

一种语音频信号处理方法和装置

本申请要求于 2012 年 03 月 01 日提交中国专利局、申请号为 201210051672.6、发明名称为“一种语音频信号处理方法和装置”的中国专利申请

申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本发明涉及数字信号处理技术领域，尤其是一种语音频信号处理方法和装置。

10 背景技术

在数字通信领域，语音、图像、音频、视频的传输有着非常广泛的应用需求，如手机通话、音视频会议、广播电视、多媒体娱乐等。音频被数字化处理，通过音频通信网络从一个终端传递到另一个终端，这里的终端可以是手机、数字电话终端或其他任何类型的音频终端，数字电话终端例如 VOIP 电话或 ISDN 电话、计算机、电缆通信电话。为了降低语音频信号存储或者传输过程中占用的资源，语音频信号在发送端进行压缩处理后传输到接收端，接收端通过解压缩处理恢复语音频信号并进行播放。

15

在目前的多速率语音频编码中，由于网络状态的不同，网络会对从编码端传输到网络的码流做不同码率的截断，在解码端就会根据截断后的码流解码出不同带宽的语音频信号，这样就使得输出的语音频信号会在不同带宽间做切换。

20

不同带宽信号间的突然切换，会造成人耳听觉上的明显不舒适感；同时，由于滤波器及时频或频时变换等状态的更新，一般需要用到前后帧间的参数，在带宽切换时，如果不做一些适当的处理，这些状态的更新将会出现错误，从而造成一些能量激变的现象，造成听觉质量变差。

25

发明内容

本发明实施例的目的在于提供一种语音频信号处理方法和装置，在语音频信号带宽切换时提高听觉舒适性。

根据本发明的一实施例，一种语音频信号处理方法包括：

语音频信号从宽频带信号到窄频带信号的切换时，获得当前帧语音频信号对应的初始高频带信号；

根据当前帧语音频信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数；

利用所述时域全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号；

合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

10 根据本发明的另一实施例，一种语音频信号处理方法包括：

当语音频信号出现带宽切换时，获得当前帧语音频信号对应的初始高频带信号；

获得所述初始高频带信号时域全局增益参数；

15 将能量比值和所述时域全局增益参数进行加权处理，得到的加权值作为预测的全局增益参数，其中，能量比值为历史帧高频带时域信号能量与当前帧初始高频带信号能量的比值；

利用预测的全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号；

合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

20

根据本发明的另一实施例，一种语音频信号处理装置包括：

预测单元，当语音频信号从宽频带信号到窄频带信号的切换时，用于获得当前帧语音频信号对应的初始高频带信号；

25 参数获得单元，用于根据当前帧语音频信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数；

修正单元，用于利用预测的全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号；

合成单元，用于合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

根据本发明的另一实施例，一种语音频信号处理装置包括：

获取单元，用于当语音频信号出现带宽切换时，获得当前帧语音频信号对应的初始高频带信号；

参数获得单元，用于获得所述初始高频带信号对应的时域全局增益参数；

5 加权处理单元，用于将能量比值和所述时域全局增益参数进行加权处理，得到的加权值作为预测的全局增益参数；其中，能量比值为历史帧高频带时域信号能量与当前帧初始高频带信号能量的比值；

修正单元，用于利用预测的全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号；

10 合成单元，用于合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

本发明实施例通过宽频带和窄频带间切换时对高频带信号的修正，使得宽频带和窄频带间高频带信号平稳的过渡，有效地去除了宽频带和窄频带间切换
15 时造成的听觉不舒适感；同时，由于带宽切换算法和切换前高频带信号的编解码算法在相同的信号域，保证了不增加额外延且算法简单的同时，还保证了输出信号的性能。

附图说明

20 为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中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性的前提下，还可以根据这些附图获得其他的附图。

图1为本发明提供的语音频信号处理方法一个实施例的流程示意图；

25 图2为本发明提供的语音频信号处理方法另一个实施例的流程示意图；

图3为本发明提供的语音频信号处理方法另一个实施例的流程示意图；

图4为本发明提供的语音频信号处理方法另一个实施例的流程示意图；

图5为本发明提供的语音频信号处理装置一个实施例的结构示意图；

图6为本发明提供的语音频信号处理装置一个实施例的结构示意图；

图7为本发明提供的参数获得单元一个实施例的结构示意图；

图8为本发明提供的全局增益参数获得单元一个实施例的结构示意图；

图9为本发明提供的获取单元一个实施例的结构示意图；

图10为本发明提供的语音频信号处理装置另一个实施例的结构示意图。

5

具体实施方式

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，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。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，都属于本发明保护的范围。

数字信号处理领域，音频编解码器、视频编解码器广泛应用于各种电子设备中，例如：移动电话，无线装置，个人数据助理（PDA），手持式或便携式计算机，GPS接收机/导航器，照相机，音频/视频播放器，摄像机，录像机，监控设备等。通常，这类电子设备中包括音频编码器或音频解码器，音频编码器或者解码器可以直接由数字电路或芯片例如DSP（digital signal processor）实现，或者由软件代码驱动处理器执行软件代码中的流程而实现。

在现有技术中，由于网络中传输的语语音频信号的带宽不同，在语语音频信号传输过程中，语语音频信号的带宽会时常发生变化，存在窄频带语语音频信号向宽频带语语音频信号切换，以及宽频带语语音频信号向窄频带语语音频信号切换的现象。这种语语音频信号在高低频带间切换的过程称为带宽切换，带宽切换包括从窄频带信号到宽频带信号的切换和从宽频带到窄频带信号的切换。本发明中提到的窄频带信号为通过上采样和低通滤波，只有低频带成分而高频带成分为空的语音信号，而宽频带语语音频信号既有低频带信号成分又有高频带信号成分。窄频带信号和宽频带信号是相对的，例如相对于窄带信号而言，宽带信号为宽频带信号；相对于宽带信号而言，超宽带信号为宽频带信号。通常，窄带信号为采样率为8kHz的语语音频信号；宽带信号为采样率为16kHz的语语音频信号；超宽带为采样率32kHz的语语音频信号。

在切换前的高频带信号的编解码算法根据信号类型不同在时域和频域的编解码算法间选择时，或当切换前的高频带信号的编码算法是时域编码算法

时,为了保证切换时输出信号的连续性,切换算法保持和切换前的高频带编解码算法在相同的信号域进行处理,即切换前高频带信号采用时域编解码算法,接下来的切换算法就采用时域的切换算法;切换前的高频带信号采用频域的编解码算法,接下来的切换算法就采用频域的切换算法。现有技术没有切换前使用5 用时域频带扩展算法切换后也使用类似的时域切换技术。

语音频编码一般以帧为单位进行处理。当前输入的需要处理的音频帧为当前帧语音频信号;当前帧语音频信号中包括窄频带信号和低频带信号,即当前帧窄频带信号和当前帧低频带信号。当前帧语音频信号之前的任意一帧语音频信号为历史帧语音频信号,也包括历史帧窄频带信号和历史帧低频带信号;当前10 帧语音频信号之前一帧语音频信号为前一帧语音频信号。

参考图1,本发明语音频信号处理方法的一个实施例包括:

S101:当语音频信号出现带宽切换时,获得当前帧语音频信号对应的初始高频带信号;

15 当前帧语音频信号是由当前帧窄频带信号和当前帧低频带时域信号组成。带宽切换包括从窄频带信号到宽频带信号的切换和从宽频带到窄频带信号的切换;对于从窄频带信号到宽频带信号的切换,当前帧语音频信号为当前帧宽频带信号,包括窄频带信号和低频带信号,当前帧语音频信号的初始高频带信号为真实的信号,可以直接从当前帧语音频信号中获得;对于从宽频带到窄频20 带信号的切换,当前帧语音频信号为当前帧窄频带信号,当前帧低频带时域信号为空,当前帧语音频信号的初始高频带信号为预测信号,需要预测当前帧窄频带信号对应的高频带信号,作为初始高频带信号。

S102:获得该初始高频带信号对应的时域全局增益参数;

对于窄频带信号到宽频带信号的切换,高频带信号的时域全局增益参数可以通过解码得到;对于宽频带信号到窄频带信号的切换,高频带信号的时域全25 局增益参数可以根据当前帧信号获得:根据窄频带信号的谱倾斜参数和当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数。

S103:将能量比值和该时域全局增益参数进行加权处理,得到的加权值作

为预测的全局增益参数；其中，能量比值为历史帧语音频信号高频带时域信号能量与当前帧语音频信号初始高频带信号能量的比值；

历史帧语音频信号使用的是历史帧最终输出的语音频信号，当前帧语音频信号使用的是指初始高频带信号；能量比值 $\text{Ratio} = \text{Esyn}(-1) / \text{Esyn_tmp}$ ；

- 5 $\text{Esyn}(-1)$ 表示历史帧输出的高频带时域信号syn的能量， Esyn_tmp 表示当前帧对应的初始高频带时域信号syn的能量。

预测的全局增益参数 $\text{gain} = \text{alfa} * \text{Ratio} + \text{beta} * \text{gain}'$ ，其中， gain' 为时域全局增益参数， $\text{alfa} + \text{beta} = 1$ ，且根据信号类型的不同， alfa 和 beta 的取值不同。

- 10 S104：利用预测的全局增益参数对该初始高频带信号进行修正，获得修正的高频带时域信号；

- 修正指信号相乘，即用预测的全局增益参数与初始高频带信号相乘。另一个实施例中，步骤S102中获得该初始高频带信号对应的时域包络参数和时域全局增益参数，则步骤S104中利用时域包络参数和预测的全局增益参数对该初始高频带信号进行修正，获得修正的高频带时域信号；即用
- 15 时域包络参数和预测的时域全局增益参数乘于该预测的高频带信号，获得高频带时域信号。

对于窄频带信号到宽频带信号的切换，高频带信号的时域包络参数可以通过解码得到；对于宽频带信号到窄频带信号的切换，高频带信号的时域包络参数可以根据当前帧信号获得：可以将预先设定好的一系列值或者历史帧高频带时域包络参数作为当前帧语音频信号的高频带时域包络参数。

- 20 S105：合成当前帧的窄频带时域信号和该修正的高频带时域信号并输出。

- 上述实施例通过宽频带和窄频带间切换时时高频带信号的修正，使得宽频带和窄频带间高频带信号平稳的过渡，有效地去除了宽频带和窄频带间切换时造成的听觉不舒适感；同时，由于带宽切换算法和切换前高频带信号的编解码算法在相同的信号域，保证了不增加额外延且算法简单的同时，还保证了输出
- 25 信号的性能。

参考图2，本发明语音频信号处理方法的另一个实施例包括：

S201：当宽频带信号向窄频带信号切换时，预测当前帧窄频带信号对应的预测高频带信号；

由宽频带信号向窄频带切换，即前一帧为宽频带信号，当前帧为窄频带信

号。预测当前帧窄频带信号对应的预测高频带信号的步骤包括：根据当前帧窄频带信号预测当前帧语音频信号高频带信号激励信号；预测当前帧语音频信号高频带信号的LPC（Linear Predictive Coding，线性预测编码）系数；合成预测的高频带激励信号和LPC系数，获得预测高频带信号syn_tmp。

5 一个实施例中，可以从窄频带信号中提取基音周期、代数码数和增益等参数，通过变采样，滤波预测到高频带的激励信号；

另一个实施例中，可以通过对窄频带时域信号或窄频带时域激励信号通过上采用、低通，然后取绝对值或取平方等操作来预测高频带激励信号。

预测高频带信号的LPC系数，可以将历史帧的高频带LPC系数或预先设定好的一系列值作为当前帧LPC系数；也可以对不同的信号类型采用不同的预测方式。

10 S202：获得所述预测高频带信号对应的时域包络参数和时域全局增益参数；

可以将预先设定好的一系列值作为当前帧的高频带时域包络参数。可以将窄带信号大体分几类，每类预先设定好一系列值，根据当前帧窄带信号的类型，选择一组预先设定好的时域包络参数；也可以就设定好一组时域包络值，例如，时域包络的个数为M，则预先设定好的值可以为M个0.3536。该实施例中，时域包络参数的获得为可选步骤，并不是必须的。

20 根据窄频带信号的谱倾斜参数和当前帧窄频带信号和历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数；一个实施例中，包括如下步骤：

S2021：根据所述当前帧语音频信号的谱倾斜参数和当前帧窄频带信号与历史帧窄频带信号的相关性，将当前帧语音频信号分为第一类信号或第二类信号；一个实施例中，第一类信号为摩擦音信号，第二类信号为非摩擦音信号；当谱倾斜参数tilt>5且相关性参数cor小于一给定值时，将窄频带信号分成摩擦音，其他的为非摩擦音。

其中，当前帧窄频带信号和历史帧窄频带信号的相关性大小参数cor的计算，可以通过相同某频段信号的能量的大小关系来确定，也可以通过几个相同频段的能量关系确定，也可以通过时域信号或时域激励信号的自相关或互相关

公式来计算。

S2022: 如果当前帧语音频信号为第一类信号, 则将谱倾斜参数限制到小于等于第一预定值, 获得谱倾斜参数限制值; 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。即当前帧语音频信号的谱倾斜参数小于等于第一预定值时, 保留谱倾斜参数原值作为谱倾斜参数限制值; 当前帧语音频信号的谱倾斜参数大于第一预定值时, 取第一预定值作为谱倾斜参数限制值。

时域全局增益参数 $gain'$ 通过以下公式获得:

$$gain' = \begin{cases} tilt, & tilt \leq \partial 1 \\ \partial 1, & tilt > \partial 1 \end{cases}, \text{ 其中, } tilt \text{ 为谱倾斜参数, } \partial 1 \text{ 为第一预订值。}$$

S2023: 如果当前帧语音频信号为第二类信号, 则将谱倾斜参数限制到属于第一区间值, 获得谱倾斜参数限制值; 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。即当前帧语音频信号的谱倾斜参数属于第一区间值时, 保留谱倾斜参数原值作为谱倾斜参数限制值; 当前帧语音频信号的谱倾斜参数大于第一区间值的上限时, 取第一区间值的上限作为谱倾斜参数限制值; 当前帧语音频信号的谱倾斜参数小于第一区间值的下限时, 取第一区间值的下限作为谱倾斜参数限制值。

时域全局增益参数 $gain'$ 通过以下公式获得:

$$gain' = \begin{cases} tilt, & tilt \in [a, b] \\ a, & tilt < a \\ b, & tilt > b \end{cases}, \text{ 其中, } tilt \text{ 为谱倾斜参数, } [a, b] \text{ 为第一区间值。}$$

一个实施例中, 获得窄频带信号的谱倾斜参数 $tilt$ 及当前帧窄频带信号和历史帧窄频带信号的相关性大小参数 cor ; 根据 $tilt$ 及 cor 将当前帧信号分为摩擦音及非摩擦音两类, 当谱倾斜参数 $tilt > 5$ 且相关性参数 cor 小于一给定值时, 将窄频带信号分成摩擦音, 其他的为非摩擦音; 将 $tilt$ 的取值范围限制到 $0.5 \leq tilt \leq 1.0$ 之间作为非摩擦音的时域全局增益参数, 将 $tilt$ 的取值范围限制到 $tilt \leq 8.0$ 作为摩擦音的时域全局增益参数。对摩擦音而言, 谱倾斜参数可以是大于5的任何值, 对非摩擦音而言, 可以小于等于5的任何值, 也可能大于5, 为了保证能将谱倾斜参数 $tilt$ 能作为估计的时域全局增益参数, 对 $tilt$ 的值的范围做限定后作为时域全局增益参数, 即当 $tilt > 8$ 时, 取 $tilt = 8$ 作为摩擦音的时域

全局增益参数，当 $\text{tilt} < 0.5$ 时，取 $\text{tilt} = 0.5$ 或 $\text{tilt} > 1.0$ 时，取 $\text{tilt} = 1.0$ 作为非摩擦音的时域全局增益参数。

S203: 将能量比值和该时域全局增益参数进行加权处理，得到的加权值作为预测的全局增益参数；其中，能量比值为历史帧语音频信号高频带时域信号能量与当前帧语音频信号初始高频带信号能量的比值；

求解能量比值 $\text{Ratio} = \text{Esyn}(-1) / \text{Esyn_tmp}$ ，将 tilt 和 Ratio 的加权值作为当前帧预测的全局增益参数 gain ，即 $\text{gain} = \text{alfa} * \text{Ratio} + \text{beta} * \text{gain}'$ ；其中， gain' 为时域全局增益参数， $\text{alfa} + \text{beta} = 1$ ，且根据信号类型的不同， alfa 和 beta 的取值不同； $\text{Esyn}(-1)$ 表示历史帧的最终输出的高频带时域信号 syn 的能量， Esyn_tmp 表示当前帧预测高频带时域信号 syn 的能量。

S204: 利用时域包络参数和预测的全局增益参数对该预测高频带信号进行修正，获得修正的高频带时域信号；

用时域包络参数和预测的时域全局增益参数乘于该预测的高频带信号，获得高频带时域信号。

该实施例中，时域包络参数为可选的，当仅包含时域全局增益参数时，则可以利用预测的全局增益参数对该预测高频带信号进行修正，获得修正的高频带时域信号；即用预测的全局增益参数乘于预测高频带信号得到修正的高频带时域信号。

S205: 合成当前帧的窄频带时域信号和该修正的高频带时域信号并输出。

高频带时域信号 syn 的能量 Esyn 用来预测下一帧时域全局增益参数，即将 Esyn 的值赋值给 $\text{Esyn}(-1)$

上述实施例通过对宽频带信号后窄频带信号高频带的修正，使得宽频带和窄频带间高频带部分平稳的过渡，有效地去除了宽频带和窄频带间切换时造成的听觉不舒适感；同时，由于对切换时的帧进行了相应的处理，间接去除了参数和状态更新时出现的问题。通过保持带宽切换算法和切换前高频带信号的编解码算法在相同的信号域，保证了不增加额外延且算法简单的同时，还保证了输出信号的性能。

参考图3，本发明语音频信号处理方法的另一个实施例包括：

S301：当窄频带信号向宽频带信号切换时，获得当前帧高频带信号；

当由窄频带信号向宽频带切换时，即前一帧为窄频带信号，当前帧为宽频带信号。

5 S302：获得所述高频带信号对应的时域包络参数和时域全局增益参数；
该时域包络参数和时域全局增益参数可以从当前帧高频带信号中直接获得。其中，时域包络参数的获得为可选步骤。

S303：将能量比值和该时域全局增益参数进行加权处理，得到的加权值作为预测的全局增益参数；其中，能量比值为历史帧语音频信号高频带时域信号
10 能量与当前帧语音频信号初始高频带信号能量的比值。；

因为当前帧是宽频带信号，所以高频带信号的各项参数都能通过解码得到，为了保证切换时能平滑过渡，通过如下方式对时域全局增益参数进行平滑：

求解能量比值 $\text{Ratio} = \text{Esyn}(-1) / \text{Esyn_tmp}$ ， $\text{Esyn}(-1)$ 表示历史帧的最终输出的高频带时域信号syn的能量； Esyn_tmp 当前帧的高频带时域信号syn的能量。

15 将解码出的时域全局增益参数gain和Ratio的加权值作为当前帧预测的全局增益参数gain，即 $\text{gain} = \text{alfa} * \text{Ratio} + \text{beta} * \text{gain}'$ ，其中， gain' 为时域全局增益参数， $\text{alfa} + \text{beta} = 1$ ，且根据信号类型的不同，alfa和beta的取值不同

如果当前音频帧与前一帧语音频信号的窄带信号具有预定相关性时，则对前一帧语音频信号对应的所述能量比值的加权因子alfa按一定的步长衰减后的
20 值作为当前音频帧对应的所述能量比值的加权因子，逐帧衰减直到alfa为0。

当前后帧间窄频带信号有相同的信号类型或相关性满足一定的条件时，即前后帧间有一定的相关性，或前后帧间信号类型相似，则对alfa按一定的步长逐帧衰减，直到alfa衰减到0；当前后帧间窄频带信号不具有相关性时，直接将alfa衰减到0，即保持当前解码结果，不做加权和修正处理。

25 S304：利用时域包络参数和预测的全局增益参数对该高频带信号进行修正，获得修正的高频带时域信号；

修正即用时域包络参数和预测的时域全局增益参数乘于该高频带信号，获得修正的高频带时域信号。

该实施例中，时域包络参数为可选的，当仅包含时域时域全局增益参数时，

则可以利用预测的全局增益参数对该高频带信号进行修正,获得修正的高频带时域信号;即用预测的全局增益参数乘于高频带信号得到修正的高频带时域信号。

S305: 合成当前帧的窄频带时域信号和该修正的高频带时域信号并输出。

- 5 上述实施例通过对窄频带信号后宽频带信号高频带的修正,使得宽频带和窄频带间高频带部分平稳的过渡,有效地去除了宽频带和窄频带间切换时造成的听觉不舒适感;同时,由于对切换时的帧进行了相应的处理,间接去除了参数和状态更新时出现的问题。通过保持带宽切换算法和切换前高频带信号的编解码算法在相同的信号域,保证了不增加额外延迟且算法简单的同时,还保证了
- 10 输出信号的性能。

参考图4,本发明语音频信号处理方法的另一个实施例包括:

S401: 语音频信号从宽频带信号到窄频带信号的切换时,获得当前帧语音频信号对应的初始高频带信号;

- 15 由宽频带信号向窄频带切换,即前一帧为宽频带信号,当前帧为窄频带信号。预测当前帧窄频带信号对应的初始高频带信号的步骤包括: 根据当前帧窄频带信号预测当前帧语音频信号高频带信号激励信号; 预测当前帧语音频信号高频带信号的LPC系数; 合成预测的高频带激励信号和LPC系数,获得初始高频带信号syn_tmp。

- 20 一个实施例中,可以从窄频带信号中提取基音周期、代数码数和增益等参数,通过变采样,滤波预测到高频带的激励信号;

另一个实施例中,可以通过对窄频带时域信号或窄频带时域激励信号通过上采样、低通,然后取绝对值或取平方等操作来预测高频带激励信号。

- 25 预测高频带信号的LPC系数,可以将历史帧的高频带LPC系数或预先设定好的一系列值作为当前帧LPC系数;也可以对不同的信号类型采用不同的预测方式。

S402: 根据当前帧语音频信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数;

一个实施例中, 包括如下步骤:

S2021: 根据所述当前帧语音频信号的谱倾斜参数和当前帧窄频带与历史帧窄频带信号的相关性, 将当前帧语音频信号分为第一类信号或第二类信号; 一个实施例中, 第一类信号为摩擦音信号, 第二类信号为非摩擦音信号。

一个实施例中, 当谱倾斜参数 $\text{tilt} > 5$ 且相关性参数 cor 小于一给定值时, 将
5 窄频带信号分成摩擦音, 其他的为非摩擦音。其中, 当前帧窄频带信号和历史帧窄频带信号的相关性大小参数 cor 的计算, 可以通过相同某频段信号的能量的大小关系来确定, 也可以通过几个相同频段的能量关系确定, 也可以通过时域信号或时域激励信号的自相关或互相关公式来计算。

S2022: 如果当前帧语音频信号为第一类信号, 则将谱倾斜参数限制到小
10 于等于第一预定值, 获得谱倾斜参数限制值; 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。即当前帧语音频信号的谱倾斜参数小于等于第一预定值时, 保留谱倾斜参数原值作为谱倾斜参数限制值; 当前帧语音频信号的谱倾斜参数大于第一预定值时, 取第一预定值作为谱倾斜参数限制值。

当前帧语音频信号为摩擦音信号时, 时域全局增益参数 gain' 通过以下公式
15 获得:

$$\text{gain}' = \begin{cases} \text{tilt}, & \text{tilt} \leq \partial 1 \\ \partial 1, & \text{tilt} > \partial 1 \end{cases}, \text{ 其中, tilt 为谱倾斜参数, } \partial 1 \text{ 为第一预订值。}$$

S2023: 如果当前帧语音频信号为第二类信号, 则将谱倾斜参数限制到属
于第一区间值, 获得谱倾斜参数限制值; 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。即当前帧语音频信号的谱倾斜参数属于第一区间值
20 时, 保留谱倾斜参数原值作为谱倾斜参数限制值; 当前帧语音频信号的谱倾斜参数大于第一区间值的上限时, 取第一区间值的上限作为谱倾斜参数限制值; 当前帧语音频信号的谱倾斜参数小于第一区间值的下限时, 取第一区间值的下限作为谱倾斜参数限制值。

当前帧语音频信号为非摩擦音信号时, 时域全局增益参数 gain' 通过以下公
25 式获得:

$$\text{gain}' = \begin{cases} \text{tilt}, & \text{tilt} \in [a, b] \\ a, & \text{tilt} < a \\ b, & \text{tilt} > b \end{cases}, \text{ 其中, tilt 为谱倾斜参数, } [a, b] \text{ 为第一区间值。}$$

一个实施例中，获得窄频带信号的谱倾斜参数tilt及当前帧窄频带信号和历史帧窄频带信号的相关性大小参数cor；根据tilt及cor将当前帧信号分为摩擦音及非摩擦音两类，当谱倾斜参数tilt>5且相关性参数cor小于一给定值时，将窄频带信号分成摩擦音，其他的为非摩擦音；将tilt的取值范围限制到

5 0.5≤tilt≤1.0之间作为非摩擦音的时域全局增益参数，将tilt的取值范围限制到tilt≤8.0作为摩擦音的时域全局增益参数。对摩擦音而言，谱倾斜参数可以是大于5的任何值，对非摩擦音而言，可以小于等于5的任何值，也可能大于5，为了保证能将谱倾斜参数tilt能作为预测的的全局增益参数，对tilt的值的范围做限定后作为时域全局增益参数，即当tilt>8时，取tilt=8作为摩擦音信号的时

10 域全局增益参数，当tilt<0.5时，取tilt=0.5或tilt>1.0时，取tilt=1.0作为非摩擦音信号的时域全局增益参数。

S403：利用时域全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号；

一个实施例中，用时域全局增益参数乘于初始高频带信号得到修正的高频带时域信号。

15

另一个实施例中，步骤S403可以包括：

将能量比值和所述时域全局增益参数进行加权处理，得到的加权值作为预测的全局增益参数，其中，能量比值为历史帧高频带时域信号能量与当前帧初始高频带信号能量的比值；

20 利用预测的全局增益参数对所述初始高频带信号进行修正得到修正的高频带时域信号；即用预测的全局增益参数乘于初始高频带信号得到修正的高频带时域信号。

可选的，在步骤S403之前还可以包括：

获得所述初始高频带信号对应的时域包络参数；

25 则利用预测的全局增益参数对所述初始高频带信号进行修正包括：

利用所述时域包络参数和时域全局增益参数对所述初始高频带信号进行修正。

S404：合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

上述实施例中，在宽频带向窄频带切换时，根据谱倾斜参数和帧间相关性获得高频带信号的时域全局增益参数，用窄频带的谱倾斜参数能相对比较准确地估计出窄频带信号和高频带信号间的能量关系，进而更好地估计出高频带信号的能量；用帧间相关性，可以很好地利用窄频带帧间的相关性，估计出高频带信号的帧间相关性，进而在加权求高频带的全局增益时，既可以很好地利用前面真实的信息，又不会引入不好的噪声。利用时域全局增益参数对高频带信号进行修正，使得宽频带和窄频带间高频带部分平稳的过渡，有效地去除了宽频带和窄频带间切换时造成的听觉不舒适感。

与上述方法实施例相关联，本发明还提供一种语音频信号处理装置，该装置可以位于终端设备，网络设备，或测试设备中。所述语音频信号处理装置可以由硬件电路来实现，或者由软件配合硬件来实现。例如，参考图5，由一个处理器调用语音频信号处理装置来实现语音频信号处理。该语音频信号处理装置可以执行上述方法实施例中的各种方法和流程。

参考图6，语音频信号处理装置的一个实施例，包括：

获取单元601，用于当语音频信号出现带宽切换时，获得当前帧语音频信号对应的初始高频带信号；

参数获得单元602，用于获得所述初始高频带信号对应时域全局增益参数；

加权处理单元603，用于将能量比值和该时域全局增益参数进行加权处理，得到的加权值作为预测的全局增益参数；其中，能量比值为历史帧高频带时域信号能量与当前帧初始高频带信号能量的比值；

修正单元604，用于利用预测的全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号；

合成单元605，用于合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

一个实施例中，带宽切换为宽频带信号到窄频带信号的切换，参数获得单元602包括：

全局增益参数获得单元，用于根据当前帧语音频信号的谱倾斜参数、当前

帧语音频信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数。

参考图7, 另一个实施例中, 带宽切换为宽频带信号到窄频带信号的切换, 则参数获得单元602包括:

- 5 时域包络获得单元701, 用于将预设一系列值作为当前帧语音频信号的高频带时域包络参数;

全局增益参数获得单元702, 用于根据当前帧语音频信号的谱倾斜参数、当前帧语音频信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数。

- 10 则修正单元604, 用于利用时域包络参数和预测的全局增益参数对所述初始高频带信号进行修正, 获得修正的高频带时域信号。

参考图8, 进一步的, 全局增益参数获得单元702的一个实施例包括:

- 15 分类单元801, 用于根据所述当前帧语音频信号的谱倾斜参数和当前帧语音频信号与历史帧窄频带信号的相关性, 将当前帧语音频信号分为第一类信号或第二类信号;

第一限制单元802, 如果当前帧语音频信号为第一类信号, 用于将谱倾斜参数限制到小于等于第一预定值, 得到谱倾斜参数限制值, 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数;

- 20 第二限制单元803, 如果当前帧语音频信号为第二类信号, 用于将谱倾斜参数限制到属于第一区间值, 得到谱倾斜参数限制值, 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。

- 25 进一步的, 一个实施例中, 第一类信号为摩擦音信号, 第二类信号为非摩擦音信号; 当谱倾斜参数 $\text{tilt} > 5$ 且相关性参数 cor 小于一给定值时, 将窄频带信号分成摩擦音; 其他的为非摩擦音; 所述第一预定值为8; 第一预定区间为 $[0.5, 1]$ 。

参考图9, 一个实施例中, 获取单元601包括:

激励信号获得单元901, 用于根据当前帧语音频信号预测高频带信号激励信号;

LPC系数获得单元902, 用于预测高频带信号的LPC系数;

生成单元903, 用于合成高频带信号激励信号和高频带信号的LPC系数, 获得所述预测高频带信号。

一个实施例中, 该带宽切换为窄频带信号到宽频带信号的切换, 则该语音频信号处理装置还包括:

- 5 加权因子设置单元, 如果当前音频帧与前一帧语音频信号的窄带信号具有预定相关性时, 用于对前一帧语音频信号对应的所述能量比值的加权因子 α 按一定的步长衰减后的值作为当前音频帧对应的所述能量比值的加权因子, 逐帧衰减直到 α 为到0。

参考图10, 语音频信号处理装置的另一个实施例, 包括:

- 10 预测单元1001, 当语音频信号从宽频带信号到窄频带信号的切换时, 用于获得当前帧语音频信号对应的初始高频带信号;

参数获得单元1002, 用于根据当前帧语音频信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数;

- 15 修正单元1003, 用于利用预测的全局增益参数对所述初始高频带信号进行修正, 获得修正的高频带时域信号;

合成单元1004, 用于合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

参考图8, 参数获得单元1002包括:

- 20 分类单元801, 用于根据所述当前帧语音频信号的谱倾斜参数和当前帧语音频信号与历史帧窄频带信号的相关性, 将当前帧语音频信号分为第一类信号或第二类信号;

- 25 第一限制单元802, 如果当前帧语音频信号为第一类信号, 用于将谱倾斜参数限制到小于等于第一预定值, 得到谱倾斜参数限制值, 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数;

第二限制单元803, 如果当前帧语音频信号为第二类信号, 用于将谱倾斜参数限制到属于第一区间值, 得到谱倾斜参数限制值, 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。

进一步的, 一个实施例中, 第一类信号为摩擦音信号, 第二类信号为非摩

擦音信号；当谱倾斜参数 $\text{tilt} > 5$ 且相关性参数 cor 小于一给定值时，将窄频带信号分成摩擦音；其他的为非摩擦音；其中，第一预定值为8；第一预定区间为 $[0.5, 1]$ 。

可选的，一个实施例中，语音频信号处理装置还包括：

- 5 加权处理单元，用于将能量比值和所述时域全局增益参数进行加权处理，得到的加权值作为预测的全局增益参数，其中，能量比值为历史帧高频带时域信号能量与当前帧初始高频带信号能量的比值；

所述修正单元用于利用预测的全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号。

- 10 另一个实施例中，参数获得单元还用于获得所述初始高频带信号对应的时域包络参数；则修正单元用于利用所述时域包络参数和时域全局增益参数对所述初始高频带信号进行修正。

- 15 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory, ROM）或随机存储记忆体（Random Access Memory, RAM）等。

- 20 以上所述仅为本发明的几个实施例，本领域的技术人员依据申请文件公开的可以对本发明进行各种改动或变型而不脱离本发明的精神和范围。

权 利 要 求

1、一种语音频信号处理方法，其特征在于，包括：

语音频信号从宽频带信号到窄频带信号的切换时，获得当前帧语音频信号对应的初始高频带信号；

5 根据当前帧语音频信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数；

利用所述时域全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号；

合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

10 2、根据权利要求 1 所述的方法，其特征在于，所述根据当前帧语音频信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数包括：

根据所述当前帧语音频信号的谱倾斜参数和当前帧窄频带信号与历史帧窄频带信号的相关性，将当前帧语音频信号分为第一类信号或第二类信号；

15 如果当前帧语音频信号为第一类信号，则将谱倾斜参数限制到小于等于第一预定值，得到谱倾斜参数限制值；

如果当前帧语音频信号为第二类信号，则将谱倾斜参数限制到属于第一区间值，得到谱倾斜参数限制值；

以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。

20 3、根据权利要求 2 所述的方法，其特征在于，所述第一类信号为摩擦音信号，第二类信号为非摩擦音信号；当谱倾斜参数 $\text{tilt} > 5$ 且相关性参数 cor 小于一给定值时，将窄频带信号分成摩擦音；其他的为非摩擦音；所述第一预定值为 8；第一预定区间为 $[0.5, 1]$ 。

4、根据权利要求 1-3 所述的任一方法，其特征在于，利用所述时域全局增益参数对所述初始高频带信号进行修正，获得修正的高频带时域信号包括：

25 将能量比值和所述时域全局增益参数进行加权处理，得到的加权值作为预测的全局增益参数，其中，能量比值为历史帧高频带时域信号能量与当前帧初始高频带信号能量的比值；

利用预测的全局增益参数对所述初始高频带信号进行修正。

30 5、根据权利要求 1-3 所述的任一方法，其特征在于，还包括：

获得所述初始高频带信号对应的时域包络参数;

其中, 利用时域全局增益参数对所述初始高频带信号进行修正包括:

利用所述时域包络参数和时域全局增益参数对所述初始高频带信号进行修正。

5 6、一种语音信号处理方法, 其特征在于, 包括:

当语音信号出现带宽切换时, 获得当前帧语音信号对应的初始高频带信号;

获得所述初始高频带信号时域全局增益参数;

10 将能量比值和所述时域全局增益参数进行加权处理, 得到的加权值作为预测的全局增益参数, 其中, 能量比值为历史帧高频带时域信号能量与当前帧初始高频带信号能量的比值;

利用预测的全局增益参数对所述初始高频带信号进行修正, 获得修正的高频带时域信号;

合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

15 7、根据权利要求 6 所述的方法, 其特征在于, 所述带宽切换为宽频带信号到窄频带信号的切换, 所述获得所述初始高频带信号对应的全局增益参数, 包括:

根据当前帧语音信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数。

20 8、根据权利要求 7 所述的方法, 其特征在于, 所述根据当前帧语音信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数包括:

根据所述当前帧语音信号的谱倾斜参数和当前帧窄频带信号与历史帧窄频带信号的相关性, 将当前帧语音信号分为第一类信号或第二类信号;

25 如果当前帧语音信号为第一类信号, 则将谱倾斜参数限制到小于等于第一预定值, 得到谱倾斜参数限制值;

如果当前帧语音信号为第二类信号, 则将谱倾斜参数限制到属于第一区间值, 得到谱倾斜参数限制值;

以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。

30 9、根据权利要求 8 所述的方法, 其特征在于, 所述第一类信号为摩擦音

信号, 第二类信号为非摩擦音信号; 当谱倾斜参数 $\text{tilt} > 5$ 且相关性参数 cor 小于一给定值时, 将窄频带信号分成摩擦音; 其他的为非摩擦音; 所述第一预定值为 8; 第一预定区间为 $[0.5, 1]$ 。

10、根据权利要求 6 所述的方法, 其特征在于, 所述带宽切换为宽频带信号到窄频带信号的切换, 所述获得当前帧语音频信号对应的初始高频带信号包括:

根据当前帧语音频信号预测高频带激励信号;

预测高频带信号的 LPC 系数;

合成高频带激励信号和高频带信号的 LPC 系数, 获得所述预测高频带信号。

11、根据权利要求 6 所述的方法, 其特征在于, 所述带宽切换为窄频带信号到宽频带信号的切换, 所述方法还包括:

如果当前帧与前一帧语音频信号的窄带信号具有预定相关性时, 则对前一帧语音频信号对应的所述能量比值的加权因子 α 按一定的步长衰减后的值作为当前音频帧对应的所述能量比值的加权因子, 逐帧衰减直到 α 为 0。

12、一种语音频信号处理装置, 其特征在于, 包括:

预测单元, 当语音频信号从宽频带信号到窄频带信号的切换时, 用于获得当前帧语音频信号对应的初始高频带信号;

参数获得单元, 用于根据当前帧语音频信号的谱倾斜参数、当前帧窄频带信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数;

修正单元, 用于利用预测的全局增益参数对所述初始高频带信号进行修正, 获得修正的高频带时域信号;

合成单元, 用于合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

13、根据权利要求 12 所述的装置, 其特征在于, 所述参数获得单元包括:

分类单元, 用于根据所述当前帧语音频信号的谱倾斜参数和当前帧语音频信号与历史帧窄频带信号的相关性, 将当前帧语音频信号分为第一类信号或第二类信号;

第一限制单元, 如果当前帧语音频信号为第一类信号, 用于将谱倾斜参数限制到小于等于第一预定值, 得到谱倾斜参数限制值, 以所述谱倾斜参数限制

值作为高频带信号的时域全局增益参数;

第二限制单元, 如果当前帧语音频信号为第二类信号, 用于将谱倾斜参数限制到属于第一区间值, 得到谱倾斜参数限制值, 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。

- 5 14、根据权利要求 13 所述的装置, 其特征在于, 所述第一类信号为摩擦音信号, 第二类信号为非摩擦音信号; 当谱倾斜参数 $\text{tilt} > 5$ 且相关性参数 cor 小于一给定值时, 将窄频带信号分成摩擦音; 其他的为非摩擦音; 所述第一预定值为 8; 第一预定区间为 $[0.5, 1]$ 。

15、根据权利要求 12-14 所述的任一装置, 其特征在于, 还包括:

- 10 加权处理单元, 用于将能量比值和所述时域全局增益参数进行加权处理, 得到的加权值作为预测的全局增益参数, 其中, 能量比值为历史帧高频带时域信号能量与当前帧初始高频带信号能量的比值;

所述修正单元用于利用预测的全局增益参数对所述初始高频带信号进行修正, 获得修正的高频带时域信号。

- 15 16、根据权利要求 12-14 所述的任一装置, 其特征在于,
所述参数获得单元还用于获得所述初始高频带信号对应的时域包络参数;
所述修正单元用于利用所述时域包络参数和时域全局增益参数对所述初始高频带信号进行修正。

17、一种语音频信号处理装置, 其特征在于, 包括:

- 20 获取单元, 用于当语音频信号出现带宽切换时, 获得当前帧语音频信号对应的初始高频带信号;

参数获得单元, 用于获得所述初始高频带信号对应的时域全局增益参数;

加权处理单元, 用于将能量比值和所述时域全局增益参数进行加权处理, 得到的加权值作为预测的全局增益参数; 其中, 能量比值为历史帧高频带时域
25 信号能量与当前帧初始高频带信号能量的比值;

修正单元, 用于利用预测的全局增益参数对所述初始高频带信号进行修正, 获得修正的高频带时域信号;

合成单元, 用于合成当前帧的窄频带时域信号和所述修正的高频带时域信号并输出。

- 30 18、根据权利要求 17 所述的装置, 其特征在于, 所述带宽切换为宽频带

信号到窄频带信号的切换, 所述参数获得单元包括:

全局增益参数获得单元, 用于根据当前帧语音频信号的谱倾斜参数、当前帧语音频信号与历史帧窄频带信号的相关性获得所述高频带信号的时域全局增益参数。

- 5 19、根据权利要求 18 所述的装置, 其特征在于, 所述全局增益参数获得单元包括:

分类单元, 用于根据所述当前帧语音频信号的谱倾斜参数和当前帧语音频信号与历史帧窄频带信号的相关性, 将当前帧语音频信号分为第一类信号或第二类信号;

- 10 第一限制单元, 如果当前帧语音频信号为第一类信号, 用于将谱倾斜参数限制到小于等于第一预定值, 得到谱倾斜参数限制值, 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数;

- 15 第二限制单元, 如果当前帧语音频信号为第二类信号, 用于将谱倾斜参数限制到属于第一区间值, 得到谱倾斜参数限制值, 以所述谱倾斜参数限制值作为高频带信号的时域全局增益参数。

20、根据权利要求 19 所述的装置, 其特征在于, 所述第一类信号为摩擦音信号, 第二类信号为非摩擦音信号; 当谱倾斜参数 $\text{tilt} > 5$ 且相关性参数 cor 小于一给定值时, 将窄频带信号分成摩擦音; 其他的为非摩擦音; 所述第一预定值为 8; 第一预定区间为 $[0.5, 1]$ 。

- 20 21、根据权利要求 17-20 所述的任一装置, 其特征在于, 所述带宽切换为窄频带信号到宽频带信号的切换, 所述装置还包括:

时域包络获得单元, 用于将预设一系列值作为当前帧语音频信号的高频带时域包络参数;

- 25 所述修正单元, 用于利用时域包络参数和预测的全局增益参数对所述初始高频带信号进行修正, 获得修正的高频带时域信号。

22、根据权利要求 17-20 所述的任一装置, 其特征在于, 所述获取单元包括:

激励信号获得单元, 用于根据当前帧语音频信号预测高频带信号激励信号;

- 30 LPC 系数获得单元, 用于预测高频带信号的 LPC 系数;

合成单元, 用于合成高频带信号激励信号和高频带信号的 LPC 系数, 获得所述预测高频带信号。

23、根据权利要求 17-20 所述的任一装置, 其特征在于, 所述带宽切换为窄频带信号到宽频带信号的切换, 所述装置还包括:

- 5 加权因子设置单元, 如果当前音频帧与前一帧语音频信号的窄带信号具有预定相关性时, 用于对前一帧语音频信号对应的所述能量比值的加权因子 α 按一定的步长衰减后的值作为当前音频帧对应的所述能量比值的加权因子, 逐帧衰减直到 α 为 0。