



(12) 发明专利

(10) 授权公告号 CN 101847404 B

(45) 授权公告日 2012.08.22

(21) 申请号 201010129147.2

(22) 申请日 2010.03.18

(73) 专利权人 北京天籁传音数字技术有限公司
地址 100098 北京市海淀区大钟寺13号院1号华杰大厦9B15

(72) 发明人 李靛 潘兴德

(74) 专利代理机构 北京德琦知识产权代理有限公司 11018

代理人 谢安昆 宋志强

US 6463406 B1, 2002.10.08,

GB 2398980 A, 2004.09.01,

CN 1256586 A, 2000.06.14,

JP 2003036085 A, 2003.02.07,

CN 101354889 A, 2009.01.28,

蔡杰. 一种改进的时域音频变调方法及其软件实现. 《电声技术》. 2006, (第09期),

李力利. 数字音频变调算法的研究. 《电声技术》. 2003, (第06期),

审查员 杨艳兰

(51) Int. Cl.

G10L 11/06 (2006.01)

G10L 13/08 (2006.01)

G10L 19/14 (2006.01)

(56) 对比文件

CN 101075429 A, 2007.11.21,

US 4797926 A, 1989.01.10,

EP 0051342 B1, 1986.01.29,

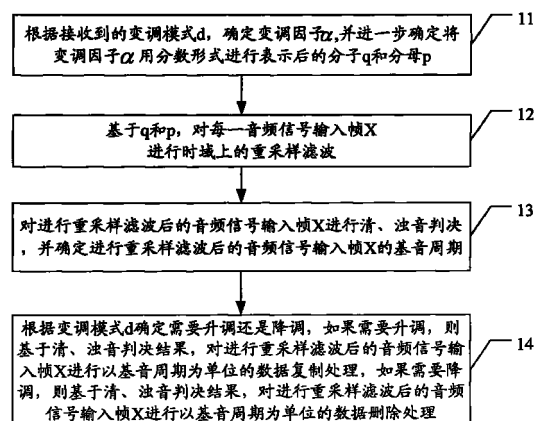
权利要求书 6 页 说明书 9 页 附图 4 页

(54) 发明名称

一种实现音频变调的方法和装置

(57) 摘要

本发明提供了一种实现音频变调的方法和装置:根据接收到的变调模式 d , 确定变调因子 α , 以及变调因子 α 用分数形式进行表示后的分子 q 和分母 p ; 针对每一音频信号输入帧 X : 基于 q 和 p , 对音频信号输入帧 X 进行时域上的重采样滤波; 对进行重采样滤波后的音频信号输入帧 X 进行清、浊音判决, 并确定进行重采样滤波后的音频信号输入帧 X 的基音周期; 根据变调模式 d 确定需要升调还是降调, 如果需要升调, 则基于清、浊音判决结果, 对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据复制处理, 如果需要降调, 则基于清、浊音判决结果, 对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据删除处理。应用本发明所述方案, 能够有效地实现升调或降调。



1. 一种实现音频变调的方法,其特征在于,该方法包括:

根据接收到的变调模式 d,确定变调因子 α ,并进一步确定将所述变调因子 α 用分数形式进行表示后的分子 q 和分母 p;

针对每一音频信号输入帧 X,分别进行以下处理:

基于所述 q 和 p,对所述音频信号输入帧 X 进行时域上的重采样滤波;

对进行重采样滤波后的音频信号输入帧 X 进行清、浊音判决,并确定所述进行重采样滤波后的音频信号输入帧 X 的基音周期;

根据所述变调模式 d 确定需要升调还是降调,如果需要升调,则基于清、浊音判决结果,对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据复制处理,如果需要降调,则基于清、浊音判决结果,对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据删除处理;

其中,所述基于所述 q 和 p,对所述音频信号输入帧 X 进行时域上的重采样滤波包括:

对所述音频信号输入帧 X 进行 p 倍的线性插值;

将进行线性插值后的音频信号输入帧 X 通过低通滤波器进行低通滤波,所述低通滤波器的增益为 p,截止频率为 π/q 和 π/p 中的较小值;

将进行低通滤波后的音频信号输入帧 X 进行 q 倍的采样。

2. 根据权利要求 1 所述的方法,其特征在于,所述对进行重采样滤波后的音频信号输入帧 X 进行清、浊音判决,并确定基音周期包括:

$$\text{计算 } \rho(\tau) = \frac{\sum_{n=0}^{N-1} s(n)s(n-\tau)}{\sqrt{\sum_{n=0}^{N-1} s^2(n) \sum_{n=0}^{N-1} s^2(n-\tau)}}; \text{ 其中, } s(n) \text{ 表示音频信号强度, } N \text{ 表示进行重采样滤波后的音频信号输入帧 X 的长度, } \tau \text{ 表示延迟参数,为正整数,取值范围预先设定};$$

波后的音频信号输入帧 X 的长度, τ 表示延迟参数,为正整数,取值范围预先设定;

确定不同的延迟参数 τ 对应的 $\rho(\tau)$ 中的最大值,并将所述最大值与预先设定的阈值进行比较,如果所述最大值小于所述阈值,则确定所述进行重采样滤波后的音频信号输入帧 X 为清音帧,否则,为浊音帧,并将所述最大值对应的延迟参数 τ 确定为所述进行重采样滤波后的音频信号输入帧 X 的基音周期。

3. 根据权利要求 1~2 中任一项所述的方法,其特征在于,所述根据所述变调模式 d 确定需要升调还是降调包括:

如果所述变调模式 d 的取值大于 0,则确定需要升调,如果所述变调模式 d 的取值小于 0,则确定需要降调。

4. 根据权利要求 1~2 中任一项所述的方法,其特征在于,所述基于清、浊音判决结果,对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据复制处理包括:

A1、将所述进行重采样滤波后的音频信号输入帧 X 平均分为 4 个子块,并计算每个子块的能量,比较每两个相邻子块的能量大小,如果存在突变,即存在后一子块的能量为与其相邻的前一子块的能量的 12 倍的情况,则确定所述进行重采样滤波后的音频信号输入帧 X 为突变帧,并记录下发生突变的子块位置;

A2、将所述进行重采样滤波后的音频信号输入帧 X 拼接在所保存的超出前一音频信号

输出帧长度的数据之后,所述前一音频信号输出帧为紧邻所述音频信号输入帧 X 的前一音频信号输入帧对应的音频信号输出帧;

A3、计算需要复制的数据的长度 CL;

A4、确定所述 CL 是否大于 0,如果是,则执行步骤 A5,否则,执行步骤 A11;

A5、如果所述进行重采样滤波后的音频信号输入帧 X 为浊音帧,则根据所述 CL 以及基音周期,计算需复制基音周期的个数 R,然后执行步骤 A6,如果为清音帧,则执行步骤 A10;

A6、如果所述进行重采样滤波后的音频信号输入帧 X 为突变帧,且发生突变的子块位置为第 4 子块,则执行步骤 A7,如果发生突变的子块位置不是第 4 子块,则执行步骤 A8;

A7、以第 3 子块的结束位置作为复制起始点,将位于所述复制起始点之前,且紧邻所述复制起始点的基音周期数据作为复制源,进行 R 次复制,并将第 4 子块的数据拼接在复制的基音周期数据之后,之后,将本步骤中拼接后的数据进一步拼接在步骤 A2 中拼接后的数据之后,然后执行步骤 A9;

A8、以第 4 子块的结束位置作为复制起始点,将位于所述复制起始点之前,且紧邻所述复制起始点的基音周期数据作为复制源,进行 R 次复制,并将复制的基音周期数据拼接在步骤 A2 中拼接后的数据之后,然后执行步骤 A9;

A9、计算当前拼接出的数据超出音频信号输出帧 X' 的长度 OL,然后执行步骤 A12;所述音频信号输出帧 X' 表示所述音频信号输入帧 X 对应的音频信号输出帧;

A10、以第 4 子块的结束位置作为复制起始点,复制所述复制起始点之前,且紧邻所述复制起始点的长度为 CL 的数据,并将复制的长度为 CL 的数据拼接在步骤 A2 中拼接后的数据之后,将 OL 设置为 0,之后执行步骤 A13;

A11、将 OL 设置为 CL 的相反数,然后执行步骤 A12;

A12、将超出音频信号输出帧 X' 长度的数据进行保存,然后执行步骤 A13;

A13、输出升调后的音频信号输出帧 X'。

5. 根据权利要求 4 所述的方法,其特征在于,所述计算需要复制的数据的长度 CL 包括:

$CL = OutLen - (OL + InLen)$,其中,InLen 表示进行重采样滤波后的音频信号输入帧 X 的长度,OutLen 表示音频信号输出帧 X' 的长度, $OutLen = \lfloor InLen * \alpha \rfloor$,符号 $\lfloor \cdot \rfloor$ 表示向下取整,OL 表示所保存的超出前一音频信号输出帧长度的数据的长度,初始值设置为 0。

6. 根据权利要求 4 所述的方法,其特征在于,所述根据所述 CL 以及基音周期,计算需复制基音周期的个数 R 包括:

$R = \lfloor CL / pitch \rfloor + 1$,其中,所述 pitch 表示基音周期。

7. 根据权利要求 4 所述的方法,其特征在于,所述计算当前拼接出的数据超出音频信号输出帧 X' 数据的长度 OL 包括:

$OL = R * pitch - CL$,其中,所述 pitch 表示基音周期。

8. 根据权利要求 1~2 中任一项所述的方法,其特征在于,所述基于清、浊音判决结果,对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据删除处理包括:

B1、将所述进行重采样滤波后的音频信号输入帧 X 平均分为 4 个子块,并计算每个子块

的能量,比较每两个相邻子块的能量大小,如果存在突变,即存在后一子块的能量为与其相邻的前一子块的能量的 12 倍的情况,则确定所述进行重采样滤波后的音频信号输入帧 X 为突变帧,并记录下发生突变的子块位置;

B2、计算需要删除的数据的长度 DL;

B3、如果所述进行重采样滤波后的音频信号输入帧 X 为浊音帧,则根据所述 DL 以及基音周期,计算需复制基音周期的个数 R,然后执行步骤 B4,如果为清音帧,则执行步骤 B7;

B4、如果所述进行重采样滤波后的音频信号输入帧 X 不是突变帧,则执行步骤 B5,否则,执行步骤 B6;

B5、将距离所述进行重采样滤波后的音频信号输入帧 X 的起始点长度为已保存的 SL 的位置作为音频信号输出帧 X' 的起始点,输出长度为 OutLen 的数据,所述 OutLen 表示所述音频信号输入帧 X 对应的音频信号输出帧 X' 的长度,删除未输出的其余数据,然后计算新的 SL,利用新计算出的 SL 更新已保存的 SL,之后执行步骤 B8;

B6、基于所述 DL 和发生突变的子块位置,确定音频信号输出帧 X' 的起始点,输出长度为 OutLen 的数据,删除未输出的其余数据,并将已保存的 SL 设置为 0,之后执行步骤 B8;

B7、将所述进行重采样滤波后的音频信号输入帧 X 的起始点作为音频信号输出帧 X' 的起始点,输出长度为 OutLen 的数据,删除未输出的其余数据,并将已保存的 SL 设置为 0,之后执行步骤 B8;

B8、利用低通滤波器对音频信号输出帧 X' 进行低通滤波,所述低通滤波器的截止频率为音频信号输入帧 X 的采样频率的 $0.5 * q/p$ 倍。

9. 根据权利要求 8 所述的方法,其特征在于,所述计算需要删除的数据的长度 DL 包括:

$DL = InLen - OutLen$,其中,所述 InLen 表示进行重采样滤波后的音频信号输入帧 X 的长度,所述 OutLen 表示音频信号输出帧 X' 的长度, $OutLen = \lfloor InLen * \alpha \rfloor$,符号 $\lfloor \rfloor$ 表示向下取整。

10. 根据权利要求 8 所述的方法,其特征在于,所述根据所述 DL 以及基音周期,计算需复制基音周期的个数 R 包括:

$R = \lfloor DL / pitch \rfloor + 1$,其中,所述 pitch 表示基音周期。

11. 根据权利要求 8 所述的方法,其特征在于,所述计算新的 SL 包括:

$SL = old_SL + R * pitch - DL$,其中,所述 old_SL 表示已保存的 SL,初始值设置为 0。

12. 根据权利要求 8 所述的方法,其特征在于,所述基于所述 DL 和发生突变的子块位置,确定音频信号输出帧 X' 的起始点包括:

如果发生突变的子块位置距离所述进行重采样滤波后的音频信号输入帧 X 的起始点的距离小于 DL,则以所述进行重采样滤波后的音频信号输入帧 X 的起始点作为音频信号输出帧 X' 的起始点,否则,以距离所述进行重采样滤波后的音频信号输入帧 X 的起始点 DL 的位置作为音频信号输出帧 X' 的起始点。

13. 一种实现音频变调的装置,其特征在于,包括:

确定单元,用于根据接收到的变调模式 d,确定变调因子 α ,并进一步确定将所述变调因子 α 用分数形式进行表示后的分子 q 和分母 p;

处理单元,用于针对每一音频信号输入帧 X,分别进行以下处理:

基于所述 q 和 p , 对所述音频信号输入帧 X 进行时域上的重采样滤波, 包括: 对所述音频信号输入帧 X 进行 p 倍的线性插值, 将进行线性插值后的音频信号输入帧 X 通过低通滤波器进行低通滤波, 所述低通滤波器的增益为 p , 截止频率为 π/q 和 π/p 中的较小值, 将进行低通滤波后的音频信号输入帧 X 进行 q 倍的采样;

对进行重采样滤波后的音频信号输入帧 X 进行清、浊音判决, 并确定所述进行重采样滤波后的音频信号输入帧 X 的基音周期;

根据所述变调模式 d 确定需要升调还是降调, 如果需要升调, 则基于清、浊音判决结果, 对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据复制处理, 如果需要降调, 则基于清、浊音判决结果, 对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据删除处理。

14. 根据权利要求 13 所述的装置, 其特征在于, 所述处理单元包括:

重采样滤波子单元, 用于对所述音频信号输入帧 X 进行 p 倍的线性插值, 将进行线性插值后的音频信号输入帧 X 通过低通滤波器进行低通滤波, 所述低通滤波器的增益为 p , 截止频率为 π/q 和 π/p 中的较小值, 将进行低通滤波后的音频信号输入帧 X 进行 q 倍的采样;

判决子单元, 用于计算 $\rho(\tau) = \frac{\sum_{n=0}^{N-1} s(n)s(n-\tau)}{\sqrt{\sum_{n=0}^{N-1} s^2(n) \sum_{n=0}^{N-1} s^2(n-\tau)}}$; 其中, $s(n)$ 表示音频信号强度, N

表示进行重采样滤波后的音频信号输入帧 X 的长度, τ 表示延迟参数, 为正整数, 取值范围预先设定; 确定不同的延迟参数 τ 对应的 $\rho(\tau)$ 中的最大值, 并将所述最大值与预先设定的阈值进行比较, 如果所述最大值小于所述阈值, 则确定所述进行重采样滤波后的音频信号输入帧 X 为清音帧, 否则, 为浊音帧, 并将所述最大值对应的延迟参数 τ 确定为所述进行重采样滤波后的音频信号输入帧 X 的基音周期;

确定子单元, 用于根据所述变调模式 d 确定需要升调还是降调, 如果所述变调模式 d 的取值大于 0, 则确定需要升调, 如果所述变调模式 d 的取值小于 0, 则确定需要降调;

变调子单元, 用于当确定需要升调时, 基于清、浊音判决结果, 对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据复制处理, 当确定需要降调时, 基于清、浊音判决结果, 对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据删除处理。

15. 根据权利要求 14 所述的装置, 其特征在于, 所述变调子单元包括:

升调子单元, 用于基于清、浊音判决结果, 对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据复制处理, 包括:

A1、将所述进行重采样滤波后的音频信号输入帧 X 平均分为 4 个子块, 并计算每个子块的能量, 比较每两个相邻子块的能量大小, 如果存在突变, 即存在后一子块的能量为与其相邻的前一子块的能量 12 倍的情况, 则确定所述进行重采样滤波后的音频信号输入帧 X 为突变帧, 并记录下发生突变的子块位置;

A2、将所述进行重采样滤波后的音频信号输入帧 X 拼接在所保存的超出前一音频信号输出帧长度的数据之后, 所述前一音频信号输出帧为紧邻所述音频信号输入帧 X 的前一音频信号输入帧对应的音频信号输出帧;

A3、计算需要复制的数据的长度 CL；

A4、确定所述 CL 是否大于 0，如果是，则执行步骤 A5，否则，执行步骤 A11；

A5、如果所述进行重采样滤波后的音频信号输入帧 X 为浊音帧，则根据所述 CL 以及基音周期，计算需复制基音周期的个数 R，然后执行步骤 A6，如果为清音帧，则执行步骤 A10；

A6、如果所述进行重采样滤波后的音频信号输入帧 X 为突变帧，且发生突变的子块位置为第 4 子块，则执行步骤 A7，如果发生突变的子块位置不是第 4 子块，则执行步骤 A8；

A7、以第 3 子块的结束位置作为复制起始点，将位于所述复制起始点之前，且紧邻所述复制起始点的基音周期数据作为复制源，进行 R 次复制，并将第 4 子块的数据拼接在复制的基音周期数据之后，之后，将本步骤中拼接后的数据进一步拼接在步骤 A2 中拼接后的数据之后，然后执行步骤 A9；

A8、以第 4 子块的结束位置作为复制起始点，将位于所述复制起始点之前，且紧邻所述复制起始点的基音周期数据作为复制源，进行 R 次复制，并将复制的基音周期数据拼接在步骤 A2 中拼接后的数据之后，然后执行步骤 A9；

A9、计算当前拼接出的数据超出音频信号输出帧 X' 的长度 OL，然后执行步骤 A12；所述音频信号输出帧 X' 表示所述音频信号输入帧 X 对应的音频信号输出帧；

A10、以第 4 子块的结束位置作为复制起始点，复制所述复制起始点之前，且紧邻所述复制起始点的长度为 CL 的数据，并将复制的长度为 CL 的数据拼接在步骤 A2 中拼接后的数据之后，降 OL 设置为 0，之后执行步骤 A13；

A11、将 OL 设置为 CL 的相反数，然后执行步骤 A12；

A12、将超出音频信号输出帧 X' 长度的数据进行保存，然后执行步骤 A13；

A13、输出升调后的音频信号输出帧 X'；

降调子单元，用于基于清、浊音判决结果，对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据删除处理，包括：

B1、将所述进行重采样滤波后的音频信号输入帧 X 平均分为 4 个子块，并计算每个子块的能量，比较每两个相邻子块的能量大小，如果存在突变，即存在后一子块的能量为与其相邻的前一子块的能量 12 倍的情况，则确定所述进行重采样滤波后的音频信号输入帧 X 为突变帧，并记录下发生突变的子块位置；

B2、计算需要删除的数据的长度 DL；

B3、如果所述进行重采样滤波后的音频信号输入帧 X 为浊音帧，则根据所述 DL 以及基音周期，计算需复制基音周期的个数 R，然后执行步骤 B4，如果为清音帧，则执行步骤 B7；

B4、如果所述进行重采样滤波后的音频信号输入帧 X 不是突变帧，则执行步骤 B5，否则，执行步骤 B6；

B5、将距离所述进行重采样滤波后的音频信号输入帧 X 的起始点长度为已知的 SL 的位置作为音频信号输出帧 X' 的起始点，输出长度为 OutLen 的数据，所述 OutLen 表示所述音频信号输入帧 X 对应的音频信号输出帧 X' 的长度，删除未输出的其余数据，然后计算新的 SL，利用新计算的 SL 更新已保存的 SL，之后执行步骤 B8；

B6、基于所述 DL 和发生突变的子块位置，确定音频信号输出帧 X' 的起始点，输出长度为 OutLen 的数据，删除未输出的其余数据，并将已保存的 SL 设置为 0，之后执行步骤 B8；

B7、将所述进行重采样滤波后的音频信号输入帧 X 的起始点作为音频信号输出帧 X' 的

起始点,输出长度为 OutLen 的数据,删除未输出的其余数据,并将已保存的 SL 设置为 0,之后执行步骤 B8;

B8、利用低通滤波器对音频信号输出帧 X' 进行低通滤波,所述低通滤波器的截止频率为音频信号输入帧 X 的采样频率的 $0.5 * q/p$ 倍。

一种实现音频变调的方法和装置

技术领域

[0001] 本发明涉及音频处理技术领域,特别涉及一种实现音频变调的方法和装置。

背景技术

[0002] 众所周知,声音的基本要素主要有:音调、音强和音色。音频变调是指在保持音频信号的播放时间不变的前提下,改变声音的音调,将音调升高或降低。音频变调可改变声音的特征,是音频处理中的一项热门技术,在许多方面都存在着广泛的应用,比如音乐合成、配音、网络视频聊天以及保密电话等等,其中一种典型的应用当属卡拉 OK 机,其伴唱系统中即可运用音频变调技术,通过调整伴奏音乐的音调以适合演唱者的嗓音。可以预见,随着互联网和多媒体技术的迅速发展,对音频变调技术的需要将会越来越多。

[0003] 现有乐音体系中,根据振动频率的不同,可将音调划分成不同的音级,世界上普遍采用的划分方式为 12 平均律。12 平均律是指将一个纯八度划分为 12 个均等的频率成分,即 12 个半音,各相邻半音的振动频率相差 $2^{1/12}$,一个纯八度内的各半音间的振动频率最大相差 2 倍。如果将各频率成分的振动频率升高 $2^{1/12}$ 倍,则相当于将音调升高了一个半音,反之,如果将各频率成分的振动频率降低 $2^{1/12}$ 倍,则相当于将音调降低了一个半音。

[0004] 基于上述介绍,假设某音频信号的原始频率为 f ,经过音频变调后的频率为 f' ,那么则有:

[0005] $f' = f * 2^{d/12}$, $d = \pm 1, \pm 2, \dots, \pm 12$;

[0006] 其中, d 称为变调模式,当 $d > 0$ 时表示升调,反之表示降调。 d 的取值每增大或减小 1,表示将音调升高或降低了一个半音。

[0007] 基于上述 $f' = f * 2^{d/12}$,可推导出变调因子 $\alpha = 2^{d/12} = \frac{f'}{f}$ 。通常将变调因子 α 表示为分数形式,即 $\alpha = \frac{f'}{f} = \frac{q}{p}$, q 和 p 作为变调因子 α 的分子和分母,均为整数,且两者之间无公约数(除 1 以外)。

[0008] 如前所述,随着互联网和多媒体技术的迅速发展,对音频变调技术的需要将会越来越多,但是,现有技术中还没有一种有效的音频变调方式,从而给用户的使用带来了很大的不便。

[0009] 发明内容

[0010] 有鉴于此,本发明的主要目的在于提供一种实现音频变调的方法,能够有效地实现升调或降调。

[0011] 本发明的另一目的在于提供一种实现音频变调的装置,能够有效地实现升调或降调。

[0012] 为达到上述目的,本发明的技术方案是这样实现的:

[0013] 一种实现音频变调的方法,该方法包括:

[0014] 根据接收到的变调模式 d ,确定变调因子 α ,并进一步确定将所述变调因子 α 用

分数形式进行表示后的分子 q 和分母 p ；

[0015] 针对每一音频信号输入帧 X ，分别进行以下处理：

[0016] 基于所述 q 和 p ，对所述音频信号输入帧 X 进行时域上的重采样滤波；

[0017] 对进行重采样滤波后的音频信号输入帧 X 进行清、浊音判决，并确定所述进行重采样滤波后的音频信号输入帧 X 的基音周期；

[0018] 根据所述变调模式 d 确定需要升调还是降调，如果需要升调，则基于清、浊音判决结果，对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据复制处理，如果需要降调，则基于清、浊音判决结果，对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据删除处理；

[0019] 其中，所述基于所述 q 和 p ，对所述音频信号输入帧 X 进行时域上的重采样滤波包括：

[0020] 对所述音频信号输入帧 X 进行 p 倍的线性插值；

[0021] 将进行线性插值后的音频信号输入帧 X 通过低通滤波器进行低通滤波，所述低通滤波器的增益为 p ，截止频率为 π/q 和 π/p 中的较小值；

[0022] 将进行低通滤波后的音频信号输入帧 X 进行 q 倍的采样。

[0023] 一种实现音频变调的装置，包括：

[0024] 确定单元，用于根据接收到的变调模式 d ，确定变调因子 α ，并进一步确定将所述变调因子 α 用分数形式进行表示后的分子 q 和分母 p ；

[0025] 处理单元，用于针对每一音频信号输入帧 X ，分别进行以下处理：

[0026] 基于所述 q 和 p ，对所述音频信号输入帧 X 进行时域上的重采样滤波，包括：对所述音频信号输入帧 X 进行 p 倍的线性插值，将进行线性插值后的音频信号输入帧 X 通过低通滤波器进行低通滤波，所述低通滤波器的增益为 p ，截止频率为 π/q 和 π/p 中的较小值，将进行低通滤波后的音频信号输入帧 X 进行 q 倍的采样；

[0027] 对进行重采样滤波后的音频信号输入帧 X 进行清、浊音判决，并确定所述进行重采样滤波后的音频信号输入帧 X 的基音周期；

[0028] 根据所述变调模式 d 确定需要升调还是降调，如果需要升调，则基于清、浊音判决结果，对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据复制处理，如果需要降调，则基于清、浊音判决结果，对所述进行重采样滤波后的音频信号输入帧 X 进行以所述基音周期为单位的数据删除处理。

[0029] 可见，采用本发明的技术方案，通过改变采样频率来改变音频信号的频率，并对改变频率后的音频信号进行数据复制和数据删除等处理，从而不但保持了音频信号的播放时间不变，而且有效地实现了升调和降调。

[0030] 附图说明

[0031] 图 1 为本发明实现音频变调的方法实施例的流程图。

[0032] 图 2 为本发明方法实施例中的重采样滤波过程示意图。

[0033] 图 3 为本发明方法实施例中的升调过程示意图。

[0034] 图 4 为本发明方法实施例中的降调过程示意图。

[0035] 图 5 为本发明实现音频变调的装置实施例的组成结构示意图。

[0036] 具体实施方式

[0037] 针对现有技术中存在的问题,本发明中提出一种实现音频变调的方案,通过改变采样频率来改变音频信号的频率,并对改变频率后的音频信号进行数据复制和数据删除等处理,从而不但保持了音频信号的播放时间不变,而且有效地实现了升调和降调。

[0038] 图1为本发明实现音频变调的方法实施例的流程图。如图1所示,包括以下步骤:

[0039] 步骤11:根据接收到的变调模式 d ,确定变调因子 α ,并进一步确定将变调因子 α 用分数形式进行表示后的分子 q 和分母 p 。

[0040] 本步骤中,用户可根据自身需要输入变调模式 d ,比如,+2或-5等;然后,系统通过查询表一,可确定出用户输入的变调模式 d 对应的变调因子 α ,以及将变调因子 α 用分数形式进行表示后的分子 q 和分母 p 。

[0041]

d	α	q	p
-12	0.5	1	2
-11	0.529732	1461	2758
-10	0.561231	857	1527
-9	0.594604	1785	3002
-8	0.629961	635	1008
-7	0.667420	1477	2213
-6	0.707107	985	1393
-5	0.749154	2434	3249
-4	0.793701	504	635
-3	0.840896	1501	1785
-2	0.890899	1527	1714
-1	0.943874	1379	1461
0	1.0	1	1
1	1.05946	1461	1379
2	1.12246	1714	1527
3	1.18921	1785	1501
4	1.25992	635	504
5	1.33484	3249	2434
6	1.41421	1393	985
7	1.49831	2213	1477
8	1.58740	1008	635
9	1.68179	3002	1785
10	1.78180	1527	857
11	1.88775	2758	1461
12	2.0	2	1

[0042] 表1变调模式 d 、变调因子 α ,以及分子 q 和分母 p 的对应关系

[0043] 表1中的各取值均为本领域公认的经验值。当变调模式 d 取值为0时,表示不进行音频变调,即既不升调也不降调。

[0044] 步骤12:基于 q 和 p ,对每一音频信号输入帧 X 进行时域上的重采样滤波。

[0045] 图2为本发明方法实施例中的重采样滤波过程示意图。如图2所示,首先,对音频信号输入帧 X 进行 p 倍的线性插值;然后,将进行线性插值后的音频信号输入帧 X 通过低通滤波器进行低通滤波,低通滤波器的增益为 p ,截止频率为 π/q 和 π/p 中的较小值,之后,将进行低通滤波后的音频信号输入帧 X 进行 q 倍的采样。

[0046] 经过本步骤的处理后,进行重采样滤波后的音频信号输入帧 X 的采样频率将变为音频信号输入帧 X 的采样频率的 q/p 倍。

[0047] 步骤 13 :对进行重采样滤波后的音频信号输入帧 X 进行清、浊音判决,并确定进行重采样滤波后的音频信号输入帧 X 的基音周期。

[0048] 人类的发音过程通常有三类不同的激励方式,因而能产生三类不同的声音,即浊音、清音和爆破音。当气流通过声门时声带的张力刚好使声带发生较低频率的张弛振动,形成准周期性的空气脉冲时,这些空气脉冲便激励声道产生“浊音”;如果声道中某处面积很小,气流高速冲过此处时产生湍流,当气流速度与横截面积之比大于某个门限(临界速度)时,便产生“摩擦音”,即“清音”;如果声道某处完全闭合建立起气压,然后突然释放而产生的声音就是爆破音。通常遇到最多的就是浊音和清音。

[0049] 基音是指产生浊音时声带振动所引起的周期性,基音周期的倒数称为基音频率。

[0050] 本步骤中,计算 $\rho(\tau) = \frac{\sum_{n=0}^{N-1} s(n)s(n-\tau)}{\sqrt{\sum_{n=0}^{N-1} s^2(n) \sum_{n=0}^{N-1} s^2(n-\tau)}}$;其中, s(n) 表示音频信号强度, N 表

示进行重采样滤波后的音频信号输入帧 X 的长度, τ 表示延迟参数,为正整数,取值范围预先设定,通常,在 8kHz 的采样频率下, τ 的取值范围为 [20, 140], 具体确定方式为现有技术,不再赘述。

[0051] 确定不同的 τ 对应的 $\rho(\tau)$ 中的最大值,并将该最大值与预先设定的阈值(经验值)进行比较,如果该最大值小于阈值,则确定进行重采样滤波后的音频信号输入帧 X 为清音帧,否则,为浊音帧,并将该最大值对应的延迟参数 τ 确定为进行重采样滤波后的音频信号输入帧 X 的基音周期。

[0052] 步骤 14 :根据变调模式 d 确定需要升调还是降调,如果需要升调,则基于清、浊音判决结果,对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据复制处理,如果需要降调,则基于清、浊音判决结果,对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据删除处理。

[0053] 如果步骤 11 中用户输入的变调模式 d 的取值大于 0,则说明需要升调,如果变调模式 d 的取值小于 0,则说明需要降调。

[0054] 根据需要升调或降调的不同,后续采用不同的处理方式,下面结合附图,对升调和降调过程的具体实现分别进行详细说明。

[0055] 图 3 为本发明方法实施例中的升调过程示意图。如图 3 所示,包括:

[0056] 步骤 31 :将进行重采样滤波后的音频信号输入帧 X 平均分为 4 个子块,确定进行重采样滤波后的音频信号输入帧 X 是否为突变帧,并记录下发生突变的子块位置。

[0057] 本步骤中,计算每个子块的能量,并比较每两个相邻子块的能量大小,如果存在突变,即存在后一子块的能量为与其相邻的前一子块的能量 12 倍的情况,则确定进行重采样滤波后的音频信号输入帧 X 为突变帧,并记录下发生突变的子块位置,比如第 4 子块的能量是第 3 子块的能量 12 倍,则记录下发生突变的子块位置为第 4 子块。通常一帧中只会发生一次突变。

[0058] 如何计算每个子块的能量为现有技术,不再赘述。

[0059] 步骤 32 :将进行重采样滤波后的音频信号输入帧 X 拼接在所保存的超出前一音频信号输出帧长度的数据之后。

[0060] 前一音频信号输出帧是指紧邻音频信号输入帧 X 的前一音频信号输入帧对应的经过变调后的音频信号输出帧。

[0061] 步骤 33 : 计算需要复制的数据的长度 CL。

[0062] $CL = OutLen - (OL + InLen)$, 其中, InLen 表示进行重采样滤波后的音频信号输入帧 X 的长度, OutLen 表示音频信号输入帧 X 对应的音频信号输出帧 X' 的长度, $OutLen = \lfloor InLen * \alpha \rfloor$, 符号 $\lfloor \rfloor$ 表示向下取整, OL 表示所保存的超出前一音频信号输出帧长度的数据的长度, 初始值设置为 0。

[0063] 步骤 34 : 确定 CL 是否大于 0, 如果是, 则执行步骤 35, 否则, 执行步骤 311。

[0064] 步骤 35 : 如果进行重采样滤波后的音频信号输入帧 X 为浊音帧, 则根据 CL 以及基音周期, 计算需复制基音周期的个数 R, 然后执行步骤 36, 如果为清音帧, 则执行步骤 310。

[0065] $R = \lfloor CL / pitch \rfloor + 1$, 其中, pitch 表示基音周期。

[0066] 步骤 36 : 如果进行重采样滤波后的音频信号输入帧 X 为突变帧, 且发生突变的子块位置为第 4 子块, 则执行步骤 37, 如果发生突变的子块位置不是第 4 子块, 则执行步骤 38。

[0067] 步骤 37 : 以第 3 子块的结束位置作为复制起始点, 将位于该复制起始点之前, 且紧邻该复制起始点的基音周期数据作为复制源, 进行 R 次复制, 并将第 4 子块的数据拼接在复制的基音周期数据之后, 之后, 将本步骤中拼接后的数据进一步拼接在步骤 32 中拼接后的数据之后, 然后执行步骤 39。

[0068] 步骤 38 : 以第 4 子块的结束位置作为复制起始点, 将位于该复制起始点之前, 且紧邻该复制起始点的基音周期数据作为复制源, 进行 R 次复制, 并将复制的基音周期数据拼接在步骤 32 中拼接后的数据之后, 然后执行步骤 39。

[0069] 步骤 39 : 计算当前拼接出的数据超出音频信号输出帧 X' 的长度 OL, 然后执行步骤 312。

[0070] $OL = R * pitch - CL$, 其中, pitch 表示基音周期。音频信号输出帧 X' 表示音频信号输入帧 X 对应的音频信号输出帧。

[0071] 本实施例中, 在针对每一音频信号输入帧进行处理时, 均需要用本次生成的 OL 替换之前所保存的处理前一音频信号输入帧时生成的 OL。

[0072] 步骤 310 : 以第 4 子块的结束位置作为复制起始点, 复制该复制起始点之前, 且紧邻该复制起始点的长度为 CL 的数据, 并将复制的长度为 CL 的数据拼接在步骤 32 中拼接后的数据之后, 然后将 OL 设置为 0, 之后执行步骤 313。

[0073] 由于本步骤中复制的数据长度就是步骤 33 中计算出的需要复制的数据长度, 所以 OL 的取值将为 0。

[0074] 步骤 311 : 将 OL 设置为 CL 的相反数, 然后执行步骤 312。

[0075] 本步骤中, 将 OL 设置为 -CL。

[0076] 步骤 34 中提到, 如果步骤 33 中计算出的需要复制的数据长度 CL 小于 0 (等于 0 可按任一方式进行处理), 则执行步骤 311, 而 CL 小于 0, 说明步骤 32 中拼接后的数据长度已经超出音频信号输出帧 X' 的长度, 超出的长度即为 CL 的相反数 (因为 CL 为负数)。

[0077] 步骤 312 : 将超出音频信号输出帧 X' 长度的数据进行保存, 然后执行步骤 313。

[0078] 步骤 313 : 输出音频信号输出帧 X', 即升调后的音频信号。

[0079] 图 4 为本发明方法实施例中的降调过程示意图。如图 4 所示,包括:

[0080] 步骤 41:将进行重采样滤波后的音频信号输入帧 X 平均分为 4 个子块,确定进行重采样滤波后的音频信号输入帧 X 是否为突变帧,并记录下发生突变的子块位置。

[0081] 本步骤中,计算每个子块的能量,并比较每两个相邻子块的能量大小,如果存在突变,即存在后一子块的能量为与其相邻的前一子块的能量 12 倍的情况,则确定进行重采样滤波后的音频信号输入帧 X 为突变帧,并记录下发生突变的子块位置。

[0082] 如何计算每个子块的能量为现有技术,不再赘述。

[0083] 步骤 42:计算需要删除的数据的长度 DL。

[0084] $DL = InLen - OutLen$, 其中, InLen 表示进行重采样滤波后的音频信号输入帧 X 的长度, OutLen 表示音频信号输入帧 X 对应的音频信号输出帧 X' 的长度,

$OutLen = \lfloor InLen * \alpha \rfloor$, 符号 $\lfloor \cdot \rfloor$ 表示向下取整。

[0085] 步骤 43:如果进行重采样滤波后的音频信号输入帧 X 为浊音帧,则根据 DL 以及基音周期,计算需复制基音周期的个数 R,然后执行步骤 44,如果为清音帧,则执行步骤 47。

[0086] $R = \lfloor DL / pitch \rfloor + 1$, 其中, pitch 表示基音周期。

[0087] 步骤 44:如果进行重采样滤波后的音频信号输入帧 X 不是突变帧,则执行步骤 45,否则,执行步骤 46。

[0088] 步骤 45:将距离进行重采样滤波后的音频信号输入帧 X 的起始点长度为已保存的 SL 的位置作为音频信号输出帧 X' 的起始点,输出长度为 OutLen 的数据, OutLen 表示音频信号输入帧 X 对应的音频信号输出帧 X' 的长度,删除未输出的其余数据(起始点之前和输出终点之后的数据),然后计算新的 SL,并利用新计算出的 SL 更新已保存的 SL,之后执行步骤 48。

[0089] $SL = old_SL + R * pitch - DL$, 其中, old_SL 表示已保存的 SL,初始值设置为 0。

[0090] 步骤 46:基于 DL 和发生突变的子块位置,确定音频信号输出帧 X' 的起始点,输出长度为 OutLen 的数据,删除未输出的其余数据,并将已保存的 SL 设置为 0,之后执行步骤 48。

[0091] 本步骤中,如果发生突变的子块位置距离进行重采样滤波后的音频信号输入帧 X 的起始点的距离小于 DL,则以进行重采样滤波后的音频信号输入帧 X 的起始点作为音频信号输出帧 X' 的起始点,否则,以距离进行重采样滤波后的音频信号输入帧 X 的起始点 DL 的位置作为音频信号输出帧 X' 的起始点。

[0092] 步骤 47:将进行重采样滤波后的音频信号输入帧 X 的起始点作为音频信号输出帧 X' 的起始点,输出长度为 OutLen 的数据,删除未输出的其余数据,并将已保存的 SL 设置为 0,之后执行步骤 48。

[0093] 步骤 48:利用低通滤波器对音频信号输出帧 X' 进行低通滤波。

[0094] 低通滤波器的截止频率为音频信号输入帧 X 的采样频率的 $0.5 * q/p$ 倍,音频信号输入帧 X 的采样频率为已知的。通过本步骤的处理,可滤除超出降调后带宽的频谱。

[0095] 至此,即完成了本发明方法实施例所示流程。

[0096] 图 5 为本发明实现音频变调的装置实施例的组成结构示意图。如图 5 所示,包括:

[0097] 确定单元 51,用于根据接收到的变调模式 d ,确定变调因子 α ,并进一步确定将变调因子 α 用分数形式进行表示后的分子 q 和分母 p ;

[0098] 处理单元 52,用于针对每一音频信号输入帧 X ,分别进行以下处理:

[0099] 基于 q 和 p ,对音频信号输入帧 X 进行时域上的重采样滤波;

[0100] 对进行重采样滤波后的音频信号输入帧 X 进行清、浊音判决,并确定进行重采样滤波后的音频信号输入帧 X 的基音周期;

[0101] 根据变调模式 d 确定需要升调还是降调,如果需要升调,则基于清、浊音判决结果,对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据复制处理,如果需要降调,则基于清、浊音判决结果,对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据删除处理。

[0102] 其中,处理单元 52 中可具体包括:

[0103] 重采样滤波子单元 521,用于对音频信号输入帧 X 进行 p 倍的线性插值,将进行线性插值后的音频信号输入帧 X 通过低通滤波器进行低通滤波,低通滤波器的增益为 p ,截止频率为 π/q 和 π/p 中的较小值,将进行低通滤波后的音频信号输入帧 X 进行 q 倍的采样;

[0104] 判决子单元 522,用于计算 $\rho(\tau) = \frac{\sum_{n=0}^{N-1} s(n)s(n-\tau)}{\sqrt{\sum_{n=0}^{N-1} s^2(n) \sum_{n=0}^{N-1} s^2(n-\tau)}}$;其中, $s(n)$ 表示音频信号强度, N 表示进行重采样滤波后的音频信号输入帧 X 的长度, τ 表示延迟参数,为正整数,取值范围预先设定;确定不同的延迟参数 τ 对应的 $\rho(\tau)$ 中的最大值,并将最大值与预先设定的阈值进行比较,如果最大值小于阈值,则确定进行重采样滤波后的音频信号输入帧 X 为清音帧,否则,为浊音帧,并将最大值对应的延迟参数 τ 确定为进行重采样滤波后的音频信号输入帧 X 的基音周期;

[0105] 确定子单元 523,用于根据变调模式 d 确定需要升调还是降调,如果变调模式 d 的取值大于 0,则确定需要升调,如果变调模式 d 的取值小于 0,则确定需要降调;

[0106] 变调子单元 524,用于当确定需要升调时,基于清、浊音判决结果,对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据复制处理,当确定需要降调时,基于清、浊音判决结果,对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据删除处理。

[0107] 变调子单元 524 中又可进一步包括(为简化附图,未图示):

[0108] 升调子单元,用于基于清、浊音判决结果,对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据复制处理,包括:

[0109] A1、将进行重采样滤波后的音频信号输入帧 X 平均分为 4 个子块,并计算每个子块的能量,比较每两个相邻子块的能量大小,如果存在突变,即存在后一子块的能量为与其相邻的前一子块的能量 12 倍的情况,则确定进行重采样滤波后的音频信号输入帧 X 为突变帧,并记录下发生突变的子块位置;

[0110] A2、将进行重采样滤波后的音频信号输入帧 X 拼接在所保存的超出前一音频信号输出帧长度的数据之后,前一音频信号输出帧为紧邻音频信号输入帧 X 的前一音频信号输入帧对应的音频信号输出帧;

- [0111] A3、计算需要复制的数据的长度 CL；
- [0112] A4、确定 CL 是否大于 0，如果是，则执行步骤 A5，否则，执行步骤 A11；
- [0113] A5、如果进行重采样滤波后的音频信号输入帧 X 为浊音帧，则根据 CL 以及基音周期，计算需复制基音周期的个数 R，然后执行步骤 A6，如果为清音帧，则执行步骤 A10；
- [0114] A6、如果进行重采样滤波后的音频信号输入帧 X 为突变帧，且发生突变的子块位置为第 4 子块，则执行步骤 A7，如果发生突变的子块位置不是第 4 子块，则执行步骤 A8；
- [0115] A7、以第 3 子块的结束位置作为复制起始点，将位于复制起始点之前，且紧邻复制起始点的基音周期数据作为复制源，进行 R 次复制，并将第 4 子块的数据拼接在复制的基音周期数据之后，之后，将本步骤中拼接后的数据进一步拼接在步骤 A2 中拼接后的数据之后，然后执行步骤 A9；
- [0116] A8、以第 4 子块的结束位置作为复制起始点，将位于复制起始点之前，且紧邻复制起始点的基音周期数据作为复制源，进行 R 次复制，并将复制的基音周期数据拼接在步骤 A2 中拼接后的数据之后，然后执行步骤 A9；
- [0117] A9、计算当前拼接出的数据超出音频信号输出帧 X' 的长度 OL，然后执行步骤 A12；
- [0118] A10、以第 4 子块的结束位置作为复制起始点，复制复制起始点之前，且紧邻复制起始点的长度为 CL 的数据，并将复制的长度为 CL 的数据拼接在步骤 A2 中拼接后的数据之后，降 OL 设置为 0，之后执行步骤 A13；
- [0119] A11、将 OL 设置为 CL 的相反数，然后执行步骤 A12；
- [0120] A12、将超出音频信号输出帧 X' 长度的数据进行保存，然后执行步骤 A13；
- [0121] A13、输出升调后的音频信号输出帧 X'；
- [0122] 降调子单元，用于基于清、浊音判决结果，对进行重采样滤波后的音频信号输入帧 X 进行以基音周期为单位的数据删除处理，包括：
- [0123] B1、将进行重采样滤波后的音频信号输入帧 X 平均分为 4 个子块，并计算每个子块的能量，比较每两个相邻子块的能量大小，如果存在突变，即存在后一子块的能量为与其相邻的前一子块的能量的 12 倍的情况，则确定进行重采样滤波后的音频信号输入帧 X 为突变帧，并记录下发生突变的子块位置；
- [0124] B2、计算需要删除的数据的长度 DL；
- [0125] B3、如果进行重采样滤波后的音频信号输入帧 X 为浊音帧，则根据 DL 以及基音周期，计算需复制基音周期的个数 R，然后执行步骤 B4，如果为清音帧，则执行步骤 B7；
- [0126] B4、如果进行重采样滤波后的音频信号输入帧 X 不是突变帧，则执行步骤 B5，否则，执行步骤 B6；
- [0127] B5、将距离进行重采样滤波后的音频信号输入帧 X 的起始点长度为已知的 SL 的位置作为音频信号输出帧 X' 的起始点，输出长度为 OutLen 的数据，OutLen 表示音频信号输入帧 X 对应的音频信号输出帧 X' 的长度，删除未输出的其余数据，然后计算新的 SL，利用新计算的 SL 更新已保存的 SL，之后执行步骤 B8；
- [0128] B6、基于 DL 和发生突变的子块位置，确定音频信号输出帧 X' 的起始点，输出长度为 OutLen 的数据，删除未输出的其余数据，并将已保存的 SL 设置为 0，之后执行步骤 B8；
- [0129] B7、将进行重采样滤波后的音频信号输入帧 X 的起始点作为音频信号输出帧 X' 的起始点，输出长度为 OutLen 的数据，删除未输出的其余数据，并将已保存的 SL 设置为 0，之

后执行步骤 B8；

[0130] B8、利用低通滤波器对音频信号输出帧 X' 进行低通滤波，低通滤波器的截止频率为音频信号输入帧 X 的采样频率的 $0.5 \cdot q/p$ 倍。

[0131] 图 5 所示装置实施例的具体工作流程请参照图 1 所示方法实施例中的相应说明，不再赘述。

[0132] 总之，采用本发明的技术方案，有效地实现了音频信号的升调和降调。

[0133] 以上所述仅为本发明的较佳实施例而已，并不用以限制本发明，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明保护的范围之内。

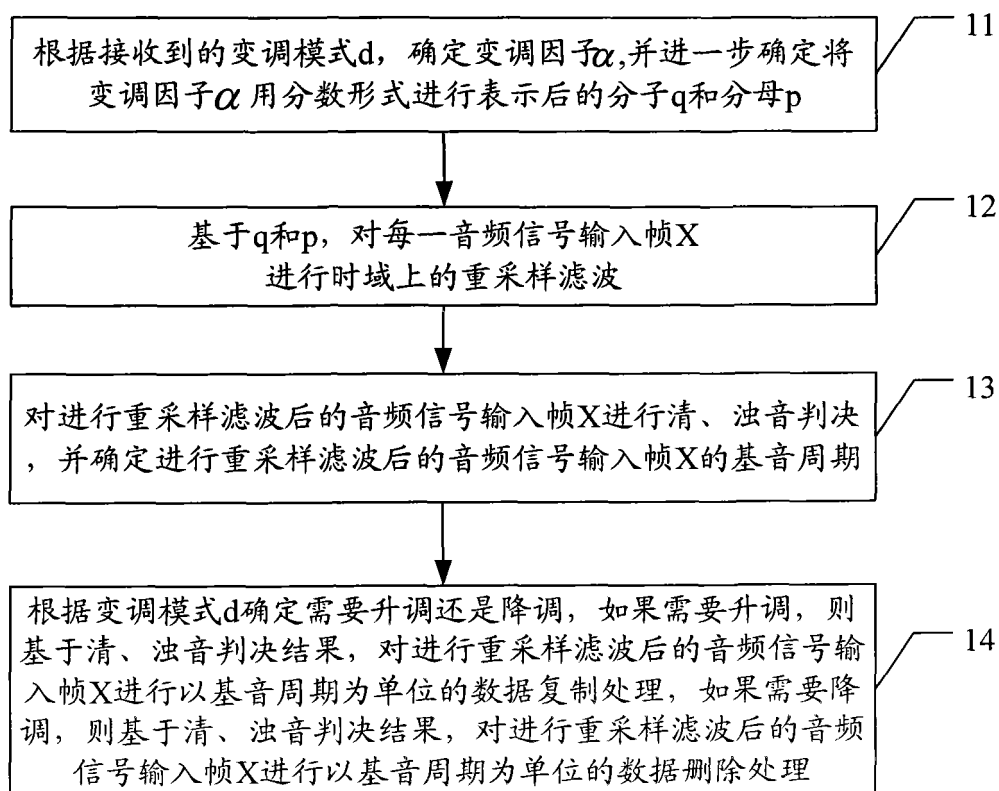


图 1

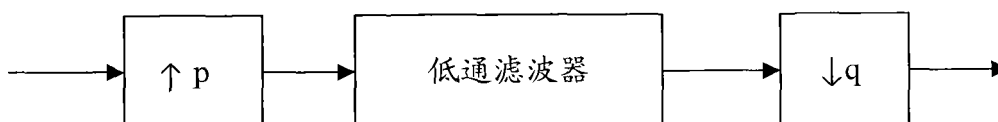


图 2

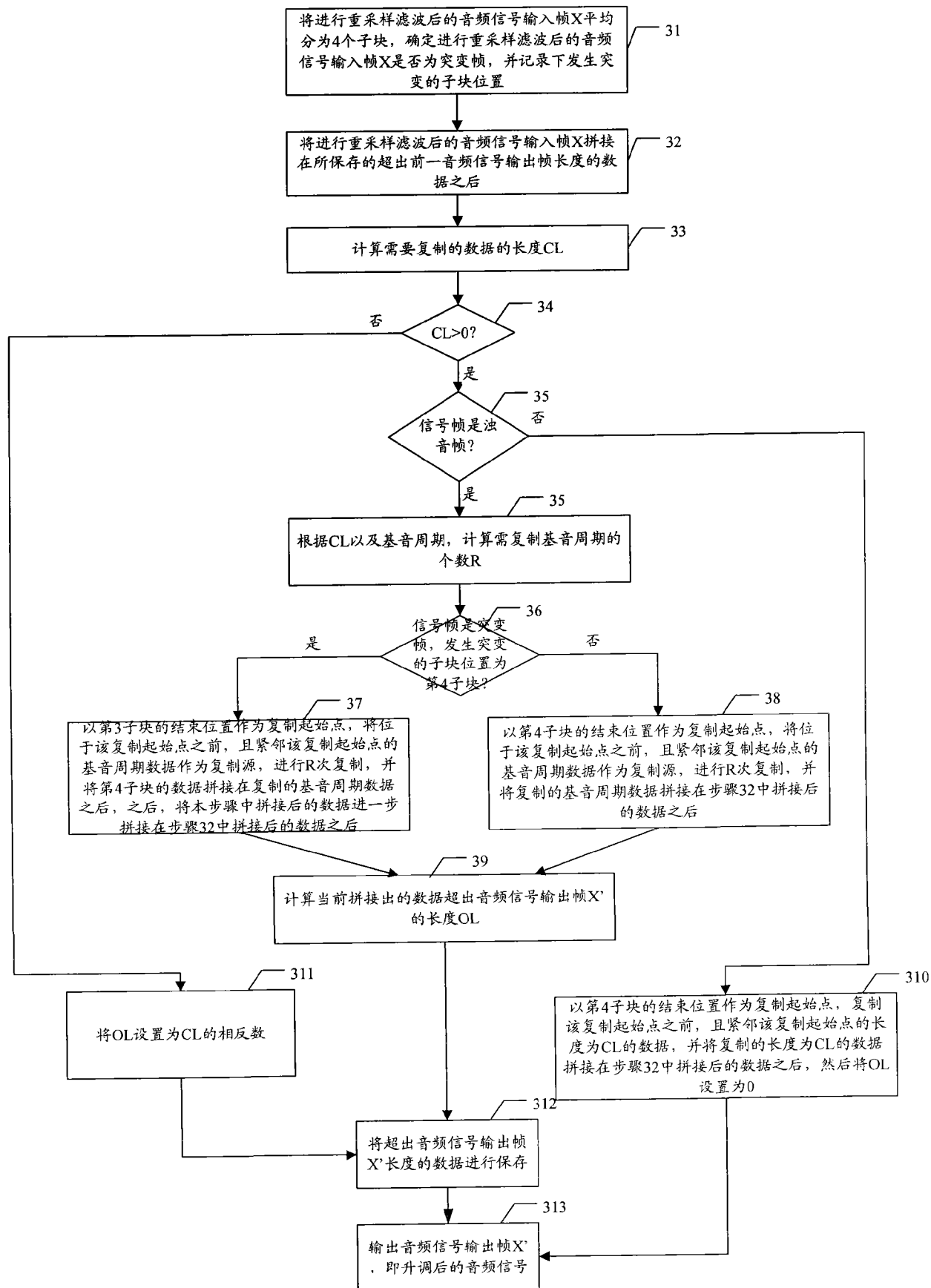


图 3

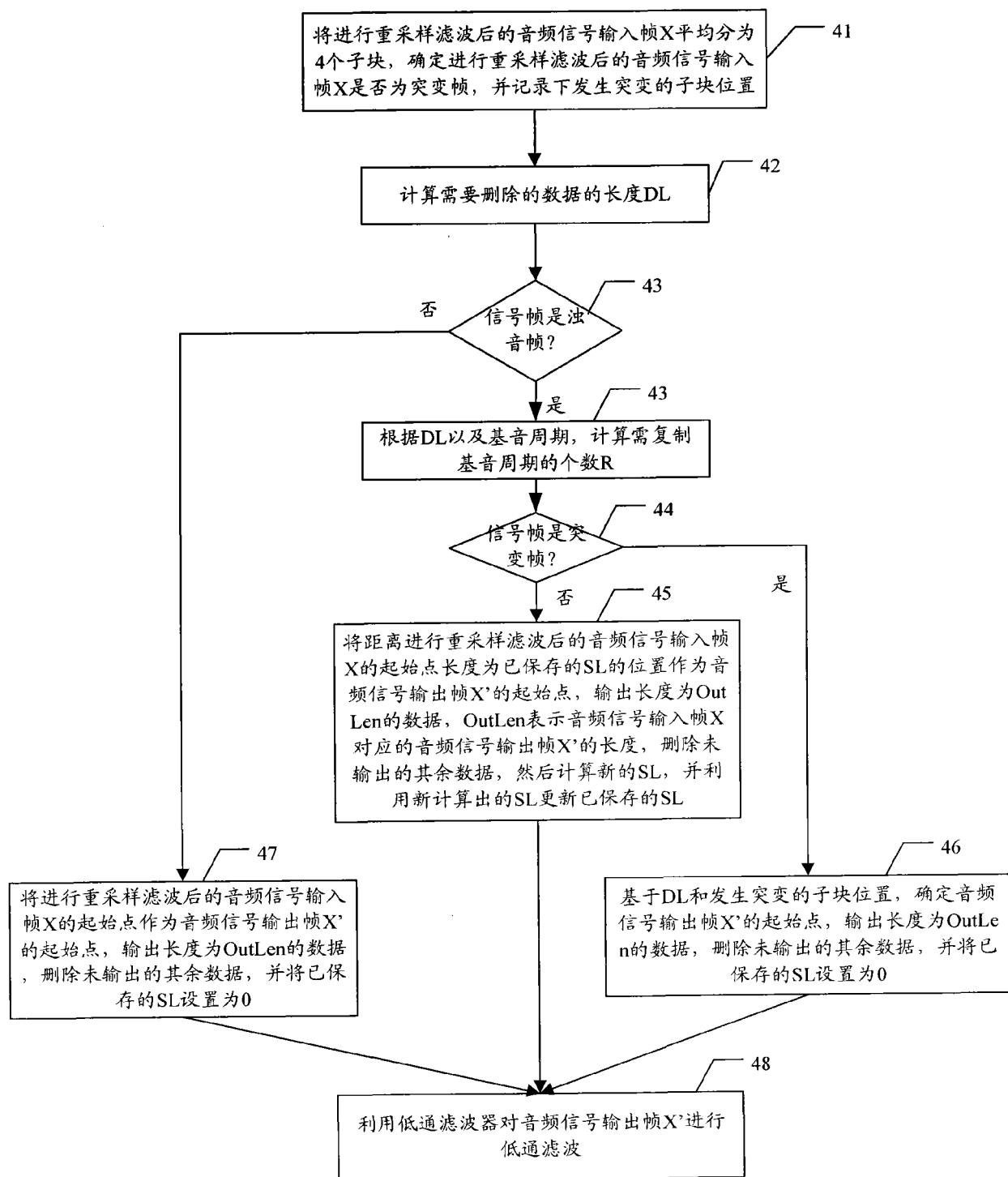


图 4

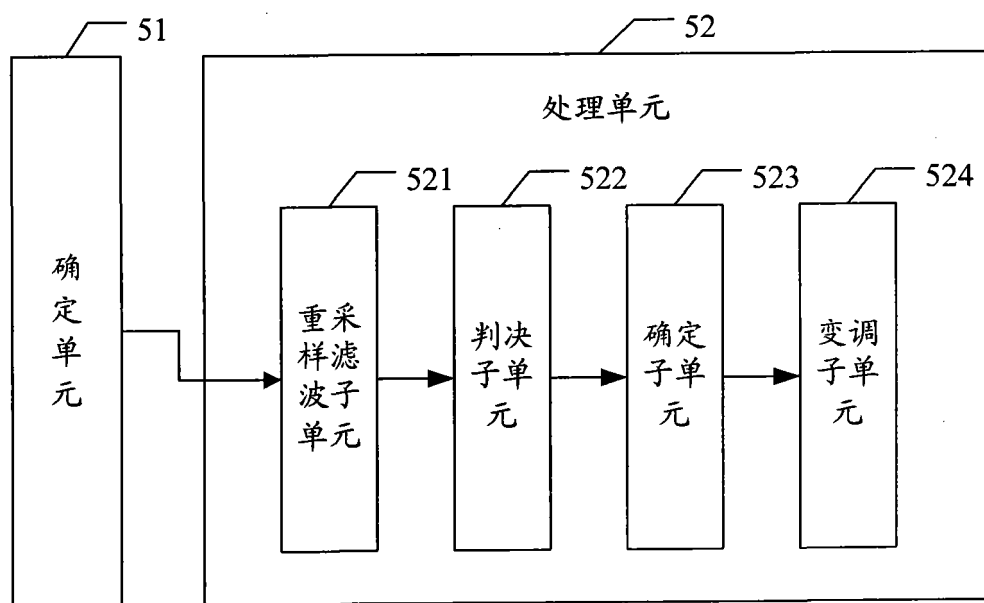


图 5