说明书

发明名称: 音频编解码方法和装置

技术领域

[0001] 本申请涉及音频处理技术，尤其涉及一种音频编解码方法和装置。

背景技术

[0002] 随着通讯系统的升级，音频传输质量得到了极大的提高。但仍存在一些场景，传输带宽还是会受到限制，有限的传输比特无法表示全带宽信号。为解决这一问题，相关技术中将音频分为低频成分和高频成分，通过编解码方式来处理低频成分，并将大部分传输比特分配给低频成分的码流，以保证低频成分的编码质量；再利用高低频之间的关联性，在解码侧通过频带扩展恢复高频成分。这种扩展频带的方法被称为频带扩展(Bandwidth Extension，BWE)。相较于窄带音频，采用BWE得到的音频的听感会明显提升。

[0003] 然而，上述BWE方法会因为高频成分的频谱与低频成分的频谱的关联性弱，而出现音频失真的问题。

[0004] 发明内容

[0005] 本申请提供一种音频编解码方法和装置，以重建出更多的高频细节，从而可以避免音频失真的问题。

[0006] 第一方面，本申请提供一种音频编码方法，包括：获取第一码流，所述第一码流是基于第一音频信号得到的，所述第一音频信号是当前帧中频率低于预设阈值的音频信号；获取边信息，所述边信息是基于第二音频信号得到的，所述第二音频信号包括所述当前帧中频率高于或等于所述预设阈值的音频信号；根据所述边信息获取第二码流；传输码流，所述码流包括所述第一码流和所述第二码流。

[0007] 本实施例，通过机器学习模型从音频中提取边信息，以指导解码端实现非盲BWE，重建出高频成分，尤其是与低频成分关联性较弱的高频成分，而这种实现方式可以将当前帧中高频成分的更多细节体现于特征信息中，这些细节例如是

高频成分的特性，高频成分与低频成分之间的关联等，有助于重建出更多的高频细节，从而可以避免音频失真的问题。

[0008] 原始音频经采样得到音频帧序列，其中包括多个音频帧，在编码端可以一帧一帧的进行压缩处理，为便于描述，本申请实施例中将当前正在处理的音频帧称作当前帧。

[0009] 当前帧包含的音频信号的特征可以通过频率、幅度和相位来表征，本申请实施例根据频率从当前帧中提取出第一音频信号和第二音频信号，其中，第一音频信号是当前帧中频率低于预设阈值的音频信号(本申请实施例中，满足该条件的音频信号亦可以称作低频信号或低频成分)，第二音频信号包括当前帧中频率高于或等于预设阈值的音频信号(本申请实施例中，满足该条件的音频信号亦可以称作高频信号或高频成分)。前述预设阈值例如可以是16kHz，本申请实施例对此不做具体限定。需要说明的是，前文将当前帧中频率等于预设阈值的音频信号归入第二音频信号，可选的，本申请实施例也可以将当前帧中频率等于预设阈值的音频信号归入第一音频信号，对此不做具体限定。

[0010] 一些可能的实施方式中，可以采用低通滤波器从当前帧中提取出第一音频信号。

[0011] 一些可能的实施方式中，可以采用核心编码器对第一音频信号进行编码，该核心编码器例如可以采用AVS3核心编码器，本申请实施例对此不做具体限定。

[0012] 本申请实施例中，第二音频信号可能存在以下三种情况：

[0013] 1)第二音频信号只包括高频成分，该情况下第二音频信号等同于当前帧的高频成分。与之对应，边信息可以只包含高频成分的特性信息；

[0014] 2)第二音频信号除了包括高频成分外，还包括低频成分中的部分音频信号，亦即，第二音频信号和第一音频信号有部分相同的音频信号，该情况下第二音频信号等同于当前帧的高频成分加部分低频成分。与之对应，边信息中除了包含高频成分的特性信息外，还可以包含高频成分与前述部分低频成分之间的关联性信息；

[0015] 3)第二音频信号除了包括高频成分外，还包括低频成分中的全部音频信号，亦即，第二音频信号包括第一音频信号，该情况下第二音频信号等同于当前帧。

与之对应，边信息中除了包含高频成分的特征信息外，还可以包含高频成分与低频成分之间的关联性信息。

[0016] 由此可见，边信息可以仅占用少量的传输比特，就可以指导解码端实现非盲BWE，以重建出高频成分，尤其是高低频之间映射关系建复杂的信号。

[0017] 一些可能的实施方式中，获取边信息可以包括将第二音频信号输入预先训练得到的第二机器学习模型，以获取边信息，该第二机器学习模型包括至少一个子模型。至少一个子模型可以分别专注于不同的特性，例如，为各个子模型分别设置限制条件，包括输入输出、模型结构、处理规则等，这样每个子模型可以专注于提取对应特性的边信息。

[0018] 本申请实施例中，第二机器学习模型可以参照上文的神经网络、深度神经网络、卷积神经网络、循环神经网络等。机器学习模型可以将当前帧中高频成分的更多细节体现于特征信息中，这些细节例如是高频成分的特性，高频成分与低频成分之间的关联等，有助于重建出更多的高频细节。

[0019] 一些可能的实施方式中，第二机器学习模型可以包括多个子模型，该多个子模型各自的输入包括当前帧中频率高于或等于预设阈值的音频信号，即多个子模型各自的输入是第二音频信号中的部分或全部音频信号。

[0020] 本申请实施例中，若子模型的输入只包含高频成分，则该子模型主要用于将高频成分压缩为特征，或者通过特征来表示高频成分；若子模型的输入除了高频成分外还包含低频成分，则该子模型还用于提取高低频之间的映射关系。采用多个子模型可以在边信息的提取中得到更多样的特征，因此本申请对子模型的数量、类型、功能等均不做具体限定。

[0021] n个子模型各自的输出，包括特征信息1、特征信息2、…、特征信息n，经合并处理后得到边信息，亦即，边信息包括n个子模型分别输出的特征信息。

[0022] 一些可能的实施方式中，第一子模型与多个子模型中的其他至少一个子模型包括相同的神经元，该相同的神经元位于第一子模型的输入阶段、输出阶段或者中间阶段，该第一子模型是多个子模型中的任意一个。

[0023] 第一子模型可以和第二机器学习模型包括的多个子模型中的一个或多个有相同的神经元(亦称作共用神经元)，该相同的神经元可以有一层或多层，而该一层或

多层神经元可以位于子模型的输入阶段、输出阶段或中间阶段，其中，输入阶段是指该子模型的输入端的第一层及其之后的一层或多层，输出阶段是指该子模型的输出端的最后一层及其之前的一层或多层，中间阶段是指不包括子模型的第一层和最后一层的一层或多层。需要说明的是，本申请实施例对前述相同的神经元的数量、类型等不做具体限定，对输入阶段、输出阶段以及中间阶段的位置也不做具体限定。

[0024] 一些可能的实施方式中，第一子模型的输入还包括第一输入；当第二机器学习模型包括多个子模型时，第一输入包括多个子模型中的其他至少一个子模型的输出，或者，第一输入包括至少一个子规则方式的输出；第一子模型是多个子模型中的任意一个。

[0025] 一些可能的实施方式中，第一子模型的输入还包括第二输入，第二输入是第一输入经冗余信息处理后得到的音频信号；当第二机器学习模型包括多个子模型时，第一输入包括多个子模型中的其他至少一个子模型的输出，或者，第一输入包括至少一个子规则方式的输出；第一子模型是多个子模型中的任意一个。

[0026] 需要说明的是，上述两个示例以两个子模型为例进行说明，但这并非是对子模型之间、子模型与子规则方式之间的串联关系构成限定，任意一个子模型的输入除了自己的输入外，还可以包括其他一个或多个子模型的输出或经冗余处理后的输出，和/或，一个或多个子规则方式的输出或经冗余处理后的输出，本申请实施例不限定串联的个数，串联方式。

[0027] 一些可能的实施方式中，获取边信息还可以包括根据第二音频信号按照预先设置的规则方式进行处理以获取边信息，该预先设置的规则方式包括至少一个子规则方式。

[0028] 本申请实施例中，按照预先设置的规则方式提取特征的方法，例如可以包括提取频谱包络和参照AVS3提取特征，其中，频谱包络是当前帧的频谱中各频带的能量，用于给解码侧进行对应频带能量的恢复；AVS3提取的特征包括频带能量，以及对各频带谐波特性和噪声特性的标识，用于给解码侧恢复出BWE基底。相比于机器学习方式，规则方式可以保留频谱能量等信息，从而避免对听感的影响。

[0029] 本申请实施例中，在第二机器学习模型的基础上，还可以设置规则方式，对第二音频信号按照前述规则方式进行处理，得到与其对应的输出(特征信息)。

[0030] 第二机器学习模型输出的特征信息1和规则方式输出的特征信息2经过合并处理后得到边信息，亦即，边信息包括第二机器学习模型和按照规则方式分别输出的特征信息。

[0031] 一些可能的实施方式中，预先设置的规则方式可以包括多个子规则方式，该多个子规则方式各自的输入包括当前帧中频率高于或等于预设阈值的音频信号，即多个子规则方式各自的输入是第二音频信号中的部分或全部音频信号。

[0032] n个子模型、m个规则方式各自的输出，包括特征信息1、特征信息2、…、特征信息n、特征信息n+1、特征信息n+2、…、特征信息n+m，经合并处理后得到边信息，亦即，边信息包括n个子模型分别输出的特征信息和按照m个规则方式输出的特征信息。

[0033] 一些可能的实施方式中，第一子规则方式的输入还包括第三输入；当预先设置的规则方式包括多个子规则方式时，第三输入包括多个子规则方式中的其他至少一个子规则方式的输出，或者，第三输入包括至少一个子模型的输出；第一子规则方式是多个子规则方式中的任意一个。

[0034] 一些可能的实施方式中，第一子规则方式的输入还包括第四输入，第四输入是第三输入经冗余信息处理后得到的音频信号；当预先设置的规则方式包括多个子规则方式时，第三输入包括多个子规则方式中的其他至少一个子规则方式的输出，或者，第三输入包括至少一个子模型的输出；第一子规则方式是多个子规则方式中的任意一个。

[0035] 需要说明的是，上述示例以两个子规则方式为例进行说明，但这并非是对子规则方式之间、子规则方式与子模型之间的串联关系构成限定，任意一个子规则方式的输入除了自己的输入外，还可以包括其他一个或多个子规则方式的输出或经冗余处理后的输出，和/或，一个或多个子模型的输出或经冗余处理后的输出，本申请实施例不限定串联的个数，串联方式。

[0036] 一些可能的实施方式中，可以对获取边信息的第二机器学习模型和/或规则方式进行联合训练。即，通过联合训练多个获取边信息的机制(该机制可以是上述

子模型或者子规则方式)来降低或消除特征组之间的信息冗余；为多个获取边信息的机制设置联合损失函数，通过误差反向传播来修正多个获取边信息的机制的权重，来达到多个特征信息的整体最优的效果，进而降低或消除特征组之间的信息冗余；每个获取边信息的机制可以具有各自的损失函数，用于保留其提取到的特征信息的独特性。

[0037] 本申请实施例中，可以通过边信息编码器对边信息进行编码，该边信息编码器可以与上文核心编码器采用相同的编码算法，也可以采用不同的编码算法，对此不做具体限定。

[0038] 编码端得到的码流包括两部分，一部分码流(第一码流)对应当前帧的低频成分，另一部分码流(第二码流)对应当前帧的边信息。编码端将码流传输至解码端，以便于解码端基于码流得到重建的低频成分和边信息，从而得到与当前帧对应的重建帧。

[0039] 第二方面，本申请提供一种音频解码方法，包括：接收码流，所述码流包括第一码流和第二码流；根据所述第一码流获取第一音频信号，所述第一音频信号的频率低于预设阈值；根据所述第二码流获取边信息，所述边信息用于表征第二音频信号的特性，所述第二音频信号的频率高于或等于所述预设阈值；根据所述第一音频信号和所述边信息获取重建帧。

[0040] 本实施例，通过机器学习模型在边信息的指导下，基于低频成分得到重建帧，尤其是高频成分与低频成分关联性较弱或者高低频成分关系紧密但映射关系复杂的场景下的重建帧获取，而这种实现方式可以将高频成分的更多细节体现于特征信息中，这些细节例如是高频成分的特性，高频成分与低频成分之间的关联等，有助于重建出更多的高频细节，从而可以避免音频失真的问题。

[0041] 本申请实施例中，接收来自编码端的码流。与编码端生成的码流相对应，接收到的码流也包括第一码流和第二码流两部分。

[0042] 本申请实施例中，可以通过核心解码器对第一码流进行解析以得到第一音频信号，

[0043] 解码端要得到的重建帧与编码端的当前帧相对应，该重建帧包含的音频信号的特征也可以通过频率、幅度和相位来表征，根据频率划分，重建帧也包括第一

音频信号和第二音频信号，其中，第一音频信号的频率低于预设阈值的音频信号(本申请实施例中，满足该条件的音频信号亦可以称作低频信号或低频成分)，第二音频信号的频率高于或等于预设阈值的音频信号(本申请实施例中，满足该条件的音频信号亦可以称作高频信号或高频成分)，即重建帧包括低频成分和高频成分。前述预设阈值例如可以是16kHz，本申请实施例对此不做具体限定。需要说明的是，前文将频率等于预设阈值的音频信号归入第二音频信号，可选的，本申请实施例也可以将频率等于预设阈值的音频信号归入第一音频信号，对此不做具体限定。

[0044] 基于此，根据第一码流得到的第一音频信号可以是重建帧的低频成分，亦即，解码端可以通过解析码流直接得到重建帧的低频成分。

[0045] 边信息用于表征第二音频信号的特性，此外，边信息还可以用于表征高低频之间的映射关系。本申请实施例中，边信息仅占用少量的传输比特，就可以指导解码端实现非盲BWE，以重建出高频成分，尤其是与低频成分关系紧密但映射关系复杂的高频成分。

[0046] 一些可能的实施方式中，根据第二码流获取边信息可以直接解析第二码流以得到边信息。即，第二码流是由编码端处理边信息得到的，因此解码端可以直接解析第二码流得到边信息。

[0047] 可选的，解析第二码流先得到处理前边信息；然后对处理前边信息进行边信息处理以得到边信息，该边信息处理包括输入预先训练得到的第三机器学习模型或者按照预先设置的第三规则方式进行处理。

[0048] 本申请实施例中，第二码流携带的信息可以是与边信息关联的处理前边信息，该信息可以是编码端考虑到进一步节省码流，而基于边信息得到的，进而将其编码入码流。相应的，解码端为了得到边信息可以先解析第二码流得到的处理前边信息，然后基于处理前边信息获取边信息。前述获取边信息的方法可以包括通过第三机器学习模型或者按照预先设置的第三规则方式获取边信息。

[0049] 一些可能的实施方式中，获取重建帧可以根据第一音频信号和边信息通过预先训练得到的第一机器学习模型，获取重建帧。由于获取重建帧参考了边信息，以实现从低频成分向高频成分的扩展，因此获取重建帧也可以称作非盲BWE。

[0050] 本申请实施例中，第一机器学习模型可以参照上文的神经网络、深度神经网络、卷积神经网络、循环神经网络等。

[0051] 可选的，重建帧是第一机器学习模型直接输出的。

[0052] 本申请实施例中，非盲BWE可以采用一个机器学习模型(第一机器学习模型)，在边信息的指导下，根据解码出的低频成分(第一音频信号)，直接输出重建帧。此外，第一机器学习模型可以完成低频成分的编解码失真修正，以及可能的音频渲染，本申请实施例对此不做具体限定。

[0053] 可选的，重建帧是基于第一机器学习模型输出的第二音频信号与第一音频信号得到的。

[0054] 本申请实施例中，非盲BWE可以采用一个机器学习模型(第一机器学习模型)，在边信息的指导下，根据解码出的低频成分(第一音频信号)，输出重建的高频成分，将重建的高频成分与解码出的低频成分合并或相加，得到重建音频。此外，第一机器学习模型可以完成低频成分的编解码失真修正，以及可能的音频渲染，本申请实施例对此不做具体限定。

[0055] 一些可能的实施方式中，第一机器学习模型可以包括第一子模型和第二子模型，非盲BWE可以将第一音频信号和第一边信息输入第一子模型，以获取第三音频信号；再将第三音频信号和第二边信息输入第二子模型，以获取重建帧；第一边信息和第二边信息分别是边信息中的部分或全部信息。

[0056] 一些可能的实施方式中，非盲BWE可以根据第一音频信号和第三边信息按照预先设置的第一规则方式进行处理以获取第四音频信号；再将第四音频信号和第四边信息输入第一机器学习模型，以获取重建帧；第三边信息和第四边信息分别是边信息中的部分或全部信息。

[0057] 一些可能的实施方式中，非盲BWE可以根据第一音频信号和第五边信息通过预先训练得到的第一机器学习模型，获取第五音频信号；再根据第五音频信号和第六边信息按照预先设置的第二规则方式进行处理以获取重建帧；第五边信息和第六边信息分别是边信息中的部分或全部信息。

[0058] 需要说明的是，多个边信息可以分别是边信息中的部分或全部信息，多个边信息可以完全相同或完全不同或部分相同、部分不同，对此不做具体限定。

[0059] 第三方面，本申请提供一种编码装置，包括：获取模块，用于获取第一码流，所述第一码流是基于第一音频信号得到的，所述第一音频信号是当前帧中频率低于预设阈值的低频信号；提取模块，用于获取边信息，所述边信息是基于第二音频信号得到的，所述第二音频信号包括所述当前帧中频率高于或等于所述预设阈值的高频信号；编码模块，用于根据所述边信息获取第二码流；传输模块，用于传输码流，所述码流包括所述第一码流和所述第二码流。

[0060] 在一种可能的实现方式中，所述提取模块，具体用于将所述第二音频信号输入预先训练得到的第二机器学习模型，以获取所述边信息，所述第二机器学习模型包括至少一个子模型。

[0061] 在一种可能的实现方式中，当所述第二机器学习模型包括多个子模型时，所述多个子模型各自的输入包括所述第二音频信号中的部分音频信号或全部音频信号，其中，所述部分音频信号包括所述高频信号。

[0062] 在一种可能的实现方式中，第一子模型与所述多个子模型中的其他至少一个子模型包括相同的神经元，所述相同的神经元位于所述第一子模型的输入阶段、输出阶段或者中间阶段，所述第一子模型是所述多个子模型中的任意一个。

[0063] 在一种可能的实现方式中，所述边信息包括所述多个子模型分别输出的特征信息。

[0064] 在一种可能的实现方式中，当所述第二机器学习模型包括一个子模型时，所述边信息包括所述一个子模型输出的特征信息。

[0065] 在一种可能的实现方式中，所述提取模块，具体用于根据所述第二音频信号按照预先设置的规则方式进行处理以获取所述边信息，所述预先设置的规则方式包括至少一个子规则方式。

[0066] 在一种可能的实现方式中，当所述预先设置的规则方式包括多个子规则方式时，所述多个子规则方式各自的输入包括所述第二音频信号中的部分音频信号或全部音频信号，其中，所述部分音频信号包括所述高频信号。

[0067] 在一种可能的实现方式中，所述边信息包括按照所述多个子规则方式分别输出的特征信息。

[0068] 在一种可能的实现方式中，当所述预先设置的规则方式包括一个子规则方式时，所述边信息包括按照所述一个子规则方式输出的特征信息。

[0069] 在一种可能的实现方式中，第一子模型的输入还包括第一输入；当所述第二机器学习模型包括多个子模型时，所述第一输入包括所述多个子模型中的其他至少一个子模型的输出，或者，所述第一输入包括所述至少一个子规则方式的输出；所述第一子模型是所述多个子模型中的任意一个。

[0070] 在一种可能的实现方式中，第一子模型的输入还包括第二输入，所述第二输入是第一输入经冗余信息处理后得到的音频信号；当所述第二机器学习模型包括多个子模型时，所述第一输入包括所述多个子模型中的其他至少一个子模型的输出，或者，所述第一输入包括所述至少一个子规则方式的输出；所述第一子模型是所述多个子模型中的任意一个。

[0071] 在一种可能的实现方式中，第一子规则方式的输入还包括第三输入；当所述预先设置的规则方式包括多个子规则方式时，所述第三输入包括所述多个子规则方式中的其他至少一个子规则方式的输出，或者，所述第三输入包括所述至少一个子模型的输出；所述第一子规则方式是所述多个子规则方式中的任意一个。

[0072] 在一种可能的实现方式中，第一子规则方式的输入还包括第四输入，所述第四输入是第三输入经冗余信息处理后得到的音频信号；当所述预先设置的规则方式包括多个子规则方式时，所述第三输入包括所述多个子规则方式中的其他至少一个子规则方式的输出，或者，所述第三输入包括所述至少一个子模型的输出；所述第一子规则方式是所述多个子规则方式中的任意一个。

[0073] 在一种可能的实现方式中，还包括：滤波模块，用于通过低通滤波的方式从所述当前帧中获取所述第一音频信号。

[0074] 在一种可能的实现方式中，所述第二音频信号还包括所述第一音频信号中的部分或全部音频信号。

[0075] 第四方面，本申请提供一种解码装置，包括：接收模块，用于接收码流，所述码流包括第一码流和第二码流；获取模块，用于根据所述第一码流获取第一音频信号，所述第一音频信号的频率低于预设阈值；提取模块，用于根据所述第

二码流获取边信息，所述边信息用于表征第二音频信号的特性，所述第二音频信号的频率高于或等于所述预设阈值；重建模块，用于根据所述第一音频信号和所述边信息获取重建帧。

[0076] 在一种可能的实现方式中，所述重建模块，具体用于根据所述第一音频信号和所述边信息通过预先训练得到的第一机器学习模型，获取所述重建帧。

[0077] 在一种可能的实现方式中，所述重建帧是所述第一机器学习模型直接输出的；或者，所述重建帧是基于所述第一机器学习模型输出的所述第二音频信号与所述第一音频信号得到的。

[0078] 在一种可能的实现方式中，当所述第一机器学习模型包括第一子模型和第二子模型时，所述重建模块，具体用于将所述第一音频信号和第一边信息输入所述第一子模型，以获取第三音频信号；将所述第三音频信号和第二边信息输入所述第二子模型，以获取所述重建帧；所述第一边信息和所述第二边信息分别是所述边信息中的部分或全部信息。

[0079] 在一种可能的实现方式中，所述重建模块，具体用于根据所述第一音频信号和第三边信息按照预先设置的第一规则方式进行处理以获取第四音频信号；
将所述第四音频信号和第四边信息输入所述第一机器学习模型，以获取所述重建帧；所述第三边信息和所述第四边信息分别是所述边信息中的部分或全部信息。

[0080] 在一种可能的实现方式中，所述重建模块，具体用于根据所述第一音频信号和所述五边信息通过预先训练得到的第一机器学习模型，获取第五音频信号；根据所述第五音频信号和第六边信息按照预先设置的第二规则方式进行处理以获取所述重建帧；所述第五边信息和所述第六边信息分别是所述边信息中的部分或全部信息。

[0081] 在一种可能的实现方式中，所述提取模块，具体用于解析所述第二码流以得到所述边信息。

[0082] 在一种可能的实现方式中，所述提取模块，具体用于解析所述第二码流以得到处理前边信息；对所述处理前边信息进行边信息处理以得到所述边信息，所述

边信息处理包括输入预先训练得到的第三机器学习模型或者按照预先设置的第三规则方式进行处理。

[0083] 第五方面，本申请提供一种设备，包括：一个或多个处理器；存储器，用于存储一个或多个程序；当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如上述第一至二方面中任一项所述的方法。

[0084] 第六方面，本申请提供一种计算机可读存储介质，包括计算机程序，所述计算机程序在计算机上被执行时，使得所述计算机执行上述第一至二方面中任一项所述的方法。

[0085] 第七方面，本申请提供一种计算机程序产品，所述计算机程序产品包括计算机程序代码，当所述计算机程序代码在计算机上运行时，使得计算机执行上述第一至二方面中任一项所述的方法。

附图说明

[0086] 图1示出了本申请实施例的一种应用场景100的示意图；

[0087] 图2示出了本申请实施例的另一种应用场景200的示意图；

[0088] 图3示出了本申请实施例的另一种应用场景300的示意图；

[0089] 图4示出了本申请实施例的另一种应用场景400的示意图；

[0090] 图5为本申请提供的一种终端设备500的结构示意图；

[0091] 图6为本申请实施例提供的一种接入网设备600的结构示意图；

[0092] 图7为本申请实施例提供的音频编码方法的过程700的流程图；

[0093] 图8a为本申请实施例的获取边信息的示意图；

[0094] 图8b为本申请实施例的获取边信息的示意图；

[0095] 图8c为本申请实施例的获取边信息的示意图；

[0096] 图8d为本申请实施例的获取边信息的示意图；

[0097] 图9a为本申请实施例的两个子模型的相同部分的示意图；

[0098] 图9b为本申请实施例的两个子模型的相同部分的示意图；

[0099] 图9c为本申请实施例的两个子模型的相同部分的示意图；

[0100] 图10a为本申请实施例的子模型串联的示意图；

[0101] 图10b为本申请实施例的子模型串联的示意图；

[0102] 图11a为本申请实施例的子规则方式串联的示意图；

[0103] 图11b为本申请实施例的子规则方式串联的示意图；

[0104] 图12为本申请实施例的联合训练的示意图；

[0105] 图13为本申请实施例提供的音频解码方法的过程1300的流程图；

[0106] 图14a为本申请实施例的非盲BWE的示意图；

[0107] 图14b为本申请实施例的非盲BWE的示意图；

[0108] 图14c为本申请实施例的非盲BWE的示意图；

[0109] 图14d为本申请实施例的非盲BWE的示意图；

[0110] 图14e为本申请实施例的非盲BWE的示意图；

[0111] 图14f为本申请实施例的非盲BWE的示意图；

[0112] 图15a为本申请实施例的获取重建帧的示意图；

[0113] 图15b为本申请实施例的获取重建帧的示意图；

[0114] 图16为本申请实施例的包含AI的非盲BWE的编解码框架的示意图；

[0115] 图17为本申请实施例的包含AI的非盲BWE的编解码框架的示意图；

[0116] 图18为本申请实施例的包含AI的非盲BWE的编解码框架的示意图；

[0117] 图19为本申请实施例的边信息提取环节的流程图；

[0118] 图20为本申请实施例的非盲BWE环节的流程图；

[0119] 图21和图22分别是对三种BWE方式得到的音频频谱的比较示意图；

[0120] 图23为本申请实施例的包含AI的非盲BWE的编解码框架的示意图；

[0121] 图24为本申请实施例的边信息提取环节的流程图；

[0122] 图25为本申请实施例的非盲BWE环节的流程图；

[0123] 图26为本申请实施例的编码装置2600的结构示意图；

[0124] 图27为本申请实施例的解码装置2700的结构示意图。

具体实施方式

[0125] 为使本申请的目的、技术方案和优点更加清楚，下面将结合本申请中的附图，对本申请中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通

技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

[0126] 本申请的说明书实施例和权利要求书及附图中的术语“第一”、“第二”等仅用于区分描述的目的，而不能理解为指示或暗示相对重要性，也不能理解为指示或暗示顺序。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元。方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0127] 应当理解，在本申请中，“至少一个(项)”是指一个或者多个，“多个”是指两个或两个以上。“和/或”，用于描述关联对象的关联关系，表示可以存在三种关系，例如，“A和/或B”可以表示：只存在A，只存在B以及同时存在A和B三种情况，其中A，B可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系。“以下至少一项(个)”或其类似表达，是指这些项中的任意组合，包括单项(个)或复数项(个)的任意组合。例如，a，b或c中的至少一项(个)，可以表示：a，b，c，“a和b”，“a和c”，“b和c”，或“a和b和c”，其中a，b，c可以是单个，也可以是多个。

[0128] 本申请涉及到的相关名词解释：

[0129] 音频帧：音频数据是流式的，在实际应用中，为了便于音频处理和传输，通常取一时长内的音频数据量作为一帧音频，该时长被称为“采样时间”，可以根据编解码器和具体应用的需求确定其值，例如该时长为2.5ms～60ms，ms为毫秒。

[0130] 音频信号：音频信号是带有语音、音乐和音效的有规律的声波的频率、幅度变化信息载体。音频是一种连续变化的模拟信号，可用一条连续的曲线来表示，称为声波。音频通过模数转换或计算机生成的数字信号即为音频信号。声波有三个重要参数：频率、幅度和相位，这也就决定了音频信号的特征。

[0131] 此外，由于本申请实施例涉及人工智能(artificial intelligence，AI)的应用，为了便于理解，下面先对本申请实施例所使用到的一些相关名词或术语进行解释说明。

[0132] (1)神经网络

[0133] 神经网络(Neural Network，NN)是机器学习模型，神经网络可以是由神经单元组成的，神经单元可以是指以xs和截距1为输入的运算单元，该运算单元的输出可以为：

$$h_{W,b}(x) = f(W^T x) = f(\sum_{s=1}^{n} W_s x_s + b)$$

[0134] 其中，s=1、2、……n，n为大于1的自然数，W_s为x_s的权重，b为神经单元的偏置。f为神经单元的激活函数(activation functions)，用于将非线性特性引入神经网络中，来将神经单元中的输入信号转换为输出信号。该激活函数的输出信号可以作为下一层卷积层的输入。激活函数可以是sigmoid函数。神经网络是将许多个上述单一的神经单元联结在一起形成的网络，即一个神经单元的输出可以是另一个神经单元的输入。每个神经单元的输入可以与前一层的局部接受域相连，来提取局部接受域的特征，局部接受域可以是由若干个神经单元组成的区域。

[0135] (2)深度神经网络

[0136] 深度神经网络(deep neural network，DNN)，也称多层神经网络，可以理解为具有很多层隐含层的神经网络，这里的“很多”并没有特别的度量标准。从DNN按不同层的位置划分，DNN内部的神经网络可以分为三类：输入层，隐含层，输出层。一般来说第一层是输入层，最后一层是输出层，中间的层数都是隐含层。层与层之间是全连接的，也就是说，第i层的任意一个神经元一定与第i+1层的任意一个神经元相连。虽然DNN看起来很复杂，但是就每一层的工作来说，其实并不复杂，简单来说就是如下线性关系表达式：$\vec{y} = \alpha(W\vec{x} + \vec{b})$,其中，$\vec{x}$是输入向量，$\vec{y}$是输出向量，$\vec{b}$是偏移向量，W是权重矩阵(也称系数)，α()是激活函数。每一层仅仅是对输入向量$\vec{x}$经过如此简单的操作得到输出向量$\vec{y}$。由于DNN层数多，则系数W和偏移向量$\vec{b}$的数量也就很多了。这些参数在DNN中的定义如下所述：以系数W为例：假设在一个三层的DNN中，第二层的第4个神经元到第三层的第2个神经元的线性系数定义为w_{24}^{3}。上标3代表系数W所在的层数，而下标对应的是输出的第三层索引2和输入的第二层索引4。总结就是：第L−1层的第k个

神经元到第L层的第j个神经元的系数定义为w_{jk}^{L}。需要注意的是，输入层是没有W参数的。在深度神经网络中，更多的隐含层让网络更能够刻画现实世界中的复杂情形。理论上而言，参数越多的模型复杂度越高，“容量”也就越大，也就意味着它能完成更复杂的学习任务。训练深度神经网络的也就是学习权重矩阵的过程，其最终目的是得到训练好的深度神经网络的所有层的权重矩阵(由很多层的向量W形成的权重矩阵)。

[0137] (3)卷积神经网络

[0138] 卷积神经网络(convolutional neuron network，CNN)是一种带有卷积结构的深度神经网络，是一种深度学习(deep learning)架构，深度学习架构是指通过机器学习的算法，在不同的抽象层级上进行多个层次的学习。作为一种深度学习架构，CNN是一种前馈(feed-forward)人工神经网络，该前馈人工神经网络中的各个神经元可以对输入其中的图像作出响应。卷积神经网络包含了一个由卷积层和池化层构成的特征抽取器。该特征抽取器可以看作是滤波器，卷积过程可以看作是使用一个可训练的滤波器与一个输入的图像或者卷积特征平面(feature map)做卷积。

[0139] 卷积层是指卷积神经网络中对输入信号进行卷积处理的神经元层。卷积层可以包括很多个卷积算子，卷积算子也称为核，其在图像处理中的作用相当于一个从输入图像矩阵中提取特定信息的过滤器，卷积算子本质上可以是一个权重矩阵，这个权重矩阵通常被预先定义，在对图像进行卷积操作的过程中，权重矩阵通常在输入图像上沿着水平方向一个像素接着一个像素(或两个像素接着两个像素……这取决于步长stride的取值)的进行处理，从而完成从图像中提取特定特征的工作。该权重矩阵的大小应该与图像的大小相关，需要注意的是，权重矩阵的纵深维度(depth dimension)和输入图像的纵深维度是相同的，在进行卷积运算的过程中，权重矩阵会延伸到输入图像的整个深度。因此，和一个单一的权重矩阵进行卷积会产生一个单一纵深维度的卷积化输出，但是大多数情况下不使用单一权重矩阵，而是应用多个尺寸(行×列)相同的权重矩阵，即多个同型矩阵。每个权重矩阵的输出被堆叠起来形成卷积图像的纵深维度，这里的维度可以理解为由上面所述的“多个”来决定。不同的权重矩阵可以用来提取图

像中不同的特征，例如一个权重矩阵用来提取图像边缘信息，另一个权重矩阵用来提取图像的特定颜色，又一个权重矩阵用来对图像中不需要的噪点进行模糊化等。该多个权重矩阵尺寸(行×列)相同，经过该多个尺寸相同的权重矩阵提取后的特征图的尺寸也相同，再将提取到的多个尺寸相同的特征图合并形成卷积运算的输出。这些权重矩阵中的权重值在实际应用中需要经过大量的训练得到，通过训练得到的权重值形成的各个权重矩阵可以用来从输入图像中提取信息，从而使得卷积神经网络进行正确的预测。当卷积神经网络有多个卷积层的时候，初始的卷积层往往提取较多的一般特征，该一般特征也可以称之为低级别的特征；随着卷积神经网络深度的加深，越往后的卷积层提取到的特征越来越复杂，比如高级别的语义之类的特征，语义越高的特征越适用于待解决的问题。

[0140] 由于常常需要减少训练参数的数量，因此卷积层之后常常需要周期性的引入池化层，可以是一层卷积层后面跟一层池化层，也可以是多层卷积层后面接一层或多层池化层。在图像处理过程中，池化层的唯一目的就是减少图像的空间大小。池化层可以包括平均池化算子和/或最大池化算子，以用于对输入图像进行采样得到较小尺寸的图像。平均池化算子可以在特定范围内对图像中的像素值进行计算产生平均值作为平均池化的结果。最大池化算子可以在特定范围内取该范围内值最大的像素作为最大池化的结果。另外，就像卷积层中用权重矩阵的大小应该与图像尺寸相关一样，池化层中的运算符也应该与图像的大小相关。通过池化层处理后输出的图像尺寸可以小于输入池化层的图像的尺寸，池化层输出的图像中每个像素点表示输入池化层的图像的对应子区域的平均值或最大值。

[0141] 在经过卷积层/池化层的处理后，卷积神经网络还不足以输出所需要的输出信息。因为如前所述，卷积层/池化层只会提取特征，并减少输入图像带来的参数。然而为了生成最终的输出信息(所需要的类信息或其他相关信息)，卷积神经网络需要利用神经网络层来生成一个或者一组所需要的类的数量的输出。因此，在神经网络层中可以包括多层隐含层，该多层隐含层中所包含的参数可以

根据具体的任务类型的相关训练数据进行预先训练得到，例如该任务类型可以包括图像识别，图像分类，图像超分辨率重建等等。

[0142] 可选的，在神经网络层中的多层隐含层之后，还包括整个卷积神经网络的输出层，该输出层具有类似分类交叉熵的损失函数，具体用于计算预测误差，一旦整个卷积神经网络的前向传播完成，反向传播就会开始更新前面提到的各层的权重值以及偏差，以减少卷积神经网络的损失，及卷积神经网络通过输出层输出的结果和理想结果之间的误差。

[0143] (4)循环神经网络

[0144] 循环神经网络(recurrent neural networks，RNN)是用来处理序列数据的。在传统的神经网络模型中，是从输入层到隐含层再到输出层，层与层之间是全连接的，而对于每一层层内之间的各个节点是无连接的。这种普通的神经网络虽然解决了很多难题，但是却仍然对很多问题却无能无力。例如，你要预测句子的下一个单词是什么，一般需要用到前面的单词，因为一个句子中前后单词并不是独立的。RNN之所以称为循环神经网路，即一个序列当前的输出与前面的输出也有关。具体的表现形式为网络会对前面的信息进行记忆并应用于当前输出的计算中，即隐含层本层之间的节点不再无连接而是有连接的，并且隐含层的输入不仅包括输入层的输出还包括上一时刻隐含层的输出。理论上，RNN能够对任何长度的序列数据进行处理。对于RNN的训练和对传统的CNN或DNN的训练一样。同样使用误差反向传播算法，不过有一点区别：即，如果将RNN进行网络展开，那么其中的参数，如W，是共享的；而如上举例上述的传统神经网络却不是这样。并且在使用梯度下降算法中，每一步的输出不仅依赖当前步的网络，还依赖前面若干步网络的状态。该学习算法称为基于时间的反向传播算法(Back propagation Through Time，BPTT)。

[0145] 既然已经有了卷积神经网络，为什么还要循环神经网络？原因很简单，在卷积神经网络中，有一个前提假设是：元素之间是相互独立的，输入与输出也是独立的，比如猫和狗。但现实世界中，很多元素都是相互连接的，比如股票随时间的变化，再比如一个人说了：我喜欢旅游，其中最喜欢的地方是云南，以后有机会一定要去。这里填空，人类应该都知道是填“云南”。因为人类会根据

上下文的内容进行推断，但如何让机器做到这一步，RNN就应运而生了。RNN旨在让机器像人一样拥有记忆的能力。因此，RNN的输出就需要依赖当前的输入信息和历史的记忆信息。

[0146] (5)损失函数

[0147] 在训练深度神经网络的过程中，因为希望深度神经网络的输出尽可能的接近真正想要预测的值，所以可以通过比较当前网络的预测值和真正想要的目标值，再根据两者之间的差异情况来更新每一层神经网络的权重向量(当然，在第一次更新之前通常会有初始化的过程，即为深度神经网络中的各层预先配置参数)，比如，如果网络的预测值高了，就调整权重向量让它预测低一些，不断的调整，直到深度神经网络能够预测出真正想要的目标值或与真正想要的目标值非常接近的值。因此，就需要预先定义“如何比较预测值和目标值之间的差异”，这便是损失函数(loss function)或目标函数(objective function)，它们是用于衡量预测值和目标值的差异的重要方程。其中，以损失函数举例，损失函数的输出值(loss)越高表示差异越大，那么深度神经网络的训练就变成了尽可能缩小这个loss的过程。

[0148] (6)反向传播算法

[0149] 卷积神经网络可以采用误差反向传播(back propagation，BP)算法在训练过程中修正初始的超分辨率模型中参数的大小，使得超分辨率模型的重建误差损失越来越小。具体地，前向传递输入信号直至输出会产生误差损失，通过反向传播误差损失信息来更新初始的超分辨率模型中参数，从而使误差损失收敛。反向传播算法是以误差损失为主导的反向传播运动，旨在得到最优的超分辨率模型的参数，例如权重矩阵。

[0150] (7)生成式对抗网络

[0151] 生成式对抗网络(generative adversarial networks，GAN)是一种深度学习模型。该模型中至少包括两个模块：一个模块是生成模型(Generative Model)，另一个模块是判别模型(Discriminative Model)，通过这两个模块互相博弈学习，从而产生更好的输出。生成模型和判别模型都可以是神经网络，具体可以是深度神经网络，或者卷积神经网络。GAN的基本原理如下：以生成图片的GAN为例，假设

有两个网络，G(Generator)和D(Discriminator)，其中G是一个生成图片的网络，它接收一个随机的噪声z，通过这个噪声生成图片，记做G(z)；D是一个判别网络，用于判别一张图片是不是“真实的”。它的输入参数是x，x代表一张图片，输出D(x)代表x为真实图片的概率，如果为1，就代表100%是真实的图片，如果为0，就代表不可能是真实的图片。在对该生成式对抗网络进行训练的过程中，生成网络G的目标就是尽可能生成真实的图片去欺骗判别网络D，而判别网络D的目标就是尽量把G生成的图片和真实的图片区分开来。这样，G和D就构成了一个动态的“博弈”过程，也即“生成式对抗网络”中的“对抗”。最后博弈的结果，在理想的状态下，G可以生成足以“以假乱真”的图片G(z)，而D难以判定G生成的图片究竟是不是真实的，即D(G(z))＝0.5。这样就得到了一个优异的生成模型G，它可以用来生成图片。

[0152] 基于此，下面结合附图，对本申请实施例的技术方案进行描述。

[0153] 本申请实施例可以应用于终端设备、无线接入网设备、核心网设备等的音频编/解码器中。

[0154] 图1示出了本申请实施例的一种应用场景100的示意图。图1所示的是本申请实施例在终端设备侧应用的系统架构示意图。如图1所示，包括第一终端设备110、第二终端设备120、无线或者有线网络通信设备130和无线或者有线网络通信设备140。其中，所述第一终端设备110和所述第二终端设备120可以为发送端设备，也可以为接收端设备。作为示例，以所述第一终端设备110为发送端设备，所述第二终端设备120为接收端设备为例进行说明。

[0155] 在音频通信中，第一终端设备110中的音频采集模块用于采集音频，音频编码器将采集到的音频信号进行音频编码，信道编码模块进行信道编码，获得码流，然后通过发送端侧的所述无线或者有线网络通信设备130进行在数字信道中传输信号。

[0156] 接收端侧的所述无线或者有线网络通信设备140通过数字信道获取所述第一终端设备110发送的信号，并将所述信号传送给所述第二终端设备120，所述第二终端设备120根据接收到的信号，在信道解码模块进行信道解码，再经过音频解

码器解码出音频信号，然后根据该解码出的音频信号在音频回放模块进音频回放。

[0157] 应理解，当所述第二终端设备120为发送端设备，所述第一终端设备110为接收端设备时可以相应参考所述第一终端设备110为发送端设备，所述第二终端设备120为接收端设备进行理解，此处不再赘述。

[0158] 应理解，所述无线或者有线网络通信设备130和所述无线或者有线网络通信设备140也可以是核心网设备。

[0159] 图2示出了本申请实施例的另一种应用场景200的示意图。图2所示的是本申请实施例在无线接入网设备或者核心网设备转码应用的系统架构示意图。如图2所示，无线接入网设备或者核心网设备包括信道解码模块、其他音频解码器、音频编码器和信道编码模块。其中，转码时需要进行相应的音频编解码处理。无线接入网设备或者核心网设备在信道解码模块对接收到的信号进行信道解码，然后通过其他音频解码器对音频码流解码，得到音频码流，音频编码器对音频码流重新编码，然后再进行信道编码进而传输音频信号。

[0160] 图3示出了本申请实施例的另一种应用场景300的示意图。图3所示的是本申请实施例在无线接入网设备或者核心网设备转码应用的系统架构示意图。如图3所示，无线接入网设备或者核心网设备包括信道解码模块、音频解码器、其他音频编码器和信道编码模块。其中，转码时需要进行相应的音频编解码处理。无线接入网设备或者核心网设备在信道解码模块对接收到的信号进行信道解码，然后通过音频解码器对音频码流解码，得到音频码流，其他音频编码器对音频码流重新编码，然后再进行信道编码进而传输音频信号。

[0161] 本申请实施例还可以适用于虚拟现实(Virtual Reality，VR)串流streaming服务中的音频编码模块(Audio Encoding)和音频解码模块(Audio Decoding)。图4示出了本申请实施例的另一种应用场景400的示意图，如图4所示中虚框部分所示。

[0162] 端到端对音频信号的处理流程为：发送端对音频信号和视频信号经过采集模块(Acquisition)处理后，分成音频信号和视频信号，音频信号进行预处理操作(Audio Preprocessing)，预处理操作包括滤除掉音频信号中的低频部分，通常是以20Hz或者50Hz为分界点，提取信号中的方位信息，之后进行音频编码处理(Audio e

ncoding)；视频信号经过视觉缝合以及投影和映射，之后进行视频编码和图像编码。将音频码流、视频码流和图像码流打包(文件File/段Segment封装encapsulation)之后发送(Delivery)到解码端。

[0163] 解码端首先进行解包(File/Segment解封装decapsulation)，之后分别进行音频解码(Audio decoding)、视频解码和图像解码，对解码后的音频信号进行音频双耳渲染(Audio rendering)处理，渲染处理后的信号映射到收听者耳机(headphones)上，可以为独立的耳机也可以是HTC VIVE等眼镜设备上的耳机，对解码后的视频信号和图像信号进行视频双耳渲染(video rendering)处理，渲染处理后的信号映射到显示器(display)上。

[0164] 本申请实施例中的终端设备也可以称为：用户设备(user equipment，UE)、移动台(mobile station，MS)、移动终端(mobile terminal，MT)、接入终端、用户单元、用户站、移动站、移动台、远方站、远程终端、移动设备、用户终端、终端、无线通信设备、用户代理或用户装置等。

[0165] 终端设备可以是无线终端也可以是有线终端，无线终端可以是指向用户提供语音和/或其他业务数据连通性的设备，具有无线连接功能的手持式设备、或连接到无线调制解调器的其他处理设备。无线终端可以经无线接入网(Radio Access Network，RAN)与一个或多个核心网进行通信，无线终端可以是移动终端，如移动电话(或称为“蜂窝”电话)和具有移动终端的计算机，例如，可以是便携式、袖珍式、手持式、计算机内置的或者车载的移动装置，它们与无线接入网交换语言和/或数据。例如，个人通信业务(Personal Communication Service，PCS)电话、无绳电话、会话发起协议(Session Initiation Protocol，SIP)话机、无线本地环路(Wireless Local Loop，WLL)站、个人数字助理(Personal Digital Assistant，PDA)等设备。无线终端也可以称为系统、订户单元(Subscriber Unit)、订户站(Subscriber Station)，移动站(Mobile Station)、移动台(Mobile)、远程站(Remote Station)、远程终端(Remote Terminal)、接入终端(Access Terminal)、用户终端(User Terminal)、用户代理(User Agent)、用户设备(User Device or User Equipment)，移动互联网设备(mobile internet device，MID)、可穿戴设备，虚拟现实(virtual reality，VR)设备、增强现实(augmented reality，AR)设备、工业控制(industrial control)中的无

线终端、无人驾驶(self driving)中的无线终端、远程手术(remote medical surgery)中的无线终端、智能电网(smart grid)中的无线终端、运输安全(transportation safety)中的无线终端、智慧城市(smart city)中的无线终端、智慧家庭(smart home)中的无线终端、车载设备、可穿戴设备，5G网络中的终端设备或者未来演进的公用陆地移动通信网络(public land mobile network，PLMN)中的终端设备等，本申请实施例对此并不限定。

[0166] 作为示例而非限定，在本申请实施例中，可穿戴设备也可以称为穿戴式智能设备，是应用穿戴式技术对日常穿戴进行智能化设计、开发出可以穿戴的设备的总称，如眼镜、手套、手表、服饰及鞋等。可穿戴设备即直接穿在身上，或是整合到用户的衣服或配件的一种便携式设备。可穿戴设备不仅仅是一种硬件设备，更是通过软件支持以及数据交互、云端交互来实现强大的功能。广义穿戴式智能设备包括功能全、尺寸大、可不依赖智能手机实现完整或者部分的功能，例如：智能手表或智能眼镜等，以及只专注于某一类应用功能，需要和其它设备如智能手机配合使用，如各类进行体征监测的智能手环、智能首饰等。

[0167] 此外，在本申请实施例中，终端设备还可以是物联网(internet of things，IoT)系统中的终端设备，IoT是未来信息技术发展的重要组成部分，其主要技术特点是将物品通过通信技术与网络连接，从而实现人机互连，物物互连的智能化网络。

[0168] 而如上介绍的各种终端设备，如果位于车辆上(例如放置在车辆内或安装在车辆内)，都可以认为是车载终端设备，车载终端设备例如也称为车载单元(on-board unit，OBU)。

[0169] 本申请实施例中，终端设备还可以包括中继(relay)。或者理解为，能够与基站进行数据通信的都可以看作终端设备。

[0170] 示例性的，图5为本申请提供的一种终端设备500的结构示意图。该终端设备500可以执行上述实施例中终端设备执行的动作。

[0171] 为了便于说明，图5仅示出了终端设备的主要部件。如图5所示，终端设备500包括处理器、存储器、控制电路、天线以及输入输出装置。

[0172] 处理器主要用于对通信协议以及通信数据进行处理，以及对整个终端设备进行控制，执行软件程序，处理软件程序的数据，例如用于支持终端设备执行上述传输预编码矩阵的指示方法实施例中所描述的动作。存储器主要用于存储软件程序和数据，例如存储上述实施例中所描述的码本。控制电路主要用于基带信号与射频信号的转换以及对射频信号的处理。控制电路和天线一起也可以叫做收发器，主要用于收发电磁波形式的射频信号。输入输出装置，例如触摸屏、显示屏，键盘等主要用于接收用户输入的数据以及对用户输出数据。

[0173] 当终端设备开机后，处理器可以读取存储单元中的软件程序，解释并执行软件程序的指令，处理软件程序的数据。当需要通过无线发送数据时，处理器对待发送的数据进行基带处理后，输出基带信号至射频电路，射频电路将基带信号进行射频处理后将射频信号通过天线以电磁波的形式向外发送。当有数据发送到终端设备时，射频电路通过天线接收到射频信号，将射频信号转换为基带信号，并将基带信号输出至处理器，处理器将基带信号转换为数据并对该数据进行处理。

[0174] 本领域技术人员可以理解，为了便于说明，图5仅示出了一个存储器和处理器。在实际的终端设备中，可以存在多个处理器和存储器。存储器也可以称为存储介质或者存储设备等，本申请实施例对此不做限制。

[0175] 例如，处理器可以包括基带处理器和中央处理器，基带处理器主要用于对通信协议以及通信数据进行处理，中央处理器主要用于对整个终端设备进行控制，执行软件程序，处理软件程序的数据。图5中的处理器集成了基带处理器和中央处理器的功能，本领域技术人员可以理解，基带处理器和中央处理器也可以是各自独立的处理器，通过总线等技术互联。本领域技术人员可以理解，终端设备可以包括多个基带处理器以适应不同的网络制式，终端设备可以包括多个中央处理器以增强其处理能力，终端设备的各个部件可以通过各种总线连接。所述基带处理器也可以表述为基带处理电路或者基带处理芯片。所述中央处理器也可以表述为中央处理电路或者中央处理芯片。对通信协议以及通信数据进行处理的功能可以内置在处理器中，也可以以软件程序的形式存储在存储单元中，由处理器执行软件程序以实现基带处理功能。

[0176] 示例性的，在本申请实施例中，可以将具有收发功能的天线和控制电路视为终端设备500的收发单元510，将具有处理功能的处理器视为终端设备500的处理单元520，处理单元520也可以实现获取单元的功能。如图5所示，终端设备500包括收发单元510和处理单元520。收发单元也可以称为收发器、收发机、收发装置等。可选的，可以将收发单元510中用于实现接收功能的器件视为接收单元，将收发单元510中用于实现发送功能的器件视为发送单元，即收发单元包括接收单元和发送单元。示例性的，接收单元也可以称为接收机、接收器、接收电路等，发送单元可以称为发射机、发射器或者发射电路等。

[0177] 本申请实施例中的接入网设备可以是用于与终端设备通信的设备，可以是基站，或者接入点，或者网络设备，或者可以是指接入网中在空中接口上通过一个或多个扇区与无线终端通信的设备。网络设备可用于将收到的空中帧与IP分组进行相互转换，作为无线终端与接入网的其余部分之间的路由器，其中接入网的其余部分可包括网际协议(IP)网络。网络设备还可协调对空中接口的属性管理。例如，接入网设备可以是全球移动通讯(Global System of Mobile communication，GSM)或码分多址(Code Division Multiple Access，CDMA)中的基站(Base Transceiver Station，BTS)，也可以是宽带码分多址(Wideband Code Division Multiple Access，WCDMA)中的基站(NodeB，NB)，也可以是LTE系统中的演进型基站(evolved NodeB，eNB或eNodeB)，还可以是云无线接入网络(cloud radio access network，CRAN)场景下的无线控制器，或者该接入设备可以为中继站、接入点、车载设备、可穿戴设备以及5G网络中的接入设备或者未来演进的PLMN网络中的网络设备等，可以是WLAN中的接入点(access point，AP),可以是新型无线系统(new radio，NR)系统中的gNB本申请实施例并不限定。需要说明的是，对于5G系统，在一个基站下，可能存在一个或多个发送接收点(Transmission Reception Point，TRP)，所有的TRP属于同一个小区，其中，每个TRP和终端都可以使用本申请实施例所述的测量上报方法。在另一种场景下，网络设备还可以分为控制单元(Control Unit，CU)和数据单元(Data Unit，DU)，在一个CU下，可以存在多个DU，其中，每个DU和终端都可以使用本申请实施例所述的测量上报方

法。CU-DU分离场景和多TRP场景的区别在于，TRP只是一个射频单元或一个天线设备，而DU中可以实现协议栈功能，例如DU中可以实现物理层功能。

[0178] 另外，在本申请实施例中，接入网设备是接入网(radio access network，RAN)中的设备，或者说，是将终端设备接入到无线网络的RAN节点。例如，作为示例而非限定，作为接入网设备，可以列举：gNB、传输接收点(transmission reception point，TRP)、演进型节点B(evolved Node B，eNB)、无线网络控制器(radio network controller，RNC)、节点B(Node B，NB)、基站控制器(base station controller，BSC)、基站收发台(base transceiver station，BTS)、家庭基站(例如，home evolved NodeB，或home Node B，HNB)、基带单元(base band unit，BBU)，或无线保真(wireless fidelity，Wifi)接入点(access point，AP)等。

[0179] 接入网设备为小区提供服务，终端设备通过该小区使用的传输资源(例如，频域资源，或者说，频谱资源)与接入网设备进行通信，该小区可以是接入网设备(例如基站)对应的小区，小区可以属于宏基站，也可以属于小小区(small cell)对应的基站，这里的小小区可以包括：城市小区(metro cell)、微小区(micro cell)、微微小区(pico cell)、毫微微小区(femto cell)等，这些小小区具有覆盖范围小、发射功率低的特点，适用于提供高速率的数据传输服务。

[0180] 示例性的，图6为本申请实施例提供的一种接入网设备600的结构示意图，可以用于实现上述实施例中的无线接入网设备的功能。接入网设备600包括一个或多个射频单元，如远端射频单元(remote radio unit,RRU)66和一个或多个基带单元(baseband unit，BBU)(也可称为数字单元，digital unit，DU)620。所述RRU66可以称为收发单元、收发机、收发电路、或者收发器等等，其可以包括至少一个天线611和射频单元612。所述RRU66部分主要用于射频信号的收发以及射频信号与基带信号的转换，例如用于向终端设备发送上述实施例中所述的信令消息。所述BBU620部分主要用于进行基带处理，对基站进行控制等。所述RRU66与BBU620可以是物理上设置在一起，也可以物理上分离设置的，即分布式基站。

[0181] 所述BBU620为基站的控制中心，也可以称为处理单元，主要用于完成基带处理功能，如信道编码，复用，调制，扩频等等。例如该BBU(处理单元)620可以用于控制接入网设备执行上述方法实施例中关于接入网设备的操作流程。

[0182] 在一个示例中，所述BBU620可以由一个或多个单板构成，多个单板可以共同支持单一接入制式的无线接入网(如LTE系统，或5G系统，或未来的无线接入网系统等等)，也可以分别支持不同接入制式的无线接入网。所述BBU620还包括存储器621和处理器622。所述存储器621用以存储必要的指令和数据。例如存储器621存储上述实施例中的码本等。所述处理器622用于控制基站进行必要的动作，例如用于控制基站执行上述方法实施例中关于网络设备的操作流程。所述存储器621和处理器622可以服务于一个或多个单板。也就是说，可以每个单板上单独设置存储器和处理器。也可以是多个单板共用相同的存储器和处理器。此外每个单板上还可以设置有必要的电路。

[0183] 在一种可能的实施方式中，随着片上系统(system-on-chip，SoC)技术的发展，可以将620部分和66部分的全部或者部分功能由SoC技术实现，例如由一颗基站功能芯片实现，该基站功能芯片集成了处理器、存储器、天线接口等器件，基站相关功能的程序存储在存储器中，由处理器执行程序以实现基站的相关功能。可选的，该基站功能芯片也能够读取该芯片外部的存储器以实现基站的相关功能。

[0184] 应理解，图6示例的接入网设备的结构仅为一种可能的形态，而不应对本申请实施例构成任何限定。本申请并不排除未来可能出现的其他形态的基站结构的可能。

[0185] 应理解，本申请实施例中，该处理器可以为中央处理单元(central processing unit，CPU)，该处理器还可以是其他通用处理器、数字信号处理器(digital signal processor，DSP)、专用集成电路(application specific integrated circuit，ASIC)、现成可编程门阵列(field programmable gate array，FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0186] 还应理解，本申请实施例中的存储器可以是易失性存储器或非易失性存储器，或可包括易失性和非易失性存储器两者。其中，非易失性存储器可以是只读存储器(read-only memory，ROM)、可编程只读存储器(programmable ROM，PROM)、可擦除可编程只读存储器(erasable PROM，EPROM)、电可擦除可编程只读存储器(electrically EPROM，EEPROM)或闪存。易失性存储器可以是随机存取存储器(random access memory，RAM)，其用作外部高速缓存。通过示例性但不是限制性说明，许多形式的随机存取存储器(random access memory，RAM)可用，例如静态随机存取存储器(static RAM，SRAM)、动态随机存取存储器(DRAM)、同步动态随机存取存储器(synchronous DRAM，SDRAM)、双倍数据速率同步动态随机存取存储器(double data rate SDRAM，DDR SDRAM)、增强型同步动态随机存取存储器(enhanced SDRAM，ESDRAM)、同步连接动态随机存取存储器(synchlink DRAM，SLDRAM)和直接内存总线随机存取存储器(direct rambus RAM，DR RAM)。

[0187] 核心网设备可以是核心网网元，例如接入与移动性管理功能(Access and Mobility Management Function，AMF)实体、会话管理功能(Session Management Function，SMF)实体、用户面功能(User Plane Function，UPF)实体、策略控制功能(Policy Control function，PCF)实体等。AMF实体提供核心网中的移动性管理功能，主要负责接入和移动性控制，包括注册管理(registration management，RM)和连接管理(connection management，CM)、接入鉴权和接入授权、可达性管理和移动性管理等。SMF实体是核心网中的会话管理功能，AMF实体在对终端设备进行移动性管理之外，还负责将从会话管理相关消息在终端设备和SMF实体之间的转发。PCF实体是核心网中的策略管理功能，负责制定对终端设备的移动性管理、会话管理、计费等相关的策略。UPF实体是核心网中的用户面功能，通过接口与外部数据网络进行数据传输，通过接口与接入网设备进行数据传输，主要提供用户面支持，包括PDU会话和数据网络的连接点、数据包路由和转发、数据包检测和用户面策略执行、为用户面处理QoS、下行数据包缓存和下行数据通知的触发等。

[0188] 应理解，上述核心网的功能单元可以独立工作，也可以组合在一起实现某些控制功能，如：AMF、SMF和PCF可以组合在一起作为管理设备，用于完成终端设备的接入鉴权、安全加密、位置注册等接入控制和移动性管理功能，以及用户面传输路径的记录、释放和更改等会话管理功能，以及分析一些切片(slice)相关的数据(如拥塞)、终端设备相关的数据的功能，UPF作为网关设备主要完成用户面数据的路由转发等功能，如：负责对终端设备的数据报文过滤、数据传输/转发、速率控制、生成计费信息等。

[0189] 本申请实施例的技术方案可以应用于各种通信系统，例如：全球移动通信(global system for mobile communications，GSM)系统、码分多址(code division multiple access，CDMA)系统、宽带码分多址(wideband code division multiple access，WCDMA)系统、通用分组无线业务(general packet radio service，GPRS)、长期演进(long term evolution，LTE)系统、LTE频分双工(frequency division duplex，FDD)系统、LTE时分双工(time division duplex，TDD)、通用移动通信系统(universal mobile telecommunication system，UMTS)、全球互联微波接入(worldwide interoperability for microwave access，WiMAX)通信系统、第五代(5th generation，5G)系统或新无线(new radio，NR)，此外，还可以适用于使用后续的演进系统，如第六代6G通信系统、甚至更高级的第七代7G通信系统等。

[0190] 核心编码器中，为了降低编码比特率，可以利用不同频带信号间的相关性进行编码。利用不同频带信号间的相关性进行编码，其基本原理是利用低频带信号，通过频带扩展(Bandwidth Extension，BWE)的方法，产生高频带信号。无论是3GPP最新的增强语音服务(Enhanced Voice Services，EVS)音频编解码器，还是动态图像专家组高效高级音频编码(Moving Picture Experts Group High-Efficiency Advanced Audio Coding，MPEG HE-AAC)、统一语音和音频编码(Unified Speech and Audio Coding，USAC)音频编解码器，都利用了不同频带信号间的相关性，采用频带扩展技术，对高频带信号进行编码，以便用较少的比特数对高频带信号进行编码，从而降低编码器的编码比特率。但真实的音频信号中，高频带的频谱中往往存在一些与低频带的频谱不相似或关联性弱的频谱，从而导致高频带信号重建失真。

[0191] 为了解决上述技术问题，本申请提出了一种音频编解码方法和装置。

[0192] 需要说明的是，下面的方法实施例可以应用在图1-图4所示的应用场景中，当然也可以应用在其他场景中，本申请实施例在此不作限制。还应理解，在本申请实施例中，执行该方法的执行主体可以是终端设备、接入网设备和核心网设备。作为示例而非限定，执行该方法的执行主体也可以是应用于终端设备、接入网设备和核心网设备的芯片、芯片系统、或处理器等。其中，终端设备、接入网设备和核心网设备具备编解码功能，也可以称为编解码设备。

[0193] 图7为本申请实施例提供的音频编码方法的过程700的流程图。过程700描述为一系列的步骤或操作，应当理解的是，过程700可以以各种顺序执行和/或同时发生，不限于图7所示的执行顺序。过程700可以包括：

[0194] 步骤701、获取第一码流，该第一码流是基于第一音频信号得到的。

[0195] 原始音频经采样得到音频帧序列，其中包括多个音频帧，在编码端可以一帧一帧的进行压缩处理，为便于描述，本申请实施例中将当前正在处理的音频帧称作当前帧。

[0196] 当前帧包含的音频信号的特征可以通过频率、幅度和相位来表征，本申请实施例根据频率从当前帧中提取出第一音频信号和第二音频信号，其中，第一音频信号是当前帧中频率低于预设阈值的音频信号(本申请实施例中，满足该条件的音频信号亦可以称作低频信号或低频成分)，第二音频信号包括当前帧中频率高于或等于预设阈值的音频信号(本申请实施例中，满足该条件的音频信号亦可以称作高频信号或高频成分)。前述预设阈值例如可以是16kHz，本申请实施例对此不做具体限定。需要说明的是，前文将当前帧中频率等于预设阈值的音频信号归入第二音频信号，可选的，本申请实施例也可以将当前帧中频率等于预设阈值的音频信号归入第一音频信号，对此不做具体限定。

[0197] 一些可能的实施方式中，可以采用低通滤波器从当前帧中提取出第一音频信号。

[0198] 一些可能的实施方式中，可以采用核心编码器对第一音频信号进行编码，该核心编码器例如可以采用AVS3核心编码器，本申请实施例对此不做具体限定。

[0199] 步骤702、获取边信息，该边信息是基于第二音频信号得到的。

[0200] 本申请实施例中，第二音频信号可能存在以下三种情况：

[0201] 1)第二音频信号只包括高频成分，该情况下第二音频信号等同于当前帧的高频成分。与之对应，边信息可以只包含高频成分的特性信息；

[0202] 2)第二音频信号除了包括高频成分外，还包括低频成分中的部分音频信号，亦即，第二音频信号和第一音频信号有部分相同的音频信号，该情况下第二音频信号等同于当前帧的高频成分加部分低频成分。与之对应，边信息中除了包含高频成分的特性信息外，还可以包含高频成分与前述部分低频成分之间的关联性信息；

[0203] 3)第二音频信号除了包括高频成分外，还包括低频成分中的全部音频信号，亦即，第二音频信号包括第一音频信号，该情况下第二音频信号等同于当前帧。与之对应，边信息中除了包含高频成分的特征信息外，还可以包含高频成分与低频成分之间的关联性信息。

[0204] 由此可见，边信息可以仅占用少量的传输比特，就可以指导解码端实现非盲BWE，以重建出高频成分，尤其是高低频之间映射关系复杂的信号。

[0205] 一些可能的实施方式中，获取边信息可以包括将第二音频信号输入预先训练得到的第二机器学习模型，以获取边信息，该第二机器学习模型包括至少一个子模型。至少一个子模型可以分别专注于不同的特性，例如，为各个子模型分别设置限制条件，包括输入输出、模型结构、处理规则等，这样每个子模型可以专注于提取对应特性的边信息。

[0206] 本申请实施例中，第二机器学习模型可以参照上文的神经网络、深度神经网络、卷积神经网络、循环神经网络等。机器学习模型可以将当前帧中高频成分的更多细节体现于特征信息中，这些细节例如是高频成分的特性，高频成分与低频成分之间的关联等，有助于重建出更多的高频细节。

[0207] 示例性的，图8a为本申请实施例的获取边信息的示意图，如图8a所示，获取边信息通过第二机器学习模型实现，该第二机器学习模型包括一个子模型，即第二机器学习模型亦即该子模型，其输入为第二音频信号，输出为边信息。

[0208] 一些可能的实施方式中，第二机器学习模型可以包括多个子模型，该多个子模型各自的输入包括当前帧中频率高于或等于预设阈值的音频信号，即多个子模型各自的输入是第二音频信号中的部分或全部音频信号。

[0209] 示例性的，图8b为本申请实施例的获取边信息的示意图，如图8b所示，获取边信息通过第二机器学习模型实现，该第二机器学习模型包括多个子模型(图8b中以n个子模型为例，n＞1)，即子模型1、子模型2、…、子模型n，其中，子模型1的输入为输入1、输出为特征信息1，子模型2的输入为输入2、输出为特征信息2，…，子模型n的输入为输入n、输出为特征信息n。

[0210] 输入1、输入2、…、输入n均来自第二音频信号，其共同之处在于均包括第二音频信号中的高频成分。此外，输入1、输入2、…、输入n分别还可以包括或不包括第二音频信号中的低频成分。当包括第二音频信号中的低频成分时，输入1、输入2、…、输入n分别包括第二音频信号中的部分或全部低频成分。例如，输入1包括第二音频信号中的高频成分，输入2包括第二音频信号中的高频成分和部分低频成分，输入3包括第二音频信号中的高频成分和全部低频成分。需要说明的是，本申请实施例中，输入1、输入2、…、输入n可以相同，也可以完全不同，还可以部分相同、部分不同，对此不做具体限定。

[0211] 本申请实施例中，若子模型的输入只包含高频成分，则该子模型主要用于将高频成分压缩为特征，或者通过特征来表示高频成分；若子模型的输入除了高频成分外还包含低频成分，则该子模型还用于提取高低频之间的映射关系。采用多个子模型可以在边信息的提取中得到更多样的特征，因此本申请对子模型的数量、类型、功能等均不做具体限定。

[0212] 上述n个子模型各自的输出，包括特征信息1、特征信息2、…、特征信息n，经合并处理后得到边信息，亦即，边信息包括n个子模型分别输出的特征信息。

[0213] 一些可能的实施方式中，第一子模型与多个子模型中的其他至少一个子模型包括相同的神经元，该相同的神经元位于第一子模型的输入阶段、输出阶段或者中间阶段，该第一子模型是多个子模型中的任意一个。

[0214] 第一子模型可以和第二机器学习模型包括的多个子模型中的一个或多个有相同的神经元(亦称作共用神经元)，该相同的神经元可以有一层或多层，而该一层或

多层神经元可以位于子模型的输入阶段、输出阶段或中间阶段，其中，输入阶段是指该子模型的输入端的第一层及其之后的一层或多层，输出阶段是指该子模型的输出端的最后一层及其之前的一层或多层，中间阶段是指不包括子模型的第一层和最后一层的一层或多层。需要说明的是，本申请实施例对前述相同的神经元的数量、类型等不做具体限定，对输入阶段、输出阶段以及中间阶段的位置也不做具体限定。

[0215] 示例性的，图9a为本申请实施例的两个子模型的相同部分的示意图，如图9a所示，以两个子模型为例，该两个子模型有相同的神经元(亦即相同部分)，各自又有不同的神经元(亦即差异部分)。

[0216] 相同部分位于子模型1和子模型2的输入阶段，子模型1的差异部分1、子模型2的差异部分2分别位于前述相同部分之后，分别输出特征信息1和特征信息2。

[0217] 示例性的，图9b为本申请实施例的两个子模型的相同部分的示意图，如图9b所示，以两个子模型为例，该两个子模型有相同的神经元(亦即相同部分)，各自又有不同的神经元(亦即差异部分)。

[0218] 相同部分位于子模型1和子模型2的输出阶段，子模型1的差异部分1、子模型2的差异部分2分别位于前述相同部分之前，两个子模型分别输出特征信息1和特征信息2。

[0219] 示例性的，图9c为本申请实施例的两个子模型的相同部分的示意图，如图9c所示，以两个子模型为例，该两个子模型有相同的神经元(亦即相同部分)，各自又有不同的神经元(亦即差异部分)。

[0220] 相同部分位于子模型1和子模型2的输入阶段和输出阶段，子模型1的差异部分1、子模型2的差异部分2分别位于两个相同部分之间，两个子模型分别输出特征信息1和特征信息2。

[0221] 需要说明的是，上述两个示例以两个子模型为例进行说明，但这并非是对多个子模型之间的相同部分构成限定，本申请实施例对具有相同部分的多个子模型的数量、相同部分的具体实施方式(包括相同部分的位置、相同部分包括的层等)均不做具体限定。

[0222] 一些可能的实施方式中，第一子模型的输入还包括第一输入；当第二机器学习模型包括多个子模型时，第一输入包括多个子模型中的其他至少一个子模型的输出，或者，第一输入包括至少一个子规则方式的输出；第一子模型是多个子模型中的任意一个。

[0223] 示例性的，图10a为本申请实施例的子模型串联的示意图，如图10a所示，以两个子模型为例，该两个子模型中，子模型2的输出可以作为子模型1的一个输入，这样子模型1包括输入1和子模型2的输出。除此之外，子规则方式的输出也可以作为子模型1的一个输入，这样子模型1包括输入1和子规则方式的输出，甚至可以再加上子模型2的输出，对此不做具体限定。这可以是子模型与子模型或子规则方式的串联。

[0224] 一些可能的实施方式中，第一子模型的输入还包括第二输入，第二输入是第一输入经冗余信息处理后得到的音频信号；当第二机器学习模型包括多个子模型时，第一输入包括多个子模型中的其他至少一个子模型的输出，或者，第一输入包括至少一个子规则方式的输出；第一子模型是多个子模型中的任意一个。

[0225] 示例性的，图10b为本申请实施例的子模型串联的示意图，如图10b所示，以两个子模型为例，该两个子模型中，子模型2的输出先经过冗余信息处理，然后可以作为子模型1的一个输入，这样子模型1包括输入1和子模型2的经冗余信息处理后输出。除此之外，子规则方式的输出也可以先经过冗余信息处理，然后作为子模型1的一个输入，这样子模型1包括输入1和子规则方式的经冗余信息处理后输出，甚至可以再加上子模型2的经冗余信息处理后输出，对此不做具体限定。这可以是子模型与子模型或子规则方式的串联。

[0226] 需要说明的是，上述两个示例以两个子模型为例进行说明，但这并非是对子模型之间、子模型与子规则方式之间的串联关系构成限定，任意一个子模型的输入除了自己的输入外，还可以包括其他一个或多个子模型的输出或经冗余处理后的输出，和/或，一个或多个子规则方式的输出或经冗余处理后的输出，
本申请实施例不限定串联的个数，串联方式。

[0227] 一些可能的实施方式中，获取边信息还可以包括根据第二音频信号按照预先设置的规则方式进行处理以获取边信息，该预先设置的规则方式包括至少一个子规则方式。

[0228] 本申请实施例中，按照预先设置的规则方式提取特征的方法，例如可以包括提取频谱包络和参照AVS3提取特征，其中，频谱包络是当前帧的频谱中各频带的能量，用于给解码侧进行对应频带能量的恢复；AVS3提取的特征包括频带能量，以及对各频带谐波特性和噪声特性的标识，用于给解码侧恢复出BWE基底。相比于机器学习方式，规则方式可以保留频谱能量等信息，从而避免对听感的影响。

[0229] 本申请实施例中，在第二机器学习模型的基础上，还可以设置规则方式，对第二音频信号按照前述规则方式进行处理，得到与其对应的输出(特征信息)。

[0230] 示例性的，图8c为本申请实施例的获取边信息的示意图，如图8c所示，获取边信息通过第二机器学习模型和预先设置的规则方式实现，该第二机器学习模型包括一个子模型，即第二机器学习模型亦即该子模型，其输入为输入1，输出为特征信息1，规则方式包括一个子规则方式，即预先设置的规则方式亦即该子规则方式，其输入为输入2，输出为特征信息2。

[0231] 输入1和输入2均来自第二音频信号，其共同之处在于均包括第二音频信号中的高频成分。此外，输入1和输入2分别还可以包括或不包括第二音频信号中的低频成分。当包括第二音频信号中的低频成分时，输入1和输入2分别包括第二音频信号中的部分或全部低频成分。例如，输入1包括第二音频信号中的高频成分，输入2包括第二音频信号中的高频成分和部分低频成分。需要说明的是，本申请实施例中，输入1和输入2可以相同，也可以完全不同，还可以部分相同、部分不同，对此不做具体限定。

[0232] 第二机器学习模型输出的特征信息1和规则方式输出的特征信息2经过合并处理后得到边信息，亦即，边信息包括第二机器学习模型和按照规则方式分别输出的特征信息。

[0233] 一些可能的实施方式中，预先设置的规则方式可以包括多个子规则方式，该多个子规则方式各自的输入包括当前帧中频率高于或等于预设阈值的音频信号，即多个子规则方式各自的输入是第二音频信号中的部分或全部音频信号。

[0234] 示例性的，图8d为本申请实施例的获取边信息的示意图，如图8d所示，获取边信息通过第二机器学习模型和预先设置的规则方式实现，该第二机器学习模型包括多个子模型(图8d中以n个子模型为例，n＞1)，即子模型1、子模型2、…、子模型n，其中，子模型1的输入为输入1、输出为特征信息1，子模型2的输入为输入2、输出为特征信息2，…，子模型n的输入为输入n、输出为特征信息n。该规则方式包括多个子规则方式(图8d中以m个子规则方式为例，m＞1)，即子规则方式1的输入为输入n+1、输出为特征信息n+1，子规则方式2的输入为输入n+2、输出为特征信息n+2，…，子规则方式m的输入为输入n+m、输出为特征信息n+m。

[0235] 输入1、输入2、…、输入n、输入n+1、输入n+2、…、输入n+m均来自第二音频信号，其共同之处在于均包括第二音频信号中的高频成分。此外，输入1、输入2、…、输入n、输入n+1、输入n+2、…、输入n+m分别还可以包括或不包括第二音频信号中的低频成分。当包括第二音频信号中的低频成分时，输入1、输入2、…、输入n、输入n+1、输入n+2、…、输入n+m分别包括第二音频信号中的部分或全部低频成分。例如，输入1包括第二音频信号中的高频成分，输入2包括第二音频信号中的高频成分和部分低频成分，输入3包括第二音频信号中的高频成分和全部低频成分。需要说明的是，本申请实施例中，输入1、输入2、…、输入n、输入n+1、输入n+2、…、输入n+m可以相同，也可以完全不同，还可以部分相同、部分不同，对此不做具体限定。

[0236] 上述n个子模型、m个规则方式各自的输出，包括特征信息1、特征信息2、…、特征信息n、特征信息n+1、特征信息n+2、…、特征信息n+m，经合并处理后得到边信息，亦即，边信息包括n个子模型分别输出的特征信息和按照m个规则方式输出的特征信息。

[0237] 一些可能的实施方式中，第一子规则方式的输入还包括第三输入；当预先设置的规则方式包括多个子规则方式时，第三输入包括多个子规则方式中的其他至

少一个子规则方式的输出，或者，第三输入包括至少一个子模型的输出；第一子规则方式是多个子规则方式中的任意一个。

[0238] 示例性的，图11a为本申请实施例的子规则方式串联的示意图，如图11a所示，以两个子规则方式为例，该两个子规则方式中，子规则方式2的输出可以作为子规则方式1的一个输入，这样子规则方式1包括输入1和子规则方式2的输出。除此之外，子模型的输出也可以作为子规则方式1的一个输入，这样子规则方式1包括输入1和子模型的输出，甚至可以再加上子规则方式2的输出，对此不做具体限定。这可以是子规则方式与子规则方式或子模型的串联。

[0239] 一些可能的实施方式中，第一子规则方式的输入还包括第四输入，第四输入是第三输入经冗余信息处理后得到的音频信号；当预先设置的规则方式包括多个子规则方式时，第三输入包括多个子规则方式中的其他至少一个子规则方式的输出，或者，第三输入包括至少一个子模型的输出；第一子规则方式是多个子规则方式中的任意一个。

[0240] 示例性的，图11b为本申请实施例的子规则方式串联的示意图，如图11b所示，以两个子规则方式为例，该两个子规则方式中，子规则方式2的输出先经过冗余信息处理，然后可以作为子规则方式1的一个输入，这样子规则方式1包括输入1和子规则方式2的经冗余信息处理后输出。除此之外，子模型的输出也可以先经过冗余信息处理，然后作为子规则方式1的一个输入，这样子规则方式1包括输入1和子模型的经冗余信息处理后输出，甚至可以再加上子规则方式2的经冗余信息处理后输出，对此不做具体限定。这可以是子规则方式与子规则方式或子模型的串联。

[0241] 需要说明的是，上述示例以两个子规则方式为例进行说明，但这并非是对子规则方式之间、子规则方式与子模型之间的串联关系构成限定，任意一个子规则方式的输入除了自己的输入外，还可以包括其他一个或多个子规则方式的输出或经冗余处理后的输出，和/或，一个或多个子模型的输出或经冗余处理后的输出，本申请实施例不限定串联的个数，串联方式。

[0242] 一些可能的实施方式中，可以对获取边信息的第二机器学习模型和/或规则方式进行联合训练。即，通过联合训练多个获取边信息的机制(该机制可以是上述

子模型或者子规则方式)来降低或消除特征组之间的信息冗余；为多个获取边信息的机制设置联合损失函数，通过误差反向传播来修正多个获取边信息的机制的权重，来达到多个特征信息的整体最优的效果，进而降低或消除特征组之间的信息冗余；每个获取边信息的机制可以具有各自的损失函数，用于保留其提取到的特征信息的独特性。

[0243] 示例性的，图12为本申请实施例的联合训练的示意图，如图12所示，以两个获取边信息的机制(该机制可以是上述子模型或者子规则方式)为例，每个获取边信息的机制可以具有各自的损失函数(损失函数1和损失函数2)，为该两个获取边信息的机制设置联合损失函数。这些损失函数通过误差反向传播来修正两个获取边信息的机制的权重，来达到两个特征信息的整体最优的效果，一方面可以保留其提取到的特征信息的独特性，另一面还可以降低或消除两特征组之间的信息冗余。

[0244] 步骤703、根据边信息获取第二码流。

[0245] 本申请实施例中，可以通过边信息编码器对边信息进行编码，该边信息编码器可以与上文核心编码器采用相同的编码算法，也可以采用不同的编码算法，对此不做具体限定。

[0246] 步骤704、传输码流，该码流包括第一码流和第二码流。

[0247] 编码端得到的码流包括两部分，一部分码流(第一码流)对应当前帧的低频成分，另一部分码流(第二码流)对应当前帧的边信息。编码端将码流传输至解码端，以便于解码端基于码流得到重建的低频成分和边信息，从而得到与当前帧对应的重建帧。

[0248] 本实施例，通过机器学习模型从音频中提取边信息，以指导解码端实现非盲BWE，重建出高频成分，尤其是与低频成分关联性较弱的高频成分，而这种实现方式可以将当前帧中高频成分的更多细节体现于特征信息中，这些细节例如是高频成分的特性，高频成分与低频成分之间的关联等，有助于重建出更多的高频细节，从而可以避免音频失真的问题。

[0249] 图13为本申请实施例提供的音频解码方法的过程1300的流程图。过程1300描述为一系列的步骤或操作，应当理解的是，过程1300可以以各种顺序执行和/或同时发生，不限于图13所示的执行顺序。过程1300可以包括：

[0250] 步骤1301、接收码流，该码流包括第一码流和第二码流。

[0251] 本申请实施例中，接收来自编码端的码流。与编码端生成的码流相对应，接收到的码流也包括第一码流和第二码流两部分。

[0252] 步骤1302、根据第一码流获取第一音频信号，该第一音频信号的频率低于预设阈值。

[0253] 本申请实施例中，可以通过核心解码器对第一码流进行解析以得到第一音频信号，

[0254] 解码端要得到的重建帧与编码端的当前帧相对应，该重建帧包含的音频信号的特征也可以通过频率、幅度和相位来表征，根据频率划分，重建帧也包括第一音频信号和第二音频信号，其中，第一音频信号的频率低于预设阈值的音频信号(本申请实施例中，满足该条件的音频信号亦可以称作低频信号或低频成分)，第二音频信号的频率高于或等于预设阈值的音频信号(本申请实施例中，满足该条件的音频信号亦可以称作高频信号或高频成分)，即重建帧包括低频成分和高频成分。前述预设阈值例如可以是16kHz，本申请实施例对此不做具体限定。需要说明的是，前文将频率等于预设阈值的音频信号归入第二音频信号，可选的，本申请实施例也可以将频率等于预设阈值的音频信号归入第一音频信号，对此不做具体限定。

[0255] 基于此，根据第一码流得到的第一音频信号可以是重建帧的低频成分，亦即，解码端可以通过解析码流直接得到重建帧的低频成分。

[0256] 步骤1303、根据第二码流获取边信息。

[0257] 边信息用于表征第二音频信号的特性，此外，边信息还可以用于表征高低频之间的映射关系。本申请实施例中，边信息仅占用少量的传输比特，就可以指导解码端实现非盲BWE，以重建出高频成分，尤其是与低频成分关系紧密但映射关系复杂的高频成分。

[0258] 一些可能的实施方式中，根据第二码流获取边信息可以直接解析第二码流以得到边信息。即，第二码流是由编码端处理边信息得到的，因此解码端可以直接解析第二码流得到边信息。

[0259] 可选的，解析第二码流先得到处理前边信息；然后对处理前边信息进行边信息处理以得到边信息，该边信息处理包括输入预先训练得到的第三机器学习模型或者按照预先设置的第三规则方式进行处理。

[0260] 本申请实施例中，第二码流携带的信息可以是与边信息关联的处理前边信息，该信息可以是编码端考虑到进一步节省码流，而基于边信息得到的，进而将其编码入码流。相应的，解码端为了得到边信息可以先解析第二码流得到的处理前边信息，然后基于处理前边信息获取边信息。前述获取边信息的方法可以包括通过第三机器学习模型或者按照预先设置的第三规则方式获取边信息。

[0261] 步骤1304、根据第一音频信号和边信息获取重建帧。

[0262] 一些可能的实施方式中，获取重建帧可以根据第一音频信号和边信息通过预先训练得到的第一机器学习模型，获取重建帧。由于获取重建帧参考了边信息，以实现从低频成分向高频成分的扩展，因此获取重建帧也可以称作非盲BWE。

[0263] 本申请实施例中，第一机器学习模型可以参照上文的神经网络、深度神经网络、卷积神经网络、循环神经网络等。

[0264] 示例性的，图14a为本申请实施例的非盲BWE的示意图，如图14a所示，非盲BWE通过第一机器学习模型实现，该第一机器学习模型包括一个子模型，即第一机器学习模型亦即该子模型，其输入为第一音频信号和边信息，输出为重建帧。

[0265] 可选的，重建帧是第一机器学习模型直接输出的。

[0266] 本申请实施例中，非盲BWE可以采用一个机器学习模型(第一机器学习模型)，在边信息的指导下，根据解码出的低频成分(第一音频信号)，直接输出重建帧。此外，第一机器学习模型可以完成低频成分的编解码失真修正，以及可能的音频渲染，本申请实施例对此不做具体限定。

[0267] 示例性的，图15a为本申请实施例的获取重建帧的示意图，如图15a所示，非盲BWE的输入包括低频成分和边信息，输出为重建帧。

[0268] 可选的，重建帧是基于第一机器学习模型输出的第二音频信号与第一音频信号得到的。

[0269] 本申请实施例中，非盲BWE可以采用一个机器学习模型(第一机器学习模型)，在边信息的指导下，根据解码出的低频成分(第一音频信号)，输出重建的高频成分，将重建的高频成分与解码出的低频成分合并或相加，得到重建音频。此外，第一机器学习模型可以完成低频成分的编解码失真修正，以及可能的音频渲染，本申请实施例对此不做具体限定。

[0270] 示例性的，图15b为本申请实施例的获取重建帧的示意图，如图15b所示，非盲BWE的输入包括低频成分和边信息，输出为重建的高频成分，再将高频成分与低频成分进行高低频合并，得到重建帧。

[0271] 一些可能的实施方式中，第一机器学习模型可以包括第一子模型和第二子模型，非盲BWE可以将第一音频信号和第一边信息输入第一子模型，以获取第三音频信号；再将第三音频信号和第二边信息输入第二子模型，以获取重建帧；第一边信息和第二边信息分别是边信息中的部分或全部信息。

[0272] 示例性的，图14b为本申请实施例的非盲BWE的示意图，如图14b所示，非盲BWE采用的第一机器学习模型可以分成第一子模型(对应低频特征提取)和第二子模型(对应高频成分重建)，这两个子模型串联，即低频特征提取的输出(第三音频信号，亦即低频特征)可以作为高频成分重建的输入；低频特征提取的输入包括第一音频信号和第一边信息，高频成分重建的输入包括第三音频信号和第二边信息；低频特征提取由第一边信息指导，高频成分重建由第二边信息指导；高频成分重建的输出为重建帧。需要说明的是，本申请实施例中，第一边信息和第二边信息可以分别是边信息中的部分或全部信息，第一边信息和第二边信息可以相同或不同，对此不做具体限定。

[0273] 一些可能的实施方式中，非盲BWE可以根据第一音频信号和第三边信息按照预先设置的第一规则方式进行处理以获取第四音频信号；再将第四音频信号和第四边信息输入第一机器学习模型，以获取重建帧；第三边信息和第四边信息分别是边信息中的部分或全部信息。

[0274] 示例性的，图14c为本申请实施例的非盲BWE的示意图，如图14c所示，与编码端相对应，当编码端按照规则方式获取边信息时，解码端也可以相应的在非盲BWE中采用规则方式。该规则方式可以设置于机器学习之前，即先按照第一规则方式处理(对应基于规则前处理)，其输入为第一音频信号和第三边信息，输出为第四音频信号；再通过第一机器学习模型，其输入为第四音频信号和第四边信息，输出为重建帧。需要说明的是，本申请实施例中，第三边信息和第四边信息可以分别是边信息中的部分或全部信息，第三边信息和第四边信息可以相同或不同，对此不做具体限定。

[0275] 一些可能的实施方式中，非盲BWE可以根据第一音频信号和第五边信息通过预先训练得到的第一机器学习模型，获取第五音频信号；再根据第五音频信号和第六边信息按照预先设置的第二规则方式进行处理以获取重建帧；第五边信息和第六边信息分别是边信息中的部分或全部信息。

[0276] 示例性的，图14d为本申请实施例的非盲BWE的示意图，如图14d所示，与编码端相对应，当编码端按照规则方式获取边信息时，解码端也可以相应的在非盲BWE中采用规则方式。该规则方式可以设置于机器学习之后，即先通过第一机器学习模型，其输入为第一音频信号和第五边信息，输出为第五音频信号；再按照第二规则方式处理，其输入为第五音频信号和第六边信息，输出为重建帧。需要说明的是，本申请实施例中，第五边信息和第六边信息可以分别是边信息中的部分或全部信息，第五边信息和第六边信息可以相同或不同，对此不做具体限定。

[0277] 需要说明的是，图14a−图14d所示实施例中，均以非盲BWE直接输出重建帧(参照图15a)为例，应理解，本申请实施例中，还可以采用图15b所示实施例的方法得到重建帧。此外，图14c和图14d所示实施例中，非盲BWE采用的第一机器学习模型还可以参照图14b所示实施例。

[0278] 示例性的，图14e为本申请实施例的非盲BWE的示意图，如图14e所示，非盲BWE先进行基于规则前处理，其输入为低频成分和边信息1；然后采用的第一机器学习模型，其输入为基于规则前处理的输出和边信息2；再进行基于规则后处

理，其输入为第一机器学习模型的输出和边信息3，输出为高频成分；最后将低频成分和高频成分进行高低频合并，输出重建帧。

[0279] 需要说明的是，上述边信息1−边信息3可以分别是边信息中的部分或全部信息，边信息1−边信息3可以完全相同或完全不同或部分相同、部分不同，对此不做具体限定。

[0280] 示例性的，图14f为本申请实施例的非盲BWE的示意图，如图14f所示，非盲BWE先进行基于规则前处理，其输入为低频成分和边信息1；然后采用的第一机器学习模型中的低频特征提取，其输入为基于规则前处理的输出和边信息2，输出为低频特征；然后采用的第一机器学习模型中的高频成分重建，其输入为低频特征和边信息3，输出为高频特征；再进行基于规则后处理，其输入为高频特征和边信息4，输出为高频成分；最后将低频成分和高频成分进行高低频合并，输出重建帧。

[0281] 需要说明的是，上述边信息1−边信息4可以分别是边信息中的部分或全部信息，边信息1−边信息4可以完全相同或完全不同或部分相同、部分不同，对此不做具体限定。

[0282] 本实施例，通过机器学习模型在边信息的指导下，基于低频成分得到重建帧，尤其是高频成分与低频成分关联性较弱或者高低频成分关系紧密但映射关系复杂的场景下的重建帧获取，而这种实现方式可以将高频成分的更多细节体现于特征信息中，这些细节例如是高频成分的特性，高频成分与低频成分之间的关联等，有助于重建出更多的高频细节，从而可以避免音频失真的问题。

[0283] 下面采用几个具体的实施例，对图7和图13所示方法实施例的技术方案进行详细说明。

[0284] 图16为本申请实施例的包含AI的非盲BWE的编解码框架的示意图，如图16所示，该框架包括以下步骤：

[0285] 1、编码端将原始音频转换为码流。

[0286] 1.1、将原始音频的低频成分输入核心编码器，核心编码器将低频成分编码为低频码流(对应上文第一码流)。参照上文步骤701。

[0287] 1.2、边信息提取环节采用机器学习模型，机器学习模型从包含原始音频的高频成分的音频中提取出有助于BWE的特征，称作边信息。参照上文步骤702。

[0288] 1.3、边信息输入边信息编码器，边信息编码器将边信息编码为边信息码流(对应上文第二码流)。参照上文步骤703。

[0289] 1.4、将边信息码流和低频码流合并得到码流，该码流经过传输网络传输到解码端。

[0290] 2、解码端根据码流重建音频。

[0291] 2.1、接收到码流后，将码流解析成低频码流和边信息码流。参照上文步骤1301。

[0292] 2.2、将低频码流输入核心解码器，核心解码器将低频码流解码为低频成分。参照上文步骤1302。

[0293] 2.3、将边信息码流输入边信息解码器，边信息解码器将边信息码流解码为边信息。参照上文步骤1303。

[0294] 2.4、非盲BWE采用机器学习模型，机器学习模型在边信息的指导下，根据解码出的低频成分，输出重建音频。此外机器学习模型还可以完成低频成分的编解码失真修正，以及可能的音频渲染。

[0295] 可选的，步骤2.4还可以参照图15b的方式先得到重建的高频成分，再获取重建音频。

[0296] 图17为本申请实施例的包含AI的非盲BWE的编解码框架的示意图，如图17所示，本实施例在图16所示实施例的基础上，编码端的边信息提取环节，除了采用机器学习模型外，还采用了规则方式，这样边信息提取环节输出的边信息包括机器学习模型输出的特征信息和按照规则方式输出的特征信息。解码端的非盲BWE环节，除了采用机器学习模型外，还采用了规则方式，这样非盲BWE环节输出的重建音频结合了机器学习模型的输出和按照规则方式的输出。

[0297] 本实施例的边信息提取环节的几种实现方式可以参照图8b-图8d、图9a-图9c、图10a-图10b、图11a-图11b所示实施例。

[0298] 图18为本申请实施例的包含AI的非盲BWE的编解码框架的示意图，如图18所示，本实施例建立在AVS3编解码框架上，实现了48kHz的音频编解码。其中，

核心编解码器通过编解码方式负责16kHz以下的低频成分的编码和重建，边信息提取环节负责提取边信息，非盲BWE环节负责重建从16kHz到48kHz的高频成分。该框架包括以下步骤：

[0299] 1、编码端将48kHz的原始音频转换为比特流。

[0300] 1.1、对48kHz的原始音频做分帧处理，每帧帧长为2048个采样点，帧间叠接1024个采样点，依次处理这些信号帧。

[0301] 1.2、低频通道编码信号帧中低于16kHz的低频成分。

[0302] 1.2.1、采用低通滤波器从信号帧中提取出低于16kHz的音频信号，即为低频成分。

[0303] 1.2.2、通过AVS3核心编码器将低于16kHz的低频成分编码为低频码流。

[0304] 1.3、边信息通道编码信号帧中从16kHz到48kHz的高频成分。

[0305] 1.3.1、边信息提取环节从信号帧中提取边信息，如图19(图19为本申请实施例的边信息提取环节的流程图)所示：

[0306] 1.3.1.1、通过傅里叶变换将信号帧变换为频谱。

[0307] 1.3.1.2、将频谱的高频区域划分为多个频带，进行包络计算，即计算每个频带的频谱能量的平均值，作为频带包络(特征组1)。

[0308] 1.3.1.3、将高频区域的每个频带的频谱除以该频带对应的频带包络，将这些频带的频谱能量变为1，消除其中包含的能量包络信息。经频带归一化后的频谱被称作归一化频谱。

[0309] 1.3.1.4、将归一化频谱输入到神经网络，该神经网络从归一化频谱中提取特征(特征组2)。

[0310] 1.3.1.5、将特征组1和特征组2合并，作为边信息。

[0311] 1.3.2、量化编码器将边信息量化为边信息码流。

[0312] 1.4、将低频码流和边信息码流合并为码流，并传输到解码端。

[0313] 2、解码端根据码流重建出48kHz的原始音频。

[0314] 2.1、接收编码端传输过来的码流，并将其拆分为低频码流和边信息码流。

[0315] 2.2、利用AVS3核心解码器从低频码流中解码出低于16kHz的低频成分。

[0316] 2.3、通过量化解码器从边信息码流中解码出边信息。

[0317] 2.4、非盲BWE环节负责在解码出的边信息的指导下，根据解码出的低于16kHz的低频信号重建出高频信号，如图20(图20为本申请实施例的非盲BWE环节的流程图)所示：

[0318] 2.4.1、解码出的低频成分经傅里叶变换后变换为低频频谱。

[0319] 2.4.2、基于低频特征提取模型从低频频谱中提取低频特征。

[0320] 2.4.3、边信息处理模型将特征组2映射为高频特征。

[0321] 2.4.4、高频成分重建模型在高频特征的指导下，根据低频特征重建高频成分，得到重建频谱。

[0322] 2.4.5、根据特征组1调整对应频带的频谱能量，保证每个频带的频谱能量与真实相同，并将16kHz以下的频谱替换为低频频谱，得到调整后的频谱。

[0323] 2.4.6、对调整后的频谱进行傅里叶反变换，得到重建音频。

[0324] 2.5、将解码出的低频成分和重建出的高频成分相加，然后进行叠接处理，得到48kHz的重建音频。

[0325] 在相同的实验条件下，比较了基于规则的BWE、基于AI的全盲BWE和基于AI的非盲BWE(本申请实施例的技术方案)的结果以及真实标的，实验结果如图21和图22(图21和图22分别是对三种BWE方式得到的音频频谱的比较示意图)所示：

[0326] 基于规则的BWE将低频频谱拷贝到高频，并根据边信息调整各频带的能量。因此，该方法无法扩展出低频成分中不存在的纹理。如图21和图22中框出的区域的频谱是从低频拷贝上来的，与真实频谱有明显差异，进而造成了听感差异。

[0327] 基于AI的全盲BWE可以根据低频成分重建出一部分高频成分。由于频率越高的频谱与低频频谱的关联性越弱，模型重建出的频谱失真也越大。这也导致随着频率的升高，模型重建出的高频能量会逐渐降低。另外，一些高频成分式独立存在的。如果没有边信息的指导，模型无法判断这些成分是否存在，也无法重建出这些成分。

[0328] 相较于基于规则的BWE和基于AI的全盲BWE，基于AI的非盲BWE重建出的高频频谱明显更接近真实，与原始音频的听感也更加一致。相较于基于规则的BWE，由AI模型发掘出的高低频映射关系更加复杂和全面。因此由基于AI的边信息提取模块所提取出的边信息能够包含更丰富的频谱特征，基于AI的非盲BWE模

块也能够据此重建出更多更准确的高频细节。其中，基于规则的前处理模块可在边信息指导下，构建高频基底，简化后续AI计算。将边信息引入低频特征提取和高频成分重建模块，可分别达到提取出符合边信息的低频特征和重建出符合边信息的高频成分的效果。而基于规则的后处理根据边信息进一步调整重建出的高频成分，以确保高频成分满足先验特性。如可根据边信息调整各频带的频谱包络能量，以保证包络能量的准确性。

[0329] 图23为本申请实施例的包含AI的非盲BWE的编解码框架的示意图，如图23所示，本实施例建立在AVS3编解码框架上，实现了48kHz的音频编解码。其中，核心编解码器通过编解码方式负责指定频率f以下的低频成分的编码和重建，边信息提取环节负责提取边信息，非盲BWE环节负责重建从f到48kHz的高频成分。该框架包括以下步骤：

[0330] 1、编码端将48kHz原始音频转换为比特流。

[0331] 1.1、对48kHz的原始音频做分帧处理，每帧帧长为2048个采样点，帧间叠接1024个采样点，依次处理这些信号帧。

[0332] 1.2、指定频率边界f，低频通道编码信号帧中低于f的低频成分。

[0333] 1.2.1、采用低通滤波器从48kHz的原始音频中提取出低于f的低频信号，即为低频成分。

[0334] 1.2.2、通过AVS3核心编码器将低于f的低频成分编码为低频码流。

[0335] 1.3、边信息通道编码从f到48kHz的高频成分。

[0336] 1.3.1、边信息提取环节从48kHz的原始音频中提取边信息，如图24(图24为本申请实施例的边信息提取环节的流程图)所示。

[0337] 1.3.1.1、通过傅里叶变换将信号帧变换为频谱。

[0338] 1.3.1.2、采用AVS3自带的边信息提取模块，将频谱的高频区域划分为多个频带，进行包络计算，即计算每个频带的频谱能量的平均值，作为频带包络，并计算每个频带采用的频带复制方式标识，将两种信息合并为特征组1。

[0339] 1.3.1.3、将高频区域的每个频带的频谱除以该频带对应的频带包络，将这些频带的频谱能量变为1，消除其中包含的能量包络信息。经频带归一化后的频谱被称作归一化频谱。

[0340] 1.3.1.4、将归一化频谱输入到神经网络，该神经网络从归一化频谱中提取特征(特征组2)。

[0341] 1.3.1.5、将特征组1和特征组2合并，作为边信息。

[0342] 1.3.2、量化编码器将边信息量化为边信息码流。

[0343] 1.4、将低频码流、边信息码流和f合并为码流，并传输到解码端。

[0344] 2、解码端根据码流重建出48kHz的原始音频。

[0345] 2.1、接收编码端传输过来的码流，并将其拆分为低频码流和边信息码流。

[0346] 2.2、利用AVS3核心解码器从低频码流中解码出低于f的低频信号。

[0347] 2.3、通过量化解码器从边信息码流中解码出边信息。

[0348] 2.4、非盲BWE环节负责在解码出的边信息的指导下，根据解码出的低于f的低频信号重建出高频信号，如图25(图25为本申请实施例的非盲BWE环节的流程图)所示。

[0349] 2.4.1、将解码出的低频成分和特征组1输入到AVS3频带扩展，利用基于规则的BWE方法扩展出高频成分，将高频成分与低频成分相加得到信号基底。

[0350] 2.4.2、边信息处理模型将特征组2映射为高频特征。

[0351] 2.4.3、在高频特征的指导下，非盲BWE模型根据信号基底，得到重建信号。

[0352] 2.4.4、对重建信号进行傅里叶变换，得到重建频谱。根据特征组1的频带包络调整对应频带的频谱能量，保证每个频带的频谱能量与真实相同。再对调整后的频谱进行傅里叶反变换，得到调整后的重建信号。

[0353] 2.5、对调整后的重建信号进行叠接处理，得到48kHz的重建音频。

[0354] 图26为本申请实施例的编码装置2600的结构示意图，如图26所示，本实施例的编码装置2600可以应用于上文编码设备。该编码装置2600可以包括：获取模块2601、提取模块2602、编码模块2603、传输模块2604和滤波模块2605。其中，

[0355] 获取模块2601，用于获取第一码流，所述第一码流是基于第一音频信号得到的，所述第一音频信号是当前帧中频率低于预设阈值的低频信号；提取模块2602，用于获取边信息，所述边信息是基于第二音频信号得到的，所述第二音频信号包括所述当前帧中频率高于或等于所述预设阈值的高频信号；编码模块2603，

用于根据所述边信息获取第二码流；传输模块2604，用于传输码流，所述码流包括所述第一码流和所述第二码流。

[0356] 在一种可能的实现方式中，所述提取模块2602，具体用于将所述第二音频信号输入预先训练得到的第二机器学习模型，以获取所述边信息，所述第二机器学习模型包括至少一个子模型。

[0357] 在一种可能的实现方式中，当所述第二机器学习模型包括多个子模型时，所述多个子模型各自的输入包括所述第二音频信号中的部分音频信号或全部音频信号，其中，所述部分音频信号包括所述高频信号。

[0358] 在一种可能的实现方式中，第一子模型与所述多个子模型中的其他至少一个子模型包括相同的神经元，所述相同的神经元位于所述第一子模型的输入阶段、输出阶段或者中间阶段，所述第一子模型是所述多个子模型中的任意一个。

[0359] 在一种可能的实现方式中，所述边信息包括所述多个子模型分别输出的特征信息。

[0360] 在一种可能的实现方式中，当所述第二机器学习模型包括一个子模型时，所述边信息包括所述一个子模型输出的特征信息。

[0361] 在一种可能的实现方式中，所述提取模块2602，具体用于根据所述第二音频信号按照预先设置的规则方式进行处理以获取所述边信息，所述预先设置的规则方式包括至少一个子规则方式。

[0362] 在一种可能的实现方式中，当所述预先设置的规则方式包括多个子规则方式时，所述多个子规则方式各自的输入包括所述第二音频信号中的部分音频信号或全部音频信号，其中，所述部分音频信号包括所述高频信号。

[0363] 在一种可能的实现方式中，所述边信息包括按照所述多个子规则方式分别输出的特征信息。

[0364] 在一种可能的实现方式中，当所述预先设置的规则方式包括一个子规则方式时，所述边信息包括按照所述一个子规则方式输出的特征信息。

[0365] 在一种可能的实现方式中，第一子模型的输入还包括第一输入；当所述第二机器学习模型包括多个子模型时，所述第一输入包括所述多个子模型中的其他至

少一个子模型的输出，或者，所述第一输入包括所述至少一个子规则方式的输出；所述第一子模型是所述多个子模型中的任意一个。

[0366] 在一种可能的实现方式中，第一子模型的输入还包括第二输入，所述第二输入是第一输入经冗余信息处理后得到的音频信号；当所述第二机器学习模型包括多个子模型时，所述第一输入包括所述多个子模型中的其他至少一个子模型的输出，或者，所述第一输入包括所述至少一个子规则方式的输出；所述第一子模型是所述多个子模型中的任意一个。

[0367] 在一种可能的实现方式中，第一子规则方式的输入还包括第三输入；当所述预先设置的规则方式包括多个子规则方式时，所述第三输入包括所述多个子规则方式中的其他至少一个子规则方式的输出，或者，所述第三输入包括所述至少一个子模型的输出；所述第一子规则方式是所述多个子规则方式中的任意一个。

[0368] 在一种可能的实现方式中，第一子规则方式的输入还包括第四输入，所述第四输入是第三输入经冗余信息处理后得到的音频信号；当所述预先设置的规则方式包括多个子规则方式时，所述第三输入包括所述多个子规则方式中的其他至少一个子规则方式的输出，或者，所述第三输入包括所述至少一个子模型的输出；所述第一子规则方式是所述多个子规则方式中的任意一个。

[0369] 在一种可能的实现方式中，滤波模块2605，用于通过低通滤波的方式从所述当前帧中获取所述第一音频信号。

[0370] 在一种可能的实现方式中，所述第二音频信号还包括所述第一音频信号中的部分或全部音频信号。

[0371] 本实施例的装置，可以用于执行图7所示方法实施例的技术方案，其实现原理和技术效果类似，此处不再赘述。

[0372] 图27为本申请实施例的解码装置2700的结构示意图，如图27所示，本实施例的解码装置2700可以应用于上文解码设备。该解码装置2700可以包括：接收模块2701、获取模块2702、提取模块2703和重建模块2704。其中，

[0373] 接收模块2701，用于接收码流，所述码流包括第一码流和第二码流；获取模块2702，用于根据所述第一码流获取第一音频信号，所述第一音频信号的频率低于

预设阈值；提取模块2703，用于根据所述第二码流获取边信息，所述边信息用于表征第二音频信号的特性，所述第二音频信号的频率高于或等于所述预设阈值；重建模块2704，用于根据所述第一音频信号和所述边信息获取重建帧。

[0374] 在一种可能的实现方式中，所述重建模块2704，具体用于根据所述第一音频信号和所述边信息通过预先训练得到的第一机器学习模型，获取所述重建帧。

[0375] 在一种可能的实现方式中，所述重建帧是所述第一机器学习模型直接输出的；或者，所述重建帧是基于所述第一机器学习模型输出的所述第二音频信号与所述第一音频信号得到的。

[0376] 在一种可能的实现方式中，当所述第一机器学习模型包括第一子模型和第二子模型时，所述重建模块2704，具体用于将所述第一音频信号和第一边信息输入所述第一子模型，以获取第三音频信号；将所述第三音频信号和第二边信息输入所述第二子模型，以获取所述重建帧；所述第一边信息和所述第二边信息分别是所述边信息中的部分或全部信息。

[0377] 在一种可能的实现方式中，所述重建模块2704，具体用于根据所述第一音频信号和第三边信息按照预先设置的第一规则方式进行处理以获取第四音频信号；将所述第四音频信号和第四边信息输入所述第一机器学习模型，以获取所述重建帧；所述第三边信息和所述第四边信息分别是所述边信息中的部分或全部信息。

[0378] 在一种可能的实现方式中，所述重建模块2704，具体用于根据所述第一音频信号和所述五边信息通过预先训练得到的第一机器学习模型，获取第五音频信号；根据所述第五音频信号和第六边信息按照预先设置的第二规则方式进行处理以获取所述重建帧；所述第五边信息和所述第六边信息分别是所述边信息中的部分或全部信息。

[0379] 在一种可能的实现方式中，所述提取模块2703，具体用于解析所述第二码流以得到所述边信息。

[0380] 在一种可能的实现方式中，所述提取模块2703，具体用于解析所述第二码流以得到处理前边信息；对所述处理前边信息进行边信息处理以得到所述边信息，

所述边信息处理包括输入预先训练得到的第三机器学习模型或者按照预先设置的第三规则方式进行处理。

[0381] 本实施例的装置，可以用于执行图13所示方法实施例的技术方案，其实现原理和技术效果类似，此处不再赘述。

[0382] 在实现过程中，上述方法实施例的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。处理器可以是通用处理器、数字信号处理器(digital signal processor,DSP)、特定应用集成电路(application-specific integrated circuit，ASIC)、现场可编程门阵列(field programmable gate array,FPGA)或其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。本申请实施例公开的方法的步骤可以直接体现为硬件编码处理器执行完成，或者用编码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器，处理器读取存储器中的信息，结合其硬件完成上述方法的步骤。

[0383] 上述各实施例中提及的存储器可以是易失性存储器或非易失性存储器，或可包括易失性和非易失性存储器两者。其中，非易失性存储器可以是只读存储器(read-only memory，ROM)、可编程只读存储器(programmable ROM，PROM)、可擦除可编程只读存储器(erasable PROM，EPROM)、电可擦除可编程只读存储器(electrically EPROM，EEPROM)或闪存。易失性存储器可以是随机存取存储器(random access memory，RAM)，其用作外部高速缓存。通过示例性但不是限制性说明，许多形式的RAM可用，例如静态随机存取存储器(static RAM，SRAM)、动态随机存取存储器(dynamic RAM，DRAM)、同步动态随机存取存储器(synchronous DRAM，SDRAM)、双倍数据速率同步动态随机存取存储器(double data rate SDRAM，DDR SDRAM)、增强型同步动态随机存取存储器(enhanced SDRAM，ESDRAM)、同步连接动态随机存取存储器(synchlink DRAM，SLDRAM)和直接内存总线随机存取存储器(direct rambus RAM，DR RAM)。应注意，本文

描述的系统和方法的存储器旨在包括但不限于这些和任意其它适合类型的存储器。

[0384] 本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

[0385] 所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统、装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

[0386] 在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

[0387] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0388] 另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

[0389] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备(个人计算机，服务器，或者网络设备等)执行本申请各

个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U盘、移动硬盘、只读存储器(read-only memory，ROM)、随机存取存储器(random access memory，RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0390] 以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以所述权利要求的保护范围为准。

权利要求书

[权利要求 1] 一种音频编码方法，其特征在于，包括：

获取第一码流，所述第一码流是基于第一音频信号得到的，所述第一音频信号是当前帧中频率低于预设阈值的低频信号；

获取边信息，所述边信息是基于第二音频信号得到的，所述第二音频信号包括所述当前帧中频率高于或等于所述预设阈值的高频信号；

根据所述边信息获取第二码流；

传输码流，所述码流包括所述第一码流和所述第二码流。

[权利要求 2] 根据权利要求1所述的方法，其特征在于，所述获取边信息，包括：

将所述第二音频信号输入预先训练得到的第二机器学习模型，以获取所述边信息，所述第二机器学习模型包括至少一个子模型。

[权利要求 3] 根据权利要求2所述的方法，其特征在于，当所述第二机器学习模型包括多个子模型时，所述多个子模型各自的输入包括所述第二音频信号中的部分音频信号或全部音频信号，其中，所述部分音频信号包括所述高频信号。

[权利要求 4] 根据权利要求3所述的方法，其特征在于，第一子模型与所述多个子模型中的其他至少一个子模型包括相同的神经元，所述相同的神经元位于所述第一子模型的输入阶段、输出阶段或者中间阶段，所述第一子模型是所述多个子模型中的任意一个。

[权利要求 5] 根据权利要求2-4中任一项所述的方法，其特征在于，所述获取边信息，包括：

根据所述第二音频信号按照预先设置的规则方式进行处理以获取所述边信息，所述预先设置的规则方式包括至少一个子规则方式。

[权利要求 6] 根据权利要求5所述的方法，其特征在于，当所述预先设置的规则方式包括多个子规则方式时，所述多个子规则方式各自的输入包括

所述第二音频信号中的部分音频信号或全部音频信号，其中，所述部分音频信号包括所述高频信号。

[权利要求 7] 根据权利要求5或6所述的方法，其特征在于，第一子模型的输入还包括第一输入；当所述第二机器学习模型包括多个子模型时，所述第一输入包括所述多个子模型中的其他至少一个子模型的输出，或者，所述第一输入包括所述至少一个子规则方式的输出；所述第一子模型是所述多个子模型中的任意一个。

[权利要求 8] 根据权利要求5或6所述的方法，其特征在于，第一子模型的输入还包括第二输入，所述第二输入是第一输入经冗余信息处理后得到的音频信号；当所述第二机器学习模型包括多个子模型时，所述第一输入包括所述多个子模型中的其他至少一个子模型的输出，或者，所述第一输入包括所述至少一个子规则方式的输出；所述第一子模型是所述多个子模型中的任意一个。

[权利要求 9] 根据权利要求5-8中任一项所述的方法，其特征在于，第一子规则方式的输入还包括第三输入；当所述预先设置的规则方式包括多个子规则方式时，所述第三输入包括所述多个子规则方式中的其他至少一个子规则方式的输出，或者，所述第三输入包括所述至少一个子模型的输出；所述第一子规则方式是所述多个子规则方式中的任意一个。

[权利要求 10] 根据权利要求5-8中任一项所述的方法，其特征在于，第一子规则方式的输入还包括第四输入，所述第四输入是第三输入经冗余信息处理后得到的音频信号；当所述预先设置的规则方式包括多个子规则方式时，所述第三输入包括所述多个子规则方式中的其他至少一个子规则方式的输出，或者，所述第三输入包括所述至少一个子模型的输出；所述第一子规则方式是所述多个子规则方式中的任意一个。

[权利要求 11] 一种音频解码方法，其特征在于，包括：

接收码流，所述码流包括第一码流和第二码流；

根据所述第一码流获取第一音频信号，所述第一音频信号的频率低于预设阈值；

根据所述第二码流获取边信息，所述边信息用于表征第二音频信号的特性，所述第二音频信号的频率高于或等于所述预设阈值；

根据所述第一音频信号和所述边信息获取重建帧。

[权利要求 12] 根据权利要求11所述的方法，其特征在于，所述根据所述第一音频信号和所述边信息获取重建帧，包括：

根据所述第一音频信号和所述边信息通过预先训练得到的第一机器学习模型，获取所述重建帧。

[权利要求 13] 根据权利要求12所述的方法，其特征在于，所述重建帧是所述第一机器学习模型直接输出的；或者，所述重建帧是基于所述第一机器学习模型输出的所述第二音频信号与所述第一音频信号得到的。

[权利要求 14] 根据权利要求12或13所述的方法，其特征在于，当所述第一机器学习模型包括第一子模型和第二子模型时，所述根据所述第一音频信号和所述边信息通过预先训练得到的第一机器学习模型，获取所述重建帧，包括：

将所述第一音频信号和第一边信息输入所述第一子模型，以获取第三音频信号；

将所述第三音频信号和第二边信息输入所述第二子模型，以获取所述重建帧；所述第一边信息和所述第二边信息分别是所述边信息中的部分或全部信息。

[权利要求 15] 根据权利要求12−14中任一项所述的方法，其特征在于，所述根据所述第一音频信号和所述边信息通过预先训练得到的第一机器学习模型，获取所述重建帧，包括：

根据所述第一音频信号和第三边信息按照预先设置的第一规则方式进行处理以获取第四音频信号；

将所述第四音频信号和第四边信息输入所述第一机器学习模型，以获取所述重建帧；所述第三边信息和所述第四边信息分别是所述边信息中的部分或全部信息。

[权利要求 16] 根据权利要求12-15中任一项所述的方法，其特征在于，所述根据所述第一音频信号和所述边信息通过预先训练得到的第一机器学习模型，获取所述重建帧，包括：

根据所述第一音频信号和所述五边信息通过预先训练得到的第一机器学习模型，获取第五音频信号；

根据所述第五音频信号和第六边信息按照预先设置的第二规则方式进行处理以获取所述重建帧；所述第五边信息和所述第六边信息分别是所述边信息中的部分或全部信息。

[权利要求 17] 根据权利要求11-16中任一项所述的方法，其特征在于，所述根据所述第二码流获取边信息，包括：

解析所述第二码流以得到所述边信息。

[权利要求 18] 根据权利要求11-16中任一项所述的方法，其特征在于，所述根据所述第二码流获取边信息，包括：

解析所述第二码流以得到处理前边信息；

对所述处理前边信息进行边信息处理以得到所述边信息，所述边信息处理包括输入预先训练得到的第三机器学习模型或者按照预先设置的第三规则方式进行处理。

[权利要求 19] 一种编码装置，其特征在于，包括：

获取模块，用于获取第一码流，所述第一码流是基于第一音频信号得到的，所述第一音频信号是当前帧中频率低于预设阈值的低频信号；

提取模块，用于获取边信息，所述边信息是基于第二音频信号得到的，所述第二音频信号包括所述当前帧中频率高于或等于所述预设阈值的高频信号；

编码模块，用于根据所述边信息获取第二码流；

传输模块，用于传输码流，所述码流包括所述第一码流和所述第二码流。

[权利要求 20] 根据权利要求19所述的装置，其特征在于，所述提取模块，具体用于将所述第二音频信号输入预先训练得到的第二机器学习模型，以获取所述边信息，所述第二机器学习模型包括至少一个子模型。

[权利要求 21] 根据权利要求20所述的装置，其特征在于，当所述第二机器学习模型包括多个子模型时，所述多个子模型各自的输入包括所述第二音频信号中的部分音频信号或全部音频信号，其中，所述部分音频信号包括所述高频信号。

[权利要求 22] 根据权利要求21所述的装置，其特征在于，第一子模型与所述多个子模型中的其他至少一个子模型包括相同的神经元，所述相同的神经元位于所述第一子模型的输入阶段、输出阶段或者中间阶段，所述第一子模型是所述多个子模型中的任意一个。

[权利要求 23] 根据权利要求20-22中任一项所述的装置，其特征在于，所述提取模块，具体用于根据所述第二音频信号按照预先设置的规则方式进行处理以获取所述边信息，所述预先设置的规则方式包括至少一个子规则方式。

[权利要求 24] 根据权利要求23所述的装置，其特征在于，当所述预先设置的规则方式包括多个子规则方式时，所述多个子规则方式各自的输入包括所述第二音频信号中的部分音频信号或全部音频信号，其中，所述部分音频信号包括所述高频信号。

[权利要求 25] 根据权利要求23或24所述的装置，其特征在于，第一子模型的输入还包括第一输入；当所述第二机器学习模型包括多个子模型时，所述第一输入包括所述多个子模型中的其他至少一个子模型的输出，或者，所述第一输入包括所述至少一个子规则方式的输出；所述第一子模型是所述多个子模型中的任意一个。

[权利要求 26] 根据权利要求23或24所述的装置，其特征在于，第一子模型的输入还包括第二输入，所述第二输入是第一输入经冗余信息处理后得

到的音频信号；当所述第二机器学习模型包括多个子模型时，所述第一输入包括所述多个子模型中的其他至少一个子模型的输出，或者，所述第一输入包括所述至少一个子规则方式的输出；所述第一子模型是所述多个子模型中的任意一个。

[权利要求 27] 根据权利要求23-26中任一项所述的装置，其特征在于，第一子规则方式的输入还包括第三输入；当所述预先设置的规则方式包括多个子规则方式时，所述第三输入包括所述多个子规则方式中的其他至少一个子规则方式的输出，或者，所述第三输入包括所述至少一个子模型的输出；所述第一子规则方式是所述多个子规则方式中的任意一个。

[权利要求 28] 根据权利要求23-26中任一项所述的装置，其特征在于，第一子规则方式的输入还包括第四输入，所述第四输入是第三输入经冗余信息处理后得到的音频信号；当所述预先设置的规则方式包括多个子规则方式时，所述第三输入包括所述多个子规则方式中的其他至少一个子规则方式的输出，或者，所述第三输入包括所述至少一个子模型的输出；所述第一子规则方式是所述多个子规则方式中的任意一个。

[权利要求 29] 一种解码装置，其特征在于，包括：

接收模块，用于接收码流，所述码流包括第一码流和第二码流；

获取模块，用于根据所述第一码流获取第一音频信号，所述第一音频信号的频率低于预设阈值；

提取模块，用于根据所述第二码流获取边信息，所述边信息用于表征第二音频信号的特性，所述第二音频信号的频率高于或等于所述预设阈值；

重建模块，用于根据所述第一音频信号和所述边信息获取重建帧。

[权利要求 30] 根据权利要求29所述的装置，其特征在于，所述重建模块，具体用于根据所述第一音频信号和所述边信息通过预先训练得到的第一机器学习模型，获取所述重建帧。

[权利要求 31] 根据权利要求30所述的装置，其特征在于，所述重建帧是所述第一机器学习模型直接输出的；或者，所述重建帧是基于所述第一机器学习模型输出的所述第二音频信号与所述第一音频信号得到的。

[权利要求 32] 根据权利要求30或31所述的装置，其特征在于，当所述第一机器学习模型包括第一子模型和第二子模型时，所述重建模块，具体用于将所述第一音频信号和第一边信息输入所述第一子模型，以获取第三音频信号；将所述第三音频信号和第二边信息输入所述第二子模型，以获取所述重建帧；所述第一边信息和所述第二边信息分别是所述边信息中的部分或全部信息。

[权利要求 33] 根据权利要求30-32中任一项所述的装置，其特征在于，所述重建模块，具体用于根据所述第一音频信号和第三边信息按照预先设置的第一规则方式进行处理以获取第四音频信号；将所述第四音频信号和第四边信息输入所述第一机器学习模型，以获取所述重建帧；所述第三边信息和所述第四边信息分别是所述边信息中的部分或全部信息。

[权利要求 34] 根据权利要求30-33中任一项所述的装置，其特征在于，所述重建模块，具体用于根据所述第一音频信号和所述五边信息通过预先训练得到的第一机器学习模型，获取第五音频信号；根据所述第五音频信号和第六边信息按照预先设置的第二规则方式进行处理以获取所述重建帧；所述第五边信息和所述第六边信息分别是所述边信息中的部分或全部信息。

[权利要求 35] 根据权利要求29-34中任一项所述的装置，其特征在于，所述提取模块，具体用于解析所述第二码流以得到所述边信息。

[权利要求 36] 根据权利要求29-34中任一项所述的装置，其特征在于，所述提取模块，具体用于解析所述第二码流以得到处理前边信息；对所述处理前边信息进行边信息处理以得到所述边信息，所述边信息处理包括输入预先训练得到的第三机器学习模型或者按照预先设置的第三规则方式进行处理。

[权利要求 37] 一种设备，其特征在于，包括：

一个或多个处理器；

存储器，用于存储一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求1-18中任一项所述的方法。

[权利要求 38] 一种计算机可读存储介质，其特征在于，包括计算机程序，所述计算机程序在计算机上被执行时，使得所述计算机执行权利要求1-18中任一项所述的方法。

[权利要求 39] 一种计算机程序产品，其特征在于，所述计算机程序产品包括计算机程序代码，当所述计算机程序代码在计算机上运行时，使得计算机执行权利要求1-18中任一项所述的方法。