字语音信号进行预加重处理；将预加重处理后的数字语音信号进行分帧处理；将分帧处理后的数字语音信号进行加窗处理；将加窗处理后的数字语音信号进行快速傅里叶变换以得到频域的声音信号；通过三角形带通滤波器组对频域的声音信号进行滤波以使三角形带通滤波器中的每个滤波器分别输出滤波结果，其中，该三角形带通滤波器包括 p 个滤波器；将每个滤波器输出的滤波结果分别取对数以得到声音信号的 p 个对数能量；将所得的 p 个对数能量进行离散余弦变化得到梅尔频率倒谱系数的 p 阶分量。其中，p 可以在 22-26 范围内取值，也可以在其他合适的范围内取值。

其中，预加载处理就是将清唱录音的数字语音信号通过一个高通滤波器，预加重的目的是提升高频部分，使得清唱录音的数字语音信号的频谱变得平坦，移除频谱倾斜，来补偿语音信号受到发音系统所抑制的高频部分；同时，也是为了消除发生过程中声带和嘴唇的效应。使用三角带通滤波器的目的是对频谱进行平滑化，并消除谐波的作用，此外还可以减少运算量。

S104，根据所述声学特征提取所述清唱录音的颤音特征。

在本实施例中，认为颤音是出现在基频序列上小幅的正弦的波动，颤音特征包括振幅和频率。根据所述声学特征提取所述清唱录音的颤音特征，包括:根据所述声学特征中的基频来提取所述清唱录音的颤音特征。

在一实施例中，步骤 S104，包括：定位所述基频所对应的基频序列中的颤音片段；计算所述颤音片段的振幅和频率，将计算出的所述颤音片段的振幅和频率作为所述颤音片段中每一帧的颤音特征。

在一实施例中，还需将所述基频中的非颤音片段的振幅和频率设置为零，如此，以得到所述清唱录音的颤音特征。可以理解地，计算颤音片段的振幅和频率，将计算出的振幅和频率作为颤音特征；对于清唱录音中的非颤音片段，即没有颤音或者静音的片段，设置振幅和频率都为零。

图 3 是本申请实施例提供的一段清唱录音所对应的基频的示意图。图 3 中

横轴是时间，单位为 0.5 毫秒；纵轴为基频，单位为森特。椭圆圈的部分即为颤音片段，具体地定位基频所对应的基频序列中的颤音片段的方法下文将具体介绍。其中，可以看出颤音是出现在基频序列上小幅的正弦的波动。其他没有颤音或者静音的片段，设置振幅和频率都为零。

在一实施例中，如图 4 所示，定位所述基频所对应的基频序列中的颤音片段，包括 S401-S405。

S401，检测所述基频序列中连续出现波谷或者波峰的次数是否达到预设次数。

其中，连续出现波谷或者波峰意味着该基频序列片段没有间断，即该基频序列片段是不间断的，在一些实施例中，也可以理解为，该基频序列片段中没有频率为 0 的频率。检测在一个不间断的基频序列片段中出现波谷或者波峰的次数是否达到预设次数。其中，连续出现波谷或者波峰的次数可以刚好是预设次数，也可以超过预设次数。刚好是预设次数，或者超过预设次数，都认为达到预设次数。其中，预设次数可以设置为 5 次，也可以设置为其他的次数。

若所述基频序列中连续出现波谷或者波峰的次数达到预设次数，执行步骤 S402；否则，执行步骤 S405。

S402，获取所对应的基频序列片段，并统计该基频序列片段中的平均频率。

基频序列片段中的平均频率通过获取该基频序列片段中每一帧的频率，将该基频序列片段中每一帧的频率之和，除以该基频序列片段中的总帧数，以得到该基频序列片段中的平均频率。

S403，检测预设次数中每次频率由波谷到波峰或者频率由波峰到波谷的过程中，波谷所对应的频率是否小于平均频率，且波峰所对应的频率是否大于平均频率。

若每次波谷所对应的频率小于平均频率，且波峰所对应的频率大于平均频率，执行步骤 S404；否则，执行步骤 S405。

S404，确定该基频序列片段为颤音片段。

S405，确定该基频序列片段为非颤音片段。

即基频序列中连续出现波谷或者波峰的次数未达到预设次数；或者达到预设次数，但是并非每次波谷所对应的频率小于平均频率、且波峰所对应的频率

大于平均频率，那么确定该频率序列片段为非颤音片段。

该实施例中通过判断基频序列中的基频是否连续预设次数或者预设次数以上由高到低，或由低到高的穿过该基频序列片段的平均基频。若是，则认为是该基频序列片段为颤音片段，否则，认为该基频序列片段为非颤音片段。如图3所示，所确定颤音片段的频率是连续预设次数或者预设次数以上由高到低穿过该基频序列片段的平均基频。

在一实施例中，预设次数的判断可通过统计该基频序列中出现频率连续由波谷到波峰或者频率连续由波峰到波谷的次数，即判断基频序列中出现频率连续由波谷到波峰或者频率连续由波峰到波谷的次数是否达到预设次数。

其中，连续波谷到波峰理解为基频的频率从一个波谷到达波峰（再达到波谷），紧接着又从波谷到达一个波峰（再达到波谷），如此连续的从波谷到达波峰的次数是否达到预设次数。需要注意的是，在该过程中，基频的频率并未间断。其中，预设次数可以设置为5次。基频连续由高到低的理解也是一致，只是统计的是基频从波峰到达波谷的次数。

由于颤音的频率范围是在一个预设赫兹频域范围内，如预设赫兹频域范围是5到8赫兹，因此如果某一基频序列片段是颤音片段，那么该基频序列片段的基频以一个5到8赫兹之间的频率f在振动。在一实施例中，如图5所示，定位所述基频所对应的基频序列中的颤音片段，包括S501-S508。

S501，对所述基频序列中的基频序列片段做短时距傅里叶变换，得到基频序列片段的功率谱。

将基频的功率谱表示为$X(f,t)$，其中，f为基频的震动频率，t为时间。则$X(f,t)$可以理解为是在时间点t时，频率为f的波形（如图3所示的基频曲线）的振幅。

S502，将所述功率谱进行正则化以得到正则化函数。

即对$X(f,t)$进行正则化以得到正则化函数$\hat{X}(f,t)$，其中，正则化公式为：$\hat{X}(f,t)=\frac{X(f,t)}{\int X(f,t)df}$。

S503，计算正则化函数在预设赫兹频域范围上的积分，以得到所述基频序列片段的功率。

若预设赫兹频域范围为5到8赫兹，对$\hat{X}(f,t)$做5到8赫兹频域上的积分，

得到该基频序列片段的功率$\psi_v(t)$，其中，计算功率的公式如下：

$\psi_v(t)=\int_{F_L}^{F_H}\hat{X}(f,t)df$，其中，$F_L$表示预设赫兹频域范围中最低的频率，$F_H$表示预设赫兹频域范围中最高的频率。

S504，计算正则化函数在所述基频序列片段中的斜率变化。

其中，斜率变化越大代表函数值的尖峰越明显。计算斜率变化的计算公式如下：

$S_v(t)=\int_{F_L}^{F_H}\left|\frac{\partial\hat{X}(f,t)}{\partial f}\right|df$，其中，$F_L$表示预设赫兹频域范围中最低的频率，$F_H$表示预设赫兹频域范围中最高的频率。

S505，根据计算出的所述基频序列片段的功率和斜率变化确定颤音概率。

其中，在时间 t 时，根据计算出的所述基频序列片段的功率和斜率变化确定颤音概率$P_v(t)$的计算公式为：$P_v(t)=S_v(t)\psi_v(t)$。

$P_v(t)$的值越大，说明该时间点 t 属于颤音片段的可能性越大。

S506，判断所述颤音概率是否超过预设数值。

若所述颤音概率超过预设数值，则认为该时间点属于颤音片段中的点。若连续多个点判断出来都是属于颤音片段，那么该基频序列片段则属于颤音片段。

若所述颤音概率超过预设数值，执行步骤 S507；若所述颤音概率未超过预设数值，执行步骤 S508。

S507，确定所述基频序列片段为颤音片段。

S508，确定所述基频序列片段为非颤音片段。

其中，需要注意的是，计算斜率变化和计算功率的步骤的先后顺序不做具体限定，也可以先计算功率再计算斜率变化。

该实施例中，可以理解为，若某一个基频序列片段中预设赫兹频域范围段（如 5 到 8 赫兹这个频率段）的功率高于其他频率段；频率为 f 的波形振幅要明显的大于 5 到 8 赫兹之间其他波形的振幅，那么认为该基频序列片段是颤音片段。

在一实施例中，也可以将以上图 4 和图 5 两个实施例中的判断方法结合起来以用来定位基频序列中的颤音片段，那么可以理解地，基频连续预设次数以上由高到低，或由低到高的穿过该片段的平均基频，且计算出的颤音概率超过预设数值，那么确认该时间点所在的基频序列片段是颤音片段。

S105，基于隐马尔可夫模型，以所述乐谱的语言学特征和音乐特征为输入，以所述清唱录音的声学特征和颤音特征为输出，训练得到颤音模型。

可以理解地，将所述乐谱的语言学特征和音乐特征作为输入，输入到隐马尔科夫模型中，训练该隐马尔科夫模型，以使得输出为所述清唱录音的声学特征和颤音特征。将训练得到的隐马尔科夫模型作为颤音模型。

以上方法实施例通过提取清唱录音的颤音特征，基于隐马尔可夫模型，将歌曲的乐谱的语言学特征和音乐特征作为输入，将该歌曲的清唱录音的声学特征和颤音特征作为输出，如此，得到输入为所述乐谱的语言学特征和音乐特征，输出为所述清唱录音的声学特征和颤音特征的颤音模型。以上方法实施例实现了颤音模型的训练。该颤音模型可有效的保留颤音特征，以在合成歌曲时提高合成歌曲的自然度。

图 6 是本申请另一实施例提供的颤音建模方法的示意流程图。如图 6 所示，该实施例包括步骤 S601-S609。其中，该实施例与图 1 实施例的区别在于：增加了步骤 S606-S609，其他步骤 S601-S605 与图 1 实施例的步骤 S101-S105 分别对应。下面将详细介绍该实施例与图 1 实施例的区别之处。

S606，提取待合成歌曲的乐谱的语言学特征和音乐特征。

其中，提取待合成歌曲的乐谱的语言学特征和音乐特征的方法请参看图 1 实施例中的提取乐谱的语言学特征和音乐特征的描述，在此不再赘述。

S607，将提取的语言学特征和音乐特征输入到训练好的颤音模型中，以得到与所述待合成歌曲的乐谱相符的清唱录音的声学特征和颤音特征。

S608，在所得到的声学特征中加入所得到的颤音特征。

首先判断待合成歌曲中的每一帧是否为静音；如果是静音，那么基频不变；如果不为静音，那么就在基频上加入所得到的颤音特征。如此，在基频上加入了颤音特征。具体地，在基频上加入所得到的颤音特征，包括：在基频上加入与所得到的颤音特征的振幅和频率对应的正弦波来模拟颤音。该处的对应理解为相同，即在基频上加入与所得到的颤音特征的振幅和频率相同的正弦波来模拟颤音。

S609，将加入了颤音特征的声学特征输入声码器，以合成歌曲。

输入到声码器中的声学特征包括：加入了颤音特征的基频和梅尔频谱系数。

声码器（vocoder)，是语音信号某种模型的语音分析合成系统。在传输中只利用模型中的特征参数，在编译码时利用模型中的特征参数估计和语音合成的语音信号编译码器，一种对话音进行分析和合成的编、译码器。

该实施例进一步通过训练好的颤音模型来得到待合成歌曲的乐谱相符的清唱录音的声学特征和颤音特征，并将得到的颤音特征加入到声学特征中的基频中，如此，声学特征（包括基频和梅尔频谱系数）中可以表达颤音。在将加入了颤音特征的声学特征合成歌曲时，歌曲中可以很好的体现颤音，如此，使生成的歌曲的自然度显著提升。

图 7 是本申请实施例提供的颤音建模装置的示意性框图。如图 7 所示，该装置包括用于执行上述颤音建模方法所对应的单元。具体地，如图 7 所示，该装置 70 包括歌曲数据获取单元 701、音乐特征提取单元 702、声学特征提取单元 703、颤音特征提取单元 704 以及模型建立单元 705。

歌曲数据获取单元 701，用于获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与所述乐谱相符的清唱录音。

音乐特征提取单元 702，用于提取所述乐谱的语言学特征和音乐特征。具体地，音乐特征提取单元 702，包括标签对获取单元、解析单元。其中，标签对获取单元，用于获取所述乐谱所对应的乐谱文件中的标签对。解析单元，用于解析所述标签对，以提取所述标签对中所对应的音乐特征以及语言学特征的值。

声学特征提取单元 703，用于提取所述清唱录音的声学特征。其中，所述声学特征包括基频和梅尔频谱倒数。

颤音特征提取单元 704，用于根据所述声学特征提取所述清唱录音的颤音特征。具体地，根据所述声学特征中的基频提取所述清唱录音的颤音特征。颤音特征包括振幅和频率。

在一实施例中，颤音特征提取单元 704，包括：颤音定位单元、颤音特征确定单元。其中，颤音定位单元，用于定位所述基频所对应的基频序列中的颤音片段。颤音特征确定单元，用于计算所述颤音片段的振幅和频率，将计算出的所述颤音片段的振幅和频率作为所述颤音片段中每一帧的颤音特征。

在一实施例中，颤音特征提取单元 704 还包括设置单元。所述设置单元，

用于将所述基频中的非颤音片段的振幅和频率设置为零。如此，以得到所述清唱录音的颤音特征。

在一实施例中，如图 8 所示，颤音定位单元 80 包括次数检测单元 801、统计单元 802、频率检测单元 803 以及第一颤音确定单元 804。其中，次数检测单元 801，用于检测所述基频序列中连续出现波谷或者波峰的次数是否达到预设次数。统计单元 802，用于若达到预设次数，获取所对应的基频序列片段，并统计该基频序列片段中的平均频率。频率检测单元 803，用于检测预设次数中每次频率由波谷到波峰或者频率由波峰到波谷的过程中，波谷所对应的频率是否小于平均频率，且波峰所对应的频率是否大于平均频率。第一颤音确定单元 804，用于若每次波谷所对应的频率小于平均频率，且波峰所对应的频率大于平均频率，确定该基频序列片段为颤音片段。第一颤音确定单元 804，还用于若基频序列中连续出现波谷或者波峰的次数未达到预设次数; 或者达到预设次数，但是并非每次波谷所对应的频率小于平均频率、且波峰所对应的频率大于平均频率，确定该基频序列片段为非颤音片段。

在一实施例中，如图 9 所示，颤音定位单元 90 包括变换单元 901、正则化单元 902、功率计算单元 903、斜率变化计算单元 904、概率计算单元 905、概率判断单元 906 以及第二颤音确定单元 907。其中，变换单元 901，用于对所述基频序列中的基频序列片段做短时距傅里叶变换，得到基频序列片段的功率谱。正则化单元 902，用于将所述功率谱进行正则化以得到正则化函数。功率计算单元 903，用于计算正则化函数在预设赫兹频域范围上的积分，以得到所述基频序列片段的功率。斜率变化计算单元 904，用于计算正则化函数在所述基频序列片段中的斜率变化。概率计算单元 905，用于根据计算出的所述基频序列片段的功率和斜率变化确定颤音概率。概率判断单元 906，用于判断所述颤音概率是否超过预设数值。第二颤音确定单元 907，用于若所述颤音概率超过预设数值，确定所述基频序列片段为颤音片段; 若所述颤音概率未超过预设数值，确定所述基频序列片段为非颤音片段。

模型建立单元 705，用于基于隐马尔可夫模型，以所述乐谱的语言学特征和音乐特征为输入，以所述清唱录音的声学特征和颤音特征为输出，训练得到颤音模型。

图 10 本申请实施例提供的颤音建模装置的示意性框图。如图 10 所示，该装置包括用于执行上述颤音建模方法所对应的单元。具体地，如图 10 所示，该装置 100 包括歌曲数据获取单元 101、音乐特征提取单元 102、声学特征提取单元 103、颤音特征提取单元 104、模型建立单元 105、模型使用单元 106、颤音加入单元 107 以及合成单元 108。其中，该实施例与图 7 所示的实施例的不同之处在于：增加了模型使用单元 106、颤音加入单元 107 以及合成单元 108。下面将描述该实施例与图 7 实施例的不同之处，其他的单元与图 7 实施例所示的单元分别对应，在此不再赘述。

音乐特征提取单元 102，还用于提取待合成歌曲的乐谱的语言学特征和音乐特征。

模型使用单元 106，用于将提取的语言学特征和音乐特征输入到训练好的颤音模型中，以得到与所述待合成歌曲的乐谱相符的清唱录音的声学特征和颤音特征。

颤音加入单元 107，用于在所得到的声学特征中加入所得到的颤音特征。

合成单元 108，用于将加入了颤音特征的声学特征输入声码器，以合成歌曲。

需要说明的是，所属领域的技术人员可以清楚地了解到，上述装置和各单元的具体实现过程，可以参考前述方法实施例中的相应描述，为了描述的方便和简洁，在此不再赘述。

上述装置可以实现为一种计算机程序的形式，计算机程序可以在如图 11 所示的计算机设备上运行。

图 11 为本申请实施例提供的一种计算机设备的示意性框图。该设备为终端等设备，如移动终端、PC终端、IPad等。该设备 110 包括通过系统总线 111 连接的处理器 112、存储器和网络接口 113，其中，存储器可以包括非易失性存储介质 114 和内存储器 115。

该非易失性存储介质 114 可存储操作系统 1141 和计算机程序 1142。该非易失性存储介质中所存储的计算机程序 1142 被处理器 112 执行时，可实现上述所述的颤音建模方法。该处理器 112 用于提供计算和控制能力，支撑整个设备的运行。该内存储器 115 为非易失性存储介质中的计算机程序的运行提供环境，

该计算机程序被处理器 112 执行时，可使得处理器 112 执行上述所述的颤音建模方法。该网络接口 113 用于进行网络通信。本领域技术人员可以理解，图 11 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的设备的限定，具体的设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

其中，所述处理器 112 用于运行存储在存储器中的计算机程序，以实现前述的颤音建模方法的任一实施例。

应当理解，在本申请实施例中，所称处理器 112 可以是中央处理单元 (Central Processing Unit，CPU)，该处理器还可以是其他通用处理器、数字信号处理器 (Digital Signal Processor，DSP)、专用集成电路 (应用程序 lication Specific Integrated Circuit，ASIC)、现成可编程门阵列 (Field-Programmable Gate Array，FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

本领域普通技术人员可以理解的是实现上述实施例的方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成。该计算机程序可存储于一存储介质中，该存储介质可以为计算机可读存储介质。该计算机程序被该计算机系统中的至少一个处理器执行，以实现上述方法的实施例的流程步骤。

因此，本申请还提供了一种存储介质。该存储介质可以为计算机可读存储介质，该计算机可读存储介质包括非易失性计算机可读存储介质。该存储介质存储有计算机程序，该计算机程序当被处理器执行时实现前述的颤音建模方法的任一实施例。

所述存储介质可以是 U 盘、移动硬盘、只读存储器（Read-Only Memory，ROM）、磁碟或者光盘等各种可以存储程序代码的计算机可读存储介质。

在本申请所提供的几个实施例中，应该理解到，所揭露的装置、设备和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，上述描述的装置、设备和单元的具体工作过程，可以参考前述方法实施例中的对

应过程，在此不再赘述。以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以权利要求的保护范围为准。

权 利 要 求 书

1.一种颤音建模方法，包括:

获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与所述乐谱相符的清唱录音;

提取所述乐谱的语言学特征和音乐特征;

提取所述清唱录音的声学特征;

根据所述声学特征提取所述清唱录音的颤音特征;

基于隐马尔可夫模型，以所述乐谱的语言学特征和音乐特征为输入，以所述清唱录音的声学特征和颤音特征为输出，训练得到颤音模型。

2.根据权利要求 1 所述的方法，其中，所述方法还包括:

提取待合成歌曲的乐谱的语言学特征和音乐特征;

将提取的语言学特征和音乐特征输入到训练好的颤音模型中，以得到与所述待合成歌曲的乐谱相符的清唱录音的声学特征和颤音特征;

在所得到的声学特征中加入所得到的颤音特征;

将加入了颤音特征的声学特征输入声码器，以合成歌曲。

3.根据权利要求 1 所述的方法，其中，所述声学特征包括基频，所述根据所述声学特征提取所述清唱录音的颤音特征，包括:

定位所述基频所对应的基频序列中的颤音片段;

计算所述颤音片段的振幅和频率，将计算出的所述颤音片段的振幅和频率作为所述颤音片段中每一帧的颤音特征。

4.根据权利要求 3 所述的方法，其中，所述定位所述基频所对应的基频序列中的颤音片段，包括:

检测所述基频序列中连续出现波谷或者波峰的次数是否达到预设次数;

若达到预设次数，获取所对应的基频序列片段，并统计该基频序列片段中的平均频率;

检测预设次数中每次频率由波谷到波峰或者频率由波峰到波谷的过程中，波谷所对应的频率是否小于平均频率，且波峰所对应的频率是否大于平均频率;

若每次波谷所对应的频率小于平均频率，且波峰所对应的频率大于平均频率，确定该基频序列片段为颤音片段;否则，确定该基频序列片段为非颤音片

段。

5.根据权利要求 3 所述的方法，其中，所述定位所述基频所对应的基频序列中的颤音片段，包括:

对所述基频序列中的基频序列片段做短时距傅里叶变换，得到基频序列片段的功率谱;

将所述功率谱进行正则化以得到正则化函数;

计算正则化函数在预设赫兹频域范围上的积分，以得到所述基频序列片段的功率;

计算正则化函数在所述基频序列片段中的斜率变化;

根据计算出的所述基频序列片段的功率和斜率变化确定颤音概率;

判断所述颤音概率是否超过预设数值;

若所述颤音概率超过预设数值，确定所述基频序列片段为颤音片段; 否则，确定所述基频序列片段为非颤音片段。

6.根据权利要求 1 所述的方法，其中，所述提取所述乐谱的语言学特征和音乐特征，包括:

获取所述乐谱所对应的乐谱文件中的标签对;

解析所述标签对，以提取所述标签对中所对应的音乐特征以及语言学特征的值。

7.根据权利要求 1 所述的方法，其中，所述声学特征包括梅尔频谱系数，所述提取所述清唱录音的声学特征，包括:

对清唱录音的语音信号进行采样处理以得到清唱录音的数字语音信号;

将清唱录音的数字语音信号进行预加重处理;

将预加重处理后的数字语音信号进行分帧处理;

将分帧处理后的数字语音信号进行加窗处理;

将加窗处理后的数字语音信号进行快速傅里叶变换以得到频域的声音信号;

通过三角形带通滤波器组对频域的声音信号进行滤波以使三角形带通滤波器中的每个滤波器分别输出滤波结果，其中，该三角形带通滤波器包括 p 个滤波器;

将每个滤波器输出的滤波结果分别取对数以得到声音信号的 p 个对数能量;

将所得的 p 个对数能量进行离散余弦变化得到梅尔频率倒谱系数的 p 阶分量。

8.根据权利要求 2 所述的方法，其中，所得到的声学特征中包括基频，所述在所得到的声学特征中加入所得到的颤音特征，包括:

判断所述待合成歌曲中的每一帧是否为静音;

若为静音，则保持所得到的声学特征中的基频不变;

若不为静音，在所得到的声学特征中的基频上加入所对应的颤音特征。

9.根据权利要求 8 所述的方法，其中，所述在所得到的声学特征中的基频上加入所得到的颤音特征，包括:

在所得到的声学特征中的基频上加入与所述颤音特征的振幅和频率对应的正弦波。

10.一种颤音建模装置，其中，所述颤音建模装置包括:

歌曲数据获取单元，用于获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与所述乐谱相符的清唱录音;

音乐特征提取单元，用于提取所述乐谱的语言学特征和音乐特征;

声学特征提取单元，用于提取所述清唱录音的声学特征;

颤音特征提取单元,用于根据所述声学特征提取所述清唱录音的颤音特征;

模型建立单元，用于基于隐马尔可夫模型，以所述乐谱的语言学特征和音乐特征为输入，以所述清唱录音的声学特征和颤音特征为输出，训练得到颤音模型。

11.一种计算机设备，所述计算机设备包括存储器，以及与所述存储器相连的处理器；其中，所述存储器用于存储计算机程序；所述处理器用于运行所述存储器中存储的计算机程序，以执行如下步骤:

获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与所述乐谱相符的清唱录音;

提取所述乐谱的语言学特征和音乐特征;

提取所述清唱录音的声学特征;

根据所述声学特征提取所述清唱录音的颤音特征;

基于隐马尔可夫模型，以所述乐谱的语言学特征和音乐特征为输入，以所述清唱录音的声学特征和颤音特征为输出，训练得到颤音模型。

12.根据权利要求 11 所述的计算机设备，其中，所述处理器还执行如下步骤:

提取待合成歌曲的乐谱的语言学特征和音乐特征;

将提取的语言学特征和音乐特征输入到训练好的颤音模型中，以得到与所述待合成歌曲的乐谱相符的清唱录音的声学特征和颤音特征;

在所得到的声学特征中加入所得到的颤音特征;

将加入了颤音特征的声学特征输入声码器，以合成歌曲。

13.根据权利要求 11 所述的计算机设备，其中，所述声学特征包括基频，所述处理器在执行所述根据所述声学特征提取所述清唱录音的颤音特征的步骤时，具体执行如下步骤:

定位所述基频所对应的基频序列中的颤音片段;

计算所述颤音片段的振幅和频率，将计算出的所述颤音片段的振幅和频率作为所述颤音片段中每一帧的颤音特征。

14.根据权利要求 13 所述的计算机设备，其中，所述处理器在执行所述定位所述基频所对应的基频序列中的颤音片段的步骤时，具体执行如下步骤:

检测所述基频序列中连续出现波谷或者波峰的次数是否达到预设次数;

若达到预设次数，获取所对应的基频序列片段，并统计该基频序列片段中的平均频率;

检测预设次数中每次频率由波谷到波峰或者频率由波峰到波谷的过程中，波谷所对应的频率是否小于平均频率，且波峰所对应的频率是否大于平均频率;

若每次波谷所对应的频率小于平均频率，且波峰所对应的频率大于平均频率，确定该基频序列片段为颤音片段; 否则，确定该基频序列片段为非颤音片段。

15.根据权利要求 13 所述的计算机设备，其中，所述处理器在执行所述定位所述基频所对应的基频序列中的颤音片段的步骤时，具体执行如下步骤:

对所述基频序列中的基频序列片段做短时距傅里叶变换，得到基频序列片

段的功率谱;

将所述功率谱进行正则化以得到正则化函数;

计算正则化函数在预设赫兹频域范围上的积分，以得到所述基频序列片段的功率;

计算正则化函数在所述基频序列片段中的斜率变化;

根据计算出的所述基频序列片段的功率和斜率变化确定颤音概率;

判断所述颤音概率是否超过预设数值;

若所述颤音概率超过预设数值，确定所述基频序列片段为颤音片段;否则，确定所述基频序列片段为非颤音片段。

16.根据权利要求 11 所述的计算机设备，其中，所述处理器在执行所述提取所述乐谱的语言学特征和音乐特征的步骤时，具体执行如下步骤:

获取所述乐谱所对应的乐谱文件中的标签对;

解析所述标签对，以提取所述标签对中所对应的音乐特征以及语言学特征的值。

17.根据权利要求 11 所述的方法，其中，所述声学特征包括梅尔频谱系数，所述处理器在执行所述提取所述清唱录音的声学特征的步骤时，具体执行如下步骤:

对清唱录音的语音信号进行采样处理以得到清唱录音的数字语音信号;

将清唱录音的数字语音信号进行预加重处理;

将预加重处理后的数字语音信号进行分帧处理;

将分帧处理后的数字语音信号进行加窗处理;

将加窗处理后的数字语音信号进行快速傅里叶变换以得到频域的声音信号;

通过三角形带通滤波器组对频域的声音信号进行滤波以使三角形带通滤波器中的每个滤波器分别输出滤波结果，其中，该三角形带通滤波器包括 p 个滤波器;

将每个滤波器输出的滤波结果分别取对数以得到声音信号的 p 个对数能量;

将所得的 p 个对数能量进行离散余弦变化得到梅尔频率倒谱系数的 p 阶分

量。

18.根据权利要求 12 所述的计算机设备，其中，所得到的声学特征中包括基频，所述处理器在执行所述在所得到的声学特征中加入所得到的颤音特征的步骤时，具体执行如下步骤：

判断所述待合成歌曲中的每一帧是否为静音；

若为静音，则保持所得到的声学特征中的基频不变；

若不为静音，在所得到的声学特征中的基频上加入所对应的颤音特征。

19.根据权利要求 18 所述的计算机设备，其中，所述处理器在执行所述在所得到的声学特征中的基频上加入所得到的颤音特征的步骤时，具体执行如下步骤：

在所得到的声学特征中的基频上加入与所述颤音特征的振幅和频率对应的正弦波。

20.一种计算机可读存储介质，其中，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时，实现如权利要求 1-9 任一项所述的方法。

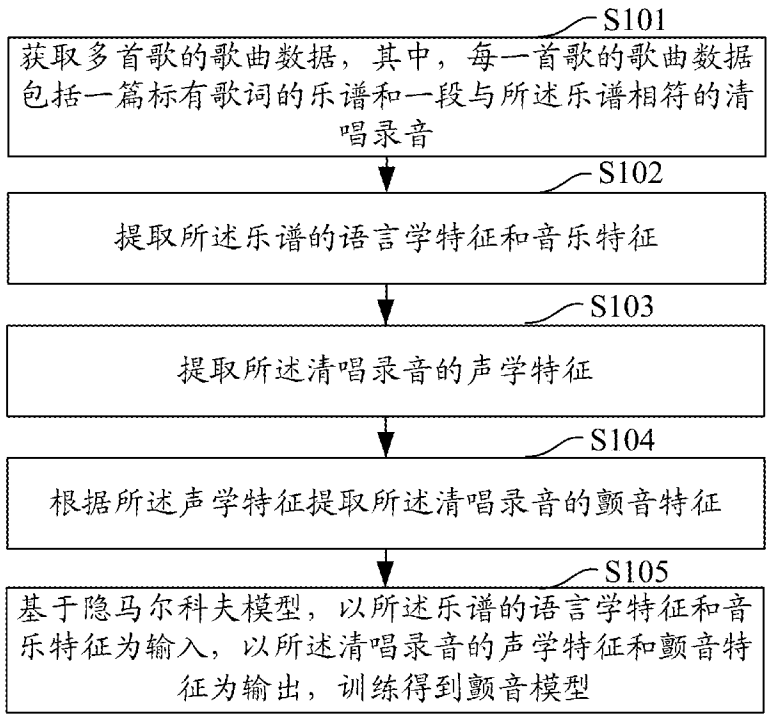
S101
获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与所述乐谱相符的清唱录音
S102
提取所述乐谱的语言学特征和音乐特征
S103
提取所述清唱录音的声学特征
S104
根据所述声学特征提取所述清唱录音的颤音特征
S105
基于隐马尔科夫模型，以所述乐谱的语言学特征和音乐特征为输入，以所述清唱录音的声学特征和颤音特征为输出，训练得到颤音模型

图 1

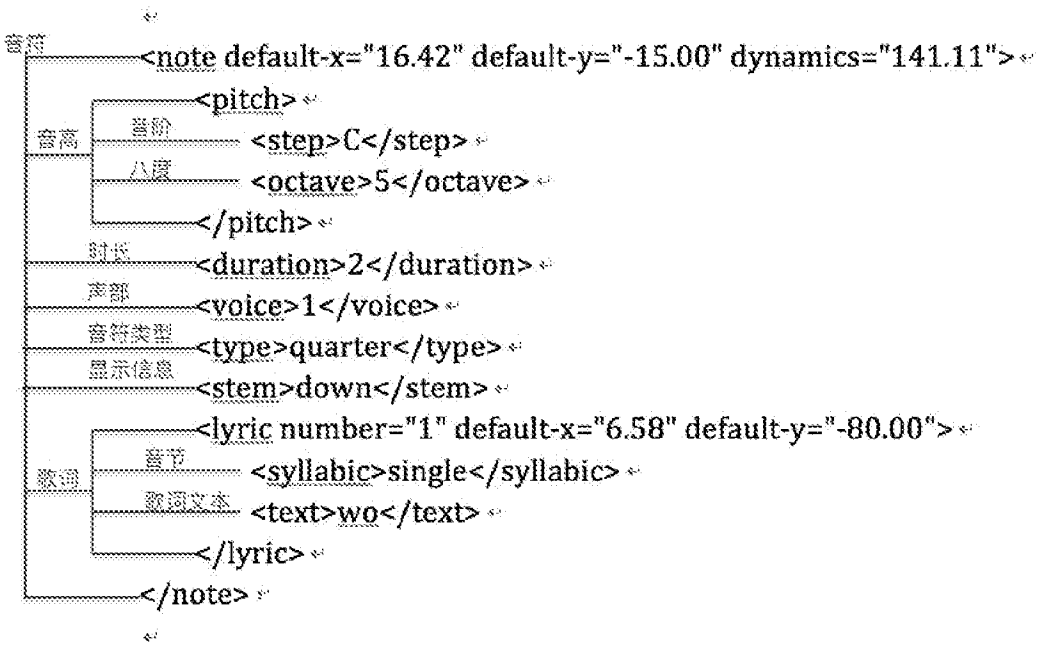
音符
<note default-x="16.42" default-y="-15.00" dynamics="141.11">
音高
音阶
<step>C</step>
八度
<octave>5</octave>
</pitch>
时长
<duration>2</duration>
声部
<voice>1</voice>
音符类型
<type>quarter</type>
显示信息
<stem>down</stem>
歌词
<lyric number="1" default-x="6.58" default-y="-80.00">
音节
<syllabic>single</syllabic>
歌词文本
<text>wo</text>
</lyric>
</note>

图 2

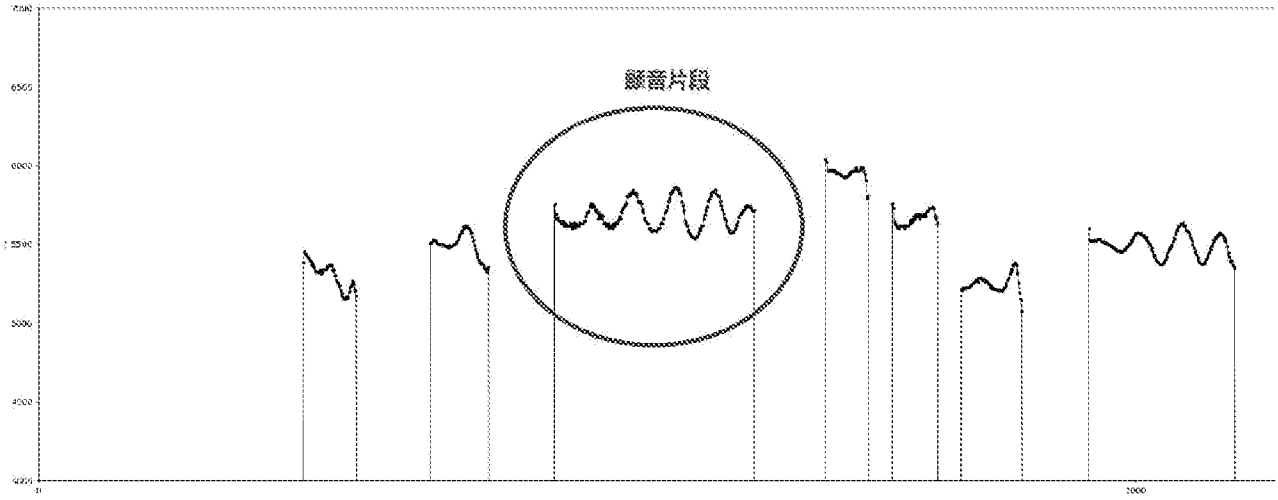
颤音片段

图 3

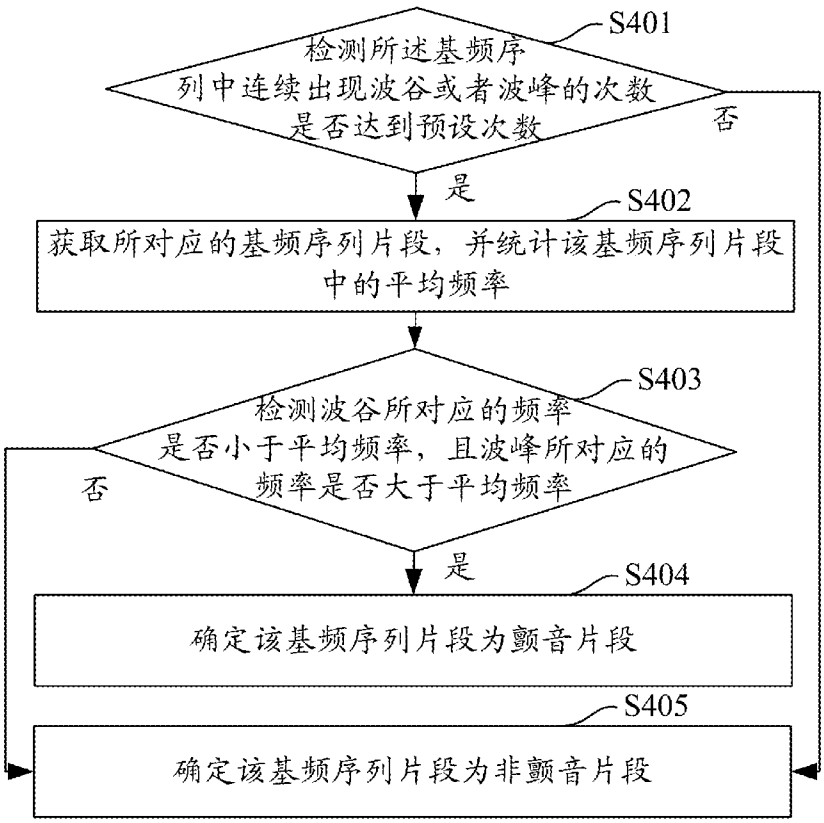
S401
检测所述基频序列中连续出现波谷或者波峰的次数是否达到预设次数
否
是
S402
获取所对应的基频序列片段，并统计该基频序列片段中的平均频率
S403
检测波谷所对应的频率是否小于平均频率，且波峰所对应的频率是否大于平均频率
否
是
S404
确定该基频序列片段为颤音片段
S405
确定该基频序列片段为非颤音片段

图 4

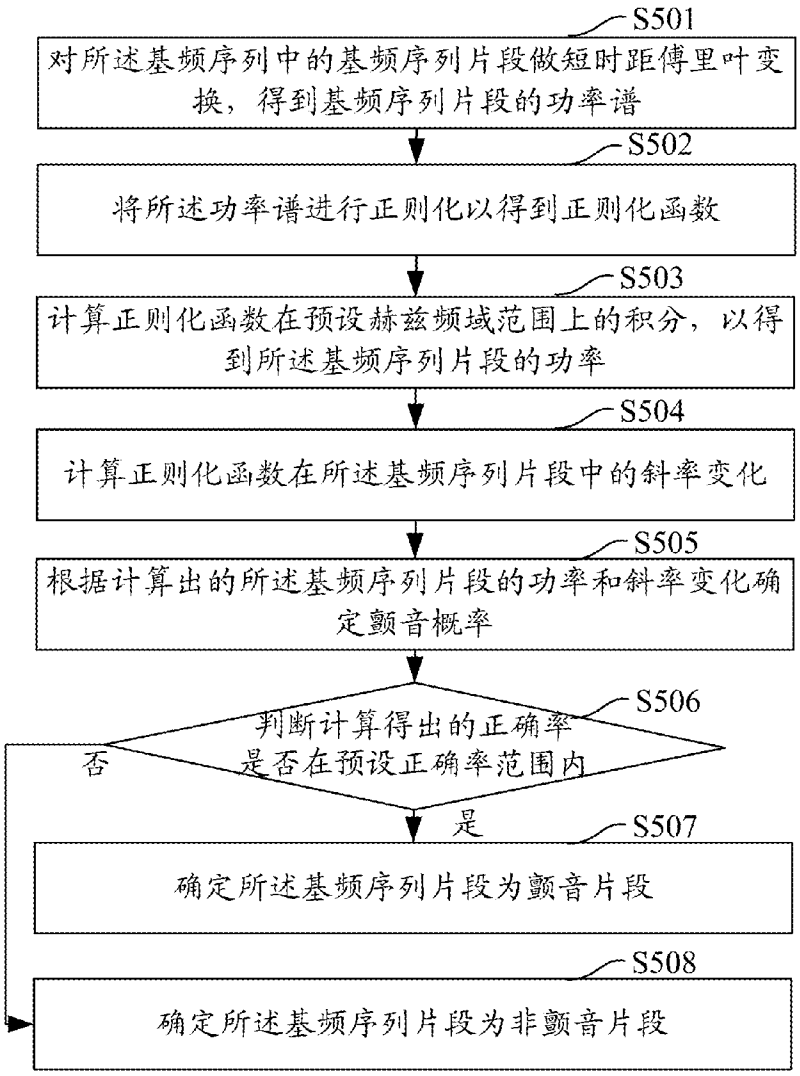
S501
对所述基频序列中的基频序列片段做短时距傅里叶变换，得到基频序列片段的功率谱
S502
将所述功率谱进行正则化以得到正则化函数
S503
计算正则化函数在预设赫兹频域范围上的积分，以得到所述基频序列片段的功率
S504
计算正则化函数在所述基频序列片段中的斜率变化
S505
根据计算出的所述基频序列片段的功率和斜率变化确定颤音概率
S506
判断计算得出的正确率是否在预设正确率范围内
否
是
S507
确定所述基频序列片段为颤音片段
S508
确定所述基频序列片段为非颤音片段

图 5

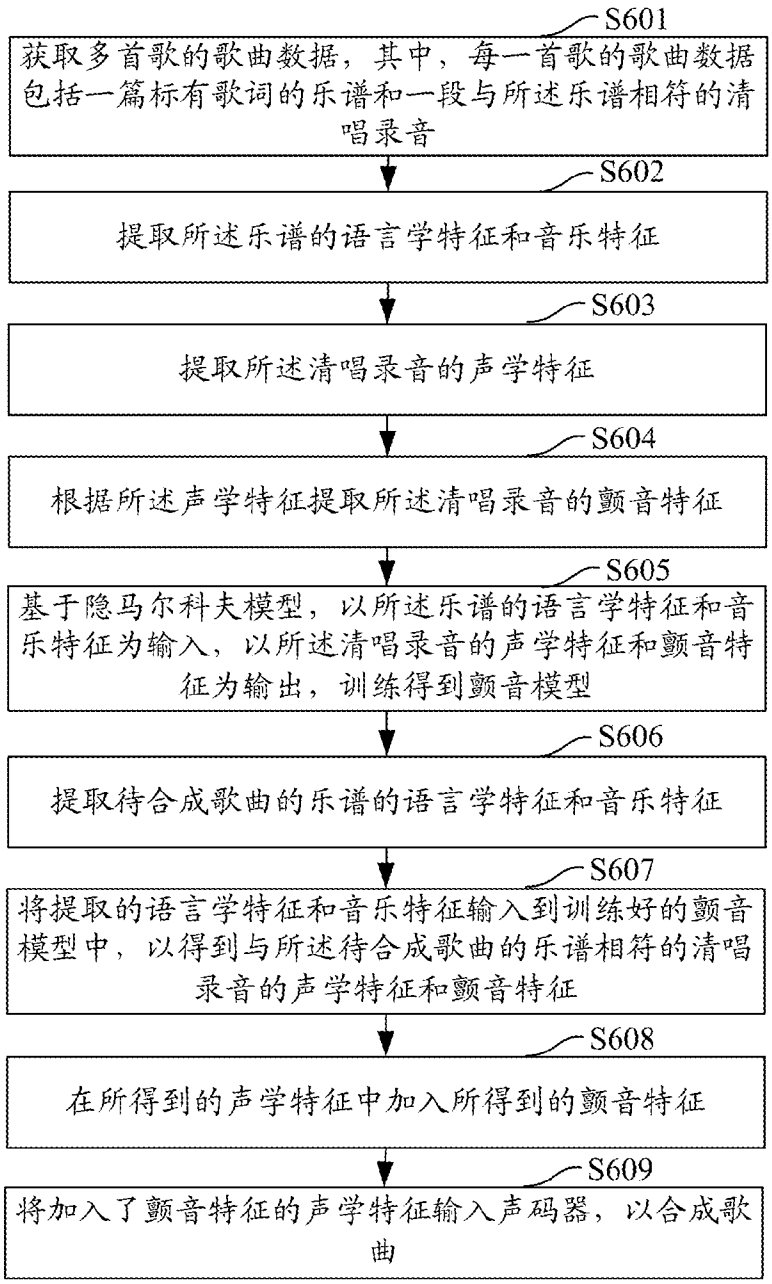
S601
获取多首歌的歌曲数据，其中，每一首歌的歌曲数据包括一篇标有歌词的乐谱和一段与所述乐谱相符的清唱录音
S602
提取所述乐谱的语言学特征和音乐特征
S603
提取所述清唱录音的声学特征
S604
根据所述声学特征提取所述清唱录音的颤音特征
S605
基于隐马尔科夫模型，以所述乐谱的语言学特征和音乐特征为输入，以所述清唱录音的声学特征和颤音特征为输出，训练得到颤音模型
S606
提取待合成歌曲的乐谱的语言学特征和音乐特征
S607
将提取的语言学特征和音乐特征输入到训练好的颤音模型中，以得到与所述待合成歌曲的乐谱相符的清唱录音的声学特征和颤音特征
S608
在所得到的声学特征中加入所得到的颤音特征
S609
将加入了颤音特征的声学特征输入声码器，以合成歌曲

图 6

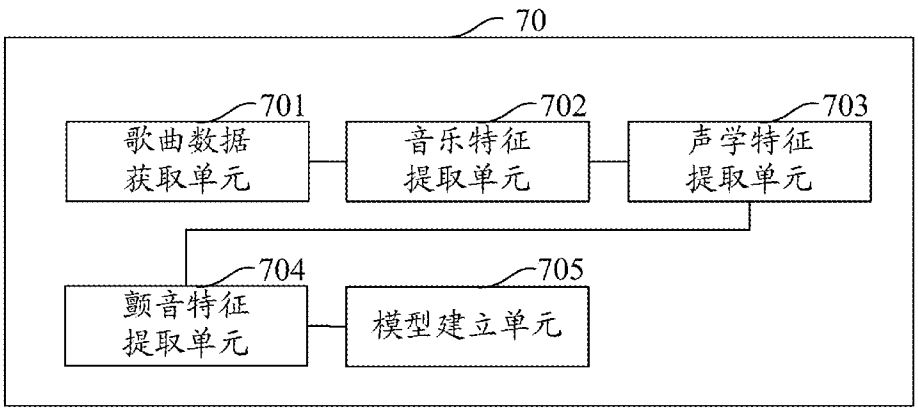
70
701
歌曲数据获取单元
702
音乐特征提取单元
703
声学特征提取单元
704
颤音特征提取单元
705
模型建立单元

图 7

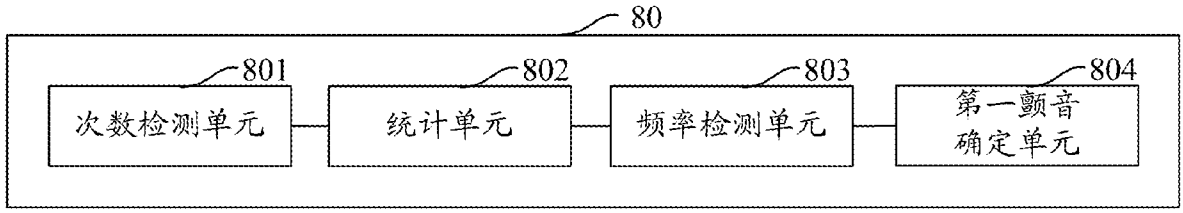
80
801
次数检测单元
802
统计单元
803
频率检测单元
804
第一颤音
确定单元

图 8

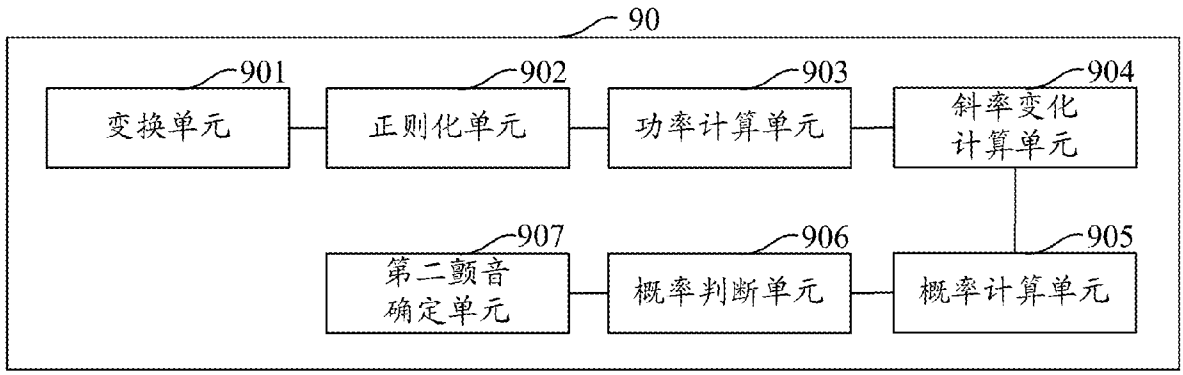
90
901
变换单元
902
正则化单元
903
功率计算单元
904
斜率变化
计算单元
907
第二颤音
确定单元
906
概率判断单元
905
概率计算单元

图 9

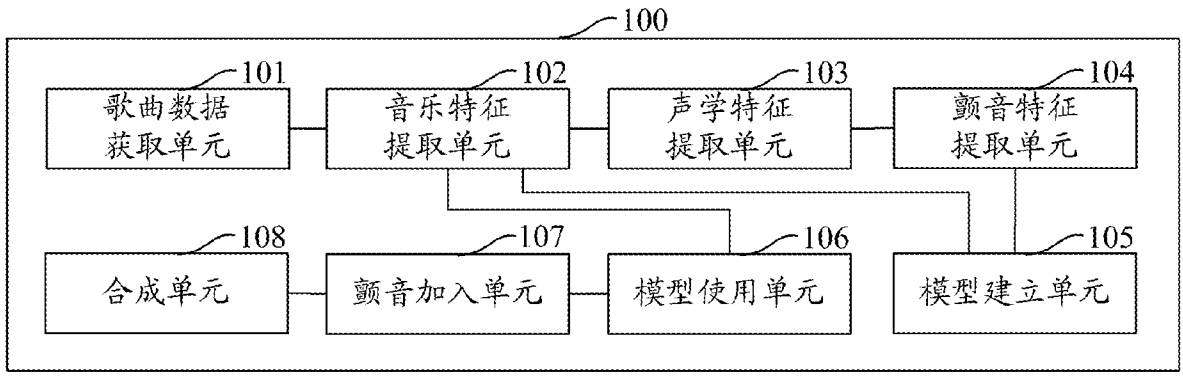
100
101
歌曲数据
获取单元
102
音乐特征
提取单元
103
声学特征
提取单元
104
颤音特征
提取单元
108
合成单元
107
颤音加入单元
106
模型使用单元
105
模型建立单元

图 10

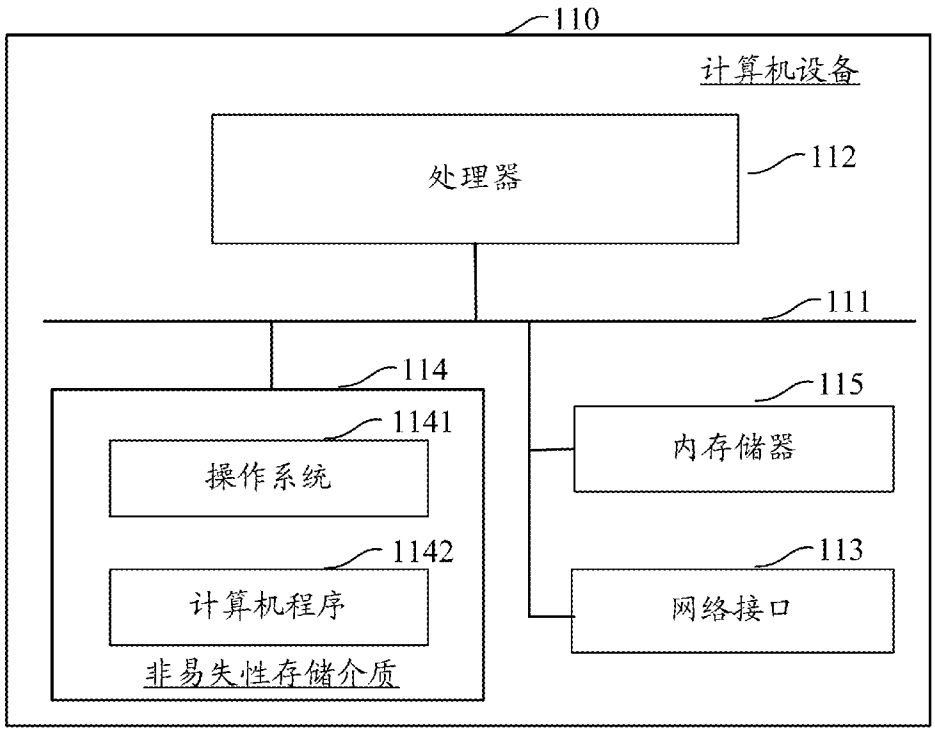
110
计算机设备
处理器
112
111
114
1141
操作系统
1142
计算机程序
非易失性存储介质
115
内存储器
113
网络接口

图 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/091093

A. CLASSIFICATION OF SUBJECT MATTER

G10H 1/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10H 1;G10L 15

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT; DWPI; CNKI; 百度学术, BAIDU SCHOLAR: 歌曲, 合成, 颤音, 训练, 模型, 建模, 检测, 识别, 定位, 波谷, 波峰, 功率, 斜率, music, sing, synthesis, vibrato, detect, recognize, peak, power, slope, gradient

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109817191 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 28 May 2019 (2019-05-28) claims 1-10	1-20
X	LI, Xian. "基于统计模型的汉语歌声合成研究 (Statistical Model Based Mandarin Chinese Singing Voice Synthesis)" *中国博士学位论文全文数据库 (信息科技辑) (Electronic Technology & Information Science, China Doctoral Dissertations Full-Text Database),* 15 March 2016 (2016-03-15), chapters 2, 3 and 5	1-3, 6-13, 16-20
X	LI, Xian et al. "A HMM-Based Mandarin Chinese Singing Voice Synthesis System" *IEEE/CAA Journal of Automatica Sinica,* Vol. 3, No. (2), 30 April 2016 (2016-04-30), sections II-IV	1-3, 6-13, 16-20
X	HONO, Y. et al. "Recent Development of the DNN-Based Singing Voice Synthesis System — Sinsy" *Proceedings, APSIPA Annual Summit and Conference 2018,* 15 November 2018 (2018-11-15), entire document	1-3, 6-13, 16-20
A	CN 106898340 A (TENCENT MUSIC ENTERTAINMENT (SHENZHEN) CO., LTD.) 27 June 2017 (2017-06-27) entire document	1-20

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 September 2019	**11 October 2019**
Name and mailing address of the ISA/CN **China National Intellectual Property Administration** **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China** Facsimile No. **(86-10)62019451**	Authorized officer Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 106971703 A (NORTHWEST NORMAL UNIVERSITY) 21 July 2017 (2017-07-21) entire document	1-20
A	US 2009314155 A1 (MICROSOFT CORP.) 24 December 2009 (2009-12-24) entire document	1-20
A	US 2009306987 A1 (NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY) 10 December 2009 (2009-12-10) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/091093

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 109817191 A	28 May 2019	None	
CN 106898340 A	27 June 2017	None	
CN 106971703 A	21 July 2017	None	
US 2009314155 A1	24 December 2009	US 7977562 B2	12 July 2011
		US 2011231193 A1	22 September 2011
US 2009306987 A1	10 December 2009	US 8244546 B2	14 August 2012
		JP 5024711 B2	12 September 2012
		JP 2010009034 A	14 January 2010

国际检索报告

国际申请号
PCT/CN2019/091093

A. 主题的分类

G10H 1/00(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10H 1;G10L 15

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNTXT;DWPI;CNKI;百度学术:歌曲，合成，颤音，训练，模型，建模，检测，识别，定位，波谷，波峰，功率，斜率，music, sing, synthesis,vibrato, detect, recognize,peak, power, slope, gradient

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 109817191 A (平安科技深圳有限公司) 2019年 5月 28日 (2019-05-28) 权利要求1-10	1-20
X	李贤. "基于统计模型的汉语歌声合成研究" 中国博士学位论文全文数据库（信息科技辑），2016年 3月 15日 (2016-03-15)， 第2，3，5章	1-3，6-13，16-20
X	Xian Li 等. "A HMM-based Mandarin Chinese Singing Voice Synthesis System" IEEE/CAA JOURNAL OF AUTOMATICA SINICA, 第3卷，第2期，2016年 4月 30日 (2016-04-30)， 第II-IV部分	1-3，6-13，16-20
X	Yukiya Hono等. "Recent Development of the DNN-based Singing Voice Synthesis System — Sinsy" Proceedings, APSIPA Annual Summit and Conference 2018, 2018年 11月 15日 (2018-11-15)， 全文	1-3，6-13，16-20
A	CN 106898340 A (腾讯音乐娱乐深圳有限公司) 2017年 6月 27日 (2017-06-27) 全文	1-20

☑ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

"A" 认为不特别相关的表示了现有技术一般状态的文件

"E" 在国际申请日的当天或之后公布的在先申请或专利

"L" 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

"O" 涉及口头公开、使用、展览或其他方式公开的文件

"P" 公布日先于国际申请日但迟于所要求的优先权日的文件

"T" 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

"X" 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

"Y" 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

"&" 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2019年 9月 20日	2019年 10月 11日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 游晓梅 电话号码 (86-10)62089539

国际检索报告

国际申请号
PCT/CN2019/091093

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 106971703 A (西北师范大学) 2017年 7月 21日 (2017-07-21) 全文	1-20
A	US 2009314155 A1 (MICROSOFT CORP.) 2009年 12月 24日 (2009-12-24) 全文	1-20
A	US 2009306987 A1 (NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY) 2009年 12月 10日 (2009-12-10) 全文	1-20

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/091093

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	109817191	A	2019年 5月 28日	无			
CN	106898340	A	2017年 6月 27日	无			
CN	106971703	A	2017年 7月 21日	无			
US	2009314155	A1	2009年 12月 24日	US	7977562	B2	2011年 7月 12日
				US	2011231193	A1	2011年 9月 22日
US	2009306987	A1	2009年 12月 10日	US	8244546	B2	2012年 8月 14日
				JP	5024711	B2	2012年 9月 12日
				JP	2010009034	A	2010年 1月 14日