



(12) 发明专利

(10) 授权公告号 CN 112634912 B

(45) 授权公告日 2024. 04. 09

(21) 申请号 202011502968.6

(22) 申请日 2020.12.18

(65) 同一申请的已公布的文献号
申请公布号 CN 112634912 A

(43) 申请公布日 2021.04.09

(73) 专利权人 北京猿力未来科技有限公司
地址 100102 北京市朝阳区广顺南大街8号
院1号楼6层F01-03、05-10单元

(72) 发明人 王晓红 陈佳路 王冬晓 杨明祺
马楠 高强 夏龙 郭常圳

(74) 专利代理机构 北京智信禾专利代理有限公司 11637

专利代理师 刘晓楠

(51) Int. Cl.

G10L 19/00 (2013.01)

G10L 25/30 (2013.01)

G10L 25/03 (2013.01)

G10L 25/24 (2013.01)

(56) 对比文件

CN 106251875 A, 2016.12.21

JP 2008111991 A, 2008.05.15

CN 109496333 A, 2019.03.19

JP 2008139661 A, 2008.06.19

KR 20160004462 A, 2016.01.13

US 2004184443 A1, 2004.09.23

CN 106788876 A, 2017.05.31

CN 111554308 A, 2020.08.18

CN 1901431 A, 2007.01.24

CN 101325631 A, 2008.12.17

CN 101894565 A, 2010.11.24

CN 108011686 A, 2018.05.08

CN 109963184 A, 2019.07.02

CN 111640442 A, 2020.09.08

Gabriel Mittag等.Non-intrusive

Estimation of Packet Loss Rates in Speech Communication Systems Using Convolutional Neural Networks.2018 IEEE International Symposium on Multimedia.2019,全文.

审查员 王静

权利要求书3页 说明书21页 附图7页

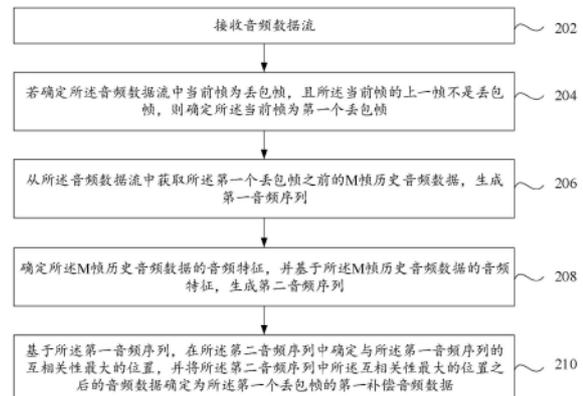
(54) 发明名称

丢包补偿方法及装置

(57) 摘要

本说明书提供丢包补偿方法及装置,其中所述丢包补偿方法包括:接收音频数据流,若确定音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;从音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音序列;确定M帧历史音频数据的音频特征,并基于M帧历史音频数据的音频特征,生成第二音序列;基于第一音序列,在第二音序列中确定与第一音序列的互相关性最大的位置,并将第二音序列中互相关性最大的位置之后的音频数据确定为第一个丢包帧的第一补偿音频数据。通过上述方法可以减少播放的音频数据机械且重复的问

题的出现。



1. 一种丢包补偿方法,其特征在于,所述方法包括:

接收音频数据流;

若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;

确定所述M帧历史音频数据的音频特征,并根据所述M帧历史音频数据的音频特征确定所述第一个丢包帧的音频数据的预测音频特征,将所述M帧历史音频数据的音频特征与所述第一个丢包帧的音频数据的预测音频特征输入音频合成模型中,获得第二音频序列,其中,所述第二音频序列中包括合成的所述M帧与所述第一个丢包帧的采样音频数据;

基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

2. 如权利要求1所述的丢包补偿方法,其特征在于,所述根据所述M帧历史音频数据的音频特征确定所述第一个丢包帧的音频数据的预测音频特征,包括:

将所述第一个丢包帧上一帧的历史音频数据的音频特征确定为所述第一个丢包帧的音频数据的预测音频特征。

3. 如权利要求1所述的丢包补偿方法,其特征在于,所述基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,包括:

基于所述第二音频序列的采样频率从所述第一音频序列中沿终点向起点方向获取N个连续的采样音频数据,得到第一采样音频数据组,其中,所述N为大于1的正整数;

以N为采样音频数据组的长度,逐次从所述第二音频序列中沿终点向起点方向获取N个连续的采样音频数据,得到第二采样音频数据组集合,其中,采样音频数据组的长度为采样音频数据组包括的采样音频数据的数量;

对每个得到的第二采样音频数据组,将所得到的第二采样音频数据组与所述第一采样音频数据组进行互相关性计算,以互相关性最大的第二采样音频数据组所在的位置确定所述互相关性最大的位置。

4. 如权利要求3所述的丢包补偿方法,其特征在于,所述以互相关性最大的第二采样音频数据组所在的位置确定所述互相关性最大的位置,包括:

将所述互相关性最大的第二采样音频数据组中距离终点最近的采样音频数据所在的位置确定为所述互相关性最大的位置。

5. 如权利要求1所述的丢包补偿方法,其特征在于,所述方法还包括:

从所述第一音频序列中沿终点向起点的方向获取X个未播放的采样音频数据,其中,所述X大于1且小于一帧采样音频数据的数量;

将所述第二音频序列中所述互相关性最大的位置之前的X个采样音频数据与所述X个未播放的采样音频数据进行融合,得到融合音频数据;

将所述第一音频序列中X个未播放的音频数据替换为所述融合音频数据,并将所述融合音频数据与所述第一个丢包帧的补偿音频数据进行拼接。

6. 如权利要求1所述的丢包补偿方法,其特征在于,所述方法还包括:

若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧为丢包帧,则确定所述当前帧为连续丢包帧,其中,所述连续丢包帧为多个连续的丢包帧中非第一个丢包帧的丢包帧;

基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述连续丢包帧的第二补偿音频数据。

7.如权利要求6所述的丢包补偿方法,其特征在于,所述基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述连续丢包帧的第二补偿音频数据,包括:

以所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征作为所述连续丢包帧的预测音频特征;

将所述连续丢包帧的预测音频特征输入音频合成模型中,得到所述连续丢包帧的第二补偿音频数据。

8.如权利要求1所述的丢包补偿方法,其特征在于,所述方法还包括:

若确定所述音频数据流中当前帧不是丢包帧,且所述当前帧的上一帧是丢包帧,则确定所述当前帧为目标帧;

获取所述目标帧的实际音频数据,并基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述目标帧的预测音频数据;

基于所述目标帧的实际音频数据,在所述目标帧的预测音频数据中确定与所述目标帧的实际音频数据互相关性最大的位置;

将所述互相关性最大的位置之后的Y个采样音频数据与所述目标帧的实际音频数据中沿起点到终点方向的Y个采样音频数据进行融合,得到所述目标帧融合后的音频数据,其中,所述Y为大于1的正整数。

9.如权利要求2或7所述的丢包补偿方法,其特征在于,所述方法还包括:

将样本音频特征和样本采样音频数据输入所述音频合成模型中,输出预测采样音频数据;

基于所述预测采样音频数据和所述样本采样音频数据对所述音频合成模型进行训练,直至达到训练停止条件。

10.一种丢包补偿装置,其特征在于,所述装置包括:

接收模块,被配置为接收音频数据流;

第一确定模块,被配置为若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

第一生成模块,被配置为从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音序列,其中,所述M为大于1的正整数;

第二生成模块,被配置为确定所述M帧历史音频数据的音频特征,并根据所述M帧历史音频数据的音频特征确定所述第一个丢包帧的音频数据的预测音频特征,将所述M帧历史音频数据的音频特征与所述第一个丢包帧的音频数据的预测音频特征输入音频合成模型中,获得第二音序列,其中,所述第二音序列中包括合成的所述M帧与所述第一个丢包帧的采样音频数据;

第二确定模块,被配置为基于所述第一音序列,在所述第二音序列中确定与所述第一音序列的互相关性最大的位置,并将所述第二音序列中所述互相关性最大的位置

之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

11. 一种计算设备,其特征在於,包括:

存储器和处理器;

所述存储器用于存储计算机可执行指令,所述处理器用于执行所述计算机可执行指令,以实现下述方法:

接收音频数据流;

若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;

确定所述M帧历史音频数据的音频特征,并根据所述M帧历史音频数据的音频特征确定所述第一个丢包帧的音频数据的预测音频特征,将所述M帧历史音频数据的音频特征与所述第一个丢包帧的音频数据的预测音频特征输入音频合成模型中,获得第二音频序列,其中,所述第二音频序列中包括合成的所述M帧与所述第一个丢包帧的采样音频数据;

基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

12. 一种计算机可读存储介质,其存储有计算机指令,该指令被处理器执行时实现权利要求1至9任意一项所述丢包补偿方法的步骤。

丢包补偿方法及装置

技术领域

[0001] 本说明书涉及音频处理技术领域,特别涉及丢包补偿方法及装置。

背景技术

[0002] 为在互联网上完成通话传输,通话中的数据通常被拆分为数据包。这些数据包通过网络从发送方传输到接收方,并在接收方重新组合,以生成连续的视频流和音频流。然而,当数据包到达接收方时,可能会出现个别数据包完全丢失的问题,进而导致通话质量变差,因此接收方需要补全丢失的数据包,以提高通话质量。处理丢包的过程被称为丢包补偿,接收方的PLC(Packet Loss Concealment,丢包补偿)模块负责创建音频(或视频),以填补由丢包造成的间隔。

[0003] 为了隐藏丢包带来的影响,现有技术通常直接将丢包帧的上一帧的历史音频数据作为丢包帧的补偿音频数据。但这种方法在连续多帧丢包的情况下,生成的每帧的补偿音频数据均与丢包帧之前最后一帧的历史音频数据相同,导致接收方播放出来的音频数据机械且重复。

发明内容

[0004] 有鉴于此,本说明书实施例提供了一种丢包补偿方法。本说明书同时涉及一种丢包补偿装置,一种计算设备,以及一种计算机可读存储介质,以解决现有技术中存在的技术缺陷。

[0005] 根据本说明书实施例的第一方面,提供了一种丢包补偿方法,包括:

[0006] 接收音频数据流;

[0007] 若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

[0008] 从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;

[0009] 确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;

[0010] 基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

[0011] 可选地,所述基于所述M帧历史音频数据的音频特征,生成第二音频序列,包括:

[0012] 将所述第一个丢包帧上一帧的历史音频数据的音频特征确定为所述第一个丢包帧的音频数据的预测音频特征;

[0013] 将所述M帧历史音频数据的音频特征与所述第一个丢包帧的音频数据的预测音频特征输入音频合成模型中,得到所述第二音频序列,其中,所述第二音频序列中包括合成的所述M帧与所述第一个丢包帧的采样音频数据。

[0014] 可选地,所述基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,包括:

[0015] 基于所述第二音频序列的采样频率从所述第一音频序列中沿终点向起点方向获取N个连续的采样音频数据,得到第一采样音频数据组,其中,所述N为大于1的正整数;

[0016] 以N为采样音频数据组的长度,逐次从所述第二音频序列中沿终点向起点方向获取N个连续的采样音频数据,得到第二采样音频数据组集合,其中,采样音频数据组的长度为采样音频数据组包括的采样音频数据的数量;

[0017] 对每个得到的第二采样音频数据组,将所得到的第二采样音频数据组与所述第一采样音频数据组进行互相关性计算,以互相关性最大的第二采样音频数据组所在的位置确定所述互相关性最大的位置。

[0018] 可选地,所述以互相关性最大的第二采样音频数据组所在的位置确定所述互相关性最大的位置,包括:

[0019] 将所述互相关性最大的第二采样音频数据组中距离终点最近的采样音频数据所在的位置确定为所述互相关性最大的位置。

[0020] 可选地,所述方法还包括:

[0021] 从所述第一音频序列中沿终点向起点的方向获取X个未播放的采样音频数据,其中,所述X大于1且小于一帧采样音频数据的数量;

[0022] 将所述第二音频序列中所述互相关性最大的位置之前的X个采样音频数据与所述X个未播放的采样音频数据进行融合,得到融合音频数据;

[0023] 将所述第一音频序列中X个未播放的音频数据替换为所述融合音频数据,并将所述融合音频数据与所述第一个丢包帧的补偿音频数据进行拼接。

[0024] 可选地,所述方法还包括:

[0025] 若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧为丢包帧,则确定所述当前帧为连续丢包帧,其中,所述连续丢包帧为多个连续的丢包帧中非第一个丢包帧的丢包帧;

[0026] 基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述连续丢包帧的第二补偿音频数据。

[0027] 可选地,所述基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述连续丢包帧的第二补偿音频数据,包括:

[0028] 以所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征作为所述连续丢包帧的预测音频特征;

[0029] 将所述连续丢包帧的预测音频特征输入音频合成模型中,得到所述连续丢包帧的第二补偿音频数据。

[0030] 可选地,所述方法还包括:

[0031] 若确定所述音频数据流中当前帧不是丢包帧,且所述当前帧的上一帧是丢包帧,则确定所述当前帧为目标帧;

[0032] 获取所述目标帧的实际音频数据,并基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述目标帧的预测音频数据;

[0033] 基于所述目标帧的实际音频数据,在所述目标帧的预测音频数据中确定与所述目

标帧的实际音频数据互相关性最大的位置；

[0034] 将所述互相关性最大的位置之后的Y个采样音频数据与所述目标帧的实际音频数据中沿起点到终点方向的Y个采样音频数据进行融合,得到所述目标帧融合后的音频数据,其中,所述Y为大于1的正整数。

[0035] 可选地,所述方法还包括:

[0036] 将样本音频特征和样本采样音频数据输入所述音频合成模型中,输出预测采样音频数据;

[0037] 基于所述预测采样音频数据和所述样本采样音频数据对所述音频合成模型进行训练,直至达到训练停止条件。

[0038] 根据本说明书实施例的第二方面,提供了一种丢包补偿装置,包括:

[0039] 接收模块,被配置为接收音频数据流;

[0040] 第一确定模块,被配置为若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

[0041] 第一生成模块,被配置为从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;

[0042] 第二生成模块,被配置为确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;

[0043] 第二确定模块,被配置为基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

[0044] 根据本说明书实施例的第三方面,提供了一种计算设备,包括:

[0045] 存储器和处理器;

[0046] 所述存储器用于存储计算机可执行指令,所述处理器用于执行所述计算机可执行指令:

[0047] 接收音频数据流;

[0048] 若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

[0049] 从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;

[0050] 确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;

[0051] 基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

[0052] 根据本说明书实施例的第四方面,提供了一种计算机可读存储介质,其存储有计算机可执行指令,该指令被处理器执行时实现所述丢包补偿方法的步骤。

[0053] 本说明书提供的丢包补偿方法,接收音频数据流,若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其

中,所述M为大于1的正整数;确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。上述方法通过第一个丢包帧的前M帧的历史音频数据以及音频特征生成第二音频序列,并将真实的第一音频序列和预测生成的第二音频序列进行比对,得到第二音频序列中与第一音频序列的互相关性最大的位置,进而确定第一个丢包帧的第一补偿音频数据,得到的补偿音频数据与上一帧的历史音频数据是不同的,因此,可以降低播放的音频数据机械且重复的问题的出现。

附图说明

- [0054] 图1是本说明书一实施例提供的一种NetEQ的结构示意图;
- [0055] 图2是本说明书一实施例提供的一种丢包补偿方法的流程图;
- [0056] 图3是本说明书一实施例提供的一种频带划分的示意图;
- [0057] 图4是本说明书一实施例提供的一种LPCNet模型的结构示意图;
- [0058] 图5是本说明书一实施例提供的一种第一音频序列和第二音频序列的示意图;
- [0059] 图6是本说明书一实施例提供的一种应用于音频传输的丢包补偿方法的处理流程图;
- [0060] 图7是本说明书一实施例提供的一种丢包补偿装置的结构示意图;
- [0061] 图8是本说明书一实施例提供的一种计算设备的结构框图。

具体实施方式

[0062] 在下面的描述中阐述了很多具体细节以便于充分理解本说明书。但是本说明书能够以很多不同于在此描述的其它方式来实施,本领域技术人员可以在不违背本说明书内涵的情况下做类似推广,因此本说明书不受下面公开的具体实施的限制。

[0063] 在本说明书一个或多个实施例中使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本说明书一个或多个实施例。在本说明书一个或多个实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本说明书一个或多个实施例中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0064] 应当理解,尽管在本说明书一个或多个实施例中可能采用术语第一、第二等来描述各种信息,但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如,在不脱离本说明书一个或多个实施例范围的情况下,第一也可以被称为第二,类似地,第二也可以被称为第一。取决于语境,如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0065] 首先,对本说明书一个或多个实施例涉及的名词术语进行解释。

[0066] 丢包补偿:在数据包丢失时产生一个相似的音频数据包或噪声包以替代丢失的数据包,基于音频的短时相似性,在丢包率较低的情况下提升播放效果。

[0067] WebRTC(Web Real-Time Communications,网络实时通信):是一项实时通讯技术,它允许网络应用或者站点,在不借助中间媒介的情况下,建立浏览器之间点对点(Peer-to-

Peer)的连接,实现视频流和(或)音频流或者其他任意数据的传输。

[0068] NetEQ:WebRTC中的模块,是一项专为IP(Internet Protocol Address,互联网协议地址)电信系统开发的高级语音质量处理技术,其能够在大幅提高语音质量的同时,有效地减少由于缓存带来的延时影响。

[0069] PLC组件:用于产生丢包帧的补偿音频数据以及实现补偿音频数据与实际音频数据之间的平滑切换。

[0070] LPCNet:是一个数字信号处理(DSP,Digital Signal Process)和神经网络(NN,Neural Network)巧妙结合的模型,可以应用于语音合成中vocoder(语音编码器)的工作,能够在普通的CPU(Central Processing Unit,中央处理器)上实时合成高质量语音。

[0071] 接下来,对本说明书实施例提供的丢包补偿方法的应用场景进行说明。

[0072] 在音频数据传输的过程中,音频数据通过网络从发送方传输到接收方的过程中,可能会出现音频数据丢失的问题,接收方则可以通过WebRTC的NetEQ模块中的PLC组件进行丢包补偿。

[0073] NetEQ包括MCU(Micro Controller Unit,微控制单元)模块和DSP模块,MCU模块主要负责做延时及抖动的计算统计,并生成对应的控制命令;DSP模块负责接收并根据MCU的控制命令进行对应的数据包处理。NetEQ的PLC组件对应图1中的DSP处理中的丢包(extend)和融合(merge)两个模块。但这种方式通常是通过重复包或噪声包来对丢包帧进行音频数据的补偿,在连续丢包情况下,针对多个丢包帧生成的补偿音频数据是相同的,导致播放的音频数据机械且重复,降低音频播放效果。

[0074] 为此,本说明书提供了一种丢包补偿方法,可以解决上述技术问题,其具体实现可以参见下述各个实施例。

[0075] 在本说明书中,提供了一种丢包补偿方法,本说明书同时涉及一种丢包补偿装置,一种计算设备,以及一种计算机可读存储介质,在下面的实施例中逐一进行详细说明。

[0076] 图2示出了根据本说明书一实施例提供的一种丢包补偿方法的流程图,具体包括以下步骤:

[0077] 步骤202,接收音频数据流。

[0078] 在实施中,音频数据流可以以数据包的形式实时发送至计算设备,则计算设备可以实时接收到音频数据流。

[0079] 步骤204,若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧。

[0080] 在实施中,计算设备实时接收到音频数据流后,可以实时确定接收到的音频数据流中当前帧的数据包是否存在,即确定当前帧是否为丢包帧。若确定当前帧是丢包帧,由于丢包帧需要进行丢包补偿以得到补偿音频数据,且补偿音频数据与真实音频数据之间可能存在无法平滑切换的问题,但补偿音频数据与补偿音频数据之间可以实现平滑切换,因此,为了判断后续是否需要解决平滑切换的问题,还需要判断当前帧的上一帧是否为丢包帧,若确定当前帧的上一帧不是丢包帧,则可以将当前帧确定为第一个丢包帧。

[0081] 作为一种示例,假设当前帧为音频数据流中的第10帧,只要该音频数据流中第9帧不是丢包帧,则可以将第10帧确定为第一个丢包帧。例如,若第9帧不是丢包帧,但第5帧和第6帧均为丢包帧,由于第5帧和第6帧与第10帧不相邻,即第5帧和第6帧不是与第10帧连续

的丢包帧,因此,可以将第10帧确定为第一个丢包帧。

[0082] 进一步地,确定当前帧为第一个丢包帧后,可以生成启用丢包补偿的命令,以便于后续执行丢包补偿的操作。

[0083] 在一些实施例中,本说明书实施例也使用NetEQ进行丢包补偿,但对NetEQ中的PLC组件进行替换,替换成音频合成模型。

[0084] 作为一种示例,参见图1,音频数据流可以以数据包的形式被发送至计算设备,计算设备接收到数据包之后,可以将数据包存储至NetEQ中MCU模块的抖动缓冲区。NetEQ的MCU模块和DSP模块可以共享内存,即MCU模块的抖动缓冲区中存储的数据包可以被DSP模块获取到。MCU模块负责命令决策,根据数据包缺失情况可以生成丢包命令并发送至DSP模块,DSP模块可以根据该丢包补偿命令对语音缓冲区(sync_buffer)的音频数据进行丢包补偿处理得到补偿音频数据,将补偿音频数据暂存至NetEQ算法缓冲区中,将NetEQ算法缓冲区中的补偿音频数据存储至语音缓冲区,声卡可以从语音缓冲区中获取补偿音频数据进行播放。另外,DSP模块对正常抵达的数据包通过解码器进行解码后,将得到的音频数据流存储至解码缓冲区,对解码缓冲区中的音频数据流进行DSP处理,将处理后的音频数据流暂存至NetEQ算法缓冲区,将NetEQ算法缓冲区中的音频数据流存储至语音缓冲区,声卡可以从语音缓冲区中获取音频数据进行播放。

[0085] 步骤206,从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列。

[0086] 其中,所述M为大于1的正整数。例如,M可以为8。

[0087] 在本申请实施例中,以使用Modified_LPCNet(调整后的LPCNet)合成预测音频数据为例。在一种可能的实现方式中,Modified_LPCNet可以以4帧音频数据作为一组输入确定音频特征,因此,M的取值可以为大于等于4且是4的倍数。

[0088] 需要说明的是,上述仅是以使用Modified_LPCNet合成预测音频数据为例对M的取值进行举例说明,在另一些实施例中,还可以采用其他方式确定M的取值范围,本申请实施例对此不作限定。

[0089] 作为一种示例,以NetEQ处理音频数据的处理单位10毫秒为例,DSP模块接收到启用丢包补偿命令后,对于第一个丢包帧,可以从sync_buffer_中获取该第一个丢包帧之前的M帧的历史音频数据,生成第一音频序列。

[0090] 其中,该第一音频序列中包括对前M帧的历史音频数据进行采样后得到的采样音频数据。例如,假设M取8,可以从语音缓冲区中取8*10毫秒的历史音频数据,构成第一音频序列。

[0091] 步骤208,确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列。

[0092] 其中,音频特征可以包括BFCC(Bark Frequency cepstrum coefficients,巴克频率倒谱系数)系数、基音周期和基因相关性。

[0093] 其中,第二音频序列是预测得到的音频序列。

[0094] 在一些实施例中,确定M帧历史音频数据的音频特征的具体实现可以包括:对M帧历史音频数据中的每帧历史音频数据进行傅里叶变换,得到每帧历史音频数据的频谱图。对于每帧历史音频数据的频谱图,均可以按照图3的方式将频谱图划分为频带,每帧历史音

频数据可以得到多个频带,并计算每个频带内的能量。然后取频谱的Log(对数)压缩动态范围得到倒谱,对倒谱进行DCT(Discrete Cosine Transformation,离散余弦变换)得到最终的BFCC。如此,可以得到每帧历史音频数据的BFCC系数。还可以根据采样频率确定基音周期(pitch period)和基音相关性(pitch correlation)。将确定的BFCC系数、基音周期和基音相关性确定为每帧音频数据的音频特征。

[0095] 例如,基音周期可以为10毫秒。

[0096] 在本说明书实施例中,NetEQ中的PLC组件被替换成音频合成模型。确定M帧历史音频数据的音频特征后,可以基于M帧历史音频数据的音频特征,通过音频合成模型,可以得到第二音频序列。

[0097] 在一些实施例中,基于所述M帧历史音频数据的音频特征,生成第二音频序列的具体实现可以包括:将所述第一个丢包帧上一帧的历史音频数据的音频特征确定为所述第一个丢包帧的音频数据的预测音频特征;将所述M帧历史音频数据的音频特征与所述第一个丢包帧的音频数据的预测音频特征输入音频合成模型中,得到所述第二音频序列。

[0098] 其中,所述第二音频序列中包括合成的所述M帧与所述第一个丢包帧的采样音频数据。

[0099] 也就是说,可以先以最邻近的一帧音频数据的音频特征作为第一个丢包帧的音频数据的预测音频特征。将该预测音频特征与M帧音频特征共同输入音频合成模型中,可以得到第二音频序列。

[0100] 其中,音频合成模型可以为Modified_LPCNet模型,即调整后的LPCNet模型。

[0101] 在一些实施例中,LPCNet模型可以包括两个子网络和一个计算模块,该两个子网络分别为帧频网络(Frame rate network)和采样率网络(Sample rate network),该一个计算模块可以为计算LPC的模块。Frame rate network主要为Sample rate network提供一个条件向量的输入,这个条件向量一帧计算一次,并在该帧时间内保持不变。LPC计算模块则从输入的音频特征中计算线性预测参数LPC,LPC也是一帧计算一次,并在帧内保持不变。

[0102] 作为一种示例,参见图4,图4是LPCNet模型的结构示意图。左侧虚线框为帧率网络,包括卷积层和全连接层,用来确定条件向量f。右侧虚线框为采样率网络,包括concat函数、稀疏矩阵GRU_A、门控循环单元GRU_B、双全连接层dual FC(Fully Connected)和激活函数层,用来生成音频数据。上方的虚线框为LPC计算模块,用来计算线性预测参数LPC。

[0103] 在本说明书实施例中,可以将Modified_LPCNet模型的卷积层调整为因果卷积层。并且,可以将Modified_LPCNet模型的采样率调整为与NetEQ的采样率相同,或者,可以根据实际情况对Modified_LPCNet模型的采样率进行调整。

[0104] 作为一种示例,调整Modified_LPCNet模型的采样率可以包括:调整频带划分,调整BFCC维数,调整FRAME_SIZE,调整LPC_ORDER,调整PITCH_MIN_PERIOD和PITCH_MAX_PERIOD等等。

[0105] 例如,参见下表,假设调整之前的采样率为16kHz,调整后的采样率为8kHz。下表为调整采样率之前的LPCNet模型的参数与调整采样率之后的Modified_LPCNet模型的参数的对比。

	原始 16kHz 采样率 LPCNet	8kHz 采样率 Modified_LPCNet	参数含义
	对应图 3 中 0~8000 的 18 个频带	对应图 3 中 0~4000 的 14 个频带	LPCNet 使用图 3 所示的频带划分
	BFCC 维数 18	14	BFCC 维数=频带的个数
	FRAME_SIZE 160	80	对应 10ms 的采样音频数据个数
[0106]	LPC_ORDER 16	12	音频采样率很高, 相邻音频数据之间有很强的相关性, 当前帧的音频数据可以近似由 LPC_ORDER 个相邻历史帧的音频数据线性表示
	PITCH_MIN_PERIOD 32	16	基音周期对应的采样音频数据的最小值
[0107]	PITCH_MAX_PERIOD 256	128	基音周期对应的采样音频数据的最大值

[0108] 在实施中, 音频合成模型的训练方法可以包括: 将样本音频特征和样本采样音频数据输入音频合成模型中, 输出预测采样音频数据; 基于所述预测采样音频数据和所述样本采样音频数据对所述音频合成模型进行训练, 直至达到训练停止条件。

[0109] 作为一种示例, 可以将样本音频源数据进行解码得到样本音频文件, 该样本音频

文件可以为wav格式(Waveform Audio File Format,波形音频文件格式)的文件,并将所有的样本音频文件重采样到工作采样率,得到采样样本音频文件,将采样样本音频文件拼接在一起构成样本输入文件,该样本输入文件可以为pcm格式的文件。在linux机器上的LPCNet对应的文件夹下,make(生成)训练数据的可执行文件。该可执行文件可按照实际项目需求进行修改,例如,静音帧是否要去掉、是否要调用双二阶滤波器、是否要加入随机噪声等等。调用该修改后的可执行文件,可以生成模型训练所需的样本音频特征和样本音频数据。然后将样本音频特征和样本音频数据输入音频合成模型中,音频合成模型会自动学习采样的非线性部分,并输出预测音频数据,基于预测音频数据和样本音频数据可以对模型进行迭代训练,直到达到迭代次数或者模型的准确率大于预设阈值,可以确定训练得到符合要求的音频合成模型,停止模型训练。

[0110] 在本说明书中,可以以第一个丢包帧上一帧的历史音频数据的音频特征确定为该第一个丢包帧的预测音频特征,将该预测音频特征和M帧预测音频特征输入上述音频合成模型中,可以输出第二音频序列,第二音频序列是预测得到的M+1帧的采样音频数据。

[0111] 例如,以M为8为例,假设第9帧为丢包帧,可以将第8帧的音频特征作为该第9帧的预测音频特征,将前8帧的音频特征和该预测音频特征输入音频合成模型中,可以得到第二音频序列,该第二音频序列包括9帧采样音频数据。参见图5,图5中第一音频序列包括8帧音频数据,第二音频序列包括9帧音频数据。假设采样频率为8KHz,每帧音频数据为10毫秒,则每帧音频数据包括的采样音频数据的个数为80个,该第一音频序列中包括640个采样音频数据,第二音频序列中包括720个采样音频数据。

[0112] 需要说明的是,上述是以上一帧的历史音频数据的音频特征为该第一个丢包帧的预测音频特征为例,在另一些实施例中,可以使用其他传统算法或者神经网络模型,通过前M帧历史音频数据的音频特征预测第一个丢包帧的预测音频特征,本说明书实施例对此不作限定。

[0113] 步骤210,基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

[0114] 在实施中,基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置的具体实现可以包括:基于所述第二音频序列的采样频率从所述第一音频序列中沿终点向起点方向获取N个连续的采样音频数据,得到第一采样音频数据组。以N为采样音频数据组的长度,逐次从所述第二音频序列中沿终点向起点方向获取N个连续的采样音频数据,得到第二采样音频数据组集合,其中,采样音频数据组的长度为采样音频数据组包括的采样音频数据的数量;对每个得到的第二采样音频数据组,将所得到的第二采样音频数据组与所述第一采样音频数据组进行互相关性计算,以互相关性最大的第二采样音频数据组所在的位置确定所述互相关性最大的位置。

[0115] 其中,所述N为大于1的正整数。

[0116] 在本申请实施例中,以采样率为4kHz为例,4kHz采样率下的一个基音周期内采样音频数据的个数为[8,64]。由于N值太小不利于做相关性计算,N值太大相关性计算的计算量也会相应增大。所以在基音周期的基础上,对N的取值会有最大值和最小值的限制,在4kHz采样率下,可以限制N的取值范围为[60,120]。

[0117] 需要说明的是,上述仅是以采样率为4kHz为例,对N的取值范围进行说明,在另一些实施例中,还可以以其他采样率或其他方式确定N的取值范围,本申请实施例对此不作限定。

[0118] 作为一种示例,可以基于第二音频序列的采样频率,将第一音频序列重采样(resample)至相同的采样频率,从重采样后的第一音频序列中沿终点至起点方向获取N个连续的采样音频数据,得到第一采样音频数据组。

[0119] 由于第二音频序列中包括通过合成的M帧采样音频数据和预测的第一个丢包帧的采样音频数据,第一音频序列中包括M帧采样音频数据,在播放音频的过程中,先播放第一音频序列,然后播放第二音频序列中预测的第一个丢包帧的采样音频数据,为了使得第一音频序列和第二音频序列之间的衔接更加平滑,可以确定第二音频序列中与第一音频序列的互相关性最大的位置。

[0120] 在一些实施例中,以互相关性最大的第二采样音频数据组所在的位置确定所述互相关性最大的位置的具体实现可以包括:将所述互相关性最大的第二采样音频数据组中距离终点最近的采样音频数据所在的位置确定为所述互相关性最大的位置。

[0121] 也就是说,可以将互相关性最大的第二音频数据组中最后一个采样音频数据所在的位置确定为该互相关性最大的位置。

[0122] 示例性地,以采样频率为8KHz,N为80为例,可以基于8KHz对第一音频序列进行采样,假设第一音频序列包括8帧音频数据,进行采样后,第一音频序列可以包括640个采样音频数据。可以在第一音频序列中沿终点向起点方向获取80个采样音频数据作为第一采样音频数据组,并以80为采样音频数据组的长度,从第二音频序列中沿终点向起点方向取80个连续的采样音频数据,得到第二采样音频数据组集合。对每个第二采样音频数据组,将得到的第二采样音频数据组与第一采样音频数据组进行互相关计算,确定与第一采样音频数据组的互相关性最大的第二采样音频数据组,称为目标采样音频组,可以将目标采样音频数据组中最后一个采样音频数据的位置确定为互相关性最大的位置。

[0123] 例如,参见图5,图5中,A为第一音频序列中的第一采样音频数据组, B_1 、 B_2 、 B_3 、 B_j 均为第二音频序列中的第二采样音频数据组,且 B_2 是将第二采样音频数据组 B_1 左移1个采样音频数据对应的音频采样点后得到的第二采样音频数据组,以此类推,可以一直将第二采样音频数据组 B_1 左移,直至左移j-1个采样音频数据后可以得到第二采样音频数据组 B_j ,如此能够得到第二采样音频数据组集合。其中,j大于1且小于P帧对应的采样点数。将A分别与 B_1 、 B_2 、 B_3 、.....、 B_j 进行互相关计算,假设采样音频数据组的长度N为80,则是将A包括的80个采样音频数据分别与 B_1 、 B_2 、 B_3 、.....、 B_j 包括的80个采样音频数据进行相关性计算。具体的,将A包括的80个采样音频数据与 B_1 包括的80个采样音频数据进行组对,可以得到80个采样音频数据对,将每对采样音频数据的值相乘,得到80个乘积,将80个乘积相加,可以得到该第一采样音频数据组A与第二采样音频数据组 B_1 的互相关度,同理,可以得到第一采样音频数据组A分别与第二采样音频数据组 B_2 、 B_3 、.....、 B_j 的互相关度,即可以得到j个互相关度,确定j个互相关度中最大的互相关度对应的目标第二采样音频数据组,假设目标第二采样音频数据组为 B_3 ,则将 B_3 中最后一个采样音频数据所在的位置确定为互相关性最大的位置。

[0124] 其中,P可以为大于1且小于M的正整数。

[0125] 作为一种示例,为了提高确定互相关性最大的位置的概率,可以在第二音频序列中获取很多第二采样音频数据组,但如此会增加很多计算量,为了减少计算量,则可减少获取的第二采样音频数据组的数量。因此,根据大数据或经验可以得到,P比较优选的取值可以为2,以1帧为10ms为例,即j大于1且小于20ms对应的采样点数。

[0126] 需要说明的是,上述是以将每对采样音频数据的值相乘确定互相关度为例说明确定互相关性的方法,在另一些实施例中,还可以对采样音频数据进行降采样,例如,以4kHz对采样音频数据进行重采样,然后从重采样后的音频数据中确定互相关性最大的位置,即做粗粒度的互相关性最大的位置的搜索,再在确定的互相关性最大的位置附近进行按照原本的采样率进行采样,即做细粒度的互相关性最大的位置的搜索,进而确定互相关性最大的位置。另外,还可以使用最小失真算法进行辅助以确定互相关性最大的位置,本说明书实施例对此不作限定。

[0127] 在一些实施例中,确定互相关性最大的位置后,可以认为该位置之前的采样音频数据是与音频数据流中原有的音频数据最相似的,则该位置之后的采样音频数据为第一个丢包帧的补偿音频数据。

[0128] 进一步地,确定互相关性最大的位置后,由于音频播放的顺序是先播放第一音频序列,再播放第一个丢包帧的补偿音频数据,为了使得第一音频序列和补偿音频数据之间的切换更加平滑,可以将第一音频序列和第二音频序列进行融合。其具体实现方法可以包括:从所述第一音频序列中沿终点向起点的方向获取X个未播放的采样音频数据;将所述第二音频序列中所述互相关性最大的位置之前的X个采样音频数据与所述X个未播放的采样音频数据进行融合,得到融合音频数据,将所述第一音频序列中X个未播放的音频数据替换为所述融合音频数据;将所述融合音频数据与所述第一个丢包帧的补偿音频数据进行拼接。

[0129] 其中,所述X大于1且小于一帧采样音频数据的数量。

[0130] 也就是说,将第一音频序列中从后往前X个未播放的采样音频数据与第二音频序列中互相关性最大的位置之前的X个采样音频数据进行融合,则该融合音频数据中比较靠前的音频数据与第一音频序列中的历史音频数据衔接比较平滑,因此,可以使用该融合音频数据将第一音频序列中X个未播放的音频数据替换掉,即将融合音频数据中靠前的数据与第一音频序列中的历史音频数据进行拼接,如此,播放音频数据时,在播放至该X个未播放的采样音频数据的前一个音频数据时,可以继续播放该融合音频数据。该融合音频数据中比较靠后的音频数据与第一个丢包帧的补偿音频数据衔接比较平滑,因此,可以将融合音频数据与第一个丢包帧的补偿音频数据进行拼接。如此,可以实现播放音频数据时,在播放融合音频数据中最后一个音频数据时,可以继续播放该第一个丢包帧的补偿音频数据。

[0131] 本申请实施例中,通过将第一音频序列中未播放的采样音频数据、第二音频序列中互相关性最大的位置之前的采样音频数据、以及第一丢包帧的补偿音频数据进行融合处理,可以实现第一音频序列与第一个丢包帧的补偿音频数据之间的平滑切换,即实现接收的真实音频数据与补偿得到的补偿音频数据之间的平滑切换,增强音频数据的播放效果。

[0132] 例如,以X为100为例,获取音频数据缓冲区中未播放的100个采样音频数据线性递减,第二音频序列中所述互相关性最大的位置之前的100个采样音频数据线性递增,再将对应的位置的采样音频数据相加,可以得到100个融合音频数据。将第二音频序列中互相关性

最大的位置之前的100个采样音频数据替换为该100个融合数据,假设第二音频序列中互相关性最大的位置之后有160个采样音频数据,即第一个丢包帧的补偿音频数据的数量为160个,可以将该100个融合音频数据与该160个补偿音频数据进行拼接,得到拼接后的音频数据,后续可以按照拼接后的音频数据进行音频播放。如此,可以实现第一音频序列和补偿音频数据之间的平滑切换。

[0133] 进一步地,上述是对当前帧为第一个丢包帧时进行丢包补偿的过程进行说明,若当前帧不是第一个丢包帧,需要采用下述方式进行丢包补偿。

[0134] 在实施中,若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧为丢包帧,则确定所述当前帧为连续丢包帧,其中,所述连续丢包帧为多个连续的丢包帧中非第一个丢包帧的丢包帧。基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述连续丢包帧的第二补偿音频数据。

[0135] 作为一种示例,为了便于描述,可以将多个连续的丢包帧中的非第一个丢包帧称为连续丢包帧,将连续丢包帧的补偿音频数据均称为第二补偿音频数据。例如,假设音频数据流中当前帧为第10帧,且第9帧和第8帧均为丢包帧,则第8帧为多个连续的丢包帧中的第一个丢包帧,则可以将第10帧称为连续丢包帧。

[0136] 作为一种示例,当前帧之前最后一个非丢包帧即是第一个丢包帧的上一帧。例如,假设音频数据流中当前帧为第10帧,且第9帧和第8帧均为丢包帧,前7帧均不是丢包帧,可以确定第7帧为当前帧之前最后一个非丢包帧,或者,可以确定第一个丢包帧是第8帧,则第一个丢包帧的上一帧为第7帧。

[0137] 也就是说,若在当前帧之前存在与当前帧连续的丢包帧,则说明已经进行过丢包补偿处理,不需要再确定互相关性最大的位置,也不需要进行音频数据的拼接,可以基于当前帧之前的最后一个非丢包帧的历史音频数据的音频特征,确定该连续丢包帧的第二补偿音频数据。

[0138] 作为一种示例,确定当前帧为连续丢包帧后,NetEQ模块会生成连续丢包补偿命令,DSP模块则接收到该连续丢包补偿命令,并基于第一个丢包帧的上一帧的历史音频数据的音频特征,确定该当前帧的第二补偿音频数据。继续上述举例,可以基于第7帧的历史音频数据的音频特征,确定连续丢包帧即第10帧的第二补偿音频数据。

[0139] 在一些实施例中,基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述连续丢包帧的第二补偿音频数据的具体实现可以包括:以所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征作为所述连续丢包帧的预测音频特征;将所述连续丢包帧的预测音频特征输入音频合成模型中,得到所述连续丢包帧的第二补偿音频数据。

[0140] 也就是说,可以持续以当前帧之前最后一个非丢包帧的历史音频数据的音频特征作为连续丢包帧的预测音频特征,使用Modified_LPCNet模型来合成新的补偿音频数据,即第二补偿音频数据。

[0141] 需要说明的是,丢包补偿的时间越长,生成的补偿音频数据与实际音频数据的相似性越差,所以在出现连续丢包的情况下,需要对生成的补偿音频数据的音量进行衰减,譬如每10毫秒音量衰减到原来的0.9倍。如此,可以在听觉上降低补偿音频数据与实际音频数据的差异。

[0142] 值得注意的是,在进行丢包补偿的过程中,可以将得到的第一补偿音频数据和第二补偿音频数据均存储在补偿缓冲区中,为每帧对应的补偿音频数据设置平滑系数,并且根据执行丢包补偿操作的次数,对平滑系数进行衰减。示例性地,平滑系数可以称为mute_slope,其初始值取决于第一个丢包帧之前的两个基音周期的能量,设第一个丢包帧之前的第一个基音周期为vector1,第二个基音周期为vector2, $ratio = \text{vector1的能量} / \text{vector2的能量}$,该ratio为两个基音周期的能量之间的比例,可以用于表明丢包之前能量变化的趋势。譬如 $ratio > 1.5$,就表示音频是onset(开始),可以根据ratio来设定平滑系数的几个档次。根据连续进行丢包补偿的次数,来衰减平滑系数的值。

[0143] 通过本申请实施例的方法,对第一个丢包帧和连续丢包帧采用不同的方法确定第一补偿音频数据和第二补偿音频数据,且由于音频合成模型本身在输出音频数据时的随机特性,连续丢包帧中每个丢包帧得到的第二补偿音频数据也是不同,因此,可以得到每个丢包帧不同的补偿音频数据,在连续丢包的情况下,生成的补偿音频数据不同,可以有效解决现有技术中生成的音频机械且重复的问题。

[0144] 进一步地,上述是对当前帧为丢包帧时进行丢包补偿的过程进行说明,若当前帧不是丢包帧,但当前帧的上一帧是丢包帧,需要将当前帧的上一帧的补偿音频数据和当前帧的实际音频数据进行拼接,以使得音频数据之间可以平滑切换。

[0145] 在实施中,若确定所述音频数据流中当前帧不是丢包帧,且所述当前帧的上一帧是丢包帧,则确定所述当前帧为目标帧;获取所述目标帧的实际音频数据,并基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述目标帧的预测音频数据;基于所述目标帧的实际音频数据,在所述目标帧的预测音频数据中确定与所述目标帧的实际音频数据互相关性最大的位置;将所述互相关性最大的位置之后的Y个采样音频数据与所述目标帧的实际音频数据中沿起点到终点方向的Y个采样音频数据进行融合,得到所述目标帧融合后的音频数据。

[0146] 其中,所述Y为大于1的正整数。

[0147] 作为一种示例,以采样率为8kHz为例,Y的取值可以为 $\min(60, 202 - \text{互相关性最大的位置})$ 。其中,202是通过大量的实验得到的经验值。例如,假设互相关性最大的位置为第100个采样音频数据所在的位置,因此,Y的取值为 $\min(60, 102) = 60$ 。

[0148] 也就是说,若当前帧不是丢包帧,则说明丢包已经结束,为了便于与当前帧为丢包帧时区分开来,在当前帧不是丢包帧的情况下将当前帧称为目标帧。若接收到当前帧的实际音频数据,说明已经接收到了最后一个丢包帧的下一帧的实际音频数据,由于该最后一个丢包帧的音频数据是丢包补偿得到的,与该目标帧的实际音频数据可能衔接不上,因此,可以根据当前帧之前最后一个非丢包帧的历史音频数据的音频特征,合成该目标帧的预测音频数据,并将预测音频数据与实际音频数据进行融合,可以确定目标帧融合后的音频数据,该融合后的音频数据中靠前的音频数据与最后一个丢包帧的补偿音频数据之间的衔接比较平滑,该融合后的音频数据中靠后的音频数据与目标帧的实际音频数据之间的衔接比较平滑,可以实现丢包帧的补偿音频数据和目标帧的实际音频数据之间的平滑切换。

[0149] 本说明书提供的丢包补偿方法,接收音频数据流,若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其

中,所述M为大于1的正整数;确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。上述方法通过第一个丢包帧的前M帧的历史音频数据以及音频特征生成第二音频序列,并将真实的第一音频序列和预测生成的第二音频序列进行比对,得到第二音频序列中与第一音频序列的互相关性最大的位置,进而确定第一个丢包帧的第一补偿音频数据,得到的补偿音频数据与上一帧的历史音频数据是不同的,因此,可以减少播放的音频数据机械且重复的问题的出现。

[0150] 下述结合附图6,以本说明书提供的丢包补偿方法在音频传输中的应用为例,对所述丢包补偿方法进行进一步说明。其中,图6示出了本说明书一实施例提供的一种应用于音频传输中的丢包补偿方法的处理流程图,具体包括以下步骤:

[0151] 步骤602,接收音频数据流。

[0152] 步骤604,若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧。

[0153] 例如,假设当前帧为第10帧,第9帧不是丢包帧,但第5帧和第6帧均为丢包帧,由于第5帧和第6帧与第10帧不相邻,即第5帧和第6帧不是与第10帧连续的丢包帧,因此,可以将第10帧确定为第一个丢包帧。

[0154] 步骤606,从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列。

[0155] 其中,M为大于1的正整数。

[0156] 现有技术中的丢包补偿通常是通过NetEQ的PLC组件来实现的,在本说明书实施例中,可以对NetEQ中的PLC组件进行替换。

[0157] 例如,参见图1,音频数据流可以以数据包的形式被发送至计算设备,计算设备接收到数据包之后,可以将数据包存储至NetEQ中MCU模块的抖动缓冲区,并对数据包进行计算分析,做出命令决策。若分析确定数据包中缺失当前帧的音频数据,可以生成丢包补偿命令发送至DSP模块,DSP模块可以基于该丢包补偿命令对语音缓冲区中的音频数据进行处理,得到补偿音频数据。

[0158] 作为一种示例,以NetEQ处理音频数据的处理单位10毫秒为例,DSP模块接收到启用丢包补偿命令后,对于第一个丢包帧,可以从语音缓冲区中获取该第一个丢包帧之前的M帧的历史音频数据,生成第一音频序列。

[0159] 需要说明的是,步骤606的具体实现可以参见步骤206的相关描述,本实施例在此不再赘述。

[0160] 步骤608,确定所述M帧历史音频数据的音频特征。

[0161] 例如,确定M帧历史音频数据的音频特征时,可以对M帧历史音频数据中的每帧历史音频数据进行傅里叶变换,得到每帧历史音频数据的频谱图。对于每帧历史音频数据的频谱图,均可以按照图4的方式对频谱图进行划分,每帧历史音频数据得到多个频带,并计算每个频带内的能量,Log压缩动态范围得到倒谱,DCT去相关得到最终的BFCC。如此,可以得到每帧历史音频数据的BFCC系数。还可以根据采样频率确定基音周期(pitch period)和基音相关性(pitch correlation)。将确定的BFCC系数、基音周期和基音相关性确定为每帧

音频数据的音频特征。

[0162] 步骤610,将所述第一个丢包帧上一帧的历史音频数据的音频特征确定为所述第一个丢包帧的音频数据的预测音频特征。

[0163] 步骤612,将所述M帧历史音频数据的音频特征与所述第一个丢包帧的音频数据的预测音频特征输入音频合成模型中,得到所述第二音序列。

[0164] 其中,所述第二音序列中包括合成的所述M帧与所述第一个丢包帧的采样音频数据。

[0165] 需要说明的是,上述步骤608-步骤612是对步骤208的下位描述,具体实现可以参见步骤208的相关描述,本实施例在此不再赘述。

[0166] 步骤614,基于所述第二音序列的采样频率从所述第一音序列中沿终点向起点方向获取N个连续的采样音频数据,得到第一采样音频数据组。

[0167] 在本申请实施例中,以采样率为4kHz为例,4kHz采样率下的一个基音周期内采样音频数据的个数为[8,64]。由于N值太小不利于做相关性计算,N值太大相关性计算的计算量也会相应增大。所以在基音周期的基础上,对N的取值会有最大值和最小值的限制,在4kHz采样率下,可以限制N的取值范围为[60,120]。

[0168] 步骤616,以N为采样音频数据组的长度,逐次从所述第二音序列中沿终点向起点方向获取N个连续的采样音频数据,得到第二采样音频数据组集合,其中,采样音频数据组的长度为采样音频数据组包括的采样音频数据的数量。

[0169] 步骤618,对每个得到的第二采样音频数据组,将所得到的第二采样音频数据组与所述第一采样音频数据组进行互相关性计算,将所述互相关性最大的第二采样音频数据组中距离终点最近的采样音频数据所在的位置确定为所述互相关性最大的位置。

[0170] 示例性地,以采样频率为8KHz,N为80为例,可以基于8KHz对第一音序列进行采样,假设第一音序列包括8帧音频数据,进行采样后,第一音序列可以包括640个采样音频数据。可以在第一音序列中沿终点向起点方向获取80个采样音频数据作为第一采样音频数据组,并以80为采样音频数据组的长度,从第二音序列中沿终点向起点方向取80个连续的采样音频数据,得到第二采样音频数据组集合。对每个第二采样音频数据组,将得到的第二采样音频数据组与第一采样音频数据组进行互相关计算,基于互相关性最大的第二采样音频数据组所在的位置确定互相关性最大的位置。例如,可以将第二采样音频数据组中最后一个采样音频数据的位置确定为互相关性最大的位置。

[0171] 示例性地,进行互相关计算时,可以将第一采样音频数据组中的80个采样音频数据与第二采样音频数据组中的80个采样音频数据进行组对,得到80个采样音频数据对,将每对采样音频数据的值相乘,得到80个乘积,将80个乘积相加,可以得到该第一采样音频数据组与第二采样音频数据组的互相关度。

[0172] 步骤620,将所述第二音序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

[0173] 在本申请实施例中,互相关性最大的位置是第一采样音频数据组中的采样音频数据和第二采样音频数据组的采样音频数据互相关性最大的位置,因此,第二音序列中互相关性最大的位置之前的音频数据可以认为是M帧音频数据,则互相关性最大的位置之后的音频数据可以认为是模型合成的第一个丢包帧的补偿音频数据。

[0174] 需要说明的是,上述步骤614-步骤620是对步骤210的下位描述,具体实现可以参见步骤206的相关描述,本实施例在此不再赘述。

[0175] 步骤622,从所述第一音频序列中沿终点向起点的方向获取X个未播放的采样音频数据。

[0176] 其中,所述X小于一帧采样音频数据的数量。

[0177] 为了实现第一音频序列中的音频数据和补偿的第一个丢包帧的补偿音频数据在播放时的平滑切换,可以从音频数据缓冲区中获取所述第一音频序列中未播放的X个采样音频数据。

[0178] 步骤624,将所述第二音频序列中所述互相关性最大的位置之前的X个采样音频数据与所述X个未播放的采样音频数据进行融合,得到融合音频数据。

[0179] 如此得到的融合数据中比较靠前的音频数据与第一音频序列中的音频数据衔接比较平滑,该融合数据中比较靠后的音频数据与第一个丢包帧的补偿音频数据衔接比较平滑。

[0180] 步骤626,将所述第一音频序列中X个未播放的音频数据替换为所述融合音频数据,并将所述融合音频数据与所述第一个丢包帧的补偿音频数据进行拼接。

[0181] 例如,以X为100为例,获取音频数据缓冲区中未播放的100个采样音频数据线性递减,第二音频序列中所述互相关性最大的位置之前的100个采样音频数据线性递增,再将对应的位置的采样音频数据相加,可以得到100个融合音频数据。将第二音频序列中互相关性最大的位置之前的100个采样音频数据替换为该100个融合数据,假设第二音频序列中互相关性最大的位置之后有160个采样音频数据,即第一个丢包帧的补偿音频数据的数量为160个,可以将该100个融合音频数据与该160个补偿音频数据进行拼接,得到拼接后的音频数据,后续可以按照拼接后的音频数据进行音频播放。如此,可以实现第一音频序列和补偿音频数据之间的平滑切换。

[0182] 需要说明的是,上述步骤622-步骤626是对步骤210中将第一音频序列和第二音频序列进行融合的下位描述,具体实现可以参见步骤210的相关描述,本实施例在此不再赘述。

[0183] 步骤628,若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧为丢包帧,则确定所述当前帧为连续丢包帧,其中,所述连续丢包帧为多个连续的丢包帧中非第一个丢包帧的丢包帧。

[0184] 步骤630,以所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征作为所述连续丢包帧的预测音频特征。

[0185] 步骤632,将所述连续丢包帧的预测音频特征输入音频合成模型中,得到所述连续丢包帧的第二补偿音频数据。

[0186] 也即是说,可以持续以第一个丢包帧上一帧的历史音频数据的音频特征作为非第一个丢包帧的预测特征,使用Modified_LPCNet模型来合成新的补偿音频数据,即第二补偿音频数据。

[0187] 需要说明的是,补偿的时间越长,生成的补偿音频数据数据与实际音频数据的相似性越差,所以连续丢包时,需要对生成的补偿音频数据的音量进行衰减,譬如每10毫秒音量衰减到原来的0.9倍。

[0188] 通过本申请实施例的方法,对第一个丢包帧和连续丢包帧采用不同的方法确定第一补偿音频数据和第二补偿音频数据,且由于音频合成模型本身在输出音频数据时的随机特性,连续丢包帧中每个丢包帧得到的第二补偿音频数据也是不同,因此,可以得到每个丢包帧不同的补偿音频数据,在连续丢包的情况下,生成的补偿音频数据不同,可以有效解决现有技术中生成的音频机械且重复的问题。

[0189] 需要说明的是,步骤628-步骤632是对步骤210中在当前帧不是丢包帧的情况下进行丢包补偿的过程的下位描述,本实施例在此不再赘述。

[0190] 步骤634,若确定所述音频数据流中当前帧不是丢包帧,且所述当前帧的上一帧是丢包帧,则确定所述当前帧为目标帧。

[0191] 步骤636,获取所述目标帧的实际音频数据,并基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述目标帧的预测音频数据。

[0192] 也就是说,若丢包已经结束,当前已经接收到最后一个丢包帧的下一帧的实际音频数据,由于该最后一个丢包帧的音频数据是丢包补偿得到的,与目标帧的实际音频数据可能衔接不上,因此,可以继续确定目标帧的预测音频数据。

[0193] 步骤638,基于所述目标帧的实际音频数据,在所述目标帧的预测音频数据中确定与所述目标帧的实际音频数据互相关性最大的位置。

[0194] 步骤640,将所述互相关性最大的位置之前的Y个采样音频数据与所述目标帧的实际音频数据中沿起点到终点方向的Y个采样音频数据进行融合,得到所述目标帧融合后的音频数据。

[0195] 其中,所述Y为大于1的正整数。

[0196] 该融合后的音频数据中靠前的音频数据与最后一个丢包帧的补偿音频数据之间的衔接比较平滑,该融合后的音频数据中靠后的音频数据与目标帧的实际音频数据之间的衔接比较平滑,如此,可以实现丢包帧的补偿音频数据和目标帧的实际音频数据之间的平滑切换。

[0197] 本说明书提供的丢包补偿方法,接收音频数据流,若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。上述方法通过第一个丢包帧的前M帧的历史音频数据以及音频特征生成第二音频序列,并将真实的第一音频序列和预测生成的第二音频序列进行比对,得到第二音频序列中与第一音频序列的互相关性最大的位置,进而确定第一个丢包帧的第一补偿音频数据,得到的补偿音频数据与上一帧的历史音频数据是不同的,因此,可以减少播放的音频数据机械且重复的问题的出现。

[0198] 与上述方法实施例相对应,本说明书还提供了丢包补偿装置实施例,图7示出了本说明书一实施例提供的一种丢包补偿装置的结构示意图。如图7所示,该装置包括:

[0199] 接收模块702,被配置为接收音频数据流;

[0200] 第一确定模块704,被配置为若确定所述音频数据流中当前帧为丢包帧,且所述当

前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

[0201] 第一生成模块706,被配置为从所述音频数据流中获取第一个丢包帧之前的M帧历史音频数据,生成第一音序列,其中,所述M为大于1的正整数;

[0202] 第二生成模块708,被配置为确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音序列;

[0203] 第二确定模块710,被配置为基于所述第一音序列,在所述第二音序列中确定与所述第一音序列的互相关性最大的位置,并将所述第二音序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

[0204] 可选地,所述第二生成模块708被配置为:

[0205] 将所述第一个丢包帧上一帧的历史音频数据的音频特征确定为所述第一个丢包帧的音频数据的预测音频特征;

[0206] 将所述M帧历史音频数据的音频特征与所述第一个丢包帧的音频数据的预测音频特征输入音频合成模型中,得到所述第二音序列,其中,所述第二音序列中包括合成的所述M帧与所述第一个丢包帧的采样音频数据。

[0207] 可选地,所述第二确定模块710被配置为:

[0208] 基于所述第二音序列的采样频率从所述第一音序列中沿终点向起点方向获取N个连续的采样音频数据,得到第一采样音频数据组,其中,所述N为大于1的正整数;

[0209] 以N为采样音频数据组的长度,逐次从所述第二音序列中沿终点向起点方向获取N个连续的采样音频数据,得到第二采样音频数据组集合,其中,采样音频数据组的长度为采样音频数据组包括的采样音频数据的数量;

[0210] 对每个得到的第二采样音频数据组,将所得到的第二采样音频数据组与所述第一采样音频数据组进行互相关性计算,以互相关性最大的第二采样音频数据组所在的位置确定所述互相关性最大的位置。

[0211] 可选地,所述第二确定模块710被配置为:

[0212] 将所述互相关性最大的第二采样音频数据组中距离终点最近的采样音频数据所在的位置确定为所述互相关性最大的位置。

[0213] 可选地,所述第二确定模块710还被配置为:

[0214] 从所述第一音序列中沿终点向起点的方向获取X个未播放的采样音频数据,其中,所述X大于1且小于一帧采样音频数据的数量;

[0215] 将所述第二音序列中所述互相关性最大的位置之前的X个采样音频数据与所述X个未播放的采样音频数据进行融合,得到融合音频数据;

[0216] 将所述第一音序列中X个未播放的音频数据替换为所述融合音频数据,并将所述融合音频数据与所述第一个丢包帧的补偿音频数据进行拼接。

[0217] 可选地,所述第二确定模块710还被配置为:

[0218] 若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧为丢包帧,则确定所述当前帧为连续丢包帧,其中,所述连续丢包帧为多个连续的丢包帧中非第一个丢包帧的丢包帧;

[0219] 基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述连续丢包帧的第二补偿音频数据。

[0220] 可选地,所述第二确定模块710还被配置为:

[0221] 以所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征作为所述连续丢包帧的预测音频特征;

[0222] 将所述连续丢包帧的预测音频特征输入音频合成模型中,得到所述连续丢包帧的第二补偿音频数据。

[0223] 可选地,所述第二确定模块710还被配置为:

[0224] 若确定所述音频数据流中当前帧不是丢包帧,且所述当前帧的上一帧是丢包帧,则确定所述当前帧为目标帧;

[0225] 获取所述目标帧的实际音频数据,并基于所述当前帧之前最后一个非丢包帧的历史音频数据的音频特征,确定所述目标帧的预测音频数据;

[0226] 基于所述目标帧的实际音频数据,在所述目标帧的预测音频数据中确定与所述目标帧的实际音频数据互相关性最大的位置;

[0227] 将所述互相关性最大的位置之后的Y个采样音频数据与所述目标帧的实际音频数据中沿起点到终点方向的Y个采样音频数据进行融合,得到所述目标帧融合后的音频数据,其中,所述Y为大于1的正整数。

[0228] 可选地,所述装置还包括:

[0229] 预测模块,用于将样本音频特征和样本采样音频数据输入所述音频合成模型中,输出预测采样音频数据;

[0230] 训练模块,用于基于所述预测采样音频数据和所述样本采样音频数据对所述音频合成模型进行训练,直至达到训练停止条件。

[0231] 本说明书提供的丢包补偿方法,接收音频数据流,若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。上述方法通过第一个丢包帧的前M帧的历史音频数据以及音频特征生成第二音频序列,并将真实的第一音频序列和预测生成的第二音频序列进行比对,得到第二音频序列中与第一音频序列的互相关性最大的位置,进而确定第一个丢包帧的第一补偿音频数据,得到的补偿音频数据与上一帧的历史音频数据是不同的,因此,可以减少播放的音频数据机械且重复的问题的出现。

[0232] 上述为本实施例的一种丢包补偿装置的示意性方案。需要说明的是,该丢包补偿装置的技术方案与上述的丢包补偿方法的技术方案属于同一构思,丢包补偿装置的技术方案未详细描述的细节内容,均可以参见上述丢包补偿方法的技术方案的描述。

[0233] 图8示出了根据本说明书一实施例提供的一种计算设备800的结构框图。该计算设备800的部件包括但不限于存储器810和处理器820。处理器820与存储器810通过总线830相连接,数据库850用于保存数据。

[0234] 计算设备800还包括接入设备840,接入设备840使得计算设备800能够经由一个或多个网络860通信。这些网络的示例包括公用交换电话网(PSTN)、局域网(LAN)、广域网

(WAN)、个域网 (PAN) 或诸如因特网的通信网络的组合。接入设备840可以包括有线或无线的任何类型的网络接口(例如,网络接口卡(NIC))中的一个或多个,诸如IEEE802.11无线局域网(WLAN)无线接口、全球微波互联接入(Wi-MAX)接口、以太网接口、通用串行总线(USB)接口、蜂窝网络接口、蓝牙接口、近场通信(NFC)接口,等等。

[0235] 在本说明书的一个实施例中,计算设备800的上述部件以及图8中未示出的其他部件也可以彼此相连接,例如通过总线。应当理解,图8所示的计算设备结构框图仅仅是出于示例的目的,而不是对本说明书范围的限制。本领域技术人员可以根据需要,增添或替换其他部件。

[0236] 计算设备800可以是任何类型的静止或移动计算设备,包括移动计算机或移动计算设备(例如,平板计算机、个人数字助理、膝上型计算机、笔记本电脑、上网本等)、移动电话(例如,智能手机)、可佩戴的计算设备(例如,智能手表、智能眼镜等)或其他类型的移动设备,或者诸如台式计算机或PC的静止计算设备。计算设备800还可以是移动式或静止式的服务器。

[0237] 其中,处理器820用于执行如下计算机可执行指令:

[0238] 接收音频数据流;

[0239] 若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

[0240] 从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;

[0241] 确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;

[0242] 基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

[0243] 上述为本实施例的一种计算设备的示意性方案。需要说明的是,该计算设备的技术方案与上述的丢包补偿方法的技术方案属于同一构思,计算设备的技术方案未详细描述的细节内容,均可以参见上述丢包补偿方法的技术方案的描述。

[0244] 本说明书一实施例还提供一种计算机可读存储介质,其存储有计算机指令,该指令被处理器执行时以用于:

[0245] 接收音频数据流;

[0246] 若确定所述音频数据流中当前帧为丢包帧,且所述当前帧的上一帧不是丢包帧,则确定所述当前帧为第一个丢包帧;

[0247] 从所述音频数据流中获取所述第一个丢包帧之前的M帧历史音频数据,生成第一音频序列,其中,所述M为大于1的正整数;

[0248] 确定所述M帧历史音频数据的音频特征,并基于所述M帧历史音频数据的音频特征,生成第二音频序列;

[0249] 基于所述第一音频序列,在所述第二音频序列中确定与所述第一音频序列的互相关性最大的位置,并将所述第二音频序列中所述互相关性最大的位置之后的音频数据确定为所述第一个丢包帧的第一补偿音频数据。

[0250] 上述为本实施例的一种计算机可读存储介质的示意性方案。需要说明的是,该存储介质的技术方案与上述的丢包补偿方法的技术方案属于同一构思,存储介质的技术方案未详细描述的细节内容,均可以参见上述丢包补偿方法的技术方案的描述。

[0251] 上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下,在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外,在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中,多任务处理和并行处理也是可以的或者可能是有利的。

[0252] 所述计算机指令包括计算机程序代码,所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括:能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、电载波信号、电信信号以及软件分发介质等。需要说明的是,所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减,例如在某些司法管辖区,根据立法和专利实践,计算机可读介质不包括电载波信号和电信信号。

[0253] 需要说明的是,对于前述的各方法实施例,为了简便描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本说明书并不受所描述的动作顺序的限制,因为依据本说明书,某些步骤可以采用其它顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本说明书所必须的。

[0254] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其它实施例的相关描述。

[0255] 以上公开的本说明书优选实施例只是用于帮助阐述本说明书。可选实施例并没有详尽叙述所有的细节,也不限制该发明仅为所述的具体实施方式。显然,根据本说明书的内容,可作很多的修改和变化。本说明书选取并具体描述这些实施例,是为了更好地解释本说明书的原理和实际应用,从而使所属技术领域技术人员能很好地理解和利用本说明书。本说明书仅受权利要求书及其全部范围和等效物的限制。

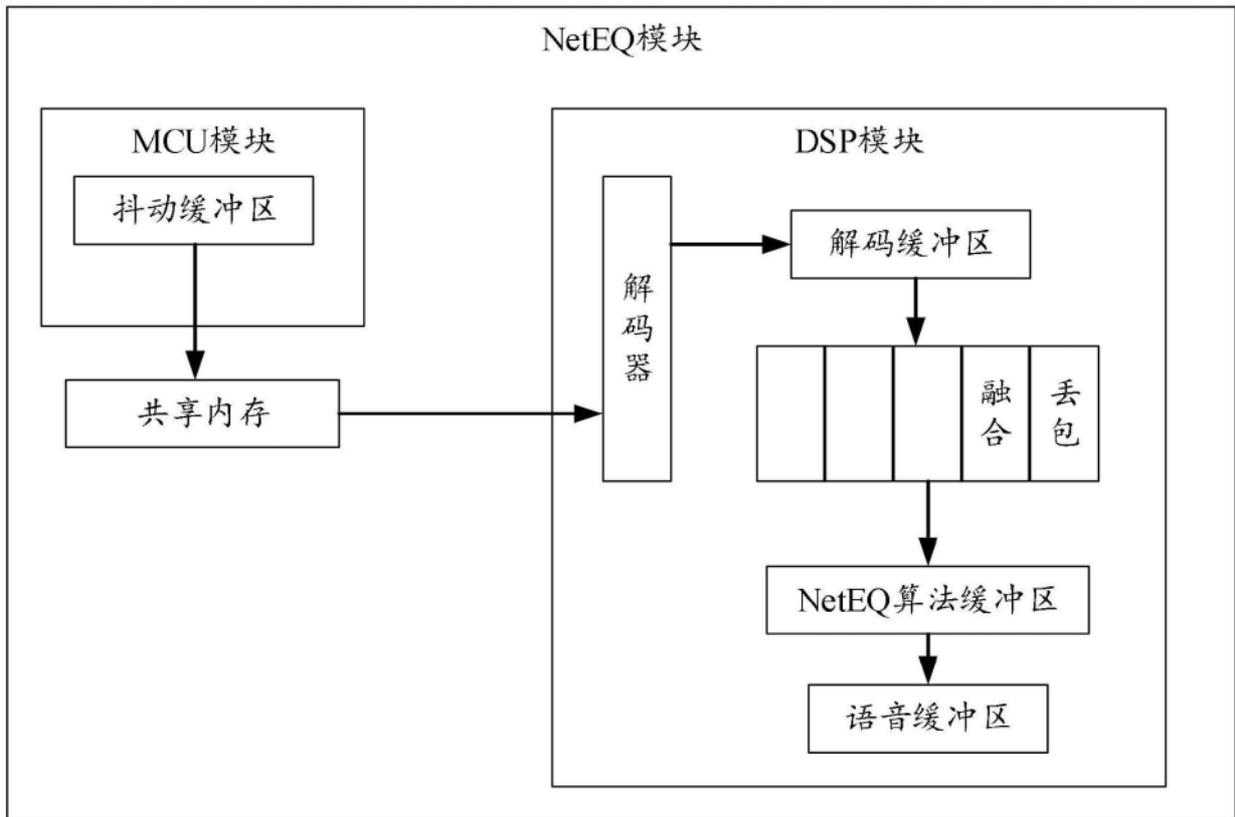


图1

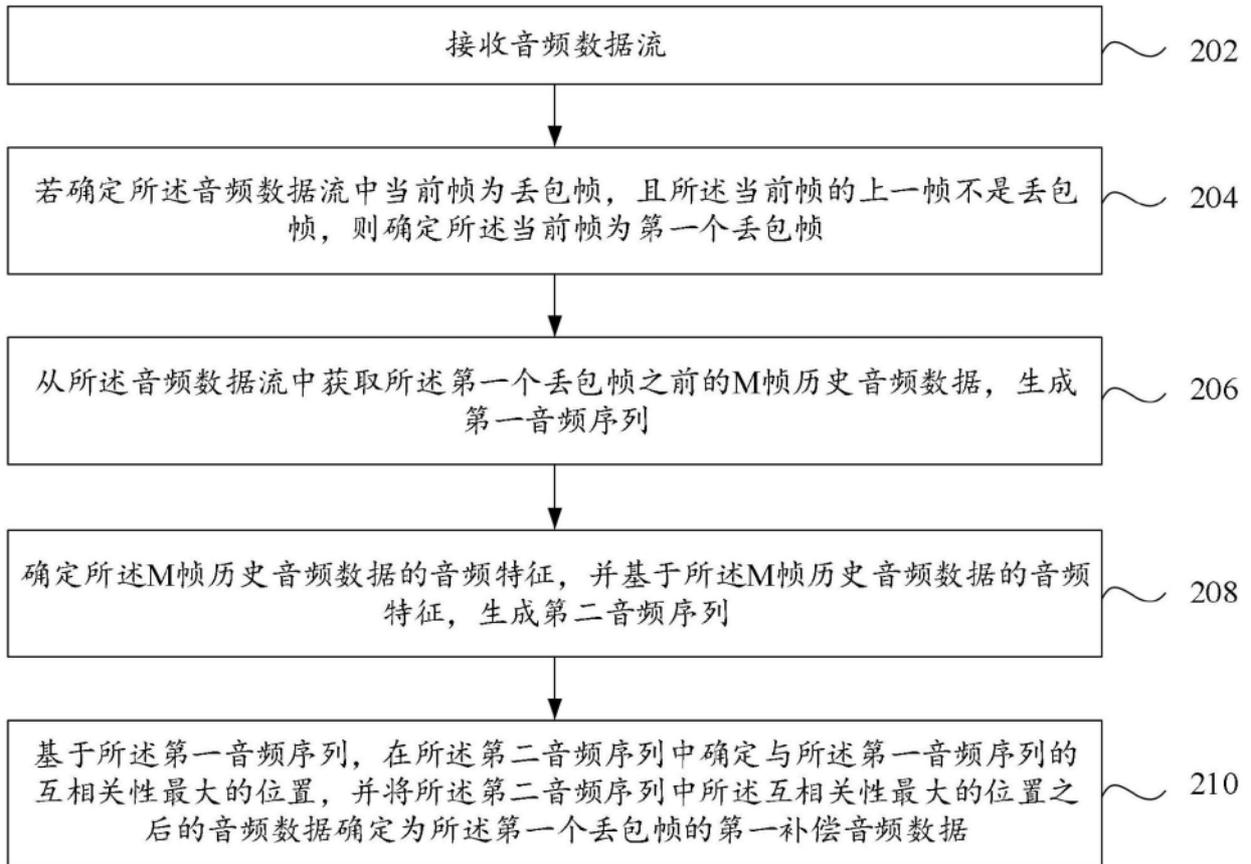


图2

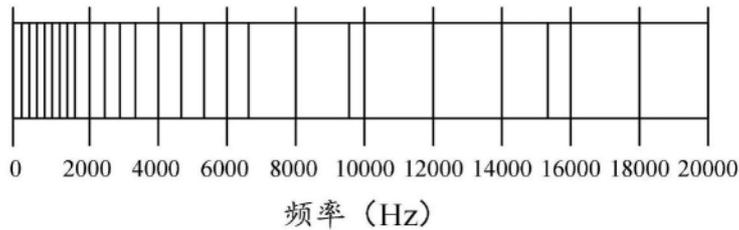


图3

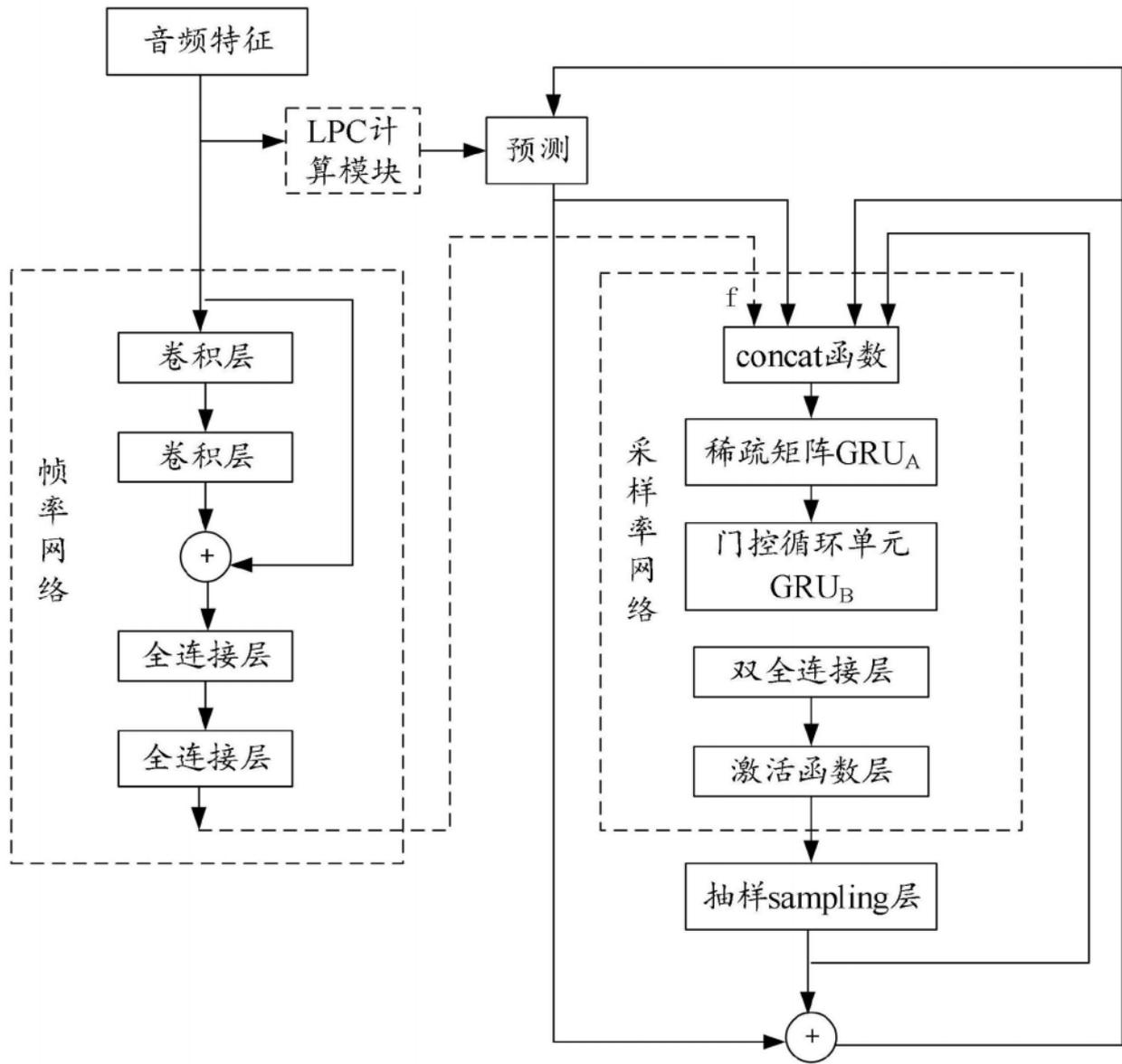


图4

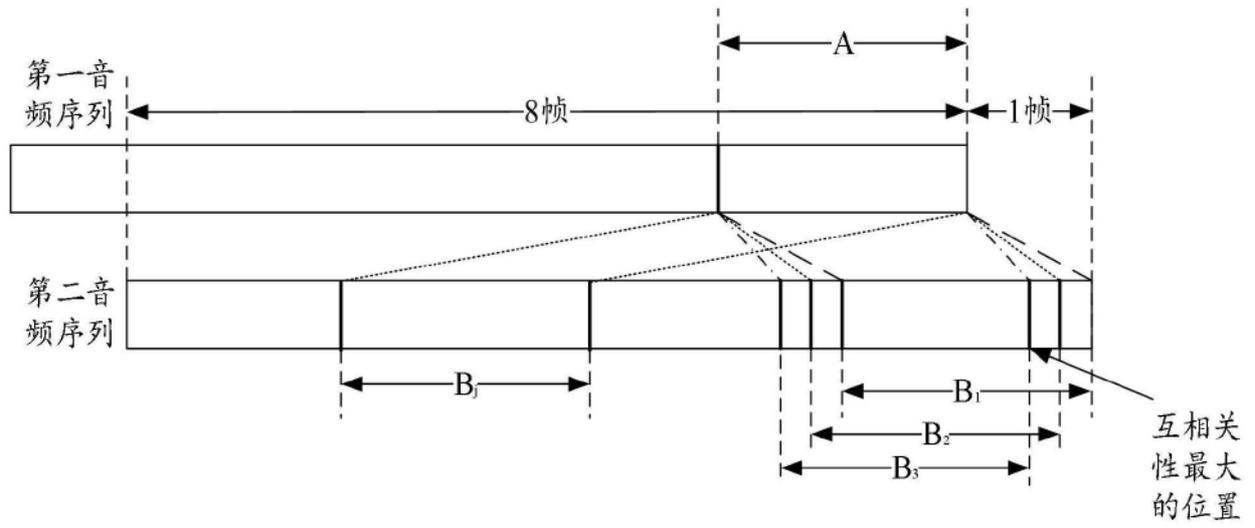


图5

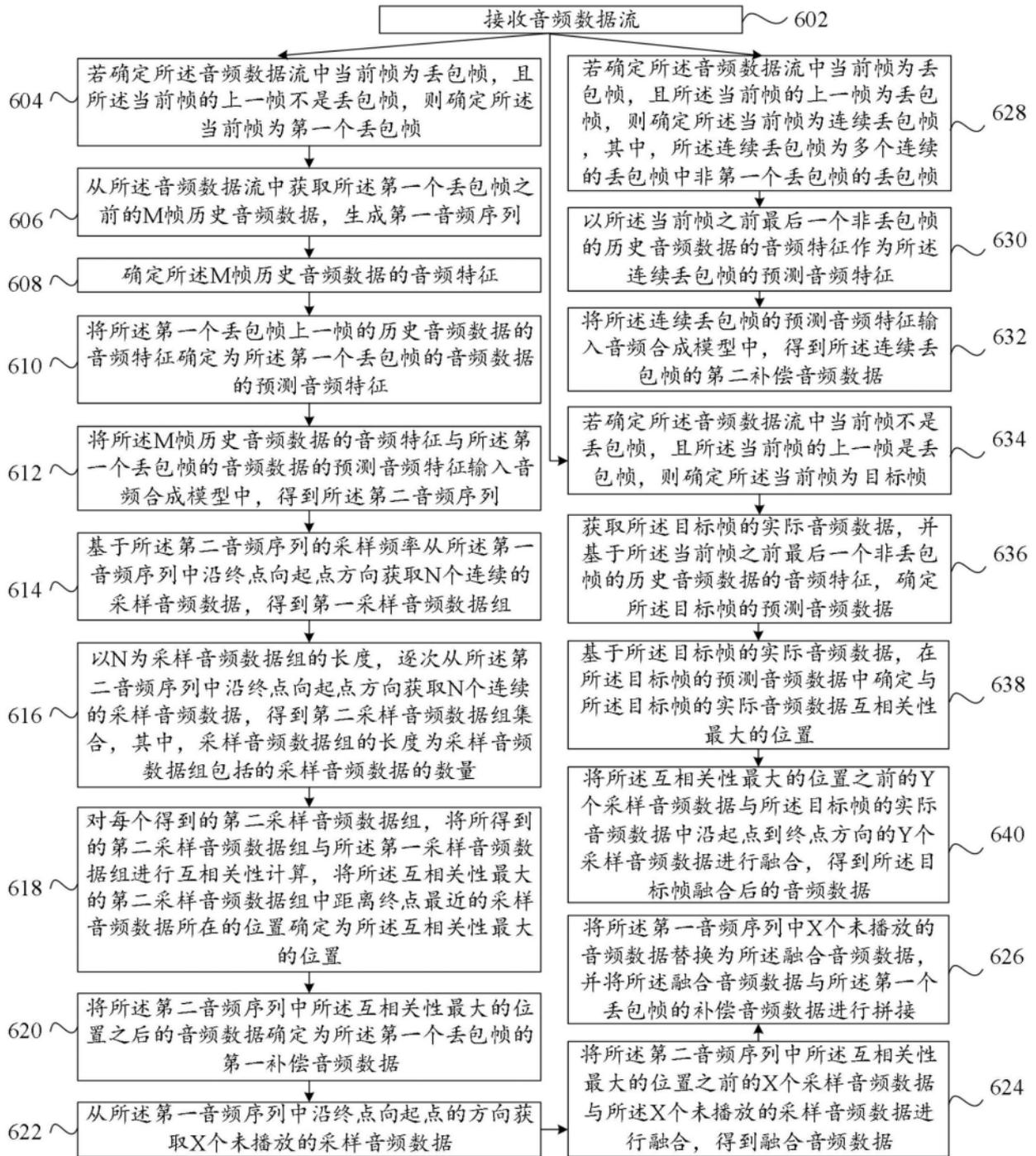


图6

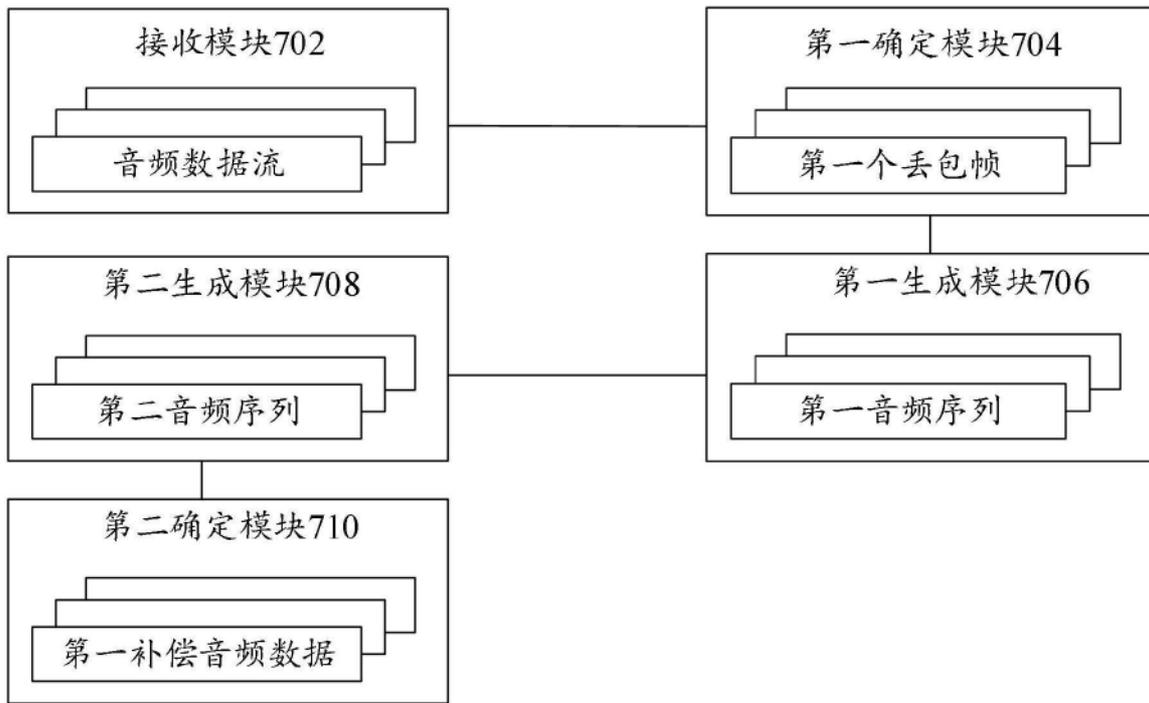


图7

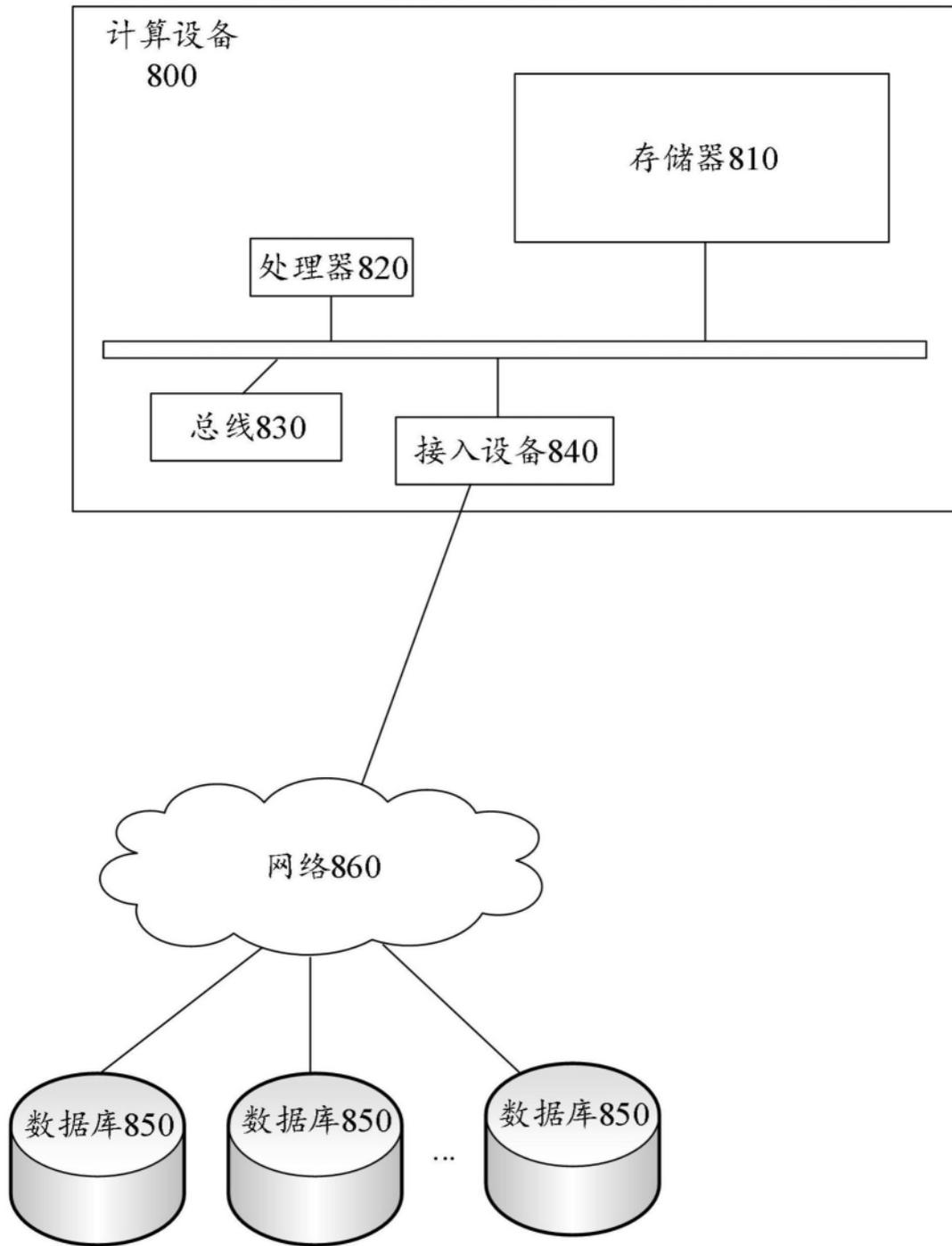


图8