



(12) 发明专利

(10) 授权公告号 CN 112802485 B

(45) 授权公告日 2021.07.02

(21) 申请号 202110387976.9

G10L 19/26 (2013.01)

(22) 申请日 2021.04.12

(56) 对比文件

(65) 同一申请的已公布的文献号

CN 103632671 A, 2014.03.12

申请公布号 CN 112802485 A

CN 108632625 A, 2018.10.09

US 2008/0312936 A1, 2008.12.18

(43) 申请公布日 2021.05.14

审查员 李召卿

(73) 专利权人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区

科技中一路腾讯大厦35层

(72) 发明人 梁俊斌

(74) 专利代理机构 北京三高永信知识产权代理

有限责任公司 11138

代理人 祝亚男

(51) Int. Cl.

G10L 19/04 (2013.01)

G10L 19/08 (2013.01)

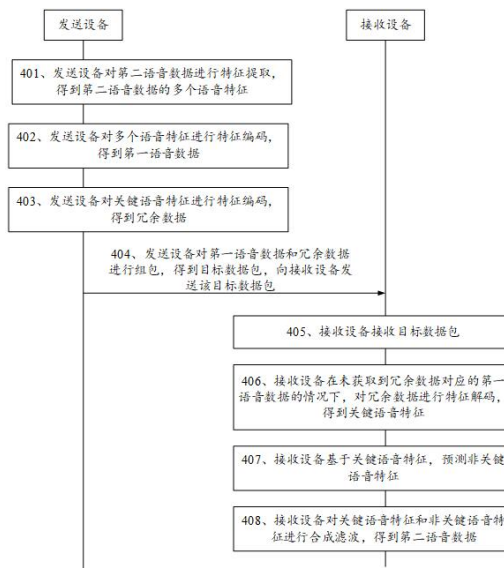
权利要求书3页 说明书29页 附图12页

(54) 发明名称

语音数据处理方法、装置、计算机设备及存储介质

(57) 摘要

本申请实施例公开了一种语音数据处理方法、装置、计算机设备及存储介质,属于计算机技术领域。该方法包括:接收目标数据包,目标数据包是发送设备对第一语音数据和冗余数据进行组包后发送的,第一语音数据是对第二语音数据的多个语音特征进行特征编码得到的,多个语音特征中包括关键语音特征和非关键语音特征,冗余数据是关键语音特征进行特征编码得到的;在目标数据包中未获取到第一语音数据的情况下,对冗余数据进行特征解码,得到关键语音特征;基于关键语音特征,预测非关键语音特征;对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据。在丢失语音数据的情况下,通过预测非关键语音特征,能够达到抗丢失语音数据的效果。



1. 一种语音数据处理方法,其特征在于,所述方法包括:

接收目标数据包,所述目标数据包是发送设备对第一语音数据和冗余数据进行组包后发送的,所述第一语音数据是对第二语音数据的多个语音特征进行特征编码得到的,所述多个语音特征中包括关键语音特征和非关键语音特征,所述冗余数据是对所述关键语音特征进行特征编码得到的,所述第二语音数据包括多个语音数据帧,所述冗余数据中包括与所述多个语音数据帧对应的多个冗余子数据;

在所述目标数据包中未获取到所述第一语音数据的情况下,对所述冗余数据进行特征解码,得到每个冗余子数据对应的关键语音特征;

基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征;

对所述每个冗余子数据对应的关键语音特征和非关键语音特征进行合成滤波,得到所述每个冗余子数据对应的语音数据帧;

按照所述多个冗余子数据的排列顺序,对得到的多个语音数据帧进行合并,得到所述第二语音数据。

2. 根据权利要求1所述的方法,其特征在于,所述关键语音特征包括线谱对参数、基音周期参数和第一增益参数,所述非关键语音特征包括第二激励参数,所述基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征,包括:

基于所述每个冗余子数据对应的所述基音周期参数和所述第一增益参数,确定所述每个冗余子数据对应的第一激励参数,所述第一激励参数用于反映所述第二语音数据中具有相关性的语音特征;

基于所述每个冗余子数据对应的所述线谱对参数和所述第一激励参数,预测所述每个冗余子数据对应的所述第二激励参数,所述第二激励参数用于反映所述第二语音数据中不具有相关性的语音特征。

3. 根据权利要求1所述的方法,其特征在于,所述基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征,包括:

对于第一个冗余子数据之后的每个冗余子数据,获取所述冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧;

基于所述冗余子数据对应的关键语音特征、所述上一个冗余子数据对应的非关键语音特征以及所述上一个冗余子数据对应的语音数据帧,预测所述冗余子数据对应的非关键语音特征。

4. 根据权利要求3所述的方法,其特征在于,所述关键语音特征包括所述第二语音数据的线谱对参数以及所述每个冗余子数据对应的基音周期参数和第一增益参数,所述基于所述冗余子数据对应的关键语音特征、所述上一个冗余子数据对应的非关键语音特征以及所述上一个冗余子数据对应的语音数据帧,预测所述冗余子数据对应的非关键语音特征,包括:

基于所述冗余子数据对应的基音周期参数和第一增益参数,确定所述冗余子数据对应的第一激励参数;

基于所述线谱对参数、所述冗余子数据对应的第一激励参数、所述上一个冗余子数据对应的非关键语音特征以及所述上一个冗余子数据对应的语音数据帧,预测所述冗余子数

据对应的非关键语音特征。

5. 根据权利要求1所述的方法,其特征在于,所述基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征,包括:

对于第一个冗余子数据,获取所述第一个冗余子数据对应的参考非关键语音特征和参考语音数据帧;

基于所述第一个冗余子数据对应的关键语音特征、所述参考非关键语音特征以及所述参考语音数据帧,预测所述第一个冗余子数据对应的非关键语音特征。

6. 根据权利要求1所述的方法,其特征在于,所述基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征,包括:

调用特征预测模型,基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征。

7. 根据权利要求6所述的方法,其特征在于,所述调用特征预测模型,基于所述关键语音特征,预测所述非关键语音特征之前,所述方法还包括:

对样本语音数据进行特征提取,得到所述样本语音数据对应的多个样本语音特征,所述多个样本语音特征中包括样本关键语音特征;

调用所述特征预测模型,基于所述样本关键语音特征进行预测,得到预测非关键语音特征;

对所述样本关键语音特征和所述预测非关键语音特征进行合成滤波,得到样本解码数据;

基于所述样本语音数据、所述样本解码数据、所述样本关键语音特征以及所述预测非关键语音特征,训练所述特征预测模型。

8. 根据权利要求7所述的方法,其特征在于,所述基于所述样本语音数据、所述样本解码数据、所述样本关键语音特征以及所述预测非关键语音特征,训练所述特征预测模型,包括:

将所述样本语音数据与所述样本解码数据之间的误差确定为残差特征;

将所述残差特征与所述样本关键语音特征之间的差值确定为样本非关键语音特征;

基于所述预测非关键语音特征与所述样本非关键语音特征之间的误差,训练所述特征预测模型。

9. 一种语音数据处理装置,其特征在于,所述装置包括:

数据包接收模块,用于接收目标数据包,所述目标数据包是发送设备对第一语音数据和冗余数据进行组包后发送的,所述第一语音数据是对第二语音数据的多个语音特征进行特征编码得到的,所述多个语音特征中包括关键语音特征和非关键语音特征,所述冗余数据是对所述关键语音特征进行特征编码得到的,所述第二语音数据包括多个语音数据帧,所述冗余数据中包括与所述多个语音数据帧对应的多个冗余子数据;

特征解码模块,用于在所述目标数据包中未获取到所述第一语音数据的情况下,对所述冗余数据进行特征解码,得到每个冗余子数据对应的关键语音特征;

特征预测模块,用于基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征;

合成滤波模块,用于对所述每个冗余子数据对应的关键语音特征和非关键语音特征进

行合成滤波,得到所述每个冗余子数据对应的语音数据帧;

所述合成滤波模块,还用于按照所述多个冗余子数据的排列顺序,对得到的多个语音数据帧进行合并,得到所述第二语音数据。

10.一种计算机设备,其特征在于,所述计算机设备包括处理器和存储器,所述存储器中存储有至少一条计算机程序,所述至少一条计算机程序由所述处理器加载并执行,以实现如权利要求1至8任一项所述的语音数据处理方法中所执行的操作。

11.一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有至少一条计算机程序,所述至少一条计算机程序由处理器加载并执行,以实现如权利要求1至8任一项所述的语音数据处理方法中所执行的操作。

语音数据处理方法、装置、计算机设备及存储介质

技术领域

[0001] 本申请实施例涉及计算机技术领域,特别涉及一种语音数据处理方法、装置、计算机设备及存储介质。

背景技术

[0002] 随着计算机技术和网络技术的发展,不同计算机设备之间能够进行语音数据的传输,而在传输语音数据的过程中,可能会出现丢包的情况,造成语音数据丢失,影响语音数据传输的效果。例如,在发送设备向接收设备发送某一帧语音数据时,该语音数据丢失,则会导致接收设备无法接收到该语音数据。因此,亟需提供一种能够达到抗丢包效果的语音数据处理方法。

发明内容

[0003] 本申请实施例提供了一种语音数据处理方法、装置、计算机设备及存储介质,能够达到数据的抗丢失效果。所述技术方案如下。

[0004] 一方面,提供了一种语音数据处理方法,所述方法包括:

[0005] 接收目标数据包,所述目标数据包是发送设备对第二语音数据和冗余数据进行组包后发送的,所述第二语音数据是对第一语音数据的多个语音特征进行特征编码得到的,所述多个语音特征中包括关键语音特征和非关键语音特征,所述冗余数据是所述关键语音特征进行特征编码得到的;

[0006] 在所述目标数据包中未获取到所述第二语音数据的情况下,对所述冗余数据进行特征解码,得到所述关键语音特征;

[0007] 基于所述关键语音特征,预测所述非关键语音特征;

[0008] 对所述关键语音特征和所述非关键语音特征进行合成滤波,得到所述第二语音数据。

[0009] 可选地,所述对所述关键语音特征和所述非关键语音特征进行合成滤波,得到所述第二语音数据,包括:

[0010] 基于所述线谱对参数,对所述第一激励参数和所述第二激励参数进行合成滤波,得到所述第二语音数据。

[0011] 可选地,所述基于所述基音周期参数和所述第一增益参数,确定第一激励参数,包括:

[0012] 在码本集合中查询与所述基音周期参数匹配的第一码本参数,所述码本集合中包括至少一组码本参数;

[0013] 将所述第一码本参数与所述第一增益参数的乘积,确定为所述第一激励参数。

[0014] 可选地,所述关键语音特征包括所述第二语音数据的线谱对参数以及所述每个冗余子数据对应的基音周期参数和第一增益参数,所述基于所述第一个冗余子数据对应的关键语音特征、所述参考非关键语音特征以及所述参考语音数据帧,预测所述第一个冗余子

数据对应的非关键语音特征,包括:

[0015] 基于所述第一个冗余子数据对应的基音周期参数和第一增益参数,确定所述第一个冗余子数据对应的第一激励参数;

[0016] 基于所述线谱对参数、所述第一个冗余子数据对应的第一激励参数、所述参考非关键语音特征以及所述参考语音数据帧,预测所述第一个冗余子数据对应的非关键语音特征。

[0017] 可选地,所述关键语音特征包括线谱对参数、基音周期参数和第一增益参数,所述非关键语音特征包括第二激励参数,所述调用特征预测模型,基于所述关键语音特征,预测所述非关键语音特征,包括:

[0018] 基于所述基音周期参数和所述第一增益参数,确定第一激励参数;

[0019] 调用所述特征预测模型,基于所述线谱对参数和所述第一激励参数,预测所述第二激励参数。

[0020] 可选地,所述第二语音数据包括多个语音数据帧,所述冗余数据中包括与所述多个语音数据帧对应的多个冗余子数据,所述调用特征预测模型,基于所述关键语音特征,预测所述非关键语音特征,包括:

[0021] 调用所述特征预测模型,基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征;

[0022] 所述对所述关键语音特征和所述非关键语音特征进行合成滤波,得到所述第二语音数据,包括:

[0023] 对所述每个冗余子数据对应的关键语音特征和非关键语音特征进行合成滤波,得到所述每个冗余子数据对应的语音数据帧;

[0024] 按照所述多个冗余子数据的排列顺序,对得到的多个语音数据帧进行合并,得到所述第二语音数据。

[0025] 可选地,所述调用所述特征预测模型,基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征,包括:

[0026] 对于第一个冗余子数据之后的每个冗余子数据,获取所述冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧;

[0027] 调用所述特征预测模型,基于所述冗余子数据对应的关键语音特征、所述上一个冗余子数据对应的非关键语音特征以及所述上一个冗余子数据对应的语音数据帧,预测所述冗余子数据对应的非关键语音特征。

[0028] 可选地,所述特征预测模型中包括参考非关键语音特征和参考语音数据帧,所述调用所述特征预测模型,基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征,包括:

[0029] 对于第一个冗余子数据,调用所述特征预测模型,基于所述第一个冗余子数据对应的关键语音特征、所述参考非关键语音特征以及所述参考语音数据帧,预测所述第一个冗余子数据对应的非关键语音特征。

[0030] 可选地,所述样本关键语音特征中包括样本线谱对参数、样本基音周期参数和第一样本增益参数,所述预测非关键语音特征包括第二样本激励参数,所述调用所述特征预测模型,基于所述样本关键语音特征进行预测,得到预测非关键语音特征,包括:

[0031] 基于所述样本基音周期参数和所述第一样本增益参数,确定第一样本激励参数,所述第一样本激励参数用于反映所述样本语音数据中具有相关性的语音特征;

[0032] 调用所述特征预测模型,基于所述样本线谱对参数和所述第一样本激励参数,预测所述第二样本激励参数,所述第二样本激励参数用于反映所述样本语音数据中不具有相关性的语音特征。

[0033] 可选地,所述样本语音数据包括多个样本语音数据帧,所述调用所述特征预测模型,基于所述样本关键语音特征进行预测,得到预测非关键语音特征,包括:

[0034] 调用所述特征预测模型,基于所述每个样本语音数据帧对应的样本关键语音特征,预测所述每个样本语音数据帧对应的预测非关键语音特征;

[0035] 所述样本解码数据中包括所述每个样本语音数据帧对应的样本解码数据帧,所述对所述样本关键语音特征和所述预测非关键语音特征进行合成滤波,得到样本解码数据,包括:

[0036] 对所述每个样本语音数据帧对应的样本关键语音特征和预测非关键语音特征进行合成滤波,得到所述每个样本语音数据帧对应的样本解码数据帧。

[0037] 可选地,所述调用所述特征预测模型,基于所述每个样本语音数据帧对应的样本关键语音特征,预测所述每个样本语音数据帧对应的预测非关键语音特征,包括:

[0038] 对于第一个样本语音数据帧之后的每个样本语音数据帧,获取所述样本语音数据帧的上一个样本语音数据帧对应的预测非关键语音特征和样本解码数据帧;

[0039] 调用所述特征预测模型,基于所述样本语音数据帧对应的样本关键语音特征、所述上一个样本语音数据帧对应的预测非关键语音特征和样本解码数据帧,预测所述样本语音数据帧对应的预测非关键语音特征。

[0040] 可选地,所述特征预测模型中包括参考非关键语音特征和参考语音数据帧,所述调用所述特征预测模型,基于所述每个样本语音数据帧对应的样本关键语音特征,预测所述每个样本语音数据帧对应的预测非关键语音特征,包括:

[0041] 对于第一个样本语音数据帧,调用所述特征预测模型,基于所述样本语音数据帧对应的样本关键语音特征、所述参考非关键语音特征以及所述参考语音数据帧,预测所述第一个样本语音数据帧对应的预测非关键语音特征。

[0042] 另一方面,提供了一种语音数据处理方法,所述方法包括:

[0043] 对第二语音数据进行特征提取,得到所述第二语音数据的多个语音特征,所述多个语音特征中包括关键语音特征和非关键语音特征;

[0044] 对所述多个语音特征进行特征编码,得到第一语音数据;

[0045] 对所述关键语音特征进行特征编码,得到冗余数据;

[0046] 对所述第一语音数据和所述冗余数据进行组包,得到目标数据包;

[0047] 向接收设备发送所述目标数据包,所述接收设备用于在所述目标数据包中未获取到所述第一语音数据的情况下,对所述冗余数据进行特征解码,得到所述关键语音特征,基于所述关键语音特征,预测所述非关键语音特征,对所述关键语音特征和所述非关键语音特征进行合成滤波,得到所述第二语音数据。

[0048] 另一方面,提供了一种语音数据处理装置,所述装置包括:

[0049] 数据包接收模块,用于接收目标数据包,所述目标数据包是发送设备对第一语音

数据和冗余数据进行组包后发送的,所述第一语音数据是对第二语音数据的多个语音特征进行特征编码得到的,所述多个语音特征中包括关键语音特征和非关键语音特征,所述冗余数据是所述关键语音特征进行特征编码得到的;

[0050] 特征解码模块,用于在所述目标数据包中未获取到所述第一语音数据的情况下,对所述冗余数据进行特征解码,得到所述关键语音特征;

[0051] 特征预测模块,用于基于所述关键语音特征,预测所述非关键语音特征;

[0052] 合成滤波模块,用于对所述关键语音特征和所述非关键语音特征进行合成滤波,得到所述第二语音数据。

[0053] 可选地,所述关键语音特征包括线谱对参数、基音周期参数和第一增益参数,所述非关键语音特征包括第二激励参数,所述特征预测模块,包括:

[0054] 第一激励确定单元,用于基于所述基音周期参数和所述第一增益参数,确定第一激励参数,所述第一激励参数用于反映所述第二语音数据中具有相关性的语音特征;

[0055] 第二激励确定单元,用于基于所述线谱对参数和所述第一激励参数,预测所述第二激励参数,所述第二激励参数用于反映所述第二语音数据中不具有相关性的语音特征。

[0056] 可选地,所述合成滤波模块,包括:

[0057] 合成滤波单元,用于基于所述线谱对参数,对所述第一激励参数和所述第二激励参数进行合成滤波,得到所述第二语音数据。

[0058] 可选地,所述第一激励确定单元,用于:

[0059] 在码本集合中查询与所述基音周期参数匹配的第一码本参数,所述码本集合中包括至少一组码本参数;

[0060] 将所述第一码本参数与所述第一增益参数的乘积,确定为所述第一激励参数。

[0061] 可选地,所述第二语音数据包括多个语音数据帧,所述冗余数据中包括与所述多个语音数据帧对应的多个冗余子数据,所述特征预测模块,包括:

[0062] 特征预测单元,用于基于所述每个冗余子数据对应的关键语音特征,预测所述每个冗余子数据对应的非关键语音特征;

[0063] 所述合成滤波模块,包括:

[0064] 数据帧获取单元,用于对所述每个冗余子数据对应的关键语音特征和非关键语音特征进行合成滤波,得到所述每个冗余子数据对应的语音数据帧;

[0065] 数据帧合并单元,用于按照所述多个冗余子数据的排列顺序,对得到的多个语音数据帧进行合并,得到所述第二语音数据。

[0066] 可选地,所述特征预测单元,用于:

[0067] 对于第一个冗余子数据之后的每个冗余子数据,获取所述冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧;

[0068] 基于所述冗余子数据对应的关键语音特征、所述上一个冗余子数据对应的非关键语音特征以及所述上一个冗余子数据对应的语音数据帧,预测所述冗余子数据对应的非关键语音特征。

[0069] 可选地,所述关键语音特征包括所述第二语音数据的线谱对参数以及所述每个冗余子数据对应的基音周期参数和第一增益参数,所述特征预测单元,用于:

[0070] 基于所述冗余子数据对应的基音周期参数和第一增益参数,确定所述冗余子数据

对应的第一激励参数；

[0071] 基于所述线谱对参数、所述冗余子数据对应的第一激励参数、所述上一个冗余子数据对应的非关键语音特征以及所述上一个冗余子数据对应的语音数据帧，预测所述冗余子数据对应的非关键语音特征。

[0072] 可选地，所述特征预测单元，用于：

[0073] 对于第一个冗余子数据，获取所述第一个冗余子数据对应的参考非关键语音特征和参考语音数据帧；

[0074] 基于所述第一个冗余子数据对应的关键语音特征、所述参考非关键语音特征以及所述参考语音数据帧，预测所述第一个冗余子数据对应的非关键语音特征。

[0075] 可选地，所述关键语音特征包括所述第二语音数据的线谱对参数以及所述每个冗余子数据对应的基音周期参数和第一增益参数，所述特征预测单元，用于：

[0076] 基于所述第一个冗余子数据对应的基音周期参数和第一增益参数，确定所述第一个冗余子数据对应的第一激励参数；

[0077] 基于所述线谱对参数、所述第一个冗余子数据对应的第一激励参数、所述参考非关键语音特征以及所述参考语音数据帧，预测所述第一个冗余子数据对应的非关键语音特征。

[0078] 可选地，所述特征预测模块，包括：

[0079] 模型调用单元，用于调用特征预测模型，基于所述关键语音特征，预测所述非关键语音特征。

[0080] 可选地，所述关键语音特征包括线谱对参数、基音周期参数和第一增益参数，所述非关键语音特征包括第二激励参数，所述模型调用单元，用于：

[0081] 基于所述基音周期参数和所述第一增益参数，确定第一激励参数；

[0082] 调用所述特征预测模型，基于所述线谱对参数和所述第一激励参数，预测所述第二激励参数。

[0083] 可选地，所述第二语音数据包括多个语音数据帧，所述冗余数据中包括与所述多个语音数据帧对应的多个冗余子数据，所述模型调用单元，用于：

[0084] 调用所述特征预测模型，基于所述每个冗余子数据对应的关键语音特征，预测所述每个冗余子数据对应的非关键语音特征；

[0085] 所述合成滤波模块，包括：

[0086] 数据帧获取单元，用于对所述每个冗余子数据对应的关键语音特征和非关键语音特征进行合成滤波，得到所述每个冗余子数据对应的语音数据帧；

[0087] 数据帧合并单元，用于按照所述多个冗余子数据的排列顺序，对得到的多个语音数据帧进行合并，得到所述第二语音数据。

[0088] 可选地，所述模型调用单元，用于：

[0089] 对于第一个冗余子数据之后的每个冗余子数据，获取所述冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧；

[0090] 调用所述特征预测模型，基于所述冗余子数据对应的关键语音特征、所述上一个冗余子数据对应的非关键语音特征以及所述上一个冗余子数据对应的语音数据帧，预测所述冗余子数据对应的非关键语音特征。

[0091] 可选地,所述特征预测模型中包括参考非关键语音特征和参考语音数据帧,所述模型调用单元,用于:

[0092] 对于第一个冗余子数据,调用所述特征预测模型,基于所述第一个冗余子数据对应的关键语音特征、所述参考非关键语音特征以及所述参考语音数据帧,预测所述第一个冗余子数据对应的非关键语音特征。

[0093] 可选地,所述装置还包括:

[0094] 特征提取模块,用于对样本语音数据进行特征提取,得到所述样本语音数据对应的多个样本语音特征,所述多个样本语音特征中包括样本关键语音特征;

[0095] 所述特征预测模块,还用于调用所述特征预测模型,基于所述样本关键语音特征进行预测,得到预测非关键语音特征;

[0096] 所述合成滤波模块,还用于对所述样本关键语音特征和所述预测非关键语音特征进行合成滤波,得到样本解码数据;

[0097] 模型训练模块,用于基于所述样本语音数据、所述样本解码数据、所述样本关键语音特征以及所述预测非关键语音特征,训练所述特征预测模型。

[0098] 可选地,所述模型训练模块,包括:

[0099] 第一特征确定单元,用于将所述样本语音数据与所述样本解码数据之间的误差确定为残差特征;

[0100] 第二特征确定单元,用于将所述残差特征与所述样本关键语音特征之间的差值确定为样本非关键语音特征;

[0101] 模型训练单元,用于基于所述预测非关键语音特征与所述样本非关键语音特征之间的误差,训练所述特征预测模型。

[0102] 可选地,所述样本关键语音特征中包括样本线谱对参数、样本基音周期参数和第一样本增益参数,所述特征预测模块,包括:

[0103] 第一激励确定单元,用于基于所述样本基音周期参数和所述第一样本增益参数,确定第一样本激励参数,所述第一样本激励参数用于反映所述样本语音数据中具有相关性的语音特征;

[0104] 第二激励确定单元,用于调用所述特征预测模型,基于所述样本线谱对参数和所述第一样本激励参数,预测所述第二样本激励参数,所述第二样本激励参数用于反映所述样本语音数据中不具有相关性的语音特征。

[0105] 可选地,所述样本语音数据包括多个样本语音数据帧,所述特征预测模块,包括:

[0106] 特征预测单元,用于调用所述特征预测模型,基于所述每个样本语音数据帧对应的样本关键语音特征,预测所述每个样本语音数据帧对应的预测非关键语音特征;

[0107] 所述样本解码数据中包括所述每个样本语音数据帧对应的样本解码数据帧,所述合成滤波模块,包括:

[0108] 数据帧获取单元,用于对所述每个样本语音数据帧对应的样本关键语音特征和预测非关键语音特征进行合成滤波,得到所述每个样本语音数据帧对应的样本解码数据帧。

[0109] 可选地,所述特征预测单元,用于:

[0110] 对于第一个样本语音数据帧之后的每个样本语音数据帧,获取所述样本语音数据帧的上一个样本语音数据帧对应的预测非关键语音特征和样本解码数据帧;

[0111] 调用所述特征预测模型,基于所述样本语音数据帧对应的样本关键语音特征、所述上一个样本语音数据帧对应的预测非关键语音特征和样本解码数据帧,预测所述样本语音数据帧对应的预测非关键语音特征。

[0112] 可选地,所述特征预测模型中包括参考非关键语音特征和参考语音数据帧,所述特征预测单元,用于:

[0113] 对于第一个样本语音数据帧,调用所述特征预测模型,基于所述样本语音数据帧对应的样本关键语音特征、所述参考非关键语音特征以及所述参考语音数据帧,预测所述第一个样本语音数据帧对应的预测非关键语音特征。

[0114] 另一方面,提供了一种语音数据处理装置,所述装置包括:

[0115] 特征提取模块,用于对第二语音数据进行特征提取,得到所述第二语音数据的多个语音特征,所述多个语音特征中包括关键语音特征和非关键语音特征;

[0116] 第一特征编码模块,用于对所述多个语音特征进行特征编码,得到第一语音数据;

[0117] 第二特征编码模块,用于对所述关键语音特征进行特征编码,得到冗余数据;

[0118] 组包模块,用于对所述第一语音数据和所述冗余数据进行组包,得到目标数据包;

[0119] 数据包发送模块,用于向接收设备发送所述目标数据包,所述接收设备用于在所述目标数据包中未获取到所述第一语音数据的情况下,对所述冗余数据进行特征解码,得到所述关键语音特征,基于所述关键语音特征,预测所述非关键语音特征,对所述关键语音特征和所述非关键语音特征进行合成滤波,得到所述第二语音数据。

[0120] 可选地,所述关键语音特征包括线谱对参数、基音周期参数和第一增益参数,所述第二特征编码模块,包括:

[0121] 特征编码单元,用于对所述线谱对参数、所述基音周期参数和所述第一增益参数进行特征编码,得到所述冗余数据。

[0122] 另一方面,提供了一种计算机设备,所述计算机设备包括处理器和存储器,所述存储器中存储有至少一条计算机程序,所述至少一条计算机程序由所述处理器加载并执行以实现如上述方面所述的语音数据处理方法中所执行的操作。

[0123] 另一方面,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中存储有至少一条计算机程序,所述至少一条计算机程序由处理器加载并执行以实现如上述方面所述的语音数据处理方法中所执行的操作。

[0124] 另一方面,提供了一种计算机程序产品或计算机程序,所述计算机程序产品或计算机程序包括计算机程序代码,所述计算机程序代码存储在计算机可读存储介质中,计算机设备的处理器从计算机可读存储介质读取所述计算机程序代码,处理器执行所述计算机程序代码,使得所述计算机设备实现如上述方面所述的语音数据处理方法中所执行的操作。

[0125] 本申请实施例提供的方法、装置、计算机设备及存储介质,在未获取到冗余数据对应的第一语音数据的情况下,对冗余数据进行特征解码得到关键语音特征,并基于关键语音特征预测出非关键语音特征,从而对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据,因此在丢失第一语音数据的情况下,依然能够恢复出第二语音数据,达到了数据的抗丢失效果。并且,由于冗余数据中仅需携带关键语音特征,无需携带非关键语音特征,从而降低了冗余数据的数据量,因此能够提高传输目标数据包过程中的带宽利用率。

附图说明

[0126] 为了更清楚地说明本申请实施例中的技术方案,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,下面描述中的附图仅仅是本申请实施例的一些实施例,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。

[0127] 图1是本申请实施例提供的一种实施环境的示意图;

[0128] 图2是本申请实施例提供的一种语音数据处理方法的流程图;

[0129] 图3是本申请实施例提供的一种语音数据处理方法的流程图;

[0130] 图4是本申请实施例提供的一种语音数据处理方法的流程图;

[0131] 图5是本申请实施例提供的一种编码方法的示意图;

[0132] 图6是本申请实施例提供的一种解码方法的示意图;

[0133] 图7是本申请实施例提供的一种预测非关键语音特征的流程图;

[0134] 图8是本申请实施例提供的一种特征预测模型的结构示意图;

[0135] 图9是本申请实施例提供的一种解码方法的示意图;

[0136] 图10是本申请实施例提供的一种模型训练方法的流程图;

[0137] 图11是本申请实施例提供的一种模型训练方法的示意图;

[0138] 图12是本申请实施例提供的一种语音数据处理装置的结构示意图;

[0139] 图13是本申请实施例提供的另一种语音数据处理装置的结构示意图;

[0140] 图14是本申请实施例提供的一种语音数据处理装置的结构示意图;

[0141] 图15是本申请实施例提供的另一种语音数据处理装置的结构示意图;

[0142] 图16是本申请实施例提供的一种终端的结构示意图;

[0143] 图17是本申请实施例提供的一种服务器的结构示意图。

具体实施方式

[0144] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合附图对本申请实施方式作进一步地详细描述。

[0145] 可以理解,本申请所使用的术语“第一”、“第二”等可在本文中用于描述各种概念,但除非特别说明,这些概念不受这些术语限制。这些术语仅用于将一个概念与另一个概念区分。举例来说,在不脱离本申请的范围的情况下,可以将第一语音数据称为第二语音数据,且类似地,可将第二语音数据称为第一语音数据。

[0146] 其中,至少一个是指一个或者一个以上,例如,至少一个语音数据帧可以是一个语音数据帧、两个语音数据帧、三个语音数据帧等任一大于等于一的整数个语音数据帧。多个是指两个或者两个以上,例如,多个语音数据帧可以是两个语音数据帧、三个语音数据帧等任一大于等于二的整数个语音数据帧。每个是指至少一个中的每一个,例如,每个语音数据帧是指多个语音数据帧中的每一个语音数据帧,若多个语音数据帧为3个语音数据帧,则每个语音数据帧是指3个语音数据帧中的每一个语音数据帧。

[0147] 人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说,人工智能是计算机科学的一个综合技术,它企图了解

智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。

[0148] 人工智能技术是一门综合学科,涉及领域广泛,既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术包括自然语言处理技术和机器学习。

[0149] 机器学习(Machine Learning,ML)是一门多领域交叉学科,涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能,重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心,是使计算机具有智能的根本途径,其应用遍及人工智能的各个领域。机器学习和深度学习包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习、示教学习等技术。

[0150] 语音技术(Speech Technology)关键技术包括自动语音识别技术(Automatic Speech Recognition)和语音合成技术(Text To Speech)以及声纹识别技术。让计算机设备能听、能看、能说、能感觉,是未来人机交互的发展方向,其中语音成为未来最被看好的人机交互方式之一。

[0151] 云技术(Cloud technology)是指在广域网或局域网内将硬件、软件、网络等系列资源统一起来,实现数据的计算、储存、处理和共享的一种托管技术。云技术是基于云计算商业模式应用的网络技术、信息技术、整合技术、管理平台技术、应用技术等的总称,可以组成资源池,按需所用,灵活便利。云计算技术将变成重要支撑。技术网络系统的后台服务需要大量的计算、存储资源,如视频网站、图片类网站和更多的门户网站。伴随着互联网行业的高度发展和应用,将来每个物品都有可能存在自己的识别标志,都需要传输到后台系统进行逻辑处理,不同程度级别的数据将会分开处理,各类行业数据皆需要强大的系统后盾支撑,只能通过云计算来实现。

[0152] 云会议是基于云计算技术的一种高效、便捷、低成本的会议形式。使用者只需要通过互联网界面,进行简单易用的操作,便可快速高效地与全球各地团队及客户同步分享语音、数据文件及视频,而会议中数据的传输、处理等复杂技术由云会议服务商帮助使用者进行操作。目前国内云会议主要集中在以SaaS (Software as a Service,软件即服务)模式为主体的服务内容,包括电话、网络、视频等服务形式,基于云计算的视频会议就叫云会议。在云会议时代,数据的传输、处理、存储全部由视频会议厂家的计算机资源处理,用户完全无需再购置昂贵的硬件和安装繁琐的软件,只需打开浏览器,登录相应界面,就能进行高效的远程会议。

[0153] 以下将基于人工智能技术和云技术,对本申请实施例提供的语音数据处理方法进行说明。

[0154] 图1是本申请实施例提供的一种实施环境的示意图,参见图1,该实施环境包括:第一设备101和第一设备102。可选地,第一设备101和第一设备102为终端,例如该终端是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表等,但并不局限于此。可选地,第一设备101和第一设备102为服务器,例如该服务器是独立的物理服务器,也可以是多个物理服务器构成的服务器集群或者分布式系统,还可以是提供云服务、云数据库、云计

算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN(Content Delivery Network,内容分发网络)、以及大数据和人工智能平台等基础云计算服务的云服务器。

[0155] 其中,第一设备101和第一设备102可以通过有线或无线通信方式进行直接或间接地连接,本申请在此不做限制。第一设备101和第一设备102能够通过本申请实施例提供的语音数据处理方法进行语音通信。例如,第一设备101为语音数据的发送方,第一设备102为语音数据的接收方,第一设备101将编码后的语音数据以及该语音数据对应的冗余数据传输给第一设备102,在编码后的语音数据被丢失的情况下,第一设备102即可通过对冗余数据进行解码,得到解码后的语音数据,从而实现第一设备101和第一设备102之间的语音通信。

[0156] 在一种可能实现方式中,第一设备101和第二设备102为区块链系统中的节点,区块链系统中的节点之间能够进行语音传输,每个节点将获取到的数据包、语音数据或者冗余数据等数据存储于区块链上。

[0157] 本申请实施例提供的语音数据处理方法可应用于传输语音数据的任一场景下。

[0158] 例如,在打电话的场景下,在第一终端与第二终端进行语音通话的过程中,第一终端对当前获取到的第二语音数据进行编码,得到第一语音数据,同时采用本申请实施例提供的方法,对该第二语音数据对应的关键语音特征进行编码,得到冗余数据,然后将该第一语音数据和冗余数据进行组包后传输给第二终端,如果在传输数据包的过程中,该第一语音数据被丢失,则第二终端根据冗余数据中的关键语音特征预测出其他的非关键语音特征,从而利用关键语音特征和非关键语音特征解码出第二语音数据,达到抗丢包的效果。

[0159] 图2是本申请实施例提供的一种语音数据处理方法的流程图。本申请实施例的执行主体为计算机设备,参见图2,该方法包括以下步骤。

[0160] 201、计算机设备接收目标数据包。

[0161] 计算机设备与其他设备进行语音通信的过程中,如果在语音数据传输过程中编码后的语音数据被丢失,则计算机设备无法接收到编码后的语音数据。为了防止由于丢包导致无法解码出语音数据的情况,在传输编码后的语音数据的过程中,还会传输该语音数据对应的冗余数据,该冗余数据又称为冗余数据,该冗余数据用于在编码后的语音数据丢失的情况下,恢复出语音数据。

[0162] 本申请实施例中,计算机设备接收目标数据包,目标数据包是发送设备对第一语音数据和冗余数据进行组包后发送的,第一语音数据是对第二语音数据的多个语音特征进行特征编码得到的,多个语音特征中包括关键语音特征和非关键语音特征,冗余数据是对关键语音特征进行特征编码得到的。其中,关键语音特征比非关键语音特征能够更精准地反映语音数据的特征。例如,关键语音特征是指具有相关性的语音特征,非关键语音特征是指不具有相关性的语音特征。或者,语音数据包括长时特征和短时特征,关键语音特征能够反映语音数据的长时特征,非关键语音特征能够反映语音数据的短时特征。除此之外,还可以采用其他方式定义关键语音特征和非关键语音特征,本申请实施例对此不做限定。

[0163] 202、计算机设备在目标数据包中未获取到第一语音数据的情况下,对冗余数据进行特征解码,得到关键语音特征。

[0164] 如果计算机设备在目标数据包中未获取到冗余数据对应的第一语音数据,说明该

第一语音数据被丢失,则计算机设备获取该目标数据包中的冗余数据,由于该冗余数据是对第二语音数据的关键语音特征进行特征编码得到的,因此计算机设备对该冗余数据进行特征解码,能够得到第二语音数据的关键语音特征。

[0165] 203、计算机设备基于关键语音特征,预测非关键语音特征。

[0166] 计算机设备获取到关键语音特征之后,基于该语音数据的关键语音特征,预测该语音数据的非关键语音特征,从而得到语音数据的关键语音特征和非关键语音特征。本申请实施例中,由于采用了基于关键语音特征预测非关键语音特征的方式,因此冗余数据中仅需携带关键语音特征,无需携带非关键语音特征,从而降低了冗余数据的数据量,因此在其他设备向计算机设备传输该冗余数据的过程中,降低了传输的数据量,能够提高带宽利用率。

[0167] 204、计算机设备对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据。

[0168] 其中,语音数据的关键语音特征和非关键语音特征是对第二语音数据进行特征提取得到的。因此,计算机设备获取到关键语音特征和非关键语音特征之后,对该关键语音特征和非关键语音特征进行合成滤波,能够恢复出第二语音数据。后续该计算机设备可以播放该第二语音数据,从而在语音数据发生丢失的情况下,依然实现了计算机设备与其他设备之间的语音通信,达到了数据的抗丢失效果。

[0169] 本申请实施例提供了一种基于语音数据的关键语音特征来预测非关键语音特征的方法,在未获取到冗余数据对应的第一语音数据的情况下,对冗余数据进行特征解码得到关键语音特征,并预测出非关键语音特征,从而对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据,因此在丢失第一语音数据的情况下,依然能够恢复出第二语音数据,达到了数据的抗丢失效果。

[0170] 并且,由于冗余数据中仅需携带关键语音特征,无需携带非关键语音特征,从而降低了冗余数据的数据量,因此能够提高传输目标数据包过程中的带宽利用率。

[0171] 图3是本申请实施例提供的一种语音数据处理方法的流程图。本申请实施例的执行主体为计算机设备,参见图3,该方法包括以下步骤。

[0172] 301、计算机设备对第二语音数据进行特征提取,得到第二语音数据的多个语音特征。

[0173] 计算机设备与其他设备进行语音通信的过程中,获取待传输的第二语音数据,并对该第二语音数据进行特征提取,得到该第二语音数据的多个语音特征。其中,对该第二语音数据进行特征提取的过程,可以看作是对该第二语音数据进行语音编码的过程,该多个语音特征中包括关键语音特征和非关键语音特征。

[0174] 关键语音特征比非关键语音特征能够更精准地反映语音数据的特征。例如,关键语音特征是指具有相关性的语音特征,非关键语音特征是指不具有相关性的语音特征。或者,语音数据包括长时特征和短时特征,关键语音特征能够反映语音数据的长时特征,非关键语音特征能够反映语音数据的短时特征。除此之外,还可以采用其他方式定义关键语音特征和非关键语音特征,本申请实施例对此不做限定。

[0175] 302、计算机设备对多个语音特征进行特征编码,得到第一语音数据。

[0176] 计算机设备对该多个语音特征进行特征编码,得到第一语音数据。对多个语音特

征进行特征编码是指将多个语音特征转换成能够在网络中进行传输的格式。

[0177] 其中,对第一语音数据进行特征解码能够得到多个语音特征,然后对该多个语音特征进行合成滤波,能够得到该第二语音数据,从而基于第一语音数据恢复出第二语音数据。

[0178] 303、计算机设备对关键语音特征进行特征编码,得到冗余数据。

[0179] 计算机设备对关键语音特征进行特征编码,得到冗余数据,该冗余数据用于在第一语音数据丢失的情况下,恢复出第二语音数据。

[0180] 304、计算机设备对第一语音数据和冗余数据进行组包,得到目标数据包。

[0181] 计算机设备得到第一语音数据和冗余数据后,对该第一语音数据和冗余数据进行组包,得到目标数据包,该目标数据包中包括该第一语音数据和冗余数据。

[0182] 305、计算机设备向接收设备发送目标数据包。

[0183] 计算机设备获取到目标数据包之后,向接收设备发送该目标数据包,该接收设备是指与该计算机设备进行语音通信的设备。

[0184] 该接收设备用于在目标数据包中获取到第一语音数据的情况下,直接对第一语音数据进行解码,得到第二语音数据。或者,该接收设备还用于在目标数据包中未获取到第一语音数据的情况下,对冗余数据进行特征解码,得到关键语音特征,基于关键语音特征,预测非关键语音特征,从而对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据。因此在丢失第一语音数据的情况下,依然实现了语音数据的传输,完成了计算机设备与接收设备之间的语音通信,达到了数据的抗丢失效果。

[0185] 本申请实施例提供的方法,传输对第二语音数据的多个语音特征进行编码得到的第一语音数据时,还会传输对该第二语音数据的关键语音特征进行编码得到的冗余数据,因此在接收设备未获取到第一语音数据时,基于冗余数据中的关键语音特征预测非关键语音特征,从而对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据,因此在丢失第一语音数据的情况下,依然能够恢复出第二语音数据,达到了数据的抗丢失效果。

[0186] 并且,由于接收设备能够基于关键语音特征预测非关键语音特征,因此冗余数据中仅需携带关键语音特征,无需携带非关键语音特征,从而降低了冗余数据的数据量,因此能够提高传输目标数据包过程中的带宽利用率。

[0187] 语音编解码技术在现代通讯系统中占有重要的地位,在语音通信的场景中,语音信号经由麦克风采集得到,通过模数转换电路,将模拟语音信号转换为数字语音信号,数字语音信号经过语音编码器进行编码,然后按照通信网络传输格式和协议打包发送到接收设备,接收设备接收到数据包后,解包输出语音编码码流,通过语音解码器后重新生成语音数字信号,最后语音数字信号通过扬声器播放出语音。其中,语音编解码技术有效地降低了语音信号传输的带宽,对于节省语音数据的存储成本以及传输成本具有重要意义,对保障通信网络传输过程中的语音数据的完整性起了决定性作用。但是在传输语音数据的过程中,可能会出现丢失语音数据的情况,影响语音数据传输的效果。因此,本申请实施例提供了一种能够达到抗丢失效果的语音数据处理方法,图4是本申请实施例提供的一种语音数据处理方法的流程图,本申请实施例的执行主体为发送设备和接收设备,参见图4,该方法包括以下步骤。

[0188] 401、发送设备对第二语音数据进行特征提取,得到第二语音数据的多个语音特

征。

[0189] 发送设备与接收设备进行语音通信的过程中,获取待传输的第二语音数据,例如,该第二语音数据为发送设备通过麦克风采集的语音数据,或者为发送设备中已存储的语音数据等。发送设备对该第二语音数据进行特征提取,得到该第二语音数据的多个语音特征。该多个语音特征中包括关键语音特征和非关键语音特征。其中,对该第二语音数据进行特征提取的过程,可以看作是对该第二语音数据进行语音编码的过程。

[0190] 关键语音特征比非关键语音特征能够更精准地反映第二语音数据的特征,基于关键语音特征能够预测出非关键语音特征。可选地,第二语音数据有多个维度的语音特征,发送设备在多个维度的语音特征中,选取至少一个目标维度的语音特征作为关键语音特征,并将其他维度的语音特征作为非关键语音特征。例如,目标维度的语音特征是指具有相关性的语音特征,发送设备将具有相关性的语音特征作为关键语音特征,将不具有相关性的语音特征作为非关键语音特征。再例如,第二语音数据的多个语音特征中包括长时特征和短时特征,目标维度的语音特征是指第二语音数据的长时特征,发送设备将第二语音数据的长时特征作为关键语音特征,将第二语音数据的短时特征作为非关键语音特征。除此之外,目标维度的语音特征还可以是其他类型的语音特征,本申请实施例对此不做限定。

[0191] 以发送设备采用CELP(Code Excited Linear Prediction,码激励线性预测编码)方法进行编码为例,则该多个语音特征中包括线谱对参数、基音周期参数、第一增益参数、第二码本参数和第二增益参数。其中,该线谱对参数用于反映第二语音数据的音谱特征,该基音周期参数用于反映第二语音数据中的声带振动频率,该第一增益参数用于调节第二语音数据中具有相关性的信号的强度,该第一增益参数又称为自适应增益参数。该第二码本参数用于反映第二语音数据中不具有相关性的信号,或者该第二码本参数用于反映第二语音数据中的短时特征,该第二码本参数又称为固定码本参数,或者还可以称为固定码书参数。该第二增益参数用于调节第二语音数据中不具有相关性的信号的强度,或者说该第二增益参数用于调节第二码本参数所反映的信号强度,该第二增益参数又称为固定增益参数。其中,该第二码本参数和第二增益参数的乘积为第二激励参数,该第二激励参数是指第二语音数据中不具有相关性的语音特征。

[0192] 或者,发送设备采用其他的编码方式对第二语音数据进行特征提取,得到其他类型的多个语音特征,本申请实施例对此不做限定。

[0193] 402、发送设备对多个语音特征进行特征编码,得到第一语音数据。

[0194] 发送设备对该多个语音特征进行特征编码,得到第一语音数据。其中,对多个语音特征进行特征编码是指将多个语音特征转换成能够在网络中进行传输的格式。

[0195] 403、发送设备对关键语音特征进行特征编码,得到冗余数据。

[0196] 发送设备对关键语音特征进行特征编码,得到冗余数据,该冗余数据用于在第一语音数据丢失的情况下,恢复出第二语音数据。

[0197] 在一种可能实现方式中,关键语音特征包括线谱对参数、基音周期参数和第一增益参数。发送设备对线谱对参数、基音周期参数和第一增益参数进行特征编码,得到冗余数据。可选地,该多个语音特征中,线谱对参数、基音周期参数、第一增益参数为关键语音特征,第二码本参数和第二增益参数为非关键语音特征,而第二码本参数和第二增益参数占用了编码数据中40%以上的带宽资源。因此为了节省带宽资源,发送设备仅对该线谱对参

数、基音周期参数和第一增益参数进行特征编码,得到冗余数据。

[0198] 404、发送设备对第一语音数据和冗余数据进行组包,得到目标数据包,向接收设备发送该目标数据包。

[0199] 发送设备获取到第一语音数据与冗余数据之后,对该第一语音数据和冗余数据进行组包,得到目标数据包,该目标数据包中包括该第一语音数据和冗余数据。发送设备向接收设备发送该目标数据包。如果在数据传输过程中第一语音数据被丢失,则接收设备无法获取到第一语音数据,因此为了防止由于第一语音数据丢失导致无法解码出第二语音数据的情况,发送设备会将第一语音数据和冗余数据放在同一数据包中进行传输。后续如果第一语音数据丢失,则还能够根据冗余数据还原出第二语音数据。

[0200] 可选地,发送设备采用带内FEC(Forward Error Correction,前向纠错)编码的方式,生成包括第一语音数据和冗余数据的目标数据包。带内FEC编码是指将语音数据以及用于进行纠错的冗余数据放在相同的数据包中。

[0201] 图5是本申请实施例提供的一种编码方法的示意图,参见图5,发送设备对第二语音数据进行特征提取,得到多个语音特征,对该多个语音特征进行特征编码,得到第一语音数据,同时发送设备在多个语音特征中提取关键语音特征,对该关键语音特征进行特征编码,得到冗余数据,从而对第一语音数据和冗余数据进行组包,得到目标数据包,传输该目标数据包。

[0202] 405、接收设备接收目标数据包。

[0203] 发送设备向接收设备发送目标数据包之后,该接收设备接收目标数据包,该目标数据包中包括第一语音数据和冗余数据。

[0204] 406、接收设备在未获取到冗余数据对应的第一语音数据的情况下,对冗余数据进行特征解码,得到关键语音特征。

[0205] 如果接收设备未获取到目标数据包中的第一语音数据,说明该第一语音数据被丢失,由于该冗余数据是对第二语音数据的关键语音特征进行特征编码得到的,因此接收设备能够对该冗余数据进行特征解码,能够得到第二语音数据的关键语音特征。可选地,特征解码包括格式转换等操作,例如该冗余数据是对关键语音特征进行格式转换得到的,则接收设备需要对冗余数据重新进行格式转换从而得到该关键语音特征。

[0206] 需要说明的是,本申请实施例仅以接收设备未获取到第一语音数据为例进行说明,在另一实施例中,接收设备在获取到第一语音数据的情况下,直接对该第一语音数据进行解码,即可得到第二语音数据,从而播放该第二语音数据,完成发送设备与接收设备之间的语音通信。

[0207] 407、接收设备基于关键语音特征,预测非关键语音特征。

[0208] 接收设备获取到第二语音数据的关键语音特征之后,基于该关键语音特征,预测第二语音数据的非关键语音特征,从而得到第二语音数据的关键语音特征和非关键语音特征。本申请实施例中,由于采用了基于关键语音特征预测非关键语音特征的方式,因此冗余数据中仅需携带关键语音特征,无需携带非关键语音特征,从而降低了冗余数据的数据量,因此在发送设备向接收设备传输该冗余数据的过程中,降低了传输的数据量,能够提高带宽利用率。可选地,接收设备采用深度学习的方式,来实现基于关键语音特征,预测非关键语音特征。

[0209] 在一种可能实现方式中,关键语音特征包括线谱对参数、基音周期参数和第一增益参数,非关键语音特征包括第二激励参数。接收设备基于基音周期参数和第一增益参数,确定第一激励参数,基于线谱对参数和第一激励参数,预测第二激励参数。其中,第一激励参数用于反映语音数据中具有相关性的语音特征;第二激励参数用于反映语音数据中不具有相关性的语音特征。该第一激励参数还用于表示语音数据中的长时特征,该第二激励参数还用于表示语音数据中的短时特征。该第一激励参数又称为自适应激励参数,该第二激励参数又称为固定激励参数。

[0210] 可选地,接收设备在码本集合中查询与基音周期参数匹配的第一码本参数,将第一码本参数与第一增益参数的乘积,确定为第一激励参数。该第一码本参数又称为自适应码本参数,其中,码本集合中包括至少一组码本参数,该至少一组码本参数是对历史语音数据进行处理得到的历史码本参数。

[0211] 在另一种可能实现方式中,语音数据包括多个语音数据帧,冗余数据中包括与多个语音数据帧对应的多个冗余子数据,接收设备基于每个冗余子数据对应的关键语音特征,预测每个冗余子数据对应的非关键语音特征。

[0212] 其中,冗余数据中包括多个冗余子数据,接收设备对冗余数据进行特征解码,得到每个冗余子数据对应的关键语音特征,则对于每个冗余子数据,接收设备基于该冗余子数据对应的关键语音特征,预测该冗余子数据对应的非关键语音特征,从而得到每个冗余子数据对应的非关键语音特征。其中,对于多个冗余子数据,接收设备采用以下第一种方式和第二种方式预测冗余子数据对应的非关键语音特征。

[0213] 第一种方式:对于第一个冗余子数据之后的每个冗余子数据,接收设备获取冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧,基于该冗余子数据对应的关键语音特征、上一个冗余子数据对应的非关键语音特征以及上一个冗余子数据对应的语音数据帧,预测冗余子数据对应的非关键语音特征。

[0214] 其中,上一个冗余子数据对应的语音数据帧的获取方法详见下述步骤408,在此暂不做说明。对于当前正在处理的冗余子数据,接收设备获取该冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧,从而基于当前的冗余子数据对应的关键语音特征以及上一个冗余子数据对应的非关键语音特征和语音数据帧进行预测,得到当前的冗余子数据对应的非关键语音特征,直至接收设备获取到多个冗余子数据中的最后一个冗余子数据对应的非关键语音特征。也即是,对于第一个冗余子数据之后的每个冗余子数据,在预测该冗余子数据对应的非关键语音特征的过程中,需要参考该冗余子数据的上一个冗余子数据对应的预测结果。

[0215] 可选地,关键语音特征包括语音数据的线谱对参数,以及每个冗余子数据对应的基音周期参数和第一增益参数。则对于第一个冗余子数据之后的每个冗余子数据,接收设备基于该冗余子数据对应的基音周期参数和第一增益参数,确定冗余子数据对应的第一激励参数,然后基于线谱对参数、冗余子数据对应的第一激励参数、上一个冗余子数据对应的非关键语音特征以及上一个冗余子数据对应的语音数据帧,预测冗余子数据对应的非关键语音特征。也即是,多个冗余子数据对应一个相同的线谱对参数,该线谱对参数反映的是整个语音数据的特征,在预测每个冗余子数据对应的非关键语音特征时,均会参考该线谱对参数。每个冗余子数据对应各自的基音周期参数和第一增益参数,某一冗余子数据对应

的基音周期参数和第一增益参数用于反映该冗余子数据对应的语音数据帧的特征。

[0216] 第二种方式:对于第一个冗余子数据,接收设备获取第一个冗余子数据对应的参考非关键语音特征和参考语音数据帧,基于第一个冗余子数据对应的关键语音特征、参考非关键语音特征以及参考语音数据帧,预测第一个冗余子数据对应的非关键语音特征。由于第一个冗余子数据没有上一个冗余子数据,因此接收设备获取对应的参考非关键语音特征和参考语音数据帧作为参考。可选地,该参考非关键语音特征和参考语音数据帧为预先设置的,例如将非关键语音特征和参考语音数据帧确定为0,或者将非关键语音特征和参考语音数据帧确定为大于0且小于1的数值等,本申请实施例对此不做限定。

[0217] 可选地,关键语音特征包括语音数据的线谱对参数以及每个冗余子数据对应的基音周期参数和第一增益参数。则接收设备基于第一个冗余子数据对应的基音周期参数和第一增益参数,确定第一个冗余子数据对应的第一激励参数,基于线谱对参数、第一个冗余子数据对应的第一激励参数、参考非关键语音特征以及参考语音数据帧,预测第一个冗余子数据对应的非关键语音特征。

[0218] 例如,以一帧语音数据包括20ms为例,该20ms的语音数据包括多个语音数据帧,每个语音数据帧为2ms,则该语音数据对应的冗余数据中包括多个2ms的冗余子数据,接收设备依次对该多个2ms的冗余子数据进行处理。

[0219] 408、接收设备对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据。

[0220] 其中,第二语音数据的关键语音特征和非关键语音特征是对第二语音数据进行特征提取得到的。因此,接收设备获取到关键语音特征和非关键语音特征之后,对该关键语音特征和非关键语音特征进行合成滤波,即可得到第二语音数据。其中,对第二语音数据进行特征提取得到第二语音数据的关键语音特征和非关键语音特征的过程,也即是对第二语音数据进行语音编码的过程,对关键语音特征和非关键语音特征进行合成滤波得到第二语音数据的过程,也即是对关键语音特征和非关键语音特征进行语音解码的过程。后续该接收设备可以播放该第二语音数据,从而在丢失第一语音数据的情况下,依然能够恢复出第二语音数据,实现了发送设备与接收设备之间的语音通信,达到了数据的抗丢失效果。

[0221] 本申请实施例提供了一种结合深度学习的抗丢包方法,在传输冗余数据时,仅需传输部分关键语音特征的冗余编码信息,当出现丢包时,利用冗余数据中的这些关键语音特征,来预测其他非关键语音特征,从而基于关键语音特征和非关键语音特征恢复丢包帧对应的数据,由于降低了冗余数据的数据量,因此在实现抗丢失效果的同时,还能提升网络带宽资源的有效利用率。

[0222] 在一种可能实现方式中,第二语音数据包括多个语音数据帧,冗余数据中包括与多个语音数据帧对应的多个冗余子数据。在上述步骤407中,接收设备对每个冗余子数据对应的关键语音特征进行处理,得到了每个冗余子数据对应的非关键语音数据。则接收设备对每个冗余子数据对应的关键语音特征和非关键语音特征进行合成滤波,得到每个冗余子数据对应的语音数据帧,然后按照多个冗余子数据的排列顺序,对得到的多个语音数据帧进行合并,得到第二语音数据。

[0223] 由于每个冗余子数据对应语音数据帧,因此多个冗余子数据的排列顺序即为多个语音数据帧的排列顺序,因此按照多个冗余子数据的排列顺序,对多个语音数据帧进行合

并,即可得到第二语音数据。或者,由于接收设备是按照多个冗余子数据的排列顺序依次获取每个冗余子数据对应的语音数据帧,因此接收设备按照多个语音数据帧的获取顺序,对多个语音数据帧进行合并,得到第二语音数据。或者,在即时语音通话的场景下,接收设备在依次获取每个冗余子数据对应的语音数据帧的过程中,依次输出得到的语音数据帧。

[0224] 在另一种可能实现方式中,关键语音特征中包括线谱对参数、基音周期参数和第一增益参数,基于基音周期参数和第一增益参数能够确定第一激励参数,非关键语音特征为第二激励参数,获取第一激励参数和第二激励参数的过程详见上述步骤407,在此不再一一赘述。则接收设备基于线谱对参数,对第一激励参数和第二激励参数进行合成滤波,得到第二语音数据。

[0225] 可选地,接收设备将第一激励参数和第二激励参数的和确定为合并激励参数,然后将该线谱对参数和该合并激励参数输入至语音合成滤波器中进行合成滤波,从而得到第二语音数据,该第二语音数据为PCM(Pulse Code Modulation,脉冲编码调制)格式的语音数据。

[0226] 图6是本申请实施例提供的一种解码方法的示意图,参见图6,接收设备判断第一语音数据是否丢失,如果第一语音数据未丢失,则接收设备直接对该第一语音数据进行解码和滤波,从而恢复出第二语音数据。如果第一语音数据丢失,则接收设备对冗余数据进行解码和滤波,从而恢复出第二语音数据,达到抗丢失的效果。

[0227] 需要说明的是,本申请实施例仅以第一语音数据和冗余数据在同一数据包中为例进行说明。在另一实施例中,发送设备生成包括第一语音数据的第一数据包,生成包括冗余数据的第二数据包,向接收设备发送第一数据包和第二数据包。也即是,发送设备将第一语音数据和冗余数据放在不同的数据包中分别进行传输,由于第一语音数据和冗余数据位于不同的数据包中,因此该第一语音数据和冗余数据之间没有相互约束的关系,减少了带宽竞争的问题。可选地,发送设备采用带外FEC编码的方式,将第一语音数据和冗余数据放在不同的数据包中,带外FEC编码是指将语音数据以及用于进行纠错的冗余数据放在不同的数据包中。

[0228] 本申请实施例提供了一种利用关键语音特征来预测非关键语音特征的方法,在未获取到冗余数据对应的第一语音数据的情况下,对冗余数据进行特征解码得到关键语音特征,并预测出非关键语音特征,从而对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据,因此在丢失第一语音数据的情况下,依然能够恢复出第二语音数据,达到了数据的抗丢失效果。

[0229] 并且,由于采用了基于关键语音特征预测非关键语音特征的方式,因此冗余数据中仅需携带关键语音特征,无需携带非关键语音特征,从而降低了冗余数据的数据量,因此能够提高传输目标数据包过程中的带宽利用率。

[0230] 需要说明的是上述图4的实施例说明了接收设备基于关键语音特征预测非关键语音特征的过程。在一种可能实现方式中,接收设备调用特征预测模型,基于关键语音特征,预测非关键语音特征。图7是本申请实施例提供的一种预测非关键语音特征的流程图,参见图7,该方法包括以下步骤。

[0231] 701、接收设备接收目标数据包。

[0232] 702、接收设备在未获取到冗余数据对应的第一语音数据的情况下,对冗余数据进

行特征解码,得到关键语音特征。

[0233] 该步骤701-702与上述步骤405-406同理,在此不再一一赘述。

[0234] 703、接收设备调用特征预测模型,基于关键语音特征,预测非关键语音特征。

[0235] 接收设备获取到关键语音特征之后,调用特征预测模型,基于该关键语音特征,预测第二语音数据的非关键语音特征,从而得到第二语音数据的关键语音特征和非关键语音特征。其中,接收设备中存储有特征预测模型,该特征预测模型为采用深度学习的方式基于样本语音数据训练得到的模型,该特征预测模型用于预测非关键语音特征,本申请实施例中通过特征预测模型来实现预测非关键语音特征的过程。可选地,该特征预测模型为接收设备训练的模型,或者为其他设备训练好之后,发送给该接收设备的模型,本申请实施例对此不做限定。其中,该特征预测模型的训练过程详见下述图10的实施例,在此暂不作说明。

[0236] 在一种可能实现方式中,关键语音特征包括线谱对参数、基音周期参数和第一增益参数,非关键语音特征包括第二激励参数。接收设备基于基音周期参数和第一增益参数,确定第一激励参数,调用特征预测模型,基于线谱对参数和第一激励参数,预测第二激励参数。

[0237] 在一种可能实现方式中,第二语音数据包括多个语音数据帧,冗余数据中包括与多个语音数据帧对应的多个冗余子数据,接收设备调用特征预测模型,基于每个冗余子数据对应的关键语音特征,预测每个冗余子数据对应的非关键语音特征。其中,对于多个冗余子数据,接收设备采用以下第三种方式和第四种方式预测冗余子数据对应的非关键语音特征。

[0238] 第三种方式:对于第一个冗余子数据之后的每个冗余子数据,接收设备获取冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧,调用特征预测模型,基于冗余子数据对应的关键语音特征、上一个冗余子数据对应的非关键语音特征以及上一个冗余子数据对应的语音数据帧,预测冗余子数据对应的非关键语音特征。其中,上一个冗余子数据对应的语音数据帧的获取方法详见下述步骤704,在此暂不做说明。

[0239] 可选地,关键语音特征包括第二语音数据的线谱对参数,以及每个冗余子数据对应的基音周期参数和第一增益参数。则对于第一个冗余子数据之后的每个冗余子数据,接收设备基于该冗余子数据对应的基音周期参数和第一增益参数,确定冗余子数据对应的第一激励参数,然后调用特征预测模型,基于线谱对参数、冗余子数据对应的第一激励参数、上一个冗余子数据对应的非关键语音特征以及上一个冗余子数据对应的语音数据帧,预测冗余子数据对应的非关键语音特征。也即是,将线谱对参数、冗余子数据对应的第一激励参数、上一个冗余子数据对应的非关键语音特征和上一个冗余子数据对应的语音数据帧输入至特征预测模型中,该特征预测模型输出该冗余子数据对应的非关键语音特征。

[0240] 第四种方式:对于第一个冗余子数据,接收设备调用特征预测模型,基于第一个冗余子数据对应的关键语音特征、参考非关键语音特征以及参考语音数据帧,预测第一个冗余子数据对应的非关键语音特征。其中,特征预测模型中包括参考非关键语音特征和参考语音数据帧。因此,接收设备将第一个冗余子数据对应的关键语音特征输入至特征预测模型中,该特征预测模型进行处理后,输出该第一个冗余子数据对应的非关键语音特征。

[0241] 可选地,关键语音特征包括第二语音数据的线谱对参数以及每个冗余子数据对应的基音周期参数和第一增益参数。则接收设备基于第一个冗余子数据对应的基音周期参数

和第一增益参数,确定第一个冗余子数据对应的第一激励参数,然后调用特征预测模型,基于线谱对参数、第一个冗余子数据对应的第一激励参数、参考非关键语音特征以及参考语音数据帧,预测第一个冗余子数据对应的非关键语音特征。也即是,将线谱对参数和冗余子数据对应的第一激励参数输入至特征预测模型中,该特征预测模型进行处理后,输出该第一个冗余子数据对应的非关键语音特征。

[0242] 该步骤703与上述步骤407中预测非关键语音特征的过程同理,不同之处在于,该步骤703中是调用特征预测模型来实现预测非关键语音特征的过程的。

[0243] 在一种可能实现方式中,特征预测模型中包括多个全连接网络和多个门控循环网络。图8是本申请实施例提供的一种特征预测模型的结构示意图,该特征预测模型中包括3个全连接网络和3个门控循环网络。当语音数据的采样率为8khz时,线谱对参数为10阶数据,第一激励参数为10阶数据、语音数据帧为10阶数据、非关键语音特征为10阶数据,所以特征预测模型输入的数据量为40阶,其中第一全连接网络801、第二全连接网络802、第三全连接网络803的神经元个数分别为64、64、10,而第一门控循环网络804、第二门控循环网络805、第三门控循环网络806的神经元个数分别为64、256、10;而当语音数据的采样率为16khz或16khz以上时,线谱对参数为16阶数据,第一激励参数为16阶数据,语音数据帧为16阶数据,非关键语音特征为16阶数据,所以特征预测模型输入的数据量为64阶,其中第一全连接网络801、第二全连接网络802、第三全连接网络803的神经元个数分别为64、64、16,而第一门控循环网络804、第二门控循环网络805、第三门控循环网络806的神经元个数分别为64、256、16。

[0244] 704、接收设备对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据。

[0245] 在一种可能实现方式中,第二语音数据包括多个语音数据帧,冗余数据中包括与多个语音数据帧对应的多个参考子数据。上述步骤703中,接收设备对每个冗余子数据对应的关键语音特征进行处理,得到了每个冗余子数据对应的非关键语音数据。则接收设备对每个冗余子数据对应的关键语音特征和非关键语音特征进行合成滤波,得到每个冗余子数据对应的语音数据帧,按照多个冗余子数据的排列顺序,对得到的多个语音数据帧进行合并,得到第二语音数据。

[0246] 该步骤704与上述步骤408中解码得到第二语音数据的过程同理,在此不再一一赘述。

[0247] 图9是本申请实施例提供的一种解码方法的示意图,如图9所示,接收设备获取线谱对参数以及当前的冗余子数据对应的基音周期参数和第一增益参数,确定基音周期参数匹配的第一码本参数,将第一码本参数与第一增益参数的乘积确定为第一激励参数,将第一激励参数、线谱对参数以及上一个冗余子数据对应的非关键语音特征和语音数据帧输入至特征预测模型中,特征预测模型输出当前的冗余子数据对应的非关键语音特征,将该非关键语音特征与第一激励参数的和,以及线谱对参数输入语音合成滤波器中进行合成滤波,得到该冗余子数据对应的语音数据帧。

[0248] 本申请实施例提供了一种调用特征预测模型来预测非关键语音特征的方法,能够对冗余数据进行特征解码得到关键语音特征,并调用特征预测模型基于关键语音特征预测出非关键语音特征,从而对关键语音特征和非关键语音特征进行合成滤波,得到第二语音

数据,因此在丢失第一语音数据的情况下,依然能够恢复出第二语音数据,达到了数据的抗丢失效果。

[0249] 并且,由于采用了基于关键语音特征预测非关键语音特征的方式,因此冗余数据中仅需携带关键语音特征,无需携带非关键语音特征,从而降低了冗余数据的数据量,因此能够提高传输目标数据包过程中的带宽利用率。

[0250] 并且,采用特征预测模型来预测非关键语音特征,由于该特征预测模型为采用深度学习的方式基于样本语音数据训练得到的模型,因此,该特征预测模型预测得到的非关键语音特征更加准确。

[0251] 图10是本申请实施例提供的一种模型训练方法的流程图,本申请实施例中训练得到的特征预测模型,能够应用于上述图7的实施例中,来预测非关键语音特征。本申请实施例的执行主体为计算机设备,参见图10,该方法包括以下步骤。

[0252] 1001、计算机设备对样本语音数据进行特征提取,得到样本语音数据对应的多个样本语音特征。

[0253] 其中,多个样本语音特征中包括样本关键语音特征。

[0254] 1002、计算机设备调用特征预测模型,基于样本关键语音特征进行预测,得到预测非关键语音特征。

[0255] 在一种可能实现方式中,样本关键语音特征中包括样本线谱对参数、样本基音周期参数和第一样本增益参数,预测非关键语音特征包括第二样本激励参数。计算机设备基于样本基音周期参数和第一样本增益参数,确定第一样本激励参数,调用特征预测模型,基于样本线谱对参数和第一样本激励参数,预测第二样本激励参数,第一样本激励参数用于反映样本语音数据中具有相关性的语音特征,第二样本激励参数用于反映样本语音数据中不具有相关性的语音特征。样本线谱对参数用于反映样本语音数据的音谱特征,样本基音周期参数用于反映样本语音数据中的声带振动频率,第一样本增益参数用于调节样本语音数据中具有相关性的信号的强度。

[0256] 在一种可能实现方式中,样本语音数据包括多个样本语音数据帧。计算机设备调用特征预测模型,基于每个样本语音数据帧对应的样本关键语音特征,预测每个样本语音数据帧对应的预测非关键语音特征。其中,对于多个样本语音数据帧,接收设备采用以下第五种方式和第六种方式预测样本语音数据帧对应的预测非关键语音特征。

[0257] 第五种方式:对于第一个样本语音数据帧之后的每个样本语音数据帧,计算机设备获取样本语音数据帧的上一个样本语音数据帧对应的预测非关键语音特征和样本解码数据帧,调用特征预测模型,基于样本语音数据帧对应的样本关键语音特征、上一个样本语音数据帧对应的预测非关键语音特征和样本解码数据帧,预测样本语音数据帧对应的预测非关键语音特征。其中,上一个样本语音数据帧对应的样本解码数据帧的获取方法详见下述步骤1003,在此暂不做说明。

[0258] 可选地,样本关键语音特征包括样本线谱对参数,以及每个样本语音数据帧对应的样本基音周期参数和第一样本增益参数。则对于第一个样本语音数据帧之后的每个样本语音数据帧,计算机设备基于该样本语音数据帧对应的样本基音周期参数和第一样本增益参数,确定样本语音数据帧对应的第一样本激励参数,然后基于样本线谱对参数、样本语音数据帧对应的第一样本激励参数、上一个样本语音数据帧对应的预测非关键语音特征以及

上一个样本语音数据帧对应的样本解码数据帧,预测样本语音数据帧对应的预测非关键语音特征。

[0259] 第六种方式:特征预测模型中包括参考非关键语音特征和参考语音数据帧,对于第一个样本语音数据帧,计算机设备调用特征预测模型,基于样本语音数据帧对应的样本关键语音特征、参考非关键语音特征以及参考语音数据帧,预测第一个样本语音数据帧对应的预测非关键语音特征。

[0260] 可选地,样本关键语音特征包括样本线谱对参数以及每个样本语音数据帧对应的样本基音周期参数和第一样本增益参数。则计算机设备基于第一个样本语音数据帧对应的样本基音周期参数和第一样本增益参数,确定第一个样本语音数据帧对应的第一样本激励参数,然后调用特征预测模型,基于样本线谱对参数、第一个样本语音数据帧对应的第一样本激励参数、参考非关键语音特征以及参考语音数据帧,预测第一个样本语音数据帧对应的预测非关键语音特征。

[0261] 该步骤1002与上述步骤703中预测非关键语音特征的过程同理,不同之处在于,该步骤1002是调用特征预测模型来预测样本语义数据对应的非关键语音特征的。

[0262] 1003、计算机设备对样本关键语音特征和预测非关键语音特征进行合成滤波,得到样本解码数据。

[0263] 在一种可能实现方式中,样本语音数据包括多个样本语音数据帧,样本解码数据中包括每个样本语音数据帧对应的样本解码数据帧,上述步骤1002中,计算机设备对每个样本语音数据帧对应的样本关键语音特征进行处理,得到了每个样本语音数据帧对应的预测非关键语音特征。则计算机设备对每个样本语音数据帧对应的样本关键语音特征和预测非关键语音特征进行合成滤波,得到每个样本语音数据帧对应的样本解码数据帧。

[0264] 该步骤1003与上述步骤408和步骤704中的过程同理,在此不再一一赘述。

[0265] 1004、计算机设备基于样本语音数据、样本解码数据、样本关键语音特征以及预测非关键语音特征,训练特征预测模型。

[0266] 计算机设备获取到样本语音数据、样本解码数据、样本关键语音特征以及预测非关键语音特征之后,基于样本语音数据、样本解码数据、样本关键语音特征以及预测非关键语音特征,训练特征预测模型。

[0267] 在一种可能实现方式中,计算机设备将样本语音数据与样本解码数据之间的误差确定为残差特征,将残差特征与样本关键语音特征之间的差值确定为样本非关键语音特征,基于预测非关键语音特征与样本非关键语音特征之间的误差,训练特征预测模型。其中,计算机设备通过训练特征预测模型,以使预测非关键语音特征与样本非关键语音特征之间的误差越来越小。例如,计算机设备采用Cross-Entropy(交叉熵)准则来训练特征预测模型。

[0268] 其中,样本解码数据为自适应加权滤波的结果,样本语音数据与样本解码数据之间的误差为残差特征,该残差特征是由样本关键语音特征和样本非关键语音特征导致的,因此残差特征与样本关键语音特征之间的差值为样本非关键语音特征。特征预测模型的训练目的即为得到该样本非关键语音特征,因此通过调整特征预测模型的模型参数,以使特征预测模型输出的预测非关键语音特征与该样本非关键语音特征之间的误差越来越小。

[0269] 可选地,样本语音数据中包括多个样本语音数据帧,每个样本语音数据帧对应

样本解码数据帧,对于每个样本语音数据帧,计算机设备将该样本语音数据帧与该样本语音数据帧的上一个样本语音数据帧对应的样本解码数据帧之间的误差确定为该样本语音数据帧对应的残差特征,将该样本语音数据帧对应的残差特征与该样本语音数据帧对应的样本关键语音特征之间的差值,确定为该样本语音数据帧对应的样本非关键语音特征,并基于该样本语音数据帧对应的预测非关键语音特征与样本非关键语音特征之间的误差,训练特征预测模型。

[0270] 图11是本申请实施例提供的一种模型训练方法的示意图,参见图11,语音数据中包括多个语音数据帧,对于任一语音数据帧,首先对语音数据帧进行预处理,例如进行高通滤波等,然后对预处理得到的数据进行LP(线性预测)分析,得到线谱对参数。将该语音数据帧与上一个语音数据帧对应的解码数据帧之间的差值输入自适应加权滤波器中,得到残差特征,基于残差特征对应的最小感知加权误差,来确定自适应码本参数(第一码本参数)和自适应增益参数(第一增益参数),将自适应码本参数与自适应增益参数的乘积确定为自适应激励参数(第一激励参数)。计算机设备将自适应激励参数、线谱对参数、上一个语音数据帧对应的解码数据帧以及上一个语音数据帧对应的固定激励参数(非关键语音特征)输入至特征预测模型中,该特征预测模型输出该语音数据帧对应的固定激励参数,然后将固定激励参数和自适应激励参数进行相加后,与线谱对参数一同输入线性预测滤波器中,由线性预测滤波器输出该语音数据帧对应的解码数据帧。其中,该线性预测滤波器用于进行合成滤波。其中,计算机设备将该残差特征与自适应激励参数之间的差值确定为样本固定激励参数,并基于特征预测模型输出的固定激励参数与该样本固定激励参数之间的误差,训练该特征预测模型。

[0271] 本申请实施例提供的方法,通过训练特征预测模型,实现通过特征预测模型来预测语音数据的非关键语音特征,后续在生成语音数据对应的冗余数据时,仅需对关键语音特征进行编码,无需对非关键语音特征进行编码,从而降低了冗余数据的数据量。

[0272] 图12是本申请实施例提供的一种语音数据处理装置的结构示意图。参见图12,该装置包括:

[0273] 数据包接收模块1201,用于接收目标数据包,目标数据包是发送设备对第一语音数据和冗余数据进行组包后发送的,第一语音数据是对第二语音数据的多个语音特征进行特征编码得到的,多个语音特征中包括关键语音特征和非关键语音特征,冗余数据是关键语音特征进行特征编码得到的;

[0274] 特征解码模块1202,用于在目标数据包中未获取到第一语音数据的情况下,对冗余数据进行特征解码,得到关键语音特征;

[0275] 特征预测模块1203,用于基于关键语音特征,预测非关键语音特征;

[0276] 合成滤波模块1204,用于对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据。

[0277] 本申请实施例提供的语音数据处理装置,基于语音数据的关键语音特征来预测非关键语音特征,在未获取到冗余数据对应的第一语音数据的情况下,对冗余数据进行特征解码得到关键语音特征,并预测出非关键语音特征,从而对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据,因此在丢失第一语音数据的情况下,依然能够恢复出第二语音数据,达到了数据的抗丢失效果。并且,由于冗余数据中仅需携带关键语音特征,

无需携带非关键语音特征,从而降低了冗余数据的数据量,因此能够提高传输目标数据包过程中的带宽利用率。

[0278] 可选地,参见图13,关键语音特征包括线谱对参数、基音周期参数和第一增益参数,非关键语音特征包括第二激励参数,特征预测模块1203,包括:

[0279] 第一激励确定单元1213,用于基于基音周期参数和第一增益参数,确定第一激励参数,第一激励参数用于反映第二语音数据中具有相关性的语音特征;

[0280] 第二激励确定单元1223,用于基于线谱对参数和第一激励参数,预测第二激励参数,第二激励参数用于反映第二语音数据中不具有相关性的语音特征。

[0281] 可选地,参见图13,合成滤波模块1204,包括:

[0282] 合成滤波单元1214,用于基于线谱对参数,对第一激励参数和第二激励参数进行合成滤波,得到第二语音数据。

[0283] 可选地,参见图13,第一激励确定单元1213,用于:

[0284] 在码本集合中查询与基音周期参数匹配的第一码本参数,码本集合中包括至少一组码本参数;

[0285] 将第一码本参数与第一增益参数的乘积,确定为第一激励参数。

[0286] 可选地,参见图13,第二语音数据包括多个语音数据帧,冗余数据中包括与多个语音数据帧对应的多个冗余子数据,特征预测模块1203,包括:

[0287] 特征预测单元1233,用于基于每个冗余子数据对应的关键语音特征,预测每个冗余子数据对应的非关键语音特征;

[0288] 合成滤波模块1204,包括:

[0289] 数据帧获取单元1224,用于对每个冗余子数据对应的关键语音特征和非关键语音特征进行合成滤波,得到每个冗余子数据对应的语音数据帧;

[0290] 数据帧合并单元1234,用于按照多个冗余子数据的排列顺序,对得到的多个语音数据帧进行合并,得到第二语音数据。

[0291] 可选地,参见图13,特征预测单元1233,用于:

[0292] 对于第一个冗余子数据之后的每个冗余子数据,获取冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧;

[0293] 基于冗余子数据对应的关键语音特征、上一个冗余子数据对应的非关键语音特征以及上一个冗余子数据对应的语音数据帧,预测冗余子数据对应的非关键语音特征。

[0294] 可选地,参见图13,关键语音特征包括语音数据的线谱对参数以及每个冗余子数据对应的基音周期参数和第一增益参数,特征预测单元1233,用于:

[0295] 基于冗余子数据对应的基音周期参数和第一增益参数,确定冗余子数据对应的第一激励参数;

[0296] 基于线谱对参数、冗余子数据对应的第一激励参数、上一个冗余子数据对应的非关键语音特征以及上一个冗余子数据对应的语音数据帧,预测冗余子数据对应的非关键语音特征。

[0297] 可选地,参见图13,特征预测单元1233,用于:

[0298] 对于第一个冗余子数据,获取第一个冗余子数据对应的参考非关键语音特征和参考语音数据帧;

- [0299] 基于第一个冗余子数据对应的关键语音特征、参考非关键语音特征以及参考语音数据帧,预测第一个冗余子数据对应的非关键语音特征。
- [0300] 可选地,参见图13,关键语音特征包括第二语音数据的线谱对参数以及每个冗余子数据对应的基音周期参数和第一增益参数,特征预测单元1233,用于:
- [0301] 基于第一个冗余子数据对应的基音周期参数和第一增益参数,确定第一个冗余子数据对应的第一激励参数;
- [0302] 基于线谱对参数、第一个冗余子数据对应的第一激励参数、参考非关键语音特征以及参考语音数据帧,预测第一个冗余子数据对应的非关键语音特征。
- [0303] 可选地,参见图13,特征预测模块1203,包括:
- [0304] 模型调用单元1243,用于调用特征预测模型,基于关键语音特征,预测非关键语音特征。
- [0305] 可选地,参见图13,关键语音特征包括线谱对参数、基音周期参数和第一增益参数,非关键语音特征包括第二激励参数,模型调用单元1243,用于:
- [0306] 基于基音周期参数和第一增益参数,确定第一激励参数;
- [0307] 调用特征预测模型,基于线谱对参数和第一激励参数,预测第二激励参数。
- [0308] 可选地,参见图13,第二语音数据包括多个语音数据帧,冗余数据中包括与多个语音数据帧对应的多个冗余子数据,模型调用单元1243,用于:
- [0309] 调用特征预测模型,基于每个冗余子数据对应的关键语音特征,预测每个冗余子数据对应的非关键语音特征;
- [0310] 合成滤波模块1204,包括:
- [0311] 数据帧获取单元1224,用于对每个冗余子数据对应的关键语音特征和非关键语音特征进行合成滤波,得到每个冗余子数据对应的语音数据帧;
- [0312] 数据帧合并单元1234,用于按照多个冗余子数据的排列顺序,对得到的多个语音数据帧进行合并,得到第二语音数据。
- [0313] 可选地,参见图13,模型调用单元1243,用于:
- [0314] 对于第一个冗余子数据之后的每个冗余子数据,获取冗余子数据的上一个冗余子数据对应的非关键语音特征和语音数据帧;
- [0315] 调用特征预测模型,基于冗余子数据对应的关键语音特征、上一个冗余子数据对应的非关键语音特征以及上一个冗余子数据对应的语音数据帧,预测冗余子数据对应的非关键语音特征。
- [0316] 可选地,参见图13,特征预测模型中包括参考非关键语音特征和参考语音数据帧,模型调用单元1243,用于:
- [0317] 对于第一个冗余子数据,调用特征预测模型,基于第一个冗余子数据对应的关键语音特征、参考非关键语音特征以及参考语音数据帧,预测第一个冗余子数据对应的非关键语音特征。
- [0318] 可选地,参见图13,装置还包括:
- [0319] 特征提取模块1205,用于对样本语音数据进行特征提取,得到样本语音数据对应的多个样本语音特征,多个样本语音特征中包括样本关键语音特征;
- [0320] 特征预测模块1203,还用于调用特征预测模型,基于样本关键语音特征进行预测,

得到预测非关键语音特征；

[0321] 合成滤波模块1204,还用于对样本关键语音特征和预测非关键语音特征进行合成滤波,得到样本解码数据；

[0322] 模型训练模块1206,用于基于样本语音数据、样本解码数据、样本关键语音特征以及预测非关键语音特征,训练特征预测模型。

[0323] 可选地,参见图13,模型训练模块1206,包括:

[0324] 第一特征确定单元1216,用于将样本语音数据与样本解码数据之间的误差确定为残差特征；

[0325] 第二特征确定单元1226,用于将残差特征与样本关键语音特征之间的差值确定为样本非关键语音特征；

[0326] 模型训练单元1236,用于基于预测非关键语音特征与样本非关键语音特征之间的误差,训练特征预测模型。

[0327] 可选地,参见图13,样本关键语音特征中包括样本线谱对参数、样本基音周期参数和第一样本增益参数,特征预测模块1203,包括:

[0328] 第一激励确定单元1213,用于基于样本基音周期参数和第一样本增益参数,确定第一样本激励参数,第一样本激励参数用于反映样本语音数据中具有相关性的语音特征；

[0329] 第二激励确定单元1223,用于调用特征预测模型,基于样本线谱对参数和第一样本激励参数,预测第二样本激励参数,第二样本激励参数用于反映样本语音数据中不具有相关性的语音特征。

[0330] 可选地,参见图13,样本语音数据包括多个样本语音数据帧,特征预测模块1203,包括:

[0331] 特征预测单元1233,用于调用特征预测模型,基于每个样本语音数据帧对应的样本关键语音特征,预测每个样本语音数据帧对应的预测非关键语音特征；

[0332] 样本解码数据中包括每个样本语音数据帧对应的样本解码数据帧,合成滤波模块1204,包括:

[0333] 数据帧获取单元1224,用于对每个样本语音数据帧对应的样本关键语音特征和预测非关键语音特征进行合成滤波,得到每个样本语音数据帧对应的样本解码数据帧。

[0334] 可选地,参见图13,特征预测单元1233,用于:

[0335] 对于第一个样本语音数据帧之后的每个样本语音数据帧,获取样本语音数据帧的上一个样本语音数据帧对应的预测非关键语音特征和样本解码数据帧；

[0336] 调用特征预测模型,基于样本语音数据帧对应的样本关键语音特征、上一个样本语音数据帧对应的预测非关键语音特征和样本解码数据帧,预测样本语音数据帧对应的预测非关键语音特征。

[0337] 可选地,参见图13,特征预测模型中包括参考非关键语音特征和参考语音数据帧,特征预测单元1233,用于:

[0338] 对于第一个样本语音数据帧,调用特征预测模型,基于样本语音数据帧对应的样本关键语音特征、参考非关键语音特征以及参考语音数据帧,预测第一个样本语音数据帧对应的预测非关键语音特征。

[0339] 需要说明的是:上述实施例提供的语音数据处理装置在处理语音数据时,仅以上

述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将计算机设备的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。另外,上述实施例提供的语音数据处理装置与语音数据处理方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0340] 图14是本申请实施例提供的一种语音数据处理装置的结构示意图。参见图14,该装置包括:

[0341] 特征提取模块1401,用于对第二语音数据进行特征提取,得到第二语音数据的多个语音特征,多个语音特征中包括关键语音特征和非关键语音特征;

[0342] 第一特征编码模块1402,用于对多个语音特征进行特征编码,得到第一语音数据;

[0343] 第二特征编码模块1403,用于对关键语音特征进行特征编码,得到冗余数据;

[0344] 组包模块1404,用于对第一语音数据和冗余数据进行组包,得到目标数据包;

[0345] 数据包发送模块1405,用于向接收设备发送目标数据包,接收设备用于在目标数据包中未获取到第一语音数据的情况下,对冗余数据进行特征解码,得到关键语音特征,基于关键语音特征,预测非关键语音特征,对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据。

[0346] 本申请实施例提供的语音数据处理装置,在传输对第二语音数据的多个语音特征进行编码得到的第一语音数据时,还会传输对该第二语音数据的关键语音特征进行编码得到的冗余数据,因此在接收设备未获取到第一语音数据时,基于冗余数据中的关键语音特征预测非关键语音特征,从而对关键语音特征和非关键语音特征进行合成滤波,得到第二语音数据,因此在丢失第一语音数据的情况下,依然能够恢复出第二语音数据,达到了数据的抗丢失效果。并且,由于能够基于关键语音特征预测非关键语音特征,因此冗余数据中仅需携带关键语音特征,无需携带非关键语音特征,从而降低了冗余数据的数据量,因此能够提高传输目标数据包过程中的带宽利用率。

[0347] 可选地,参见图15,关键语音特征包括线谱对参数、基音周期参数和第一增益参数,第二特征编码模块1403,包括:

[0348] 特征编码单元1413,用于对线谱对参数、基音周期参数和第一增益参数进行特征编码,得到冗余数据。

[0349] 需要说明的是:上述实施例提供的语音数据处理装置在处理语音数据时,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将计算机设备的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。另外,上述实施例提供的语音数据处理装置与语音数据处理方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0350] 本申请实施例还提供了一种计算机设备,该计算机设备包括处理器和存储器,存储器中存储有至少一条计算机程序,该至少一条计算机程序由处理器加载并执行,以实现上述实施例的语音数据处理方法中所执行的操作。

[0351] 可选地,该计算机设备提供为终端。图16示出了本申请一个示例性实施例提供的终端1600的结构示意图。

[0352] 终端1600包括有:处理器1601和存储器1602。

[0353] 处理器1601可以包括一个或多个处理核心,比如4核心处理器、8核心处理器等。处

处理器1601可以采用DSP(Digital Signal Processing,数字信号处理)、FPGA(Field Programmable Gate Array,现场可编程门阵列)、PLA(Programmable Logic Array,可编程逻辑阵列)中的至少一种硬件形式来实现。处理器1601也可以包括主处理器和协处理器,主处理器是用于对在唤醒状态下的数据进行处理的处理单元,也称CPU(Central Processing Unit,中央处理器);协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中,处理器1601可以集成有GPU(Graphics Processing Unit,图像处理的交互器),GPU用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中,处理器1601还可以包括AI(Artificial Intelligence,人工智能)处理器,该AI处理器用于处理有关机器学习的计算操作。

[0354] 存储器1602可以包括一个或多个计算机可读存储介质,该计算机可读存储介质可以是非暂态的。存储器1602还可包括高速随机存取存储器,以及非易失性存储器,比如一个或多个磁盘存储设备、闪存存储设备。在一些实施例中,存储器1602中的非暂态的计算机可读存储介质用于存储至少一条计算机程序,该至少一条计算机程序用于被处理器1601所具有以实现本申请中方法实施例提供的语音数据处理方法。

[0355] 在一些实施例中,终端1600还可选包括有:外围设备接口1603和至少一个外围设备。处理器1601、存储器1602和外围设备接口1603之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外围设备接口1603相连。可选地,外围设备包括:射频电路1604、显示屏1605、摄像头组件1606、音频电路1607、定位组件1608和电源1609中的至少一种。

[0356] 外围设备接口1603可被用于将I/O(Input /Output,输入/输出)相关的至少一个外围设备连接到处理器1601和存储器1602。在一些实施例中,处理器1601、存储器1602和外围设备接口1603被集成在同一芯片或电路板上;在一些其他实施例中,处理器1601、存储器1602和外围设备接口1603中的任意一个或两个可以在单独的芯片或电路板上实现,本实施例对此不加以限定。

[0357] 射频电路1604用于接收和发射RF(Radio Frequency,射频)信号,也称电磁信号。射频电路1604通过电磁信号与通信网络以及其他通信设备进行通信。射频电路1604将电信号转换为电磁信号进行发送,或者,将接收到的电磁信号转换为电信号。可选地,射频电路1604包括:天线系统、RF收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频电路1604可以通过至少一种无线通信协议来与其它设备进行通信。该无线通信协议包括但不限于:城域网、各代移动通信网络(2G、3G、4G及5G)、无线局域网和/或WiFi(Wireless Fidelity,无线保真)网络。在一些实施例中,射频电路1604还可以包括NFC(Near Field Communication,近距离无线通信)有关的电路,本申请对此不加以限定。

[0358] 显示屏1605用于显示UI(User Interface,用户界面)。该UI可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏1605是触摸显示屏时,显示屏1605还具有采集在显示屏1605的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器1601进行处理。此时,显示屏1605还可以用于提供虚拟按钮和/或虚拟键盘,也称软按钮和/或软键盘。在一些实施例中,显示屏1605可以为一个,设置在终端1600的前面板;在另一些实施例中,显示屏1605可以为至少两个,分别设置在终端1600的不同表面或呈折叠设

计;在另一些实施例中,显示屏1605可以是柔性显示屏,设置在终端1600的弯曲表面上或折叠面上。甚至,显示屏1605还可以设置成非矩形的不规则图形,也即异形屏。显示屏1605可以采用LCD(Liquid Crystal Display,液晶显示屏)、OLED(Organic Light-Emitting Diode,有机发光二极管)等材质制备。

[0359] 摄像头组件1606用于采集图像或视频。可选地,摄像头组件1606包括前置摄像头和后置摄像头。前置摄像头设置在终端1600的前面板,后置摄像头设置在终端1600的背面。在一些实施例中,后置摄像头为至少两个,分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种,以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及VR(Virtual Reality,虚拟现实)拍摄功能或者其它融合拍摄功能。在一些实施例中,摄像头组件1606还可以包括闪光灯。闪光灯可以是单色温闪光灯,也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合,可以用于不同色温下的光线补偿。

[0360] 音频电路1607可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波,并将声波转换为电信号输入至处理器1601进行处理,或者输入至射频电路1604以实现语音通信。出于立体声采集或降噪的目的,麦克风可以为多个,分别设置在终端1600的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器1601或射频电路1604的电信号转换为声波。扬声器可以是传统的薄膜扬声器,也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时,不仅可以将电信号转换为人类可听见的声波,也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中,音频电路1607还可以包括耳机插孔。

[0361] 定位组件1608用于定位终端1600的当前地理位置,以实现导航或LBS(Location Based Service,基于位置的服务)。定位组件1608可以是基于美国的GPS(Global Positioning System,全球定位系统)、中国的北斗系统、俄罗斯的格雷纳斯系统或欧盟的伽利略系统的定位组件。

[0362] 电源1609用于为终端1600中的各个组件进行供电。电源1609可以是交流电、直流电、一次性电池或可充电电池。当电源1609包括可充电电池时,该可充电电池可以支持有线充电或无线充电。该可充电电池还可以用于支持快充技术。

[0363] 本领域技术人员可以理解,图16中示出的结构并不构成对终端1600的限定,可以包括比图示更多或更少的组件,或者组合某些组件,或者采用不同的组件布置。

[0364] 可选地,该计算机设备提供为服务器。图17是本申请实施例提供的一种服务器的结构示意图,该服务器1700可因配置或性能不同而产生比较大的差异,可以包括一个或一个以上处理器(Central Processing Units,CPU)1701和一个或一个以上的存储器1702,其中,所述存储器1702中存储有至少一条计算机程序,所述至少一条计算机程序由所述处理器1701加载并执行以实现上述各个方法实施例提供的方法。当然,该服务器还可以具有有线或无线网络接口、键盘以及输入输出接口等部件,以便进行输入输出,该服务器还可以包括其他用于实现设备功能的部件,在此不做赘述。

[0365] 本申请实施例还提供了一种计算机可读存储介质,该计算机可读存储介质中存储有至少一条计算机程序,该至少一条计算机程序由处理器加载并执行,以实现上述实施例的语音数据处理方法中所执行的操作。

[0366] 本申请实施例还提供了一种计算机程序产品或计算机程序,计算机程序产品或计算机程序包括计算机程序代码,计算机程序代码存储在计算机可读存储介质中,计算机设备的处理器从计算机可读存储介质读取计算机程序代码,处理器执行计算机程序代码,使得计算机设备实现如上述实施例的语音数据处理方法中所执行的操作。

[0367] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0368] 以上所述仅为本申请实施例的可选实施例,并不用以限制本申请实施例,凡在本申请实施例的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

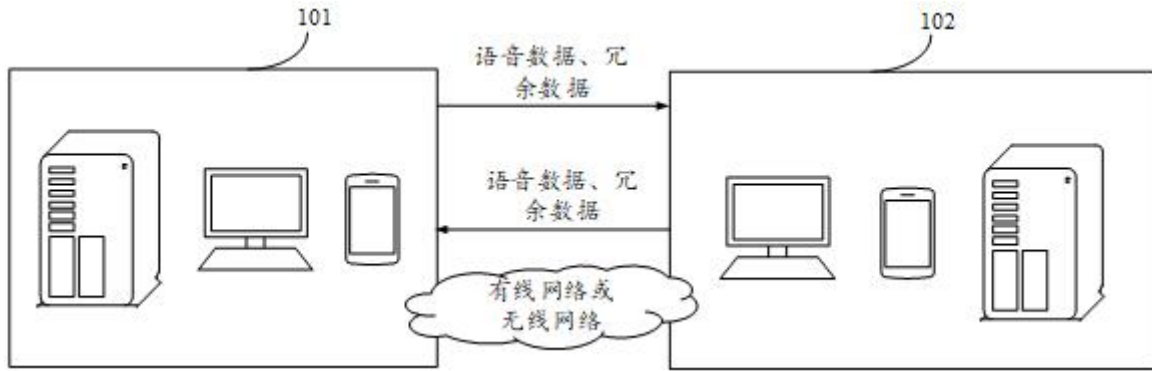


图1

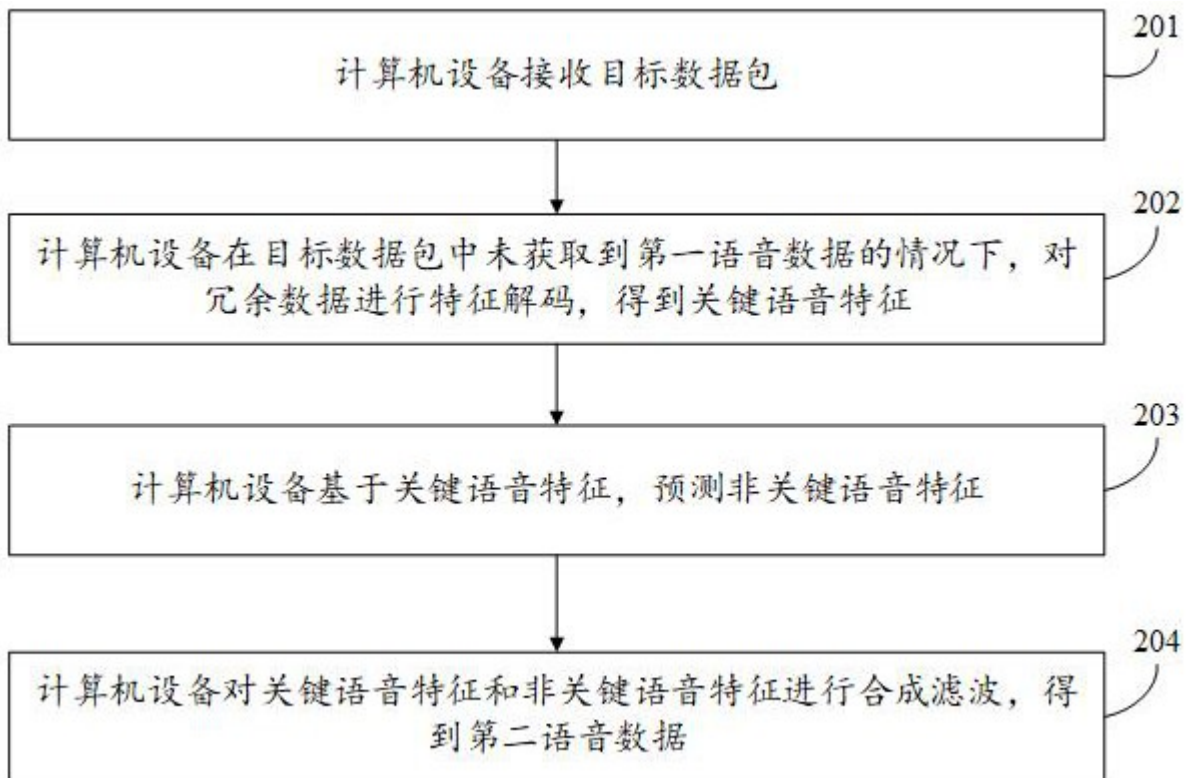


图2

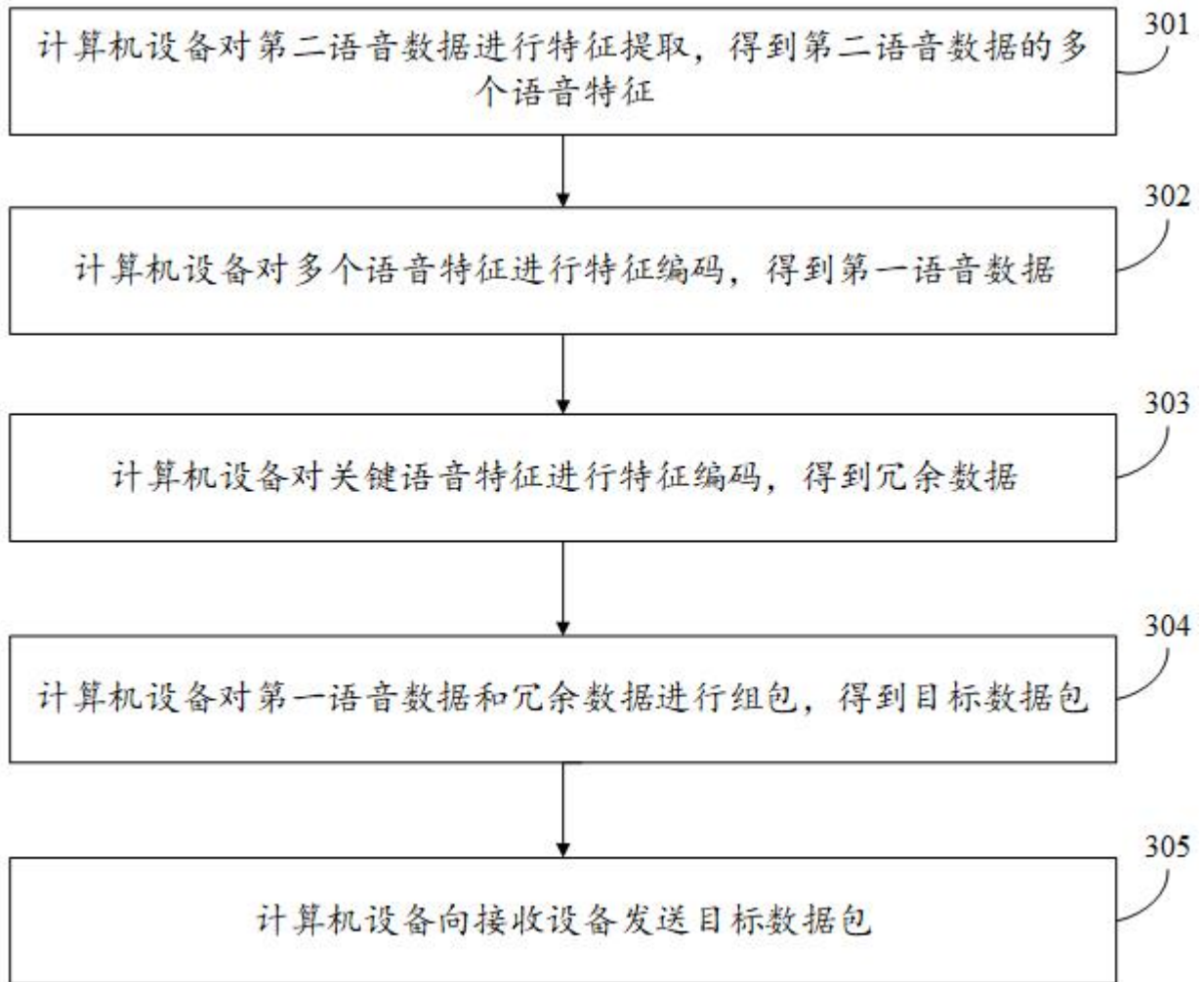


图3

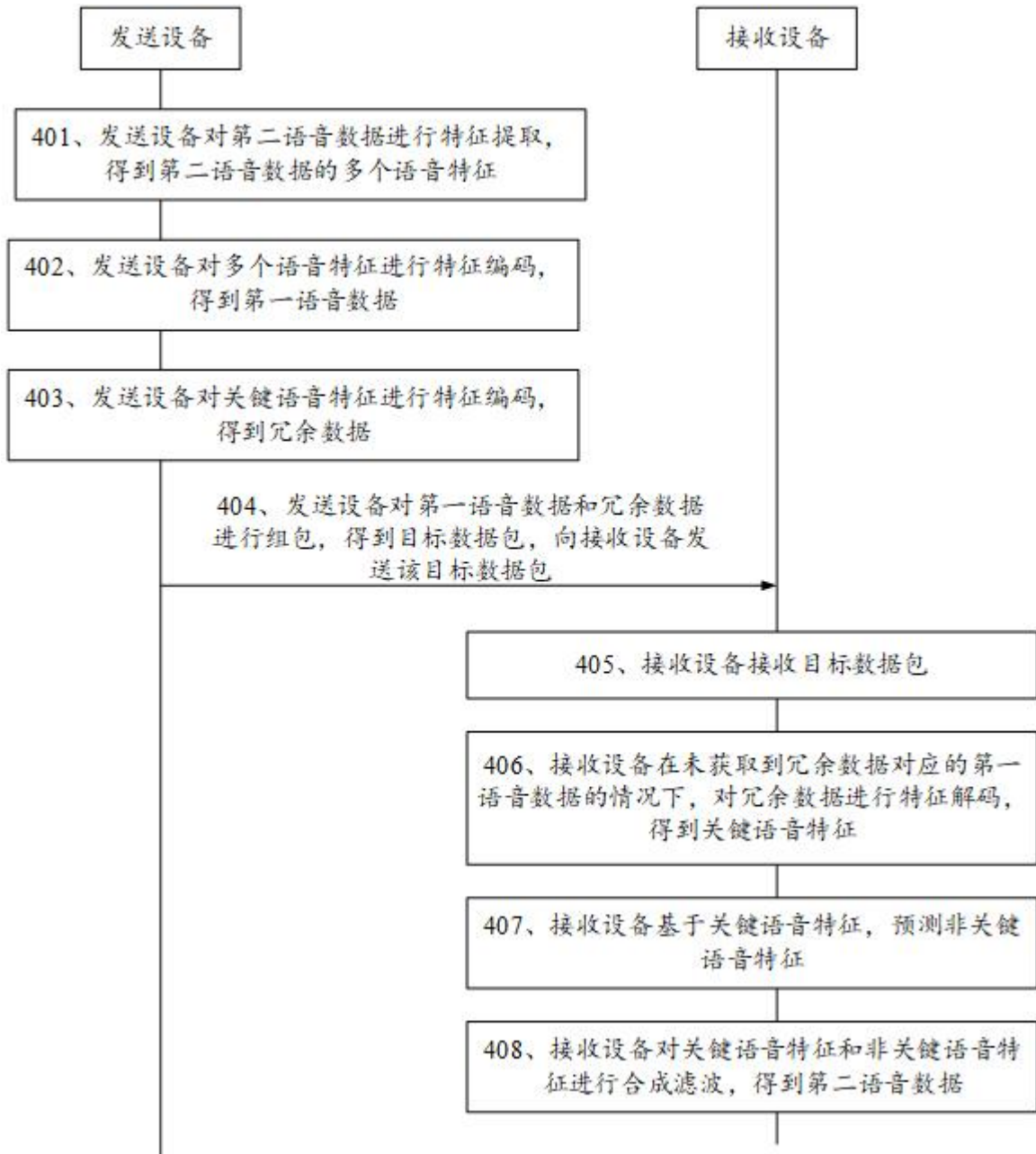


图4

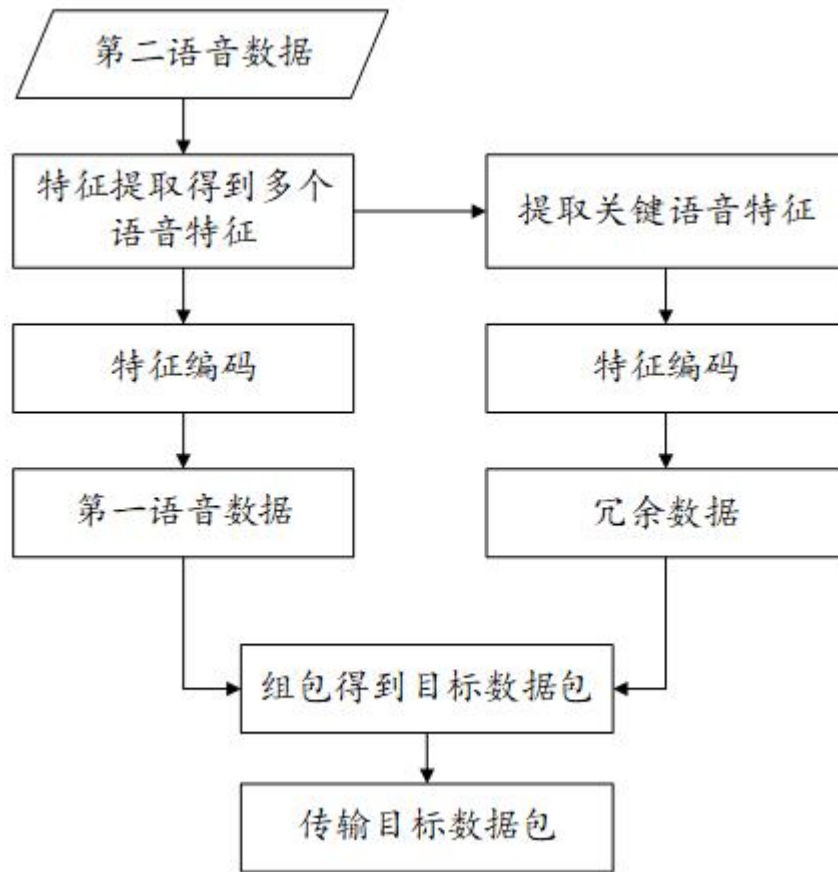


图5

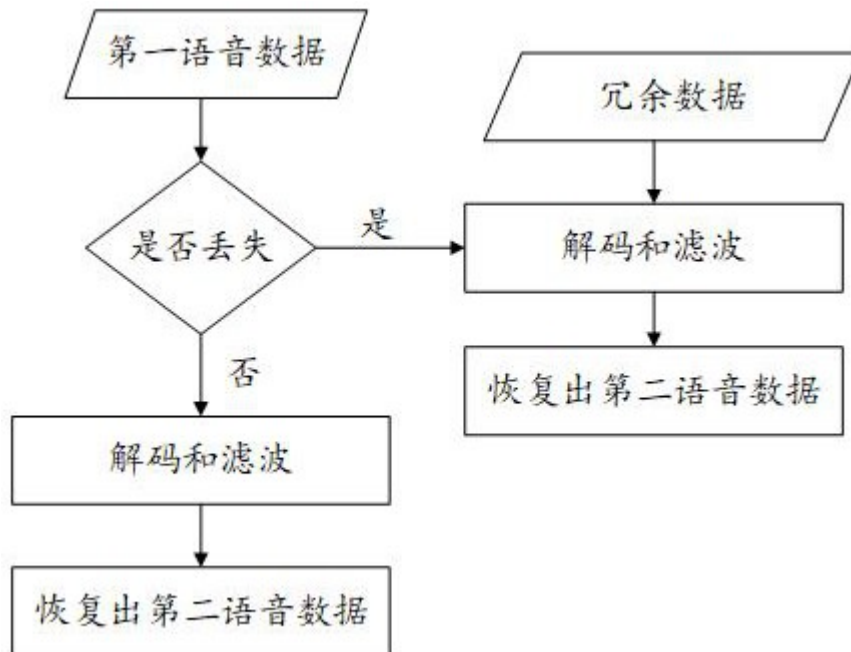


图6

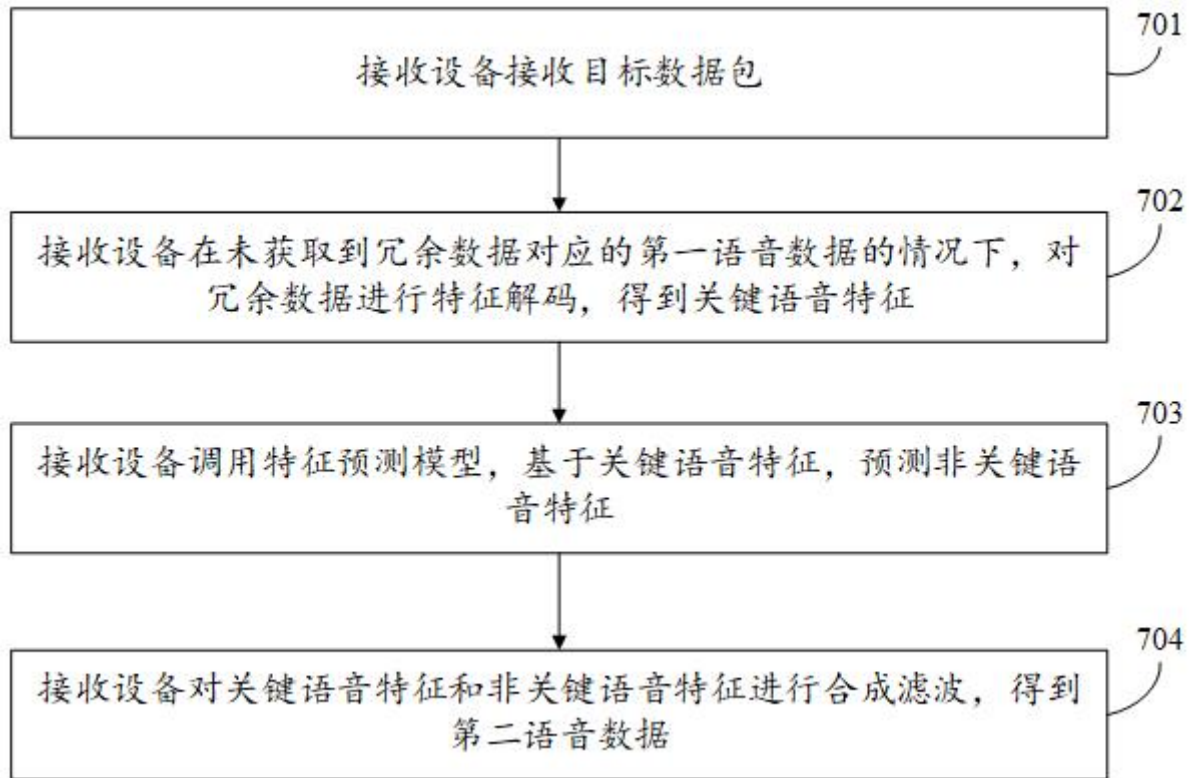


图7

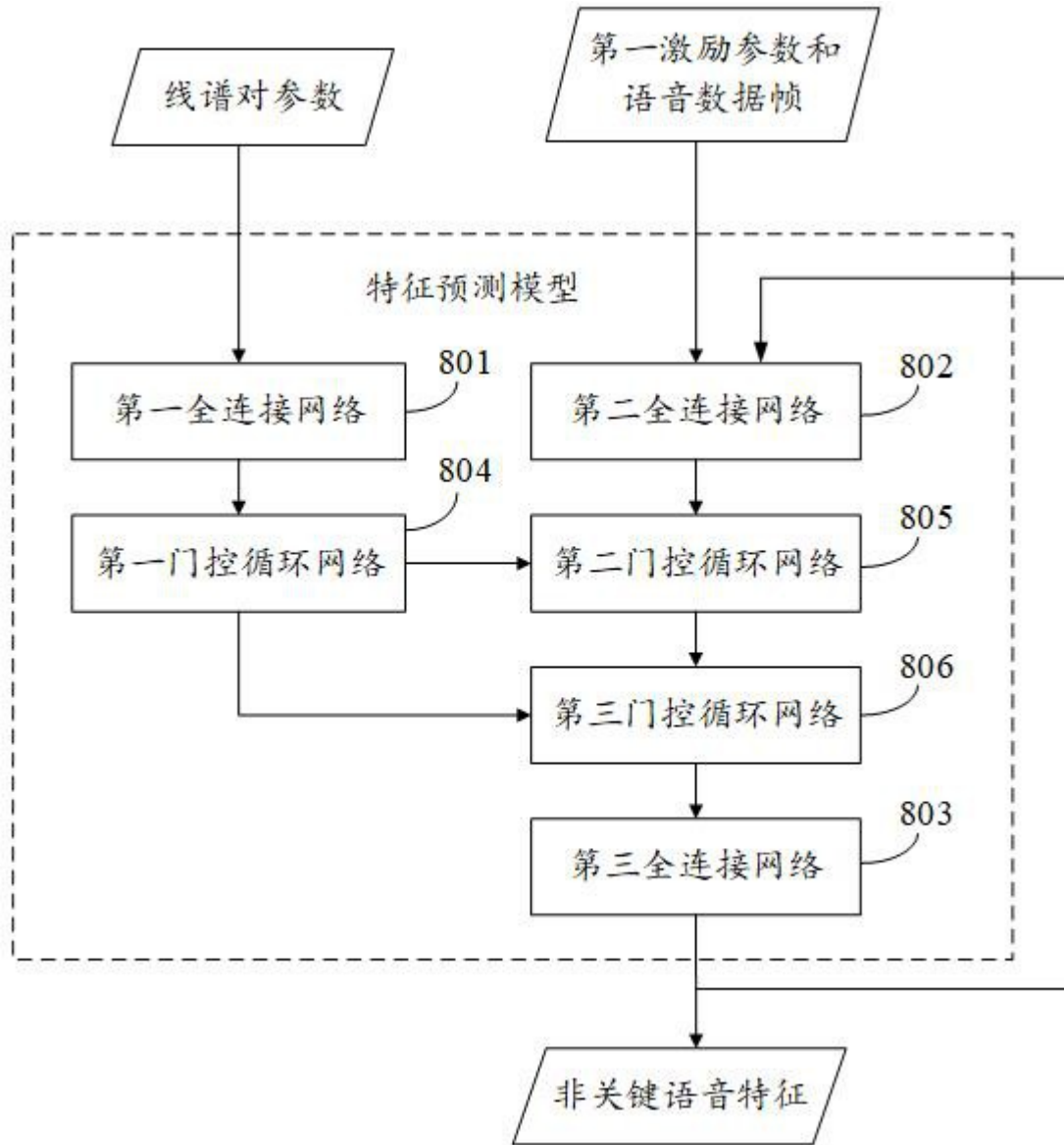


图8

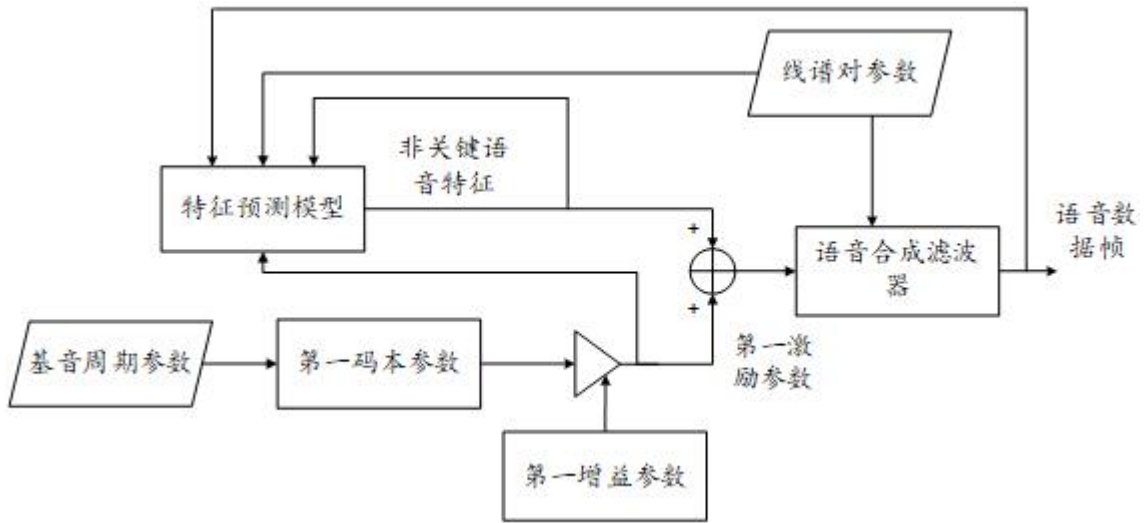


图9

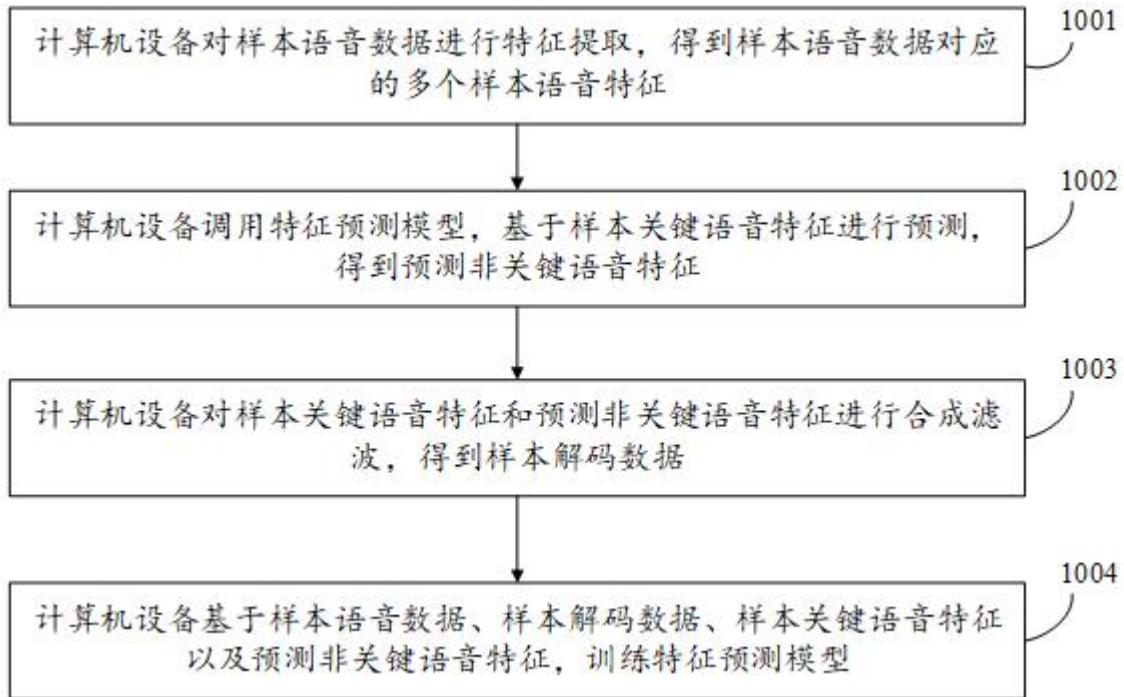


图10

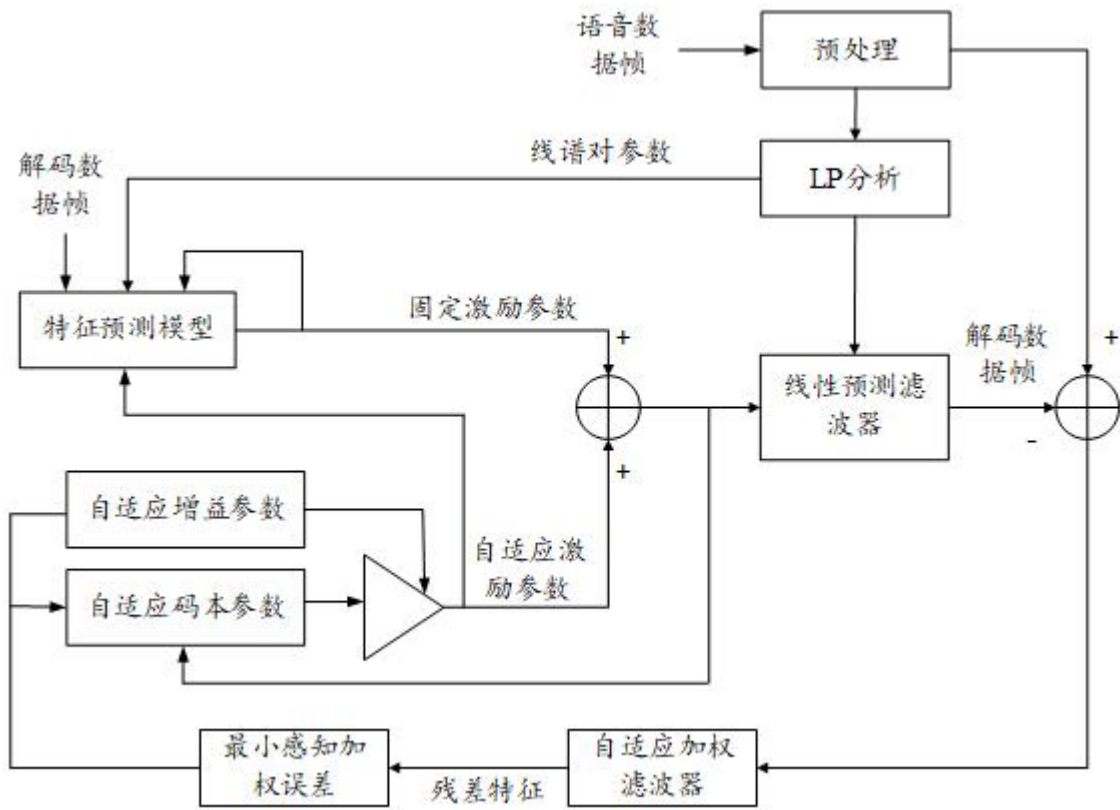


图11

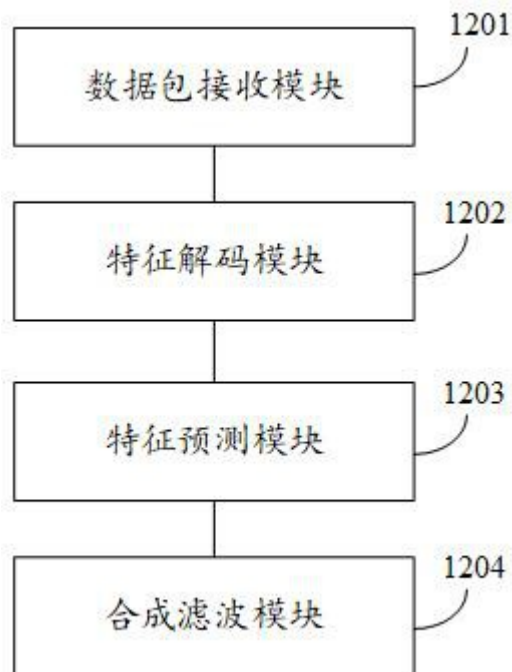


图12

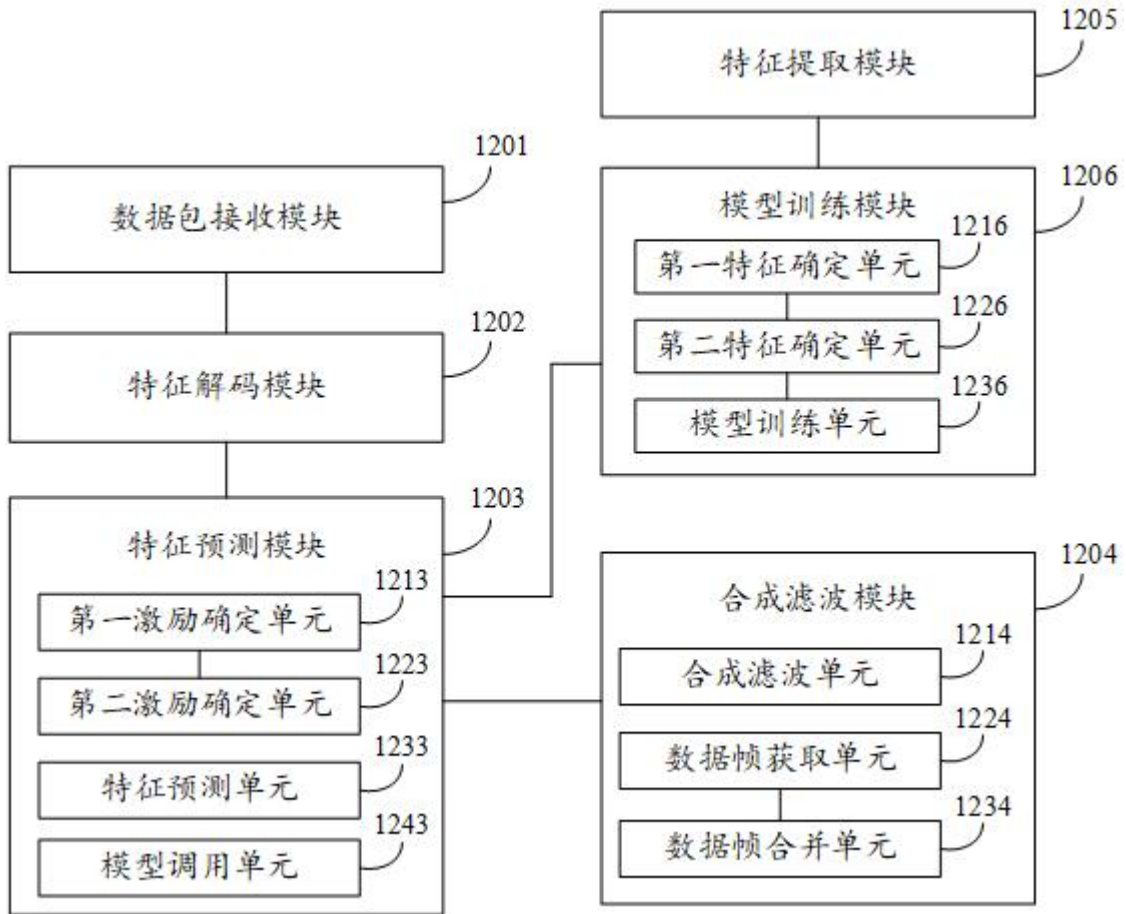


图13

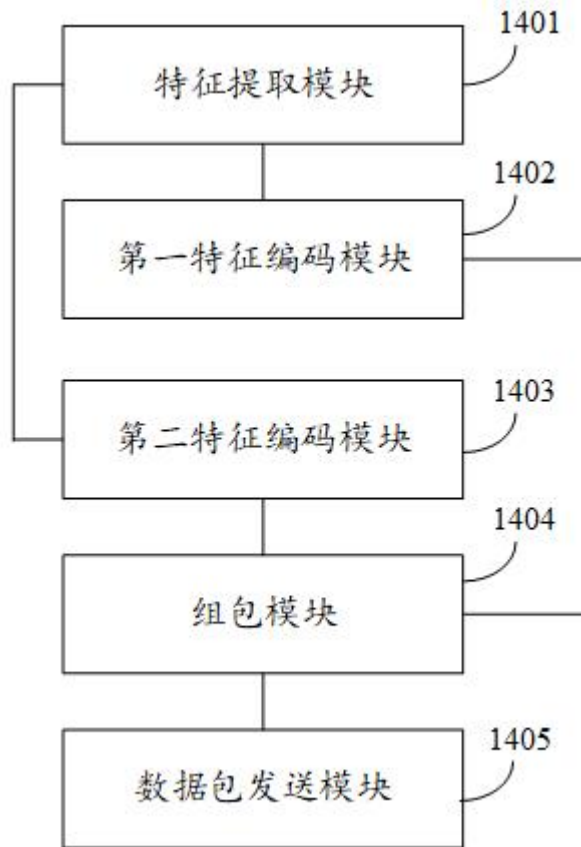


图14

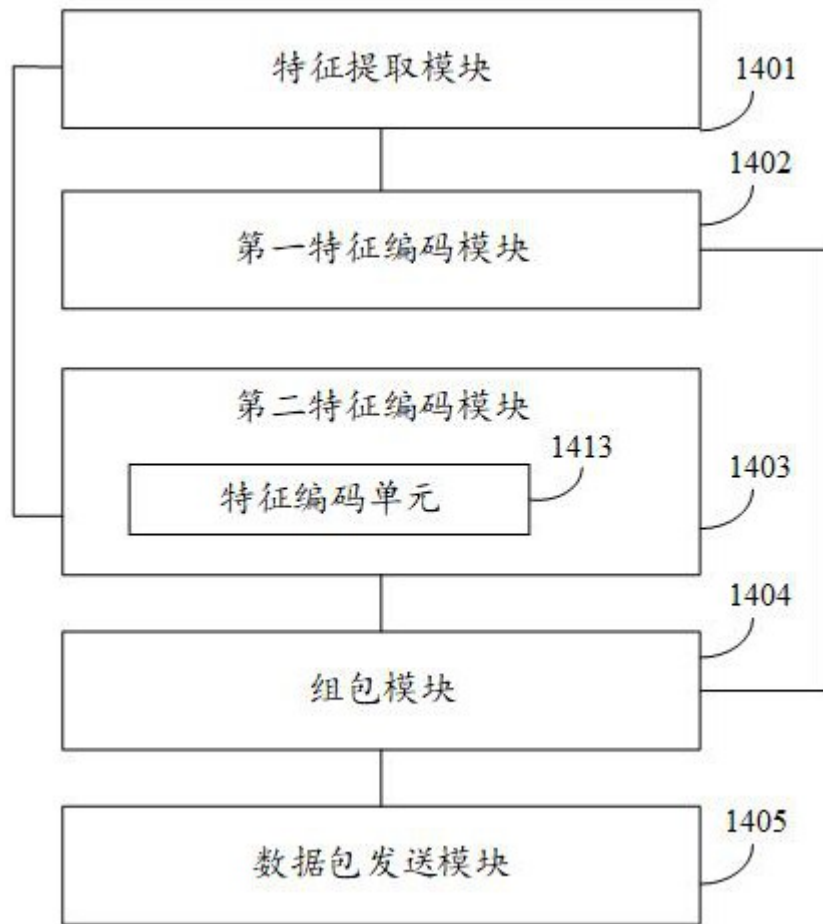


图15

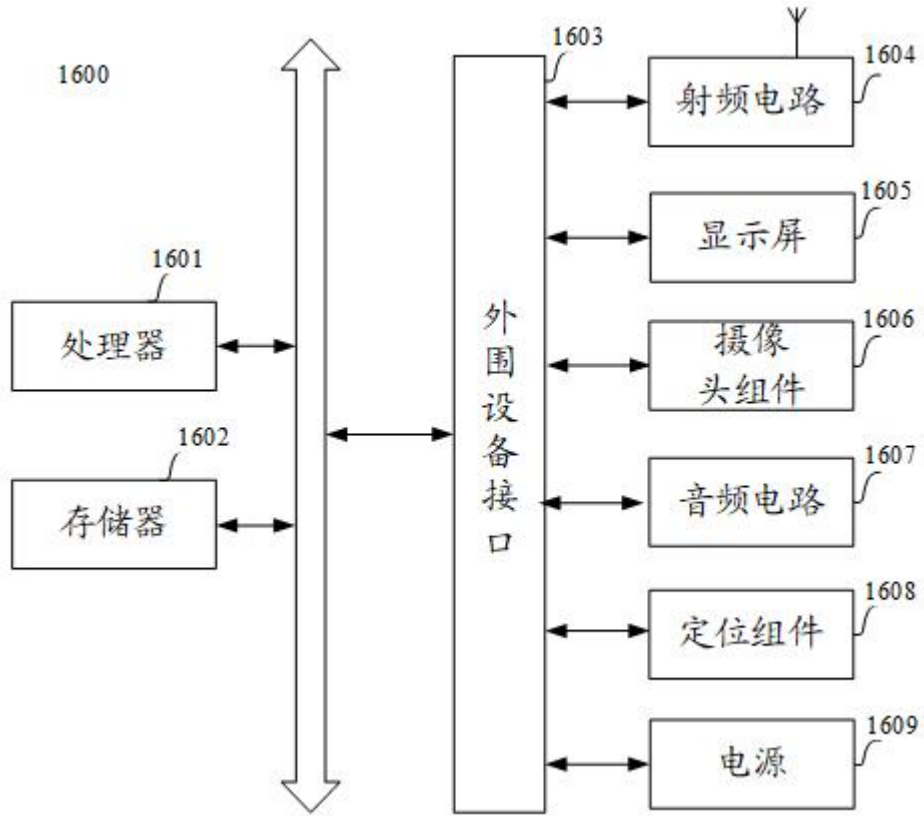


图16

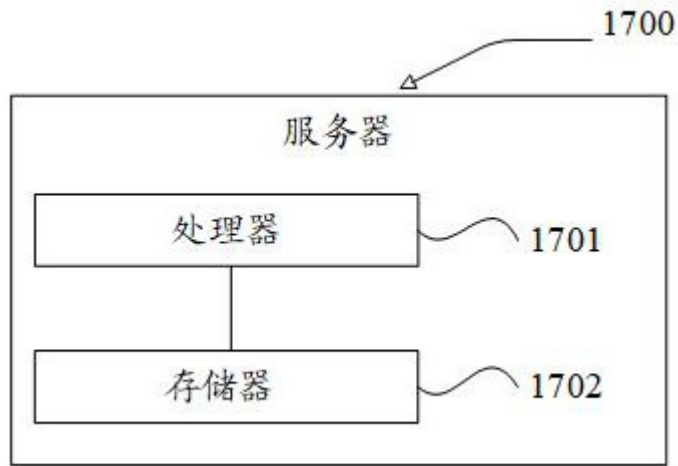


图17