

(12) 发明专利申请

(10) 申请公布号 CN 103050121 A

(43) 申请公布日 2013. 04. 17

(21) 申请号 201210592909. 1

(22) 申请日 2012. 12. 31

(71) 申请人 北京迅光达通信技术有限公司  
地址 100081 北京市海淀区四道口路甲 5 号  
4 层 401 (住宅)

(72) 发明人 洪小斌

(74) 专利代理机构 北京路浩知识产权代理有限公司 11002

代理人 王莹

(51) Int. Cl.  
G10L 19/04 (2013. 01)

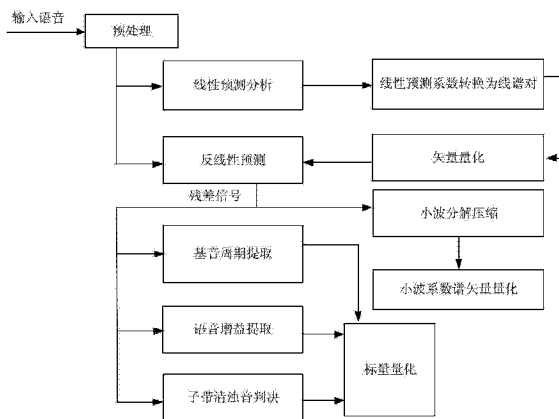
权利要求书 2 页 说明书 9 页 附图 2 页

(54) 发明名称

线性预测语音编码方法及语音合成方法

(57) 摘要

本发明公开了一种线性预测语音编码方法及一种语音合成方法。所述语音编码方法包括：对语音进行预处理；对预处理后的语音进行二阶反线性预测，得到残差信号；对残差信号进行小波分解压缩，得到小波系数，并对小波系数进行矢量量化，同时，计算残差信号的基音周期、增益参数及各子带的清浊音特性并分别进行标量量化。所述语音合成方法基于所述语音编码方法。采用本发明的技术方案，能够减小噪声对解码语音质量的影响，抑制清浊音判决错误时的音质恶化，并改善对清音语音或背景噪声的编码性能。



1. 一种线性预测语音编码方法,其特征在于,所述语音编码方法包括以下步骤:  
S101、对语音进行预处理以去除直流成分和工频干扰;  
S102、对预处理后的语音进行二阶反线性预测,得到残差信号;  
S103、对所述残差信号进行小波分解压缩,得到小波系数,并对所述小波系数进行矢量量化,  
计算所述残差信号的基音周期,并对所述基音周期进行标量量化,  
计算所述残差信号的增益参数,并对所述增益参数进行标量量化,  
将所述残差信号划分为若干个子带,对各子带进行清浊音判决,得到各子带的清浊音特性并进行标量量化。
2. 根据权利要求1所述的语音编码方法,其特征在于,步骤S102进一步包括:  
对所述预处理后的语音进行线性预测分析,得到线性预测系数,然后将所述线性预测系数转换为线谱频率对,并对所述线谱频率对进行矢量量化。
3. 根据权利要求2所述的语音编码方法,其特征在于,步骤S102中,所述线性预测分析具体包括:  
对所述预处理后的语音采用汉明窗进行加窗处理,并将加窗后的语音信号进行自相关计算,利用 Levinson-Durbin 算法计算出 10 阶线性预测系数,然后将所述 10 阶线性预测系数乘以  $0.994^{i+1}$  ( $i = 1, 2, \dots, 10$ ) 以获得带宽扩展的线性预测系数。
4. 根据权利要求1所述的语音编码方法,其特征在于,步骤S103中,所述小波分解压缩具体包括:  
选取所述残差信号的样点进行一阶小波分解,采用 dB10 小波基,得到小波系数,并将前 100 个小波系数进行压缩分析。
5. 根据权利要求1或4所述的语音编码方法,其特征在于,步骤S103中,对所述小波系数进行矢量量化具体包括:  
先将所述小波系数转换为小波激励幅度谱,然后将所述小波激励幅度谱进行矢量量化,量化时码本搜索采用全搜索算法,失真度量采用加权欧式距离。
6. 根据权利要求1所述的语音编码方法,其特征在于,步骤S103中,计算所述残差信号的基音周期具体包括:  
对所述残差信号采用傅里叶变换进行频谱分析,并对频谱幅值进行反傅里叶变换,将得到的残差信号的自相关峰值作为整数基音;  
在所述整数基音  $\pm 1$  的范围内进行搜索,通过对所述残差信号进行插值及局部相关,得到基音周期。
7. 根据权利要求6所述的语音编码方法,其特征在于,步骤S103中,计算所述残差信号的基音周期进一步包括:  
利用所述基音周期对所述残差信号进行基音峰值及其谐振峰值搜索,并将各峰值的差的平均值作为最终的基音周期。
8. 根据权利要求1所述的语音编码方法,其特征在于,步骤S103中,对各子带进行清浊音判决具体包括:  
计算各子带信号在所述基音周期附近的最大归一化自相关值;  
计算各子带包络信号在所述基音周期附近的最大归一化自相关值;

采用阈值比较法,根据所述各子带信号在所述基音周期附近的最大归一化自相关值,以及所述各子带包络信号在所述基音周期附近的最大归一化自相关值,对各子带进行清浊音判决。

9. 一种基于权利要求 2 或 3 所述的线性预测语音编码方法的语音合成方法,其特征在于,所述语音合成方法包括以下步骤:

S201、对量化的线谱频率对、小波系数、基音周期、增益参数及各子带的清浊音特性进行解码,得到线谱频率对、小波激励幅度谱、基音周期、增益参数及各子带的清浊音特性;

S202、利用所述小波激励幅度谱、所述基音周期及所述各子带的清浊音特性,合成小波激励信号;

S203、利用所述线谱频率对,对所述小波激励信号进行反线性预测,得到合成语音;

S204、对所述合成语音进行谱增强和相位调整。

10. 根据权利要求 9 所述的语音合成方法,其特征在于,步骤 S202 具体包括:

根据所述各子带的清浊音特性,对各子带的清音成分和浊音成分进行滤波混合,得到小波频谱;

对所述小波频谱进行反傅里叶变换,得到小波系数,并利用 dB10 小波基复合得到所述小波激励信号。

## 线性预测语音编码方法及语音合成方法

### 技术领域

[0001] 本发明涉及语音编码技术,特别涉及一种线性预测语音编码方法及一种语音合成方法。

### 背景技术

[0002] 随着信息社会和通信技术的高速发展,频率资源显得愈加宝贵。在数字移动通信和语音存储领域中,为了有效利用通信带宽或存储空间,使用语音编码装置压缩语音信号的传输带宽或降低电话通道的传输码率,提高效率进行编码,一直是人们追求的目标。随着通信网络用户数量的增加、网络业务综合化与多样化,网络带宽与系统容量、服务质量的矛盾日益突出,传统的语音压缩编码技术已不能满足不断变得拥挤的传输信道要求。因此,如何在不牺牲语音通话质量的前提下尽可能降低其传输的比特速率是重要的研究课题。近十年来,中比特率(4.8kbps ~ 16kbps)语音编码算法研究已取得了长足的进步并有了广泛应用,同时低比特率特别是 2.4kbps 以下语音编码算法逐渐成为研究焦点。随着运行编码算法的处理芯片运算速度的迅速提高,基于线性预测混合编码技术的算法渐渐成为了低比特率语音编码算法的主流。

[0003] 线性预测编码(Linear Prediction Coding, LPC)的基础是假设声音信号(浊音)是音管末端的蜂鸣器产生的,偶尔伴随有嘶嘶声与爆破声(齿擦音与爆破音),声带之间的声门产生不同强度(音量)与频率(音调)的声音,喉咙与嘴唇组成共鸣声道。嘶嘶声与爆破声通过舌头、嘴唇以及喉咙的作用产生出来。线性预测编码通过估计共振峰、剔除它们在语音信号中的作用、估计保留的蜂鸣音强度与频率来分析语音信号。剔除共振峰的过程称为逆滤波,经过这个过程剩余的信号称为残差信号。描述共振峰、残差信号参数以及线性预测系数可以保存、发送到接收方。接收方通过逆向的过程合成语音信号,共振峰、残差信号作为激励源产生源信号,使用线性预测系数作为声道的滤波器,源信号经过滤波器的处理就得到语音信号。

[0004] 根据对激励信号描述方式的不同,线性预测语音编码方法主要分为 LPC-10、码激励线性预测编码(Code Excited Linear Prediction, CELP)、混合激励(Mixed Excited Linear Prediction, MELP)、正弦激励(Sinusoidal Excited Linear Prediction, SELP)和多带激励(Multi-BandExcitation)等。这些语音编码方式是将语音划分为某个帧长(20ms ~ 50ms 左右),对每个帧进行语音的线性预测,用已知码本对线性预测矢量和每帧的通过线性预测所得的预测残差(激励信号)进行编码。

[0005] 图 1 是现有的基于线性预测的语音编码方法的基本框图,这些方法除了残差参数的提取方法不同外,其它参数的提取都基本相同。在图 1 中,激励信号用残差参数、原始语音的基音周期、原始语音的增益和原始语音的各子带的清浊音来表示,残差参数用来描述残差中浊音的谐波分量,清音用噪声代替。

[0006] 现有的基于线性预测的声码器语音质量强烈地依赖于原始语音的噪声强度,原始语音信噪比较差时,清浊音判决错误、基音提取错误会导致严重的变调失真,使合成自然度

下降。这些技术中产生激励信号的基音周期、增益和子带的清浊音都是从原始语音提取的，接收端还原激励信号时部分参数来源于原始语音，部分参数来源于残差信号，使解码的语音质量受到限制。

## 发明内容

[0007] (一)所要解决的技术问题

[0008] 本发明的目的在于提供一种线性预测语音编码方法及一种语音合成方法，能够减小噪声对解码语音质量的影响，抑制清浊音判决错误时的音质恶化，并改善对清音语音或背景噪声的编码性能。

[0009] (二)技术方案

[0010] 为了解决上述技术问题，本发明提出了一种线性预测语音编码方法，所述语音编码方法包括以下步骤：

[0011] S101、对语音进行预处理以去除直流成分和工频干扰；

[0012] S102、对预处理后的语音进行二阶反线性预测，得到残差信号；

[0013] S103、对所述残差信号进行小波分解压缩，得到小波系数，并对所述小波系数进行矢量量化，

[0014] 计算所述残差信号的基音周期，并对所述基音周期进行标量量化，

[0015] 计算所述残差信号的增益参数，并对所述增益参数进行标量量化，

[0016] 将所述残差信号划分为若干个子带，对各子带进行清浊音判决，得到各子带的清浊音特性并进行标量量化。

[0017] 可选的，步骤 S102 进一步包括：

[0018] 对所述预处理后的语音进行线性预测分析，得到线性预测系数，然后将所述线性预测系数转换为线谱频率对，并对所述线谱频率对进行矢量量化。

[0019] 可选的，步骤 S102 中，所述线性预测分析具体包括：

[0020] 对所述预处理后的语音采用汉明窗进行加窗处理，并将加窗后的语音信号进行自相关计算，利用 Levinson-Durbin 算法计算出 10 阶线性预测系数，然后将所述 10 阶线性预测系数乘以  $0.994^{i+1}$  ( $i = 1, 2, \dots, 10$ ) 以获得带宽扩展的线性预测系数。

[0021] 可选的，步骤 S103 中，所述小波分解压缩具体包括：

[0022] 选取所述残差信号的样点进行一阶小波分解，采用 dB10 小波基，得到小波系数，并将前 100 个小波系数进行压缩分析。

[0023] 可选的，步骤 S103 中，对所述小波系数进行矢量量化具体包括：

[0024] 先将所述小波系数转换为小波激励幅度谱，然后将所述小波激励幅度谱进行矢量量化，量化时码本搜索采用全搜索算法，失真度量采用加权欧式距离。

[0025] 可选的，步骤 S103 中，计算所述残差信号的基音周期具体包括：

[0026] 对所述残差信号采用傅里叶变换进行频谱分析，并对频谱幅值进行反傅里叶变换，将得到的残差信号的自相关峰值作为整数基音；

[0027] 在所述整数基音  $\pm 1$  的范围内进行搜索，通过对所述残差信号进行插值及局部相关，得到基音周期。

[0028] 可选的，步骤 S103 中，计算所述残差信号的基音周期进一步包括：

[0029] 利用所述基音周期对所述残差信号进行基音峰值及其谐振峰值搜索,并将各峰值的差的平均值作为最终的基音周期。

[0030] 可选的,步骤 S103 中,对各子带进行清浊音判决具体包括:

[0031] 计算各子带信号在所述基音周期附近的最大归一化自相关值;

[0032] 计算各子带包络信号在所述基音周期附近的最大归一化自相关值;

[0033] 采用阈值比较法,根据所述各子带信号在所述基音周期附近的最大归一化自相关值,以及所述各子带包络信号在所述基音周期附近的最大归一化自相关值,对各子带进行清浊音判决。

[0034] 本发明同时提出了一种基于所述语音编码方法的语音合成方法,所述语音合成方法包括以下步骤:

[0035] S201、对量化的线谱频率对、小波系数、基音周期、增益参数及各子带的清浊音特性进行解码,得到线谱频率对、小波激励幅度谱、基音周期、增益参数及各子带的清浊音特性;

[0036] S202、利用所述小波激励幅度谱、所述基音周期及所述各子带的清浊音特性,合成小波激励信号;

[0037] S203、利用所述线谱频率对,对所述小波激励信号进行反线性预测,得到合成语音;

[0038] S204、对所述合成语音进行谱增强和相位调整。

[0039] 可选的,步骤 S202 具体包括:

[0040] 根据所述各子带的清浊音特性,对各子带的清音成分和浊音成分进行滤波混合,得到小波频谱;

[0041] 对所述小波频谱进行反傅里叶变换,得到小波系数,并利用 dB10 小波基复合得到所述小波激励信号。

[0042] (三) 有益效果

[0043] 本发明所述技术方案具有如下优点:

[0044] 1、利用小波压缩去除了语音信号的背景噪声,同时去除了冗余信息,采用小波系数谱作为激励源能更好地描述原始语音信号。由于采用小波分解压缩方法来产生激励信号,在同样的量化比特数下,能够比现有技术更准确地描述残差信号,从而能够提高解码后的语音质量。

[0045] 2、提取整数基音时对残差信号采用傅里叶变换进行频谱分析,对频谱幅值进行反 FFT 变换,得到残差信号的自相关峰值位置作为整数基音,相对于背景技术中提取的整数基音更加准确,从而能够显著提高合成后的语音质量。

[0046] 3、产生小波激励信号的基音周期、增益参数和各子带的清浊音特性均从残差信号提取,提高了解码的语音质量。

## 附图说明

[0047] 图 1 是现有的基于线性预测的语音编码方法的基本框图。

[0048] 图 2 是本发明所述线性预测语音编码方法的基本框图。

[0049] 图 3 是本发明所述语音合成方法的基本框图。

## 具体实施方式

[0050] 下面结合附图和实施例,对本发明的具体实施方式作进一步详细描述。

[0051] 本发明提供了一种线性预测语音编码方法,如图 2 所示,所述语音编码方法包括以下步骤:

[0052] S101、对语音进行预处理以去除直流成分和工频干扰;

[0053] S102、对预处理后的语音进行线性预测分析,得到线性预测系数,然后将所述线性预测系数转换为线谱频率对,并对所述线谱频率对进行矢量量化,

[0054] 对所述预处理后的语音进行二阶反线性预测,得到残差信号;

[0055] S103、对所述残差信号进行小波分解压缩,得到小波系数,并对所述小波系数进行矢量量化,

[0056] 计算所述残差信号的基音周期,并对所述基音周期进行标量量化,

[0057] 计算所述残差信号的增益参数,并对所述增益参数进行标量量化,

[0058] 将所述残差信号划分为若干个子带,对各子带进行清浊音判决,得到各子带的清浊音特性并进行标量量化。

[0059] 优选的,步骤 S102 中,所述线性预测分析具体包括:

[0060] 对所述预处理后的语音采用汉明窗进行加窗处理,并将加窗后的语音信号进行自相关计算,利用 Levinson-Durbin 算法计算出 10 阶线性预测系数,然后将所述 10 阶线性预测系数乘以  $0.994^{i+1}$  ( $i = 1, 2, \dots, 10$ ) 以获得带宽扩展的线性预测系数。

[0061] 优选的,步骤 S102 中对所述线谱频率对进行矢量量化时,采用 3 级码本,码本搜索采用加权欧式距离准则。

[0062] 优选的,步骤 S103 中,所述小波分解压缩具体包括:

[0063] 选取所述残差信号的样点进行一阶小波分解,采用 dB10 小波基,得到小波系数,并将前 100 个小波系数进行压缩分析。

[0064] 优选的,步骤 S103 中,对所述小波系数进行矢量量化具体包括:

[0065] 先将所述小波系数转换为小波激励幅度谱,然后将所述小波激励幅度谱进行矢量量化,量化时码本搜索采用全搜索算法,失真度量采用加权欧式距离。

[0066] 优选的,步骤 S103 中,计算所述残差信号的基音周期具体包括:

[0067] 对所述残差信号采用傅里叶变换进行频谱分析,并对频谱幅值进行反傅里叶变换,将得到的残差信号的自相关峰值作为整数基音;

[0068] 在所述整数基音  $\pm 1$  的范围内进行搜索,通过对所述残差信号进行插值及局部相关,得到基音周期。

[0069] 优选的,步骤 S103 中,计算所述残差信号的基音周期进一步包括:

[0070] 利用所述基音周期对所述残差信号进行基音峰值及其谐振峰值搜索,并将各峰值的差的平均值作为最终的基音周期。

[0071] 优选的,步骤 S103 中,对各子带进行清浊音判决具体包括:

[0072] 计算各子带信号在所述基音周期附近的最大归一化自相关值;

[0073] 计算各子带包络信号在所述基音周期附近的最大归一化自相关值;

[0074] 采用阈值比较法,根据所述各子带信号在所述基音周期附近的最大归一化自相关

值,以及所述各子带包络信号在所述基音周期附近的最大归一化自相关值,对各子带进行清浊音判决。

[0075] 本发明提出的技术方案可由一种小波激励线性预测(WaveletExcited Linear Prediction,WELP)语音编码/解码装置,简称 WELP 声码器,来加以实现。WELP 声码器需要提取的参数主要包括线谱频率(Line Spectrum Frequency,LSF)、残差小波激励系数(即小波系数)、小波系数周期(即基音周期)、小波增益(即增益参数)以及小波系数周期性标志(即各子带的清浊音特性)。

[0076] 下面结合 WELP 声码器,分别对各步骤的具体实现方法进行详细的举例说明。

[0077] 在步骤 S101 中,WELP 声码器将输入语音经过高通滤波器以去掉工频干扰,预处理由一个截止频率为 60Hz 的高通滤波器完成,用以去除直流成分和 50Hz 的工频干扰并进行高频提升,高通滤波器的频率响应函数如公式(1)所示:

$$[0078] \quad H(z) = \frac{1-z^{-1}}{1-0.95z^{-1}} \quad (1)$$

[0079] 公式(1)中  $z$  为  $Z$  变换的频率变量。上述高通滤波器的频率响应在 50Hz 处有 50dB 的抑制。

[0080] 在步骤 S102 中,线性预测分析用于提取线性预测滤波器的系数(即 LPC 系数),对选取的语音帧信号,滤波器系数选取 10 阶,其传递函数如公式(2)所示:

$$[0081] \quad A(z) = 1 - \sum_{i=1}^{10} a_i z^{-i} \quad (2)$$

[0082] 其中,  $\{a_i\}$ , ( $i = 1, 2, \dots, 10$ ) 即为 LPC 系数。

[0083] WELP 声码器将经过预处理的语音信号采用 200 点汉明窗进行加窗处理,汉明窗  $w(n)$  的表达式如公式(3)所示:

$$[0084] \quad w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{199}\right) \quad (3)$$

[0085] 其中,样点序号  $n=0, 1, \dots, 199$ 。加窗后的语音信号  $s_w(n)$  用于自相关函数的计算,如公式(4)所示:

$$[0086] \quad r(k) = \sum_{n=k}^{199} s_w(n) s_w(n-k) \quad (4)$$

[0087] 其中,  $s_w(n)$  为加窗后的语音信号,  $r(k)$  为自相关函数,  $n$  和  $k$  分别为序号。LPC 系数通过求解公式(5)得到:

$$[0088] \quad \sum_{i=1}^{10} a_i r(|i-k|) = -r(k) \quad k=1, 2, \dots, 10 \quad (5)$$

[0089] 公式(5)中  $a_i$  为 LPC 系数,  $r$  为自相关函数,利用 Levinson-Durbin 算法解上式构成的方程组即可得到  $a_i$ ,  $i = 1, 2, \dots, 10$ 。将得到的 10 阶线性预测系数  $a_i$  乘以  $0.994^{i+1}$ ,进行带宽扩展,这有助于改善共振峰结构和便于 LSF 参数量化。

[0090] WELP 声码器将提取的 LPC 系数转化为线谱频率(LSF),LSF 把圆内随机分布的 LPC 变换到圆周上,适合用于量化。LSF 参数定义为公式(6)和公式(7)的根:

$$[0091] \quad F_1(z) = \frac{A(z) + z^{-11} A(z^{-1})}{1 + z^{-1}} \quad (6)$$



$$[0092] \quad F_2(z) = \frac{A(z) - z^{-11}A(z^{-1})}{1 - z^{-1}} \quad (7)$$

[0093] 其中,  $A(z)$  即为公式(2)所述线性预测滤波器的传递函数。

[0094] 接下来, WELP 声码器对 10 维的线谱频率参数使用矢量量化, 采用 3 级码本, 各级所用比特数分别为 7、7、6, 总计 20 比特。码本搜索采用加权欧式距离准则, 如公式(8)所示:

[0095]

$$d_{\text{isp}}(\mathbf{l}, \hat{\mathbf{l}}) = \sum_{i=1}^{10} w(i) [l(i) - \hat{l}(i)]^2 \quad (8)$$

[0096] 其中,  $d_{\text{isp}}$  为欧氏加权距离,  $\mathbf{l}$  表示待量化的 LSF 矢量,  $\hat{\mathbf{l}}$  表示码本矢量,  $l(i)$  表示待量化的第  $i$  维 LSF,  $\hat{l}(i)$  表示码本中第  $i$  维 LSF,  $w(i)$  是每一维的权重,  $w(i)$  如公式(9)所示:

$$[0097] \quad w(i) = \begin{cases} p(f_i)^{0.3}, & i=1-8 \\ 0.64 \cdot p(f_i)^{0.3}, & i=9 \\ 0.16 \cdot p(f_i)^{0.3}, & i=10 \end{cases} \quad (9)$$

[0098] 其中,  $p(f_i)$  表示线性预测滤波器在第  $i$  个线谱频率参数对应频率处的功率谱密度。

[0099] 同时, WELP 声码器采用截止频率为 1000Hz 的低通滤波器对语音信号进行滤波, 去除高频信号对基音估值的影响; 然后对该信号进行二阶反线性预测, 去除共振峰对基音谐波的影响, 从而得到残差信号。

[0100] 在步骤 S103 中, 首先对残差信号进行小波分解和压缩。残差信号是去掉谱包络后的信号, 对浊音信号, 频谱具有周期性。取当前帧的 160 点和前一帧的 40 个残差信号样点进行 1 阶小波分解, 采用 dB10 小波基, 得到 218 个小波系数, 然后取前 100 个小波系数进行压缩分析。

[0101] 接下来, 将小波系数变换到频域, 并将基频乘 2, 计算小波系数功率谱在倍频后基频及其谐波附近的 10 个频谱峰值, 当峰值数小于 10 时, 对剩余的数填充为 1。

[0102] WELP 声码器对 10 维的小波激励幅度谱采用矢量量化方法, 码本搜索采用全搜索算法, 失真度量采用加权欧式距离, 如公式(10)所示:

[0103]

$$d(\mathbf{A}, \hat{\mathbf{A}}) = \sum_{i=1}^{10} w(i) \cdot [A(i) - \hat{A}(i)]^2 \quad (10)$$

[0104] 其中,  $d(\mathbf{A}, \hat{\mathbf{A}})$  表示失真度量,  $A(i)$  表示待量化的的小波激励幅度谱,  $\hat{A}(i)$  表示码本矢量,  $w(i)$  表示加权系数。

[0105] 在计算基音周期时, 采用傅里叶变换对去除共振峰的信号进行频谱分析, 并对频谱幅值进行反 FFT 变换, 得到该信号的自相关峰值, 如公式(11)所示:

$$[0106] \quad r(\tau) = \int \left[ \int x(t) e^{-j\omega t} dt \right]^2 e^{-j\omega\tau} d\omega \quad (11)$$

[0107] 其中,  $r(\tau)$  表示自相关函数,  $x(t)$  表示去除共振峰的信号值,  $\omega$  为信号角频率,  $t$  为时间,  $r(\tau)$  在时间轴上的相关曲线峰值与时间轴中点的差即为基音周期。

[0108] 上述方法得到的基音周期值为整数,而语音的基音周期可能不是一个整数,需要在该整数  $\pm 1$  的范围内进行搜索,得到准确的基音值,通过对去除共振峰的信号进行插值并局部相关,搜索得到准确的基音周期。

[0109] 由于基音周期对声码器语音质量影响很大,为了更加准确地估算基音周期,利用上述得到的基音周期对去除共振峰信号频谱进行基音峰值及其谐振峰值搜索,以各峰值的差的平均值作为最后的基音周期。

[0110] 基音周期的范围为  $18 \sim 145$ ,采用 7 比特进行标量量化, $18 \sim 145$  分别映射到 7 比特的  $0 \sim 127$ 。

[0111] 增益参数(Gain)表示语音信号的增益值,如公式(12)所示:

$$[0112] \quad G = 10 \log_{10} \left[ 0.01 + \sqrt{\frac{1}{L} \sum_{n=0}^{L-1} r(n)^2} \right] \quad (12)$$

[0113] 其中,L 为基音周期, $r(n)$  为小波系数,对帧头和帧尾分别提取增益参数  $G_1$  和  $G_2$ 。

[0114] WELP 声码器在对数域进行增益参数的量化。增益参数先变换到对数域,并且限定范围  $10 \sim 77$ dB, $G_1$  采用 3 比特进行量化, $G_2$  采用 5 比特进行量化,其量化索引如公式(13)所示:

[0115]

$$index1 = \left\lfloor (g1' - 10) \cdot \frac{7}{77 - 10} \right\rfloor \quad (13)$$

[0116]

$$index2 = \left\lfloor (g2' - 10) \cdot \frac{31}{77 - 10} \right\rfloor$$

[0117] 其中, $g'$  表示对数域的增益。

[0118] 另外还需要对各子带进行清浊音判决,得到各子带的清浊音特性并进行标量量化。WELP 声码器对  $0 \sim 4000$ Hz 的语音信号划分为多个子带进行清 / 浊音分析。下面以 5 个子带为例进行描述,把语音信号划分为  $0 \sim 500$ Hz,  $500 \sim 1000$ Hz,  $1000 \sim 2000$ Hz,  $2000 \sim 3000$ Hz 和  $3000 \sim 4000$ Hz 共 5 个子带。采用 6 阶 Butterworth IIR 带通滤波器将预处理后的语音信号进行滤波,分别得到各子带信号。然后,对各个带通滤波信号在基音周期值附近计算最大归一化自相关函数值,如公式(14)所示:

$$[0119] \quad r_i(p) = \frac{\sum_{n=0}^{L-1} s_i(n) s_i(n-p)}{\left[ \sum_{n=0}^{L-1} s_i(n)^2 \sum_{n=0}^{L-1} s_i(n-p)^2 \right]^{1/2}} \quad i=1, \dots, 5 \quad (14)$$

[0120] 其中, $s_i(n)$  为带通滤波信号, $p$  为基音周期附近样点值。在这些归一化自相关值中,搜索最大的自相关值  $r_1(i)$ ,  $i = 1, \dots, 5$ 。

[0121] 然后对各子带信号的包络信号计算归一化自相关函数。包络信号通过对子带信号进行二阶低通滤波得到。对包络信号在基音周期附近计算最大归一化自相关函数值  $r_2(i)$ ,  $i = 1, \dots, 5$ 。

[0122] 采用域值比较法,根据  $r_1(i)$ ,  $i = 1, \dots, 5$  和  $r_2(i)$ ,  $i = 1, \dots, 5$  联合判断各个子

带清浊音。

[0123] WELP 声码器第一个子带的清 / 浊音用基音周期表示, 0 表示清音, 非 0 为浊音, 其它 4 个子带用 4 个比特表示, 1 表示浊音, 0 表示清音。

[0124] 基于上述线性预测语音编码方法, 本发明同时提出了一种语音合成方法, 如图 3 所示, 所述语音合成方法包括以下步骤:

[0125] S201、对量化的线谱频率对、小波系数、基音周期、增益参数及各子带的清浊音特性进行解码, 得到线谱频率对、小波激励幅度谱、基音周期、增益参数及各子带的清浊音特性;

[0126] S202、利用所述小波激励幅度谱、所述基音周期及所述各子带的清浊音特性, 合成小波激励信号;

[0127] S203、利用所述线谱频率对, 对所述小波激励信号进行反线性预测, 得到合成语音;

[0128] S204、对所述合成语音进行谱增强和相位调整。

[0129] 优选的, 步骤 S202 具体包括:

[0130] 根据所述各子带的清浊音特性, 对各子带的清音成分和浊音成分进行滤波混合, 得到小波频谱;

[0131] 对所述小波频谱进行反傅里叶变换, 得到小波系数, 并利用 dB10 小波基复合得到所述小波激励信号。

[0132] 下面以 5 个子带为例描述语音合成过程。根据 5 个子带的清浊音特性、基音周期和小波谱进行小波激励信号合成, 各子带信号根据 0~500Hz、500~1000Hz、1000~2000Hz、2000~3000Hz 和 3000~4000Hz 共 5 个子带的清浊音判决结果对浊音成分和清音成分在各个频带按比例进行滤波混合, 其中浊音子带采用浊音成分采用小波系数频谱幅度进行反傅里叶变换, 清音子带采用白噪声进行拟合。

[0133] 根据编码器对 5 个子带的清浊音判决结果对浊音成分和清音成分在各个频带按比例进行滤波混合, 得到小波频谱  $c_f(k)$  :

$$[0134] \quad c_f(k) = V_k A_k + (1 - V_k) e_k \quad 0 \leq k \leq 9 \quad (15)$$

[0135] 其中,  $A_k$  为小波系数频谱幅度,  $V_k$  为第  $k$  个小波系数频谱幅度所处子带的清浊音标记, 清音为 0, 浊音为 1,  $e_k$  为随机噪声。

[0136] 通过对  $c_f(k)$  进行反傅里叶变换即可恢复小波系数, 并利用 dB10 小波基复合形成激励信号。

[0137] 将激励信号通过反线性预测合成滤波器  $H(z)$  可以得到合成语音:

$$[0138] \quad H(z) = \frac{1}{1 + \sum_{i=1}^{10} a_i z^{-i}} \quad (16)$$

[0139] 其中,  $a_i$ ,  $i = 1, 2, \dots, 10$  为线性预测系数。

[0140] 为了使合成语音更加自然, 需要对合成语音进行谱增强和相位调整, 谱增强采用线性滤波器  $A(z)$  来构造:

$$[0141] \quad H(z) = \frac{A(z/\alpha)}{A(z/\beta)} (1 + \mu z^{-1}) \quad (17)$$

$$[0142] \quad A(z) = \frac{1}{H(z)} \quad (18)$$

[0143] 其中,  $\alpha$ ,  $\beta$  为短时后滤波调整因子,  $\mu$  为倾斜补偿因子。  $\alpha = 0.5p$ ,  $\beta = 0.8p$ ,

$$p = \frac{G_{\text{int}} - G_n - 12}{18}, \quad G_{\text{int}} \text{ 为插值后的当前增益, } G_n \text{ 为背景噪声增益, } \mu \text{ 为 } \max(0.5k1, 0)p \text{ 插值}$$

后的倾斜因子。

[0144] 下面给出一个本发明的应用实例。

[0145] 对于 WELP 声码器, 输入语音为 8kHz 采样的线性 PCM 信号, 语音信号分析帧长为 20ms, 当前帧共 160 个样点, 并缓存之前的 240 个样点。对这 400 个样点进行参数提取, 提取的参数包括 10 个线性预测系数生成 10 个线谱对、10 个小波系数、1 个基音周期、2 个增益参数、4 个子带语音的清 / 浊音 (第 1 个子带的清浊音特性由基音周期来描述, 基音周期为 0 表示清音, 否则为浊音)。这些参数的比特分配如表 1 所示:

[0146] 表 1 参数的比特分配表

[0147]

参数	浊音	清音
线谱对	20	20
基音周期	8	8
小波系数	8	0
增益	8	8
子带清浊音标记	4	0
FEC	0	13

[0148] 为了衡量本发明的效果, 采用 ITU-TP. 862 建议书提供的客观 MOS 值评价方法进行语音编码评价, 评价软件 PESQ 的 MOS 取值为 0~4.5, MOS 值越高, 语音质量越好, 测试结果如表 2 所示:

[0149] 表 2 测试结果

[0150]

	英语	汉语
男声	3.3	3.18
女声	3.4	3.21

[0151] 以上所述仅是本发明的优选实施方式, 应当指出, 对于本领域的技术人员来说, 在不脱离本发明技术原理的前提下, 还可以做出若干改进和替换, 这些改进和替换也应视为本发明的保护范围。

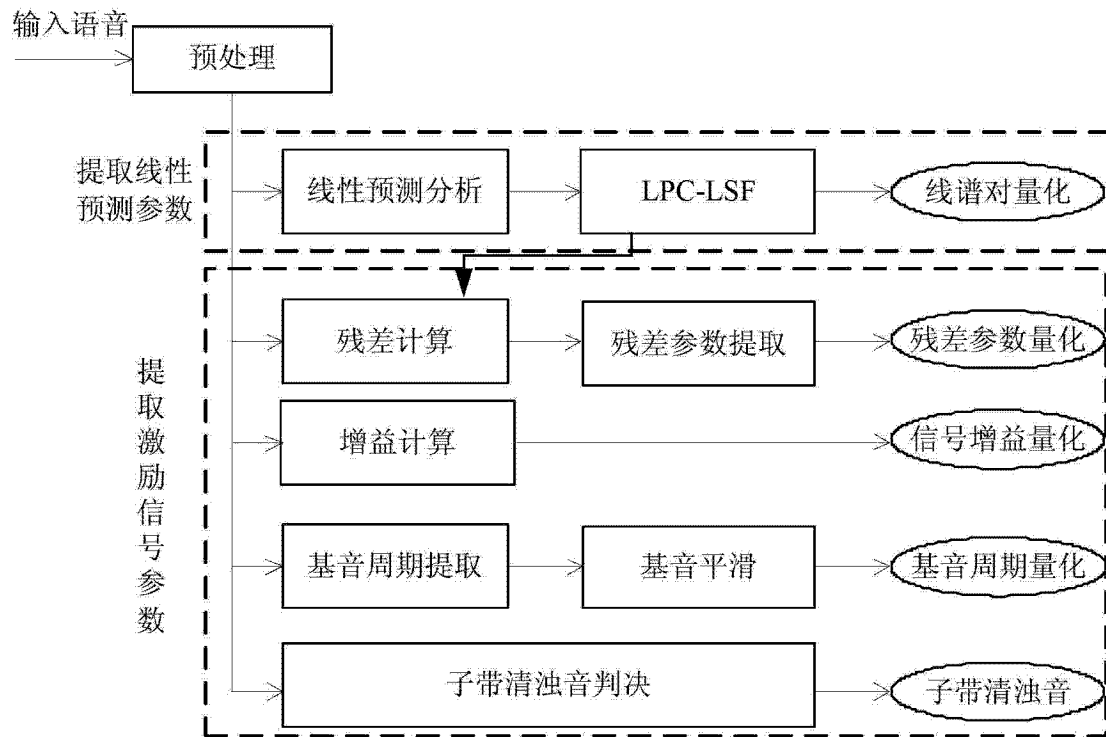


图 1

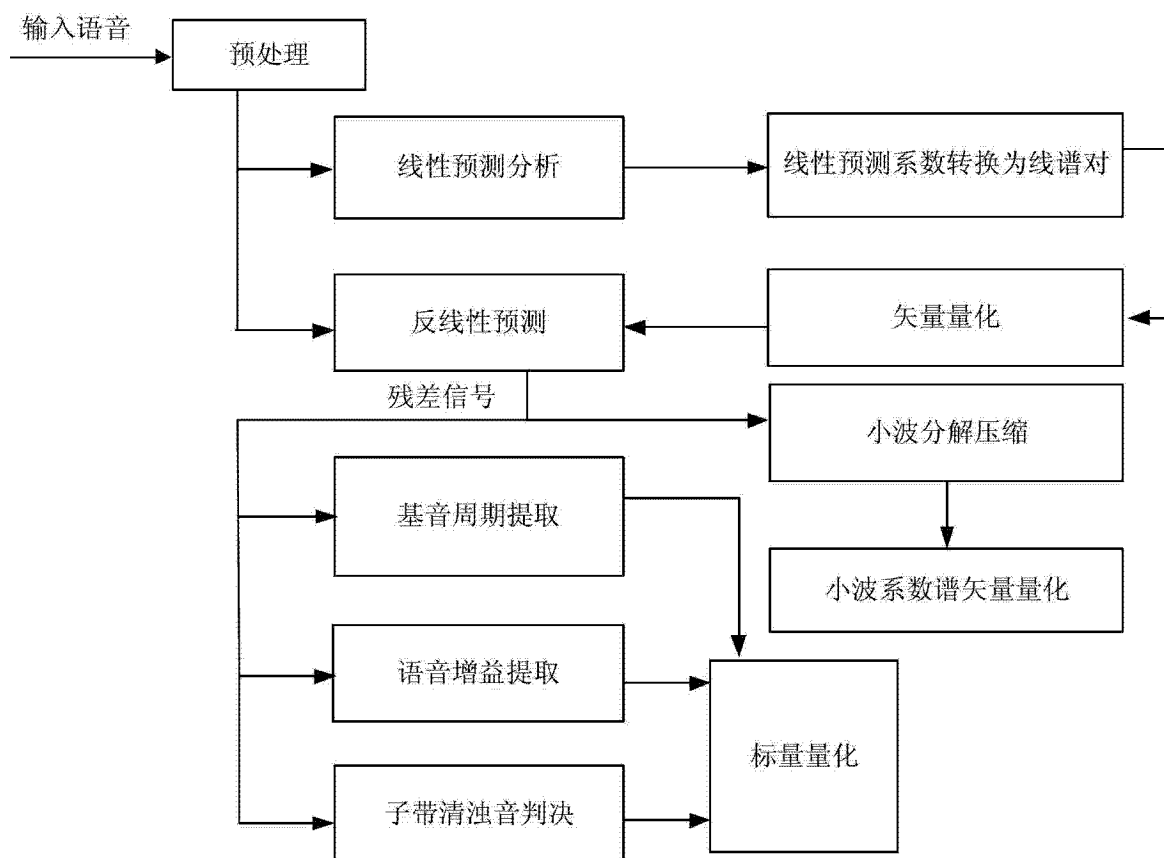


图 2

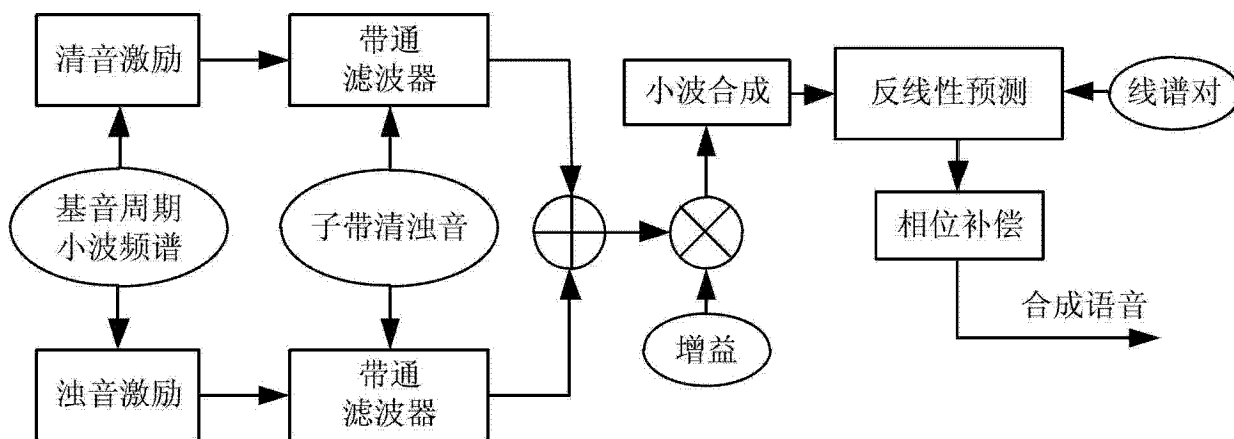


图 3