

一种媒体数据的传输方法、装置和系统

本申请要求于2013年5月30日提交中国专利局、申请号为201310209760.9, 发明名称为“信号编码方法及设备”的中国专利申请优先权, 上述专利的全部内容通过引用结合在本申请中。

5

技术领域

本发明涉及信号处理领域, 并且具体地, 涉及信号编码方法及设备。

背景技术

非连续传输系统 (Discontinuous Transmission, DTX) 是一种被广泛应用的
10 语音通信系统, 能够在语音通信的静音期间采用非连续的编码和传输语音帧的方式减少信道带宽的占用, 同时仍能够保证足够的主观通话质量。

语音信号通常可以分为两类, 即活动语音信号和静音信号。活动语音信号是指包含有通话语音的信号, 而静音信号则是指不含有通话语音的信号。在 DTX 系统中, 对活动语音信号采用连续传输的方法进行传输, 对静音信号采用非连续传输的方法进行传输。这种对静音信号的非连续传输, 是通过编码端间歇地
15 编码并发送一种叫静音描述帧 (Silence Descriptor, SID) 的特殊编码帧来实现的, 两个相邻的 SID 帧之间 DTX 系统不会编码任何其它的信号帧。解码端根据非连续接收到的 SID 帧自主地生成令用户主观听觉舒适的噪声。这种舒适噪声 (Comfort Noise, CN) 并非以如实的还原原始静音信号为目的, 而是为了满足
20 解码端用户的主观听觉质量要求, 不要有不适感。

为了在解码端获得更好的主观听觉质量, 由语音活动段到 CN 段的过渡质量是至关重要的。为了获得更平滑的过渡, 一种有效的方法是: 在由语音活动段过渡到静音段时, 编码端不立即过渡到非连续传输状态, 而是额外延迟一段时间。在这段时间内, 静音段开头的部分静音帧仍然被视作语音活动帧连续的编
25 码和发送, 即设置一个连续传输的拖尾区间。这样做的好处是解码端可以充分地利用这段拖尾区间内的静音信号更好地估计和提取静音信号的特征, 以生成更优的 CN。

但是, 在现有技术中没有对拖尾机制进行高效的控制。拖尾机制的触发条件是比较简单的, 即通过简单地统计在语音活动结束时是否有足够数量的语音
30 活动帧被连续地编码和发送来确定是否触发拖尾机制, 而触发拖尾机制后, 固

定长度的拖尾区间就会被强制执行。然而，并非有足够数量的语音活动帧被连续地编码和发送就一定需要执行固定长度的拖尾区间，例如在通信环境的背景噪声比较平稳时，即使不设置拖尾区间或设置较短的拖尾区间，解码端也能获得优质的 CN。因此，这种对拖尾机制的简单控制模式造成了通信带宽的浪费。

5 发明内容

本发明实施例提供信号编码方法及设备，能够节省通信带宽。

第一方面，提供了一种信号编码方法，包括：在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在所述当前输入帧被编码为静音描述 SID 帧的情况下解码器根据所述当前输入帧生成的舒适噪声，并确定实际静音信号，
10 其中所述当前输入帧为静音帧；确定所述舒适噪声与所述实际静音信号的偏离程度；根据所述偏离程度，确定所述当前输入帧的编码方式，所述当前输入帧的编码方式包括拖尾帧编码方式或 SID 帧编码方式；根据所述当前输入帧的编码方式，对所述当前输入帧进行编码。

结合第一方面，在第一种可能的实现方式中，所述预测在所述当前输入帧
15 被编码为 SID 帧的情况下解码器根据所述当前输入帧生成的舒适噪声，并确定实际静音信号，包括：预测所述舒适噪声的特征参数，并确定所述实际静音信号的特征参数，其中所述舒适噪声的特征参数与所述实际静音信号的特征参数是一一对应的；

所述确定所述舒适噪声与所述实际静音信号的偏离程度，包括：确定所述
20 舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离。

结合第一方面的第一种可能的实现方式，在第二种可能的实现方式中，所述根据所述偏离程度，确定所述当前输入帧的编码方式，包括：在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述 SID 帧编码方式，其中

所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离与所述阈值集合中的阈值是一一对应的；在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离大于或等于所述阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述拖尾帧编码方式。

- 5 结合第一方面的第一种可能的实现方式或第二种可能的实现方式，在第三种可能的实现方式中，所述舒适噪声的特征参数用于表征以下至少一种信息：能量信息，谱信息。

结合第一方面的第三种可能的实现方式，在第四种可能的实现方式中，所述能量信息包括码激励线性预测 CELP 激励能量；

- 10 所述谱信息包括以下至少一种：线性预测滤波器系数，快速傅立叶变换 FFT 系数，修正离散余弦变换 MDCT 系数；

所述线性预测滤波器系数包括以下至少一种：线谱频率 LSF 系数，线谱对 LSP 系数，导抗谱频率 ISF 系数，导谱对 ISP 系数，反射系数，线性预测编码 LPC 系数。

- 15 结合第一方面的第一种可能的实现方式至第四种可能的实现方式中任一实现方式，在第五种可能的实现方式中，所述预测所述舒适噪声的特征参数，包括：根据所述当前输入帧的前一帧的舒适噪声参数和所述当前输入帧的特征参数，预测所述舒适噪声的特征参数；或者，根据所述当前输入帧之前的 L 个拖尾帧的特征参数和所述当前输入帧的特征参数，预测所述舒适噪声的特征参数，
- 20 其中 L 为正整数。

结合第一方面的第一种可能的实现方式至第五种可能的实现方式中任一实现方式，在第六种可能的实现方式中，所述确定所述实际静音信号的特征参数，包括：确定所述当前输入帧的特征参数作为所述实际静音信号的特征参数；或

者，对 M 个静音帧的特征参数进行统计处理，以确定所述实际静音信号的特征参数。

结合第一方面的第六种可能的实现方式，在第七种可能的实现方式中，所述 M 个静音帧包括所述当前输入帧以及所述当前输入帧之前的 $(M-1)$ 个静音
5 帧， M 为正整数。

结合第一方面的第二种可能的实现方式，在第八种可能的实现方式中，所述舒适噪声的特征参数包括所述舒适噪声的码激励线性预测 CELP 激励能量和所述舒适噪声的线谱频率 LSF 系数，所述实际静音信号的特征参数包括所述实际静音信号的 CELP 激励能量和所述实际静音信号的 LSF 系数；

10 所述确定所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离，包括：确定所述舒适噪声的 CELP 激励能量与所述实际静音信号的 CELP 激励能量之间的距离 D_e ，并确定所述舒适噪声的 LSF 系数与所述实际静音信号的 LSF 系数之间的距离 D_{lsf} 。

结合第一方面的第八种可能的实现方式，在第九种可能的实现方式中，所述
15 述在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述 SID 帧编码方式，包括：在所述距离 D_e 小于第一阈值，且所述距离 D_{lsf} 小于第二阈值的情况下，确定所述当前输入帧的编码方式为所述 SID 帧编码方式；

所述在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距
20 离大于或等于所述阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述拖尾帧编码方式，包括：在所述距离 D_e 大于或等于第一阈值，或者所述距离 D_{lsf} 大于或等于第二阈值的情况下，确定所述当前输入帧的编码方式为所述拖尾帧编码方式。

结合第一方面的第九种可能的实现方式，在第十种可能的实现方式中，还包括：获取预设的所述第一阈值和预设的所述第二阈值；或者，根据所述当前输入帧之前的 N 个静音帧的 CELP 激励能量确定所述第一阈值，并根据所述 N 个静音帧的 LSF 系数确定所述第二阈值，其中 N 为正整数。

5 结合第一方面或第一方面的第一种可能的实现方式至第十种可能的实现方式中任一实现方式，在第十一种可能的实现方式中，所述预测在所述当前输入帧被编码为 SID 帧的情况下解码器根据所述当前输入帧生成的舒适噪声，包括：采用第一预测方式，预测所述舒适噪声，其中所述第一预测方式与所述解码器生成所述舒适噪声的方式相同。

10 第二方面，提供了一种信号处理方法，包括：确定 P 个静音帧中每个静音帧的组加权谱距离，其中所述 P 个静音帧中每个静音帧的组加权谱距离为所述 P 个静音帧中所述每个静音帧与其它 (P-1) 个静音帧之间的加权谱距离之和，P 为正整数；根据所述 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，其中所述第一谱参数用于生成舒适噪声。

15 结合第二方面，在第一种可能的实现方式中，所述每个静音帧与一组加权系数相对应，其中在所述一组加权系数中，对应于第一组子带的加权系数大于对应于第二组子带的加权系数，其中所述第一组子带的感知重要性大于所述第二组子带的感知重要性。

结合第二方面或第二方面的第一种可能的实现方式，在第二种可能的实现
20 方式中，所述根据所述 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，包括：从所述 P 个静音帧中选择第一静音帧，使得在所述 P 个静音帧中所述第一静音帧的组加权谱距离最小；将所述第一静音帧的谱参数确定为所述第一谱参数。

结合第二方面或第二方面的第一种可能的实现方式，在第三种可能的实现方式中，所述根据所述 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，包括：从所述 P 个静音帧中选择至少一个静音帧，使得在所述 P 个静音帧中所述至少一个静音帧的组加权谱距离均小于第三阈值；根据所述至少一个静音帧的谱参数，确定所述第一谱参数。

结合第二方面或第二方面的第一种可能的实现方式至第三种可能的实现方式中任一实现方式，在第四种可能的实现方式中，所述 P 个静音帧包括所述当前输入静音帧以及所述当前输入静音帧之前的 $(P-1)$ 个静音帧。

结合第二方面的第四种可能的实现方式，在第五种可能的实现方式中，还包括：将当前输入静音帧编码为静音描述 SID 帧，其中所述 SID 帧包括所述第一谱参数。

第三方面，提供了一种信号处理方法，包括：将输入信号的频带划分为 R 个子带，其中 R 为正整数；在所述 R 个子带中的每个子带上，确定 S 个静音帧中每个静音帧的子带组谱距离，所述 S 个静音帧中每个静音帧的子带组谱距离为在所述每个子带上所述 S 个静音帧中所述每个静音帧与其它 $(S-1)$ 个静音帧之间的谱距离之和， S 为正整数；在所述每个子带上根据所述 S 个静音帧中每个静音帧的子带组谱距离，确定所述每个子带的第一谱参数，其中所述每个子带的第一谱参数用于生成舒适噪声。

结合第三方面，在第一种可能的实现方式中，所述在所述每个子带上，根据所述 S 个静音帧中每个静音帧的子带组谱距离，确定所述每个子带的第一谱参数，包括：在所述每个子带上，从所述 S 个静音帧中选择第一静音帧，使得在所述每个子带上所述 S 个静音帧中所述第一静音帧的子带组谱距离最小；在所述每个子带上，将所述第一静音帧的谱参数确定为所述每个子带的第一谱参

数。

结合第三方面，在第二种可能的实现方式中，所述在所述每个子带上，根据所述 S 个静音帧中每个静音帧的子带组谱距离，确定所述每个子带的第一谱参数，包括：在所述每个子带上，从所述 S 个静音帧中选择至少一个静音帧，
5 使得所述至少一个静音帧的子带组谱距离均小于第四阈值；在所述每个子带上，根据所述至少一个静音帧的谱参数，确定所述每个子带的第一谱参数。

结合第三方面或第三方面的第一种可能的实现方式或第二种可能的实现方式，在第三种可能的实现方式中，所述 S 个静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 (S-1) 个静音帧。

10 结合第三方面的第三种可能的实现方式，在第四种可能的实现方式中，还包括：将所述当前输入静音帧编码为静音描述 SID 帧，其中所述 SID 帧包括所述每个子带的第一谱参数。

第四方面，提供了一种信号处理方法，包括：确定 T 个静音帧中每个静音帧的第一参数，所述第一参数用于表征谱熵，T 为正整数；根据所述 T 个静音
15 帧中每个静音帧的第一参数，确定第一谱参数，其中所述第一谱参数用于生成舒适噪声。

结合第四方面，在第一种可能的实现方式中，所述根据所述 T 个静音帧中每个静音帧的第一参数，确定第一谱参数，包括：在确定能够按照聚类准则将所述 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，根据所述第一组
20 静音帧的谱参数，确定所述第一谱参数，其中所述第一组静音帧的第一参数所表征的谱熵均大于所述第二组静音帧的第一参数所表征的谱熵；在确定不能够按照聚类准则将所述 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，对所述 T 个静音帧的谱参数进行加权平均处理，以确定所述第一谱参数，其中

所述第一组静音帧的第一参数所表征的谱熵均大于所述第二组静音帧的第一参数所表征的谱熵。

结合第四方面的第一种可能的实现方式，在第二种可能的实现方式中，所述聚类准则包括：所述第一组静音帧中每个静音帧的第一参数与第一均值之间的距离小于或等于所述第一组静音帧中每个静音帧的第一参数与第二均值之间的距离；所述第二组静音帧中每个静音帧的第一参数与第二均值之间的距离小于或等于所述第二组静音帧中每个静音帧的第一参数与第一均值之间的距离；所述第一均值与第二均值之间的距离大于所述第一组静音帧的第一参数与第一均值之间的平均距离；所述第一均值与第二均值之间的距离大于所述第二组静音帧的第一参数与第二均值之间的平均距离；其中，所述第一均值为所述第一组静音帧的第一参数的平均值，所述第二均值为所述第二组静音帧的第一参数的平均值。

结合第四方面，在第三种可能的实现方式中，所述根据所述 T 个静音帧中每个静音帧的第一参数，确定第一谱参数，包括：

15 对所述 T 个静音帧的谱参数进行加权平均处理，以确定所述第一谱参数；其中，对于所述 T 个静音帧中任意不同的第 i 个静音帧和第 j 个静音帧，所述第 i 个静音帧对应的加权系数大于或等于所述 j 个静音帧对应的加权系数；在所述第一参数与所述谱熵正相关时，所述第 i 个静音帧的第一参数大于所述第 j 个静音帧的第一参数；在所述第一参数与所述谱熵负相关时，所述第 i 个静音帧的第一参数小于所述第 j 个静音帧的第一参数， i 和 j 均为正整数，且 $1 \leq i \leq T$ ， $1 \leq j \leq T$ 。

结合第四方面或第四方面的第一种可能的实现方式至第三种可能的实现方式中任一实现方式，在第四种可能的实现方式中，所述 T 个静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 $(T-1)$ 个静音帧

结合第四方面的第四种可能的实现方式，在第五种可能的实现方式中，还包括：将所述当前输入静音帧编码为静音描述 SID 帧，其中所述 SID 帧包括所述第一谱参数。

第五方面，提供了一种信号编码设备，包括：第一确定单元，用于在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在所述当前输入帧被编码为静音描述 SID 帧的情况下解码器根据所述当前输入帧生成的舒适噪声，并确定实际静音信号，其中所述当前输入帧为静音帧；第二确定单元，用于确定所述第一确定单元确定的所述舒适噪声与所述第一确定单元确定的所述实际静音信号的偏离程度；第三确定单元，用于根据所述第二确定单元确定的所述偏离程度，确定所述当前输入帧的编码方式，所述当前输入帧的编码方式包括拖尾帧编码方式或 SID 帧编码方式；编码单元，用于根据所述第三确定单元确定的所述当前输入帧的编码方式，对所述当前输入帧进行编码。

结合第五方面，在第一种可能的实现方式中，所述第一确定单元具体用于预测所述舒适噪声的特征参数，并确定所述实际静音信号的特征参数，其中所述舒适噪声的特征参数与所述实际静音信号的特征参数是一一对应的；所述第二确定单元具体用于确定所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离。

结合第五方面的第一种可能的实现方式，在第二种可能的实现方式中，所述第三确定单元具体用于：在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述 SID 帧编码方式，其中所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离与所述阈值集合中的阈值是一一对应的；在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离大于或等于

所述阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述拖尾帧编码方式。

结合第五方面的第一种可能的实现方式或第二种可能的实现方式，在第三种可能的实现方式中，所述第一确定单元具体用于：根据所述当前输入帧的前一帧的舒适噪声参数和所述当前输入帧的特征参数，预测所述舒适噪声的特征参数；或者，根据所述当前输入帧之前的 L 个拖尾帧的特征参数和所述当前输入帧的特征参数，预测所述舒适噪声的特征参数，其中 L 为正整数。

结合第五方面的第一种可能的实现方式或第二种可能的实现方式或第三种可能的实现方式，在第四种可能的实现方式中，所述第一确定单元具体用于：
10 确定所述当前输入帧的特征参数作为所述实际静音信号的特征参数；或者，对 M 个静音帧的特征参数进行统计处理，以确定所述实际静音信号的特征参数。

结合第五方面的第二种可能的实现方式，在第五种可能的实现方式中，所述舒适噪声的特征参数包括所述舒适噪声的码激励线性预测 CELP 激励能量和所述舒适噪声的线谱频率 LSF 系数，所述实际静音信号的特征参数包括所述实际静音信号的 CELP 激励能量和所述实际静音信号的 LSF 系数；所述第二确定单元具体用于确定所述舒适噪声的 CELP 激励能量与所述实际静音信号的 CELP 激励能量之间的距离 D_e ，并确定所述舒适噪声的 LSF 系数与所述实际静音信号的 LSF 系数之间的距离 D_{lsf} 。

结合第五方面的第五种可能的实现方式，在第六种可能的实现方式中，所述第三确定单元具体用于在所述距离 D_e 小于第一阈值，且所述距离 D_{lsf} 小于第二阈值的情况下，确定所述当前输入帧的编码方式为所述 SID 帧编码方式；所述第三确定单元具体用于在所述距离 D_e 大于或等于第一阈值，或者所述距离 D_{lsf} 大于或等于第二阈值的情况下，确定所述当前输入帧的编码方式为所述拖

尾帧编码方式。

结合第五方面的第六种可能的实现方式，在第七种可能的实现方式中，还包括：第四确定单元，用于：获取预设的所述第一阈值和预设的所述第二阈值；或者，根据所述当前输入帧之前的 N 个静音帧的 CELP 激励能量确定所述第一
5 阈值，并根据所述 N 个静音帧的 LSF 系数确定所述第二阈值，其中 N 为正整数。

结合第五方面或第五方面的第一种可能的实现方式至第七种可能的实现方式中任一实现方式，在第八种可能的实现方式中，所述第一确定单元具体用于采用第一预测方式，预测所述舒适噪声，其中所述第一预测方式与所述解码器生成所述舒适噪声的方式相同。

10 第六方面，提供了一种信号处理设备，包括：第一确定单元，用于确定 P 个静音帧中每个静音帧的组加权谱距离，其中所述 P 个静音帧中每个静音帧的组加权谱距离为所述 P 个静音帧中所述每个静音帧与其它 $(P-1)$ 个静音帧之间的加权谱距离之和， P 为正整数；第二确定单元，用于根据所述第一确定单元确定的所述 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，所述第
15 一谱参数用于生成舒适噪声。

结合第六方面，在第一种可能的实现方式中，所述第二确定单元具体用于：从所述 P 个静音帧中选择第一静音帧，使得在所述 P 个静音帧中所述第一静音帧的组加权谱距离最小；将所述第一静音帧的谱参数确定为所述第一谱参数。

结合第六方面，在第二种可能的实现方式中，所述第二确定单元具体用于：
20 从所述 P 个静音帧中选择至少一个静音帧，使得在所述 P 个静音帧中所述至少一个静音帧的组加权谱距离均小于第三阈值；根据所述至少一个静音帧的谱参数，确定所述第一谱参数。

结合第六方面或第六方面的第一种可能的实现方式或第二种可能的实现方

式，在第三种可能的实现方式中，所述 P 个静音帧包括所述当前输入静音帧以及所述当前输入静音帧之前的 $(P-1)$ 个静音帧；

所述设备还包括：编码单元，用于将当前输入静音帧编码为静音描述 SID 帧，其中所述 SID 帧包括所述第二确定单元确定的所述第一谱参数。

- 5 第七方面，提供了一种信号处理设备，包括：划分单元，用于将输入信号的频带划分为 R 个子带，其中 R 为正整数；第一确定单元，用于在所述划分单元划分的所述 R 个子带中每个子带上，确定 S 个静音帧中每个静音帧的子带组谱距离，所述 S 个静音帧中每个静音帧的子带组谱距离为在所述每个子带上所述 S 个静音帧中所述每个静音帧与其它 $(S-1)$ 个静音帧之间的谱距离之和， S
- 10 为正整数；第二确定单元，用于在所述划分单元划分的所述每个子带上根据所述第一确定单元确定的 S 个静音帧中每个静音帧的子带组谱距离，确定所述每个子带的第一谱参数，其中所述每个子带的第一谱参数用于生成舒适噪声。

结合第七方面，在第一种可能的实现方式中，所述第二确定单元具体用于：在所述每个子带上，从所述 S 个静音帧中选择第一静音帧，使得在所述每个子

15 带上的所述 S 个静音帧中所述第一静音帧的子带组谱距离最小；在所述每个子带上，将所述第一静音帧的谱参数确定为所述每个子带的第一谱参数。

结合第七方面，在第二种可能的实现方式中，所述第二确定单元具体用于：在所述每个子带上，从所述 S 个静音帧中选择至少一个静音帧，使得所述至少一个静音帧的子带组谱距离均小于第四阈值；在所述每个子带上，根据所述至

20 少一个静音帧的谱参数，确定所述每个子带的第一谱参数。

结合第七方面或第七方面的第一种可能的实现方式或第二种可能的实现方式，在第三种可能的实现方式中，所述 S 个静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 $(S-1)$ 个静音帧；

所述设备还包括：编码单元，用于将所述当前输入静音帧编码为静音描述SID帧，其中所述SID帧包括所述每个子带的谱参数。

- 第八方面，提供了一种信号处理设备，包括：第一确定单元，用于确定T个静音帧中每个静音帧的第一参数，所述第一参数用于表征谱熵，T为正整数；
- 5 第二确定单元，用于根据所述第一确定单元确定的所述T个静音帧中每个静音帧的第一参数，确定第一谱参数，其中所述第一谱参数用于生成舒适噪声。

- 结合第八方面，在第一种可能的实现方式中，所述第二确定单元具体用于：在确定能够按照聚类准则将所述T个静音帧分为所述第一组静音帧和所述第二组静音帧的情况下，根据所述第一组静音帧的谱参数，确定所述第一谱参数，
- 10 其中所述第一组静音帧的第一参数所表征的谱熵均大于所述第二组静音帧的第一参数所表征的谱熵；在确定不能够按照聚类准则将所述T个静音帧分为所述第一组静音帧和所述第二组静音帧的情况下，对所述T个静音帧的谱参数进行加权平均处理，以确定所述第一谱参数，其中所述第一组静音帧的第一参数所表征的谱熵均大于所述第二组静音帧的第一参数所表征的谱熵。

- 15 结合第八方面，在第二种可能的实现方式中，所述第二确定单元具体用于：对所述T个静音帧的谱参数进行加权平均处理，以确定所述第一谱参数；

- 其中，对于所述T个静音帧中任意不同的第i个静音帧和第j个静音帧，所述第i个静音帧对应的加权系数大于或等于所述j个静音帧对应的加权系数；在所述第一参数与所述谱熵正相关时，所述第i个静音帧的第一参数大于所述第j
- 20 个静音帧的第一参数；在所述第一参数与所述谱熵负相关时，所述第i个静音帧的第一参数小于所述第j个静音帧的第一参数，i和j均为正整数，且 $1 \leq i \leq T$ ， $1 \leq j \leq T$ 。

结合第八方面或第八方面的第一种可能的实现方式或第二种可能的实现方

式，在第三种可能的实现方式中，所述 T 个静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 (T-1) 个静音帧；

所述设备还包括：编码单元，用于将所述当前输入静音帧编码为静音描述 SID 帧，其中所述 SID 帧包括所述第一谱参数。

- 5 本发明实施例中，通过在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧生成的舒适噪声，并确定舒适噪声与实际静音信号的偏离程度，根据该偏离程度确定当前输入帧的编码方式为拖尾帧编码方式或 SID 帧编码方式，而非简单地根据统计得到的语音活动帧的数量将当前输入帧编码为拖尾帧，从而能够节省通信带宽。
- 10

附图说明

- 为了更清楚地说明本发明实施例的技术方案，下面将对本发明实施例中所需要使用的附图作简单地介绍，显而易见地，下面所描述的附图仅仅是本发明的一些实施例，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。
- 15

图 1 是根据本发明一个实施例的语音通信系统的示意框图。

图 2 是根据本发明实施例的信号编码方法的示意性流程图。

图 3a 是根据本发明一个实施例的信号编码方法的过程的示意性流程图。

- 20 图 3b 是根据本发明另一实施例的信号编码方法的过程的示意性流程图。

图 4 是根据本发明一个实施例的信号处理方法的示意性流程图。

图 5 是根据本发明另一实施例的信号处理方法的示意性流程图。

图 6 是根据本发明另一实施例的信号处理方法的示意性流程图。

图 7 是根据本发明一个实施例的信号编码设备的示意框图。

图 8 是根据本发明另一实施例的信号处理设备的示意框图。

图 9 是根据本发明另一实施例的信号处理设备的示意框图。

图 10 是根据本发明另一实施例的信号处理设备的示意框图。

5 图 11 是根据本发明另一实施例的信号编码设备的示意框图。

图 12 是根据本发明另一实施例的信号处理设备的示意框图。

图 13 是根据本发明另一实施例的信号处理设备的示意框图。

图 14 是根据本发明另一实施例的信号处理设备的示意框图。

具体实施方式

10 下面将结合本发明实施例中的附图，对本发明实施例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发明的一部分实施例，而不是全部实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例，都应属于本发明保护的范围。

图 1 是根据本发明一个实施例的语音通信系统的示意框图。

15 图 1 的系统 100 可以是 DTX 系统。系统 100 可以包括编码器 110 和解码器 120。

编码器 110 可以将输入的时域语音信号截断为语音帧，并对语音帧进行编码，然后将编码后的语音帧发送给解码器 120。解码器 120 可以从编码器 110 接收编码后的语音帧，并对编码后的语音帧进行解码，然后输出解码后的时域语音信号。

20

编码器 110 还可以包括语音活动性检测器（Voice Activity Detector, VAD）110a。VAD 110a 可以检测当前输入语音帧为语音活动帧还是静音帧。其中，语音活动帧可以表示含有通话语音信号的帧，静音帧可以表示不含有通话语音信

号的帧。此处，静音帧可以包括能量低于静音门限的无声帧，也可以包括背景噪声帧。编码器 110 可以有两种工作状态，即连续传输状态和非连续传输状态。当编码器 110 工作在连续传输状态时，编码器 110 可以对每个输入语音帧均进行编码并发送。当编码器 110 工作在非连续传输状态时，编码器 110 可以不对输入语音帧编码，或者可以将其编码为 SID 帧。通常，只有在输入语音帧为静音帧时，编码器 110 才会工作在非连续传输状态下。

如果当前输入的静音帧是语音活动段结束后的第一帧时，此处语音活动段包括可能存在的拖尾区间，那么编码器 110 可以将该静音帧编码为 SID 帧，此处可以用 SID_FIRST 表示该 SID 帧。如果当前输入的静音帧为上一个 SID 帧之后的第 n 帧，此处 n 为正整数，且与上一个 SID 帧之间没有语音活动帧时，那么编码器 110 可以将该静音帧编码为 SID 帧，此处可以用 SID_UPDATE 表示该 SID 帧。

SID 帧可以包括一些描述静音信号的特征的信息。解码器根据这些特征信息能够生成舒适噪声。例如 SID 帧可以包括静音信号的能量信息和谱信息。进一步地，例如，静音信号的能量信息可以包括码激励线性预测 (Code Excited Linear Prediction, CELP) 模型中激励信号的能量，或者静音信号的时域能量。谱信息可以包括线谱频率 (Line Spectral Frequency, LSF) 系数、线谱对 (Line Spectrum Pair, LSP) 系数、导抗谱频率 (Immittance Spectral Frequencies, ISF) 系数、导谱对 (Immittance Spectral Pairs, ISP) 系数、线性预测编码 (Linear Predictive Coding, LPC) 系数、快速傅立叶变换 (Fast Fourier Transform, FFT) 系数或修正离散余弦变换 (Modified Discrete Cosine Transform, MDCT) 系数等。

编码后的语音帧可以包括三种类型：语音编码帧、SID 帧和 NO_DATA 帧。其中语音编码帧为编码器 110 在连续传输状态下编码的帧，NO_DATA 帧可以表

示没有任何编码比特的帧，即物理上并不存在的帧，如 SID 帧之间的未编码静音帧等。

解码器 120 可以从编码器 110 接收编码后的语音帧，并对编码后的语音帧进行解码。当接收到语音编码帧时，解码器可以直接解码该帧并输出时域语音
5 帧。当接收到 SID 帧时，解码器可以解码 SID 帧，并获得 SID 帧中的拖尾长度、能量和谱信息。具体地，当 SID 帧为 SID_UPDATE 时，解码器可以根据当前 SID 帧中的信息，或者根据当前 SID 帧中的信息并结合其它信息，获得静音信号的能量信息和谱信息，也就是获得 CN 参数，从而根据 CN 参数生成时域 CN 帧。
当 SID 帧为 SID_FIRST 时，解码器根据 SID 帧中的拖尾长度信息获得该帧之前
10 m 帧中能量和谱的统计信息，并结合该 SID 帧中解码得到的信息获得 CN 参数，从而生成时域 CN 帧，其中 m 为正整数。当解码器的输入为 NO_DATA 帧时，解码器根据最近接收到的 SID 帧并结合其它信息，获得 CN 参数，从而生成时域 CN 帧。

图 2 是根据本发明实施例的信号编码方法的示意性流程图。图 2 的方法由
15 编码器执行，例如可以由图 1 中的编码器 110 执行。

210，在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧生成的舒适噪声，并确定实际静音信号，其中当前输入帧为静音帧。

本发明实施例中，实际静音信号可以是指输入编码器的实际静音信号。

20 220，确定舒适噪声与实际静音信号的偏离程度。

230，根据偏离程度，确定当前输入帧的编码方式，当前输入帧的编码方式包括拖尾帧编码方式或 SID 帧编码方式。

具体地，拖尾帧编码方式可以是指连续编码方式。编码器可以以连续编码

方式对处于拖尾区间的静音帧进行编码，编码得到的帧可以称为拖尾帧。

240，根据当前输入帧的编码方式，对当前输入帧进行编码。

在步骤 210 中，编码器可以根据不同的因素，确定以连续编码方式对当前输入帧的前一帧进行编码，例如，如果编码器中的 VAD 确定前一帧处于语音活动段或者编码器确定前一帧处于拖尾区间，那么编码器会以连续编码方式对前一帧进行编码。

由于输入语音信号进入静音段后，编码器可以根据实际情况决定工作在连续传输状态还是非连续传输状态。因此对于作为静音帧的当前输入帧而言，编码器需要确定如何编码当前输入帧。

10 当前输入帧可以是输入语音信号进入静音段后的第一个静音帧，也可以是输入语音信号进入静音段后的第 n 帧，此处 n 为大于 1 的正整数。

如果当前输入帧为第一个静音帧，那么在步骤 230 中，编码器确定当前输入帧的编码方式也就是确定是否需要设置拖尾区间，如果需要设置拖尾区间，则编码器可以将当前输入帧编码为拖尾帧；如果不需要设置拖尾区间，则编码器
15 器可以将当前输入帧编码为 SID 帧。

如果当前输入帧为第 n 个静音帧且编码器能够确定当前输入帧正处于拖尾区间，即当前输入帧前面的静音帧被连续地编码，那么在步骤 230 中，编码器确定当前输入帧的编码方式也就是确定是否结束拖尾区间。如果需要结束拖尾区间，则编码器可以将当前输入帧编码为 SID 帧；如果需要继续延长拖尾区间，
20 则编码器可以将当前输入帧编码为拖尾帧。

如果当前输入帧为第 n 个静音帧，并且也不存在拖尾机制，那么在步骤 230 中，编码器需要确定当前输入帧的编码方式，使得解码器对编码后的当前输入帧进行解码能够得到优质的舒适噪声信号。

可见，本发明实施例既可以应用于拖尾机制的触发场景，也可以应用于拖尾机制的执行场景，还可以应用于不存在拖尾机制的场景中。具体地，本发明实施例既可以确定是否触发拖尾机制，也可以确定是否提前结束拖尾机制。或者对于不存在拖尾机制的场景，本发明实施例可以确定静音帧的编码方式从而
5 达到更好的编码效果和解码效果。

具体地，编码器可以假设当前输入帧编码为 SID 帧，如果解码器接收到该 SID 帧，将根据 SID 帧生成舒适噪声，而编码器可以预测到该舒适噪声。然后，编码器可以估计该舒适噪声与输入编码器的实际静音信号的偏离程度。此处的偏离程度也可以理解为近似程度。如果预测到的舒适噪声与实际静音信号足够
10 接近，那么编码器可以认为无需设置拖尾区间或者无需继续延长拖尾区间。

在现有技术中，通过简单地统计语音活动帧的数量来确定是否执行固定长度的拖尾区间。也就是，如果有足够数量的语音活动帧被连续编码，那么就设置固定长度的拖尾区间。不论当前输入帧为第一个静音帧还是处于拖尾区间的第 n 个静音帧，当前输入帧均会被编码为拖尾帧。然而，不必要的拖尾帧会造成通信带宽的浪费。而本发明实施例中，通过根据预测的舒适噪声与实际静音信号的偏离程度确定当前输入帧的编码方式，而非简单地依据语音活动帧的数量来确定当前输入帧编码为拖尾帧，因此能够节省通信带宽。
15

本发明实施例中，通过在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧生成的舒适噪声，并确定舒适噪声与实际静音信号的偏离程度，根据该偏离程度确定当前输入帧的编码方式为拖尾帧编码方式或 SID 帧编码方式，而非简单地根据统计得到的语音活动帧的数量将当前输入帧编码为拖尾帧，从而能够节省通信带宽。
20

可选地，作为一个实施例，在步骤 210 中，编码器可以采用第一预测方式，预测舒适噪声，其中第一预测方式与解码器用于生成舒适噪声的方式相同。

具体地，编码器与解码器可以采用相同的方式确定舒适噪声。或者，编码器与解码器也可以分别采用不同的方式确定舒适噪声。本发明实施例对此不做
5 限定。

可选地，作为一个实施例，在步骤 210 中，编码器可以预测舒适噪声的特征参数，并确定实际静音信号的特征参数，其中舒适噪声的特征参数与实际静音信号的特征参数是一一对应的。在步骤 220 中，编码器可以确定舒适噪声的特征参数与实际静音信号的特征参数之间的距离。

10 具体地，编码器可以比较舒适噪声的特征参数与实际静音信号的特征参数之间的距离，从而确定舒适噪声与实际静音信号的偏离程度。舒适噪声的特征参数与实际静音信号的特征参数应当是一一对应的。也就是说，舒适噪声的特征参数的类型与实际静音信号的特征参数的类型是相同的。例如，编码器可以将舒适噪声的能量参数与实际静音信号的能量参数进行比较，也可以将舒适噪
15 声的谱参数与实际静音信号的谱参数进行比较。

本发明实施例中，当特征参数为标量时，特征参数之间的距离可以指特征参数之间的差值的绝对值，即标量距离。当特征参数为矢量时，特征参数之间的距离可以是指特征参数之间对应元素的标量距离的和。

可选地，作为另一实施例，在步骤 230 中，编码器可以在舒适噪声的特征
20 参数与实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下，确定当前输入帧的编码方式为 SID 帧编码方式，其中舒适噪声的特征参数与实际静音信号的特征参数之间的距离与阈值集合中的阈值是一一对应的。编码器也可以在舒适噪声的特征参数与实际静音信号的特征参数之间的距离大于

或等于阈值集合中对应阈值的情况下，确定当前输入帧的编码方式为拖尾帧编码方式。

具体地，舒适噪声的特征参数和实际静音信号的特征参数均可以包括至少一个参数，因此，舒适噪声的特征参数与实际静音信号的特征参数之间的距离也可以包括至少一种参数之间的距离。阈值集合也可以包括至少一个阈值。每种参数之间的距离可以对应于一个阈值。在确定当前输入帧的编码方式时，编码器可以分别将至少一种参数之间的距离与阈值集合中对应的阈值进行比较。阈值集合中的至少一个阈值可以是预先设定的，也可以是由编码器根据当前输入帧之前的多个静音帧的特征参数确定的。

10 如果舒适噪声的特征参数与实际静音信号的特征参数之间的距离小于阈值集合中对应阈值，编码器可以认为舒适噪声与实际静音信号足够接近，从而可以将当前输入帧编码为 SID 帧。如果舒适噪声的特征参数与实际静音信号的特征参数之间的距离大于或等于阈值集合中对应阈值，那么编码器可以认为舒适噪声与实际静音信号偏离较大，从而可以将当前输入帧编码为拖尾帧。

15 可选地，作为另一实施例，上述舒适噪声的特征参数可以用于表征以下至少一种信息：能量信息，谱信息。

可选地，作为另一实施例，上述能量信息可以包括 CELP 激励能量。上述谱信息可以包括以下至少一种：线性预测滤波器系数，FFT 系数，MDCT 系数。线性预测滤波器系数可以包括以下至少一种：LSF 系数，LSP 系数，ISF 系数，
20 ISP 系数，反射系数，LPC 系数。

可选地，作为另一实施例，在步骤 210 中，编码器可以确定当前输入帧的特征参数作为实际静音信号的特征参数。或者，编码器可以对 M 个静音帧的特征参数进行统计处理，以确定实际静音信号的特征参数。

可选地，作为另一实施例，上述 M 个静音帧可以包括当前输入帧以及当前输入帧之前的 $(M-1)$ 个静音帧， M 为正整数。

例如，如果当前输入帧为第一个静音帧，那么实际静音信号的特征参数可以是当前输入帧的特征参数；如果当前输入帧为第 n 个静音帧，那么实际静音信号的特征参数可以是编码器对包含当前输入帧在内的 M 个静音帧的特征参数进行统计处理得到的。 M 个静音帧可以是连续的，也可以是不连续的，本发明实施例对此不做限定。

可选地，作为另一实施例，在步骤 210 中，编码器可以根据当前输入帧的前一帧的舒适噪声参数和当前输入帧的特征参数，预测舒适噪声的特征参数。

10 或者，编码器可以根据当前输入帧之前的 L 个拖尾帧的特征参数和当前输入帧的特征参数，预测舒适噪声的特征参数， L 为正整数。

例如，如果当前输入帧为第一个静音帧，那么编码器可以根据前一帧的舒适噪声参数和当前输入帧的特征参数预测舒适噪声的特征参数。编码器对每一帧进行编码时，会在编码器内部保存每一帧的舒适噪声参数。通常只有在输入

15 帧为静音帧时，这个保存的舒适噪声参数才会较前一帧时发生变化，因为编码器可能会根据当前输入静音帧的特征参数对保存的舒适噪声参数进行更新，而在当前输入帧为语音活动帧时通常不对舒适噪声参数进行更新。因此，编码器可以获取内部存储的前一帧的舒适噪声参数。例如，舒适噪声参数可以包括静音信号的能量参数和谱参数。

20 此外，如果当前输入帧正处于拖尾区间，编码器可以根据当前输入帧之前的 L 个拖尾帧的参数进行统计，根据统计得到的结果和当前输入帧的特征参数，得到舒适噪声的特征参数。

可选地，作为另一实施例，舒适噪声的特征参数可以包括舒适噪声的 CELP

激励能量和舒适噪声的 LSF 系数，实际静音信号的特征参数可以包括实际静音信号的 CELP 激励能量和实际静音信号的 LSF 系数。在步骤 220 中，编码器可以确定舒适噪声的 CELP 激励能量与实际静音信号的 CELP 激励能量之间的距离 D_e ，并可以确定舒适噪声的 LSF 系数与实际静音信号的 LSF 系数之间的距离 D_{lsf} 。

应注意，此处距离 D_e 和距离 D_{lsf} 可以包含一个变量，也可以包含一组变量。例如，距离 D_{lsf} 可以包含两个变量，一个可以是平均的 LSF 系数的距离，即每个对应 LSF 系数的距离的均值。另一个可以是 LSF 系数间的最大距离，即距离最大的那对 LSF 系数之间的距离。

10 可选地，作为另一实施例，在步骤 230 中，在距离 D_e 小于第一阈值，且距离 D_{lsf} 小于第二阈值的情况下，编码器可以确定当前输入帧的编码方式为 SID 帧编码方式。在距离 D_e 大于或等于第一阈值，或者距离 D_{lsf} 大于或等于第二阈值的情况下，编码器可以确定当前输入帧的编码方式为拖尾帧编码方式。其中，第一阈值和第二阈值均属于上述阈值集合。

15 可选地，作为另一实施例，当 D_e 或 D_{lsf} 包含一组变量时，编码器将一组变量中的每个变量与其相对应的阈值做比较，从而确定以何种方式编码当前输入帧。

具体地，编码器可以根据距离 D_e 和距离 D_{lsf} ，确定当前输入帧的编码方式。如果距离 $D_e < \text{第一阈值}$ ，并且距离 $D_{lsf} < \text{第二阈值}$ ，则可以表明预测的舒适噪声
20 的 CELP 激励能量和 LSF 系数与实际静音信号的 CELP 激励能量和 LSF 系数差别都不大，则编码器可以认为舒适噪声和实际静音信号足够接近，可以将当前输入帧编码为 SID 帧。否则，可以将当前输入帧编码为拖尾帧。

可选地，作为另一实施例，在步骤 230 中，编码器可以获取预设的第一阈

值和预设的第二阈值。或者，编码器可以根据当前输入帧之前的 N 个静音帧的 CELP 激励能量确定第一阈值，并根据 N 个静音帧的 LSF 系数确定第二阈值，其中 N 为正整数。

具体地，第一阈值和第二阈值均可以是预设的固定值。或者，第一阈值和
5 第二阈值均可以是自适应的变量。例如，第一阈值可以是编码器对当前输入帧之前的 N 个静音帧的 CELP 激励能量统计得到的。第二阈值可以是编码器对当前输入帧之前的 N 个静音帧的 LSF 系数统计得到的。 N 个静音帧可以是连续的，也可以是不连续的。

下面将结合具体例子详细描述上述图 2 的具体过程。在下面图 3a 和图 3b
10 的例子中，将以本发明实施例可应用的两个场景来进行描述。应理解，这些例子只是为了帮助本领域技术人员更好地理解本发明实施例，而非限制本发明实施例的范围。

图 3a 是根据本发明一个实施例的信号编码方法的过程的示意性流程图。在图 3a 中，假设当前输入帧的前一帧的编码方式为连续编码方式，编码器内部的
15 VAD 确定当前输入帧为输入语音信号进入静音段后的第一个静音帧。那么，编码器将需要确定是否设置拖尾区间，也就是需要确定是将当前输入帧编码为拖尾帧还是 SID 帧。下面将详细描述该过程。

301a, 确定实际静音信号的 CELP 激励能量和 LSF 系数。

具体地，编码器可以将当前输入帧的 CELP 激励能量 e 作为实际静音信号的
20 CELP 激励能量 e_{SI} ，可以将当前输入帧的 LSF 系数 $lsf(i)$ 作为实际静音信号的 LSF 系数 $lsf_{SI}(i)$ ， $i=0, 1, \dots, K-1$ ， K 为滤波器阶数。编码器可以参照现有技术，确定当前输入帧的 CELP 激励能量以及 LSF 系数。

302a, 预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧

生成的舒适噪声的 CELP 激励能量和 LSF 参数。

编码器可以假设当前输入帧编码为 SID 帧，那么解码器将根据该 SID 帧生成舒适噪声。对于编码器而言，其能够预测该舒适噪声的 CELP 激励能量 e_{CN} 和 LSF 系数 $lsf_{CN}(i)$, $i=0, 1, \dots, K-1$, K 为滤波器阶数。编码器可以根据编码器
5 内部存储的前一帧的舒适噪声参数和当前输入帧的 CELP 激励能量和 LSF 系数，分别确定舒适噪声的 CELP 激励能量和 LSF 系数。

例如，编码器可以按照等式 (1) 预测舒适噪声的 CELP 激励能量 e_{CN} :

$$e_{CN} = 0.4 * e_{CN}^{[-1]} + 0.6 * e \quad (1)$$

其中， $e_{CN}[-1]$ 可以表示前一帧的 CELP 激励能量， e 可以表示当前输入帧
10 的 CELP 激励能量。

编码器可以按照等式 (2) 预测舒适噪声的 LSF 系数 $lsf_{CN}(i)$, $i=0, 1, \dots, K-1$, K 为滤波器阶数。

$$lsf_{CN}(i) = 0.4 * lsf_{CN}^{[-1]}(i) + 0.6 * lsf(i) \quad (2)$$

其中， $lsf_{CN}[-1](i)$ 可以表示前一帧的 LSF 系数， $lsf(i)$ 可以表示当前输入帧
15 的第 i 个 LSF 系数。

303a, 确定舒适噪声的 CELP 激励能量与实际静音信号的 CELP 激励能量之间的距离 De , 并确定舒适噪声的 LSF 系数与实际静音信号的 LSF 系数之间的距离 $Dlsf$ 。

具体地，编码器可以根据等式 (3) 确定舒适噪声的 CELP 激励能量与实际
20 静音信号的 CELP 激励能量之间的距离 De :

$$De = |\log_2 e_{CN} - \log_2 e| \quad (3)$$

编码器可以根据等式 (4) 确定舒适噪声的 LSF 系数与实际静音信号的 LSF 系数之间的距离 $Dlsf$:

$$Dlsf = \sum_{i=0}^{K-1} |lsfCN(i) - lsf(i)| \quad (4)$$

304a, 确定距离 De 是否小于第一阈值, 并且距离 Dlsf 是否小于第二阈值。

具体地, 第一阈值和第二阈值均可以是预设的固定值。

或者, 第一阈值和第二阈值可以是自适应的变量。编码器可以根据当前输

5 入帧之前的 N 个静音帧的 CELP 激励能量确定第一阈值, 例如, 编码器可以按照等式 (5) 确定第一阈值 thr1:

$$thr1 = \frac{\sum_{n=0}^{N-1} (\log_2 e^n - \log_2 \frac{1}{N} \sum_{m=0}^{N-1} e^{[m]})}{N} \quad (5)$$

编码器可以根据 N 个静音帧的 LSF 系数确定第二阈值, 例如, 编码器可以按照等式 (6) 确定第二阈值 thr2:

$$thr2 = \frac{\sum_{n=0}^{N-1} \sum_{i=0}^{K-1} (lsf^{[n]}(i) - \frac{1}{N} \sum_{p=0}^{N-1} lsf^{[p]}(i))}{N} \quad (6)$$

其中, 在等式 (5) 和等式 (6) 中, [x] 可以表示第 x 帧, x 可以为 n、m 或 p。例如, $e[m]$ 可以表示第 m 帧的 CELP 激励能量。 $lsf^{[n]}(i)$ 可以表示第 n 帧的第 i 个 LSF 系数, $lsf^{[p]}(i)$ 可以表示第 p 帧的第 i 个 LSF 系数。

305a, 如果距离 De 小于第一阈值并且距离 Dlsf 小于第二阈值, 则确定不设置拖尾区间, 将当前输入帧编码为 SID 帧。

如果距离 De 小于第一阈值并且距离 Dlsf 小于第二阈值, 则编码器可以认为解码器能够生成的舒适噪声与实际的静音信号足够接近, 那么可以不设置拖尾区间, 那么将当前输入帧编码为 SID 帧。

306a, 如果距离 De 大于或等于第一阈值, 或者距离 Dlsf 大于或等于第二阈值, 则确定设置拖尾区间, 将当前输入帧编码为拖尾帧。

本发明实施例中, 通过在根据在当前输入帧被编码为 SID 帧的情况下解码

器根据当前输入帧生成的舒适噪声与实际静音信号的偏离程度，确定当前输入帧的编码方式为拖尾帧编码方式或 SID 帧编码方式，而非简单地根据统计得到的语音活动帧的数量将当前输入帧编码为拖尾帧，从而能够节省通信带宽。

图 3b 是根据本发明另一实施例的信号编码方法的过程的示意性流程图。在

5 图 3b 中，假设当前输入帧已处于拖尾区间。那么，编码器需要确定是否结束拖尾区间，也就是需要确定是将当前输入帧编码继续为拖尾帧还是编码为 SID 帧。下面将详细描述该过程。

301b，确定实际静音信号的 CELP 激励能量和 LSF 系数。

可选地，类似于步骤 301a，编码器可以将当前输入帧的 CELP 激励能量和
10 LSF 系数作为实际静音信号的 CELP 激励能量和 LSF 系数。

可选地，编码器可以对包括当前输入帧在内的 M 个静音帧的 CELP 激励能量进行统计处理，得到实际静音信号的 CELP 激励能量。其中， $M \leq$ 拖尾区间内当前输入帧之前的拖尾帧的数目。

例如，编码器可以按照等式 (7) 确定实际静音信号的 CELP 激励能量 eSI:

$$15 \quad eSI = \log_2 \left(\frac{1}{\sum_{j=0}^M w(j)} \cdot \sum_{j=0}^M w(j) \cdot e^{[-j]} \right) \quad (7)$$

再例如，编码器可以按照等式 (8) 确定实际静音信号的 LSF 系数 lsfSI(i), $i=0, 1, \dots, K-1$, K 为滤波器阶数。

$$lsfSI(i) = \frac{1}{\sum_{j=0}^M w(j)} \cdot \sum_{j=0}^M w(j) \cdot lsf(i)^{[-j]} \quad (8)$$

其中，在上述等式 (7) 和等式 (8) 中，w(j) 可以表示加权系数， $e^{[-j]}$ 可以
20 表示当前输入帧之前的第 j 个静音帧的 CELP 激励能量。

302b，预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧

生成的舒适噪声的 CELP 激励能量和 LSF 系数。

具体地, 编码器可以根据当前输入帧之前的 L 个拖尾帧的 CELP 激励能量和 LSF 系数, 分别确定舒适噪声的 CELP 激励能量 eCN 和 LSF 系数 lsfCN(i), i=0, 1, ..., K-1, K 为滤波器阶数。

5 例如, 编码器可以按照等式 (9) 确定舒适噪声的 CELP 激励能量 eCN:

$$eCN = 0.4 * \left(\frac{1}{\sum_{j=0}^L w(j)} \cdot \sum_{j=0}^L w(j) \cdot eHO^{[-j]} \right) + 0.6 * e \quad (9)$$

其中, eHO[-j] 可以表示当前输入帧之前的第 j 个拖尾帧的激励能量。

再例如, 编码器可以按照等式 (10) 确定舒适噪声的 LSF 系数 lsfCN(i), i=0, 1, ..., K-1, K 为滤波器阶数。

$$lsfCN(i) = 0.4 * \left(\frac{1}{\sum_{j=1}^L w(j)} \cdot \sum_{j=1}^L w(j) \cdot lsfHO(i)^{[-j]} \right) + 0.6 * lsf(i) \quad (10)$$

其中, lsfHO(i)[-j] 可以表示当前输入帧之前的第 j 个拖尾帧的第 i 个 lsf 系数。

在等式 (9) 和 (10) 中, w(j) 可以表示加权系数。

303b, 确定舒适噪声的 CELP 激励能量与实际静音信号的 CELP 激励能量之间的距离 De, 并确定舒适噪声的 LSF 系数与实际静音信号的 LSF 系数之间的距离 Dlsf。
15 离 Dlsf。

例如, 编码器可以按照等式 (3) 确定舒适噪声的 CELP 激励能量与实际静音信号的 CELP 激励能量之间的距离 De。编码器可以根据等式 (4) 确定舒适噪声的 LSF 系数与实际静音信号的 LSF 系数之间的距离 Dlsf。

304b, 确定距离 De 是否小于第一阈值, 并且距离 Dlsf 是否小于第二阈值。

20 具体地, 第一阈值和第二阈值均可以是预设的固定值。

或者，第一阈值和第二阈值可以是自适应的变量。例如，编码器可以按照等式（5）确定第一阈值 thr1 ，可以按照等式（6）确定第二阈值 thr2 。

305b，如果距离 D_e 小于第一阈值并且距离 D_{lsf} 小于第二阈值，则确定结束拖尾区间，将当前输入帧编码为 SID 帧。

5 306b，如果距离 D_e 大于或等于第一阈值，或者距离 D_{lsf} 大于或等于第二阈值，则确定继续延长拖尾区间，将当前输入帧编码为拖尾帧。

本发明实施例中，通过根据在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧生成的舒适噪声与实际静音信号的偏离程度，确定当前输入帧的编码方式为拖尾帧编码方式或 SID 帧编码方式，而非简单地根据统计得到的
10 语音活动帧的数量将当前输入帧编码为拖尾帧，从而能够节省通信带宽。

由上述可知，在编码器进入非连续传输状态后，会间歇地编码 SID 帧。SID 帧通常包括一些描述静音信号的能量和频谱信息等。解码器从编码器接收到 SID 帧后，会根据 SID 帧中的信息生成舒适噪声。目前，由于 SID 帧是每隔若干帧才编码和发送一次，因此在编码 SID 帧时，SID 帧的信息通常都是编码器对当
15 前输入静音帧及其之前的若干静音帧统计得到的。例如，在一段连续的静音区间内，当前编码的 SID 帧的信息通常是在当前 SID 帧以及当前 SID 帧与上一 SID 帧之间的多个静音帧中统计得到的。又例如，在一段语音活动段之后的第一个 SID 帧的编码信息通常是编码器对当前输入静音帧及与其相邻的语音活动段末尾的若干拖尾帧统计得到的，也就是对位于拖尾区间内的静音帧进行统计得到
20 的。为了便于描述，将用于统计 SID 帧编码参数的多个静音帧称为分析区间。

具体地，在编码 SID 帧时，SID 帧的参数都是对分析区间的多个静音帧的参数取平均或取中值得到的。然而，实际的背景噪声频谱会夹杂各种突发的瞬态的频谱成份。一旦分析区间内包含了这样的频谱成份，求均值的方法会把这些成

份也混入 SID 帧中，取中值的方法甚至有可能错误地将含有这类频谱成份的静音谱编码入 SID 帧中，从而造成解码端根据 SID 帧生成的舒适噪声的质量下降。

图 4 是根据本发明一个实施例的信号处理方法的示意性流程图。图 4 的方法由编码器或解码器执行，例如可以由图 1 中的编码器 110 或解码器 120 执行。

- 5 410，确定 P 个静音帧中每个静音帧的组加权谱距离（Group Weighted Spectral Distance），其中 P 个静音帧中每个静音帧的组加权谱距离为 P 个静音帧中每个静音帧与其它（P-1）个静音帧之间的加权谱距离之和，P 为正整数。

例如，编码器或解码器可以将当前输入静音帧之前的多个静音帧的参数存储在某个缓存中。该缓存的长度可以是固定的或变化的。上述 P 个静音帧可以
10 是由编码器或解码器从该缓存中选择的。

420，根据 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，第一谱参数用于生成舒适噪声。

本发明实施例中，通过根据 P 个静音帧中每个静音帧的组加权谱距离确定用于生成舒适噪声的第一谱参数，而非简单地对多个静音帧的谱参数取平均或
15 取中值得到用于生成舒适噪声的谱参数，从而能够提升舒适噪声的质量。

可选地，作为一个实施例，在步骤 410 中，可以根据 P 个静音帧中每个静音帧的谱参数，确定每个静音帧的组加权谱距离。例如，可以按照等式（11）确定 P 个静音帧中的第 x 帧的组加权谱距离 $swd[x]$ ，

$$swd^{[x]} = \sum_{j=0, j \neq x}^{P-1} \sum_{i=0}^{K-1} w(i) [U^{[x]}(i) - U^{[j]}(i)] \quad (11)$$

- 20 其中， $U^{[x]}(i)$ 可以表示第 x 帧的第 i 个谱参数， $U^{[j]}(i)$ 可以表示第 j 帧的第 i 个谱参数， $w(i)$ 可以为加权系数，K 为谱参数的系数数目。

例如，上述每个静音帧的谱参数可以包括 LSF 系数、LSP 系数、ISF 系数、ISP 系数、LPC 系数、反射系数、FFT 系数或 MDCT 系数等。因此，相应地，

在步骤 420 中, 第一谱参数可以包括 LSF 系数、LSP 系数、ISF 系数、ISP 系数、LPC 系数、反射系数、FFT 系数或 MDCT 系数等。

下面以谱参数为 LSF 系数为例说明步骤 420 的过程。例如, 可以确定每个静音帧的 LSF 系数与其它 (P-1) 个静音帧的 LSF 系数之间的加权谱距离之和, 5 即每个静音帧的 LSF 系数的组加权谱距离 swd, 比如可以按照等式 (12) 确定这 P 个静音帧中第 x 帧 LSF 系数的组加权谱距离 $\text{swd}^{[x]}$, 其中 $x=0, 1, 2, \dots, P-1$:

$$\text{swd}^{[x]} = \sum_{j=0, j \neq x}^{P-1} \sum_{i=0}^{K'-1} w'(i) [\text{lsf}^{[x]}(i) - \text{lsf}^{[j]}(i)] \quad (12)$$

其中, $w'(i)$ 为加权系数, K' 为滤波器阶数。

可选地, 作为一个实施例, 每个静音帧可以与一组加权系数相对应, 其中 10 在这一组加权系数中, 对应于第一组子带的加权系数大于对应于第二组子带的加权系数, 其中第一组子带的感知重要性大于第二组子带的感知重要性。

子带可以是基于对频谱系数的划分得到的, 具体过程可以参照现有技术。子带的感知重要性可以按照现有技术确定。通常, 低频子带的感知重要性大于高频子带的感知重要性, 因此在一个简化的实施例中, 低频子带的加权系数可 15 以大于高频子带的加权系数。

例如, 在等式 (12) 中, $w'(i)$ 为加权系数, $i=0, 1, \dots, K'-1$ 。每个静音帧对应于一组加权系数, 即 $w'(0)$ 至 $w'(K'-1)$ 。在这组加权系数中, 低频子带的 lsf 系数的加权系数大于高频子带的 lsf 系数的加权系数。由于通常背景噪声的能量更多地集中在低频带, 因此, 解码器生成的舒适噪声的质量更多地是由低频带 20 的信号的质量决定的。因此, 高频带的 lsf 系数的谱距离对最终加权谱距离的影响应当适当减弱。

可选地, 作为另一实施例, 在步骤 420 中, 可以从 P 个静音帧中选择第一静音帧, 使得在 P 个静音帧中第一静音帧的组加权谱距离最小, 并可以将第一

静音帧的谱参数确定为第一谱参数。

具体地，组加权谱距离最小，可以表明第一静音帧的谱参数最能表征这 P 个静音帧谱参数的共性。因此，可以将第一静音帧的谱参数编码入 SID 帧。例如，对于每个静音帧的 LSF 系数的组加权谱距离，第一静音帧的 LSF 系数的组
5 加权谱距离最小，那么可以表明第一静音帧的 LSF 谱是最能够表征这 P 个静音帧的 LSF 谱的共性的 LSF 谱。

可选地，作为另一实施例，在步骤 420 中，可以从 P 个静音帧中选择至少一个静音帧，使得在 P 个静音帧中至少一个静音帧的组加权谱距离均小于第三阈值，然后可以根据至少一个静音帧的谱参数，确定第一谱参数。

10 例如，在一个实施例中，可以将至少一个静音帧的谱参数的均值确定为第一谱参数。在另一个实施例中，可以将至少一个静音帧的谱参数的中值确定为第一谱参数。在另一个实施例中，也可以使用本发明实施例中的其它方法根据上述至少一个静音帧的谱参数确定第一谱参数。

下面仍以谱参数为 LSF 系数为例进行说明，那么第一谱参数可以为第一 LSF
15 系数。例如，可以按照等式 (12) 得到 P 个静音帧中每个静音帧的 LSF 系数的组加权谱距离。从 P 个静音帧中选择 LSF 系数的组加权谱距离小于第三阈值的至少一个静音帧。然后可以将至少一个静音帧的 LSF 系数的均值作为第一 LSF 系数。例如，可以按照等式 (13) 确定第一 LSF 系数 $lsfSID(i)$, $i=0, 1, \dots, K'-1$, K' 为滤波器阶数。

$$lsfSID(i) = \frac{1}{\sum_{j=0, j \neq \{A\}}^{P-1} 1} \cdot \sum_{j=0, j \neq \{A\}}^{P-1} lsf^{[j]}(i) \quad (13)$$

20

其中， $\{A\}$ 可以表示 P 个静音帧中除了上述至少一个静音帧之外的静音帧。
 $lsf^{[j]}(i)$ 可以表示第 j 帧的第 i 个 LSF 系数。

此外，上述第三阈值可以是预先设定的。

可选地，作为另一实施例，在图 4 的方法由编码器执行时，上述 P 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 (P-1) 个静音帧。

在图 4 的方法由解码器执行时，上述 P 个静音帧可以为 P 个拖尾帧。

5 可选地，作为另一实施例，在图 4 的方法由编码器执行时，编码器可以将当前输入静音帧编码为 SID 帧，其中 SID 帧包括第一谱参数。

本发明实施例中，编码器可以将当前输入帧编码为 SID 帧，使得 SID 帧中包括第一谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到 SID 帧中的谱参数，从而能够提升解码器根据该 SID 帧生成的舒适噪声的质量。

10 图 5 是根据本发明另一实施例的信号处理方法的示意性流程图。图 5 的方法由编码器或解码器执行，例如可以由图 1 中的编码器 110 或解码器 120 执行。

510，将输入信号的频带划分为 R 个子带，其中 R 为正整数。

520，在 R 个子带中的每个子带上，确定 S 个静音帧中每个静音帧的子带组谱距离，S 个静音帧中每个静音帧的子带组谱距离为在每个子带上 S 个静音帧中
15 每个静音帧与其它 (S-1) 个静音帧之间的谱距离之和，S 为正整数。

530，在每个子带上，根据 S 个静音帧中每个静音帧的子带组谱距离，确定每个子带的第一谱参数，每个子带的第一谱参数用于生成舒适噪声。

本发明实施例中，通过在 R 个子带中每个子带上根据 S 个静音帧中每个静音帧的子带组谱距离确定用于生成舒适噪声的每个子带的第一谱参数，而非简
20 单地对多个静音帧的谱参数取平均或取中值得到用于生成舒适噪声的谱参数，从而能够提升舒适噪声的质量。

在步骤 530 中，对于每个子带，可以根据 S 个静音帧每个静音帧的谱参数，确定每一个子带上的每个静音帧的子带组谱距离。可选地，作为一个实施例，

可以按照等式 (14) 确定第 k 个子带上第 y 个静音帧的子带组谱距离 $\text{ssd}_k^{[y]}$, 其中, $k=1,2,\dots,R$, $y=0, 1, \dots, S-1$ 。

$$\text{ssd}_k^{[y]} = \sum_{j=0, j \neq y}^{S-1} \sum_{i=0}^{L(k)-1} [U_k^{[y]}(i) - U_k^{[j]}(i)] \quad (14)$$

其中, $L(k)$ 可以表示第 k 个子带所包括的谱参数的系数数目, $U_k^{[y]}(i)$ 可以表示第 k 个子带上第 y 个静音帧的谱参数的第 i 个系数, $U_k^{[j]}(i)$ 可以表示第 k 个子带上第 j 个静音帧的谱参数的第 i 个系数。

例如, 上述每个静音帧的谱参数可以包括 LSF 系数、LSP 系数、ISF 系数、ISP 系数、LCP 系数、反射系数、FFT 系数或 MDCT 系数等。

下面以谱参数为 LSF 系数为例进行说明。例如, 可以确定每个静音帧的 LSF 系数的子带组谱距离。每个子带可以包括一个 LSF 系数, 也可以包括多个 LSF 系数。例如, 可以按照等式 (15) 确定第 k 个子带上第 y 个静音帧的 LSF 系数的子带组谱距离 $\text{ssd}_k^{[y]}$, 其中, $k=1,2,\dots,R$, $y=0, 1, \dots, S-1$ 。

$$\text{ssd}_k^{[y]} = \sum_{j=0, j \neq k}^{S-1} \sum_{i=0}^{L(k)-1} [\text{lsf}_k^{[y]}(i) - \text{lsf}_k^{[j]}(i)] \quad (15)$$

其中, $L(k)$ 可以表示第 k 个子带所包括的 LSF 系数的数目。 $\text{lsf}_k^{[y]}(i)$ 可以表示第 k 个子带上第 y 个静音帧的第 i 个 LSF 系数, $\text{lsf}_k^{[j]}(i)$ 可以表示第 k 个子带上第 j 个静音帧的第 i 个 LSF 系数。

相应地, 每个子带的第一谱参数也可以包括 LSF 系数、LSP 系数、ISF 系数、ISP 系数、LCP 系数、反射系数、FFT 系数或 MDCT 系数等。

可选地, 作为另一实施例, 在步骤 530 中, 可以在每个子带上, 从 S 个静音帧中选择第一静音帧, 使得在每个子带上 S 个静音帧中第一静音帧的子带组谱距离最小。然后可以在每个子带上, 将第一静音帧的谱参数作为每个子带的第一谱参数。

具体地，编码器可以确定每个子带上的第一静音帧，将该第一静音帧的谱参数作为该子带的第一谱参数。

下面仍以谱参数为 LSF 系数为例进行说明，相应地，每个子带的第一谱参数为每个子带的第一 LSF 系数。例如，可以按照等式 (15) 确定每个子带上的各个静音帧的 LSF 系数的子带组谱距离。对于每个子带，可以选择子带组谱距离最小的帧的 LSF 系数作为该子带的第一 LSF 系数。

可选地，作为另一实施例，在步骤 530 中，可以在每个子带上，从 S 个静音帧中选择至少一个静音帧，使得至少一个静音帧的子带组谱距离均小于第四阈值。然后可以在每个子带上，根据至少一个静音帧的谱参数，确定每个子带的第一谱参数。

例如，在一个实施例中，可以将每个子带上的 S 个静音帧中的至少一个静音帧的谱参数的均值确定为每个子带的第一谱参数。在另一个实施例中，可以将每个子带上的 S 个静音帧中的至少一个静音帧的谱参数的中值确定为每个子带的第一谱参数。在另一个实施例中也可以使用本发明中的其它方法根据上述至少一个静音帧的谱参数确定每个子带的第一谱参数。

以 LSF 系数为例，可以按照等式 (15) 确定每个子带上的各个静音帧的 LSF 系数的子带组谱距离。对于每个子带，可以选择子带组谱距离均小于第四阈值的至少一个静音帧，将至少一个静音帧的 LSF 系数的均值确定为该子带的第一 LSF 系数。上述第四阈值可以是预先设定的。

可选地，作为另一实施例，在图 5 的方法由编码器执行时，上述 S 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 (S-1) 个静音帧。

在图 5 的方法由解码器执行时，上述 S 个静音帧可以是 S 个拖尾帧。

可选地，作为另一实施例，在图 5 的方法由编码器执行时，编码器可以将

当前输入静音帧编码为 SID 帧，其中 SID 帧包括每个子带的第一谱参数。

本发明实施例中，编码器可以在编码 SID 帧时，使 SID 帧包括各个子带的第一谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到 SID 帧中的谱参数，从而能够提升解码器根据该 SID 帧生成的舒适噪声的质量。

5 图 6 是根据本发明另一实施例的信号处理方法的示意性流程图。图 6 的方法由编码器或解码器执行，例如可以由图 1 中的编码器 110 或解码器 120 执行。

610，确定 T 个静音帧中每个静音帧的第一参数，第一参数用于表征谱熵，T 为正整数。

例如，在静音帧的谱熵能够直接确定时，第一参数可以为谱熵。某些情况
10 下，遵循严格定义的谱熵不一定能被直接确定，此时，第一参数可以为能够表征谱熵的其它参数，例如能够反映频谱结构性强弱的参数等。

例如，可以根据每个静音帧的 LSF 系数确定每个静音帧的第一参数。比如，可以按照等式 (16) 确定第 z 个静音帧的第一参数，其中 $z=1, 2, \dots, T$ 。

$$C^{[z]} = \sum_{i=0}^{K-2} \left[\text{lsf}(i+1) - \text{lsf}(i) - \frac{1}{K-1} \sum_{j=0}^{K-2} [\text{lsf}(j+1) - \text{lsf}(j)] \right]^2 \quad (16)$$

15 其中，K 为滤波器阶数。

此处，C 是能够反映频谱结构性强弱的参数，并不严格遵循谱熵的定义，C 越大，可以表示谱熵越小。

620，根据 T 个静音帧中每个静音帧的第一参数，确定第一谱参数，第一谱参数用于生成舒适噪声。

20 本发明实施例中，通过根据 T 个静音帧的用于表征谱熵的第一参数确定用于生成舒适噪声的第一谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到用于生成舒适噪声的谱参数，从而能够提升舒适噪声的质量。

可选地，作为一个实施例，可以在确定能够按照聚类准则将 T 个静音帧分

为第一组静音帧和第二组静音帧的情况下，可以根据第一组静音帧的谱参数，确定第一谱参数，其中第一组静音帧的第一参数所表征的谱熵均大于第二组静音帧的第一参数所表征的谱熵。在确定不能够按照聚类准则将 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，可以对 T 个静音帧的谱参数进行加权平均处理，以确定第一谱参数，其中第一组静音帧的第一参数所表征的谱熵均大于第二组静音帧的第一参数所表征的谱熵。

一般而言，普通噪声谱的结构性相对较弱，而非噪声信号谱或包含有瞬态成份的噪声谱的结构性相对较强。谱的结构性强弱直接对应谱熵的大小。相对而言，普通噪声的谱熵会较大，而非噪声信号或含有瞬态成份的噪声的谱熵会较小。因此，在 T 个静音帧能够被分为第一组静音帧和第二组静音帧的情况下，编码器可以根据静音帧的谱熵，选择不包含瞬态成份的第一组静音帧的谱参数来确定第一谱参数。

例如，在一个实施例中可以将第一组静音帧的谱参数的均值确定为第一谱参数。在另一个实施例中，可以将第一组静音帧的谱参数的中值确定为第一谱参数。在另一个实施例中，也可以使用本发明中的其它方法根据上述第一组静音帧的谱参数确定第一谱参数。

如果 T 个静音帧不能被分为第一组静音帧和第二组静音帧，那么可以对 T 个静音帧的谱参数进行加权平均处理来得到第一谱参数。可选地，作为另一实施例，上述聚类准则可以包括：第一组静音帧中每个静音帧的第一参数与第一均值之间的距离小于或等于第一组静音帧中每个静音帧的第一参数与第二均值之间的距离；第二组静音帧中每个静音帧的第一参数与第二均值之间的距离小于或等于第二组静音帧中每个静音帧的第一参数与第一均值之间的距离；第一均值与第二均值之间的距离大于第一组静音帧的第一参数与第一均值之间的平

均距离；第一均值与第二均值之间的距离大于第二组静音帧的第一参数与第二均值之间的平均距离。

其中，第一均值为第一组静音帧的第一参数的平均值，第二均值为第二组静音帧的第一参数的平均值。

- 5 可选地，作为另一实施例，编码器可以对 T 个静音帧的谱参数进行加权平均处理，以确定第一谱参数；其中，对于 T 个静音帧中任意不同的第 i 个静音帧和第 j 个静音帧，第 i 个静音帧对应的加权系数大于或等于 j 个静音帧对应的加权系数；在第一参数与谱熵正相关时，第 i 个静音帧的第一参数大于第 j 个静音帧的第一参数；在第一参数与谱熵负相关时，第 i 个静音帧的第一参数小于第 j 个静音帧的第一参数， i 和 j 均为正整数，且 $1 \leq i \leq T$ ， $1 \leq j \leq T$ 。

具体地，编码器可以对 T 个静音帧的谱参数进行加权平均，从而得到第一谱参数。如上所述，普通噪声的谱熵会较大，而非噪声信号或含有瞬态成份的噪声的谱熵会较小。因此，在 T 个静音帧中，谱熵较大的静音帧对应的加权系数可以大于或等于谱熵较小的静音帧对应的加权系数。

- 15 可选地，作为另一实施例，在图 6 的方法由编码器执行时，上述 T 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 $(T-1)$ 个静音帧。

在图 6 的方法由解码器执行时，上述 T 个静音帧可以为 T 个拖尾帧。

可选地，作为另一实施例，在图 6 的方法由编码器执行时，编码器可以将当前输入静音帧编码为 SID 帧，其中 SID 帧包括第一谱参数。

- 20 本发明实施例中，编码器可以在编码 SID 帧时，使 SID 帧包括各个子带的第一谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到 SID 帧中的谱参数，从而能够提升解码器根据该 SID 帧生成的舒适噪声的质量。

图 7 是根据本发明一个实施例的信号编码设备的示意框图。图 7 的设备 700

的一个例子为编码器，例如图 1 所示的编码器 110。设备 700 包括第一确定单元 710、第二确定单元 720、第三确定单元 730 和编码单元 740。

第一确定单元 710 在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧生成 5 的舒适噪声，并确定实际静音信号，其中当前输入帧为静音帧。第二确定单元 720 确定第一确定单元 710 确定的舒适噪声与第一确定单元 710 确定的实际静音信号的偏离程度。第三确定单元 730 根据第二确定单元确定的偏离程度，确定当前输入帧的编码方式，当前输入帧的编码方式包括拖尾帧编码方式或 SID 帧编码方式。编码单元 740 根据第三确定单元 730 确定的当前输入帧的编码方式， 10 对当前输入帧进行编码。

本发明实施例中，通过在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧生成的舒适噪声，并确定舒适噪声与实际静音信号的偏离程度，根据该偏离程度确定当前输入帧的编码方式为拖尾帧编码方式或 SID 帧编码方式，而非简单 15 地根据统计得到的语音活动帧的数量将当前输入帧编码为拖尾帧，从而能够节省通信带宽。

可选地，作为一个实施例，第一确定单元 710 可以预测舒适噪声的特征参数，并确定实际静音信号的特征参数，其中舒适噪声的特征参数与实际静音信号的特征参数是一一对应的。第二确定单元 720 可以确定舒适噪声的特征参数 20 与实际静音信号的特征参数之间的距离。

可选地，作为另一实施例，第三确定单元 730 可以在舒适噪声的特征参数与实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下，确定当前输入帧的编码方式为 SID 帧编码方式，其中舒适噪声的特征参数与实际

静音信号的特征参数之间的距离与阈值集合中的阈值是一一对应的。第三确定单元 730 可以在舒适噪声的特征参数与实际静音信号的特征参数之间的距离大于或等于阈值集合中对应阈值的情况下，确定当前输入帧的编码方式为拖尾帧编码方式。

- 5 可选地，作为另一实施例，上述舒适噪声的特征参数可以用于表征以下至少一种信息：能量信息，谱信息。

可选地，作为另一实施例，上述能量信息可以包括 CELP 激励能量。上述谱信息可以包括以下至少一种：线性预测滤波器系数，FFT 系数，MDCT 系数。

- 线性预测滤波器系数可以包括以下至少一种：LSF 系数，LSP 系数，ISF 系
10 数，ISP 系数，反射系数，LPC 系数。

可选地，作为另一实施例，第一确定单元 710 可以根据当前输入帧的前一帧的舒适噪声参数和当前输入帧的特征参数，预测舒适噪声的特征参数。或者，第一确定单元 710 可以根据当前输入帧之前的 L 个拖尾帧的特征参数和当前输入帧的特征参数，预测舒适噪声的特征参数，其中 L 为正整数。

- 15 可选地，作为另一实施例，第一确定单元 710 可以确定当前输入帧的特征参数作为实际静音信号的特征参数。或者，第一确定单元 710 可以对 M 个静音帧的特征参数进行统计处理，以确定实际静音信号的特征参数。

可选地，作为另一实施例，上述 M 个静音帧可以包括当前输入帧以及当前输入帧之前的 (M-1) 个静音帧，M 为正整数。

- 20 可选地，作为另一实施例，舒适噪声的特征参数可以包括舒适噪声的码激励线性预测 CELP 激励能量和舒适噪声的线谱频率 LSF 系数，实际静音信号的特征参数可以包括实际静音信号的 CELP 激励能量和实际静音信号的 LSF 系数。第二确定单元 720 可以确定舒适噪声的 CELP 激励能量与实际静音信号的 CELP

激励能量之间的距离 D_e ，并确定舒适噪声的 LSF 系数与实际静音信号的 LSF 系数之间的距离 D_{lsf} 。

可选地，作为另一实施例，第三确定单元 730 可以在距离 D_e 小于第一阈值，且距离 D_{lsf} 小于第二阈值的情况下，确定当前输入帧的编码方式为 SID 帧编码方式。第三确定单元 730 可以在距离 D_e 大于或等于第一阈值，或者距离 D_{lsf} 大于或等于第二阈值的情况下，确定当前输入帧的编码方式为拖尾帧编码方式。

可选地，作为另一实施例，设备 700 还可以包括第四确定单元 750。第四确定单元 750 可以获取预设的第一阈值和预设的第二阈值。或者，第四确定单元 750 可以根据当前输入帧之前的 N 个静音帧的 CELP 激励能量确定第一阈值，并根据 N 个静音帧的 LSF 系数确定第二阈值，其中 N 为正整数。

可选地，作为另一实施例，第一确定单元 710 可以采用第一预测方式，预测舒适噪声，其中第一预测方式与解码器生成舒适噪声的方式相同。

设备 700 的其它功能和操作可以参照上面图 1 至图 3b 的方法实施例的过程，为了避免重复，此处不再赘述。

图 8 是根据本发明另一实施例的信号处理设备的示意框图。图 8 的设备 800 的例子为编码器或解码器，如图 1 所示的编码器 110 或解码器 120。设备 800 包括第一确定单元 810 和第二确定单元 820。

第一确定单元 810 确定 P 个静音帧中每个静音帧的组加权谱距离，其中 P 个静音帧中每个静音帧的组加权谱距离为 P 个静音帧中每个静音帧与其它 $(P-1)$ 个静音帧之间的加权谱距离之和， P 为正整数。第二确定单元 820 根据第一确定单元 810 确定的 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，其中第一谱参数用于生成舒适噪声。

本发明实施例中，通过根据 P 个静音帧中每个静音帧的组加权谱距离确定

用于生成舒适噪声的第一谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到用于生成舒适噪声的谱参数，从而能够提升舒适噪声的质量。

可选地，作为一个实施例，每个静音帧可以与一组加权系数相对应，其中在这一组加权系数中，对应于第一组子带的加权系数大于对应于第二组子带的加权系数，其中第一组子带的感知重要性大于第二组子带的感知重要性。

可选地，作为另一实施例，第二确定单元 820 可以从 P 个静音帧中选择第一静音帧，使得在 P 个静音帧中第一静音帧的组加权谱距离最小，并可以将第一静音帧的谱参数确定为第一谱参数。

可选地，作为另一实施例，第二确定单元 820 可以从 P 个静音帧中选择至少一个静音帧，使得在 P 个静音帧中至少一个静音帧的组加权谱距离均小于第三阈值，并根据至少一个静音帧的谱参数，确定第一谱参数。

可选地，作为另一实施例，在设备 800 为编码器时，设备 800 还可以包括编码单元 830。

上述 P 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 (P-1) 个静音帧。编码单元 830 可以将当前输入静音帧编码为 SID 帧，其中 SID 帧包括第二确定单元 820 确定的第一谱参数。

设备 800 的其它功能和操作可以参照上面图 4 的方法实施例的过程，为了避免重复，此处不再赘述。

图 9 是根据本发明另一实施例的信号处理设备的示意框图。图 9 的设备 900 的例子为编码器或解码器，如图 1 所示的编码器 110 或解码器 120。设备 900 包括划分单元 910、第一确定单元 920 和第二确定单元 930。

划分单元 910 将输入信号的频带划分为 R 个子带，其中 R 为正整数。第一确定单元 920 在划分单元 910 划分的 R 个子带中每个子带上，确定 S 个静音帧

中每个静音帧的子带组谱距离, S 个静音帧中每个静音帧的子带组谱距离为在每个子带上 S 个静音帧中每个静音帧与其它 (S-1) 个静音帧之间的谱距离之和, S 为正整数。第二确定单元 930 在每个子带上根据第一确定单元 920 确定的 S 个静音帧中每个静音帧的谱距离, 确定每个子带的第一谱参数, 其中每个子带 5 的第一谱参数用于生成舒适噪声。

本发明实施例中, 通过在 R 个子带中每个子带上根据 S 个静音帧中每个静音帧的谱距离确定用于生成舒适噪声的每个子带的谱参数, 而非简单地对多个静音帧的谱参数取平均或取中值得到用于生成舒适噪声的的谱参数, 从而能够提升舒适噪声的质量。

10 可选地, 作为一个实施例, 第二确定单元 930 可以在每个子带上, 从 S 个静音帧中选择第一静音帧, 使得在每个子带上的 S 个静音帧中第一静音帧的子带组谱距离最小, 并在每个子带上将第一静音帧的谱参数确定为每个子带的第一谱参数。

15 可选地, 作为另一实施例, 第二确定单元 930 可以在每个子带上, 从 S 个静音帧中选择至少一个静音帧, 使得至少一个静音帧的子带组谱距离均小于第四阈值, 并在每个子带上, 根据至少一个静音帧的谱参数确定每个子带的第一谱参数。

可选地, 作为另一实施例, 在设备 900 为编码器时, 设备 900 还可以包括编码单元 940。

20 上述 S 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 (S-1) 个静音帧。编码单元 940 可以将当前输入静音帧编码为 SID 帧, 其中 SID 帧包括每个子带的第一谱参数。

设备 900 的其它功能和操作可以参照上面图 5 的方法实施例的过程, 为了

避免重复，此处不再赘述。

图 10 是根据本发明另一实施例的信号处理设备的示意框图。图 10 的设备 1000 的一个例子为编码器或解码器，如图 1 所示的编码器 110 或解码器 120。设备 1000 包括第一确定单元 1010 和第二确定单元 1020。

- 5 第一确定单元 1010 确定 T 个静音帧中每个静音帧的第一参数，第一参数用于表征谱熵，T 为正整数。第二确定单元 1020 根据第一确定单元 1010 确定的 T 个静音帧中每个静音帧的第一参数，确定第一谱参数，其中第一谱参数用于生成舒适噪声。

10 本发明实施例中，通过根据 T 个静音帧的用于表征谱熵的第一参数确定用于生成舒适噪声的第一谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到用于生成舒适噪声的谱参数，从而能够提升舒适噪声的质量。

 可选地，作为一个实施例，第二确定单元 1020 可以在确定能够按照聚类准则将 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，根据第一组静音帧的谱参数，确定第一谱参数，其中第一组静音帧的第一参数所表征的谱熵均
15 大于第二组静音帧的第一参数所表征的谱熵；在确定不能够按照聚类准则将 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，对 T 个静音帧的谱参数进行加权平均处理，以确定第一谱参数，其中第一组静音帧的第一参数所表征的谱熵均大于第二组静音帧的第一参数所表征的谱熵。

 可选地，作为另一实施例，上述聚类准则可以包括：第一组静音帧中每个
20 静音帧的第一参数与第一均值之间的距离小于或等于第一组静音帧中每个静音帧的第一参数与第二均值之间的距离；第二组静音帧中每个静音帧的第一参数与第二均值之间的距离小于或等于第二组静音帧中每个静音帧的第一参数与第一均值之间的距离；第一均值与第二均值之间的距离大于第一组静音帧的第一

参数与第一均值之间的平均距离；第一均值与第二均值之间的距离大于第二组静音帧的第一参数与第二均值之间的平均距离。

其中，第一均值为第一组静音帧的第一参数的平均值，第二均值为第二组静音帧的第一参数的平均值。

- 5 可选地，作为另一实施例，第二确定单元 1020 可以对 T 个静音帧的谱参数进行加权平均处理，以确定第一谱参数。其中，对于 T 个静音帧中任意不同的第 i 个静音帧和第 j 个静音帧，第 i 个静音帧对应的加权系数大于或等于 j 个静音帧对应的加权系数；在第一参数与谱熵正相关时，第 i 个静音帧的第一参数大于第 j 个静音帧的第一参数；在第一参数与谱熵负相关时，第 i 个静音帧的第一参数小于第 j 个静音帧的第一参数， i 和 j 均为正整数，且 $1 \leq i \leq T$ ， $1 \leq j \leq T$ 。

可选地，作为另一实施例，在设备 1000 为编码器时，设备 1000 还可以包括编码单元 1030。

- 上述 T 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 $(T-1)$ 个静音帧。编码单元 1030 可以将当前输入静音帧编码为 SID 帧，其中 SID 帧包
15 括第一谱参数。

设备 1000 的其它功能和操作可以参照上面图 6 的方法实施例的过程，为了避免重复，此处不再赘述。

图 11 是根据本发明另一实施例的信号编码设备的示意框图。图 7 的设备 1100 的一个例子为编码器。设备 1100 包括存储器 1110 和处理器 1120。

- 20 存储器 1110 可以包括随机存储器、闪存、只读存储器、可编程只读存储器、非易失性存储器或寄存器等。处理器 1120 可以是中央处理器 (Central Processing Unit, CPU)。

存储器 1110 用于存储可执行指令。处理器 1120 可以执行存储器 1110 中存

储的可执行指令，用于：在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧生成的舒适噪声，并确定实际静音信号，其中当前输入帧为静音帧；确定舒适噪声与实际静音信号的偏离程度；根据偏离程度，确定当前输入帧的编码方式，

5 当前输入帧的编码方式包括拖尾帧编码方式或 SID 帧编码方式；根据当前输入帧的编码方式，对当前输入帧进行编码。

本发明实施例中，通过在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在当前输入帧被编码为 SID 帧的情况下解码器根据当前输入帧生成的舒适噪声，并确定舒适噪声与实际静音信号的偏离程度，根据该偏离程

10 度确定当前输入帧的编码方式为拖尾帧编码方式或 SID 帧编码方式，而非简单地根据统计得到的语音活动帧的数量将当前输入帧编码为拖尾帧，从而能够节省通信带宽。

可选地，作为一个实施例，处理器 1120 可以预测舒适噪声的特征参数，并确定实际静音信号的特征参数，其中舒适噪声的特征参数与实际静音信号的特

15 征参数是一一对应的。处理器 1120 可以确定舒适噪声的特征参数与实际静音信号的特征参数之间的距离。

可选地，作为另一实施例，处理器 1120 可以在舒适噪声的特征参数与实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下，确定当前输入帧的编码方式为 SID 帧编码方式，其中舒适噪声的特征参数与实际静音信

20 号的特征参数之间的距离与阈值集合中的阈值是一一对应的。处理器 1120 可以在舒适噪声的特征参数与实际静音信号的特征参数之间的距离大于或等于阈值集合中对应阈值的情况下，确定当前输入帧的编码方式为拖尾帧编码方式。

可选地，作为另一实施例，上述舒适噪声的特征参数可以用于表征以下至

少一种信息：能量信息，谱信息。

可选地，作为另一实施例，上述能量信息可以包括 CELP 激励能量。上述谱信息可以包括以下至少一种：线性预测滤波器系数，FFT 系数，MDCT 系数。线性预测滤波器系数可以包括以下至少一种：LSF 系数，LSP 系数，ISF 系数，5 ISP 系数，反射系数，LPC 系数。

可选地，作为另一实施例，处理器 1120 可以根据当前输入帧的前一帧的舒适噪声参数和当前输入帧的特征参数，预测舒适噪声的特征参数。或者，处理器 1120 可以根据当前输入帧之前的 L 个拖尾帧的特征参数和当前输入帧的特征参数，预测舒适噪声的特征参数，其中 L 为正整数。

10 可选地，作为另一实施例，处理器 1120 可以确定当前输入帧的特征参数作为实际静音信号的参数。或者，处理器 1120 可以对 M 个静音帧的特征参数进行统计处理，以确定实际静音信号的参数。

可选地，作为另一实施例，上述 M 个静音帧可以包括当前输入帧以及当前输入帧之前的 (M-1) 个静音帧，M 为正整数。

15 可选地，作为另一实施例，舒适噪声的特征参数可以包括舒适噪声的码激励线性预测 CELP 激励能量和舒适噪声的线谱频率 LSF 系数，实际静音信号的特征参数可以包括实际静音信号的 CELP 激励能量和实际静音信号的 LSF 系数。处理器 1120 可以确定舒适噪声的 CELP 激励能量与实际静音信号的 CELP 激励能量之间的距离 D_e ，并确定舒适噪声的 LSF 系数与实际静音信号的 LSF 系数之
20 间的距离 D_{lsf} 。

可选地，作为另一实施例，处理器 1120 可以在距离 D_e 小于第一阈值，且距离 D_{lsf} 小于第二阈值的情况下，确定当前输入帧的编码方式为 SID 帧编码方式。处理器 1120 可以在距离 D_e 大于或等于第一阈值，或者距离 D_{lsf} 大于或等

于第二阈值的情况下，确定当前输入帧的编码方式为拖尾帧编码方式。

可选地，作为另一实施例，处理器 1120 还可以获取预设的第一阈值和预设的第二阈值。或者，处理器 1120 还可以根据当前输入帧之前的 N 个静音帧的 CELP 激励能量确定第一阈值，并根据 N 个静音帧的 LSF 系数确定第二阈值，

5 其中 N 为正整数。

可选地，作为另一实施例，处理器 1120 可以采用第一预测方式，预测舒适噪声，其中第一预测方式与解码器生成舒适噪声的方式相同。

设备 1100 的其它功能和操作可以参照上面图 1 至图 3b 的方法实施例的过程，为了避免重复，此处不再赘述。

10 图 12 是根据本发明另一实施例的信号编码设备的示意框图。图 12 的设备 1200 的例子为编码器或解码器，如图 1 所示的编码器 110 或解码器 120。设备 1200 包括存储器 1210 和处理器 1220。

存储器 1210 可以包括随机存储器、闪存、只读存储器、可编程只读存储器、非易失性存储器或寄存器等。处理器 1220 可以是 CPU。

15 存储器 1210 用于存储可执行指令。处理器 1220 可以执行存储器 1210 中存储的可执行指令，用于：确定 P 个静音帧中每个静音帧的组加权谱距离，其中 P 个静音帧中每个静音帧的组加权谱距离为 P 个静音帧中每个静音帧与其它 (P-1) 个静音帧之间的加权谱距离之和，P 为正整数；根据 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，其中第一谱参数用于生成舒适噪声。

20 本发明实施例中，通过根据 P 个静音帧中每个静音帧的组加权谱距离确定用于生成舒适噪声的第一谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到用于生成舒适噪声的谱参数，从而能够提升舒适噪声的质量。

可选地，作为一个实施例，每个静音帧可以与一组加权系数相对应，其中

在这一组加权系数中，对应于第一组子带的加权系数大于对应于第二组子带的加权系数，其中第一组子带的感知重要性大于第二组子带的感知重要性。

可选地，作为另一实施例，处理器 1220 可以从 P 个静音帧中选择第一静音帧，使得在 P 个静音帧中第一静音帧的组加权谱距离最小，并将第一静音帧的谱参数确定为第一谱参数。

可选地，作为另一实施例，处理器 1220 可以从 P 个静音帧中选择至少一个静音帧，使得在 P 个静音帧中至少一个静音帧的组加权谱距离均小于第三阈值，并根据至少一个静音帧的谱参数，确定第一谱参数。

可选地，作为另一实施例，在设备 1200 为编码器时，上述 P 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 (P-1) 个静音帧。处理器 1220 可以将当前输入静音帧编码为 SID 帧，其中 SID 帧包括第一谱参数。

设备 1200 的其它功能和操作可以参照上面图 4 的方法实施例的过程，为了避免重复，此处不再赘述。

图 13 是根据本发明另一实施例的信号处理设备的示意框图。图 13 的设备 1300 的例子为编码器或解码器，如图 1 所示的编码器 110 或解码器 120。设备 1300 包括存储器 1310 和处理器 1320。

存储器 1310 可以包括随机存储器、闪存、只读存储器、可编程只读存储器、非易失性存储器或寄存器等。处理器 1320 可以是 CPU。

存储器 1310 用于存储可执行指令。处理器 1320 可以执行存储器 1310 中存储的可执行指令，用于：将输入信号的频带划分为 R 个子带，其中 R 为正整数；在 R 个子带中的每个子带上，确定 S 个静音帧中每个静音帧的子带组谱距离，S 个静音帧中每个静音帧的子带组谱距离为在每个子带上 S 个静音帧中每个静音帧与其它 (S-1) 个静音帧之间的谱距离之和，S 为正整数；在每个子带上根据

S 个静音帧中每个静音帧的子带组谱距离，确定每个子带的第一谱参数，其中每个子带的第一谱参数用于生成舒适噪声。

本发明实施例中，通过根据 R 个子带中每个子带上的 S 个静音帧中每个静音帧的谱距离确定用于生成舒适噪声的每个子带的谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到用于生成舒适噪声的谱参数，从而能够提升舒适噪声的质量。

可选地，作为一个实施例，处理器 1320 可以在每个子带上，从 S 个静音帧中选择第一静音帧，使得在每个子带上 S 个静音帧中第一静音帧的子带组谱距离最小，并在每个子带上将第一静音帧的谱参数确定为每个子带的第一谱参数。

10 可选地，作为另一实施例，处理器 1320 可以在每个子带上，从 S 个静音帧中选择至少一个静音帧，使得至少一个静音帧的子带组谱距离均小于第四阈值，并在每个子带上，根据至少一个静音帧的谱参数确定每个子带的第一谱参数。

可选地，作为另一实施例，在设备 1300 为编码器时，上述 S 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 (S-1) 个静音帧。处理器 1320
15 可以将当前输入静音帧编码为 SID 帧，其中 SID 帧包括每个子带的第一谱参数。

设备 1300 的其它功能和操作可以参照上面图 5 的方法实施例的过程，为了避免重复，此处不再赘述。

图 14 是根据本发明另一实施例的信号处理设备的示意框图。图 14 的设备 1400 的例子为编码器或解码器，如图 1 所示的编码器 110 或解码器 120。设备
20 1400 包括存储器 1410 和处理器 1420。

存储器 1410 可以包括随机存储器、闪存、只读存储器、可编程只读存储器、非易失性存储器或寄存器等。处理器 1420 可以是 CPU。

存储器 1410 用于存储可执行指令。处理器 1420 可以执行存储器 1410 中存

储的可执行指令，用于：确定 T 个静音帧中每个静音帧的第一参数，第一参数用于表征谱熵， T 为正整数；根据 T 个静音帧中每个静音帧的第一参数，确定第一谱参数，其中第一谱参数用于生成舒适噪声。

本发明实施例中，通过根据 T 个静音帧的用于表征谱熵的第一参数确定用于生成舒适噪声的第一谱参数，而非简单地对多个静音帧的谱参数取平均或取中值得到用于生成舒适噪声的谱参数，从而能够提升舒适噪声的质量。

可选地，作为一个实施例，处理器 1420 可以在确定能够按照聚类准则将 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，根据第一组静音帧的谱参数，确定第一谱参数，其中第一组静音帧的第一参数所表征的谱熵均大于第二组静音帧的第一参数所表征的谱熵；在确定不能够按照聚类准则将 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，对 T 个静音帧的谱参数进行加权平均处理，以确定第一谱参数，其中第一组静音帧的第一参数所表征的谱熵均大于第二组静音帧的第一参数所表征的谱熵。

可选地，作为另一实施例，上述聚类准则可以包括：第一组静音帧中每个静音帧的第一参数与第一均值之间的距离小于或等于第一组静音帧中每个静音帧的第一参数与第二均值之间的距离；第二组静音帧中每个静音帧的第一参数与第二均值之间的距离小于或等于第二组静音帧中每个静音帧的第一参数与第一均值之间的距离；第一均值与第二均值之间的距离大于第一组静音帧的第一参数与第一均值之间的平均距离；第一均值与第二均值之间的距离大于第二组静音帧的第一参数与第二均值之间的平均距离。

其中，第一均值为第一组静音帧的第一参数的平均值，第二均值为第二组静音帧的第一参数的平均值。

可选地，作为另一实施例，处理器 1420 可以对 T 个静音帧的谱参数进行加

权平均处理，以确定第一谱参数。其中，对于 T 个静音帧中任意不同的第 i 个静音帧和第 j 个静音帧，第 i 个静音帧对应的加权系数大于或等于 j 个静音帧对应的加权系数；在第一参数与谱熵正相关时，第 i 个静音帧的第一参数大于第 j 个静音帧的第一参数；在第一参数与谱熵负相关时，第 i 个静音帧的第一参数小于第 j 个静音帧的第一参数， i 和 j 均为正整数，且 $1 \leq i \leq T$ ， $1 \leq j \leq T$ 。

可选地，作为另一实施例，在设备 1400 为编码器时，上述 T 个静音帧可以包括当前输入静音帧以及当前输入静音帧之前的 $(T-1)$ 个静音帧。处理器 1420 可以将当前输入静音帧编码为 SID 帧，其中 SID 帧包括第一谱参数。

设备 1400 的其它功能和操作可以参照上面图 6 的方法实施例的过程，为了避免重复，此处不再赘述。

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发明的范围。

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。

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。

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
5 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。

另外，在本发明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。

10 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，本发明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（可以是个人计算机，服务器，或者网络设备
15 等）执行本发明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（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

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

权 利 要 求

1. 一种信号编码方法，其特征在于，包括：

在当前输入帧的前一帧的编码方式为连续编码方式的情况下，预测在所述当前输入帧被编码为静音描述 SID 帧的情况下解码器根据所述当前输入帧生成
5 的舒适噪声，并确定实际静音信号，其中所述当前输入帧为静音帧；

确定所述舒适噪声与所述实际静音信号的偏离程度；

根据所述偏离程度，确定所述当前输入帧的编码方式，所述当前输入帧的编码方式包括拖尾帧编码方式或 SID 帧编码方式；

根据所述当前输入帧的编码方式，对所述当前输入帧进行编码。

10 2. 根据权利要求 1 所述的方法，其特征在于，所述预测在所述当前输入帧被编码为 SID 帧的情况下解码器根据所述当前输入帧生成的舒适噪声，并确定实际静音信号，包括：

预测所述舒适噪声的特征参数，并确定所述实际静音信号的特征参数，其中所述舒适噪声的特征参数与所述实际静音信号的特征参数是一一对应的；

15 所述确定所述舒适噪声与所述实际静音信号的偏离程度，包括：

确定所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离。

3. 根据权利要求 2 所述的方法，其特征在于，所述根据所述偏离程度，确定所述当前输入帧的编码方式，包括：

20 在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述 SID 帧编码方式，其中所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离与所述阈值集合中的阈值是一一对应的；

在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离大于或等于所述阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述拖尾帧编码方式。

4. 根据权利要求 2 或 3 所述的方法，其特征在于，所述舒适噪声的特征参数用于表征以下至少一种信息：能量信息，谱信息。

5. 根据权利要求 4 所述的方法，其特征在于，所述能量信息包括码激励线性预测 CELP 激励能量；

所述谱信息包括以下至少一种：线性预测滤波器系数，快速傅立叶变换 FFT 系数，修正离散余弦变换 MDCT 系数；

10 所述线性预测滤波器系数包括以下至少一种：线谱频率 LSF 系数，线谱对 LSP 系数，导抗谱频率 ISF 系数，导谱对 ISP 系数，反射系数，线性预测编码 LPC 系数。

6. 根据权利要求 2 至 5 中任一项所述的方法，其特征在于，所述预测所述舒适噪声的特征参数，包括：

15 根据所述当前输入帧的前一帧的舒适噪声参数和所述当前输入帧的特征参数，预测所述舒适噪声的特征参数；或者，

根据所述当前输入帧之前的 L 个拖尾帧的特征参数和所述当前输入帧的特征参数，预测所述舒适噪声的特征参数，其中 L 为正整数。

7. 根据权利要求 2 至 6 中任一项所述的方法，其特征在于，所述确定所述
20 实际静音信号的特征参数，包括：

将所述当前输入帧的特征参数作为所述实际静音信号的特征参数；或者，

对 M 个静音帧的特征参数进行统计处理，以确定所述实际静音信号的特征参数。

8. 根据权利要求 7 所述的方法, 其特征在于, 所述 M 个静音帧包括所述当前输入帧以及所述当前输入帧之前的 $(M-1)$ 个静音帧, M 为正整数。

9. 根据权利要求 3 所述的方法, 其特征在于, 所述舒适噪声的特征参数包括所述舒适噪声的码激励线性预测 CELP 激励能量和所述舒适噪声的线谱频率 LSF 系数, 所述实际静音信号的特征参数包括所述实际静音信号的 CELP 激励能量和所述实际静音信号的 LSF 系数;

所述确定所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离, 包括:

确定所述舒适噪声的 CELP 激励能量与所述实际静音信号的 CELP 激励能量之间的距离 D_e , 并确定所述舒适噪声的 LSF 系数与所述实际静音信号的 LSF 系数之间的距离 D_{lsf} 。

10. 根据权利要求 9 所述的方法, 其特征在于, 所述在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下, 确定所述当前输入帧的编码方式为所述 SID 帧编码方式, 包括:

15 在所述距离 D_e 小于第一阈值, 且所述距离 D_{lsf} 小于第二阈值的情况下, 确定所述当前输入帧的编码方式为所述 SID 帧编码方式;

所述在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离大于或等于所述阈值集合中对应阈值的情况下, 确定所述当前输入帧的编码方式为所述拖尾帧编码方式, 包括:

20 在所述距离 D_e 大于或等于第一阈值, 或者所述距离 D_{lsf} 大于或等于第二阈值的情况下, 确定所述当前输入帧的编码方式为所述拖尾帧编码方式。

11. 根据权利要求 10 所述的方法, 其特征在于, 还包括:

获取预设的所述第一阈值和预设的所述第二阈值; 或者,

根据所述当前输入帧之前的 N 个静音帧的 CELP 激励能量确定所述第一阈值，并根据所述 N 个静音帧的 LSF 系数确定所述第二阈值，其中 N 为正整数。

12. 根据权利要求 1 至 11 中任一项所述的方法，其特征在于，所述预测在所述当前输入帧被编码为 SID 帧的情况下解码器根据所述当前输入帧生成的舒

5 适噪声，包括：

采用第一预测方式，预测所述舒适噪声，其中所述第一预测方式与所述解码器生成所述舒适噪声的方式相同。

13. 一种信号处理方法，其特征在于，包括：

确定 P 个静音帧中每个静音帧的组加权谱距离，其中所述 P 个静音帧中每
10 个静音帧的组加权谱距离为所述 P 个静音帧中所述每个静音帧与其它 (P-1) 个静音帧之间的加权谱距离之和，P 为正整数；

根据所述 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，其中所述第一谱参数用于生成舒适噪声。

14. 根据权利要求 13 所述的方法，其特征在于，所述每个静音帧与一组加
15 权系数相对应，其中在所述一组加权系数中，对应于第一组子带的加权系数大于对应于第二组子带的加权系数，其中所述第一组子带的感知重要性大于所述第二组子带的感知重要性。

15. 根据权利要求 13 或 14 所述的方法，其特征在于，所述根据所述 P 个静音帧中每个静音帧的组加权谱距离，确定第一谱参数，包括：

20 从所述 P 个静音帧中选择第一静音帧，使得在所述 P 个静音帧中所述第一静音帧的组加权谱距离最小；

将所述第一静音帧的谱参数确定为所述第一谱参数。

16. 根据权利要求 13 或 14 所述的方法，其特征在于，所述根据所述 P 个

静音帧中每个静音帧的组加权谱距离，确定第一谱参数，包括：

从所述 P 个静音帧中选择至少一个静音帧，使得在所述 P 个静音帧中所述至少一个静音帧的组加权谱距离均小于第三阈值；

根据所述至少一个静音帧的谱参数，确定所述第一谱参数。

5 17. 根据权利要求 13 至 16 中任一项所述的方法，其特征在于，所述 P 个静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 (P-1) 个静音帧。

18. 根据权利要求 17 所述的方法，其特征在于，还包括：

将所述当前输入静音帧编码为静音描述 SID 帧，其中所述 SID 帧包括所述第一谱参数。

10 19. 一种信号处理方法，其特征在于，包括：

将输入信号的频带划分为 R 个子带，其中 R 为正整数；

在所述 R 个子带中的每个子带上，确定 S 个静音帧中每个静音帧的子带组谱距离，所述 S 个静音帧中每个静音帧的子带组谱距离为在所述每个子带上所述 S 个静音帧中所述每个静音帧与其它 (S-1) 个静音帧之间的谱距离之和，S

15 为正整数；

在所述每个子带上根据所述 S 个静音帧中每个静音帧的子带组谱距离，确定所述每个子带的第一谱参数，其中所述每个子带的第一谱参数用于生成舒适噪声。

20 20. 根据权利要求 19 所述的方法，其特征在于，所述在所述每个子带上，根据所述 S 个静音帧中每个静音帧的子带组谱距离，确定所述每个子带的第一谱参数，包括：

在所述每个子带上，从所述 S 个静音帧中选择第一静音帧，使得在所述每个子带上所述 S 个静音帧中所述第一静音帧的子带组谱距离最小；

在所述每个子带上，将所述第一静音帧的谱参数确定为所述每个子带的第一谱参数。

21. 根据权利要求 19 所述的方法，其特征在于，所述在所述每个子带上，根据所述 S 个静音帧中每个静音帧的子带组谱距离，确定所述每个子带的第一谱参数，包括：

在所述每个子带上，从所述 S 个静音帧中选择至少一个静音帧，使得所述至少一个静音帧的子带组谱距离均小于第四阈值；

在所述每个子带上，根据所述至少一个静音帧的谱参数，确定所述每个子带的第一谱参数。

22. 根据权利要求 19 至 21 中任一项所述的方法，其特征在于，所述 S 个静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 (S-1) 个静音帧。

23. 根据权利要求 22 所述的方法，其特征在于，还包括：

将所述当前输入静音帧编码为静音描述 SID 帧，其中所述 SID 帧包括所述每个子带的第一谱参数。

24. 一种信号处理方法，其特征在于，包括：

确定 T 个静音帧中每个静音帧的第一参数，所述第一参数用于表征谱熵，T 为正整数；

根据所述 T 个静音帧中每个静音帧的第一参数，确定第一谱参数，其中所述第一谱参数用于生成舒适噪声。

25. 根据权利要求 24 所述的方法，其特征在于，所述根据所述 T 个静音帧中每个静音帧的第一参数，确定第一谱参数，包括：

在确定能够按照聚类准则将所述 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，根据所述第一组静音帧的谱参数，确定所述第一谱参数，其中

所述第一组静音帧的第一参数所表征的谱熵均大于所述第二组静音帧的第一参数所表征的谱熵；

在确定不能够按照聚类准则将所述 T 个静音帧分为第一组静音帧和第二组静音帧的情况下，对所述 T 个静音帧的谱参数进行加权平均处理，以确定所述
5 第一谱参数，其中所述第一组静音帧的第一参数所表征的谱熵均大于所述第二组静音帧的第一参数所表征的谱熵。

26. 根据权利要求 25 所述的方法，其特征在于，所述聚类准则包括：

所述第一组静音帧中每个静音帧的第一参数与第一均值之间的距离小于或
等于所述第一组静音帧中每个静音帧的第一参数与第二均值之间的距离；所述
10 第二组静音帧中每个静音帧的第一参数与所述第二均值之间的距离小于或等于
所述第二组静音帧中每个静音帧的第一参数与所述第一均值之间的距离；所述
第一均值与所述第二均值之间的距离大于所述第一组静音帧的第一参数与所
述第一均值之间的平均距离；所述第一均值与所述第二均值之间的距离大于所
述第二组静音帧的第一参数与所述第二均值之间的平均距离；

15 其中，所述第一均值为所述第一组静音帧的第一参数的平均值，所述第二
均值为所述第二组静音帧的第一参数的平均值。

27. 根据权利要求 24 所述的方法，其特征在于，所述根据所述 T 个静音帧
中每个静音帧的第一参数，确定第一谱参数，包括：

对所述 T 个静音帧的谱参数进行加权平均处理，以确定所述第一谱参数；

20 其中，对于所述 T 个静音帧中任意不同的第 i 个静音帧和第 j 个静音帧，所
述第 i 个静音帧对应的加权系数大于或等于所述 j 个静音帧对应的加权系数；

在所述第一参数与所述谱熵正相关时，所述第 i 个静音帧的第一参数大于所
述第 j 个静音帧的第一参数；在所述第一参数与所述谱熵负相关时，所述第 i 个

静音帧的第一参数小于所述第 j 个静音帧的第一参数, i 和 j 均为正整数, 且 $1 \leq i \leq T$, $1 \leq j \leq T$ 。

28. 根据权利要求 24 至 27 中任一项所述的方法, 其特征在于, 所述 T 个静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 $(T-1)$ 个静音帧

5 29. 根据权利要求 28 所述的方法, 其特征在于, 还包括:

将所述当前输入静音帧编码为静音描述 SID 帧, 其中所述 SID 帧包括所述第一谱参数。

30. 一种信号编码设备, 其特征在于, 包括:

10 第一确定单元, 用于在当前输入帧的前一帧的编码方式为连续编码方式的情况下, 预测在所述当前输入帧被编码为静音描述 SID 帧的情况下解码器根据所述当前输入帧生成的舒适噪声, 并确定实际静音信号, 其中所述当前输入帧为静音帧;

第二确定单元, 用于确定所述第一确定单元确定的所述舒适噪声与所述第一确定单元确定的所述实际静音信号的偏离程度;

15 第三确定单元, 用于根据所述第二确定单元确定的所述偏离程度, 确定所述当前输入帧的编码方式, 所述当前输入帧的编码方式包括拖尾帧编码方式或 SID 帧编码方式;

编码单元, 用于根据所述第三确定单元确定的所述当前输入帧的编码方式, 对所述当前输入帧进行编码。

20 31. 根据权利要求 30 所述的设备, 其特征在于, 所述第一确定单元具体用于预测所述舒适噪声的特征参数, 并确定所述实际静音信号的特征参数, 其中所述舒适噪声的特征参数与所述实际静音信号的特征参数是一一对应的;

所述第二确定单元具体用于确定所述舒适噪声的特征参数与所述实际静音

信号的特征参数之间的距离。

32. 根据权利要求 31 所述的设备，其特征在于，所述第三确定单元具体用于：在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离小于阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述 SID 5 帧编码方式，其中所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离与所述阈值集合中的阈值是一一对应的；在所述舒适噪声的特征参数与所述实际静音信号的特征参数之间的距离大于或等于所述阈值集合中对应阈值的情况下，确定所述当前输入帧的编码方式为所述拖尾帧编码方式。

33. 根据权利要求 31 或 32 所述的设备，其特征在于，所述第一确定单元 10 具体用于：根据所述当前输入帧的前一帧的舒适噪声参数和所述当前输入帧的特征参数，预测所述舒适噪声的特征参数；或者，根据所述当前输入帧之前的 L 个拖尾帧的特征参数和所述当前输入帧的特征参数，预测所述舒适噪声的特征参数，其中 L 为正整数。

34. 根据权利要求 31 至 33 中任一项所述的设备，其特征在于，所述第一 15 确定单元具体用于：确定所述当前输入帧的特征参数作为所述实际静音信号的特征参数；或者，对 M 个静音帧的特征参数进行统计处理，以确定所述实际静音信号的特征参数。

35. 根据权利要求 32 所述的设备，其特征在于，所述舒适噪声的特征参数包括所述舒适噪声的码激励线性预测 CELP 激励能量和所述舒适噪声的线谱频 20 率 LSF 系数，所述实际静音信号的特征参数包括所述实际静音信号的 CELP 激励能量和所述实际静音信号的 LSF 系数；

所述第二确定单元具体用于确定所述舒适噪声的 CELP 激励能量与所述实际静音信号的 CELP 激励能量之间的距离 D_e ，并确定所述舒适噪声的 LSF 系数

与所述实际静音信号的 LSF 系数之间的距离 D_{lsf} 。

36. 根据权利要求 35 所述的设备, 其特征在于, 所述第三确定单元具体用于在所述距离 D_e 小于第一阈值, 且所述距离 D_{lsf} 小于第二阈值的情况下, 确定所述当前输入帧的编码方式为所述 SID 帧编码方式;

5 所述第三确定单元具体用于在所述距离 D_e 大于或等于第一阈值, 或者所述距离 D_{lsf} 大于或等于第二阈值的情况下, 确定所述当前输入帧的编码方式为所述拖尾帧编码方式。

37. 根据权利要求 36 所述的设备, 其特征在于, 还包括:

第四确定单元, 用于: 获取预设的所述第一阈值和预设的所述第二阈值;

10 或者, 根据所述当前输入帧之前的 N 个静音帧的 CELP 激励能量确定所述第一阈值, 并根据所述 N 个静音帧的 LSF 系数确定所述第二阈值, 其中 N 为正整数。

38. 根据权利要求 30 至 37 中任一项所述的设备, 其特征在于, 所述第一确定单元具体用于采用第一预测方式, 预测所述舒适噪声, 其中所述第一预测方式与所述解码器生成所述舒适噪声的方式相同。

15 39. 一种信号处理设备, 其特征在于, 包括:

第一确定单元, 用于确定 P 个静音帧中每个静音帧的组加权谱距离, 其中所述 P 个静音帧中每个静音帧的组加权谱距离为所述 P 个静音帧中所述每个静音帧与其它 $(P-1)$ 个静音帧之间的加权谱距离之和, P 为正整数;

第二确定单元, 用于根据所述第一确定单元确定的所述 P 个静音帧中每个
20 静音帧的组加权谱距离, 确定第一谱参数, 所述第一谱参数用于生成舒适噪声。

40. 根据权利要求 39 所述的设备, 其特征在于, 所述第二确定单元具体用于: 从所述 P 个静音帧中选择第一静音帧, 使得在所述 P 个静音帧中所述第一静音帧的组加权谱距离最小; 将所述第一静音帧的谱参数确定为所述第一谱参

数。

41. 根据权利要求 39 所述的设备，其特征在于，所述第二确定单元具体用于：从所述 P 个静音帧中选择至少一个静音帧，使得在所述 P 个静音帧中所述至少一个静音帧的组加权谱距离均小于第三阈值；根据所述至少一个静音帧的
5 谱参数，确定所述第一谱参数。

42. 根据权利要求 39 至 41 中任一项所述的设备，其特征在于，所述 P 个静音帧包括所述当前输入静音帧以及所述当前输入静音帧之前的 (P-1) 个静音帧；

所述设备还包括：

10 编码单元，用于将所述当前输入静音帧编码为静音描述 SID 帧，其中所述 SID 帧包括所述第二确定单元确定的所述第一谱参数。

43. 一种信号处理设备，其特征在于，包括：

划分单元，用于将输入信号的频带划分为 R 个子带，其中 R 为正整数；

第一确定单元，用于在所述划分单元划分的所述 R 个子带中每个子带上，
15 确定 S 个静音帧中每个静音帧的子带组谱距离，所述 S 个静音帧中每个静音帧的子带组谱距离为在所述每个子带上所述 S 个静音帧中所述每个静音帧与其它 (S-1) 个静音帧之间的谱距离之和，S 为正整数；

第二确定单元，用于在所述划分单元划分的所述每个子带上根据所述第一确定单元确定的 S 个静音帧中每个静音帧的子带组谱距离，确定所述每个子带
20 的第一谱参数，其中所述每个子带的第一谱参数用于生成舒适噪声。

44. 根据权利要求 43 所述的设备，其特征在于，所述第二确定单元具体用于：在所述每个子带上，从所述 S 个静音帧中选择第一静音帧，使得在所述每个子带上所述 S 个静音帧中所述第一静音帧的子带组谱距离最小；在所述每个

子带上, 将所述第一静音帧的谱参数确定为所述每个子带的第一谱参数。

45. 根据权利要求 43 所述的设备, 其特征在于, 所述第二确定单元具体用于: 在所述每个子带上, 从所述 S 个静音帧中选择至少一个静音帧, 使得所述至少一个静音帧的子带组谱距离均小于第四阈值; 在所述每个子带上, 根据所述至少一个静音帧的谱参数, 确定所述每个子带的第一谱参数。

46. 根据权利要求 43 至 45 中任一项所述的设备, 其特征在于, 所述 S 个静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 (S-1) 个静音帧;

所述设备还包括:

编码单元, 用于将所述当前输入静音帧编码为静音描述 SID 帧, 其中所述 SID 帧包括所述每个子带的谱参数。

47. 一种信号处理设备, 其特征在于, 包括:

第一确定单元, 用于确定 T 个静音帧中每个静音帧的第一参数, 所述第一参数用于表征谱熵, T 为正整数;

第二确定单元, 用于根据所述第一确定单元确定的所述 T 个静音帧中每个静音帧的第一参数, 确定第一谱参数, 其中所述第一谱参数用于生成舒适噪声。

48. 根据权利要求 47 所述的设备, 其特征在于, 所述第二确定单元具体用于: 在确定能够按照聚类准则将所述 T 个静音帧分为所述第一组静音帧和所述第二组静音帧的情况下, 根据所述第一组静音帧的谱参数, 确定所述第一谱参数, 其中所述第一组静音帧的第一参数所表征的谱熵均大于所述第二组静音帧的第一参数所表征的谱熵; 在确定不能够按照聚类准则将所述 T 个静音帧分为所述第一组静音帧和所述第二组静音帧的情况下, 对所述 T 个静音帧的谱参数进行加权平均处理, 以确定所述第一谱参数, 其中所述第一组静音帧的第一参数所表征的谱熵均大于所述第二组静音帧的第一参数所表征的谱熵。

49. 根据权利要求 47 所述的设备, 其特征在于, 所述第二确定单元具体用于: 对所述 T 个静音帧的谱参数进行加权平均处理, 以确定所述第一谱参数;

其中, 对于所述 T 个静音帧中任意不同的第 i 个静音帧和第 j 个静音帧, 所述第 i 个静音帧对应的加权系数大于或等于所述 j 个静音帧对应的加权系数; 在
5 所述第一参数与所述谱熵正相关时, 所述第 i 个静音帧的第一参数大于所述第 j 个静音帧的第一参数; 在所述第一参数与所述谱熵负相关时, 所述第 i 个静音帧的第一参数小于所述第 j 个静音帧的第一参数, i 和 j 均为正整数, 且 $1 \leq i \leq T$, $1 \leq j \leq T$ 。

50. 根据权利要求 47 至 49 中任一项所述的设备, 其特征在于, 所述 T 个
10 静音帧包括当前输入静音帧以及所述当前输入静音帧之前的 $(T-1)$ 个静音帧;

所述设备还包括:

编码单元, 用于将所述当前输入静音帧编码为静音描述 SID 帧, 其中所述
SID 帧包括所述第一谱参数。