



(12) 发明专利

(10) 授权公告号 CN 1971707 B

(45) 授权公告日 2010.09.29

(21) 申请号 200610161727.3

(22) 申请日 2006.12.13

(73) 专利权人 北京中星微电子有限公司
地址 100083 北京市海淀区学院路 35 号世
宁大厦 15 层

(72) 发明人 邓昊 冯宇红 张晨

(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 宋志强 麻海明

(51) Int. Cl.

- G10L 11/00 (2006.01)
- G10L 11/06 (2006.01)
- G10L 19/00 (2006.01)

(56) 对比文件

- US 6526376 B1, 2003.02.25, 全文.
- 同上.
- CN 1412742 A, 2003.04.23, 摘要, 说明书第

2 页第 4 段 - 第 6 页倒数第 3 段、图 5.
CN 1424712 A, 2003.06.18, 全文.
王秀君. 一种有效的语音基因周期提取算
法. 应用科技 33 1. 2006, 33(1), 7-9.

审查员 张亚峰

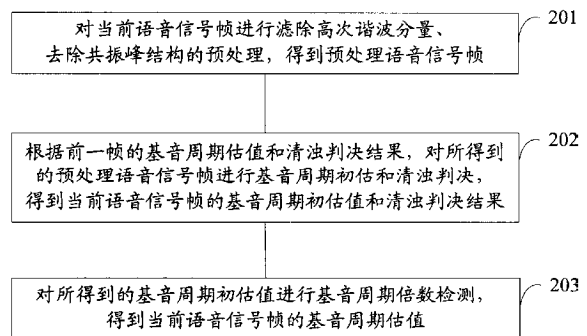
权利要求书 3 页 说明书 15 页 附图 2 页

(54) 发明名称

一种进行基音周期估计和清浊判决的方法及装置

(57) 摘要

本发明公开了一种进行基音周期估计和清浊判决的方法,该方法包括以下步骤:A、对当前语音信号帧进行滤除高次谐波分量、去除共振峰结构的预处理,得到预处理语音信号帧;B、对所得到的预处理语音信号帧进行基音周期估计和清浊判决,得到所述当前语音信号帧的基音周期估值和清浊判决结果。本发明还公开了一种进行基音周期估计和倍数检测的装置,该装置包括:预处理模块和基音周期估计/清浊判决模块。应用本发明能够提高语音信号处理中基音周期估值和清浊判决结果的准确度。



1. 一种进行基音周期估计和清浊判决的方法,其特征在于,该方法包括以下步骤:

A、对当前语音信号帧进行滤除高次谐波分量、去除共振峰结构的预处理,得到预处理语音信号帧;

B1、计算所述预处理语音信号帧的归一化自相关;

B2、根据前一语音信号帧的基音周期估值和清浊判决结果确定所述当前语音信号帧的基音周期待选值的范围,并确定所述范围对应的归一化自相关值中的最大值,以及所述最大值对应的基音周期待选值;

B3、根据所述前一语音信号帧的清浊判决结果确定所述当前语音信号帧的清浊判决阈值;

B4、根据所述归一化自相关值中的最大值以及所述当前语音信号帧的清浊判决阈值,确定当前语音信号帧的清浊判决结果;

B5、根据所确定的当前语音信号帧的清浊判决结果以及所述归一化自相关值中的最大值对应的基音周期待选值,确定当前语音信号帧的基音周期估值。

2. 根据权利要求 1 所述的方法,其特征在于,在所述步骤 B5 之后,进一步包括:

C、当所述当前语音信号帧的清浊判决结果为浊音帧时,将所得到的基音周期估值作为基音周期初估值,并对所述基音周期初估值进行基音周期倍数检测,将所述基音周期倍数检测的结果作为当前语音信号帧的基音周期估值。

3. 根据权利要求 1 或 2 所述的方法,其特征在于,步骤 A 所述预处理为:对所述当前语音信号帧进行频谱展平,并采用线性相位低通滤波器滤除经所述频谱展平处理过的所述当前语音信号帧的高频部分。

4. 根据权利要求 1 或 2 所述的方法,其特征在于,步骤 B2 所述确定所述当前语音信号帧的基音周期待选值的范围为:

当所述当前语音信号帧为第一个语音信号帧时,或者当所述前一语音信号帧为清音帧时,将所述基音周期候选值的范围确定为 20 到 140 之间;

否则,将小于所述前一语音信号帧的基音周期估值、且大于等于 20 的整数确定为所述基音周期候选值范围的最小值,将大于所述前一语音信号帧的基音周期估值、且小于等于 140 的整数确定为所述基音周期候选值范围的最大值。

5. 根据权利要求 4 所述的方法,其特征在于,进一步预先设置基音周期变化率因子;

所述最小值为:所述前一语音信号帧的基音周期估值减所述前一语音信号帧的基音周期估值与所述基音周期变化率因子之积所得到的值;

所述最大值为:所述前一语音信号帧的基音周期估值加所述前一语音信号帧的基音周期估值与所述基音周期变化率因子之积所得到的值。

6. 根据权利要求 1 或 2 所述的方法,其特征在于,所述步骤 B3 之前进一步包括:设置第一阈值和第二阈值,其中,第一阈值大于第二阈值,且第一阈值和第二阈值的取值均在 0 到 1 之间;

所述步骤 B3 为:若前一语音信号帧的清浊判决结果为清音帧,则将第一阈值作为当前语音信号帧的清浊判决阈值;若前一语音信号帧的清浊判决结果为浊音帧,则将第二阈值作为当前语音信号帧的清浊判决阈值。

7. 根据权利要求 1 或 2 所述的方法,其特征在于,所述步骤 B4 为:

若所述归一化自相关值中的最大值小于所述清浊判决阈值,则判定当前语音信号帧为清音帧;否则,判定当前语音信号帧为浊音帧。

8. 根据权利要求 1 或 2 所述的方法,其特征在于,所述步骤 B5 为:

若所述当前语音信号帧为清音帧,则将所述基音周期估值确定为 0;若所述当前语音信号帧为浊音帧,则将所述基音周期估值确定为所述归一化自相关值中的最大值对应的基音周期待选值。

9. 根据权利要求 2 所述的方法,其特征在于,步骤 C 所述基音周期倍数检测包括:

C1、设置当前基音周期下限、基音周期暂定认可估值、分频数,并设定基音周期检测下限、第一补偿因子和第二补偿因子的值;

其中,当前基音周期下限表示所述当前语音信号帧的基音周期估值下限,其值根据前一语音信号帧的基音周期估值估算得到;

基音周期暂定认可估值表示在对所述基音周期初估值进行所述基音周期倍数检测的过程中,暂时被认可的基音周期估值;

基音周期检测下限表示执行所述基音周期倍数检测的检测范围下限;

第一补偿因子的值小于 1,第二补偿因子的值大于 1;

C2、对所述当前语音信号帧的基音周期初估值与所述分频数之商取整,得到所述基音周期初估值的分数值,判断所述基音周期初估值的分数值是否大于等于所述基音周期检测下限,当所述判断的结果为是时,继续执行步骤 C3;否则,将所述基音周期暂定认可估值确定为所述当前语音信号帧的基音周期估值,结束本步骤中倍数检测的流程;

C3、若所述基音周期初估值的分数值大于等于所述当前基音周期下限,则将所述第一补偿因子的值确定为补偿因子;否则,将所述第二补偿因子的值确定为补偿因子;

C4、将信号延时长度取为所述基音周期暂定认可估值的归一化自相关值与所述补偿因子之积、与信号延时长度取为所述基音周期初估值的分数值的归一化自相关值作比较,当前者小于后者时,将所述基音周期初估值的分数值作为所述基音周期暂定认可估值;

C5、增大分频数,继续执行步骤 C2。

10. 根据权利要求 9 所述的方法,其特征在于,步骤 C1 所述基音周期下限为:小于所述前一语音信号帧的基音周期估值、且大于等于 20 的整数。

11. 根据权利要求 10 所述的方法,其特征在于,进一步预先设置基音周期变化率因子;

所述小于所述前一语音信号帧的基音周期估值、且大于等于 20 的整数为:所述前一语音信号帧的基音周期估值减所述前一语音信号帧的基音周期估值与所述基音周期变化率因子之积所得到的值。

12. 根据权利要求 10 所述的方法,其特征在于,步骤 C1 所述基音周期暂定认可估值的初值为:所述基音周期初估值。

13. 根据权利要求 9 所述的方法,其特征在于,步骤 C1 所述分频数的初值为 2;

步骤 C5 所述增大分频数为:将所述分频数加 1。

14. 一种进行基音周期估计和清浊判决的装置,其特征在于,该装置包括:预处理模块和基音周期估计/清浊判决模块;

所述预处理模块,用于对当前语音信号帧进行滤除高次谐波分量、去除共振峰结构的预处理,并将所得到的预处理语音信号帧发送给所述基音周期估计/清浊判决模块;

所述基音周期估计 / 清浊判决模块,用于对来自于所述预处理模块的预处理语音信号帧进行基音周期估计和清浊判决,得到所述当前语音信号帧的基音周期估值和清浊判决结果;

所述基音周期估计 / 清浊判决模块进一步包括:归一化自相关计算单元和基音周期估计 / 清浊判决单元,其中:

所述归一化自相关计算单元,用于计算来自于所述预处理模块的预处理语音信号帧的归一化自相关值,并将计算结果发送给所述基音周期估计 / 清浊判决单元;

所述基音周期估计 / 清浊判决单元,用于接收前一语音信号帧的基音周期估值和清浊判决结果反馈,并根据来自于所述归一化自相关计算单元的归一化自相关值、所述前一语音信号帧的基音周期估值和清浊判决结果进行基音周期估计和清浊判决,其中,根据所述前一语音信号帧的清浊判决结果确定所述当前语音信号帧的清浊判决阈值。

15. 根据权利要求 14 所述的装置,其特征在于,所述装置进一步包括倍数检测模块;

所述倍数检测模块,用于对来自于所述基音周期估计 / 清浊判决模块的基音周期估值进行基音周期倍数检测,得到当前语音信号帧的基音周期估值。

16. 根据权利要求 15 所述的装置,其特征在于,所述基音周期估计 / 清浊判决单元,用于根据来自于所述归一化自相关计算单元的归一化自相关。

一种进行基音周期估计和清浊判决的方法及装置

技术领域

[0001] 本发明涉及语音信号处理技术,特别涉及一种进行基音周期估计和清浊判决的方法及装置。

背景技术

[0002] 在语音信号处理技术领域中,语音信号的离散时域生成模型被广泛应用于参数语音编码、语音合成和变调 (pitch scaling) 等应用中。

[0003] 图 1 为现有通用的语音信号离散时域生成模型示意图。参见图 1,该模型包括三个部分:激励源 110、声道模型 120 和辐射模型 130,其中,激励源 110 包括浊音激励单元 111 和清音激励单元 112 两个分支,以及一个浊音 / 清音开关 113。

[0004] 图 1 所示模型中,浊音激励单元 111,用于根据输入的基音频率产生浊音的激励;这里,所输入的基音频率即基音周期的倒数,是根据输入语音信号得到的基音周期估计值;清音激励单元 112,用于产生清音的激励,清音的激励为随机噪声,即与基音频率无关;

[0005] 浊音 / 清音开关 113,用于根据清浊判决的结果接通浊音激励单元 111 或清音激励单元 112;这里,清浊判决的结果是判断输入语音信号是浊音信号还是清音信号的判决结果;

[0006] 声道模型 120 给出了离散时域的声道传输函数;

[0007] 辐射模型 130 负责与唇形有关的处理。

[0008] 由图 1 所示模型可见,基音周期估计和清浊判决的准确性对于各种语音应用的最终效果起着至关重要的作用。

[0009] 基音周期估计和清浊判决一般同步进行,而且,由于语音信号具有短时平稳性,基音周期估计和清浊判决一般针对每个语音信号帧,分帧进行。目前,通常都利用浊音信号的时域波形相似性,通过计算输入语音信号帧的归一化自相关来进行基音周期估计和清浊判决。下面简要介绍现有技术中典型的进行基音周期估计和清浊判决的方法,该方法包括以下步骤:

[0010] 步骤 1:对输入语音信号进行分帧,帧长取为 10 ~ 20 毫秒 (ms) 左右,得到当前语音信号帧。

[0011] 步骤 2:计算包含当前语音信号帧的一段语音信号的归一化自相关;

[0012] 通常,在计算自相关时,所用到的信号不止是当前语音信号帧这一段,而是以当前所分析的语音信号帧为中心的数帧信号减处理延时进行计算。假设,以 $s(n)$ 表示输入语音信号在第 n 个样本点的幅值、 $R(\tau)$ 表示信号延时长度为 τ 时的归一化自相关值,则可以

$$\text{根据 } R(\tau) = \frac{\sum_{n=0}^{N-1} s(n)s(n+\tau)}{\sum_{n=0}^{N-1} s^2(n)} \quad (1) \text{ 计算输入语音信号 } s(n) \text{ 的归一化自相关 } R(\tau)。 (1)$$

式中, N 为计算自相关时使用的信号样本点长度,一般应至少包含两个基音周期,当信号采

样频率为 8 千赫兹 (kHz) 时, N 的典型值为 320, 即上式中 n 的取值范围为 0 ~ 319; Σ 表示求和操作; τ 表示计算自相关时的信号延时长度, 即当前语音信号帧的基音周期待选值, 取值范围为 20 ~ 140, τ 与基音周期估值的动态范围对应。本步骤中, 按照 (1) 式分别计算信号延时长度 τ 的取值为 20 ~ 140 之间的各基音周期待选值的归一化自相关 $R(20) \sim R(140)$ 。

[0013] 步骤 3: 求 $R(20) \sim R(140)$ 中的最大值, 记为 $R(\tau_{\max}) = \max\{R(\tau) \mid \tau = 20 \sim 140\}$, 其中, τ_{\max} 表示 $R(20) \sim R(140)$ 中最大值所对应的 τ 的值。

[0014] 步骤 4: 由于从理论上来说, 对于浊音帧, τ 的取值越接近于基音周期 P, 则 $R(\tau)$ 越大; 而对于清音帧, 无论 τ 取为何值, $R(\tau)$ 均较小, 因此, 可以根据 $R(\tau_{\max})$, 按照

$$\begin{cases} \text{清音帧} & R(\tau_{\max}) < R_{\text{Th}} \\ \text{浊音帧, } p' = \tau_{\max} & R(\tau_{\max}) \geq R_{\text{Th}} \end{cases} \quad (2) \text{进行基音周期估计和清浊判决。}$$

(2) 式中, R_{Th} 为预先设定的阈值, 典型值为 0.65, p' 表示实际基音周期 P 的估值, (2) 式表示 $R(\tau_{\max})$ 小于预先设定的阈值时, 判定该语音帧为清音帧, 否则为浊音帧。

[0015] 至此, 得到基音周期估值和清浊判决结果。

[0016] 在实际应用中, 由于输入语音信号中存在共振峰结构, 基音周期估值将可能被误估为实际值的倍数或分数; 而上述方法没有对输入语音信号进行任何去除共振峰结构的处理, 因此, 上述方法将导致所估算的基音周期估值不准确。此外, 由于上述方法过于简单, 在清浊音的过渡部分, 无法准确进行基音周期估计和清浊判决, 导致使用上述方法得到的基音周期估值和清浊判决结果的准确度很差。

[0017] 发明内容

[0018] 有鉴于此, 本发明的主要目的在于提供一种进行基音周期估计和清浊判决的方法, 以提高语音信号处理中基音周期估值和清浊判决结果的准确度。

[0019] 本发明的另一个目的在于提供一种进行基音周期估计和清浊判决的装置, 以提高语音信号处理中基音周期估值和清浊判决结果的准确度。

[0020] 为达到上述目的, 本发明的技术方案具体是这样实现的:

[0021] 一种进行基音周期估计和清浊判决的方法, 该方法包括以下步骤:

[0022] A、对当前语音信号帧进行滤除高次谐波分量、去除共振峰结构的预处理, 得到预处理语音信号帧;

[0023] B1、计算所述预处理语音信号帧的归一化自相关;

[0024] B2、根据所述前一语音信号帧的基音周期估值和清浊判决结果确定所述当前语音信号帧的基音周期待选值的范围, 并确定所述范围对应的归一化自相关值中的最大值, 以及所述最大值对应的基音周期待选值;

[0025] B3、根据所述前一语音信号帧的清浊判决结果确定所述当前语音信号帧的清浊判决阈值;

[0026] B4、根据所述归一化自相关值中的最大值以及所述当前语音信号帧的清浊判决阈值, 确定当前语音信号帧的清浊判决结果;

[0027] B5、根据所确定的当前语音信号帧的清浊判决结果以及所述归一化自相关值中的最大值对应的基音周期待选值, 确定当前语音信号帧的基音周期估值。

- [0028] 进一步地,在所述步骤 B5 之后可以包括:
- [0029] C、当所述当前语音信号帧的清浊判决结果为浊音帧时,将所得到的基音周期估值作为基音周期初估值,并对所述基音周期初估值进行基音周期倍数检测,将所述基音周期倍数检测的结果作为当前语音信号帧的基音周期估值。
- [0030] 其中,步骤 A 所述预处理可以为:对所述当前语音信号帧进行频谱展平,并采用线性相位低通滤波器滤除经所述频谱展平处理过的所述当前语音信号帧的高频部分。
- [0031] 其中,步骤 B2 所述确定所述当前语音信号帧的基音周期候选值的范围可以为:
- [0032] 当所述当前语音信号帧为第一个语音信号帧时,或者当所述前一语音信号帧为清音帧时,将所述基音周期候选值的范围确定为 20 到 140 之间;
- [0033] 否则,将小于所述前一语音信号帧的基音周期估值、且大于等于 20 的整数确定为所述基音周期候选值范围的最小值,将大于所述前一语音信号帧的基音周期估值、且小于等于 140 的整数确定为所述基音周期候选值范围的最大值。
- [0034] 进一步地,可以预先设置基音周期变化率因子;
- [0035] 所述最小值可以为:所述前一语音信号帧的基音周期估值减所述前一语音信号帧的基音周期估值与所述基音周期变化率因子之积所得到的值;
- [0036] 所述最大值可以为:所述前一语音信号帧的基音周期估值加所述前一语音信号帧的基音周期估值与所述基音周期变化率因子之积所得到的值。
- [0037] 进一步地,在所述步骤 B3 之前可以包括:设置第一阈值和第二阈值,其中,第一阈值大于第二阈值,且第一阈值和第二阈值的取值均在 0 到 1 之间;
- [0038] 所述步骤 B3 可以为:若前一语音信号帧的清浊判决结果为清音帧,则将第一阈值作为当前语音信号帧的清浊判决阈值;若前一语音信号帧的清浊判决结果为浊音帧,则将第二阈值作为当前语音信号帧的清浊判决阈值。
- [0039] 其中,所述步骤 B4 可以为:
- [0040] 若所述归一化自相关值中的最大值小于所述清浊判决阈值,则判定当前语音信号帧为清音帧;否则,判定当前语音信号帧为浊音帧。
- [0041] 其中,所述步骤 B5 可以为:
- [0042] 若所述当前语音信号帧为清音帧,则将所述基音周期估值确定为 0;若所述当前语音信号帧为浊音帧,则将所述基音周期估值确定为所述归一化自相关值中的最大值对应的基音周期候选值。
- [0043] 其中,步骤 C 所述基音周期倍数检测可以包括:
- [0044] C1、设置当前基音周期下限、基音周期暂定认可估值、分频数,并设定基音周期检测下限、第一补偿因子和第二补偿因子的值;
- [0045] 其中,当前基音周期下限表示所述当前语音信号帧的基音周期估值下限,其值根据前一语音信号帧的基音周期估值估算得到;
- [0046] 基音周期暂定认可估值表示在对所述基音周期初估值进行所述基音周期倍数检测的过程中,暂时被认可的基音周期估值;
- [0047] 基音周期检测下限表示执行所述基音周期倍数检测的检测范围下限;
- [0048] 第一补偿因子的值小于 1,第二补偿因子的值大于 1;
- [0049] C2、对所述当前语音信号帧的基音周期初估值与所述分频数之商取整,得到所述

基音周期初估值的分数值,判断所述基音周期初估值的分数值是否大于等于所述基音周期检测下限,当所述判断的结果为是时,继续执行步骤 C3;否则,将所述基音周期暂定认可估值确定为所述当前语音信号帧的基音周期估值,结束本步骤中倍数检测的流程;

[0050] C3、若所述基音周期初估值的分数值大于等于所述当前基音周期下限,则将所述第一补偿因子的值确定为补偿因子;否则,将所述第二补偿因子的值确定为补偿因子;

[0051] C4、将信号延时长度取为所述基音周期暂定认可估值的归一化自相关值与所述补偿因子之积、与信号延时长度取为所述基音周期初估值的分数值的归一化自相关值作比较,当前者小于后者时,将所述基音周期初估值的分数值作为所述基音周期暂定认可估值;

[0052] C5、增大分频数,继续执行步骤 C2。

[0053] 其中,步骤 C1 所述基音周期下限可以为:小于所述前一语音信号帧的基音周期估值、且大于等于 20 的整数。

[0054] 进一步地,可以预先设置基音周期变化率因子;

[0055] 所述小于所述前一语音信号帧的基音周期估值、且大于等于 20 的整数可以为:所述前一语音信号帧的基音周期估值减所述前一语音信号帧的基音周期估值与所述基音周期变化率因子之积所得到的值。

[0056] 其中,步骤 C1 所述基音周期暂定认可估值的初值可以为:所述基音周期初估值。

[0057] 其中,步骤 C1 所述分频数的初值可以为 2;

[0058] 步骤 C5 所述增大分频数可以为:将所述分频数加 1。

[0059] 一种进行基音周期估计和清浊判决的装置,该装置包括:预处理模块和基音周期估计/清浊判决模块;

[0060] 所述预处理模块,用于对当前语音信号帧进行滤除高次谐波分量、去除共振峰结构的预处理,并将所得到的预处理语音信号帧发送给所述基音周期估计/清浊判决模块;

[0061] 所述基音周期估计/清浊判决模块,用于对来自于所述预处理模块的预处理语音信号帧进行基音周期估计和清浊判决,得到所述当前语音信号帧的基音周期估值和清浊判决结果;

[0062] 所述基音周期估计/清浊判决模块进一步包括:归一化自相关计算单元和基音周期估计/清浊判决单元,其中:

[0063] 所述归一化自相关计算单元,用于计算来自于所述预处理模块的预处理语音信号帧的归一化自相关值,并将计算结果发送给所述基音周期估计/清浊判决单元;

[0064] 所述基音周期估计/清浊判决单元,用于接收前一语音信号帧的基音周期估值和清浊判决结果反馈,并根据来自于所述归一化自相关计算单元的归一化自相关值、所述前一语音信号帧的基音周期估值和清浊判决结果进行基音周期估计和清浊,其中,根据所述前一语音信号帧的清浊判决结果确定所述当前语音信号帧的清浊判决阈值。

[0065] 进一步地,所述装置可以包括倍数检测模块;

[0066] 所述倍数检测模块,可以用于对来自于所述基音周期估计/清浊判决模块的基音周期估值进行基音周期倍数检测,得到当前语音信号帧的基音周期估值。

[0067] 其中,所述基音周期估计/清浊判决单元,用于根据来自于所述归一化自相关计算单元的归一化自相关值,以及来自于所述倍数检测模块的前一语音信号帧的基音周期估

值和清浊判决结果反馈,进行当前语音信号帧的基音周期初估和清浊判决,并将当前语音信号帧的基音周期初估值和清浊判决结果发送给所述倍数检测模块;

[0068] 所述倍数检测模块,可以进一步用于接收来自于所述基音周期估计/清浊判决单元的基音周期初估值和清浊判决结果,以及来自于所述归一化自相关计算单元的所述归一化自相关值,并将前一语音信号帧的基音周期估值和清浊判决结果反馈给所述基音周期估计/清浊判决单元;

[0069] 所述归一化自相关计算单元,可以进一步用于将所述归一化自相关值发送给所述倍数检测模块。

[0070] 由上述技术方案可见,本发明技术方案在计算归一化自相关之前,先对当前语音信号帧进行了频谱展平预处理,消除了语音信号中的共振峰结构,如此,可以有效避免由于共振峰的影响而将基音周期估值误估为实际值的倍数或分数,同时也能够提高清浊判决的准确度。

[0071] 进一步的,本发明对基音周期估值进行了倍数检测,以防止将基音周期估计为实际值的倍数,这样,也能够有效提高基音周期估计和清浊判决的准确度。

[0072] 此外,本发明在确定归一化自相关值中的最大值时,充分利用了语音信号的连续性,采用了根据前一语音信号帧的基音周期估值缩小当前语音信号帧的基音周期待选值范围的方式,同时,根据前一语音信号帧的清浊判决结果,适当调整清浊判决的阈值,如此,不仅提高了基音周期估值和清浊判决结果的准确度,而且保证了合成语音或解码语音的特性平稳。

附图说明

[0073] 图 1 为现有通用的语音信号离散时域生成模型示意图。

[0074] 图 2 为本发明实施例一中基音周期估计和清浊判决方法的示例性流程图。

[0075] 图 3 为本发明实施例二中基音周期估计和清浊判决方法的流程示意图。

[0076] 图 4 为本发明实施例三中基音周期估计和清浊判决装置的结构示意图。

[0077] 图 5 为本发明实施例四中基音周期估计和清浊判决装置的结构示意图。

具体实施方式

[0078] 为使本发明的目的、技术方案及优点更加清楚明白,以下参照附图并举实施例,对本发明作进一步详细说明。

[0079] 本发明的主要思想是:在计算归一化自相关之前,先对输入语音信号帧进行频谱展平预处理,消除语音信号中的共振峰结构;然后,对所得到的预处理语音信号帧进行基音周期估计和清浊判决,得到当前语音信号帧的基音周期估值和清浊判决结果。

[0080] 本发明对输入语音信号进行频谱展平预处理得到预处理语音信号之后,对预处理语音信号帧进行基音周期估计和清浊判决的方式有以下四种,下面分别予以介绍:

[0081] 第一种方式是:首先,计算预处理语音信号帧在信号延时长度取为 20 ~ 140 之间的归一化自相关;然后,确定所得到的归一化自相关值中的最大值;最后,比较预先设定的某个阈值与所得到的归一化自相关值中最大值的相对大小,当前者大于后者时,判定当前语音信号帧为清音帧;否则,判定当前语音信号帧为浊音帧,并将归一化自相关值中最大值

所对应的信号延时长度确定为当前帧的基音周期估值。

[0082] 第二种方式是：首先，计算预处理语音信号帧在信号延时长度取为 20 ~ 140 之间的归一化自相关；然后，根据前一语音信号帧的基音周期估值确定当前语音信号帧的基音周期待选值的范围，并确定该范围对应的归一化自相关值中的最大值，以及该最大值对应的基音周期待选值；最后，根据前一语音信号帧的清浊判决结果确定当前语音信号帧的清浊判决阈值，并比较该清浊判决阈值与所得到的归一化自相关值中最大值的相对大小，当前者大于后者时，判定当前语音信号帧为清音帧，否则，判定当前语音信号帧为浊音帧，并将归一化自相关值中最大值对应的基音周期待选值确定为基音周期估值。

[0083] 相对于第一种方式，第二种方式充分利用了语音信号的连续性，采用了根据前一语音信号帧的基音周期估值缩小当前语音信号帧的基音周期待选值范围的方式，同时，根据前一语音信号帧的清浊判决结果，适当调整清浊判决的阈值，如此，不仅提高了基音周期估值和清浊判决结果的准确度，而且保证了合成语音或解码语音的特性平稳。

[0084] 第三种方式是在第一种方式的基础上，对所得到的基音周期估值进行倍数检测。具体而言，当采用第一种方式确定清浊判决结果为浊音帧时，将所得到的基音周期估值作为基音周期初估值和基音周期暂定认可估值，并依次判断对该基音周期初估值进行除以 2、除以 3 等操作所得到的分数值是否超出预先设定的基音周期检测下限，若超出基音周期检测下限，则将该基音周期暂定认可估值作为最终的基音周期估值，否则，根据该分数值是否超出预先设定的当前基音周期下限，确定不同的补偿因子，并将信号延时长度取为基音周期暂定认可估值的归一化自相关值与所述补偿因子之积、与信号延时长度取为该分数值的归一化自相关值作比较，当前者小于后者时，将该分数值作为新的基音周期暂定认可估值，增加分频数，并进行下一轮的判断；这里，补偿因子可以预先设置，设置补偿因子的原则是：在分数值没有超出当前基音周期下限时，使得分数值越小、该分数值被确定为最终的基音周期估值的概率越大。

[0085] 第四种方式是在第二种方式的基础上，对所得到的基音周期估值进行倍数检测。这种方式与第三种方式类似，不同之处在于，本方式中，将采用第二种方式所得到的基音周期估值作为基音周期初估值和基音周期暂定认可估值，进行倍数检测，其具体实施方式请参照第三种方式进行，在此不再赘述。

[0086] 由上述分析可见，在上述四种方式中，以第四种方式为最佳，因此，在下面的实施例中将以第四种方式为例进行说明，其他几种方式的实施请参照下面给出的实施例进行。

[0087] 本发明可适用于采用各种信号采样率的场景中，由于在实际应用中，以信号采样率 8kHz 最为普遍，因此，下面的实施例中，默认情况下均以信号采样率为 8kHz 的情况为例进行说明，对于其他采样速率的情况，参照本发明所提供是实施例作相应修改即可。

[0088] 实施例一：

[0089] 下面通过一个示例性流程，说明如何采用本发明方法进行基音周期估计和清浊判决。

[0090] 图 2 为本发明实施例一中基音周期估计和清浊判决方法的示例性流程图。参见图 2，该方法包括以下步骤：

[0091] 步骤 201：对当前语音信号帧进行滤除高次谐波分量、去除共振峰结构的预处理，

得到预处理语音信号帧。

[0092] 在执行本步骤之前,可以按照现有技术对输入语音信号进行分帧,得到帧大小合适的当前语音信号帧。本步骤中,可以首先对当前语音信号帧进行频谱展平处理,然后采用线性相位低通滤波器滤除经频谱展平处理过的语音信号帧的高频部分,这样,就能够滤除输入语音信号中的高次谐波分量、去除当前语音信号帧中的共振峰结构,使得基音周期估值和清浊判决结果的准确度更高。

[0093] 步骤 202:根据前一语音信号帧的基音周期估值和清浊判决结果,对所得到的预处理语音信号帧进行基音周期初估和清浊判决,得到当前语音信号帧的基音周期初估值和清浊判决结果。

[0094] 本步骤中,可以首先按照现有技术中的公式(1),对所得到的预处理语音信号帧计算归一化自相关,其中, N 取值为 320, τ 的取值范围是 20 ~ 140。

[0095] 然后,确定归一化自相关的最大值以及与该最大值对应的基音周期候选值。由于相邻浊音帧之间的基音周期是缓慢变化的,本发明在确定归一化自相关的最大值时,可以根据前一语音信号帧的基音周期估值适当缩小当前语音信号帧的基音周期候选值的范围:

[0096] 例如,对于浊音帧,可以将小于前一语音信号帧的基音周期估值、且大于等于 20 的整数确定为该基音周期候选值范围的最小值,将大于前一语音信号帧的基音周期估值、且小于等于 140 的整数确定为该基音周期候选值范围的最大值;较佳地,可以根据预先设定的基音周期变化率因子,将前一语音信号帧的基音周期估值减前一语音信号帧的基音周期估值与基音周期变化率因子之积所得到的值作为最小值,并将前一语音信号帧的基音周期估值加前一语音信号帧的基音周期估值与基音周期变化率因子之积所得到的值作为最大值;

[0097] 对于当前语音信号帧为第一个语音信号帧,或者前一语音信号帧为清音帧的情况,将基音周期候选值的范围设为最大范围,即 20 到 140 之间,且包括 20 和 140;

[0098] 然后从缩小的范围所对应的归一化自相关值中选择最大值,进行基音周期初估和清浊判决,以提高基音周期估值和清浊判决结果的准确度;

[0099] 再根据前一语音信号帧的清浊判决结果,适当调整当前语音信号帧的清浊判决阈值,使得当前语音信号帧与前一语音信号帧的清浊判决结果相同的概率大于当前语音信号帧与前一语音信号帧的清浊判决结果不相同的概率;

[0100] 最后,根据归一化自相关值中的最大值以及当前语音信号帧的清浊判决阈值,确定当前语音信号帧清浊判决结果,并根据所确定的当前语音信号帧的清浊判决结果以及归一化自相关值中的最大值对应的基音周期候选值,确定当前语音信号帧的基音周期初估值,即比较清浊判决阈值与归一化自相关值中最大值的相对大小,当前者大于后者时,判定当前语音信号帧为清音帧,否则,判定当前语音信号帧为浊音帧,并将归一化自相关值中最大值对应的基音周期候选值确定为基音周期估值。

[0101] 本发明应用于语音合成或语音编码时,本步骤中根据前一语音信号帧的清浊判决结果适当调整当前语音信号帧的清浊判决阈值的操作,可以保证合成语音或解码语音的特性平稳。

[0102] 步骤 203:对所得到的基音周期初估值进行基音周期倍数检测,得到当前语音信

号帧的基音周期估值。

[0103] 由于对大多数语音信号而言,在信号采样率为 8kHz 时,基音周期长度的取值通常在 40 ~ 80 之间,而且,对于多数应用来说,将基音周期估计为实际值的倍数比将基音周期估计为实际值的分数,给应用系统性能带来的负面影响要大得多,因此,本步骤中采用对基音周期初估值进行基音周期倍数检测的方式来防止将基音周期估计为实际值的倍数。

[0104] 本步骤中进行基音周期倍数检测的方式是:将步骤 202 所得到的基音周期初估值作为基音周期暂定认可估值,并依次判断对该基音周期初估值进行除以 2、除以 3 等操作所得到的分数值是否超出预先设定的基音周期检测下限,若超出基音周期检测下限,则将该基音周期暂定认可估值作为最终的基音周期估值,否则,根据该分数值是否超出预先设定的当前基音周期下限,确定不同的补偿因子,并将信号延时长度取为基音周期暂定认可估值的归一化自相关值与所述补偿因子之积、与信号延时长度取为该分数值的归一化自相关值作比较,当前者小于后者时,将该分数值作为新的基音周期暂定认可估值,增加分频数,并进行下一轮的判断;这里,补偿因子可以预先设置,设置补偿因子的原则是:在分数值没有超出当前基音周期下限的前提下,使得分数值越小、该分数值被确定为最终的基音周期估值的概率越大。

[0105] 本步骤中的基音周期倍数检测能够有效防止将基音周期估计为实际值的倍数,能够提高基音周期估计和清浊判决的准确度。

[0106] 至此,得到当前语音信号帧的基音周期估值和清浊判决结果,结束本发明基音周期估计和清浊判决方法的示例性流程。

[0107] 下面结合附图,通过一个更具体的实施例说明本发明基音周期估计和清浊判决方法的具体实施方式。

[0108] 实施例二:

[0109] 图 3 为本发明实施例二中基音周期估计和清浊判决方法的流程示意图。参见图 3,该方法包括以下步骤:

[0110] 步骤 301:对当前语音信号帧进行预处理,得到预处理语音信号帧。

[0111] 在执行本步骤之前,可以按照现有技术对输入语音信号进行分帧,得到帧大小合适的当前语音信号帧。本步骤中,首先对当前语音信号帧进行频谱展平,然后采用线性相位低通滤波器滤除经频谱展平处理过的语音信号帧的高频部分,以滤除输入语音信号中的高次谐波分量、去除输入语音信号中的共振峰结构。

[0112] 常用的频谱展平方式主要有两种:对当前语音信号帧进行线性预测编码(LPC)逆滤波或中心削波,下面分别予以介绍。假设,以 $s(n)$ 表示当前输入语音信号帧、 $s_w(n)$ 表示经频谱展平处理之后所得到的语音信号帧、 $s'(n)$ 表示经本步骤预处理之后得到的预处理语音信号帧。

[0113] LPC 逆滤波方式中,首先对当前输入语音信号帧 $s(n)$ 进行 10 阶 LPC 分析,求得 $s(n)$ 的 LPC 系数,记为 α_i ,其中, $i = 1 \sim 10$;

[0114] 然后,将 LPC 预测残差信号作为 $s_w(n)$,即有如下关系:

$$[0115] \quad s_w(n) = s(n) - \sum_{i=1}^{10} \alpha_i \cdot s(n-i)$$

[0116] 如此,得到了采用 LPC 逆滤波对当前输入语音信号帧进行频谱展平之后的语音信

号帧。

[0117] 对当前语音信号帧进行中心削波处理的方式中,按照如下所示的式(3):

$$[0118] \quad s_w(n) = \begin{cases} s(n) - C_L & s(n) > C_L \\ 0 & |s(n)| \leq C_L \\ s(n) + C_L & s(n) < -C_L \end{cases} \quad (3)$$

[0119] 将中心削波处理后得到的信号帧作为对当前语音信号帧进行频谱展平之后的语音信号帧。(3)式中, C_L 表示削波电平,是预先设置的一个常量。

[0120] 关于 LPC 滤波和中心削波的原理和实施方式请参照现有技术的有关方式进行,在此不再赘述。

[0121] 本步骤中,较佳地可以使用通带截止频率为 900Hz 的线性相位低通滤波器,滤除经频谱展平处理过的语音信号帧中的高频部分,这里,低通滤波器 $h(k)$ 的系数可以通过 matlab 等软件附带的滤波器设计工具得到。例如,设其抽头数为 N ,则有:

$$[0122] \quad s'(n) = \sum_{k=0}^{N-1} h(k) s_w(n-k) \quad (4)$$

[0123] 步骤 302:对所得到的预处理语音信号帧计算归一化自相关。

[0124] 本步骤中,可以采用与现有技术相同的方式计算预处理语音信号帧的归一化自相关。以 $s'(n)$ 表示预处理语音信号帧在第 n 个样本点的幅值、 $R(\tau)$ 表示信号延时长度为 τ 时的归一化自相关值,则可以根据:

$$[0125] \quad R(\tau) = \frac{\sum_{n=0}^{N-1} s'(n)s'(n+\tau)}{\sum_{n=0}^{N-1} s'^2(n)} \quad (5)$$

[0126] 计算预处理语音信号帧 $s'(n)$ 的归一化自相关 $R(\tau)$ 。(5)式中, N 为计算自相关时使用的信号样本点长度,按照现有技术,一般应至少包含两个基音周期,信号采样频率为 8kHz 时, N 的典型值为 320,即上式中 n 的取值范围为 $0 \sim 319$;

[0127] Σ 表示求和操作;

[0128] τ 表示计算自相关时的信号延时长度,即当前语音信号帧的基音周期候选值,取值范围为 $20 \sim 140$, τ 与基音周期估值的动态范围对应。

[0129] 本步骤中,按照(5)式分别计算信号延时 τ 的取值为 $20 \sim 140$ 之间的各基音周期候选值的归一化自相关 $R(20) \sim R(140)$ 。

[0130] 步骤 303:根据计算得到的归一化自相关、前一语音信号帧的基音周期估值和前一语音信号帧的清浊判决结果,进行基音周期初估和清浊判决,得到当前语音信号帧的基音周期初估值和清浊判决结果。

[0131] 本步骤中,首先需要确定归一化自相关值的最大值,这里,对于前一语音信号帧为清音帧或者当前语音信号帧是第一个语音信号帧的情况,直接将基音周期候选值的范围设为最大范围,即 20 到 140 之间,且包括 20 和 140 ;对于前一语音信号帧为浊音帧的情况,可以利用相邻浊音帧之间的基音周期是缓慢变化的这一特性,根据前一语音信号帧的基音周

期估值适当缩小当前语音信号帧的基音周期待选值的范围。具体而言,其实施方式为:

[0132] 设前一语音信号帧的基音周期估值为 p'_{pre} , 则可以将当前语音信号帧的基音周期待选值范围设置为 p_{min} 到 p_{max} 之间, 其中, p_{min} 表示当前语音信号帧基音周期候选值的最小值、 p_{max} 表示当前语音信号帧基音周期候选值的最大值, p_{min} 和 p_{max} 的取值分别为:

[0133]

$$p_{min} = \lfloor p'_{pre} * (1 - \beta) \rfloor \quad (6.a)$$

[0134]

$$p_{max} = \lceil p'_{pre} * (1 + \beta) \rceil \quad (6.b)$$

[0135] (6. a) 式和 (6. b) 式中, β 表示基音周期变化率因子, 其典型值为 0.25;

[0136] $\lfloor \cdot \rfloor$ 表示向下取整;

[0137] $\lceil \cdot \rceil$ 表示向上取整。

[0138] 在具体应用中, 可以根据实际需要, 对 p_{min} 和 p_{max} 采取其他方式取整, β 的值也可以取为其他值; 而且, p_{min} 和 p_{max} 也可以用其他方式确定。

[0139] 然后, 搜索 $R(p_{min}) \sim R(p_{max})$, 找到最大值, 并将最大值记为 $R(\tau_{max})$, 即:

$$[0140] \quad R(\tau_{max}) = \max \{ R(\tau) \mid \tau = p_{min} \sim p_{max} \} \quad (7)$$

[0141] (7) 式中, \max 表示求最大值。

[0142] 再根据前一语音信号帧的清浊判决结果确定当前语音信号帧的清浊判决阈值。假设以 $voice_flag_pre$ 表示前一语音信号帧的清浊判决结果: $voice_flag_Pre$ 取值为 0 表示前一语音信号帧的清浊判决结果为清音帧, $voice_flag_Pre$ 取值为 1 表示前一语音信号帧的清浊判决结果为浊音帧;

[0143] 以 R_Th 表示当前语音信号帧的清浊判决阈值;

[0144] 以 $Th1$ 和 $Th2$ 表示预先设定的两个阈值, 其中, $0 < Th1 < 1$ 、 $0 < Th2 < 1$ 、且 $Th1 > Th2$, 则本步骤中的清浊判决阈值是根据:

$$[0145] \quad R_Th = \begin{cases} Th1 & voice_flag_pre = 0 \\ Th2 & voice_flag_pre = 1 \end{cases} \quad (8.a)$$

[0146] 确定的。(8. a) 式中, $Th1$ 的典型值为 0.70, $Th2$ 的典型值为 0.60, 当然 $Th1$ 和 $Th2$ 可以设置为其他值, 只要满足 $Th1 > Th2$ 这一关系即可。因为, 只要 $Th1 > Th2$, 就可以使得当前语音信号帧与前一语音信号帧的清浊判决结果相同的概率大于当前语音信号帧与前一语音信号帧的清浊判决结果不相同的概率, 如此, 不仅可以充分利用语音信号的连续性, 提高清浊判决的准确度, 而且可以保证合成语音或解码语音的特性平稳。

[0147] 最后, 比较 $R(\tau_{max})$ 与 R_Th , 得到清浊判决结果和基音周期初估值。假设以 $voice_flag$ 表示当前语音信号帧的清浊判决结果: $voice_flag$ 取值为 0 表示当前语音信号帧的清浊判决结果为清音帧, $voice_flag$ 取值为 1 表示当前语音信号帧的清浊判决结果为浊音帧, 则可以按照如下所示的式 (8. b) 确定清浊判决结果:

$$[0148] \quad voice_flag = \begin{cases} 0 & R(\tau_{max}) < R_Th \\ 1 & R(\tau_{max}) \geq R_Th \end{cases} \quad (8.b)$$

[0149] 假设以 p'_t 表示当前语音信号帧的基音周期初估值, 则 p'_t 可以按照如下所示的式 (9) 确定:

$$[0150] \quad p'_t = \begin{cases} 0 & \text{voice_flag} = 0 \\ \tau_{\max} & \text{voice_flag} = 1 \end{cases} \quad (9)$$

[0151] (9) 式中, 以 p'_t 取值为 0 表示当前语音信号帧为清音帧。

[0152] 步骤 304: 对所得到的基音周期初估值进行基音周期倍数检测, 得到基音周期估值。

[0153] 本步骤对基音周期初估值进行基音周期倍数检测, 以防止将基音周期估计为实际值的倍数。

[0154] 由于在实际应用中, 对于清音帧, 无论信号延时长度取为何值, 与该信号延时长度相对应的归一化自相关值均较小, 因此, 本步骤中的倍数检测主要针对浊音帧, 而根据步骤 303, 当清浊判决的结果为清音帧时, 当前语音信号帧的基音周期初估值为 0, 因此, 在进行基音周期倍数检测之前, 首先判断当前语音信号帧的基音周期初估值是否为 0, 当为 0 时, 无需进行基音周期倍数检测。

[0155] 具体而言, 本步骤中的基音周期倍数检测可以按照如下步骤进行:

[0156] 第 1 步: 设置当前基音周期下限、基音周期暂定认可估值、分频数, 并设定基音周期检测下限、第一补偿因子和第二补偿因子的值;

[0157] 其中, 当前基音周期下限表示当前语音信号帧的基音周期估值下限, 其值根据前一语音信号帧的基音周期估值估算得到, 其值可以为小于前一语音信号帧的基音周期估值、且大于等于 20 的任意整数, 例如, 可以根据前一语音信号帧的基音周期估值以及基音周期变化率因子估算当前基音周期下限;

[0158] 基音周期暂定认可估值表示在对基音周期初估值进行基音周期倍数检测的过程中, 暂时被认可的基音周期估值, 在倍数检测结束后, 基音周期暂定认可估值将作为最终的基音周期估值; 可以将其初值设为当前语音信号帧的基音周期初估值;

[0159] 分频数的初值可以设为 2, 基音周期初估值与分频数之商即为基音周期初估值的某个分数值;

[0160] 基音周期检测下限表示执行基音周期倍数检测的检测范围下限, 只有当基音周期初估值的某个分数在该下限定义的检测范围内时, 才需要继续执行基音周期倍数检测的后续步骤;

[0161] 第一补偿因子和第二补偿因子可以根据实际需要设定, 只要满足: 第一补偿因子的值小于 1、且第二补偿因子的值大于 1 即可; 如此, 在后续的基音周期倍数检测过程中, 当分数值没有超出当前基音周期下限时, 可以使得分数值越小、该分数值被确定为最终的基音周期估值的概率越大, 同时使得基音周期的估值具有一定的连续性。

[0162] 第 2 步: 对当前语音信号帧的基音周期初估值与分频数之商取整, 得到该基音周期初估值的分数值, 判断基音周期初估值的分数值是否大于等于基音周期检测下限, 当基音周期初估值的分数值大于等于基音周期检测下限时, 即当基音周期初估值的分数在检测范围内时, 继续执行第 3 步, 否则, 将基音周期暂定认可估值确定为当前语音信号帧的基音周期估值, 结束本步骤中倍数检测的流程;

[0163] 第 3 步: 如果该基音周期初估值的分数值大于等于当前基音周期下限, 则将第一

补偿因子的值确定为补偿因子,使得在第 4 步的判断中,该分数值被确定为基音周期暂定认可估值的概率增加;否则,将第二补偿因子的值确定为补偿因子,使得在第 4 步的判断中,该分数值被确定为基音周期暂定认可估值的概率降低;

[0164] 第 4 步:将信号延时长度取为基音周期暂定认可估值的归一化自相关值与补偿因子之积,与信号延时长度取为该基音周期初估值的分数值的归一化自相关值作比较,当前者小于后者时,将该基音周期初估值的分数值作为基音周期暂定认可估值;

[0165] 第 5 步:增大分频数,继续执行第 2 步。

[0166] 循环执行上述 5 个步骤的操作,直至循环条件不满足为止,即直至基音周期初估值的某个分数值小于基音周期检测下限为止;这样,最终确定的基音周期估值就是基音周期暂定认可估值。

[0167] 上述基音周期倍数检测也可以用伪代码来表示。例如,若以步骤 303 所确定的 p_{\min} 表示:根据前一语音信号帧的基音周期估值所设定的当前基音周期下限;

[0168] 以 p_{pre} 表示基音周期暂定认可估值,并将 p_{pre} 的初值置为步骤 303 所确定的 τ_{\max} 的值;

[0169] 以 p'_{t} 表示当前语音信号帧的基音周期初估值,其值为 τ_{\max} ;

[0170] 以 div_num 表示分频数,并将分频数的初值置为 2;

[0171] 以 p_{div} 表示基音周期初估值的分数值,其值为对当前语音信号帧的基音周期初估值与分频数之商取整,设其初值为 $\text{int}[p'_{\text{t}}/\text{div_num}]$,其中 $\text{int}[]$ 表示取整操作;

[0172] 以 P_{TH} 表示基音周期检测下限,用于定义检测范围,表示当基音周期初估值的某个分数在该下限定义的检测范围内时,对其进行基音周期倍数检测;由于对大多数语音信号而言,基音周期的取值在 40 ~ 80 之间,因此,本实施例中,将基音周期下限设为 20;

[0173] 以 FAC_1 表示第一补偿因子,其典型值为 0.80;以 FAC_2 表示第二补偿因子,其典型值为 1.2;不管 FAC_1 和 FAC_2 的取值如何变化,只要满足 $\text{FAC}_1 < \text{FAC}_2$ 即可,如此,可以保证较小的基音周期候选者更有可能被确定为最终的基音周期估值,同时使得基音周期的估值具有一定的连续性;以 comp_fac 表示补偿因子;

[0174] 以 p_{est} 表示最终确定的基音周期估值;

[0175] 则本步骤中的基音周期倍数检测可以用伪代码表示如下:

[0176] $P_{\text{TH}} = 20$; $\text{FAC}_1 = 0.8$; $\text{FAC}_2 = 1.2$; $\beta = 0.25$;

[0177] $p_{\min} = \lfloor p'_{\text{pre}} * (1 - \beta) \rfloor$;

[0178] $\text{div_num} = 2$; // div_num 为整数,且 $\text{div_num} \geq 2$

[0179] $p_{\text{pre}} = p'_{\text{t}} = \tau_{\max}$;

[0180] $p_{\text{div}} = \text{int}[p'_{\text{t}}/\text{div_num}]$;

[0181] while($p_{\text{div}} \geq P_{\text{TH}}$) // 判断基音周期初估值的分数是否在检测范围内

[0182] {

[0183] /* 确定补偿因子的值 */

[0184] if($p_{\text{div}} \geq p_{\min}$) $\text{comp_fac} = \text{FAC}_1$;

[0185] else $\text{comp_fac} = \text{FAC}_2$;

[0186] /* 比较归一化自相关,确定是否选用基音周期初估值的分数值更新基音周期暂定认可估值 */


```
[0187]         if(comp_fac*R(p_pre) < R(p_div))
[0188]             p_pre = p_div ;
[0189]         ++div_num ;// 增加分频数
[0190]         p_div = int[p' _t/div_num] ;// 更新基音周期初估值的分数
[0191]     }
[0192]     p_est = pre ;// 确定最终的基音周期估值
```

[0193] 上述伪代码中,将分频数的初值设为 2,并在每一轮判断之后,令分频数递增 1,可以比较全面地分析基音周期初估值的所有符合条件的分数的情况,使最终的基音周期估值更加准确;

[0194] 此外,设置了两个不同的补偿因子,小于 1 的补偿因子可以使判决成立的可能性增加,则以 p_div 更新 p_pre 的概率增加;大于 1 的补偿因子可以使判决不成立的可能性增加,则维持 p_pre、拒绝以 p_div 更新 p_pre 的概率增加,因为此时 p_div 已经小于由前一帧基音周期估值确定的检测范围的下限 p_min,表明 p_div 是正确估值的可能性较低,可见,维持 p_pre 不变将提高最终的基音周期估值的准确性。

[0195] 由于对大多数语音信号而言,将基音周期估计为实际值的倍数比将基音周期估计为实际值的分数,给应用系统性能带来的负面影响要大得多,因此,本步骤中对基音周期初估值进行基音周期倍数检测之后,可以有效防止将基音周期估计为实际值的倍数,能够提高基音周期估计和清浊判决的准确度。

[0196] 至此,得到当前语音信号帧的基音周期估值和清浊判决结果,结束本实施例基音周期估计和清浊判决方法的流程。

[0197] 由上述实施例可见,本实施例在计算归一化自相关之前,先对当前语音信号帧进行了频谱展平预处理,消除了语音信号中的共振峰结构,如此,可以有效避免由于共振峰的影响而将基音周期估值误估为实际值的倍数或分数,同时也能够提高清浊判决的准确度;

[0198] 而且,本实施例的基音周期估计分两步进行,先对基音周期进行初估得到基音周期初估值,然后对基音周期初估值进行基音周期估值的倍数检测,这样,可以防止将基音周期估计为实际值的倍数,同时能够有效提高基音周期估计和清浊判决的准确度;

[0199] 此外,本实施例充分利用了语音信号的连续性,采用了根据前一语音信号帧的基音周期估值缩小基音周期初估时的当前语音信号帧的基音周期待选值范围,同时,根据前一语音信号帧的清浊判决结果,适当调整清浊判决的阈值,如此,不仅提高了基音周期估值和清浊判决结果的准确度,而且保证了合成语音或解码语音的特性平稳。

[0200] 下面通过两个装置实施例对本发明的基音周期估计和清浊判决装置进行说明。

[0201] 实施例三:

[0202] 本实施例中的基音周期估计和清浊判决装置对应于本发明基音周期估计和清浊判决的第一种方式。

[0203] 图 4 为本发明实施例三中基音周期估计和清浊判决装置的结构示意图。

[0204] 参见图 4,该装置包括:预处理模块 410 和基音周期估计/清浊判决单元。

[0205] 图 4 所示装置中,预处理模块 410,用于对当前语音信号帧进行滤除高次谐波分量、去除共振峰结构的预处理,并将所得到的预处理语音信号帧发送给基音周期估计/清浊判决模块 420;

[0206] 基音周期估计 / 清浊判决模块 420, 用于对来自于预处理模块 410 的预处理语音信号帧进行基音周期估计和清浊判决, 得到所述当前语音信号帧的基音周期估值和清浊判决结果。

[0207] 由上述本实施例可见, 本实施例的预处理模块可以在进行基音周期估计和清浊判决之前, 先对当前语音信号帧进行频谱展平预处理, 消除语音信号中的共振峰结构, 如此, 可以有效避免由于共振峰的影响而将基音周期估值误估为实际值的倍数或分数, 同时也能提高清浊判决的准确度。

[0208] 本实施例的基音周期估计 / 清浊判决模块 420 可以进一步划分为归一化自相关计算单元和基音周期估计 / 清浊判决单元, 并且, 为了对基音周期进行倍数检测, 可以在图 4 所示装置中增加进行基音周期倍数检测的模块, 在下面的实施例四中将对这种装置的具体实施方式进行说明。

[0209] 实施例四:

[0210] 图 5 为本发明实施例四中基音周期估计和清浊判决装置的结构示意图。参见图 5, 该装置包括: 预处理模块 510、基音周期估计 / 清浊判决模块 520 和倍数检测模块 530, 其中, 基音周期估计 / 清浊判决模块 520 中进一步包括归一化自相关计算单元 521 和基音周期估计 / 清浊判决单元 522。

[0211] 图 5 所示装置中, 预处理模块 510, 用于对当前语音信号帧进行滤除高次谐波分量、去除共振峰结构的预处理, 并将所得到的预处理语音信号帧发送给基音周期估计 / 清浊判决模块 520 中的归一化自相关计算单元 521;

[0212] 基音周期估计 / 清浊判决模块 520 中的归一化自相关计算单元 521, 用于计算来自于预处理模块 510 的预处理语音信号帧的归一化自相关值, 并将计算结果发送给基音周期估计 / 清浊判决单元 522 和倍数检测模块 530;

[0213] 基音周期估计 / 清浊判决模块 520 中的基音周期估计 / 清浊判决单元 522, 接收到来自于归一化自相关计算单元 521 的归一化自相关值之后, 根据来自于倍数检测模块 530 的前一语音信号帧的基音周期估值和清浊判决结果反馈, 进行当前语音信号帧的基音周期估计和清浊判决, 并将当前语音信号帧的基音周期初估值和清浊判决结果发送给倍数检测模块 530; 更具体地, 基音周期估计 / 清浊判决模块 520 可以根据来自于倍数检测模块 530 的前一语音信号帧的基音周期估值反馈, 缩小当前语音信号帧的基音周期待选值的范围;

[0214] 倍数检测模块 530, 用于根据来自于归一化自相关计算单元 521 的归一化自相关值、对来自于基音周期估计 / 清浊判决单元 522 的基音周期初估值进行基音周期倍数检测, 确定当前语音信号帧的基音周期估值。

[0215] 由上述实施例可见, 本实施例的预处理模块可以在计算归一化自相关之前, 先对当前语音信号帧进行频谱展平预处理, 消除语音信号中的共振峰结构, 如此, 可以有效避免由于共振峰的影响而将基音周期估值误估为实际值的倍数或分数, 同时也能提高清浊判决的准确度;

[0216] 而且, 本实施例的基音周期估计分两步进行, 先由基音周期估计 / 清浊判决模块对基音周期进行估计得到基音周期初估值, 然后由倍数检测模块对基音周期初估值进行倍数检测, 这样, 可以防止将基音周期估计为实际值的倍数, 同时能够有效提高基音周期估计和清浊判决的准确度;

[0217] 此外,本实施例充分利用了语音信号的连续性,从倍数检测模块向基音周期估计/清浊判决模块反馈前一语音信号帧的基音周期估值和清浊判决结果,这样,基音周期估计/清浊判决模块可以根据前一语音信号帧的基音周期估值和清浊判决结果,进行基音周期估计和适当调整清浊判决的阈值,如此,不仅提高了基音周期估值和清浊判决结果的准确度,而且保证了合成语音或解码语音的特性平稳。

[0218] 以上所述仅为本发明的较佳实施例而已,并非用于限定本发明的保护范围。凡在本发明的精神和原则之内所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

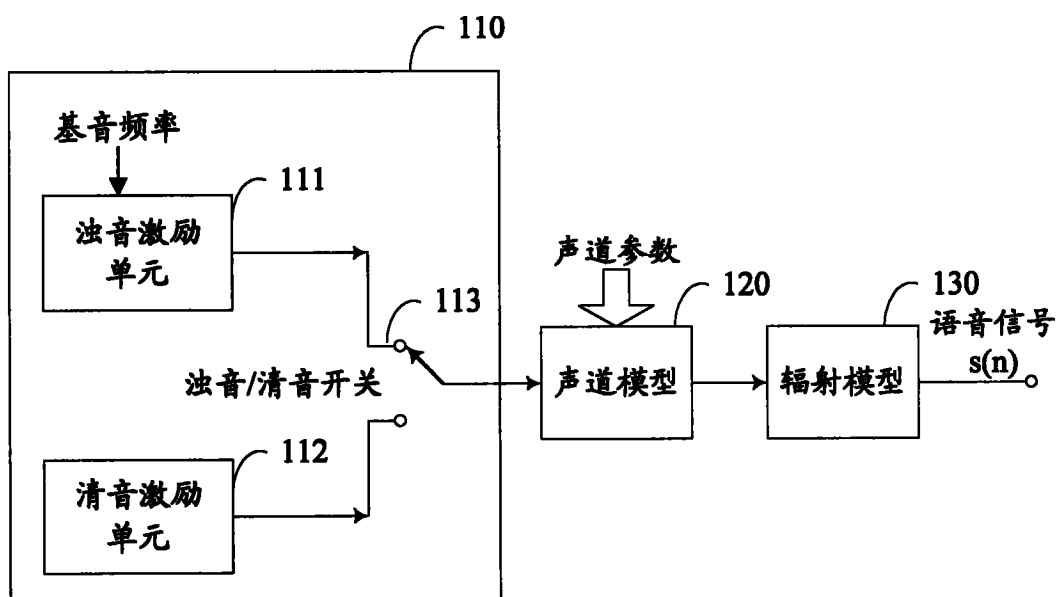


图 1

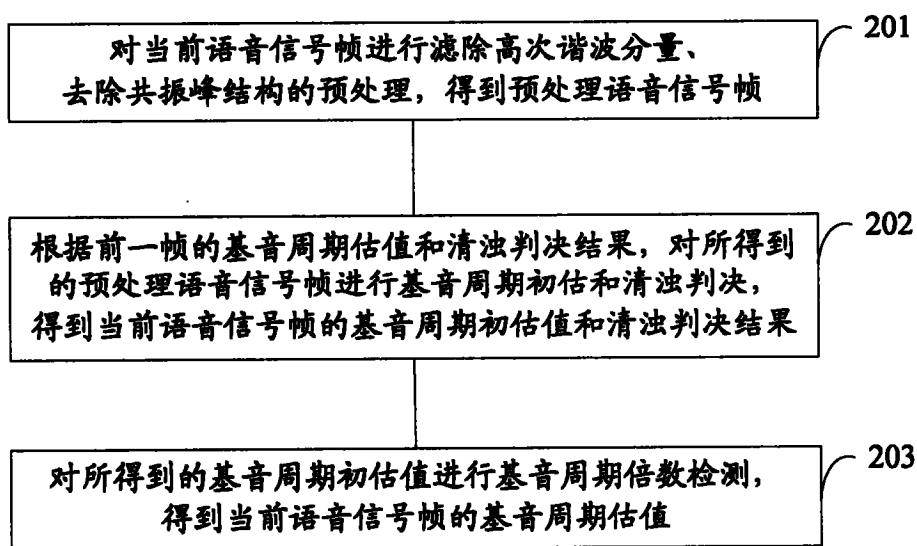


图 2

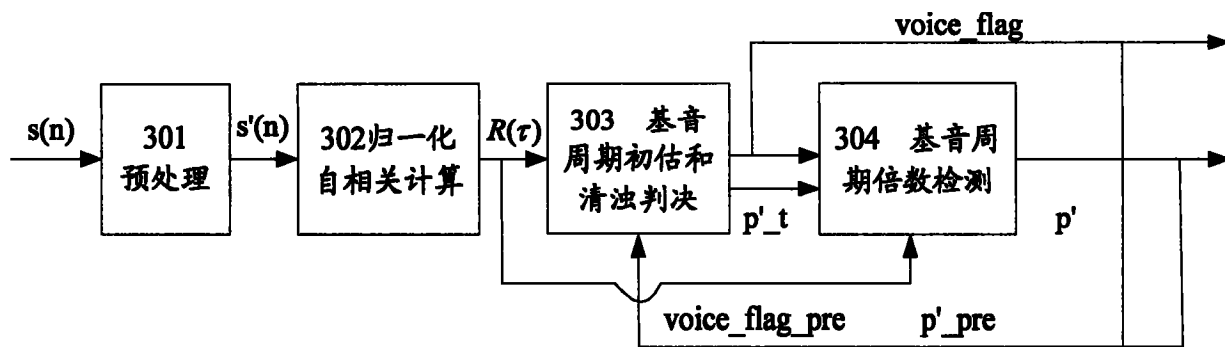


图 3

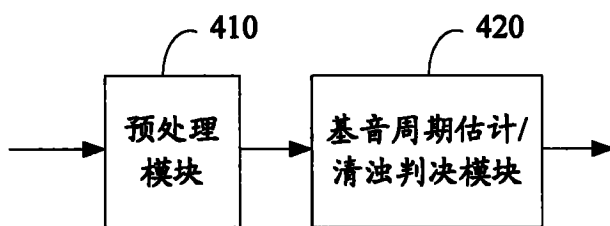


图 4

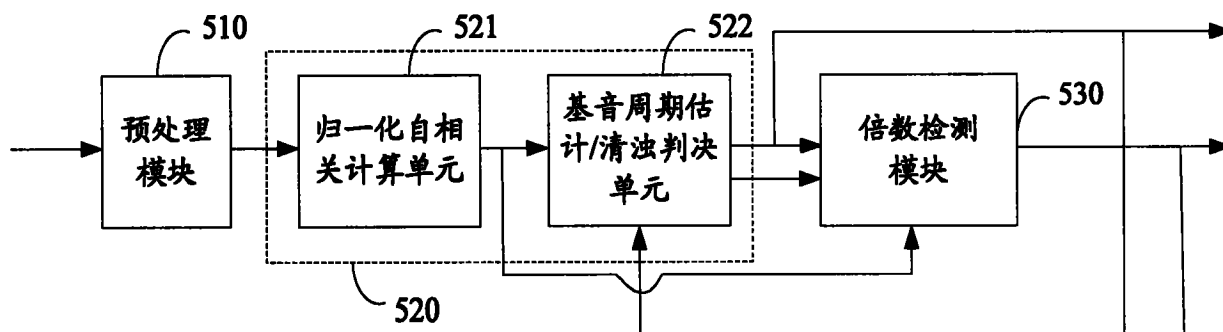


图 5