

音频信号的处理方法及其装置

本申请要求于 2022 年 07 月 27 日提交的申请号为 202210892836.1、发明名称为“音频信号的处理方法及其装置”的中国专利申请的优先权，及于 2022 年 09 月 19 日提交的申请号为 202211139722.6、发明名称为“音频信号的处理方法及其装置”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本申请涉及音频信号处理技术领域，特别涉及一种音频信号的处理方法及其装置。

背景技术

随着生活质量的提高，人们对高质量音频的需求不断增大。为了利用有限的带宽更好地传输音频信号，通常需要先对音频信号进行数据压缩，然后将经过压缩的码流传输到解码端。解码端对接收到的码流进行解码处理，得到解码后的音频信号，解码后的音频信号用于回放。

然而，对音频信号进行压缩的过程中可能会对音频信号的音质造成影响。因此，如何在保证音频信号的音质效果的同时，提升对音频信号的压缩效率，成为一个亟需解决的技术问题。

发明内容

本申请提供了一种音频信号的处理方法及其装置，在保证音质效果的同时，提升对音频信号进行编码的压缩效率。所述技术方案如下：

第一方面，本申请提供了一种音频信号的处理方法。该方法包括：获取音频信号的多个子带及每个子带的标度因子；基于多个子带的标度因子，确定用于对音频信号的频谱包络进行整形的基准值；以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，调节因子用于对音频信号的频谱值进行量化，和/或，调节因子用于对频谱值的码值进行反量化。

在本申请提供的音频信号的处理方法中，在获取音频信号的多个子带及每个子带的标度因子后，可以基于多个子带的标度因子，确定对音频信号的频谱包络进行整形的基准值，并以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子。由于调节因子用于对音频信号的频谱值进行量化，因此通过采用该方法在基准值的基础上对音频信号的频谱包络进行整形，使得基于整形得到的调节因子对音频信号的频谱值进行量化时，能够在保证音质效果的同时，提升对音频信号进行编码的压缩效率。

在一种实现方式中，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，包括：获取子带的标度因子与基准值的差距；基于差距对子带的标度因子进行调整，得到调节因子。

能量比较高的子带对能量比较低的子带有声学掩蔽效应。即当相邻的子带的能量不同时，相邻的子带之间存在掩蔽效应，在对音频信号进行整形时，可以对多个子带的标度因子执行掩蔽处理，以便于获取较好的音质。则可选地，在以基准值为基线，对音频信号的频谱包络

进行整形,得到经过整形的频谱包络对应的每个子带的调节因子之前,该方法还包括:对子带的标度因子进行掩蔽处理,使用子带经过掩蔽处理的标度因子更新子带的标度因子。则差距可以基于基准值和子带经过掩蔽处理的标度因子得到。

在一种实现方式中,当音频信号为双声道信号时,基于差距对子带的标度因子进行调整,得到调节因子,包括:对差距进行缩小处理,得到经过缩小处理的差距;基于经过缩小处理的差距和基准值,更新子带的标度因子;基于子带更新后的标度因子,得到调节因子。

其中,差距的缩小倍数基于差距的大小确定。当音频信号的强度大于基准值时,人耳对该音频信号更敏感,当音频信号的强度小于或等于基准值时,人耳对该音频信号的敏感度较弱,则当差距指示子带的标度因子大于基准值时,其缩小倍数可以小于差距指示子带的标度因子小于或等于基准值时的缩小倍数。

在另一种实现方式中,当音频信号为单声道信号时,可以以放大较大的标度因子,去掉较小的标度因子为原则,对子带的标度因子进行调整,则基于差距对子带的标度因子进行调整,得到调节因子,包括:将差距确定为调节因子。

可选地,在获取子带的标度因子与基准值的差距之前,该方法还包括:对子带的标度因子进行信号增强处理,使用子带经过信号增强处理的标度因子更新子带的标度因子。此时,差距基于基准值和子带经过信号增强处理的标度因子得到。

在一种实现方式中,当音频信号为双声道信号时,基准值基于多个子带的标度因子的均值得到;当音频信号为单声道信号时,基准值基于多个子带的标度因子中的最大值得到。

可选地,在基于多个子带的标度因子,确定用于对音频信号的频谱包络进行整形的基准值之前,该方法还包括:对子带的标度因子进行掩蔽处理,使用子带经过掩蔽处理的标度因子更新子带的标度因子。则基准值基于多个子带经过掩蔽处理的标度因子得到。

在一种实现方式中,当音频信号为单声道信号时,在基于多个子带的标度因子,确定用于对音频信号的频谱包络进行整形的基准值之前,该方法还包括:对子带的标度因子进行信号增强处理,使用子带经过信号增强处理的标度因子更新子带的标度因子。

可选地,对子带的标度因子进行信号增强处理的强度基于子带的频率和多个子带的总个数确定。在一种实现方式中,该强度可以基于该子带的频率在音频信号的频率中的占比大小确定。可选地,可以基于该子带的频率在音频信号的频率中的占比大小,对该子带的标度因子进行叠加,得到该子带经过信号增强处理的标度因子。

在一种实现方式中,对子带的标度因子进行掩蔽处理,包括:获取子带的相邻子带对子带的掩蔽系数及相邻子带的标度因子,掩蔽系数用于指示掩蔽程度;基于子带的标度因子、相邻子带的标度因子、及相邻子带对子带的掩蔽系数,获取子带经过掩蔽处理的标度因子。

可选地,当音频信号为双声道信号时,掩蔽系数基于子带的标度因子与基准值的大小关系确定;当音频信号为单声道信号时,掩蔽系数基于子带与相邻子带的频率关系确定。

本申请实施例提供的音频信号的处理方法可以在满足指定条件时执行,即以基准值为基线,对音频信号的频谱包络进行整形,得到经过整形的频谱包络对应的每个子带的调节因子,包括:当音频信号的码率小于码率阈值,和/或,音频信号的能量集中度小于集中度阈值时,以基准值为基线,对音频信号的频谱包络进行整形,得到经过整形的频谱包络对应的每个子带的调节因子。

码率是数据传输时单位时间传送的数据位数。传输音频信号的场景可以包括低码率的场

景和高码率的场景。低码率的场景通常发生在干扰较大的情况下，例如，在地铁、机场和火车站等信号较易受到干扰的环境中。高码率的场景通常发生在干扰较小的情况下，例如安静的室内等对信号干扰较小的环境中。而频域噪声整形是利用人耳听觉掩蔽原理对编解码器所产生的量化噪声谱进行整形。因此，可以根据码率确定是否对音频信号进行整形。

能量集中度用于表示音频内容在音频信号中的分布情况。根据音频信号的能量集中度能够区分音频信号是否包含实质内容，当音频信号包含实质内容时，可以对音频信号进行整形，以提高传输至音频接收设备的音频信号的音质。当音频信号不包括实质内容时，则无需对音频信号进行整形。

第二方面，本申请提供了一种音频信号的处理装置，该装置包括：获取模块，用于获取音频信号的多个子带及每个子带的标度因子；确定模块，用于基于多个子带的标度因子，确定用于对音频信号的频谱包络进行整形的基准值；处理模块，用于以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，调节因子用于对音频信号的频谱值进行量化，和/或，调节因子用于对频谱值的码值进行反量化。

可选地，处理模块具体用于：获取子带的标度因子与基准值的差距；基于差距对子带的标度因子进行调整，得到调节因子。

可选地，处理模块还用于：对子带的标度因子进行掩蔽处理，使用子带经过掩蔽处理的标度因子更新子带的标度因子。

可选地，当音频信号为双声道信号时，处理模块具体用于：对差距进行缩小处理，得到经过缩小处理的差距；基于经过缩小处理的差距和基准值，更新子带的标度因子；基于子带更新后的标度因子，得到调节因子。

可选地，差距的缩小倍数基于差距的大小确定。

可选地，当音频信号为单声道信号时，处理模块具体用于：将差距确定为调节因子。

可选地，处理模块还用于：对子带的标度因子进行信号增强处理，使用子带经过信号增强处理的标度因子更新子带的标度因子。

可选地，当音频信号为双声道信号时，基准值基于多个子带的标度因子的均值得到；当音频信号为单声道信号时，基准值基于多个子带的标度因子中的最大值得到。

可选地，处理模块还用于：对子带的标度因子进行掩蔽处理，使用子带经过掩蔽处理的标度因子更新子带的标度因子。

可选地，当音频信号为单声道信号时，处理模块还用于：对子带的标度因子进行信号增强处理，使用子带经过信号增强处理的标度因子更新子带的标度因子。

可选地，对子带的标度因子进行信号增强处理的强度基于子带的频率和多个子带的总个数确定。

可选地，处理模块具体用于：获取子带的相邻子带对子带的掩蔽系数及相邻子带的标度因子，掩蔽系数用于指示掩蔽程度；基于子带的标度因子、相邻子带的标度因子、及相邻子带对子带的掩蔽系数，获取子带经过掩蔽处理的标度因子。

可选地，当音频信号为双声道信号时，掩蔽系数基于子带的标度因子与基准值的大小关系确定；当音频信号为单声道信号时，掩蔽系数基于子带与相邻子带的频率关系确定。

可选地，处理模块具体用于：当音频信号的码率小于码率阈值，和/或，音频信号的能量集中度小于集中度阈值时，以基准值为基线，对音频信号的频谱包络进行整形，得到经过整

形的频谱包络对应的每个子带的调节因子。

第三方面，本申请提供了一种计算机设备，包括存储器和处理器，存储器存储有程序指令，处理器运行程序指令以执行本申请第一方面以及其任一种可能的实现方式中提供的方法。

第四方面，本申请提供了一种计算机可读存储介质，该计算机可读存储介质为非易失性计算机可读存储介质，该计算机可读存储介质包括程序指令，当程序指令在计算机设备上运行时，使得计算机设备执行本申请第一方面以及其任一种可能的实现方式中提供的方法。

第五方面，本申请提供了一种包含指令的计算机程序产品，当计算机程序产品在计算机上运行时，使得计算机执行本申请第一方面以及其任一种可能的实现方式中提供的方法。

附图说明

图1是本申请实施例提供的一种短距传输场景的示意图；

图2是本申请实施例提供的音频信号的处理方法所涉及的一种系统框架图；

图3是本申请实施例提供的一种音频编解码整体框架图；

图4是本申请实施例提供的一种计算机设备的结构示意图；

图5是本申请实施例提供的一种音频信号的处理方法的流程图；

图6是本申请实施例提供的一种对该子带的标度因子进行掩蔽处理的流程图；

图7是本申请实施例提供的一种以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子的流程图；

图8是本申请实施例提供的一种基于差距对子带的标度因子进行调整，得到调节因子的流程图；

图9是本申请实施例提供的一种音频信号的处理装置的结构示意图。

具体实施方式

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

首先对本申请实施例涉及的实施环境和背景知识进行介绍。

随着真无线立体声(true wireless stereo, TWS)耳机、智能音箱和智能手表等短距传输设备(如蓝牙设备)在人们日常生活中的广泛普及和使用，人们在各种场景下对追求高质量音频播放体验的需求也变得越来越迫切，尤其是在地铁、机场、火车站等蓝牙信号易受干扰的环境中。在短距传输场景中，由于连通音频发送设备与音频接收设备的信道对数据传输大小的限制，在进行音频信号传输时，为了减少音频信号传输时占用的带宽，通常使用音频发送设备中的音频编码器对音频信号进行编码，然后向音频接收设备传输。音频接收设备接收到经过编码的音频信号后，需要使用音频接收设备中的音频解码器对经过编码后的音频信号进行解码后，然后才能进行播放。可见，在短距传输设备普及的同时，也促使了各种音频编解码器的蓬勃发展。其中，短距传输场景可以包括蓝牙传输场景和无线传输场景等，本申请实施例以蓝牙传输场景为例，对本申请实施例提供的音频信号的处理方法进行说明。

目前，蓝牙音频编解码器有子带编码器(sub-band coding, SBC)、动态影像专家组(Moving Picture Experts Group, MPEG)的蓝牙高级音频编码器(advanced audio coding, AAC)系列(如AAC-LC、AAC-LD、AAC-HE、AAC-HEv2等)、LDAC、aptX系列(如aptX、aptX HD、

aptX low latency) 编码器、低延迟高清音频编解码器 (low-latency hi-definition audio codec, LHDC)、低功耗低延迟的 LC3 音频编解码器以及 LC3plus 等。

而编码会导致音频信号的高频成分丢失, 音质有所降低, 解码出的音频信号的听感较差。尤其在码率较低的场景下, 音频编解码器会通过降低带宽来节省码率, 这使得音频信号的高频成分丢失较多。所以, 为了提升音质, 音频发送设备可以先对音频信号进行处理, 然后根据经过处理的音频信号进行编码, 再向音频接收设备发送经过编码的音频信号。例如, 为了提高解码后音频信号的主观听觉质量, 音频发送设备可以对音频信号进行频域噪声整形处理, 然后向音频接收设备发送经过频域噪声整形处理和编码的音频信号。其中, 频域噪声整形是一种利用人耳听觉掩蔽原理对编解码器所产生的量化噪声谱进行整形的技术, 即把信号中的噪声谱调整为近似语音谱的形状, 利用人耳听觉掩蔽效应, 使得信号中的噪声不容易被察觉。

针对此, 本申请实施例提供了一种音频信号的处理方法。该音频信号的处理方法可视为一种频域噪声整形方法。该方法包括: 获取音频信号的多个子带及每个子带的标度因子; 基于多个子带的标度因子, 确定用于对音频信号的频谱包络进行整形的基准值; 以基准值为基线, 对音频信号的频谱包络进行整形, 得到经过整形的频谱包络对应的每个子带的调节因子, 调节因子用于对音频信号的频谱值进行量化, 和/或, 调节因子用于对频谱值的码值进行反量化。其中, 音频信号可以为语音信号或音乐信号等以音频形式呈现的信号。

通过在基准值的基础上对音频信号的频谱包络进行整形得到调节因子, 使得基于该调节因子对音频信号的频谱值进行量化, 和/或, 基于该调节因子对频谱值的码值进行反量化时, 能够在保证音质效果的同时, 提升对音频信号进行编码的压缩效率。

图 1 是本申请实施例提供的一种音频信号的处理方法涉及的应用场景的示意图。参见图 1, 该应用场景包括音频发送设备和音频接收设备。音频发送设备配置有音频编码器。音频接收设备配置有音频解码器。可选的, 音频发送设备可以是手机、电脑 (如笔记本电脑、台式电脑)、平板 (如手持式平板、车载式平板)、智能可穿戴设备等能够发送音频数据流的设备。音频接收设备可以是耳机 (如 TWS 耳机、无线头戴式耳机、无线颈圈式耳机)、音箱 (如智能音箱)、智能可穿戴设备 (如智能手表、智能眼镜)、智能车载设备等能够接收音频数据流并进行播放的设备。在一些场景中, 短距传输场景中的音频接收设备也可以是手机、电脑、平板等。

图 2 是本申请实施例提供的音频信号的处理方法所涉及的一种系统框架图。参见图 2, 该系统包括编码端和解码端。其中, 编码端包括输入模块、编码模块和发送模块。解码端包括接收模块、输入模块、解码模块和播放模块。

在编码端, 用户根据使用场景从两种编码模式中确定一种编码模式, 这两种编码模式为低延迟编码模式和高音质编码模式。这两种编码模式的编码帧长分别为 5ms 和 10ms。比如使用场景为打游戏、直播、通话等, 则用户可选择低延迟编码模式, 使用场景为通过耳机或音响欣赏音乐等, 则用户可选择高音质编码模式。用户还需要提供待编码的音频信号 (如图 2 所示的脉冲编码调制 (pulse code modulation, PCM) 数据) 给编码端。此外, 用户还需要设定

编码所得到的码流的目标码率，即音频信号的编码码率。其中，目标码率越高表示音质相对越好，但是在短距传输过程中码流的抗干扰性越差；目标码率越低，音质相对越差，但是在短距传输中码流的抗干扰性越高。简单来讲，编码端的输入模块获取用户提交的编码帧长、编码码率以及待编码的音频信号。

编码端的输入模块将用户提交的数据输入到编码模块的频域编码器中。

编码模块的频域编码器基于接收到的数据，通过编码以得到码流。其中，频域编码端对待编码的音频信号进行分析，以得到信号特点（包括单声道/双声道、平稳/非平稳、满带宽/窄带宽信号、主观/客观等），根据信号特点以及码率档位（即编码码率）进入相应的编码处理子模块，通过编码处理子模块来编码音频信号，以及打包码流的包头（包括采样率、声道数、编码模式、帧长等），最终得到码流。

编码端的发送模块将码流发送给解码端。可选地，该发送模块为如图 2 所示的发送模块或其他类型的发送模块，本申请实施例对此不作限定。

在解码端，解码端的接收模块接收到码流之后，将码流发送给解码模块的频域解码器中，并通知解码端的输入模块获取配置的位深和声道解码模式等。可选地，该接收模块为如图 2 所示的接收模块或其他类型的接收模块，本申请实施例对此不作限定。

解码端的输入模块将获取的位深和声道解码模式等信息输入到解码模块的频域解码器中。

解码模块的频域解码器基于位深、声道解码模式等来解码码流，以得到所需的音频数据（如图 2 所示的 PCM 数据），将得到的音频数据发送给播放模块，播放模块进行音频播放。其中，声道解码模式指示所需解码的声道。

图 3 是本申请实施例提供的一种音频编解码整体框架图。参见图 3，编码端的编码流程包括如下步骤：

（1）PCM 输入模块

输入 PCM 数据，该 PCM 数据为单声道数据或双声道数据，位深可以是 16 比特（bit）、24bit、32bit 浮点或 32bit 定点。可选地，PCM 输入模块将输入的 PCM 数据变换到同一位深，比如 24bit 位深，并对 PCM 数据进行解交织后按照左声道和右声道放置。

（2）加低延迟分析窗&改进离散余弦变换（modified discrete cosine transform, MDCT）变换模块

对步骤（1）处理后的 PCM 数据加低延迟分析窗以及进行 MDCT 变换后得到 MDCT 域的频谱数据。加窗的作用是防止频谱泄漏。

（3）MDCT 域信号分析模块&自适应带宽检测模块

MDCT 域信号分析模块在全码率场景下生效，自适应带宽检测模块在低码率（如码率 $\leq 150\text{kbps/声道}$ ）下激活。首先，根据上述步骤（2）得到的 MDCT 域的频谱数据，进行带宽检测，以得到截止频率或者说有效带宽。其次，对有效带宽内的频谱数据进行信号分析，即分析频点分布是集中的还是均匀的，以得到能量集中度，基于能量集中度得到指示待编码的音频信号是客观信号还是主观信号的标志（flag）（客观信号的标志为 1，主观信号的标志为 0）。如果是客观信号，在低码率下不对标度因子进行频域噪声整形（spectral noise shaping, SNS）处理和 MDCT 谱的平滑，因为这样会降低客观信号的编码效果。然后，基于带宽检测结果和主客观信号标志来确定是否进行 MDCT 域的子带截止操作。如果音频信号是客观信号，则不

做子带截止操作：如果音频信号是主观信号且带宽检测结果标识为 0（满带宽的），则子带截止操作由码率决定；如果音频信号是主观信号且带宽检测结果标识非 0（即带宽小于采样率的一半的有限带宽），则子带截止操作由带宽检测结果决定。

（4）子带划分选取和标度因子计算模块

根据码率档位以及上述步骤（3）得到的主客观信号标志和截止频率，从多种子带划分方式中选取最佳的子带划分方式，并得到编码该音频信号所需要的子带总个数。同时计算得到频谱的包络线，即计算所选取的划分子带方式对应的标度因子。

（5）MS 声道变换模块

针对双声道的 PCM 数据，根据上述步骤（4）计算得到的标度因子进行联合编码判别，即判别是否对左右声道数据进行 MS 声道变换。

（6）谱平滑模块和基于标度因子的频域噪声整形模块

谱平滑模块根据低码率的设定（如码率 $\leq 150\text{kbps/声道}$ ）进行 MDCT 谱平滑，频域噪声整形模块基于标度因子对经过谱平滑的数据进行频域噪声整形，得到调节因子，调节因子用于对音频信号的频谱值进行量化。其中，低码率的设定由低码率判别模块进行控制，当不满足低码率的设定时，无需进行谱平滑和频域噪声整形。

（7）标度因子编码模块

根据标度因子的分布对多个子带的标度因子进行差分编码或者熵编码。

（8）比特分配&MDCT 谱量化和熵编码模块

基于步骤（4）得到的标度因子和步骤（6）得到的调节因子，通过粗估和精估的比特分配策略来控制编码为恒定码率（constant bit rate, CBR）编码模式，并对 MDCT 谱值进行量化和熵编码。

（9）残余编码模块

若步骤（8）的比特消耗还没有达到目标比特，则进一步对子带进行重要性排序，将比特优先分配到重要子带的 MDCT 谱值的编码上。

（10）流包头信息打包模块

包头信息包括音频采样率（如 44.1kHz/48kHz/88.2kHz/96kHz）、声道信息（如单声道和双声道）、编码帧长（如 5ms 和 10ms）、编码模式（如时域、频域、时域切频域或频域切时域模式）等。

（11）比特流（即码流）发送模块

码流包含包头、边信息、载荷等。其中，包头携带包头信息，该包头信息如上述步骤（10）中的描述。边信息包括标度因子的编码码流、选取的子带划分方式的信息、截止频率信息、低码率标志、联合编码判别信息（即 MS 变换标志）、量化步长等信息。载荷包括 MDCT 频谱的编码码流和残余编码码流。

解码端的解码流程包括如下步骤：

（1）流包头信息解析模块

从接收的码流中解析出包头信息，包头信息包括音频信号的采样率、声道信息、编码帧长、编码模式等信息，根据码流大小、采样率和编码帧长计算得到编码码率，即得到码率档位信息。

（2）标度因子解码模块

从码流中解码出边信息,包括选取的子带划分方式的信息、截止频率信息、低码率标志、联合编码判别信息、量化步长等信息,以及各个子带的标度因子。

(3) 基于标度因子的频域噪声整形模块

在低码率(如编码码率小于 150kbps/声道)下,还需要基于标度因子做频域噪声整形,得到调节因子,调节因子用于对频谱值的码值进行反量化。其中,低码率的设定由低码率判别模块进行控制,当不满足低码率的设定时,无需进行频域噪声整形。

(4) MDCT 谱解码模块和残余解码模块

MDCT 谱解码模块根据上述步骤(2)得到的子带划分方式的信息、量化步长信息以及标度因子,解码码流中的 MDCT 频谱数据。在低码率档位下进行空洞补全,如计算得到比特还有剩余,则残余解码模块进行残余解码,以得到其他子带的 MDCT 频谱数据,进而最终的 MDCT 频谱数据。

(5) LR 声道变换模块

根据步骤(2)得到的边信息,如果根据联合编码判别判定是双声道联合编码模式且不是解码低功耗模式(如编码码率大于 150kbps/声道且采样率大于 88.2kHz),则对步骤(4)得到的 MDCT 频谱数据进行 LR 声道变换。

(6) 逆 MDCT 变换模块&加低延迟合成窗模块和交叠相加模块

逆 MDCT 变换模块在步骤(4)和步骤(5)的基础上,对得到的 MDCT 频谱数据进行 MDCT 逆变换,以得到时域混叠信号,然后加低延迟合成窗模块对时域混叠信号加低延迟合成窗,交叠相加模块将当前帧与上一帧的时域混叠缓存信号叠加得到 PCM 信号,即通过交叠相加得到最终的 PCM 数据。

(7) PCM 输出模块

根据配置的位深和声道解码模式,输出相应声道的 PCM 数据。

需要说明的是,图 3 所示的音频编解码框架仅作为本申请实施例终端一个示例,并不用于限制本申请实施例,本领域技术人员可以在图 3 的基础上得到其他的编解码框架。

请参考图 4,图 4 是根据本申请实施例示出的一种计算机设备的结构示意图。可选地,该计算机设备可以为图 1 中所示的任一设备。例如,该计算机设备可以是音频发送设备。此时,该计算机设备能够实现本申请实施例提供的音频信号的处理方法中的部分或全部功能。如图 4 所示,计算机设备 20 包括处理器 201、存储器 202、通信接口 203 和总线 204。其中,处理器 201、存储器 202、通信接口 203 通过总线 204 实现彼此之间的通信连接。

计算机设备 20 可以包括多个处理器,如图 4 中所示的处理器 201 和处理器 205。这些处理器中的每一个为一个单核处理器,或者一个多核处理器。可选地,这里的处理器指一个或多个设备、电路、和/或用于处理数据(如计算机程序指令)的处理核。处理器 201 可以包括通用处理器和/或专用硬件芯片。通用处理器可以包括:中央处理器(central processing unit, CPU)、微处理器或图形处理器(graphics processing unit, GPU)。CPU 例如是一个单核处理器(single-CPU),又如是一个多核处理器(multi-CPU)。专用硬件芯片是一个高性能处理的硬件模块。专用硬件芯片包括数字信号处理器、专用集成电路(application-specific integrated circuit, ASIC)、现场可编程逻辑门阵列(field-programmable gate array, FPGA)或者网络处理器(network processor, NP)中的至少一项。处理器 201 还可以是一种集成电路芯片,具有

信号的处理能力。在实现过程中，本申请的音频信号的处理方法的部分或全部功能，可以通过处理器 201 中的硬件的集成逻辑电路或者软件形式的指令完成。

存储器 202 用于存储计算机程序，计算机程序包括操作系统 202a 和可执行代码（即程序指令）202b。存储器 202 例如是只读存储器（read-only memory, ROM）或可存储静态信息和指令的其它类型的静态存储设备，又如是随机存取存储器（random access memory, RAM）或者可存储信息和指令的其它类型的动态存储设备，又如是电可擦可编程只读存储器（electrically erasable programmable read-only memory, EEPROM）、只读光盘（compact disc read-only memory, CD-ROM）或其它光盘存储、光碟存储（包括压缩光碟、激光碟、光碟、数字通用光碟、蓝光光碟等）、磁盘存储介质或者其它磁存储设备，或者是能够用于携带或存储具有指令或数据结构形式的期望的可执行代码并能够由计算机存取的任何其它介质，但不限于此。例如存储器 202 用于存放出端口队列等。存储器 202 例如是独立存在，并通过总线 204 与处理器 201 相连接。或者存储器 202 和处理器 201 集成在一起。存储器 202 可以存储可执行代码，当存储器 202 中存储的可执行代码被处理器 201 执行时，处理器 201 用于执行本申请实施例提供的音频信号的处理方法的部分或全部功能。且处理器 201 执行对应功能的实现方式请相应参考方法实施例中的相关描述。存储器 202 中还可以包括操作系统等其他运行进程所需的软件模块和数据等。

通信接口 203 使用例如但不限于收发器一类的收发模块，来实现与其他设备或通信网络之间的通信。通信接口 204 包括有线通信接口，可选地，还包括无线通信接口。其中，有线通信接口例如以太网接口等。可选地，以太网接口为光接口、电接口或其组合。无线通信接口为无线局域网（wireless local area networks, WLAN）接口、蜂窝网络通信接口或其组合等。

总线 204 是任何类型的，用于实现计算机设备的内部器件（例如，存储器 202、处理器 201、通信接口 203）互连的通信总线。例如系统总线。本申请实施例以计算机设备内部的上述器件通过总线 204 互连为例说明。可选地，计算机设备 20 内部的上述器件还可以采用除了总线 204 之外的其他连接方式彼此通信连接。例如，计算机设备 20 内部的上述器件通过内部的逻辑接口互连。

可选地，计算机设备还包括输出设备和输入设备。输出设备和处理器 201 通信，能够以多种方式来显示信息。例如，输出设备为液晶显示器（liquid crystal display, LCD）、发光二级管（light emitting diode, LED）显示设备、阴极射线管（cathode ray tube, CRT）显示设备或投影仪（projector）等。输入设备和处理器 201 通信，能够以多种方式接收用户的输入。例如，输入设备是鼠标、键盘、触摸屏设备或传感设备等。

需要说明的是，上述多个器件可以分别设置在彼此独立的芯片上，也可以至少部分的或者全部的设置在同一块芯片上。将各个器件独立设置在不同的芯片上，还是整合设置在一个或者多个芯片上，往往取决于产品设计的需要。本申请实施例对上述器件的具体实现形式不做限定。且上述各个附图对应的流程的描述各有侧重，某个流程中没有详述的部分，可以参见其他流程的相关描述。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。提供程序开发平台的计算机程序产品包括一个或多个计算机指令，在计算机设备上加载和执行这些计算机程序指令时，全部或部分地实现本申请实施例提供的音频信号的处理方法的流程或功能。

并且，计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。计算机可读存储介质存储有提供程序开发平台的计算机程序指令。

在一种可实现方式中，本申请实施例提供的音频信号的处理方法可以通过部署在计算机设备上的一个或多个功能模块实现。该一个或多个功能模块具体可以通过计算机设备执行可执行程序实现。当本申请实施例提供的音频信号的处理方法通过部署在计算机设备上的多个功能模块实现时，该多个功能模块可以采用集中式部署，或者采用分布式部署。并且，该多个功能模块具体可以通过一个或多个计算机设备执行计算机程序实现。该一个或多个计算机设备中的每个计算机设备能够实现本申请实施例提供的音频信号的处理方法中的部分或全部功能。

应当理解的是，以上内容是对本申请实施例提供的音频信号的处理方法的应用场景的示例性说明，并不构成对于该音频信号的处理方法的应用场景的限定。例如，本申请实施例在对音频信号的处理方法的实现过程进行说明时，是以该音频信号的处理方法应用于蓝牙传输场景的短距传输为例，但并不排除该音频信号的处理方法还可以应用于其他短距传输场景，如该音频信号的处理方法还可以应用于无线传输的短距传输场景和其他场景。则本申请实施例提供的音频信号的处理方法可以应用于短距传输场景中的音频发送设备，即短距传输场景中的编码端，也可以应用于其他的传输场景中的编码端，并且，还可以应用于短距传输场景中的音频接收设备，即短距传输场景中的解码端，也可以应用于其他的传输场景中的解码端。换种方式来讲，本申请实施例提供的音频信号的处理方法可以应用于所有编码音频信号相关的场景中。并且，对本领域普通技术人员可知，随着业务需求的改变，应用场景可以根据应用需求进行调整，本申请实施例对其不做一一列举。

图 5 是本申请实施例提供的一种音频信号的处理方法的流程图，该方法可以应用于如图 1 所示的音频发送设备和音频接收设备。下面以该方法应用于音频发送设备为例对其进行说明，如图 5 所示，该方法包括如下步骤：

步骤 301、获取音频信号的多个子带及每个子带的标度因子。

音频发送设备获取到待发送至音频接收设备的音频信号后，可以先对音频信号加低延迟分析窗，将加低延迟分析窗的音频信号变换至频域，得到音频信号的频域信号，然后基于该频域信号划分得到多个子带（如 32 个子带），并得到每个子带的标度因子（scale factor, SF）。其中，子带的标度因子用于指示子带中频点的最大幅值。例如，子带的标度因子可以为表示该最大幅值所需的比特数。

本申请实施例提供的音频信号的处理方法可以在满足以下条件时执行：音频信号的码率小于码率阈值，和/或，音频信号的能量集中度小于集中度阈值。例如，当音频信号的码率小于码率阈值，和/或，音频信号的能量集中度小于集中度阈值时，执行下面步骤 304。并且，当音频信号不满足该条件时，可以不执行本申请实施例提供的音频信号的处理方法，并将每个子带的调节因子初始化为 0。

码率是数据传输时单位时间传送的数据位数。传输音频信号的场景可以包括低码率的场

景和高码率的场景。低码率的场景通常发生在干扰较大的情况下，例如，在地铁、机场和火车站等信号较易受到干扰的环境中。高码率的场景通常发生在干扰较小的情况下，例如安静的室内等对信号干扰较小的环境中。而频域噪声整形是利用人耳听觉掩蔽原理对编解码器所产生的量化噪声谱进行整形。因此，可以根据码率确定是否对音频信号进行整形。码率阈值可以根据对音质的要求等需求进行调整。例如，码率阈值可以为 150 千位每秒 (kbps)。当码率小于码率阈值时，传输音频信号的码率为低码率，码率大于或等于码率阈值时，传输音频信号的码率为高码率。

能量集中度用于表示音频内容在音频信号中的分布情况。根据音频信号的能量集中度能够区分音频信号是否包含实质内容，当音频信号包含实质内容时，可以对音频信号进行整形，以提高传输至音频接收设备的音频信号的音质。当音频信号不包括实质内容时，则无需对音频信号进行整形。包含实质内容的音频信号可以称为主观信号，不包含实质内容的音频信号可以称为客观信号。集中度阈值可以根据对音质的要求等需求进行调整。例如，集中度阈值可以为 0.6。当能量集中度小于集中度阈值时，音频信号为主观信号，当能量集中度大于或等于集中度阈值时，音频信号为客观信号。

步骤 302、对每个子带的标度因子进行处理，使用子带经过掩蔽处理的标度因子更新子带的标度因子。

可选地，对标度因子的处理过程包括以下一种或多种：对多个子带的标度因子执行掩蔽处理、对多个子带的标度因子执行信号增强处理。

能量比较高的子带对能量比较低的子带有声学掩蔽效应。即当相邻的子带的能量不同时，相邻的子带之间存在掩蔽效应，在对音频信号进行整形时，可以对多个子带的标度因子执行掩蔽处理，以便于获取较好的音质。在一种实现方式中，如图 6 所示，对于多个子带中的任一子带，对该子带的标度因子进行掩蔽处理的过程，包括：

步骤 3021、获取子带的相邻子带对子带的掩蔽系数及相邻子带的标度因子。

每个子带具有索引值，根据当前子带的索引值可以确定其相邻子带的索引值，进而确定当前子带的相邻子带，然后根据步骤 301 中获取的多个子带的标度因子，可以得到相邻子带的标度因子。

当音频信号为双声道和单声道信号时，相邻子带对当前子带的掩蔽程度不同。下面分别对其进行说明：

当音频信号为双声道信号时，掩蔽程度可以基于子带的标度因子与对音频信号的频谱包络进行整形使用的基准值的大小关系确定。一般来说，由于基准值基于多个子带的标度因子得到，是整形时参考的基准值，则当音频信号为双声道信号时，子带的标度因子大于基准值时的掩蔽程度，通常大于子带的标度因子小于或等于基准值时的掩蔽程度。例如，当子带的标度因子大于基准值时，掩蔽系数可以为 0.375，当子带的标度因子小于或等于基准值时，掩蔽系数可以为 0.25。其中，掩蔽系数用于指示掩蔽程度。

当音频信号为单声道信号时，掩蔽程度可以基于子带与相邻子带的频率关系确定。一般来说，当音频信号为单声道信号时，频率大于子带的相邻子带对子带的掩蔽程度，小于频率小于子带的相邻子带对子带的掩蔽程度。例如，频率大于子带的相邻子带对子带的掩蔽系数可以为 0.125，频率小于子带的相邻子带对子带的掩蔽系数可以为 0.175。

步骤 3022、基于子带的标度因子、相邻子带的标度因子、及相邻子带对子带的掩蔽系数，

获取子带经过掩蔽处理的标度因子。

子带（即当前子带）经过掩蔽处理的标度因子可以基于使用掩蔽程度加权处理后的相邻子带与当前子带的标度因子之差得到。在一种实现方式中，该步骤 3022 的实现过程可以包括：获取相邻子带与当前子带的标度因子的差值，使用相邻子带对当前子带的掩蔽系数对该差值进行加权。例如，可以在当前子带的标度因子的基础上，叠加相邻子带对当前子带的加权值，以得到当前子带经过掩蔽处理的标度因子。并且，相邻子带的标度因子可以大于也可以小于当前子带的标度因子，则相邻子带与当前子带的标度因子的差值可能大于 0，也可能小于 0，但为了保证掩蔽效果，在使用掩蔽系数对差值进行加权时，可以使用掩蔽系数对差值和 0 中的较大值进行加权。

示例的，当音频信号为双声道信号时，第 b 个子带的标度因子 $E(b)$ 、相邻子带的标度因子 $E(b-1)$ 和 $E(b+1)$ ，当前子带经过掩蔽处理的标度因子 $E_{\text{new}}(b)$ ，可以满足：

$$E_{\text{new}}(b) = E(b) + c \times \text{MAX}(E(b-1) - E(b), 0) + c \times \text{MAX}(E(b+1) - E(b), 0);$$

$$E_{\text{new}}(0) = E(0) + c \times \text{MAX}(E(1) - E(0), 0);$$

$$E_{\text{new}}(B-1) = E(B-1) + c \times \text{MAX}(E(B-2) - E(B-1), 0);$$

其中， c 为相邻子带对第 b 个子带的掩蔽系数，其取值可以参考步骤 3021 中的相关说明， $b+1$ 用于指示相邻子带中比当前子带的频率大的子带， $b-1$ 用于指示相邻子带中比当前子带的频率小的子带， B 为音频信号具有的子带的总个数， b 、 $b-1$ 和 $b+1$ 的取值为 $[0, B-1]$ 中的整数。并且，该 B 个子带具体可以为音频信号子带中需要编码的子带，如根据音频信号的截止频率得到的需要编码的子带。

当音频信号为单声道信号时，第 b 个子带的标度因子 $E(b)$ 、相邻子带的标度因子 $E(b-1)$ 和 $E(b+1)$ ，当前子带经过掩蔽处理的标度因子 $E_{\text{new}}(b)$ ，可以满足：

$$E_{\text{new}}(b) = E(b) + c1 \times \text{MAX}(E(b-1) - E(b), 0) + c2 \times \text{MAX}(E(b+1) - E(b), 0);$$

$$E_{\text{new}}(0) = E(0) + c2 \times \text{MAX}(E(1) - E(0), 0);$$

$$E_{\text{new}}(B-1) = E(B-1) + c1 \times \text{MAX}(E(B-2) - E(B-1), 0);$$

其中， $b+1$ 用于指示相邻子带中比当前子带的频率大的子带， $b-1$ 用于指示相邻子带中比当前子带的频率小的子带， $c1$ 为第 $b-1$ 个子带对第 b 个子带的掩蔽系数， $c2$ 为第 $b+1$ 个子带对第 b 个子带的掩蔽系数， $c1$ 和 $c2$ 的取值可以参考步骤 3021 中的相关说明， B 为音频信号具有的子带的总个数， b 、 $b-1$ 和 $b+1$ 的取值为 $[0, B-1]$ 中的整数。并且，该 B 个子带具体可以为音频信号子带中需要编码的子带，如根据音频信号的截止频率得到的需要编码的子带。

当音频信号为单声道信号时，可以对多个子带的标度因子执行信号增强处理，得到多个子带经过信号增强处理的标度因子。并且，当需要对子带的标度因子执行掩蔽处理和信号增强处理时，可以先对子带的标度因子执行掩蔽处理，然后对经过掩蔽处理的标度因子执行信号增强处理。可选地，对子带的标度因子进行信号增强处理的强度基于子带的频率和多个子带的总个数确定。在一种实现方式中，该强度可以基于该子带的频率在音频信号的频率中的占比大小确定。可选地，可以基于该子带的频率在音频信号的频率中的占比大小，对该子带的标度因子进行叠加，得到该子带经过信号增强处理的标度因子。

示例的，第 b 个子带的标度因子 $E(b)$ 和该子带经过信号增强处理的标度因子 $E_{\text{inc}}(b)$ ，可以满足：

$$E_{inc}(b)=E(b)+3b/(B-1);$$

其中，B 为音频信号具有的子带的总个数，b、b-1 和 b+1 的取值为[0, B-1]中的整数。

步骤 303、基于多个子带的标度因子，确定用于对音频信号的频谱包络进行整形的基准值。

当音频信号为双声道信号和单声道信号时，获取该基准值的实现方式不同。下面分别对其进行说明：

当音频信号为双声道信号时，基准值基于多个子带的标度因子的均值得到。例如，基准值 E_{avg} 和多个子带的标度因子 $E(i)$ 可以满足：

$$E_{avg} = \frac{1}{B} \times \sum_{i=0}^{B-1} E(i) + 1;$$

其中，B 为音频信号具有的子带的总个数，i 的取值为[0, B-1]中的整数。

可选地，当对多个子带的标度因子执行了掩蔽处理，该基准值可以基于多个子带经过掩蔽处理的标度因子得到。例如，当基准值基于多个子带的标度因子的均值得到时，用于计算均值的标度因子可以为经过掩蔽处理的标度因子。

当音频信号为单声道信号时，基准值基于多个子带的标度因子中的最大值得到。可选地，当对多个子带的标度因子执行了掩蔽处理，该基准值可以基于多个子带经过掩蔽处理的标度因子得到。例如，当基准值基于多个子带的标度因子中的最大值得到时，该最大值为多个子带的经过掩蔽处理的标度因子中的最大值。可选地，当对多个子带的标度因子执行了信号增强处理，该基准值可以基于多个子带经过信号增强处理的标度因子得到。例如，当基准值基于多个子带的标度因子中的最大值得到时，该最大值为多个子带的经过信号增强处理的标度因子中的最大值。

步骤 304、以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子。

当本申请实施例提供的音频信号的处理方法应用于音频发送设备时，音频发送设备可以根据该调节因子对音频信号的频谱值进行量化。当本申请实施例提供的音频信号的处理方法应用于音频接收设备时，音频接收设备可以根据该调节因子对频谱值的码值进行反量化。

在一种可实现方式中，可以基于子带的标度因子和对音频信号的频谱包络进行整形的基准值，对音频信号的频谱包络进行整形。如图 7 所示，步骤 304 的实现过程可以包括：

步骤 3041、获取子带的标度因子与基准值的差距。

子带的标度因子与基准值的差距，可以通过子带的标度因子与基准值的差值表示。并且，当对子带的标度因子执行了掩蔽处理，则该差距可以基于基准值和子带经过掩蔽处理的标度因子得到。当音频信号为单声道信号时，若对子带的标度因子执行了信号增强处理，则该差距可以基于基准值和子带经过信号增强处理的标度因子得到。当音频信号为单声道信号时，若对子带的标度因子执行了掩蔽处理和信号增强处理，则该差距可以基于基准值和子带经过信号掩蔽处理和增强处理的标度因子得到。

步骤 3042、基于差距对子带的标度因子进行调整，得到调节因子。

当音频信号为双声道信号和单声道信号时，基于差距对子带的标度因子进行调整的方式不同，下面分别对其进行说明：

当音频信号为单声道信号时，可以以放大较大的标度因子，去掉较小的标度因子为原则，对子带的标度因子进行调整。此时，该步骤 3042 的实现过程包括：将差距确定为调节因子。

在一种实现方式中,当基准值基于多个子带的标度因子中的最大值得到时,第 b 个子带经过信号掩蔽处理和增强处理的标度因子 $E_{inc}(b)$ 、多个子带的标度因子中的最大值 E_{max} 和第 b 个子带的调节因子 $dr_{adjust}(b)$ 满足:

$$dr_{adjust}(b) = E_{inc}(b) - E_{max}$$

根据对单声道信号的标度因子的调整过程可以看出:对单声道信号的标度因子进行整形实际是放大较大的标度因子,去掉较小的标度因子的过程。通过放大较大的标度因子,能够有效保留中高频的信号,通过去掉较小的标度因子,能够删除人耳不易察觉的信号,能够节省量化比特数,降低码率。因此,通过该调整方式能够保留更多的中高频的信息,并在保证音质效果的同时,提升对音频信号进行编码的压缩效率。

当音频信号为双声道信号时,可以以保持音频信号的频谱形状,并将频谱整体缩小的原则,基于差距对子带的标度因子进行调整,得到调节因子。此时,如图8所示,该步骤3042的实现过程包括:

步骤a1、对差距进行缩小处理,得到经过缩小处理的差距。

可选地,差距的缩小倍数可以基于差距的大小确定。当音频信号的强度大于基准值时,人耳对该音频信号更敏感,当音频信号的强度小于或等于基准值时,人耳对该音频信号的敏感度较弱,则当差距指示子带的标度因子大于基准值时,其缩小倍数可以小于差距指示子带的标度因子小于或等于基准值时的缩小倍数。并且,该缩小倍数的具体取值可以根据对音质的需求确定。例如,当差距指示子带的标度因子大于基准值时,其缩小倍数可以为0.375,当差距指示子带的标度因子小于或等于基准值时,其缩小倍数可以为0.5。

步骤a2、基于经过缩小处理的差距和基准值,更新子带的标度因子。

在一种可实现方式中,可以在基准值的基础上叠加经过缩小处理的差距,得到子带的更新后的标度因子。当对子带的标度因子执行了掩蔽处理,则第 b 个子带的更新后的标度因子 $E_z(b)$ 、子带经过掩蔽处理的标度因子 $E_{new}(b)$ 、基准值 E_{avg} 满足:

$$E_z(b) = \begin{cases} E_{avg} + 0.375 \times (E_{new}(b) - E_{avg}), & E_{new}(b) > E_{avg} \\ E_{avg} + 0.5 \times (E_{new}(b) - E_{avg}), & E_{new}(b) \leq E_{avg} \end{cases}$$

步骤a3、基于子带更新后的标度因子,得到调节因子。

在基于经过缩小处理的差距和基准值更新子带的标度因子后,可以基于子带更新后的标度因子和子带的原始标度因子,确定该子带的调节因子。在一种可实现方式中,可以将子带的原始标度因子与更新后的标度因子的差值,确定为该子带的调节因子。例如,第 b 个子带的原始标度因子 $E(b)$ 、更新后的标度因子 $E_z(b)$ 和调节因子 $dr_{adjust}(b)$ 满足:

$$dr_{adjust}(b) = E(b) - E_z(b)$$

根据对双声道信号的标度因子的调整过程可以看出:对双声道信号的标度因子进行整形实际是保持音频信号的频谱形状,并将频谱整体缩小的过程。双声道信号包括左声道信号和右声道信号,两个声道的信号具有能量差异,通过保持音频信号的频谱形状,并将频谱整体缩小,能够在有效减小中高频信号的损失同时,去掉人耳不易察觉的信号,能够节省量化比特数,降低码率。因此,通过该调整方式能够保留更多的中高频的信息,并在保证音质效果的同时,提升对音频信号进行编码的压缩效率。并且,当双声道信号中两个声道的信号能量差异较大时,该作用体现的尤为明显。

综上所述，在本申请实施例提供的音频信号的处理方法中，在获取音频信号的多个子带及每个子带的标度因子后，可以基于多个子带的标度因子，确定对音频信号的频谱包络进行整形的基准值，并以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子。由于调节因子用于对音频信号的频谱值进行量化，因此通过采用该方法在基准值的基础上对音频信号的频谱包络进行整形，使得基于整形得到的调节因子对音频信号的频谱值进行量化时，能够在保证音质效果的同时，提升对音频信号进行编码的压缩效率。

需要说明的是，本申请实施例提供的方法的步骤先后顺序可以进行适当调整，步骤也可以根据情况进行相应增减。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，可轻易想到变化的方法，都应涵盖在本申请的保护范围之内，因此不再赘述。

本申请提供了一种音频信号的处理装置。如图9所示，音频信号的处理装置90包括：获取模块901，用于获取音频信号的多个子带及每个子带的标度因子。

确定模块902，用于基于多个子带的标度因子，确定用于对音频信号的频谱包络进行整形的基准值。

处理模块903，用于以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，调节因子用于对音频信号的频谱值进行量化，和/或，调节因子用于对频谱值的码值进行反量化。

可选地，处理模块903具体用于：获取子带的标度因子与基准值的差距；基于差距对子带的标度因子进行调整，得到调节因子。

可选地，处理模块903还用于：对子带的标度因子进行掩蔽处理，使用子带经过掩蔽处理的标度因子更新子带的标度因子。

可选地，当音频信号为双声道信号时，处理模块903具体用于：对差距进行缩小处理，得到经过缩小处理的差距；基于经过缩小处理的差距和基准值，更新子带的标度因子；基于子带更新后的标度因子，得到调节因子。

可选地，差距的缩小倍数基于差距的大小确定。

可选地，当音频信号为单声道信号时，处理模块903具体用于：将差距确定为调节因子。

可选地，处理模块903还用于：对子带的标度因子进行信号增强处理，使用子带经过信号增强处理的标度因子更新子带的标度因子。

可选地，当音频信号为双声道信号时，基准值基于多个子带的标度因子的均值得到。

当音频信号为单声道信号时，基准值基于多个子带的标度因子中的最大值得到。

可选地，处理模块903还用于：对子带的标度因子进行掩蔽处理，使用子带经过掩蔽处理的标度因子更新子带的标度因子。

可选地，当音频信号为单声道信号时，处理模块903还用于：对子带的标度因子进行信号增强处理，使用子带经过信号增强处理的标度因子更新子带的标度因子。

可选地，对子带的标度因子进行信号增强处理的强度基于子带的频率和多个子带的总个数确定。

可选地，处理模块903具体用于：获取子带的相邻子带对子带的掩蔽系数及相邻子带的标度因子，掩蔽系数用于指示掩蔽程度；基于子带的标度因子、相邻子带的标度因子、及相

邻子带对子带的掩蔽系数，获取子带经过掩蔽处理的标度因子。

可选地，当音频信号为双声道信号时，掩蔽系数基于子带的标度因子与基准值的大小关系确定。

当音频信号为单声道信号时，掩蔽系数基于子带与相邻子带的频率关系确定。

可选地，处理模块 903 具体用于：当音频信号的码率小于码率阈值，和/或，音频信号的能量集中度小于集中度阈值时，以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子。

综上所述，在本申请实施例提供的音频信号的处理装置中，在获取音频信号的多个子带及每个子带的标度因子后，可以基于多个子带的标度因子，确定对音频信号的频谱包络进行整形的基准值，并以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子。由于调节因子用于对音频信号的频谱值进行量化，因此通过采用该装置在基准值的基础上对音频信号的频谱包络进行整形，使得基于整形得到的调节因子对音频信号的频谱值进行量化时，能够在保证音质效果的同时，提升对音频信号进行编码的压缩效率。

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，上述描述的装置和模块的具体工作过程，可以参考前述方法实施例中的对应内容，在此不再赘述。

本申请实施例提供了一种计算机设备。该计算机设备包括存储器和处理器，存储器存储有程序指令，处理器运行程序指令以执行本申请实施例提供的方法。例如，执行以下过程：获取音频信号的多个子带及每个子带的标度因子；基于多个子带的标度因子，确定用于对音频信号的频谱包络进行整形的基准值；以基准值为基线，对音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，调节因子用于对音频信号的频谱值进行量化，和/或，调节因子用于对频谱值的码值进行反量化。并且，计算机设备通过执行存储器中的程序指令，执行本申请实施例提供的方法的步骤的实现过程可以相应参考上述方法实施例中对应的描述。可选地，图 4 为本申请实施例提供一种计算机设备的结构示意图。

本申请实施例还提供了一种计算机可读存储介质，该计算机可读存储介质为非易失性计算机可读存储介质，该计算机可读存储介质包括程序指令，当程序指令在计算机设备上运行时，使得计算机设备执行如本申请实施例提供的方法。

本申请实施例还提供了一种包含指令的计算机程序产品，当计算机程序产品在计算机上运行时，使得计算机执行本申请实施例提供的方法。

应当理解的是，本文提及的“至少一个”是指一个或多个，“多个”是指两个或两个以上。在本申请实施例的描述中，除非另有说明，“/”表示或的意思，例如，A/B 可以表示 A 或 B；本文中的“和/或”仅仅是一种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：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。另外，为了便于清楚描述本申请实施例的技术方案，在本申请的实施例中，采用了“第一”、“第二”等字样对功能和作用基本相同的相同项或相似项进行区分。本领域技术人员可以理解“第一”、“第二”等字

样并不对数量和执行次序进行限定，并且“第一”、“第二”等字样也并不限定一定不同。

需要说明的是，本申请实施例所涉及的信息（包括但不限于用户设备信息、用户个人信息等）、数据（包括但不限于用于分析的数据、存储的数据、展示的数据等）以及信号，均为经用户授权或者经过各方充分授权的，且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准。例如，本申请实施例中涉及到的音频信号都是在充分授权的情况下获取的。

以上所述为本申请提供的实施例，并不用以限制本申请，凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

权利要求书

- 1.一种音频信号的处理方法，其特征在于，所述方法包括：
获取音频信号的多个子带及每个子带的标度因子；
基于所述多个子带的标度因子，确定用于对所述音频信号的频谱包络进行整形的基准值；
以所述基准值为基线，对所述音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，所述调节因子用于对所述音频信号的频谱值进行量化，和/或，所述调节因子用于对所述频谱值的码值进行反量化。
- 2.如权利要求1所述的方法，其特征在于，所述以所述基准值为基线，对所述音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，包括：
获取所述子带的标度因子与所述基准值的差距；
基于所述差距对所述子带的标度因子进行调整，得到所述调节因子。
- 3.如权利要求2所述的方法，其特征在于，在所述以所述基准值为基线，对所述音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子之前，所述方法还包括：
对所述子带的标度因子进行掩蔽处理，使用所述子带经过掩蔽处理的标度因子更新所述子带的标度因子。
- 4.如权利要求2或3所述的方法，其特征在于，当所述音频信号为双声道信号时，所述基于所述差距对所述子带的标度因子进行调整，得到所述调节因子，包括：
对所述差距进行缩小处理，得到经过缩小处理的差距；
基于经过缩小处理的差距和所述基准值，更新所述子带的标度因子；
基于所述子带更新后的标度因子，得到所述调节因子。
- 5.如权利要求4所述的方法，其特征在于，所述差距的缩小倍数基于所述差距的大小确定。
- 6.如权利要求2或3所述的方法，其特征在于，当所述音频信号为单声道信号时，所述基于所述差距对所述子带的标度因子进行调整，得到所述调节因子，包括：
将所述差距确定为所述调节因子。
- 7.如权利要求6所述的方法，其特征在于，在所述获取所述子带的标度因子与所述基准值的差距之前，所述方法还包括：
对所述子带的标度因子进行信号增强处理，使用所述子带经过信号增强处理的标度因子更新所述子带的标度因子。
- 8.如权利要求1至3任一所述的方法，其特征在于，
当所述音频信号为双声道信号时，所述基准值基于所述多个子带的标度因子的均值得到；
当所述音频信号为单声道信号时，所述基准值基于所述多个子带的标度因子中的最大值得到。
- 9.如权利要求8所述的方法，其特征在于，在所述基于所述多个子带的标度因子，确定用于对所述音频信号的频谱包络进行整形的基准值之前，所述方法还包括：
对所述子带的标度因子进行掩蔽处理，使用所述子带经过掩蔽处理的标度因子更新所述子带的标度因子。
- 10.如权利要求8或9所述的方法，其特征在于，当所述音频信号为单声道信号时，在所

述基于所述多个子带的标度因子，确定用于对所述音频信号的频谱包络进行整形的基准值之前，所述方法还包括：

对所述子带的标度因子进行信号增强处理，使用所述子带经过信号增强处理的标度因子更新所述子带的标度因子。

11.如权利要求 7 或 10 所述的方法，其特征在于，对所述子带的标度因子进行信号增强处理的强度基于所述子带的频率和所述多个子带的总个数确定。

12.如权利要求 3 或 9 所述的方法，其特征在于，所述对所述子带的标度因子进行掩蔽处理，包括：

获取所述子带的相邻子带对所述子带的掩蔽系数及所述相邻子带的标度因子，所述掩蔽系数用于指示掩蔽程度；

基于所述子带的标度因子、所述相邻子带的标度因子、及所述相邻子带对所述子带的掩蔽系数，获取所述子带经过掩蔽处理的标度因子。

13.如权利要求 12 所述的方法，其特征在于，

当所述音频信号为双声道信号时，所述掩蔽系数基于所述子带的标度因子与所述基准值的大小关系确定；

当所述音频信号为单声道信号时，所述掩蔽系数基于所述子带与所述相邻子带的频率关系确定。

14.如权利要求 1 至 13 任一所述的方法，其特征在于，所述以所述基准值为基线，对所述音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，包括：

当所述音频信号的码率小于码率阈值，和/或，所述音频信号的能量集中度小于集中度阈值时，以所述基准值为基线，对所述音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子。

15.一种音频信号的处理装置，其特征在于，所述装置包括：

获取模块，用于获取音频信号的多个子带及每个子带的标度因子；

确定模块，用于基于所述多个子带的标度因子，确定用于对所述音频信号的频谱包络进行整形的基准值；

处理模块，用于以所述基准值为基线，对所述音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子，所述调节因子用于对所述音频信号的频谱值进行量化，和/或，所述调节因子用于对所述频谱值的码值进行反量化。

16.如权利要求 15 所述的装置，其特征在于，所述处理模块具体用于：

获取所述子带的标度因子与所述基准值的差距；

基于所述差距对所述子带的标度因子进行调整，得到所述调节因子。

17.如权利要求 16 所述的装置，其特征在于，所述处理模块还用于：

对所述子带的标度因子进行掩蔽处理，使用所述子带经过掩蔽处理的标度因子更新所述子带的标度因子。

18.如权利要求 16 或 17 所述的装置，其特征在于，当所述音频信号为双声道信号时，所述处理模块具体用于：

对所述差距进行缩小处理，得到经过缩小处理的差距；

基于经过缩小处理的差距和所述基准值，更新所述子带的标度因子；

基于所述子带更新后的标度因子，得到所述调节因子。

19.如权利要求 18 所述的装置，其特征在于，所述差距的缩小倍数基于所述差距的大小确定。

20.如权利要求 16 或 17 所述的装置，其特征在于，当所述音频信号为单声道信号时，所述处理模块具体用于：

将所述差距确定为所述调节因子。

21.如权利要求 20 所述的装置，其特征在于，所述处理模块还用于：

对所述子带的标度因子进行信号增强处理，使用所述子带经过信号增强处理的标度因子更新所述子带的标度因子。

22.如权利要求 15 至 17 任一所述的装置，其特征在于，

当所述音频信号为双声道信号时，所述基准值基于所述多个子带的标度因子的均值得到；

当所述音频信号为单声道信号时，所述基准值基于所述多个子带的标度因子中的最大值得到。

23.如权利要求 22 所述的装置，其特征在于，所述处理模块还用于：

对所述子带的标度因子进行掩蔽处理，使用所述子带经过掩蔽处理的标度因子更新所述子带的标度因子。

24.如权利要求 22 或 23 所述的装置，其特征在于，当所述音频信号为单声道信号时，所述处理模块还用于：

对所述子带的标度因子进行信号增强处理，使用所述子带经过信号增强处理的标度因子更新所述子带的标度因子。

25.如权利要求 21 或 24 所述的装置，其特征在于，对所述子带的标度因子进行信号增强处理的强度基于所述子带的频率和所述多个子带的总个数确定。

26.如权利要求 17 或 23 所述的装置，其特征在于，所述处理模块具体用于：

获取所述子带的相邻子带对所述子带的掩蔽系数及所述相邻子带的标度因子，所述掩蔽系数用于指示掩蔽程度；

基于所述子带的标度因子、所述相邻子带的标度因子、及所述相邻子带对所述子带的掩蔽系数，获取所述子带经过掩蔽处理的标度因子。

27.如权利要求 26 所述的装置，其特征在于，

当所述音频信号为双声道信号时，所述掩蔽系数基于所述子带的标度因子与所述基准值的大小关系确定；

当所述音频信号为单声道信号时，所述掩蔽系数基于所述子带与所述相邻子带的频率关系确定。

28.如权利要求 15 至 27 任一所述的装置，其特征在于，所述处理模块具体用于：

当所述音频信号的码率小于码率阈值，和/或，所述音频信号的能量集中度小于集中度阈值时，以所述基准值为基线，对所述音频信号的频谱包络进行整形，得到经过整形的频谱包络对应的每个子带的调节因子。

29.一种计算机设备，其特征在于，包括存储器和处理器，所述存储器存储有程序指令，所述处理器运行所述程序指令以执行权利要求 1 至 14 任一所述的方法。

30.一种计算机可读存储介质，其特征在于，所述存储介质内存储有计算机程序，所述计算机程序被处理器执行时实现权利要求1至14任一所述的方法的步骤。

31.一种计算机程序产品，其特征在于，所述计算机程序产品内存储有计算机指令，所述计算机指令被处理器执行时实现权利要求1至14任一所述的方法的步骤。