



(12)发明专利

(10)授权公告号 CN 105321526 B

(45)授权公告日 2020.07.24

(21)申请号 201510612358.4

(22)申请日 2015.09.23

(65)同一申请的已公布的文献号  
申请公布号 CN 105321526 A

(43)申请公布日 2016.02.10

(73)专利权人 联想(北京)有限公司  
地址 100085 北京市海淀区上地西路6号

(72)发明人 王少敏 陈文辉

(74)专利代理机构 北京市柳沈律师事务所  
11105  
代理人 安之斐

(56)对比文件

- CN 102592607 A, 2012.07.18,
- CN 102592607 A, 2012.07.18,
- CN 104916288 A, 2015.09.16,
- CN 104078051 A, 2014.10.01,
- CN 101354889 A, 2009.01.28,
- CN 102982809 A, 2013.03.20,
- CN 103310796 A, 2013.09.18,
- CN 103280215 A, 2013.09.04,
- JP 特开2001-306100 A, 2001.11.02,
- CN 104704558 A, 2015.06.10,

审查员 薛彦杰

(51) Int. Cl.

G10L 25/48(2013.01)

G10L 25/18(2013.01)

G10L 21/003(2013.01)

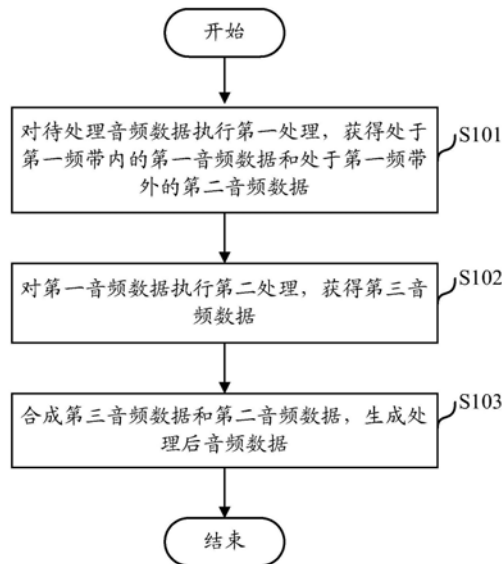
权利要求书3页 说明书9页 附图7页

(54)发明名称

音频处理方法和电子设备

(57)摘要

本申请公开了一种音频处理方法和使用该音频处理方法的电子设备。所述音频处理方法,包括:对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于所述第一频带外的第二音频数据;对所述第一音频数据执行第二处理,获得第三音频数据;以及合成所述第三音频数据和所述第二音频数据,生成处理后音频数据。



1. 一种音频处理方法,包括:

对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于所述第一频带外的第二音频数据;

对所述第一音频数据执行第二处理,获得第三音频数据;以及

合成所述第三音频数据和所述第二音频数据,生成处理后音频数据,

其中,所述第一音频数据和所述第三音频数据具有相同的时长,

其中所述合成所述第三音频数据和所述第二音频数据,生成处理后音频数据包括:

提取所述第三音频数据中的时间戳标记;

基于所述时间戳标记,确定所述第三音频数据相对于所述第二音频数据的起始时间点和终止时间点;以及

基于所述起始时间点和终止时间点,对准并且组合所述第三音频数据和所述第二音频数据,以生成处理后音频数据,

其中,所述对所述第一音频数据执行第二处理包括:

在所述第二处理是升调处理的情况下,将经所述升调处理后的所述第一音频数据分为一个个帧,在每一帧后面补偿该帧最后一部分的数据,获得第三音频数据,

其中,所述对所述第一音频数据执行第二处理还包括:

基于所述第一音频数据的特征,识别所述第一音频数据中具有第一特征的第一子音频数据和具有第二特征的第二子音频数据;以及

对所述第一子音频数据执行第一子处理,并且对所述第二子音频数据执行第二子处理。

2. 如权利要求1所述的音频处理方法,其中所述第一频带为特定的频率范围内的频带,所述特定的频率范围对应于人发声的频率范围。

3. 如权利要求1或2所述的音频处理方法,其中所述对待处理音频数据执行第一处理以及所述对所述第一音频数据执行第二处理,获得频率改变的所述第一音频数据以及保持不变的所述第二音频数据。

4. 如权利要求1或2所述的音频处理方法,其中所述对待处理的音频数据执行第一处理包括:

对待处理的音频数据执行第一滤波处理,以获得所述第一频带内的所述第一音频数据,并且对待处理的音频数据执行第二滤波处理,以获得所述第一频带外的所述第二音频数据。

5. 如权利要求1或2所述的音频处理方法,其中所述对所述第一音频数据执行第二处理包括:

改变所述第一音频数据的频率。

6. 如权利要求5所述的音频处理方法,其中所述改变所述第一音频数据的频率包括:

对所述第一音频数据执行第一变换,以获得对应于所述第一音频数据的第一音频频率数据;

改变所述第一音频频率数据的频率值,以获得第三音频频率数据;以及

对所述第三音频频率数执行第二变换,以获得所述第三音频数据,

其中所述第二变换是所述第一变换的逆变换。

7. 如权利要求1或2所述的音频处理方法,其中所述对所述第一音频数据执行第二处理还包括:

对对应于预定时间段的所述第一音频数据,执行所述第二处理。

8. 如权利要求1所述的音频处理方法,其中所述第一音频数据的特征、所述第一特征和所述第二特征包括声纹特征、音色特征和/或音调特征。

9. 一种电子设备,包括:

滤波单元,用于对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于所述第一频带外的第二音频数据;

变调单元,用于对所述第一音频数据执行第二处理,获得第三音频数据;以及

合成单元,用于合成所述第三音频数据和所述第二音频数据,生成处理后音频数据,

其中,所述第一音频数据和所述第三音频数据具有相同的时长,

其中所述合成单元提取所述第三音频数据中的时间戳标记;

基于所述时间戳标记,确定所述第三音频数据相对于所述第二音频数据的起始时间点和终止时间点;以及

基于所述起始时间点和终止时间点,对准并且组合所述第三音频数据和所述第二音频数据,以生成处理后音频数据,

其中,所述变调单元在所述第二处理是升调处理的情况下,将经所述升调处理后的所述第一音频数据分为一个个帧,在每一帧后面补偿该帧最后一部分的数据,获得第三音频数据,

其中,所述变调单元基于所述第一音频数据的特征,识别所述第一音频数据中具有第一特征的第一子音频数据和具有第二特征的第二子音频数据;以及

对所述第一子音频数据执行第一子处理,并且对所述第二子音频数据执行第二子处理。

10. 如权利要求9所述的电子设备,其中所述第一频带为特定的频率范围内的频带,所述特定的频率范围对应于人发声的频率范围。

11. 如权利要求9或10所述的电子设备,其中所述合成单元合成频率改变的所述第一音频数据以及保持不变的所述第二音频数据。

12. 如权利要求9或10所述的电子设备,其中所述滤波单元包括第一滤波子单元,用于对待处理的音频数据执行第一滤波处理,以获得所述第一频带内的所述第一音频数据,并且对待处理的音频数据执行第二滤波处理,以获得所述第一频带外的所述第二音频数据。

13. 如权利要求9或10所述的电子设备,其中所述变调单元改变所述第一音频数据的频率。

14. 如权利要求13所述的电子设备,其中所述变调单元对所述第一音频数据执行第一变换,以获得对应于所述第一音频数据的第一音频频率数据;

改变所述第一音频频率数据的频率值,以获得第三音频频率数据;以及

对所述第三音频频率数执行第二变换,以获得所述第三音频数据。

15. 如权利要求9或10所述的电子设备,其中所述变调单元对对应于预定时间段的所述第一音频数据,执行所述第二处理。

16. 如权利要求9所述的电子设备,其中所述第一音频数据的特征、所述第一特征和所

述第二特征包括声纹特征、音色特征和/或音调特征。

## 音频处理方法和电子设备

### 技术领域

[0001] 本发明涉及音频处理领域,更具体地,本发明涉及一种音频处理方法和使用该音频处理方法的电子设备。

### 背景技术

[0002] 在当前使用的诸如变声处理的音频处理中,通常是针对整个音频数据进行音调转换。这样的音频处理缺乏精准的通道变声,对于音频数据中包含的人声信号和背景音信号都不加区分进行了处理。

[0003] 因此,希望提供一种音频处理方法和使用该音频处理方法的电子设备,其能够针对音频数据中包含的目标音频信号(诸如人声信号)执行所需的音频处理,而保持其他非目标音频信号不变,从而实现了针对目标音频信号的精准音频处理。

### 发明内容

[0004] 有鉴于上述情况,本发明提供了一种音频处理方法和使用该音频处理方法的电子设备。

[0005] 根据本公开的一个实施例,提供了一种音频处理方法,包括:对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于所述第一频带外的第二音频数据;对所述第一音频数据执行第二处理,获得第三音频数据;以及合成所述第三音频数据和所述第二音频数据,生成处理后音频数据。

[0006] 此外,根据本公开的一个实施例的音频处理方法,其中所述第一频带为特定的频率范围内的频带,所述的特定频率范围对应于人发声的频率范围。

[0007] 此外,根据本公开的一个实施例的音频处理方法,其中所述对待处理音频数据执行第一处理以及所述对所述第一音频数据执行第二处理,获得频率改变的所述第一音频数据以及保持不变的所述第二音频数据。

[0008] 此外,根据本公开的一个实施例的音频处理方法,其中所述对待处理的音频数据执行第一处理包括:对待处理的音频数据执行第一滤波处理,以获得所述第一频带内的所述第一音频数据,并且对待处理的音频数据执行第二滤波处理,以获得所述第一频带外的所述第二音频数据。

[0009] 此外,根据本公开的一个实施例的音频处理方法,其中所述对所述第一音频数据执行第二处理包括:改变所述第一音频数据的频率。

[0010] 此外,根据本公开的一个实施例的音频处理方法,其中所述改变所述第一音频数据的频率包括:对所述第一音频数据执行第一变换,以获得对应于所述第一音频数据的第一音频频率数据;改变所述第一音频频率数据的频率值,以获得第三音频频率数据;以及对所述第三音频频率数执行第二变换,以获得所述第三音频数据,其中所述第一变换是所述第一变换的逆变换。

[0011] 此外,根据本公开的一个实施例的音频处理方法,其中所述对所述第一音频数据

执行第二处理还包括:对对应于预定时间段的所述第一音频数据,执行所述第二处理。

[0012] 此外,根据本公开的一个实施例的音频处理方法,其中所述对所述第一音频数据执行第二处理还包括:基于所述第一音频数据的特征,识别所述第一音频数据中具有第一特征的第一子音频数据和具有第二特征的第二子音频数据;以及对所述第一子音频数据执行第一子处理,并且对所述第二子音频数据执行第二子处理。

[0013] 此外,根据本公开的一个实施例的音频处理方法,其中所述特征包括声纹特征、音色特征和/或音调特征。

[0014] 此外,根据本公开的一个实施例的音频处理方法,其中所述合成所述第三音频数据和所述第二音频数据,生成处理后音频数据包括:提取所述第三音频数据中的时间戳标记;基于所述时间戳标记,确定所述第三音频数据相对于所述第二音频数据的起始时间点和终止时间点;以及基于所述起始时间点和终止时间点,对准并且组合所述第三音频数据和所述第二音频数据,以生成处理后音频数据。

[0015] 根据本发明的另一实施例,提供了一种电子设备,包括:滤波单元,用于对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于所述第一频带外的第二音频数据;变调单元,用于对所述第一音频数据执行第二处理,获得第三音频数据;以及合成单元,用于合成所述第三音频数据和所述第二音频数据,生成处理后音频数据。

[0016] 此外,根据本发明的另一实施例的电子设备,其中所述第一频带为特定的频率范围内的频带,所述的特定频率范围对应于人发声的频率范围。

[0017] 此外,根据本发明的另一实施例的电子设备,其中所述合成单元合成频率改变的所述第一音频数据以及保持不变的所述第二音频数据。

[0018] 此外,根据本发明的另一实施例的电子设备,其中所述滤波单元包括第一滤波单元,用于对待处理的音频数据执行第一滤波处理,以获得所述第一频带内的所述第一音频数据,并且对待处理的音频数据执行第二滤波处理,以获得所述第一频带外的所述第二音频数据。

[0019] 此外,根据本发明的另一实施例的电子设备,其中所述变调单元改变所述第一音频数据的频率。

[0020] 此外,根据本发明的另一实施例的电子设备,其中所述变调单元对所述第一音频数据执行第一变换,以获得对应于所述第一音频数据的第一音频频率数据;改变所述第一音频频率数据的频率值,以获得第三音频频率数据;以及对所述第三音频频率数据执行第二变换,以获得所述第三音频数据,

[0021] 此外,根据本发明的另一实施例的电子设备,其中所述变调单元对对应于预定时间段的所述第一音频数据,执行所述第二处理。

[0022] 此外,根据本发明的另一实施例的电子设备,其中所述变调单元基于所述第一音频数据的特征,识别所述第一音频数据中具有第一特征的第一子音频数据和具有第二特征的第二子音频数据;以及对所述第一子音频数据执行第一子处理,并且对所述第二子音频数据执行第二子处理。

[0023] 此外,根据本发明的另一实施例的电子设备,其中所述特征包括声纹特征、音色特征和/或音调特征。

[0024] 此外,根据本发明的另一实施例的电子设备,其中所述合成单元提取所述第三音

频数据中的时间戳标记;基于所述时间戳标记,确定所述第三音频数据相对于所述第二音频数据的起始时间点和终止时间点;以及基于所述起始时间点和终止时间点,对准并且组合所述第三音频数据和所述第二音频数据,以生成处理后音频数据。

[0025] 要理解的是,前面的一般描述和下面的详细描述两者都是示例性的,并且意图在于提供要求保护的技术的进一步说明。

## 附图说明

[0026] 通过结合附图对本发明实施例进行更详细的描述,本发明的上述以及其它目的、特征和优势将变得更加明显。附图用来提供对本发明实施例的进一步理解,并且构成说明书的一部分,与本发明实施例一起用于解释本发明,并不构成对本发明的限制。在附图中,相同的参考标号通常代表相同部件或步骤。

[0027] 图1是概述根据本发明实施例的音频处理方法的流程图;

[0028] 图2是图示根据本发明实施例的电子设备的结构框图;

[0029] 图3是图示根据本发明实施例的音频处理方法的第二示例的流程图;

[0030] 图4是图示根据本发明实施例的音频处理过程的框图;

[0031] 图5是图示根据本发明实施例的音频处理方法的第二示例的流程图;

[0032] 图6是图示根据本发明实施例的音频处理的示意图;以及

[0033] 图7是图示根据本发明实施例的音频处理方法的第三示例的流程图。

## 具体实施方式

[0034] 为了使得本发明的目的、技术方案和优点更为明显,下面将参照附图详细描述根据本发明的示例实施例。显然,所描述的实施例仅仅是本发明的一部分实施例,而不是本发明的全部实施例,应理解,本发明不受这里描述的示例实施例的限制。基于本公开中描述的本发明实施例,本领域技术人员在没有付出创造性劳动的情况下所得到的所有其它实施例都应落入本发明的保护范围之内。

[0035] 以下,将参考附图详细描述本发明的优选实施例。

[0036] 图1是概述根据本发明实施例的音频处理方法的流程图。根据本发明实施例的音频处理方法包括以下步骤。

[0037] 在步骤S101中,对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于第一频带外的第二音频数据。如下将详细描述的,待处理音频数据是由根据本发明实施例的电子设备采集的音频数据。在本发明的一个实施例中,在所述待处理音频数据中,包括分别处于不同频带的音频信号。一般地,声音是由物体振动产生的,声音的基本特征要素包括:音调、音强和音色。具体地,不同的振动频率即为不同的音调。在采集声音所获取的音频数据中,人声和背景具有不同特征,处于不同的频带中。在本发明的一个优选实施例中,所述第一频带为特定的频率范围内的频带,所述的特定频率范围对应于人发声的频率范围,例如64Hz-523Hz。以下,将参照附图进一步详细描述如何通过执行第一处理,获得处于第一频带内的第一音频数据和处于第一频带外的第二音频数据。此后,处理进到步骤S102。

[0038] 在步骤S102中,对第一音频数据执行第二处理,获得第三音频数据。如下将详细描

述的,对第一音频数据执行第二处理,获得频率改变的所述第一音频数据。也就是说,所述第二处理是用于改变第一音频数据的处理。如上参照步骤S101所述,第一音频数据是处于第一频带内(例如,对应于人发声的频率范围)的音频数据,那么通过所述第二处理将改变第一音频数据中人声的频率。

[0039] 例如,通常采用12平均律,即将声音分成音级CDEFGAB。某音级与网上数8个音级之间的距离称为“纯八度”。12平均律将一个纯八度分成12个均等的半音,相邻半音间的物理振动频率相差 $2^{1/12}$ 倍。也就是说,各半音之间的频率成等比关系。

[0040] 在本发明的一个实施例中,第一音频数据的信号频率例如为 $f$ ,那么通过所述第二处理获得的第三音频数据的信号频率为 $f'$

[0041]  $f' = f \times 2^{d/12}$ ,  $d = \pm 1, \pm 2, \pm 3 \dots$  (1)

[0042] 当 $d > 0$ 时,代表第二处理升高第一音频数据的频率,即将第一音频数据升调;当 $d < 0$ 时,代表第二处理降低第一音频数据的频率,即将第一音频数据降调。

[0043] 在获得频率改变的所述第一音频数据(即,第三音频数据)和保持不变的所述第二音频数据之后,处理进到步骤S103。

[0044] 在步骤S103中,合成第三音频数据和第二音频数据,生成处理后音频数据。具体地,将频率改变的所述第一音频数据与保持频率不变的所述第二音频数据合成,获得处理后的音频数据。以下,将参照附图进一步详细描述根据本发明实施例的音频处理方法中的音频数据合成。

[0045] 如上所述,与在根据本发明实施例的音频处理方法开始初输入的对待处理音频数据相比,在处理后的音频数据中,对应于特定的频率范围内(例如,对应于人声的64Hz-523Hz)的第一音频数据出现频率的改变(例如,升调或降调,即女声转换为男生,或者老人声音转换为儿童声音),而非特定的频率范围内的第二音频数据则保持频率不变,从而在合成的处理后的音频数据中仅仅对人声进行了精确地变声变调处理,而保持背景声不变。

[0046] 图2是图示根据本发明实施例的电子设备的结构框图。如图2所示的电子设备10用于执行参照图1描述的根据本发明实施例的音频处理方法。

[0047] 所述电子设备10优选地例如是具有音频处理能力的电子设备,包括但不限于便携式电子设备(诸如,智能电话、个人数字助理、平板电脑)、个人计算机、家庭影院系统、商用卡拉OK娱乐系统等。

[0048] 如图2所示,根据本发明实施例的电子设备10包括处理模块100和音频输入输出模块200。容易理解的是,图1中为了简化描述仅仅示出与本发明紧密相关的组件,根据本发明实施例的电子设备10当然还可以包括其它组件,诸如显示模块、存储模块等。

[0049] 所述处理模块100用于执行根据本发明实施例的音频处理。在本发明的一个实施例中,所述处理模块100可以由电子设备10的中央处理单元(CPU)配置。可替代地,所述处理模块100可以由电子设备10的专用音频处理单元(APU)配置。

[0050] 所述音频输入输出模块200用于获取待处理音频数据,并且将经由所述处理模块100执行音频处理之后的处理后音频数据输出。在本发明的一个实施例中,所述音频输入输出模块200可以利用诸如麦克风等音频采集单元采集待处理音频数据。可替代地,所述音频输入输出模块200可以从所述电子设备10中的存储模块(未示出)检索并且提取待处理音频数据,或者经由有线或者无线通信信道从另一电子设备接收待处理音频数据。在经由所述



处理模块100执行音频处理之后,所述音频输入输出模块200可以经由诸如扬声器等音频输出单元输出处理后音频数据。可替代地,所述音频输入输出模块200可以将处理后音频数据存储在所述电子设备10中的存储模块,或者经由有线或者无线通信信道将处理后音频数据发送到另一电子设备。

[0051] 更具体地,如图2所示,所述处理模块100包括滤波单元110、变调单元120和合成单元130。所述滤波单元110用于对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于所述第一频带外的第二音频数据。所述变调单元120用于对所述第一音频数据执行第二处理,获得第三音频数据。所述合成单元130用于合成所述第三音频数据和所述第二音频数据,生成处理后音频数据。以下,将参照附图进一步详述由包括滤波单元110、变调单元120和合成单元130的所述处理模块100执行的根据本发明实施例的音频处理方法。

[0052] 参照图3和图4描述根据本发明实施例的音频处理方法的第一示例。图3是图示根据本发明实施例的音频处理方法的第一示例的流程图。图4是图示根据本发明实施例的音频处理过程的框图。

[0053] 如图3所示,根据本发明实施例的音频处理方法的第一示例包括以下步骤。

[0054] 在步骤S301中,对待处理的音频数据执行第一滤波处理,以获得第一频带内的第一音频数据。

[0055] 在步骤S301之后,或者与步骤S301同时,在步骤S302中,对待处理的音频数据执行第二滤波处理,以获得第一频带外的第二音频数据。

[0056] 如图4所示,为了获得第一音频数据和第二音频数据,待处理的音频数据 $A_0$ 分别输入滤波单元110中的第一滤波子单元111和第二滤波子单元。具体地,所述第一滤波子单元111执行第一滤波处理(例如,第一带通滤波处理),获得并且输出第一频带内(例如,对应于人声的64Hz-523Hz)的第一音频数据 $A_1$ 。所述第二滤波子单元112执行第二滤波处理(例如,第二带通滤波处理),获得第一频带外的第二音频数据 $A_2$ 。

[0057] 返回参照图3,在步骤S302后,处理进到步骤S303。在步骤S303中,对第一音频数据执行第一变换,以获得对应于第一音频数据的第一音频频率数据。

[0058] 如图4所示,从所述第一滤波子单元111输出的第一音频数据 $A_1$ 进入变调单元120中的第一变换子单元121。所述第一变换子单元121对第一音频数据执行第一变换。在本发明的一个实施例中,所述第一变换为快速傅氏变换,用于将第一音频数据 $A_1$ 转换为频域的第一音频频率数据 $A_{1f}$ 。

[0059] 返回参照图3,在步骤S303后,处理进到步骤S304。在步骤S304中,改变第一音频频率数据的频率值,以获得第三音频频率数据。

[0060] 如图4所示,从所述第一变换子单元121输出的第一音频频率数据 $A_{1f}$ 进入变调单元120中的频率转换子单元122。所述频率转换子单元122改变第一音频频率数据 $A_{1f}$ 的频率,以获得第三音频频率数据 $A_{3f}$ 。例如,如果提高第一音频频率数据 $A_{1f}$ 的频率,即执行升调处理,则可能获得将男声变为女声,将老人声音变为儿童声音的效果;如果降低第一音频频率数据 $A_{1f}$ 的频率,即执行降调处理,则可能获得将女声变为男声,将儿童声音变为老人声音的效果。

[0061] 具体地,对于升调处理,将第一音频频率数据 $A_{1f}$ 的谱线向高频伸展。

[0062] 例如,在第一音频频率数据 $A_{1f}$ 的频谱信号中,通过每隔一定距离执行抽选,以实现降低采样从而升调。类似地,对于降调处理,将第一音频频率数据 $A_{1f}$ 的谱线向低频收缩。例如,在第一音频频率数据 $A_{1f}$ 的频谱信号中,通过在相邻两个数据点之间内插新的数据点,以实现增采样从而降调。

[0063] 例如,假设变调因子为:

$$[0064] \quad \alpha = \frac{f'}{f} = \frac{M}{L} \quad (2)$$

[0065] 其中 $M$ 和 $L$ 为正整数,任意有理数倍频率的变调,可以表示为:

$$[0066] \quad X_1(n) = X\left(\left[\frac{n}{L}\right]\right) + \frac{\text{mod}(n, L)}{L} \times \{x(\lfloor \frac{n}{L} \rfloor + 1) - x(\lfloor \frac{n}{L} \rfloor)\}, \quad n = 0, 1, 2, \dots, (N-1) \times L + 1 \quad (3)$$

[0067] 其中, $N$ 是帧长, $[\ ]$ 表示取整运算, $\text{mod}$ 为取模运算。当 $M > L$ 时实现升调, $M < L$ 时实现降调。

[0068] 主要注意的是,变调后,为了保持总体音频数据的时长不变,需要对变调后的数据进行适当的处理。例如,在升调的情况下,由于升调可能会使原始数据变短,因此可以先将经变调后处理后的数据分为一个个帧,在每一帧后面补偿该帧最后一部分的数据。在本发明的一个实施例中,补偿的数据的具体长度可以为帧长的 $(1 - \frac{M}{L})$ 倍。

[0069] 返回参照图3,在步骤S304后,处理进到步骤S305。在步骤S305中,对第三音频频率数执行第二变换,以获得第三音频数据。

[0070] 如图4所示,从频率转换子单元122输出的第三音频频率数据 $A_{3f}$ 进入变调单元120中的第二变换子单元123。所述第二变换子单元123对第三音频频率数据 $A_{3f}$ 执行第二变换。在本发明的一个实施例中,所述第二变换为快速傅氏逆变换,用于将第三音频频率数据 $A_{3f}$ 转换为时域的第三音频数据 $A_3$ 。

[0071] 返回参照图3,在步骤S305后,处理进到步骤S306。在步骤S306中,合成第三音频数据和第二音频数据,生成处理后音频数据。

[0072] 如图4所示,从变调单元120中的第二变换子单元123输出的第三音频数据 $A_3$ 与从所述第二滤波子单元112输出的第二音频数据 $A_2$ 进入合成单元130。所述合成单元130将其合成以生成处理后音频数据 $A_p$ 。

[0073] 在参照图3和图4描述的根据本发明实施例的音频处理方法的第一示例中,对应于特定的频率范围内(例如,对应于人声的64Hz-523Hz)的第一音频数据执行改变频率的变调处理,同时保持处理后的音频数据的总体长度不变以避免失真。从而在合成的处理后的音频数据中仅仅对人声进行了精确地变声变调处理,而保持背景声不变。

[0074] 参照图5和图6描述根据本发明实施例的音频处理方法的第二示例。图5是图示根据本发明实施例的音频处理方法的第二示例的流程图。图6是图示根据本发明实施例的音频处理的示意图。

[0075] 如图5所示,根据本发明实施例的音频处理方法的第二示例包括以下步骤。

[0076] 在步骤S501中,对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于第一频带外的第二音频数据。步骤S501中的处理等价于参照图3描述的步骤S301和S302以及参照图4描述的在第一滤波子单元111和第二滤波子单元112中执行的处理。此后,处理进到步骤S502。

[0077] 在步骤S502中,对对应于预定时间段的第一音频数据,执行第二处理,获得第三音频数据。不同于参照图3描述的根据本发明实施例的音频处理方法的第一示例,在根据本发明实施例的音频处理方法的第二示例中,仅对于对应于预定时间段的第一音频数据 $A_1$ 执行第二处理。在本发明的一个实施例中,所述预定时间段可以由用户指定,从而仅对该预定时间段的人声执行变调,而对于第一音频数据 $A_1$ 中其他时段的人声不执行变调。在此情况下,在预定时间段对单个人声进行变调(例如,男声变为女声),从而预定时间段的变调声与预定时间段外的未变调声组合获得男女声对唱的效果。此后,处理进到步骤S503。

[0078] 步骤S503到S505对应于参照图3的步骤S306描述的合成过程。在此将通过三个步骤S503到S505以及图6的示意图进一步详细描述该合成过程。

[0079] 在步骤S503中,提取第三音频数据中的时间戳标记。

[0080] 如图6所示,所述合成单元130提取经由所述变调单元120输出的第三音频数据 $A_3$ 的时间戳标记 $T_1$ 和 $T_2$ ,以及经由所述滤波单元110输出的第二音频数据 $A_2$ 的时间戳标记 $T_0$ 和 $T_T$ ,

[0081] 此后,处理进到步骤S504。在步骤S504中,基于时间戳标记,确定第三音频数据相对于第二音频数据的起始时间点和终止时间点。

[0082] 如图6所示,所述合成单元130基于在步骤S503中提取的第三音频数据 $A_3$ 的时间戳标记 $T_1$ 和 $T_2$ 以及第二音频数据 $A_2$ 的时间戳标记 $T_0$ 和 $T_T$ ,确定第三音频数据 $A_3$ 相对于第二音频数据 $A_2$ 的起始时间点和终止时间点。

[0083] 此后,处理进到步骤S505。在步骤S505中,基于起始时间点和终止时间点,对准并且组合第三音频数据和第二音频数据,以生成处理后音频数据。

[0084] 如图6所示,所述合成单元130基于在步骤S504中获取的起始时间点和终止时间点,对准并且组合第三音频数据 $A_3$ 和第二音频数据 $A_2$ ,以生成处理后音频数据 $A_p$ 。

[0085] 在参照图5和图6描述的根据本发明实施例的音频处理方法的第二示例中,对对应于预定时间段的第一音频数据执行改变频率的变调处理,实现了同时包括变调的人声和未变调的人声的组合。此外,在合成的处理中基于各个音频数据中的时间戳标记,精准地实现了各个音频数据的对准,避免了由于失准导致的不必要的杂音。

[0086] 参照图7描述根据本发明实施例的音频处理方法的第三示例。如图7所示,根据本发明实施例的音频处理方法的第三示例包括以下步骤。

[0087] 在步骤S701中,对待处理音频数据执行第一处理,获得处于第一频带内的第一音频数据和处于第一频带外的第二音频数据。步骤S701中的处理等价于参照图3描述的步骤S301和S302参照图5描述的步骤S501、以及参照图4描述的在第一滤波单元111和第二滤波单元112中执行的处理。此后,处理进到步骤S702。

[0088] 步骤S702和步骤S703由所述变调单元120执行。不同于参照图3描述的步骤S303到S305以及参照图5描述的步骤S502,在根据本发明实施例的音频处理方法的第三示例中,在步骤S702中,基于第一音频数据的特征,识别第一音频数据中具有第一特征的第一子音频数据和具有第二特征的第二子音频数据。在本发明的一个实施例中,所述特征包括声纹特征、音色特征和/或音调特征。

[0089] 例如,所述变调单元120基于第一音频数据的声纹特征,识别第一音频数据中的不同用户。并且,当识别到满足第一特征(即,具有特定第一声纹)的第一子音频数据和/或满

足第二特征(即,具有特定第二声纹)的第二子音频数据时,处理进到步骤S703。

[0090] 在步骤S703中,所述变调单元120对第一子音频数据执行第一子处理,并且对第二子音频数据执行第二子处理,获得第三音频数据。也就是说,对于具有不同声纹特征的第一子音频数据和第二子音频数据,所述变调单元120分别执行不同的第一子处理和第二子处理。即,实现了对于具有特定声纹特征的特定的用户执行特定的变调处理。

[0091] 类似地,所述变调单元120基于第一音频数据的音调特征,识别第一音频数据中的处于不同音调(即,频带)的音频数据。当识别到满足第一特征(即,具有特定第一频带)的第一子音频数据和/或满足第二特征(即,具有特定第二频带)的第二子音频数据时,处理进到步骤S703。

[0092] 在步骤S703中,所述变调单元120对第一子音频数据执行第一子处理,并且对第二子音频数据执行第二子处理,获得第三音频数据。也就是说,对于具有不同音调特征的第一子音频数据和第二子音频数据,所述变调单元120分别执行不同的第一子处理和第二子处理。例如,对于特定第一频带的第一子音频数据执行升调处理,而对于特定第二频带的第二子音频数据执行降调处理。

[0093] 此后,处理进到步骤S704。步骤S704到S706的合成处理分别与参照图5描述的步骤S503到S505相同,在此将省略其重复描述。

[0094] 在参照图7描述的根据本发明实施例的音频处理方法的第三示例中,基于第一音频数据的特征,对于满足不同特征的子音频数据执行不同的子处理,从而实现了基于采集的音频信号的具体特征,执行相应的音频子处理。

[0095] 以上参照图1到图7描述了根据本发明实施例的音频处理方法和使用该音频处理方法的电子设备,其能够针对音频数据中包含的目标音频信号(诸如人声信号)执行所需的音频处理,而保持其他非目标音频信号不变,从而实现了针对目标音频信号的精准音频处理。

[0096] 需要说明的是,在本说明书中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0097] 最后,还需要说明的是,上述一系列处理不仅包括以这里所述的顺序按时间序列执行的处理,而且包括并行或分别地、而不是按时间顺序执行的处理。

[0098] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到本发明可借助软件加必需的硬件平台的方式来实现,当然也可以全部通过硬件来实施。基于这样的理解,本发明的技术方案对背景技术做出贡献的全部或者部分可以以软件产品的形式体现出来,该计算机软件产品可以存储在存储介质中,如ROM/RAM、磁碟、光盘等,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行本发明各个实施例或者实施例的某些部分所述的方法。

[0099] 以上对本发明进行了详细介绍,本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;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。

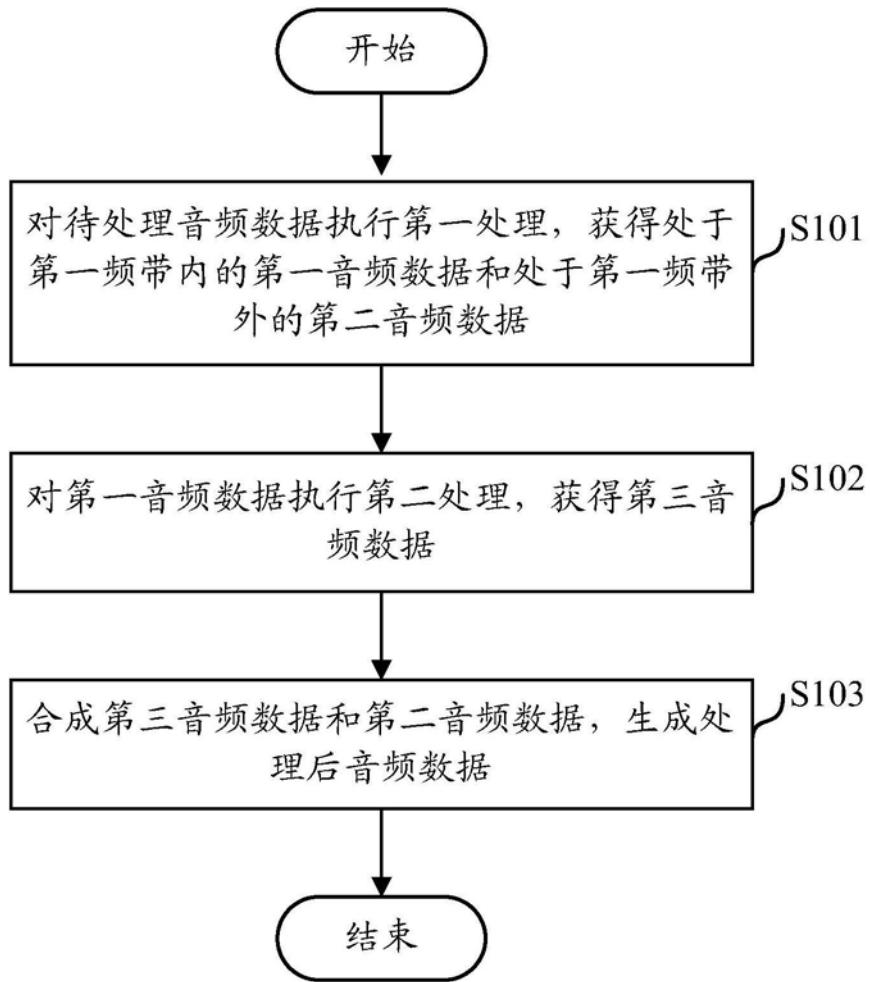


图1

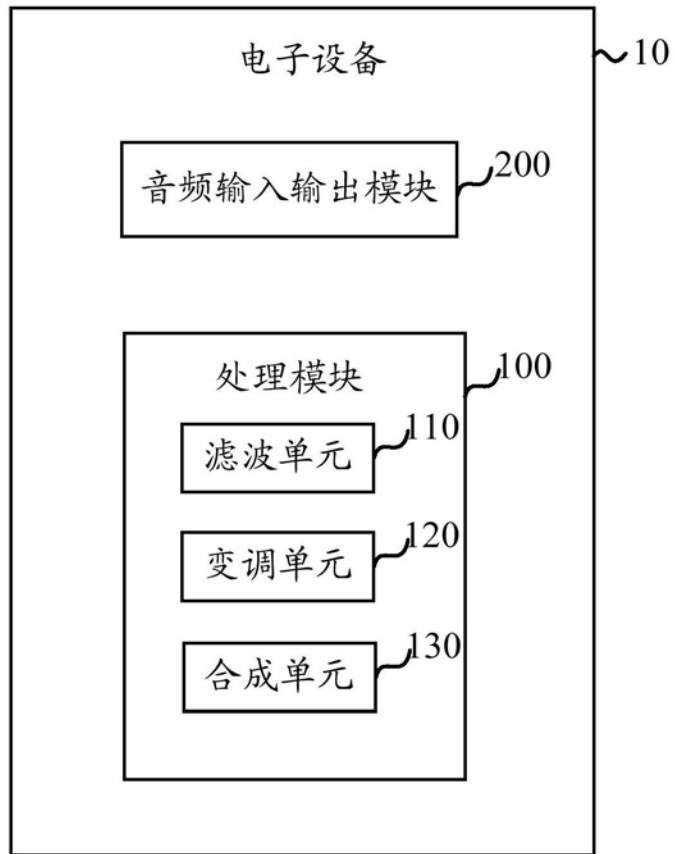


图2

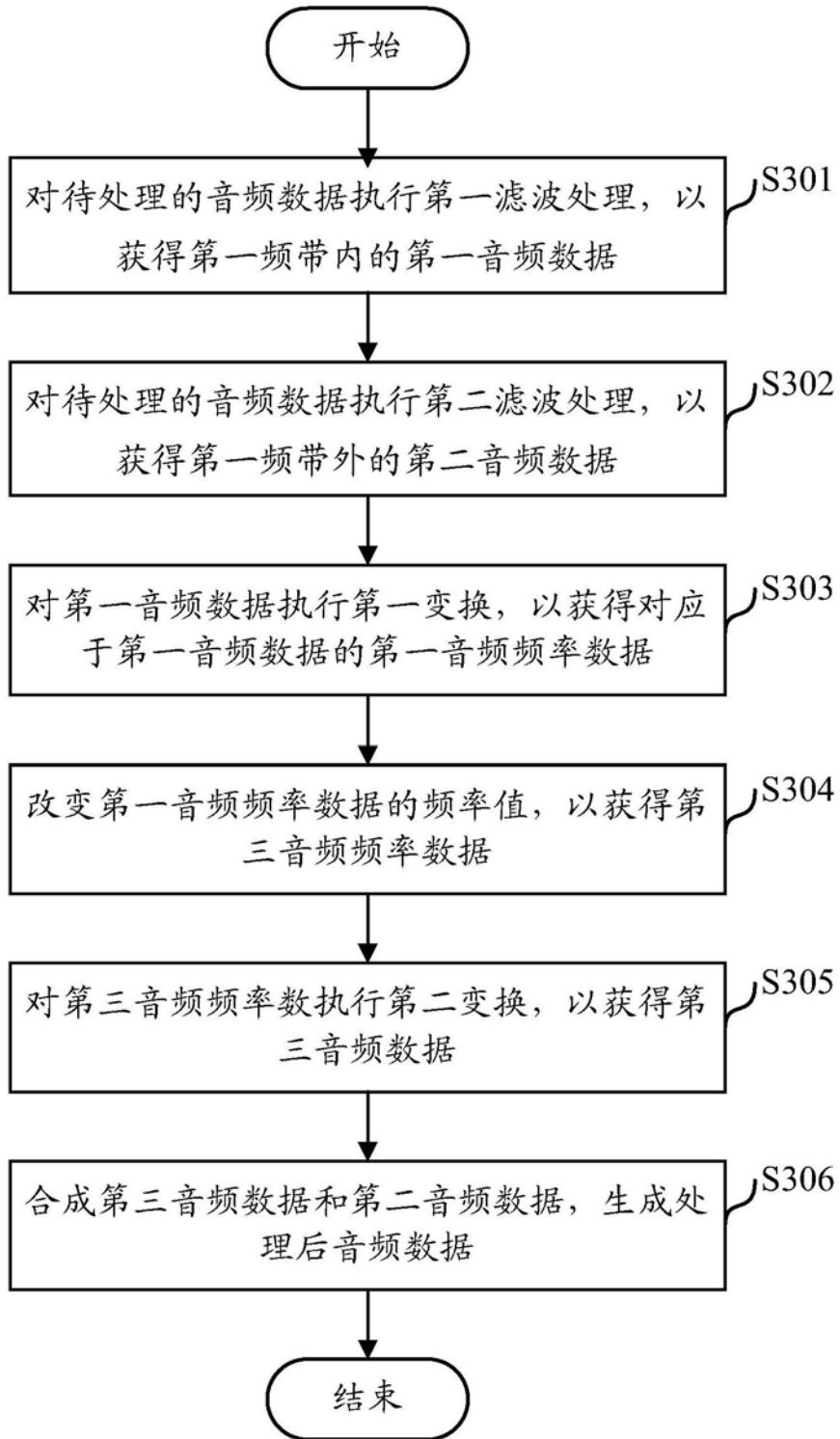


图3



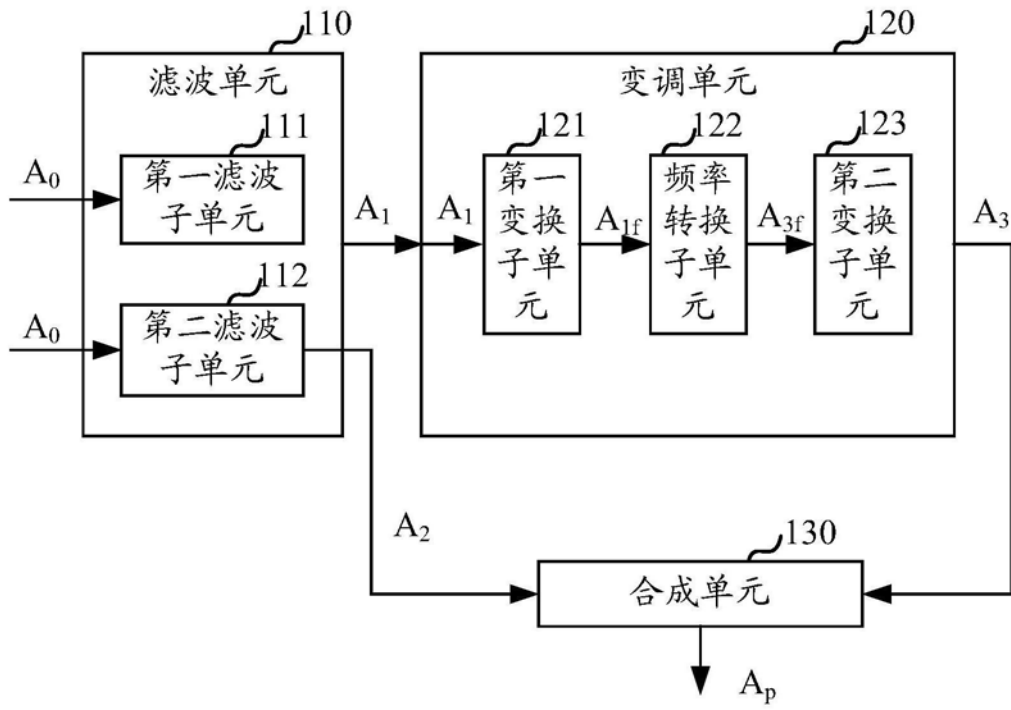


图4

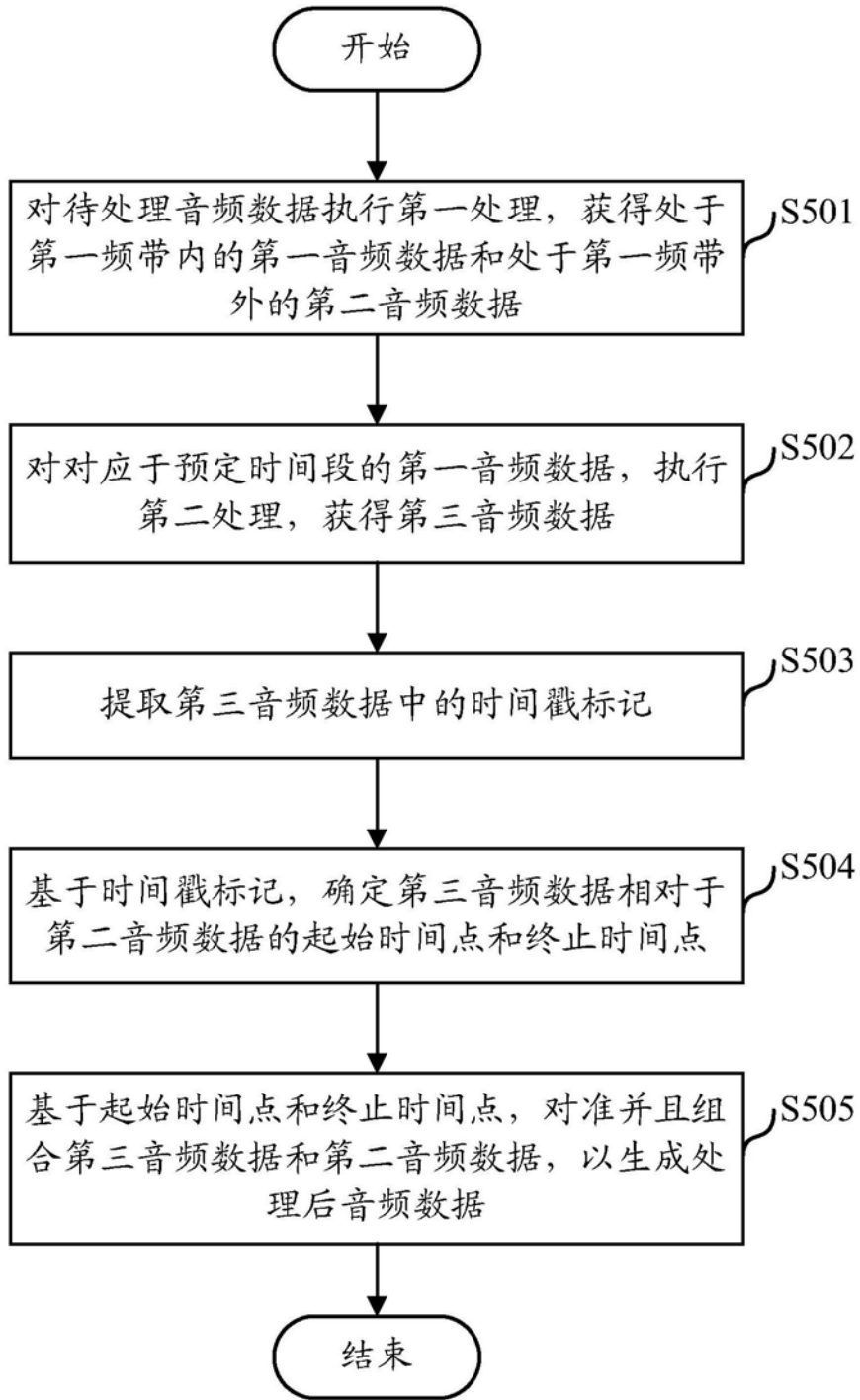


图5

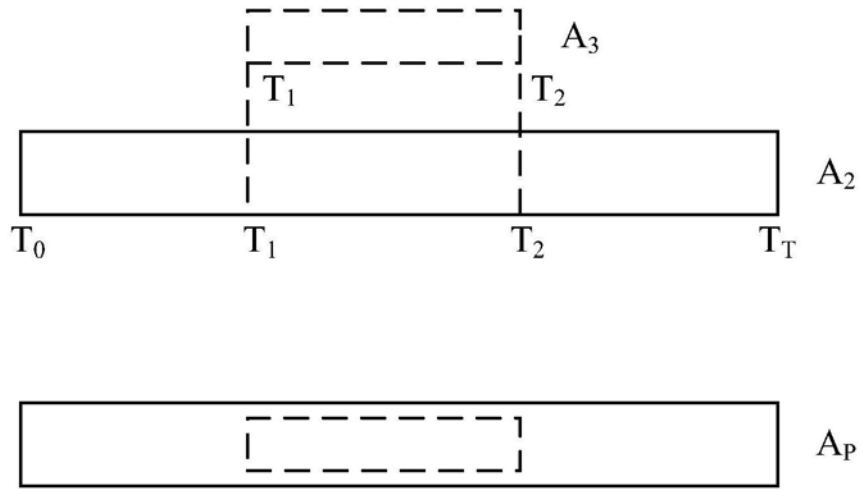


图6

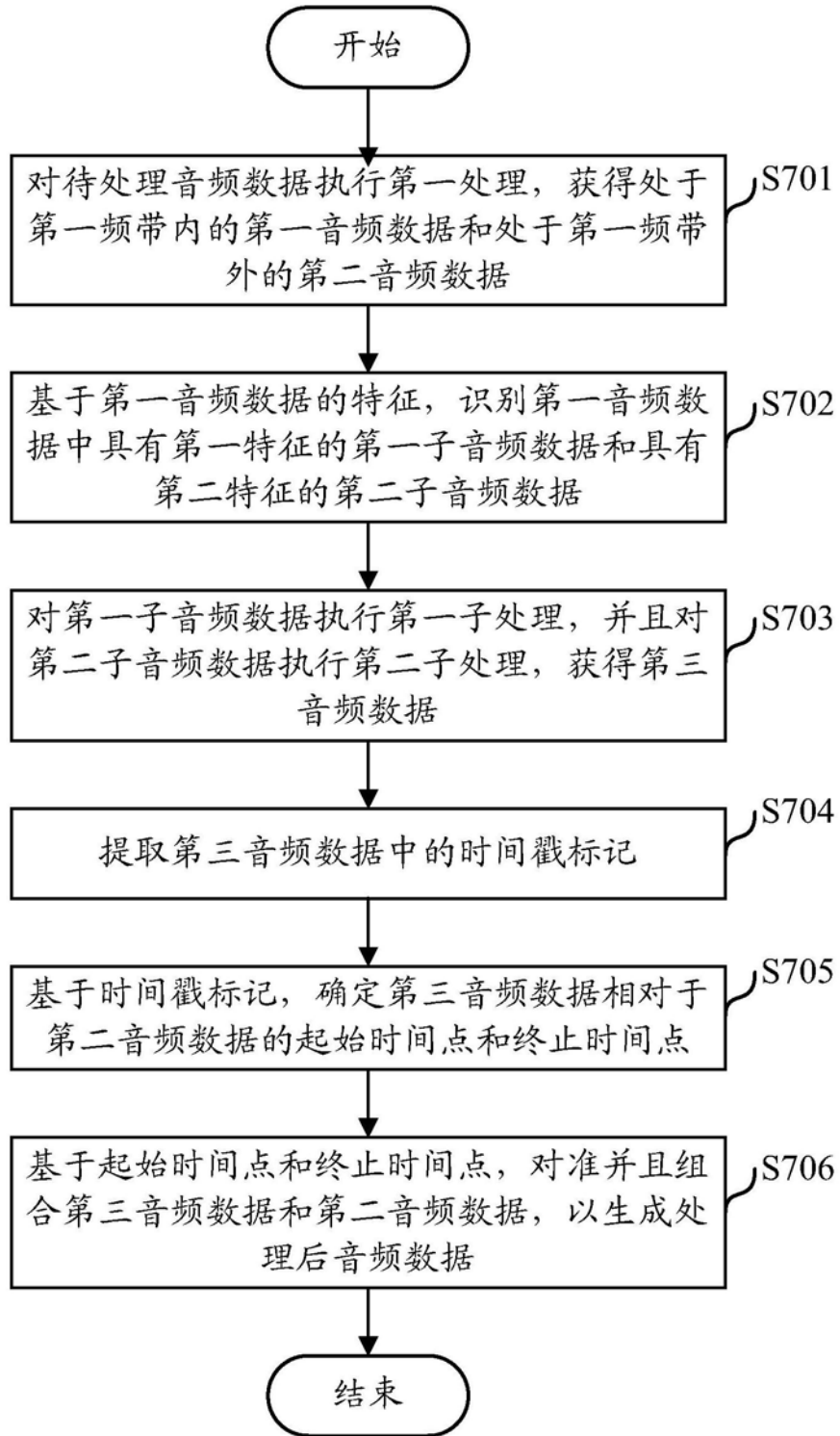


图7