



(12) 发明专利申请

(10) 申请公布号 CN 112767954 A

(43) 申请公布日 2021.05.07

(21) 申请号 202010592469.4

(22) 申请日 2020.06.24

(71) 申请人 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 深圳市隆天联鼎知识产权代
理有限公司 44232

代理人 叶虹

(51) Int. Cl.

G10L 19/00 (2013.01)

G10L 19/02 (2013.01)

G10L 19/04 (2013.01)

G10L 19/08 (2013.01)

G10L 25/30 (2013.01)

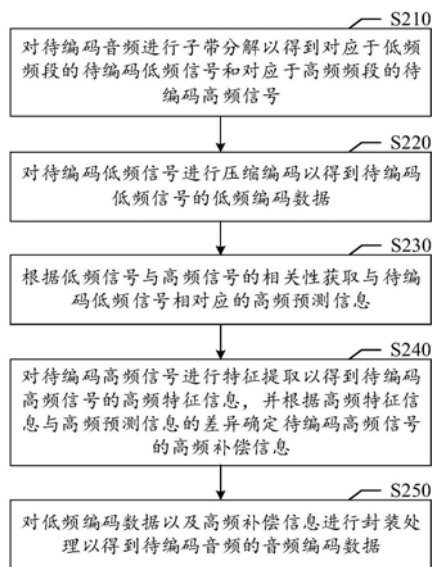
权利要求书5页 说明书23页 附图11页

(54) 发明名称

音频编解码方法、装置、介质及电子设备

(57) 摘要

本申请属于音频编解码技术领域,具体涉及一种音频编解码方法、装置、介质以及电子设备。音频编码方法包括:对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号;对待编码低频信号进行压缩编码以得到待编码低频信号的低频编码数据;根据低频信号与高频信号的相关性获取与待编码低频信号相对应的高频预测信息;对待编码高频信号进行特征提取以得到待编码高频信号的高频特征信息,并根据高频特征信息与高频预测信息的差异确定待编码高频信号的高频补偿信息;对低频编码数据以及高频补偿信息进行封装处理以得到待编码音频的音频编码数据。该方法可以压缩高频信号的编码码率并保证高频信号的准确性。



1. 一种音频编码方法,其特征在于,包括:

对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号;

对所述待编码低频信号进行压缩编码以得到所述待编码低频信号的低频编码数据;

根据低频信号与高频信号的相关性获取与所述待编码低频信号相对应的高频预测信息;

对所述待编码高频信号进行特征提取以得到所述待编码高频信号的高频特征信息,并根据所述高频特征信息与所述高频预测信息的差异确定所述待编码高频信号的高频补偿信息;

对所述低频编码数据以及所述高频补偿信息进行封装处理以得到所述待编码音频的音频编码数据。

2. 根据权利要求1所述的音频编码方法,其特征在于,所述根据低频信号与高频信号的相关性获取与所述待编码低频信号相对应的高频预测信息,包括:

对所述待编码音频进行分类处理以得到所述待编码音频的音频类别信息;

根据所述音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络;

通过所述高频预测神经网络对所述待编码低频信号进行映射处理以得到与所述待编码低频信号相对应的高频预测信息。

3. 根据权利要求2所述的音频编码方法,其特征在于,所述对所述待编码音频进行分类处理以得到所述待编码音频的音频类别信息,包括:

获取音频数据样本,并对所述音频数据样本进行逐帧标注以得到所述音频数据样本中每个数据帧的音频类别标识;

从多个特征维度对所述音频数据样本进行特征提取以得到所述音频数据样本的多维样本特征;

以所述多维样本特征作为输入值并以所述音频类别标识作为与所述输入值相对应的目标值训练得到用于对音频数据进行分类处理的音频分类神经网络;

通过所述音频分类神经网络对所述待编码音频进行分类处理以得到所述待编码音频的音频类别信息。

4. 根据权利要求2所述的音频编码方法,其特征在于,所述根据所述音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络,包括:

获取与所述音频类别信息相对应的音频数据样本,并对所述音频数据样本进行压缩变换以得到所述音频数据样本的频谱特征样本;

对所述频谱特征样本按照频点的数值进行划分以得到低频特征样本和高频特征样本;

以所述低频特征样本作为输入值并以所述高频特征样本作为与所述输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

5. 根据权利要求2所述的音频编码方法,其特征在于,所述根据所述音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络,包括:

获取与所述音频类别信息相对应的音频数据样本,并将所述音频数据样本按照所在频段的高低分解为低频数据样本和高频数据样本;

分别对所述低频数据样本和所述高频数据样本进行压缩变换以得到相应的低频特征样本和高频特征样本；

以所述低频特征样本作为输入值并以所述高频特征样本作为与所述输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

6. 根据权利要求2所述的音频编码方法,其特征在於,所述通过所述高频预测神经网络对所述待编码低频信号进行映射处理以得到与所述待编码低频信号相对应的高频预测信息,包括:

对所述低频编码数据进行解码处理以得到与所述待编码低频信号相对应的低频解码信号;

通过所述高频预测神经网络对所述低频解码信号进行映射处理以得到与所述待编码低频信号相对应的高频预测信息。

7. 根据权利要求6所述的音频编码方法,其特征在於,所述通过所述高频预测神经网络对所述低频解码信号进行映射处理以得到与所述待编码低频信号相对应的高频预测信息,包括:

对所述低频解码信号进行压缩变换以得到所述低频解码信号的低频频谱特征;

通过所述高频预测神经网络对所述低频频谱特征进行映射处理以得到与所述待编码低频信号相对应的高频预测信息。

8. 根据权利要求1所述的音频编码方法,其特征在於,所述对待编码音频进行子带分解以得到对应於低频频段的待编码低频信号和对应於高频频段的待编码高频信号,包括:

获取由对应於低频频段的低通滤波器和对应於高频频段的高通滤波器组成的正交镜像滤波器组;

通过所述正交镜像滤波器组对待编码音频进行子带分解以得到对应於低频频段的待编码低频信号和对应於高频频段的待编码高频信号。

9. 根据权利要求1所述的音频编码方法,其特征在於,所述对所述待编码高频信号进行特征提取以得到所述待编码高频信号的高频特征信息,包括:

对所述待编码高频信号进行压缩变换以得到所述待编码高频信号的高频特征信息。

10. 根据权利要求1所述的音频编码方法,其特征在於,所述根据所述高频特征信息与所述高频预测信息的差异确定所述待编码高频信号的高频补偿信息,包括:

将所述高频特征信息从线性频率域映射至临界频带域以得到与所述高频特征信息相对应的特征频谱信息;

将所述高频预测信息从线性频率域映射至临界频带域以得到与所述高频预测信息相对应的预测频谱信息;

根据所述特征频谱信息和所述预测频谱信息的差异确定所述待编码高频信号的高频补偿信息。

11. 根据权利要求10所述的音频编码方法,其特征在於,所述根据所述特征频谱信息和所述预测频谱信息的差异确定所述待编码高频信号的高频补偿信息,包括:

分别对所述特征频谱信息和所述预测频谱信息进行对数变换以得到特征频谱对数值和预测频谱对数值;

根据所述特征频谱对数值和预测频谱对数值的差值查询增益码表以得到增益量化值,

并将所述增量量化值确定为所述待编码高频信号的高频补偿信息。

12. 根据权利要求1所述的音频编码方法,其特征在在于,所述对所述低频编码数据以及所述高频补偿信息进行封装处理以得到所述待编码音频的音频编码数据,包括:

对所述待编码音频进行分类处理以得到所述待编码音频的音频类别信息;

对所述音频类别信息、低频编码数据以及所述高频补偿信息进行封装处理以得到所述待编码音频的音频编码数据。

13. 一种音频解码方法,其特征在在于,包括:

对待解码的音频编码数据进行封装解析以得到所述音频编码数据中的低频编码数据和高频补偿信息;

对所述低频编码数据进行解码处理以得到还原低频信号;

根据低频信号与高频信号的相关性获取与所述还原低频信号相对应的高频预测信息;

根据所述高频补偿信息对所述高频预测信息进行增益补偿以得到高频特征信息,并对所述高频特征信息进行特征还原以得到还原高频信号;

对所述还原低频信号和所述还原高频信号进行子带合成以得到所述音频编码数据的还原音频。

14. 根据权利要求13所述的音频解码方法,其特征在在于,所述根据低频信号与高频信号的相关性获取与所述还原低频信号相对应的高频预测信息,包括:

对所述音频编码数据进行封装解析以得到所述音频编码数据中的音频类别信息;

根据所述音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络;

通过所述高频预测神经网络对所述还原低频信号进行映射处理以得到与所述还原低频信号相对应的高频预测信息。

15. 根据权利要求14所述的音频解码方法,其特征在在于,所述根据所述音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络,包括:

获取与所述音频类别信息相对应的音频数据样本,并对所述音频数据样本进行压缩变换以得到所述音频数据样本的频谱特征样本;

对所述频谱特征样本按照频点的数值进行划分以得到低频特征样本和高频特征样本;

以所述低频特征样本作为输入值并以所述高频特征样本作为与所述输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

16. 根据权利要求14所述的音频解码方法,其特征在在于,所述根据所述音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络,包括:

获取与所述音频类别信息相对应的音频数据样本,并将所述音频数据样本按照所在频段的高低分解为低频数据样本和高频数据样本;

分别对所述低频数据样本和所述高频数据样本进行压缩变换以得到相应的低频特征样本和高频特征样本;

以所述低频特征样本作为输入值并以所述高频特征样本作为与所述输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

17. 根据权利要求14所述的音频解码方法,其特征在在于,所述通过所述高频预测神经网络对所述还原低频信号进行映射处理以得到与所述还原低频信号相对应的高频预测信息,

包括：

对所述还原低频信号进行压缩变换以得到所述还原低频信号的低频频谱特征；

通过所述高频预测神经网络对所述低频频谱特征进行映射处理以得到与所述还原低频信号相对应的高频预测信息。

18. 根据权利要求13所述的音频解码方法，其特征在于，所述根据所述高频补偿信息对所述高频预测信息进行增益补偿以得到高频特征信息，包括：

将所述高频预测信息从线性频率域映射至临界频带域以得到与所述高频预测信息相对应的预测频谱信息；

根据所述高频补偿信息对所述预测频谱信息进行增益补偿以得到特征频谱信息；

将所述特征频谱信息从临界频带域映射至线性频率域以得到与所述特征频谱信息相对应的高频特征信息。

19. 根据权利要求18所述的音频解码方法，其特征在于，所述根据所述高频补偿信息对所述预测频谱信息进行增益补偿以得到特征频谱信息，包括：

对所述预测频谱信息进行对数变换以得到预测频谱对数值；

根据所述高频补偿信息对所述预测频谱对数值进行增益补偿以得到特征频谱对数值；

对所述特征频谱对数值进行指数还原以得到特征频谱信息。

20. 根据权利要求13所述的音频解码方法，其特征在于，所述对所述高频特征信息进行特征还原以得到还原高频信号，包括：

对所述高频特征信息进行解压缩变换以得到还原高频信号。

21. 根据权利要求13所述的音频解码方法，其特征在于，所述对所述还原低频信号和所述还原高频信号进行子带合成以得到所述音频编码数据的还原音频，包括：

获取由对应于低频频段的低通滤波器和对应于高频频段的高通滤波器组成的正交镜像滤波器组；

通过所述正交镜像滤波器组对所述还原低频信号和所述还原高频信号进行子带合成以得到所述音频编码数据的还原音频。

22. 根据权利要求13所述的音频解码方法，其特征在于，所述对待解码的音频编码数据进行封装解析以得到所述音频编码数据中的低频编码数据和高频补偿信息，包括：

在待解码的音频编码数据中搜索码流单元分隔信息；

根据搜索到的码流单元分隔信息从所述音频编码数据中分离得到待解码的码流单元；

对所述码流单元进行字段解析以得到所述码流单元中封装的低频编码数据和高频补偿信息。

23. 一种音频编码装置，其特征在于，包括：

音频分解模块，被配置为对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段待编码高频信号；

低频编码模块，被配置为对所述待编码低频信号进行压缩编码以得到所述待编码低频信号的低频编码数据；

高频预测模块，被配置为根据低频信号与高频信号的相关性获取与所述待编码低频信号相对应的高频预测信息；

高频补偿模块，被配置为对所述待编码高频信号进行特征提取以得到所述待编码高频

信号的高频特征信息,并根据所述高频特征信息与所述高频预测信息的差异确定所述待编码高频信号的高频补偿信息;

编码封装模块,被配置为对所述低频编码数据以及所述高频补偿信息进行封装处理以得到所述待编码音频的音频编码数据。

24. 一种音频解码装置,其特征在于,包括:

封装解析模块,被配置为对待解码的音频编码数据进行封装解析以得到所述音频编码数据中的低频编码数据和高频补偿信息;

低频解码模块,被配置为对所述低频编码数据进行解码处理以得到还原低频信号;

高频预测模块,被配置为根据低频信号与高频信号的相关性获取与所述还原低频信号相对应的高频预测信息;

高频还原模块,被配置为根据所述高频补偿信息对所述高频预测信息进行增益补偿以得到高频特征信息,并对所述高频特征信息进行特征还原以得到还原高频信号;

音频合成模块,被配置为对所述还原低频信号和所述还原高频信号进行子带合成以得到所述音频编码数据的原始音频。

25. 一种计算机可读介质,其上存储有计算机程序,该计算机程序被处理器执行时实现权利要求1至22中任意一项所述的方法。

26. 一种电子设备,其特征在于,包括:

处理器;以及

存储器,用于存储所述处理器的可执行指令;

其中,所述处理器配置为经由执行所述可执行指令来执行权利要求1至22中任意一项所述的方法。

音频编解码方法、装置、介质及电子设备

技术领域

[0001] 本申请属于音频编解码技术领域,并涉及人工智能技术。具体而言,本申请涉及一种音频编码方法、音频解码方法、音频编码装置、音频解码装置、计算机可读介质以及电子设备。

背景技术

[0002] 音频编解码在现代通讯系统中占有重要的地位,通过对音频数据进行压缩编码处理可以降低音频数据在网络传输中的网络带宽压力,节省音频数据的存储成本和传输成本。

[0003] 音乐、语音等音频数据的频谱主要集中在低频段,高频段幅度很小。如果对整个频段编码,若是为了保护高频段数据就会造成低频段数据编码过细而导致编码文件数据量巨大,难以取得理想的压缩编码效果;若是保存了低频段数据的主要成分而丢弃高频段数据成分则会丧失音质,导致音频在解码后出现严重的失真问题。因此,如何在音频数据的编解码过程中尽可能地保证高频段数据的准确传输是目前亟待解决的问题。

[0004] 需要说明的是,在上述背景技术部分公开的信息仅用于加强对本申请的背景的理解,因此可以包括不构成对本领域普通技术人员已知的现有技术的信息。

发明内容

[0005] 本申请的目的在于提供一种音频编码方法、音频解码方法、音频编码装置、音频解码装置、计算机可读介质以及电子设备,至少在一定程度上克服音频编解码技术中存在的高频段音频数据成分丢失、传输准确性差等技术问题。

[0006] 本申请的其他特性和优点将通过下面的详细描述变得显然,或部分地通过本申请的实践而习得。

[0007] 根据本申请实施例的一个方面,提供一种音频编码方法,该方法包括:对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号;对所述待编码低频信号进行压缩编码以得到所述待编码低频信号的低频编码数据;根据低频信号与高频信号的相关性获取与所述待编码低频信号相对应的高频预测信息;对所述待编码高频信号进行特征提取以得到所述待编码高频信号的高频特征信息,并根据所述高频特征信息与所述高频预测信息的差异确定所述待编码高频信号的高频补偿信息;对所述低频编码数据以及所述高频补偿信息进行封装处理以得到所述待编码音频的音频编码数据。

[0008] 根据本申请实施例的一个方面,提供一种音频编码装置,该装置包括:音频分解模块,被配置为对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号;低频编码模块,被配置为对所述待编码低频信号进行压缩编码以得到所述待编码低频信号的低频编码数据;高频预测模块,被配置为根据低频信号与高频信号的相关性获取与所述待编码低频信号相对应的高频预测信息;高频补偿模块,

被配置为对所述待编码高频信号进行特征提取以得到所述待编码高频信号的高频特征信息,并根据所述高频特征信息与所述高频预测信息的差异确定所述待编码高频信号的高频补偿信息;编码封装模块,被配置为对所述低频编码数据以及所述高频补偿信息进行封装处理以得到所述待编码音频的音频编码数据。

[0009] 在本申请的一些实施例中,基于以上技术方案,所述高频预测模块包括:音频分类单元,被配置为对所述待编码音频进行分类处理以得到所述待编码音频的音频类别信息;编码端网络获取单元,被配置为根据所述音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络;编码端网络映射单元,被配置为通过所述高频预测神经网络对所述待编码低频信号进行映射处理以得到与所述待编码低频信号相对应的高频预测信息。

[0010] 在本申请的一些实施例中,基于以上技术方案,所述音频分类单元包括:样本类别标注子单元,被配置为获取音频数据样本,并对所述音频数据样本进行逐帧标注以得到所述音频数据样本中每个数据帧的音频类别标识;样本特征提取子单元,被配置为从多个特征维度对所述音频数据样本进行特征提取以得到所述音频数据样本的多维样本特征;分类网络训练子单元,被配置为以所述多维样本特征作为输入值并以所述音频类别标识作为与所述输入值相对应的目标值训练得到用于对音频数据进行分类处理的音频分类神经网络;分类网络处理子单元,被配置为通过所述音频分类神经网络对所述待编码音频进行分类处理以得到所述待编码音频的音频类别信息。

[0011] 在本申请的一些实施例中,基于以上技术方案,所述编码端网络获取单元包括:第一样本变换子单元,被配置为获取与所述音频类别信息相对应的音频数据样本,并对所述音频数据样本进行压缩变换以得到所述音频数据样本的频谱特征样本;第一频段划分子单元,被配置为对所述频谱特征样本按照频点的数值进行划分以得到低频特征样本和高频特征样本;第一网络获取子单元,被配置为以所述低频特征样本作为输入值并以所述高频特征样本作为与所述输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0012] 在本申请的一些实施例中,基于以上技术方案,所述编码端网络获取单元包括:第二频段划分子单元,被配置为获取与所述音频类别信息相对应的音频数据样本,并将所述音频数据样本按照所在频段的高低分解为低频数据样本和高频数据样本;第二样本变换子单元,被配置为分别对所述低频数据样本和所述高频数据样本进行压缩变换以得到相应的低频特征样本和高频特征样本;第二网络获取子单元,被配置为以所述低频特征样本作为输入值并以所述高频特征样本作为与所述输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0013] 在本申请的一些实施例中,基于以上技术方案,所述编码端网络映射单元包括:编码端低频解码子单元,被配置为对所述低频编码数据进行解码处理以得到与所述待编码低频信号相对应的低频解码信号;编码端低频映射子单元,被配置为通过所述高频预测神经网络对所述低频解码信号进行映射处理以得到与所述待编码低频信号相对应的高频预测信息。

[0014] 在本申请的一些实施例中,基于以上技术方案,所述编码端低频映射子单元包括:编码端压缩变换子单元,被配置为对所述低频解码信号进行压缩变换以得到所述低频解码

信号的低频频谱特征;编码端特征映射子单元,被配置为通过所述高频预测神经网络对所述低频频谱特征进行映射处理以得到与所述待编码低频信号相对应的高频预测信息。

[0015] 在本申请的一些实施例中,基于以上技术方案,所述音频分解模块包括:滤波器获取单元,被配置为获取由对应于低频频段的低通滤波器和对应于高频频段的高通滤波器组成的正交镜像滤波器组;子带分解单元,被配置为通过所述正交镜像滤波器组对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号。

[0016] 在本申请的一些实施例中,基于以上技术方案,所述高频补偿模块包括:高频压缩变换单元,被配置为对所述待编码高频信号进行压缩变换以得到所述待编码高频信号的高频特征信息。

[0017] 在本申请的一些实施例中,基于以上技术方案,所述高频补偿模块还包括:特征频谱转换单元,被配置为将所述高频特征信息从线性频率域映射至临界频带域以得到与所述高频特征信息相对应的特征频谱信息;预测频谱转换单元,被配置为将所述高频预测信息从线性频率域映射至临界频带域以得到与所述高频预测信息相对应的预测频谱信息;补偿信息确定单元,被配置为根据所述特征频谱信息和所述预测频谱信息的差异确定所述待编码高频信号的高频补偿信息。

[0018] 在本申请的一些实施例中,基于以上技术方案,所述补偿信息确定单元包括:第一对数变换子单元,被配置为分别对所述特征频谱信息和所述预测频谱信息进行对数变换以得到特征频谱对数值和预测频谱对数值;增益量化子单元,被配置为根据所述特征频谱对数值和预测频谱对数值的差值查询增益码表以得到增益量化值,并将所述增量量化值确定为所述待编码高频信号的高频补偿信息。

[0019] 在本申请的一些实施例中,基于以上技术方案,所述编码封装模块包括:音频分类单元,被配置为对所述待编码音频进行分类处理以得到所述待编码音频的音频类别信息;编码封装单元,被配置为对所述音频类别信息、低频编码数据以及所述高频补偿信息进行封装处理以得到所述待编码音频的音频编码数据。

[0020] 根据本申请实施例的一个方面,提供一种音频解码方法,该方法包括:对待解码的音频编码数据进行封装解析以得到所述音频编码数据中的低频编码数据和高频补偿信息;对所述低频编码数据进行解码处理以得到还原低频信号;根据低频信号与高频信号的相关性获取与所述还原低频信号相对应的高频预测信息;根据所述高频补偿信息对所述高频预测信息进行增益补偿以得到高频特征信息,并对所述高频特征信息进行特征还原以得到还原高频信号;对所述还原低频信号和所述还原高频信号进行子带合成以得到所述音频编码数据的还原音频。

[0021] 根据本申请实施例的一个方面,提供一种音频解码装置,该装置包括:封装解析模块,被配置为对待解码的音频编码数据进行封装解析以得到所述音频编码数据中的低频编码数据和高频补偿信息;低频解码模块,被配置为对所述低频编码数据进行解码处理以得到还原低频信号;高频预测模块,被配置为根据低频信号与高频信号的相关性获取与所述还原低频信号相对应的高频预测信息;高频还原模块,被配置为根据所述高频补偿信息对所述高频预测信息进行增益补偿以得到高频特征信息,并对所述高频特征信息进行特征还原以得到还原高频信号;音频合成模块,被配置为对所述还原低频信号和所述还原高频信

号进行子带合成以得到所述音频编码数据的原始音频。

[0022] 在本申请的一些实施例中,基于以上技术方案,所述高频预测模块包括:类别获取单元,被配置为对所述音频编码数据进行封装解析以得到所述音频编码数据中的音频类别信息;解码端网络获取单元,被配置为根据所述音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络;解码端网络映射单元,被配置为通过所述高频预测神经网络对所述还原低频信号进行映射处理以得到与所述还原低频信号相对应的高频预测信息。

[0023] 在本申请的一些实施例中,基于以上技术方案,所述解码端网络获取单元包括:第一样本变换子单元,被配置为获取与所述音频类别信息相对应的音频数据样本,并对所述音频数据样本进行压缩变换以得到所述音频数据样本的频谱特征样本;第一频段划分子单元,被配置为对所述频谱特征样本按照频点的数值进行划分以得到低频特征样本和高频特征样本;第一网络获取子单元,被配置为以所述低频特征样本作为输入值并以所述高频特征样本作为与所述输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0024] 在本申请的一些实施例中,基于以上技术方案,所述解码端网络获取单元包括:第二频段划分子单元,被配置为获取与所述音频类别信息相对应的音频数据样本,并将所述音频数据样本按照所在频段的高低分解为低频数据样本和高频数据样本;第二样本变换子单元,被配置为分别对所述低频数据样本和所述高频数据样本进行压缩变换以得到相应的低频特征样本和高频特征样本;第二网络获取子单元,被配置为以所述低频特征样本作为输入值并以所述高频特征样本作为与所述输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0025] 在本申请的一些实施例中,基于以上技术方案,所述解码端网络映射单元包括:解码端压缩变换子单元,被配置为对所述还原低频信号进行压缩变换以得到所述还原低频信号的低频频谱特征;解码端特征映射子单元,被配置为通过所述高频预测神经网络对所述低频频谱特征进行映射处理以得到与所述还原低频信号相对应的高频预测信息。

[0026] 在本申请的一些实施例中,基于以上技术方案,所述高频还原模块包括:频谱信息预测单元,被配置为将所述高频预测信息从线性频率域映射至临界频带域以得到与所述高频预测信息相对应的预测频谱信息;频谱信息补偿单元,被配置为根据所述高频补偿信息对所述预测频谱信息进行增益补偿以得到特征频谱信息;特征信息确定单元,被配置为将所述特征频谱信息从临界频带域映射至线性频率域以得到与所述特征频谱信息相对应的高频特征信息。

[0027] 在本申请的一些实施例中,基于以上技术方案,所述频谱信息补偿单元包括:第二对数变换子单元,被配置为对所述预测频谱信息进行对数变换以得到预测频谱对数值;对数值补偿子单元,被配置为根据所述高频补偿信息对所述预测频谱对数值进行增益补偿以得到特征频谱对数值;指数还原子单元,被配置为对所述特征频谱对数值进行指数还原以得到特征频谱信息。

[0028] 在本申请的一些实施例中,基于以上技术方案,所述高频还原模块还包括:特征信息解压缩单元,被配置为对所述高频特征信息进行解压缩变换以得到还原高频信号。

[0029] 在本申请的一些实施例中,基于以上技术方案,所述音频合成模块包括:滤波器获

取单元,被配置为获取由对应于低频频段的低通滤波器和对应于高频频段的高通滤波器组成的正交镜像滤波器组;子带合成单元,被配置为通过所述正交镜像滤波器组对所述还原低频信号和所述还原高频信号进行子带合成以得到所述音频编码数据的还原音频。

[0030] 在本申请的一些实施例中,基于以上技术方案,所述封装解析模块包括:码流搜索单元,被配置为在待解码的音频编码数据中搜索码流单元分隔信息;码流分离单元,被配置为根据搜索到的码流单元分隔信息从音频编码数据中分离得到待解码的码流单元;码流解析单元,被配置为对码流单元进行字段解析以得到码流单元中封装的低频编码数据和高频补偿信息。

[0031] 根据本申请实施例的一个方面,提供一种计算机可读介质,其上存储有计算机程序,该计算机程序被处理器执行时实现如以上技术方案中的音频编码方法或音频解码方法。

[0032] 根据本申请实施例的一个方面,提供一种电子设备,该电子设备包括:处理器;以及存储器,用于存储所述处理器的可执行指令;其中,所述处理器被配置为经由执行所述可执行指令来执行如以上技术方案中的音频编码方法或音频解码方法。

[0033] 根据本申请实施例的一个方面,提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读介质中。计算机设备的处理器从计算机可读介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行以上技术方案中的音频编码方法或音频解码方法。

[0034] 在本申请实施例提供的技术方案中,根据低频信号与高频信号的相关性可以利用待编码低频信号获取相应的高频预测信息,然后基于高频预测信息与待编码高频信号之间的特征差异确定高频补偿信息,在音频编码数据的传输中可以仅对高频补偿信息进行传输,从而可以极大地压缩高频信号的编码码率,降低网络传输的带宽压力。同时,在音频编码数据的接收端,可以基于高频补偿信息对高频信号进行重建还原,从而保证了高频信号的完整性和准确性,避免了因数据压缩丢失而导致的音频失真、音质差等问题。

[0035] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本申请。

附图说明

[0036] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本申请的实施例,并与说明书一起用于解释本申请的原理。显而易见地,下面描述中的附图仅仅是本申请的一些实施例,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。在附图中:

[0037] 图1示意性地示出了应用本申请技术方案的示例性系统架构框图。

[0038] 图2示意性地示出了本申请一些实施例中音频编码方法的步骤流程图。

[0039] 图3示意性地示出了本申请一些实施例中编码端获取高频预测信息的方法步骤流程图。

[0040] 图4示意性地示出了本申请一些实施例中对待编码音频进行分类处理的方法步骤流程图。

[0041] 图5示意性地示出了本申请一些实施例中基于先特征提取后频段分割的预处理过

程训练高频预测神经网络的方法步骤流程图。

[0042] 图6示意性地示出了本申请一些实施例中基于先频段分割后特征提取的预处理过程训练高频预测神经网络的方法步骤流程图。

[0043] 图7示意性地示出了本申请一些实施例中确定高频补偿信息的方法步骤流程图。

[0044] 图8示意性地示出了本申请实施例在一应用场景中对输入信号进行编码处理的方法流程图。

[0045] 图9示意性地示出了本申请一些实施例中音频解码方法的步骤流程图。

[0046] 图10示意性地示出了本申请一些实施例中解码端获取高频预测信息的方法步骤流程图。

[0047] 图11示意性地示出了本申请一些实施例中通过增益补偿得到高频特征信息的方法步骤流程图。

[0048] 图12示意性地示出了本申请实施例在一应用场景中对输入码流进行解码处理的方法流程图。

[0049] 图13示意性地示出了本申请一些实施例中提供的音频编码装置的结构框图。

[0050] 图14示意性地示出了本申请一些实施例中提供的音频解码装置的结构框图。

[0051] 图15示意性地示出了适于用来实现本申请实施例的电子设备的计算机系统结构框图。

具体实施方式

[0052] 现在将参考附图更全面地描述示例实施方式。然而，示例实施方式能够以多种形式实施，且不应被理解为限于在此阐述的范例；相反，提供这些实施方式使得本申请将更加全面和完整，并将示例实施方式的构思全面地传达给本领域的技术人员。

[0053] 此外，所描述的特征、结构或特性可以以任何合适的方式结合在一个或更多实施例中。在下面的描述中，提供许多具体细节从而给出对本申请的实施例的充分理解。然而，本领域技术人员将意识到，可以实践本申请的技术方案而没有特定细节中的一个或更多，或者可以采用其它的方法、组元、装置、步骤等。在其它情况下，不详细示出或描述公知方法、装置、实现或者操作以避免模糊本申请的各方面。

[0054] 附图中所示的方框图仅仅是功能实体，不一定必须与物理上独立的实体相对应。即，可以采用软件形式来实现这些功能实体，或在一个或多个硬件模块或集成电路中实现这些功能实体，或在不同网络和/或处理器装置和/或微控制器装置中实现这些功能实体。

[0055] 附图中所示的流程图仅是示例性说明，不是必须包括所有的内容和操作/步骤，也不是必须按所描述的顺序执行。例如，有的操作/步骤还可以分解，而有的操作/步骤可以合并或部分合并，因此实际执行的顺序有可能根据实际情况改变。

[0056] 在对本申请提供的音频编码方法、音频解码方法、音频编码装置、音频解码装置、计算机可读介质以及电子设备等技术方案进行说明之前，首先对本申请技术方案中涉及的人工智能技术做简要说明。

[0057] 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。

[0058] 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。人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

[0059] 语音技术(Speech Technology)的关键技术有自动语音识别技术(ASR)和语音合成技术(TTS)以及声纹识别技术。让计算机能听、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。

[0060] 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通常包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习、式教学习等技术。

[0061] 随着人工智能技术研究和进步,人工智能技术在多个领域展开研究和应用,例如常见的智能家居、智能穿戴设备、虚拟助理、智能音箱、智能营销、无人驾驶、自动驾驶、无人机、机器人、智能医疗、智能客服等,相信随着技术的发展,人工智能技术将在更多的领域得到应用,并发挥越来越重要的价值。

[0062] 图1示意性地示出了应用本申请技术方案的示例性系统架构框图。

[0063] 如图1所示,系统架构100可以包括终端设备110、网络120和服务器130。终端设备110可以包括智能手机、平板电脑、笔记本电脑、台式电脑等各种电子设备。服务端130可以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,还可以是提供云计算服务的云服务器。网络120可以是能够在终端设备110和服务器130之间提供通信链路的各种连接类型的通信介质,例如可以是有线通信链路或者无线通信链路。

[0064] 根据实现需要,本申请实施例中的系统架构可以具有任意数目的终端设备、网络和服务器。例如,服务器130可以由多个服务器设备组成的服务器群组。另外,本申请实施例提供的技术方案可以应用于终端设备110,也可以应用于服务器130,或者可以由终端设备110和服务器130共同实施,本申请对此不做特殊限定。

[0065] 举例而言,在语音通话应用中,用户A作为语音发送端,可以通过终端设备110的麦克风采集得到模拟声音信号,通过模数转换电路将模拟声音信号转换为数字声音信号,数字声音信号经过语音编码器进行压缩,而后按照通信网络传输格式和协议打包发送至服务器130。服务器130将语音编码数据发送至作为接收端的用户B。用户B再通过终端设备110接收到数据包后解包输出语音编码压缩码流,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声音。语音编解码有效地降低语音信号传输的带宽,对于节省语音信息存储传输成本、保障通信网络传输过程中的语音信息完整性方面起了决定性作用。

[0066] 下面结合图2~图8对本申请提供的音频编码方法做出详细说明。

[0067] 图2示意性地示出了本申请一些实施例中音频编码方法的步骤流程图,该音频编

码方法可以由终端设备执行,也可以由服务器执行,或者可以由终端设备和服务器共同执行。本申请实施例以音频编码方法由终端设备执行为例进行说明。如图2所示,该音频编码方法主要可以包括如下的步骤S210~步骤S250。

[0068] 步骤S210.对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号。

[0069] 步骤S220.对待编码低频信号进行压缩编码以得到待编码低频信号的低频编码数据。

[0070] 步骤S230.根据低频信号与高频信号的相关性获取与待编码低频信号相对应的高频预测信息。

[0071] 步骤S240.对待编码高频信号进行特征提取以得到待编码高频信号的高频特征信息,并根据高频特征信息与高频预测信息的差异确定待编码高频信号的高频补偿信息。

[0072] 步骤S250.对低频编码数据以及高频补偿信息进行封装处理以得到待编码音频的音频编码数据。

[0073] 在本申请实施例提供的音频编码方法中,根据低频信号与高频信号的相关性可以利用待编码低频信号获取相应的高频预测信息,然后基于高频预测信息与待编码高频信号之间的特征差异确定高频补偿信息,在音频编码数据的传输中可以仅对高频补偿信息进行传输,从而可以极大地压缩高频信号的编码码率,降低网络传输的带宽压力。同时,在音频编码数据的接收端,可以基于高频补偿信息对高频信号进行重建还原,从而保证了高频信号的完整性和准确性,避免了因数据压缩丢失而导致的音频失真、音质差等问题。

[0074] 下面分别对以上实施例中音频编码方法的各个方法步骤做详细说明。

[0075] 在步骤S210中,对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号。

[0076] 子带分解是将原始的待编码音频由时间域转变为频率域,然后再按照频率大小将其由完整频带分解为若干个连续频段的方法,其中每一个频段即称为一个子带。

[0077] 在本申请的一些实施例中,可以获取由对应于低频频段的低通滤波器和对应于高频频段的高通滤波器组成的正交镜像滤波器组,然后通过正交镜像滤波器组对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号。

[0078] 正交镜像滤波器组(Quadrature Mirror Filter,QMF)是由两个或两个以上的滤波器通过共用输入接口或者共用输出接口的方式组合而成。本申请实施例可以将一个对应于低频频段的低通滤波器和一个对应于高频频段的高通滤波器通过共用输入接口的方式组成正交镜像滤波器组,当把待编码音频输入至该正交镜像滤波器组时,可以经过子带分解后得到由低通滤波器输出的待编码低频信号以及由高通滤波器输出的待编码高频信号。使用正交镜像滤波器组的优点是可以抵消由于子带分解而带来的频谱混叠效应。

[0079] 在步骤S220中,对待编码低频信号进行压缩编码以得到待编码低频信号的低频编码数据。

[0080] 针对子带分解得到的待编码低频信号可以通过编码器对其进行压缩编码以得到相应的低频编码数据。音频压缩编码的基本思想是通过编码器去除音频信号存在的时域冗余、频域冗余、听觉冗余以达到音频信号压缩的目的。现有的音频压缩编码方法主要可以包

括基于LTP长时预测、LPC短时预测、基音周期搜索、频带复制等方法对冗余信息进行压缩。

[0081] 在本申请的一些实施例中,可以采用CELP、SILK或者ACC等编码算法对待编码低频信号进行压缩编码。其中,CELP编码算法,即码激励线性预测编码(Code Excited Linear Prediction),是一种有效的中低码率语音压缩编码技术,它是以码本为激励源,具有码率低、合成语音质量高、抗噪能力强等优点,在4.8~16kbps码率上得到广泛应用,采用CELP技术的语音编码器有G.723、G.728、G.729、G.722.2等等。SILK编码算法是基于即时通讯软件Skype开发并向第三方开发人员和硬件制造商提供的宽带音频编码器,SILK编码算法对于音频带宽、网络带宽和算法复杂度都具有很好的弹性。ACC编码算法,即高级音频编码技术(Advanced Audio Coding),是基于MPEG-2的具有高压缩比的音频压缩算法,由于采用多声道和使用低复杂性的描述方式,ACC编码算法可以在大幅压缩音频数据的前提下提供保留较好的音质。

[0082] 在步骤S230中,根据低频信号与高频信号的相关性获取与待编码低频信号相对应的高频预测信息。

[0083] 图3示意性地示出了本申请一些实施例中编码端获取高频预测信息的方法步骤流程图。如图3所示,在以上实施例的基础上,步骤S230.根据低频信号与高频信号的相关性获取与待编码低频信号相对应的高频预测信息,主要可以包括如下的步骤S310~步骤S330。

[0084] 步骤S310.对待编码音频进行分类处理以得到待编码音频的音频类别信息。

[0085] 步骤S320.根据音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络。

[0086] 步骤S330.通过高频预测神经网络对待编码低频信号进行映射处理以得到与待编码低频信号相对应的高频预测信息。

[0087] 由于音频数据的低频信号和高频信号是具有相关性的,而神经网络(尤其是深度神经网络)可以较好地描述这里的相关性,因此本申请实施例采用神经网络将低频信号作为输入、高频信号作为目标进行预测。但是不同类型的音频数据,其高低频信号的相关性是不同的。例如浊音信号,由于存在明显的谐波结构,所以低频信号和低频信号都有相类似的频谱谐波结构;而轻音信号则没有谐波成分,其特征是中高频能量呈现块式分布,高频能量远高于低频能量;音乐信号跟不同发声乐器的发声特征有关。由此可见,针对不同类型的音频数据,其低频信号与高频信号的具体映射关系会有差异,因此本申请实施例针对这个问题提出了先对音频数据做信号分类,并且基于分类结果用不同类型的神经网络进行训练和预测,以达到更稳定、更准确的预测结果。

[0088] 本申请实施例中使用的高频预测神经网络可以采用各种不同的网络架构来实现,例如卷积神经网络(Convolutional Neural Networks,CNN)、循环神经网络(Recurrent Neural Networks,RNN)、生成对抗网络(Generative Adversarial Networks,GAN)等等。另外,在一些可选的实施方式中,也可以采用除神经网络以外的其他机器学习模型对待编码低频信号进行映射处理可以得到相应的高频预测信息,本申请对此不做特殊限定。

[0089] 图4示意性地示出了本申请一些实施例中对待编码音频进行分类处理的方法步骤流程图。如图4所示,在以上各实施例的基础上,步骤S310.对待编码音频进行分类处理以得到待编码音频的音频类别信息,可以包括如下的步骤S410~步骤S440。

[0090] 步骤S410.获取音频数据样本,并对音频数据样本进行逐帧标注以得到音频数据

样本中每个数据帧的音频类别标识。

[0091] 音频数据样本可以通过麦克风等音频输入设备采集得到的真实音频数据,也可以通过音频合成软件合成的人工构造数据。音频数据样本中包括有多种不同类别的数据,如语音浊音、语音轻音、非语音、音乐等等。以数据帧为单位可以对音频数据样本进行标注以得到每个数据帧的音频类别标识。

[0092] 步骤S420.从多个特征维度对音频数据样本进行特征提取以得到音频数据样本的多维样本特征。

[0093] 为了准确获取决定音频数据样本的类别决定因素,本步骤从多个特征维度对其进行特征提取,从而得到具有较强表征能力的多维样本特征。例如,在一些可选的实施方式中,提取得到的多维样本特征例如可以包括谱平坦特征、谱斜度特征、基音周期、MDCT(改进离散余弦变换,Modified Discrete Cosine Transform)及其一阶/二阶导数等多个维度的特征。

[0094] 步骤S430.以多维样本特征作为输入值并以音频类别标识作为与输入值相对应的目标值训练得到用于对音频数据进行分类处理的音频分类神经网络。

[0095] 音频分类神经网络的训练目标是针对一个音频数据样本,当输入其多维样本特征时能够输出得到正确的音频类别标识,在训练过程中可以通过更新神经网络的网络参数以提高其输出正确音频类别标识的预测能力。当训练到达预设的收敛目标时便可以结束训练过程,收敛目标例如可以是损失函数的误差小于误差阈值或者迭代训练的次数超过次数阈值。

[0096] 步骤S440.通过音频分类神经网络对待编码音频进行分类处理以得到待编码音频的音频类别信息。

[0097] 针对待编码音频,采用与音频数据样本相同的特征提取方式得到相应的多维音频特征,然后将提取得到的多维音频特征输入至训练完成的音频分类神经网络中,由音频分类神经网络进行特征映射后可以输出预测概率最高的音频类别信息。

[0098] 通过执行步骤S410~步骤S440,可以采用音频数据样本训练得到音频分类神经网络,从而利用音频分类神经网络对待编码音频进行准确地音频类别预测。音频分类神经网络一经训练完成,可以将其网络结构和网络参数保存在作为编码端的终端设备上或者保存在服务器上,当需要对新的待编码音频进行分类处理时,便可以直接调用该音频分类神经网络,快速且准确地输出其音频类别信息。

[0099] 针对不同类别的待编码音频,可以相应地训练并选用不同类型的高频预测神经网络,以实现基于待编码低频信号进行高频预测的方案。例如,在本申请的一些实施例中,将待编码音频分类为语音浊音、语音轻音、非语音和音乐四种类型,相应地可以训练得到四种类型的高频预测神经网络,每种高频预测神经网络专门用于对相应类别的待编码音频进行由低频信号至高频信号的信息预测。另外,本申请实施例还可以进一步根据频谱能量分布对待编码音频细分为更多的音频类别,类别划分越细致,相应的高频预测神经网络将具有更加准确的表征和预测能力。

[0100] 在训练高频预测神经网络时,涉及音频数据样本的频段分割和特征提取两个预处理过程。图5和图6分别示出了两种基于不同预处理顺序训练高频预测神经网络的方法。

[0101] 图5示意性地示出了本申请一些实施例中基于先特征提取后频段分割的预处理过

程训练高频预测神经网络的方法步骤流程图。如图5所示,在以上实施例的基础上,步骤S320.根据音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络,主要可以包括如下的步骤S510~步骤S530。

[0102] 步骤S510.获取与音频类别信息相对应的音频数据样本,并对音频数据样本进行压缩变换以得到音频数据样本的频谱特征样本。

[0103] 通过压缩变换的方式对音频数据样本进行特征提取可以得到音频数据样本的频谱特征样本,压缩变换的方法例如可以采用改进离散余弦变换(Modified Discrete Cosine Transform,MDCT)。MDCT是一种线性正交交叠变换算法,使用了一种时域混叠抵消技术(Time Domain Alias Cancellation,TDAC),包含50%的时域交叠窗,在不降低编码性能的情况下能够有效地克服由边缘效应产生的周期化噪声。在本申请的另一些实施例中,也可以采用离散傅里叶变换(Discrete Fourier Transform,DFT)来对音频数据样本进行压缩变换。

[0104] 步骤S520.对频谱特征样本按照频点的数值进行划分以得到低频特征样本和高频特征样本。

[0105] 以MDCT算法为例,由步骤S510对音频数据样本进行压缩变换后得到的频谱特征样本是全带数据,可以根据对应的物理频点值划分为高频MDCT部分和低频MDCT部分,例如包括1024个频点的音频数据样本经过MDCT处理后可以得到512个频点的频谱特征样本,其中1~256点数据可以划分为低频特征样本,而257~512点数据可以划分为高频特征样本。

[0106] 步骤S530.以低频特征样本作为输入值并以高频特征样本作为与输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0107] 高频预测神经网络的训练目标是通过更新网络参数以提高对低频信号与高频信号之间相关性的表征和预测能力。高频预测神经网络可以选用CNN、RNN、GAN等各种类型的网络架构。另外,本申请实施例也可以训练除神经网络以外的其他机器学习模型,用以根据低频信号对高频信号进行预测,本申请对此不做特殊限定。

[0108] 通过执行步骤S510~步骤S530,可以先对全频段的音频数据样本进行压缩变换,然后再基于频点数值进行频段分割得到低频特征样本和高频特征样本,这种预处理方案的优点在于,针对每个音频数据样本仅需要进行一次压缩变换和频段分割操作,能够降低数据计算成本,提高样本的处理效率。

[0109] 图6示意性地示出了本申请一些实施例中基于先频段分割后特征提取的预处理过程训练高频预测神经网络的方法步骤流程图。如图6所示,在以上实施例的基础上,步骤S320.根据音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络,主要可以包括如下的步骤S610~步骤S630。

[0110] 步骤S610.获取与音频类别信息相对应的音频数据样本,并将音频数据样本按照所在频段的高低分解为低频数据样本和高频数据样本。

[0111] 步骤S620.分别对低频数据样本和高频数据样本进行压缩变换以得到相应的低频特征样本和高频特征样本。

[0112] 步骤S630.以低频特征样本作为输入值并以高频特征样本作为与输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0113] 本申请实施例与图5对应实施例的不同之处在于,本申请实施例采用了先对音频

数据样本进行频段分割,然后再进行压缩变换的预处理方案。与图5对应实施例相比,本申请实施例的预处理方案针对每个音频数据样本都增加了一次压缩变换过程,即针对低频数据样本和低频数据样本需要分别进行一次压缩变换。这种预处理方案的优点在于,保持了高频预测神经网络在训练过程和使用过程的一致性,能够在一定程度上提高高频信号预测的准确性。

[0114] 基于以上实施例训练得到高频预测神经网络后,可以将高频预测神经网络的网络结构和网络参数保存在编码端以及解码端的终端设备上或者保存在服务器上,当需要对低频信号进行高频预测时,便可以调用该高频预测神经网络。

[0115] 在本申请的一些实施例中,可以在步骤S330中直接通过高频预测神经网络对待编码低频信号进行映射处理得到相应的高频预测信号。而在另一些实施例中,为了保持编码端与解码端的数据同步性,可以先对低频编码数据(由待编码低频信号压缩编码得到)进行解码处理以得到与待编码低频信号相对应的低频解码信号,然后再通过高频预测神经网络对低频解码信号进行映射处理以得到与待编码低频信号相对应的高频预测信息。通过对待编码低频信号进行一次编码和解码操作后,再进行高频预测的方案能够保持编码端和解码端的操作一致性,从而提高高频预测的准确性。

[0116] 在本申请的一些实施例中,通过高频预测神经网络对低频解码信号进行映射处理的方法可以包括:对低频解码信号进行压缩变换以得到低频解码信号的低频频谱特征,然后通过高频预测神经网络对低频频谱特征进行映射处理以得到与待编码低频信号相对应的高频预测信息。其中,对低频解码信号进行压缩变换的方法例如可以是改进离散余弦变换MDCT或者其他形式的离散傅里叶变换DFT。

[0117] 在步骤S240中,对待编码高频信号进行特征提取以得到待编码高频信号的高频特征信息,并根据高频特征信息与高频预测信息的差异确定待编码高频信号的高频补偿信息。

[0118] 对待编码高频信号进行特征提取的方法可以是与待编码低频信号(或者低频解码信号)相同的压缩变换的方法,即对待编码高频信号进行压缩变换以得到待编码高频信号的高频特征信息。采用相同的特征提取方法,可以使高频特征信息与高频预测信息具有一致的特征属性,从而方便确定二者之间的特征差异。基于二者之间的特征差异可以确定待编码高频信号的高频补偿信息,图7示意性地示出了本申请一些实施例中确定高频补偿信息的方法步骤流程图。如图7所示,在以上各实施例的基础上,步骤S240中的根据高频特征信息与高频预测信息的差异确定待编码高频信号的高频补偿信息,主要可以包括如下的步骤S710~步骤S730。

[0119] 步骤S710.将高频特征信息从线性频率域映射至临界频带域以得到与高频特征信息相对应的特征频谱信息。

[0120] 步骤S720.将高频预测信息从线性频率域映射至临界频带域以得到与高频预测信息相对应的预测频谱信息。

[0121] 步骤S730.根据特征频谱信息和预测频谱信息的差异确定待编码高频信号的高频补偿信息。

[0122] 临界频带域,即Bark域,是听觉学和心理声学的专业名词。临界频带指的是由于听觉传感器(如人耳结构中的耳蜗)的构造产生的听觉滤波器的频率带宽。概况地说,临界

频带是声音频率带,在临界频带中第一个单音感知性会被第二单音的听觉掩蔽所干扰。在声学研究中,人们使用听觉滤波器来模拟不同的临界频带,人耳结构一般会对24个频率点产生共振,因此在临界频带域中的音频信号也呈现出24个临界频带,分别从1到24。Bark域相对于线性域更符合人耳声学频率的感知特性,而且子带数比较少,有利于编码压缩。

[0123] 为了实现对特征频谱信息和预测频谱信息之间差异的量化表征,本申请实施例可以分别对特征频谱信息和预测频谱信息进行对数变换以得到特征频谱对数值和预测频谱对数值,然后根据特征频谱对数值和预测频谱对数值的差值查询增益码表以得到增益量化值,并将增量量化值确定为待编码高频信号的高频补偿信息。其中,增益码表是一个大小为N的数值依次递增的量化表格,基于增益码表可以查询得到增益量化值。

[0124] 举例而言,在对高频特征信息或者高频预测信息从线性频率域映射至临界频带域后,可以得到相应的Bark域频谱信息 $E(k)$,然后可以对其做对数变换得到相应的频谱对数值 $20 \cdot \log_{10}(E(k)^2)$,从而确定特征频谱对数值和预测频谱对数值之间的差值 $\Delta E(k)$,其中k表示高频子带序号。

[0125] 通过查询增益码表对差值 $\Delta E(k)$ 进行数值量化的查询逻辑如下:

```

Index = 0;
For i=0:N-1
    If Table[i] <=  $\Delta E(k)$  && Table[i+1] >  $\Delta E(k)$ 
[0126]     Index = i;
    End
End
End

```

[0127] 其中,Table为数值递增的增益码表;N为增益码表的大小,表示增益码表中包括0~N-1共计N个量化值;Index为最终量化得到的增益量化值。

[0128] 基于以上查询逻辑可知,在获取到特征频谱对数值和预测频谱对数值之间的差值 $\Delta E(k)$ 之后,可以遍历增益码表中的各个量化值,比较差值 $\Delta E(k)$ 与第i个量化编码Table[i]以及第i+1个量化编码Table[i+1]之间的数值大小。若差值 $\Delta E(k)$ 大于或等于第i个量化编码Table[i],并且差值 $\Delta E(k)$ 小于第i+1个量化编码Table[i+1],那么可以确定增益量化值为i。

[0129] 通过增益码表对增益补偿进行量化处理,可以将原本连续的增益补偿信息离散化,降低对高频信号部分进行编码传输的计算量。

[0130] 在步骤S250中,对低频编码数据以及高频补偿信息进行封装处理以得到待编码音频的音频编码数据。

[0131] 封装处理是对各种编码内容组合形成指定的音频文件的过程,例如,封装得到的音频编码数据可以是MP3、AAC、WAV、FLAC、DSD等各种格式的音频文件。

[0132] 在本申请的一些实施例中,在进行音频编码时,首先对待编码音频进行了分类处理以得到相应的音频类别信息,然后基于音频类别信息选用不同类型的高频预测神经网络来对待编码低频信号进行针对性地高频预测。在此基础上,为了能够在解码端对高频部分的信号进行准确地预测和重建,在步骤S250中同样可以获取分类处理得到的音频类别信

息,然后将音频类别信息、低频编码数据以及高频补偿信息共同进行封装处理以得到待编码音频的音频编码数据,从而将音频类别信息一起传输至解码端。

[0133] 图8示意性地示出了本申请实施例在一应用场景中对输入信号进行编码处理的方法流程图。如图8所示,在该应用场景中由编码端对输入信号进行音频编码的方法主要可以包括如下的步骤S801~步骤S811。

[0134] 步骤S801.对输入信号进行信号分类得到信号分类结果,这里的分类类别可以包括四个类型:语音浊音、语音轻音、非语音和音乐。信号分类结果可以指导编解码的神经网络的选用。其中,每种信号类型将对应一个神经网络,而神经网络也是通过大量音频数据经过分类后,每个神经网络用相应的同类型数据进行独立训练的,在实际编解码中用到的是已经训练好的神经网络。

[0135] 步骤S802.输入信号经过QMF(正交镜像滤波器组)进行高低频信号分解,采用QMF的优点是可以抵消由于子带分割带来的混叠效应。通过QMF将输入信号分解为低频信号和高频信号。

[0136] 步骤S803.由步骤S802分解得到的低频信号将经过语音编码器进行压缩编码得到相应的低频信号的低频编码参数,本步骤使用的语音编码器可以是基于CELP、SILK、AAC等算法的编码器。

[0137] 步骤S804.为了让神经网络的输入在编码器和解码器中同步,对低频信号编码后的码流做一次语音解码得到还原的低频信号。

[0138] 步骤S805.由步骤S804进行语音解码还原得到的低频信号经过MDCT(Modified Discrete Cosine Transform,改进离散余弦变换)得到相关的低频频谱信息。

[0139] 步骤S806.将MDCT变换得到的低频频谱信息输入至步骤S801中根据信号分类结果选取的神经网络,通过该神经网络进行预测得到高频频谱预测信息。

[0140] 步骤S807.为了与人耳声学感知频带相对应,这里将高频频谱预测信息由线性频域转换到Bark域得到Bark域高频频谱预测值(可以以对数形式表示)。

[0141] 步骤S808.由步骤S802中通过QMF分解得到的真实高频信号经过MDCT变换得到相关的高频频谱信息。

[0142] 步骤S809.对步骤S808得到的高频频谱信息进行Bark域转换得到Bark域高频频谱真实值(可以以对数形式表示)。

[0143] 步骤S810.将步骤S809中得到的Bark域高频频谱真实值与步骤S807中得到的Bark域高频频谱预测值相减后得到子带增益补偿值并进一步增益量化得到高频编码参数。

[0144] 步骤S811.将步骤S801得到的信号分类结果、步骤S803中得到的低频编码参数以及步骤S810中得到的高频编码参数进行封装处理后组成编码参数进行输出。

[0145] 由以上编码流程获得的编码参数可以通过网络传输至其他作为音频数据接收端的终端设备或者服务器,以使接收端对其进行解码处理得到解码信号。

[0146] 下面结合图9~图12对本申请提供的音频解码方法做出详细说明。

[0147] 图9示意性地示出了本申请一些实施例中音频解码方法的步骤流程图,该音频解码方法可以由终端设备执行,也可以由服务器执行,或者可以由终端设备和服务器共同执行。本申请实施例以音频解码方法由终端设备执行为例进行说明。如图9所示,该音频解码方法主要可以包括如下的步骤S910~步骤S950。

[0148] 步骤S910.对待解码的音频编码数据进行封装解析以得到音频编码数据中的低频编码数据和高频补偿信息。

[0149] 步骤S920.对低频编码数据进行解码处理以得到还原低频信号。

[0150] 步骤S930.根据低频信号与高频信号的相关性获取与还原低频信号相对应的高频预测信息。

[0151] 步骤S940.根据高频补偿信息对高频预测信息进行增益补偿以得到高频特征信息,并对高频特征信息进行特征还原以得到还原高频信号。

[0152] 步骤S950.对还原低频信号和还原高频信号进行子带合成以得到音频编码数据的还原音频。

[0153] 在本申请实施例提供的音频解码方法中,根据低频信号与高频信号的相关性可以利用解码得到的还原低频信号获取相应的高频预测信息,然后利用高频补偿信息对高频预测信息进行增益补偿以得到高频特征信息,在对高频特征信息进行特征还原后可以得到预测重建的还原高频信号。该音频解码方法通过在编码端和解码端使用相同的高频信号预测方法和高频增益补偿方法,可以保证信号传输过程中高频信号的完整性和准确性,避免了因数据压缩丢失而导致的音频失真、音质差等问题。

[0154] 下面分别对以上实施例中音频解码方法的各个方法步骤做详细说明。

[0155] 在步骤S910中,对待解码的音频编码数据进行封装解析以得到音频编码数据中的低频编码数据和高频补偿信息。

[0156] 待解码的音频编码数据可以由一个一个连续的码流单元组成,每两个相邻的码流单元之间通过码流单元分隔信息来进行分隔。举例而言,在采用AAC编码标准(Advanced Audio Coding)时,音频编码数据由多个连续的ADTS单元(Audio Data Transport Stream)组成,每个ADTS单元作为一个音频内容的封装单元。每两个ADTS单元之间通过同步字(syncword)进行分隔,同步字可以是0xFFF(二进制“111111111111”)。

[0157] 在本申请的一些实施方式中,待解码的音频编码数据进行封装解析的方法可以包括:首先在待解码的音频编码数据中搜索码流单元分隔信息;然后根据搜索到的码流单元分隔信息从音频编码数据中分离得到待解码的码流单元;再对码流单元进行字段解析以得到码流单元中封装的低频编码数据和高频补偿信息。以AAC编码标准为例,解码端在接收到待解码的音频编码数据后,可以在原始码流中搜索字段0x0FFF,以该字段为间隔可以分离得到ADTS单元,再对ADTS单元进行字段解析可以获得封装在其中的低频编码数据和高频补偿信息。

[0158] 通过对音频编码数据进行封装解析可以获取其中包括的对应于低频信号部分的低频编码数据和对应于高频信号部分的高频补偿信息。除此之外,如果在编码端对音频数据进行了分类处理,那么本步骤进行封装解析后还可以获得其中包括的音频类别信息,从而根据音频类别信息选用与编码端一致的处理方案。

[0159] 在步骤S920中,对低频编码数据进行解码处理以得到还原低频信号。

[0160] 针对封装解析得到的低频编码数据可以通过解码器对其进行解码处理以得到相应的还原低频信号。本步骤中使用的解码器与编码端使用的编码器相对应。例如,如果编码端使用CELP算法进行压缩编码,那么本步骤也将采用相应的CELP算法进行解码处理;如果编码端使用SILK或者ACC等算法进行压缩编码,那么本步骤也将采用相应的SILK或者ACC等

算法进行解码处理。

[0161] 在步骤S930中,根据低频信号与高频信号的相关性获取与还原低频信号相对应的高频预测信息。

[0162] 图10示意性地示出了本申请一些实施例中解码端获取高频预测信息的方法步骤流程图。如图10所示,在以上实施例的基础上,步骤S930.根据低频信号与高频信号的相关性获取与还原低频信号相对应的高频预测信息,主要可以包括如下的步骤S1010~步骤S1030。

[0163] 步骤S1010.对音频编码数据进行封装解析以得到音频编码数据中的音频类别信息。

[0164] 步骤S1020.根据音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络。

[0165] 步骤S1030.通过高频预测神经网络对还原低频信号进行映射处理以得到与还原低频信号相对应的高频预测信息。

[0166] 通过对音频编码数据进行封装解析可以获得编码端对音频数据进行分类处理得到的音频类别信息。该音频类别信息在编码端用于指导高频预测神经网络的选用,而在本申请实施例适用的解码端可以基于该音频类别信息选用与编码端相同的高频预测神经网络,从而确保解码端与编码端能够保持高频信号预测的一致性。

[0167] 在本申请的一些实施例中,可以在编码端对高频预测神经网络进行训练,训练得到高频预测神经网络后可以将其网络结构和网络参数保存在编码端,同时也可以将相关数据传送至解码端。如此一来,解码端基于接收到的网络结构加载网络参数后可以得到与编码端一致的高频预测神经网络。

[0168] 在本申请的另一实施例中,也可以在解码端对高频预测神经网络进行训练,训练得到高频预测神经网络后可以将其网络结构和网络参数保存在解码端,同时也可以将相关数据传送至编码端,以使得编码端和解码端能够使用相同的高频预测神经网络对高频信号进行预测。解码端训练网络预测神经网络的方法与编码端相似或者相同,可以参考图5和图6中的相关方法步骤,此处不再赘述。

[0169] 在本申请的另一实施例中,也可以在服务器上对高频预测神经网络进行训练,训练得到高频预测神经网络后可以将其网络结构和网络参数保存在服务器上,并同时可以由服务器将相关数据传送至编码端和解码端,使得编码端和解码端能够使用相同的高频预测神经网络对高频信号进行预测。

[0170] 在解码端通过高频预测神经网络对还原低频信号进行映射处理以实现高频信号预测时,可以先对还原低频信号进行压缩变换以得到还原低频信号的低频频谱特征,然后再通过高频预测神经网络对低频频谱特征进行映射处理以得到与还原低频信号相对应的高频预测信息。其中,对还原低频信号进行压缩变换的方法例如可以是改进离散余弦变换MDCT或者其他形式的离散傅里叶变换DFT。

[0171] 在步骤S940中,根据高频补偿信息对高频预测信息进行增益补偿以得到高频特征信息,并对高频特征信息进行特征还原以得到还原高频信号。

[0172] 在解码端根据高频补偿信息对高频预测信息进行增益补偿以得到高频特征信息的方法与编码端根据高频特征信息与高频预测信息的差异确定待编码高频信号的高频补

偿信息的方法为两个相反的流程。同时,在解码端对高频特征信息进行特征还原的方法与编码端对待编码高频信号进行特征提取的方法也是两个相反的流程。

[0173] 图11示意性地示出了本申请一些实施例中通过增益补偿得到高频特征信息的方法步骤流程图。如图11所示,在以上各实施例的基础上,步骤S940中的根据高频补偿信息对高频预测信息进行增益补偿以得到高频特征信息,主要可以包括如下的步骤S1110~步骤S1130。

[0174] 步骤S1110.将高频预测信息从线性频率域映射至临界频带域以得到与高频预测信息相对应的预测频谱信息。

[0175] 步骤S1120.根据高频补偿信息对预测频谱信息进行增益补偿以得到特征频谱信息。

[0176] 步骤S1130.将特征频谱信息从临界频带域映射至线性频率域以得到与特征频谱信息相对应的高频特征信息。

[0177] 为了获取符合人耳声学频率的感知特性的特征信息并且使用较少的子带数进行编码压缩,在编码端会进行从线性频率域向临界频带域的映射变换。与之相适应地,在解码端通过高频补偿信息对预测频谱信息进行增益补偿后需要再将得到的特征频谱信息从临界频带域映射回到线性频率域,从而得到线性频率域下的高频特征信息,以便于在线性频率域下对高频特征信息进行特征还原。

[0178] 在编码端对特征频谱信息和预测频谱信息之间的差异进行量化表征时可以使用对数变换后的对数值进行计算。与之相适应地,本申请实施例在解码端对预测频谱信息进行增益补偿时,可以首先对预测频谱信息进行对数变换以得到预测频谱对数值,然后根据高频补偿信息对预测频谱对数值进行增益补偿以得到特征频谱对数值,再对特征频谱对数值进行指数还原便可以得到特征频谱信息。其中,指数还原的方法与对数变换的方法互为逆过程。

[0179] 在步骤S940中,通过Bark域变换得到与特征频谱信息相对应的高频特征信息后,可以再对其进行特征还原以得到还原高频信号。编码端可以采用压缩变换的方式对待编码高频信号进行特征提取,与之相适应地,解码端可以采用解压缩变换对高频特征信息进行特征还原。例如,在编码端使用改进离散余弦变换(Modified Discrete Cosine Transform, MDCT)进行特征提取,那么在解码端可以相应地适用改进离散余弦逆变换(Inverse Modified Discrete Cosine Transform, IMDCT)进行特征还原。

[0180] 在步骤S950中,对还原低频信号和还原高频信号进行子带合成以得到音频编码数据的还原音频。

[0181] 解码端的子带合成是编码端子带分解的逆过程,用于将多个不同频段的信号整合至一个完整频段中。在本申请的一些实施例中,可以获取由对应于低频频段的低通滤波器 and 对应于高频频段的高通滤波器组成的正交镜像滤波器组,然后通过正交镜像滤波器组对还原低频信号和还原高频信号进行子带合成以得到音频编码数据的还原音频。

[0182] 正交镜像滤波器组(Quadrature Mirror Filter, QMF)是由两个或两个以上的滤波器通过共用输入接口或者共用输出接口的方式组合而成。本申请实施例可以将一个对应于低频频段的低通滤波器和一个对应于高频频段的高通滤波器通过共用输出接口的方式组成正交镜像滤波器组,当把还原低频信号输入至低通滤波器并把还原高频信号输入至高

通滤波器时,可以经过子带合成后得到该正交镜像滤波器组输出的完整频段下的还原音频。

[0183] 图12示意性地示出了本申请实施例在一应用场景中对输入码流进行解码处理的方法流程图。如图12所示,在该应用场景中由解码端对输入码流进行音频解码的方法主要可以包括如下的步骤S1201~步骤S1207。

[0184] 步骤S1201.对接收到的输入码流进行封装解析,分别得到每个数据帧对应的低频语音编码参数、高频增益补偿参数以及信号分类参数。其中,信号分类参数用于选定当前数据帧使用的神经网络。

[0185] 步骤S1202.对步骤S1201中封装解析得到的低频语音编码参数经过与编码端对应的解码器进行解码处理得到低频信号。

[0186] 步骤S1203.低频信号经过MDCT变换得到低频频谱信息。

[0187] 步骤S1204.将步骤S1203变换得到的低频频谱信息输入至步骤S1201中根据信号分类参数选取的神经网络,由神经网络输出预测到的高频线性频谱信息。

[0188] 步骤S1205.将步骤S1204得到的高频线性频谱信息转换到Bark域,并通过步骤S1201中封装解析得到的高频增益补偿参数进行Bark子带频谱能量调整,经过调整后再由Bark域转换回线性域得到高频频谱信息。

[0189] 步骤S1206.对步骤S1205得到的高频频谱信息经IMDCT变换得到重建的高频信号。

[0190] 步骤S1207.将步骤S1202得到的低频信号和步骤S1206得到的高频信号通过QMF合成滤波器合成为全带解码信号并对其进行输出。

[0191] 本申请实施例提供的音频编解码方法通过神经网络预测音频子带编解码的方式,提升了高频信号的预测能力,从而进一步压缩高频编码码率,同时本申请实施例可以通过对输入信号进行分类,在不同类别下对应不同的神经网络,因此本申请提供的技术方案不仅适用于谐波结构信号,也适用于其它类型的信号,能够较好地实现不同输入信号的高频信号预测拟合。

[0192] 应当注意,尽管在附图中以特定顺序描述了本申请中方法的各个步骤,但是,这并非要求或者暗示必须按照该特定顺序来执行这些步骤,或是必须执行全部所示的步骤才能实现期望的结果。附加的或备选的,可以省略某些步骤,将多个步骤合并为一个步骤执行,以及/或者将一个步骤分解为多个步骤执行等。

[0193] 以下介绍本申请的装置实施例,可以用于执行本申请上述实施例中的音频编码方法及音频解码方法。

[0194] 图13示意性地示出了本申请一些实施例中提供的音频编码装置的结构框图。如图13所示,音频编码装置1300主要可以包括:音频分解模块1310、低频编码模块1320、高频预测模块1330、高频补偿模块1340和编码封装模块1350。

[0195] 音频分解模块1310,被配置为对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段段的待编码高频信号。

[0196] 低频编码模块1320,被配置为对待编码低频信号进行压缩编码以得到待编码低频信号的低频编码数据。

[0197] 高频预测模块1330,被配置为根据低频信号与高频信号的相关性获取与待编码低频信号相对应的高频预测信息。

[0198] 高频补偿模块1340,被配置为对待编码高频信号进行特征提取以得到待编码高频信号的高频特征信息,并根据高频特征信息与高频预测信息的差异确定待编码高频信号的高频补偿信息。

[0199] 编码封装模块1350,被配置为对低频编码数据以及高频补偿信息进行封装处理以得到待编码音频的音频编码数据。

[0200] 在本申请的一些实施例中,基于以上各实施例,高频预测模块1330包括:音频分类单元,被配置为对待编码音频进行分类处理以得到待编码音频的音频类别信息;编码端网络获取单元,被配置为根据音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络;编码端网络映射单元,被配置为通过高频预测神经网络对待编码低频信号进行映射处理以得到与待编码低频信号相对应的高频预测信息。

[0201] 在本申请的一些实施例中,基于以上各实施例,音频分类单元包括:样本类别标注子单元,被配置为获取音频数据样本,并对音频数据样本进行逐帧标注以得到音频数据样本中每个数据帧的音频类别标识;样本特征提取子单元,被配置为从多个特征维度对音频数据样本进行特征提取以得到音频数据样本的多维样本特征;分类网络训练子单元,被配置为以多维样本特征作为输入值并以音频类别标识作为与输入值相对应的目标值训练得到用于对音频数据进行分类处理的音频分类神经网络;分类网络处理子单元,被配置为通过音频分类神经网络对待编码音频进行分类处理以得到待编码音频的音频类别信息。

[0202] 在本申请的一些实施例中,基于以上各实施例,编码端网络获取单元包括:第一样本变换子单元,被配置为获取与音频类别信息相对应的音频数据样本,并对音频数据样本进行压缩变换以得到音频数据样本的频谱特征样本;第一频段划分子单元,被配置为对频谱特征样本按照频点的数值进行划分以得到低频特征样本和高频特征样本;第一网络获取子单元,被配置为以低频特征样本作为输入值并以高频特征样本作为与输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0203] 在本申请的一些实施例中,基于以上各实施例,编码端网络获取单元包括:第二段划分子单元,被配置为获取与音频类别信息相对应的音频数据样本,并将音频数据样本按照所在频段的高低分解为低频数据样本和高频数据样本;第二样本变换子单元,被配置为分别对低频数据样本和高频数据样本进行压缩变换以得到相应的低频特征样本和高频特征样本;第二网络获取子单元,被配置为以低频特征样本作为输入值并以高频特征样本作为与输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0204] 在本申请的一些实施例中,基于以上各实施例,编码端网络映射单元包括:编码端低频解码子单元,被配置为对低频编码数据进行解码处理以得到与待编码低频信号相对应的低频解码信号;编码端低频映射子单元,被配置为通过高频预测神经网络对低频解码信号进行映射处理以得到与待编码低频信号相对应的高频预测信息。

[0205] 在本申请的一些实施例中,基于以上各实施例,编码端低频映射子单元包括:编码端压缩变换子单元,被配置为对低频解码信号进行压缩变换以得到低频解码信号的低频频谱特征;编码端特征映射子单元,被配置为通过高频预测神经网络对低频频谱特征进行映射处理以得到与待编码低频信号相对应的高频预测信息。

[0206] 在本申请的一些实施例中,基于以上各实施例,音频分解模块1310包括:滤波器获

取单元,被配置为获取由对应于低频频段的低通滤波器和对应于高频频段的高通滤波器组成的正交镜像滤波器组;子带分解单元,被配置为通过正交镜像滤波器组对待编码音频进行子带分解以得到对应于低频频段的待编码低频信号和对应于高频频段的待编码高频信号。

[0207] 在本申请的一些实施例中,基于以上各实施例,高频补偿模块1340包括:高频压缩变换单元,被配置为对待编码高频信号进行压缩变换以得到待编码高频信号的高频特征信息。

[0208] 在本申请的一些实施例中,基于以上各实施例,高频补偿模块1340还包括:特征频谱转换单元,被配置为将高频特征信息从线性频率域映射至临界频带域以得到与高频特征信息相对应的特征频谱信息;预测频谱转换单元,被配置为将高频预测信息从线性频率域映射至临界频带域以得到与高频预测信息相对应的预测频谱信息;补偿信息确定单元,被配置为根据特征频谱信息和预测频谱信息的差异确定待编码高频信号的高频补偿信息。

[0209] 在本申请的一些实施例中,基于以上各实施例,补偿信息确定单元包括:第一对数变换子单元,被配置为分别对特征频谱信息和预测频谱信息进行对数变换以得到特征频谱对数值和预测频谱对数值;增益量化子单元,被配置为根据特征频谱对数值和预测频谱对数值的差值查询增益码表以得到增益量化值,并将增量量化值确定为待编码高频信号的高频补偿信息。

[0210] 在本申请的一些实施例中,基于以上各实施例,编码封装模块1350包括:音频分类单元,被配置为对待编码音频进行分类处理以得到待编码音频的音频类别信息;编码封装单元,被配置为对音频类别信息、低频编码数据以及高频补偿信息进行封装处理以得到待编码音频的音频编码数据。

[0211] 图14示意性地示出了本申请一些实施例中提供的音频解码装置的结构框图。如图14所示,音频解码装置1400主要可以包括:封装解析模块1410、低频解码模块1420、高频预测模块1430、高频还原模块1440以及音频合成模块1450。

[0212] 封装解析模块1410,被配置为对待解码的音频编码数据进行封装解析以得到音频编码数据中的低频编码数据和高频补偿信息。

[0213] 低频解码模块1420,被配置为对低频编码数据进行解码处理以得到还原低频信号。

[0214] 高频预测模块1430,被配置为根据低频信号与高频信号的相关性获取与还原低频信号相对应的高频预测信息。

[0215] 高频还原模块1440,被配置为根据高频补偿信息对高频预测信息进行增益补偿以得到高频特征信息,并对高频特征信息进行特征还原以得到还原高频信号。

[0216] 音频合成模块1450,被配置为对还原低频信号和还原高频信号进行子带合成以得到音频编码数据的原始音频。

[0217] 在本申请的一些实施例中,基于以上各实施例,高频预测模块1430包括:类别获取单元,被配置为对音频编码数据进行封装解析以得到音频编码数据中的音频类别信息;解码端网络获取单元,被配置为根据音频类别信息获取基于低频信号与高频信号的相关性训练得到的高频预测神经网络;解码端网络映射单元,被配置为通过高频预测神经网络对还原低频信号进行映射处理以得到与还原低频信号相对应的高频预测信息。

[0218] 在本申请的一些实施例中,基于以上各实施例,解码端网络获取单元包括:第一样本变换子单元,被配置为获取与音频类别信息相对应的音频数据样本,并对音频数据样本进行压缩变换以得到音频数据样本的频谱特征样本;第一频段划分子单元,被配置为对频谱特征样本按照频点的数值进行划分以得到低频特征样本和高频特征样本;第一网络获取子单元,被配置为以低频特征样本作为输入值并以高频特征样本作为与输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0219] 在本申请的一些实施例中,基于以上各实施例,解码端网络获取单元包括:第二段划分子单元,被配置为获取与音频类别信息相对应的音频数据样本,并将音频数据样本按照所在频段的高低分解为低频数据样本和高频数据样本;第二样本变换子单元,被配置为分别对低频数据样本和高频数据样本进行压缩变换以得到相应的低频特征样本和高频特征样本;第二网络获取子单元,被配置为以低频特征样本作为输入值并以高频特征样本作为与输入值相对应的目标值训练得到用于表征低频信号与高频信号的相关性的高频预测神经网络。

[0220] 在本申请的一些实施例中,基于以上各实施例,解码端网络映射单元包括:解码端压缩变换子单元,被配置为对还原低频信号进行压缩变换以得到还原低频信号的低频频谱特征;解码端特征映射子单元,被配置为通过高频预测神经网络对低频频谱特征进行映射处理以得到与还原低频信号相对应的高频预测信息。

[0221] 在本申请的一些实施例中,基于以上各实施例,高频还原模块1440包括:频谱信息预测单元,被配置为将高频预测信息从线性频率域映射至临界频带域以得到与高频预测信息相对应的预测频谱信息;频谱信息补偿单元,被配置为根据高频补偿信息对预测频谱信息进行增益补偿以得到特征频谱信息;特征信息确定单元,被配置为将特征频谱信息从临界频带域映射至线性频率域以得到与特征频谱信息相对应的高频特征信息。

[0222] 在本申请的一些实施例中,基于以上各实施例,频谱信息补偿单元包括:第二对数变换子单元,被配置为对预测频谱信息进行对数变换以得到预测频谱对数值;对数值补偿子单元,被配置为根据高频补偿信息对预测频谱对数值进行增益补偿以得到特征频谱对数值;指数还原子单元,被配置为对特征频谱对数值进行指数还原以得到特征频谱信息。

[0223] 在本申请的一些实施例中,基于以上各实施例,高频还原模块还包括:特征信息解压缩单元,被配置为对高频特征信息进行解压缩变换以得到还原高频信号。

[0224] 在本申请的一些实施例中,基于以上各实施例,音频合成模块1450包括:滤波器获取单元,被配置为获取由对应于低频频段的低通滤波器 and 对应于高频频段的高通滤波器组成的正交镜像滤波器组;子带合成单元,被配置为通过正交镜像滤波器组对还原低频信号和还原高频信号进行子带合成以得到音频编码数据的还原音频。

[0225] 在本申请的一些实施例中,基于以上各实施例,所述封装解析模块1410包括:码流搜索单元,被配置为在待解码的音频编码数据中搜索码流单元分隔信息;码流分离单元,被配置为根据搜索到的码流单元分隔信息从音频编码数据中分离得到待解码的码流单元;码流解析单元,被配置为对码流单元进行字段解析以得到码流单元中封装的低频编码数据和高频补偿信息。

[0226] 本申请各实施例中提供的音频编码装置以及音频解码装置的具体细节已经在对应的方法实施例中进行了详细的描述,此处不再赘述。

[0227] 图15示意性地示出了用于实现本申请实施例的电子设备的计算机系统结构框图。

[0228] 需要说明的是,图15示出的电子设备的计算机系统1500仅是一个示例,不应对本申请实施例的功能和使用范围带来任何限制。

[0229] 如图15所示,计算机系统1500包括中央处理单元(Central Processing Unit, CPU) 1501,其可以根据存储在只读存储器(Read-Only Memory, ROM) 1502中的程序或者从存储部分1508加载到随机访问存储器(Random Access Memory, RAM) 1503中的程序而执行各种适当的动作和处理。在RAM 1503中,还存储有系统操作所需的各种程序和数据。CPU 1501、ROM 1502以及RAM 1503通过总线1504彼此相连。输入/输出(Input/Output, I/O)接口1505也连接至总线1504。

[0230] 以下部件连接至I/O接口1505:包括键盘、鼠标等的输入部分1506;包括诸如阴极射线管(Cathode Ray Tube, CRT)、液晶显示器(Liquid Crystal Display, LCD)等以及扬声器等的输出部分1507;包括硬盘等的存储部分1508;以及包括诸如LAN(Local Area Network, 局域网)卡、调制解调器等的网络接口卡的通信部分1509。通信部分1509经由诸如因特网的网络执行通信处理。驱动器1510也根据需要连接至I/O接口1505。可拆卸介质1511,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器1510上,以便于从其上读出的计算机程序根据需要被安装入存储部分1508。

[0231] 特别地,根据本申请的实施例,各个方法流程图中所描述的过程可以被实现为计算机软件程序。例如,本申请的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分1509从网络上被下载和安装,和/或从可拆卸介质1511被安装。在该计算机程序被中央处理单元(CPU) 1501执行时,执行本申请的系统中限定的各种功能。

[0232] 需要说明的是,本申请实施例所示的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(Erasable Programmable Read Only Memory, EPROM)、闪存、光纤、便携式紧凑磁盘只读存储器(Compact Disc Read-Only Memory, CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本申请中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本申请中,计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:无线、有线等等,或者上述的任意合适的组合。

[0233] 附图中的流程图和框图,图示了按照本申请各种实施例的系统、方法和计算机程

序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的是,框图或流程图中的每个方框、以及框图或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0234] 应当注意,尽管在上文详细描述中提及了用于动作执行的设备的若干模块或者单元,但是这种划分并非强制性的。实际上,根据本申请的实施方式,上文描述的两个或更多模块或者单元的特征和功能可以在一个模块或者单元中具体化。反之,上文描述的一个模块或者单元的特征和功能可以进一步划分为由多个模块或者单元来具体化。

[0235] 通过以上的实施方式的描述,本领域的技术人员易于理解,这里描述的示例实施方式可以通过软件实现,也可以通过软件结合必要的硬件的方式来实现。因此,根据本申请实施方式的技术方案可以以软件产品的形式体现出来,该软件产品可以存储在一个非易失性存储介质(可以是CD-ROM,U盘,移动硬盘等)中或网络上,包括若干指令以使得一台计算设备(可以是个人计算机、服务器、触控终端、或者网络设备等)执行根据本申请实施方式的方法。

[0236] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本申请的其它实施方案。本申请旨在涵盖本申请的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本申请的一般性原理并包括本申请未公开的本技术领域中的公知常识或惯用技术手段。

[0237] 应当理解的是,本申请并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本申请的范围仅由所附的权利要求来限制。

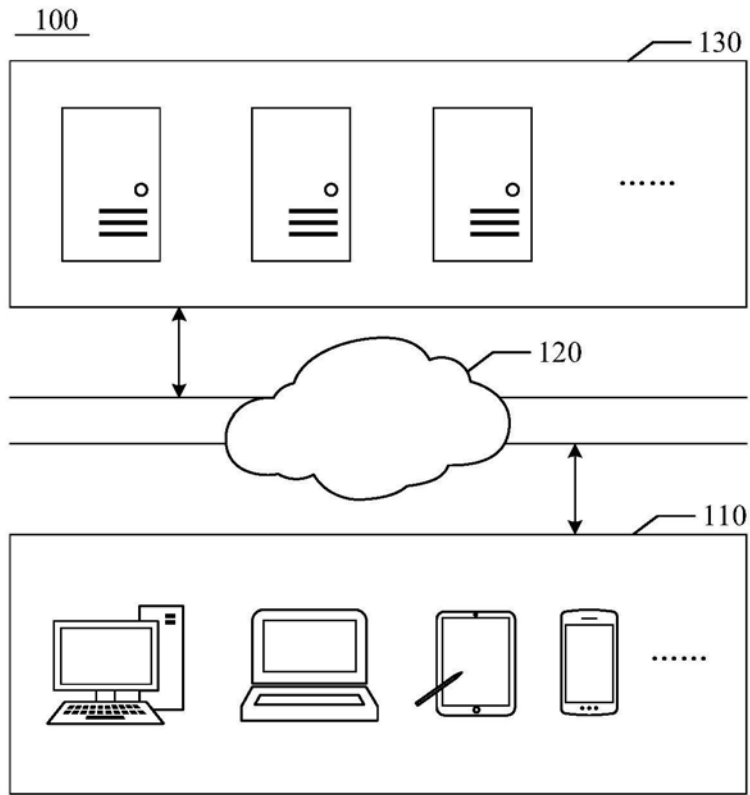


图1

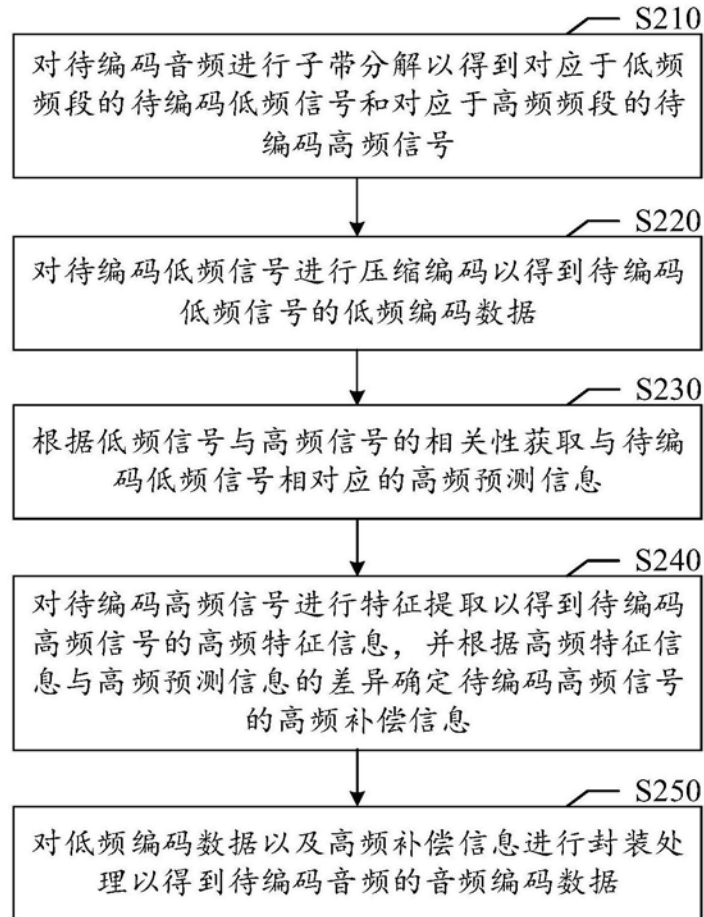


图2

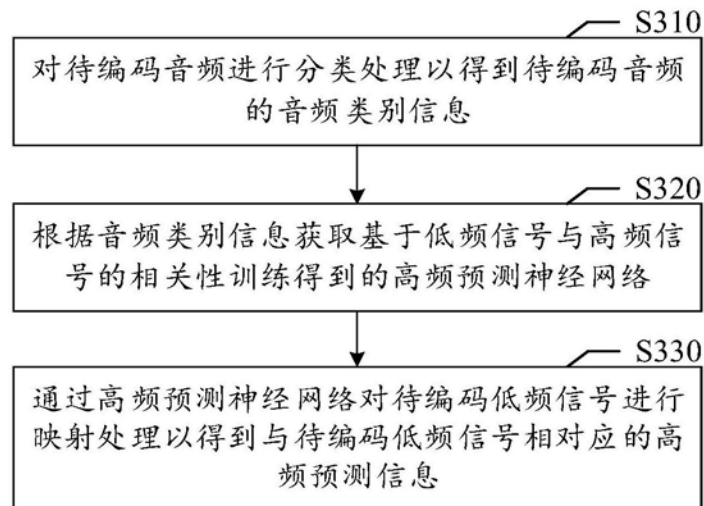


图3

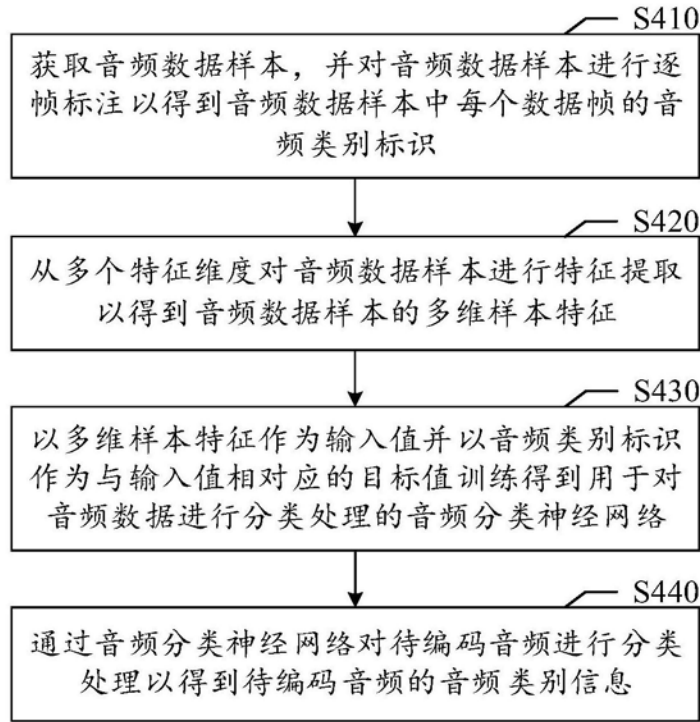


图4

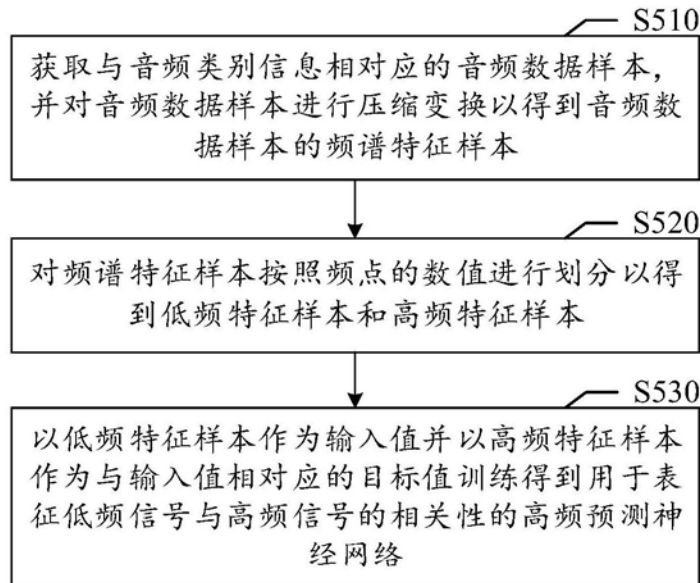


图5

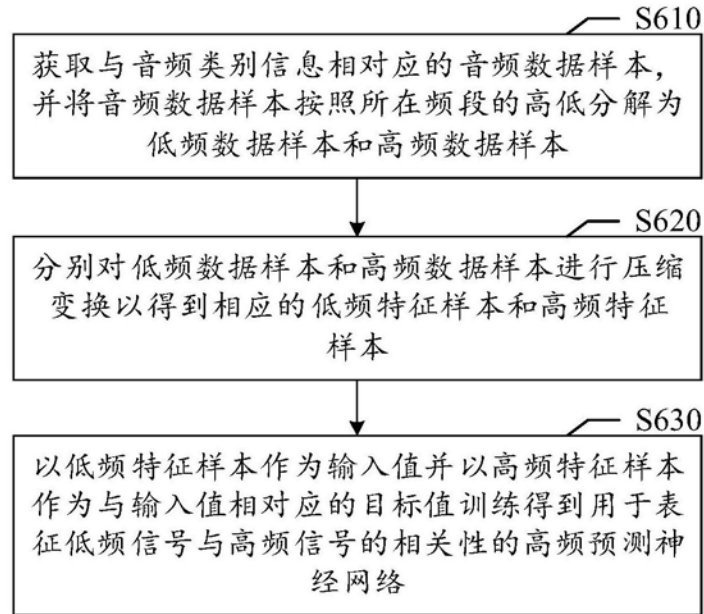


图6

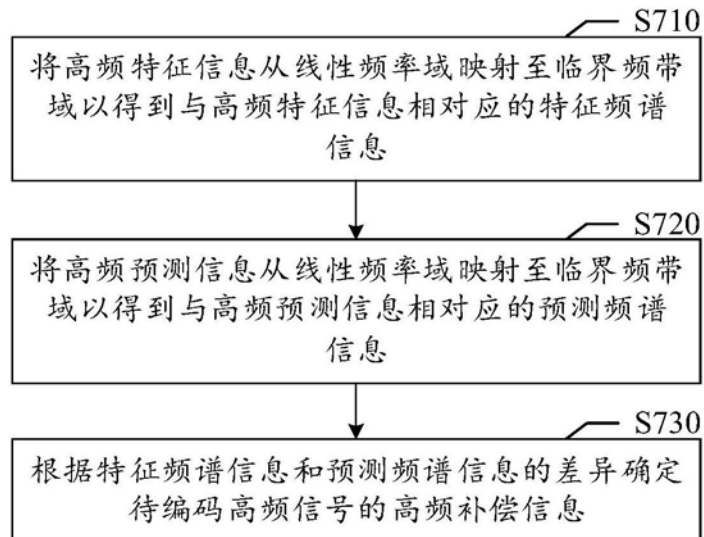


图7

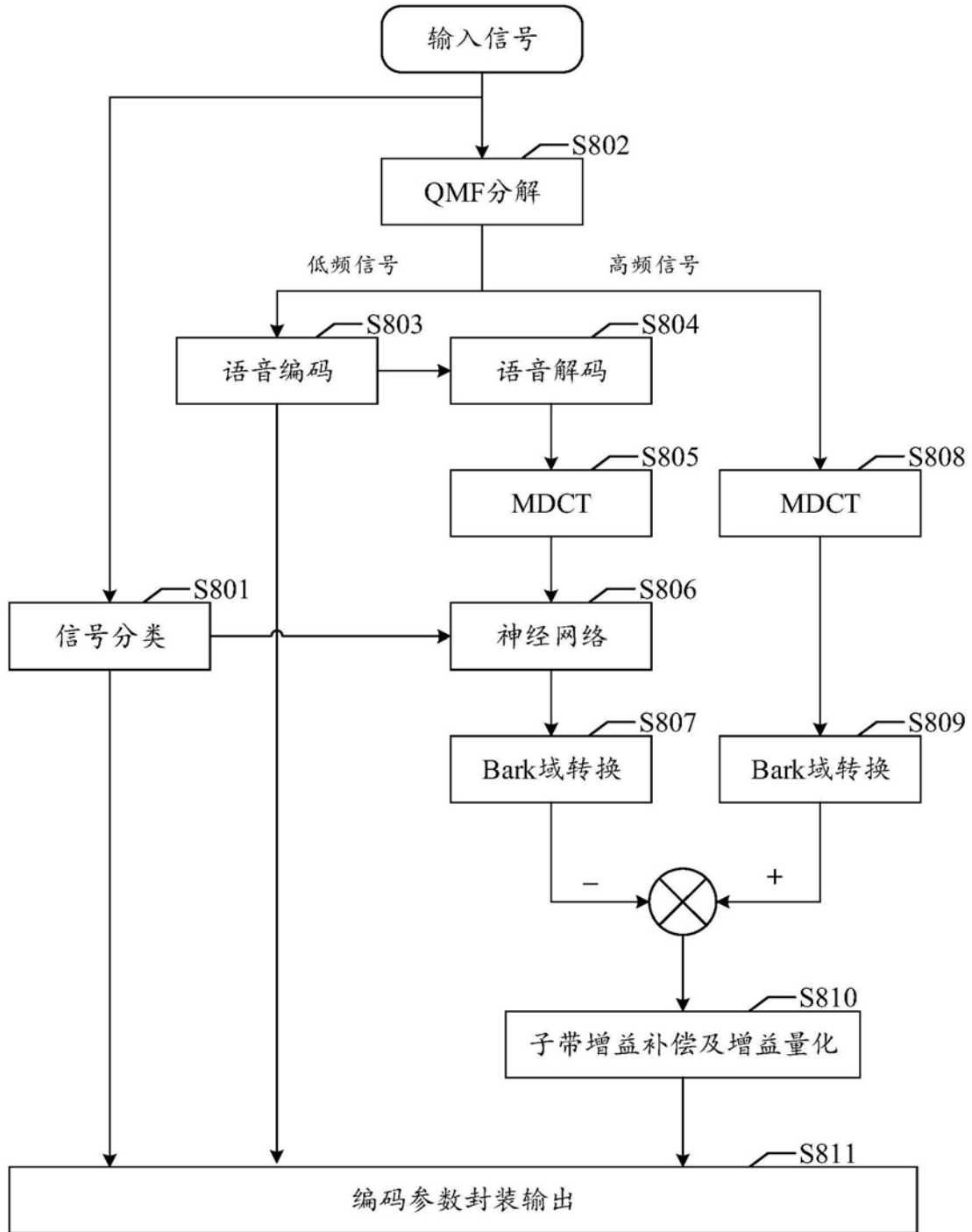


图8

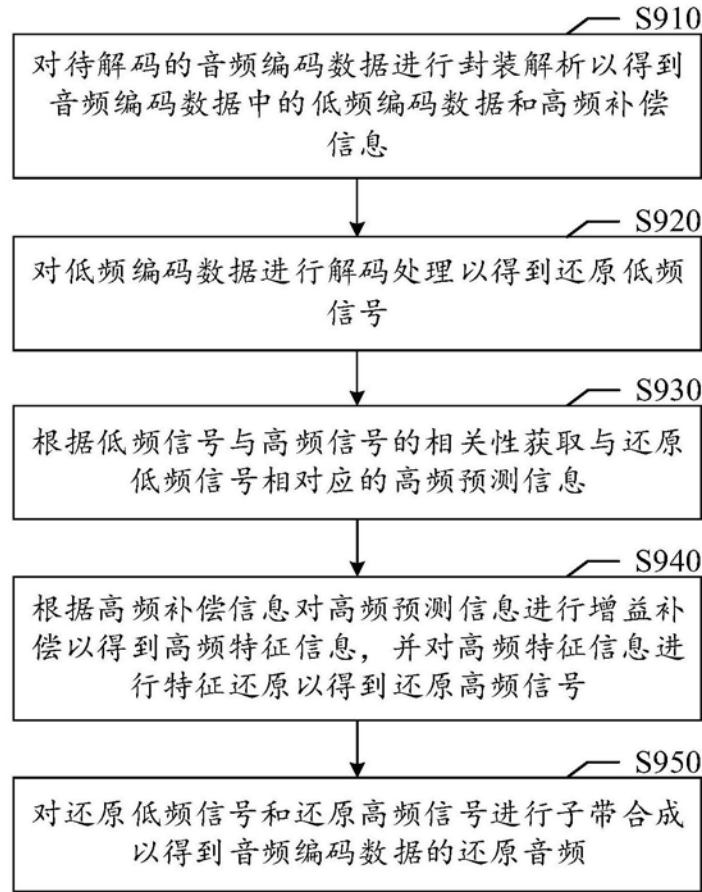


图9

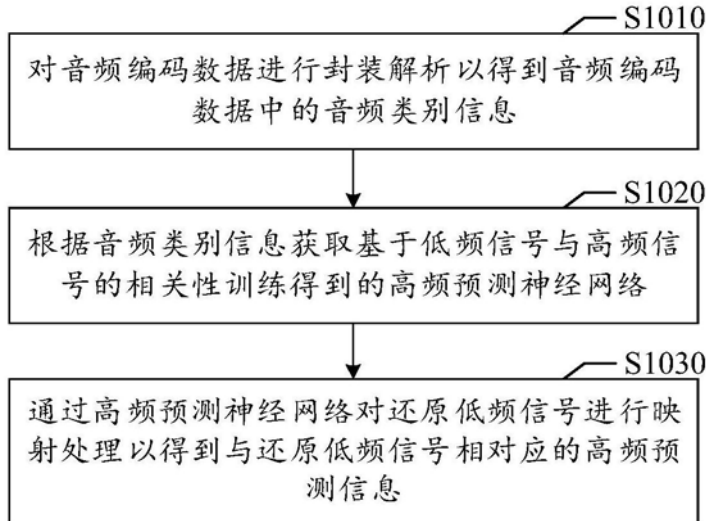


图10

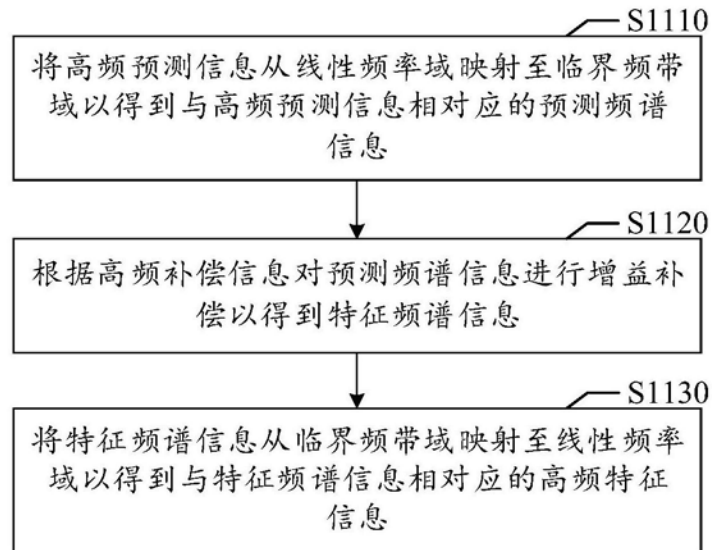


图11

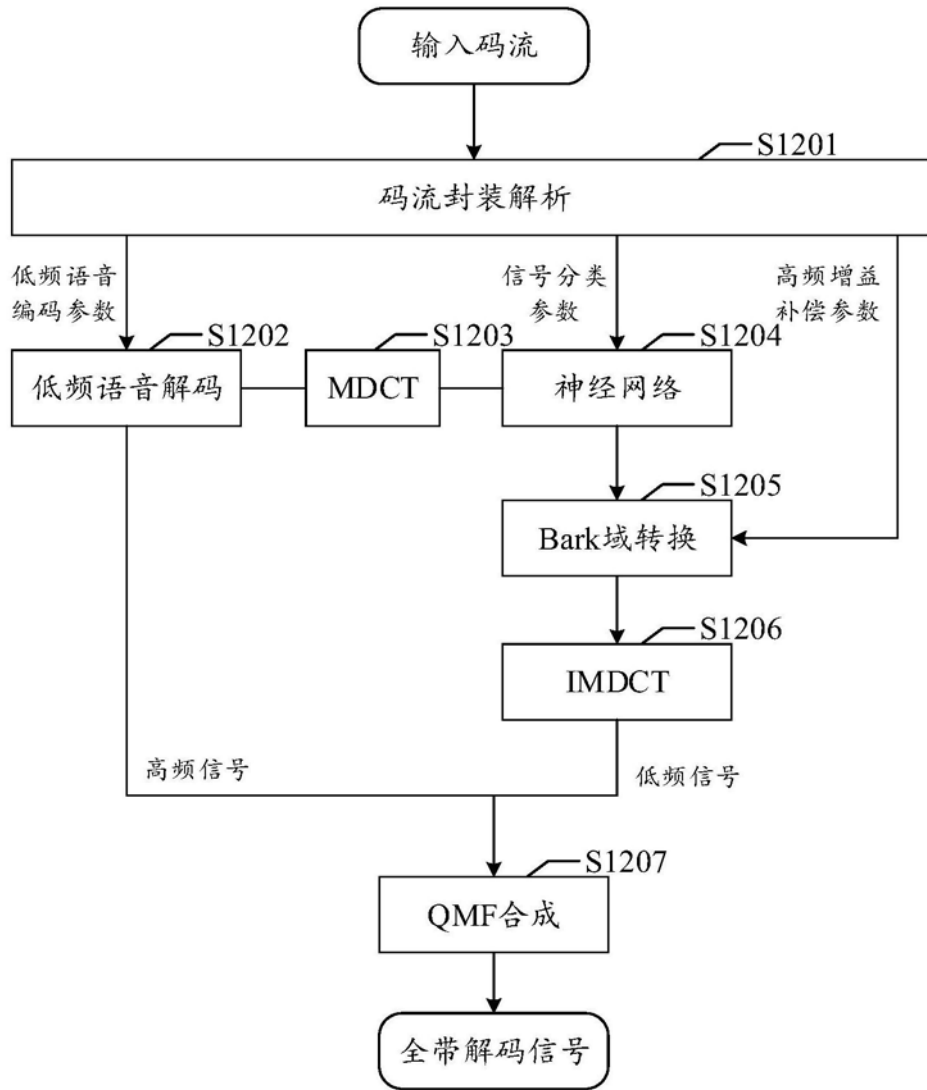


图12

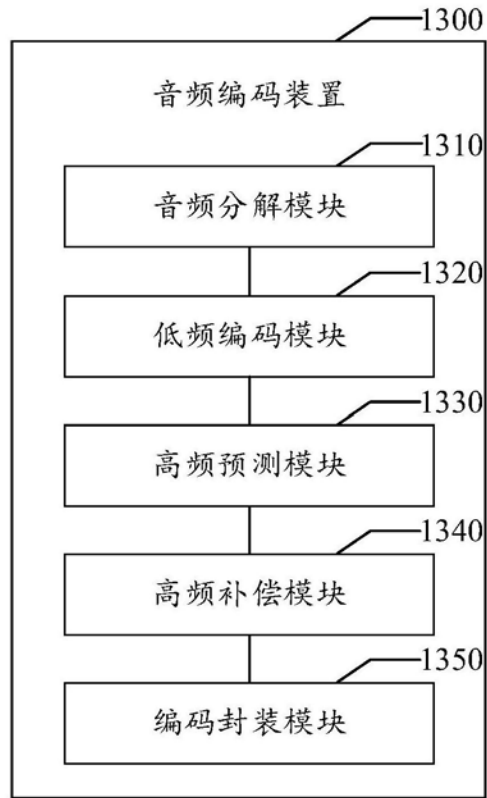


图13

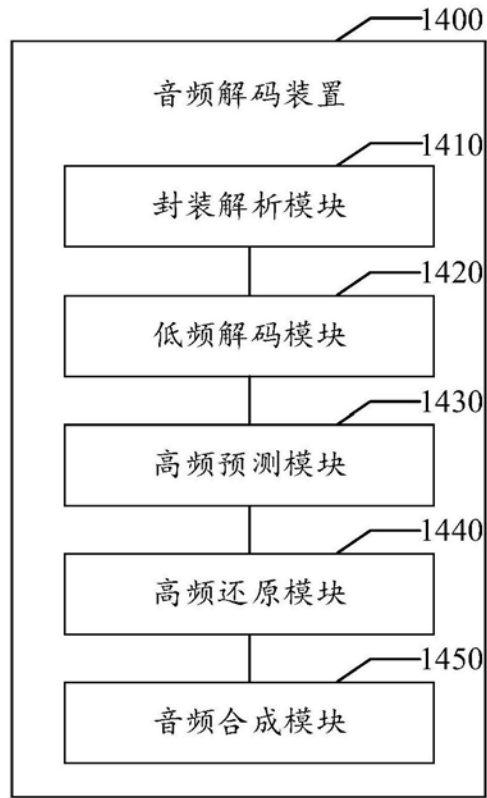


图14

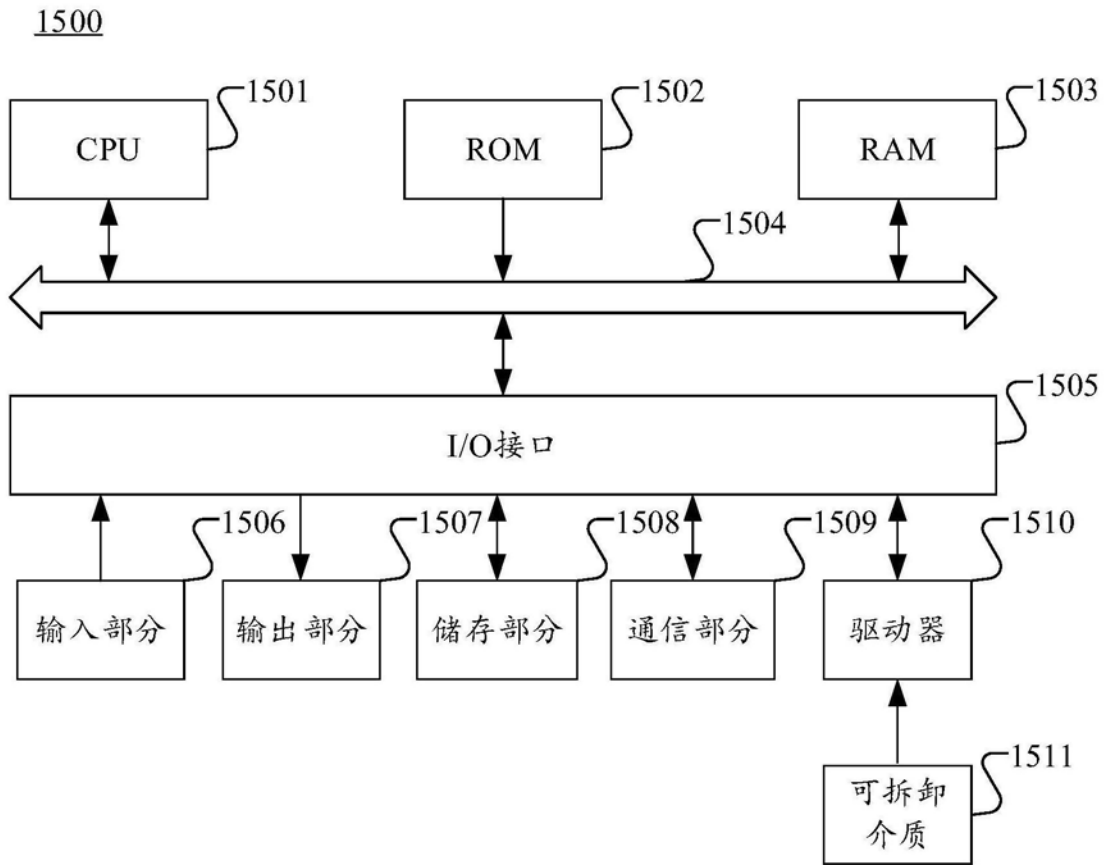


图15