



(12) 发明专利

(10) 授权公告号 CN 102623015 B

(45) 授权公告日 2015.05.06

(21) 申请号 201210082801.8

(22) 申请日 1999.12.21

(30) 优先权数据
09/217,341 1998.12.21 US

(62) 分案原申请数据
99814819.9 1999.12.21

(73) 专利权人 高通股份有限公司
地址 美国加利福尼亚州

(72) 发明人 S·曼朱那什 W·加德纳

(74) 专利代理机构 上海专利商标事务所有限公司 31100
代理人 陈斌

(51) Int. Cl.
G10L 19/24(2013.01)
G10L 19/20(2013.01)

(56) 对比文件
CN 1131994 A, 1996.09.25,

Erdal Psksoy, Allen Gersho. Variable rate speech coding for multiple access wireless networks. 《Electrotechnical Conference, 1994. Proceedings., 7th Mediterranean》. 1995, 47-50.

W. Bastiaan Kleijn. Encoding speech using prototype waveforms. 《Speech and Audio Processing, IEEE Transactions on》. 1993, 第1卷(第4期), 386-399.

Erdal Paksoy, K. Srinivasan, Allen Gersho. variable rate speech coding with phonetic segmentation. 《Acoustics, Speech, and Signal Processing, 1993. ICASSP-93., 1993 IEEE International Conference on》. 1994, 第2卷 155-158.

审查员 贾杨

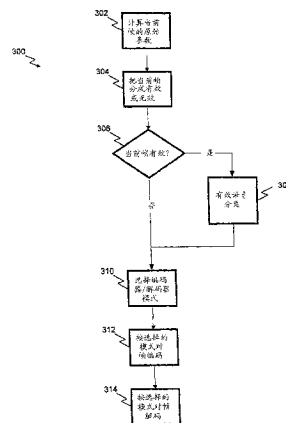
权利要求书1页 说明书28页 附图22页

(54) 发明名称

可变速率语音编码

(57) 摘要

一种用于对语音信号进行可变速率编码的方法和装置。对输入语音信号进行归类,根据此归类选择适当的编码模式。对于每个归类,选择以可接受的语音再现质量实现最低位速率的编码模式。通过仅利用高保真模式(即,广泛适用于不同类型的语音的高的位速率),在可接受的输出需要此保真度的语音部分中实现低的平均位速率。在这些模式产生可接受的输出的语音部分中使用较低的位速率。把输入语音信号归类为有效和无效区。有效区被进一步归类为语音、非语音和过渡区。可依据所需的保真水平把各种编码模式应用于不同类型的有效语音。可依据每个特定模式的强弱来利用编码模式。在语音信号的性质随时间而变化时,本设备在这些模式之间的动态切换。在适当时,把语音的区域模拟成伪随机噪声,从而获得明显较低的位速率。本发明以动态的方式使用此编码,而无论检测到非语音语音还是背景噪声。



CN 102623015 B

1. 一种用于对语音信号进行编码的方法,包括:
 - 原型音调周期 PPP 编码模式接收语音帧的共振峰剩余信号;
 - 从所述共振峰剩余信号中提取原型剩余;
 - 从当前原型剩余和前一原型剩余中计算旋转参数集,其中所述计算包括循环滤波所述原型剩余以确定目标信号;以及
 - 基于所述目标信号生成代码簿参数集。
2. 一种用于对语音信号进行编码的系统,包括:
 - 用于原型音调周期 PPP 编码模式接收语音帧的共振峰剩余信号的滤波装置;
 - 用于从所述共振峰剩余信号提取原型剩余的提取装置;
 - 用于从当前原型剩余和前一原型剩余中计算旋转参数集的相关装置,其中所述计算包括循环滤波所述原型剩余以确定目标信号;以及
 - 用于基于所述目标信号生成代码簿参数集的搜索装置。
3. 一种用于对语音信号进行解码的方法,包括:
 - 接收多个代码簿参数和多个旋转参数;
 - 从所述代码簿参数和所述旋转参数重构当前原型剩余信号;
 - 接收所重构的原型剩余信号和前一重构的原型剩余信号;
 - 内插所述重构的原型剩余信号和所述前一重构的原型剩余信号以形成合成的语音信号;以及
 - 输出所述合成的语音信号。
4. 一种用于对语音信号进行解码的系统,包括:
 - 用于接收多个代码簿参数和多个旋转参数的接收装置;
 - 用于从所述代码簿参数和所述旋转参数重构当前原型剩余信号的重构装置;
 - 用于接收所重构的原型剩余信号和前一重构的原型剩余信号的装置;
 - 用于内插所述重构的原型剩余信号和所述前一重构的原型剩余信号以形成合成的语音信号;以及
 - 用于输出所述合成的语音信号的输出装置。

可变速率语音编码

[0001] 本申请是申请日为 1999 年 12 月 21 日、申请号为 200710162109.5、发明名称为“可变速率语音编码”的发明专利申请的分案申请,该分案申请是申请日为 1999 年 12 月 21 日、申请号为 99814819.9、发明名称为“可变速率语音编码”的发明专利申请的分案申请。

技术领域

[0002] 本发明涉及对语音信号的编码。具体来说,本发明涉及对语音信号进行归类并根据此归类而利用多种编码模式中的一种。

背景技术

[0003] 当今的许多通信系统,特别是远距与数字无线电话应用中,都把话音当作数字信号发射。这类系统的性能部分取决于以最少的位数精确地代表话音信号。简单地通过采样与数字化来发送语音,为了达到普通模拟电话的语音质量,要求数据速率为每秒 64kb(kbps)。然而,现有的编码技术可明显减少正常语音再现所需的数据速率。

[0004] 术语“声码器”一般指根据人类语音发生模型通过提取诸参数来压缩发出的语音的装置。声码器包括编码器与解码器,编码器分析送入的语音并提取相关的参数,解码器用经传输信道接收自编码器的诸参数合成语音。通常把语音信号分成几帧数据与字块供声码器处理。

[0005] 围绕基于线性预测的时域编码方案建立的编码器,在数量上远远超过了其它各类编码器。这类技术从语音信号里提取诸相关的单元,只编码不相关的单元。基本的线性预测滤波器预测的当前样本作为过去样本的一种线性组合。Thomas E. Tremain 等人撰写的论文“一种 4.8kbps 码受激线性预测编码器”(移动卫星会议录,1998),描述了一例这类特定的编码算法。

[0006] 这类编码方案除去语音中固有的所有自然冗余度(即相关单元),把数字化语音信号压缩成低位速率信号。语音一般呈现出唇与舌的机械动作造成的短期冗余度和声带振动造成的长期冗余度。线性预测方案把这些动作模拟成滤波器,除去冗余度,再将得到的剩余(residual)信号模拟成白高斯噪声。因此,线性预测编码器通过发送滤波器系数和量化噪声而不是发送全带宽的语音信号,可以减小位速率。

[0007] 然而,即使是这些减小的位速率,也往往超过了有效带宽,其中语音信号必须远距传播(如地面到卫星),或在拥挤的信道中与许多其它信号共存。因此,要求有一种改进的编码方案,以实现比线性预测方案更低的位速率。

发明内容

[0008] 本发明是一种对语音信号进行可变速率编码的改进的新方法和设备。

[0009] 本发明的一方面提供了一种用于语音信号的可变速率编码的方法,其特征在于,包括以下步骤:(a)把语音信号归类为有效或无效;(b)把所述有效语音归类为多个有效语音类型中的一个;(c)从多个并行编码器模式中选择一编码器模式,其中根据语音信号

是有效还是无效来选择编码器模式,如果有效,则进一步根据所述有效语音类型来选择编码器模式;(d)依据所述所选的编码器模式对语音信号进行编码,从而形成经编码的语音信号。

[0010] 本发明的另一方面提供了一种用于对语音信号进行编码的可变速率编码系统,包括:归类装置,用于把语音信号归类为有效或无效,如果有效,则把所述有效语音归类为多个有效语音类型中的一个;以及多个并行编码装置,用于把语音信号编码成为经编码的语音信号,其中根据语音信号有效还是无效,如果有效,则进一步根据所述有效语音类型而动态选择所述并行编码装置对语音信号进行编码。

[0011] 本发明的又一方面提供了一种语音信号的可变速率编码方法,包括:把语音信号归类为有效或无效,其中把语音归类为有效或无效包括基于两个能带的阈值处理方案;把所述有效语音归类为多个有效语音类型中的一个,其中所述多个有效语音类型包括话音、非话音和过渡有效语音;根据语音信号是有效还是无效来选择一个编码器模式,如果有效,则进一步根据所述有效语音类型来选择编码器模式,其中所述所选的编码器模式由一编码位速率或由一编码算法来表征,或者由一编码位速率和一编码算法来表征;以及依据所述编码器模式对语音信号进行编码,从而形成经编码的语音信号。

[0012] 本发明的再一方面提供了一种语音信号的可变速率编码方法,包括:把语音信号归类为有效或无效,其中把语音归类为有效或无效包括:如果前 N_{h0} 个帧被归类为有效,则把后 M 个帧归类为有效;把所述有效语音归类为多个有效语音类型中的一个,其中所述多个有效语音类型包括话音、非话音和过渡有效语音;根据语音信号是有效还是无效来选择一个编码器模式,如果有效,则进一步根据所述有效语音类型来选择编码器模式,其中所述所选的编码器模式由一编码位速率或由一编码算法来表征,或者由一编码位速率和一编码算法来表征;以及依据所述编码器模式对语音信号进行编码,从而形成经编码的语音信号。

[0013] 本发明的还有一方面提供了一种对语音信号进行编码的可变速率编码系统,包括:归类装置,用于根据两个能带的阈值处理方案把语音信号归类为有效或无效,如果有效,把所述有效语音归类为多个有效语音类型中的一个;以及多个编码装置,用于把所述语音信号编码为经编码的语音信号,其中所述编码装置被动态选择以根据语音信号是有效还是无效来对语音信号进行编码,如果有效,则进一步根据所述有效语音类型来编码。

[0014] 本发明的再有一方面提供了一种对语音信号进行编码的可变速率编码系统,包括:归类装置,用于把语音信号归类为有效或无效,其中如果前 N_{h0} 个帧被归类为有效,所述归类装置就把后 M 个帧归类为有效,如果有效,还用于把所述有效语音归类为多个有效语音类型中的一个;以及多个编码装置,用于把所述语音信号编码为经编码的语音信号,其中所述编码装置被动态选择以根据语音信号是有效还是无效来对语音信号进行编码,如果有效,则进一步根据所述有效语音类型来编码。

[0015] 本发明对输入语音信号进行归类并根据此归类选择适当的编码模式。对于每个归类,本发明选择以可接受的语音再现质量实现最低位速率的编码模式。本发明通过仅利用高保真模式(即,广泛适用于不同类型的语音的高的位速率),在可接受的输出需要此保真度的语音部分中实现低的平均位速率。本发明在这些模式产生可接受的输出的语音部分中切换到较低的位速率。

[0016] 本发明的一个优点在于,以低的位速率对语音进行编码。低的位速率转化成较高

的容量、较大的范围以及较低的功率需求。

[0017] 本发明的一个特征在于,把输入语音信号归类为有效(active)和无效(inactive)区。有效区被进一步归类为话音(voiced)、非话音(unvoiced)和过渡(transient)区。因此,本发明可依据所需的保真水平把各种编码模式应用于不同类型的有效语音。

[0018] 本发明的另一个特征在于,可依据每个特定模式的强弱来利用编码模式。在语音信号的性质随时间而变化时,本发明在这些模式之间的动态切换。

[0019] 本发明的又一个特征在于,在适当时,把语音的区域模拟成伪随机噪声,从而获得明显较低的位速率。本发明以动态的方式使用此编码,而无论检测到非话音语音还是背景噪声。

[0020] 从以下结合附图的详细描述,将使本发明的特征、目的和优点变得更加明显起来,图中类似的标号指示相同或在功能上相似的元件。此外,标号最左边的数字识别最早出现该标号的图。

[0021] 附图概述

[0022] 图 1 是表示信号传输环境的图;

[0023] 图 2 是详细示出编码器 102 和解码器 104 的图;

[0024] 图 3 是表示本发明可变速率语音编码的流程图;

[0025] 图 4A 是表示一帧话音语音分割为若干子帧的图;

[0026] 图 4B 是表示一帧非话音语音分割为若干子帧的图;

[0027] 图 4C 是表示一帧过渡语音分为若干子帧的图;

[0028] 图 5 是描绘原始参数计算的流程图;

[0029] 图 6 是描绘语音分类为有效或无效的流程图;

[0030] 图 7A 是表示 CELP 编码器的图;

[0031] 图 7B 是表示 CELP 解码器的图;

[0032] 图 8 是表示音调滤波器模块的图;

[0033] 图 9A 是表示 PPP 编码器的图;

[0034] 图 9B 是表示 PPP 解码器的图;

[0035] 图 10 是表示 PPP 编码法(包括编解码)步骤的流程图;

[0036] 图 11 是描述原型剩余周期提取流程图;

[0037] 图 12 是示出从当前帧剩余信号提取的原型剩余周期和从前一帧提取的原型剩余周期的图;

[0038] 图 13 是计算旋转参数的流程图;

[0039] 图 14 是表明编码代码簿工作的流程图;

[0040] 图 15A 是表示第一滤波器更新模块实施例的图;

[0041] 图 15B 是表示第一周期内插器模块实施例的图;

[0042] 图 16A 是表示第二滤波器更新模块实施例的图;

[0043] 图 16B 是表示第二周期内插器模块实施例的图;

[0044] 图 17 是描述第一滤波器更新模块实施例的工作的流程图;

[0045] 图 18 是描述第二滤波器更模块实施例的工作的流程图;

- [0046] 图 19 是描述原型剩余周期对准与内插的流程图；
- [0047] 图 20 是描述第一实施例根据原型剩余周期重建语音信号的流程图；
- [0048] 图 21 是描述第二实施例根据原型剩余周期重建语音信号的流程图；
- [0049] 图 22A 是表示 NELP 编码器的图；
- [0050] 图 22B 是表示 NELP 解码器的图；和
- [0051] 图 23 是描述 NELP 编码法的流程图。
- [0052] 本发明的较佳实施方式
- [0053] I. 环境概述
- [0054] II. 发明概述
- [0055] III. 原始参数确定
- [0056] A. 计算 LPC 系数
- [0057] B. LSI 计算
- [0058] C. NACF 计算
- [0059] D. 音调轨迹与滞后计算
- [0060] E. 计算带能与零交叉率
- [0061] F. 计算元音共振峰 (formant) 余量
- [0062] IV. 有效 / 无效语音分类
- [0063] A. 拖尾 (hangover) 帧
- [0064] V. 有效语音帧分类
- [0065] VI. 编码器 / 解码器模式选择
- [0066] VII. 代码受激的线性预测 (CELP) 编码模式
- [0067] A. 音调编码模块
- [0068] B. 编码代码簿
- [0069] C. CELP 解码器
- [0070] D. 滤波器更新模块
- [0071] VIII. 原型音调周期 (PPP) 编码模式
- [0072] A. 提取模式
- [0073] B. 旋转相关器
- [0074] C. 编码代码簿
- [0075] D. 滤波器更新模块
- [0076] E. PPP 解码器
- [0077] F. 周期内插器
- [0078] IX. 噪声激励的线性预测 (NELP) 编码模式
- [0079] X. 结论。
- [0080] I. 环境概述
- [0081] 要发明针对可变速率语音编码的新颖改进的方法和设备。图 1 示出信号传输环境 100, 它包括编码器 102、解码器 104 和信号传输媒体 106。编码器 102 对语音信号 $s(n)$ 编码, 形成的编码语音信号 $s_{\text{enc}}(n)$ 通过传输媒体 106 传输给解码器 104, 后者对 $s_{\text{enc}}(n)$ 解码而生成合成的语音信号 $\hat{S}(n)$ 。

[0082] 这里的“编码”一般指包括编码二者的方法。一般而言,编码方法和设备试图将通过传输媒体 106 发送的位数减至最少(即将 $s_{\text{enc}}(n)$ 的带宽减至最少),同时保持可接受的语音再现(即 $\hat{S}(n) \approx s(n)$)。编码语音信号的成分随具体的语音编码方法而不同。下面描述根据它们工作的各种编码器 102、解码器 104 和编码方法。

[0083] 下述编码器 102 和解码器 104 的元件,可用电子硬件,计算机软件或二者的组合构成,下面按其功能描述这些元件。功能用硬件实施还是用软件实施,将取决于具体应用和对整个系统的设计限制性。熟练的技术人员应该知道硬软件在这些场合中的互换性以及如何最佳地实施对每个具体应用描述的功能。

[0084] 本领域的技术人员应明白,传输媒体 106 可以代表许多不同的传输媒体,包括(但不限于)陆基通信线路、基站与卫星间的链路、蜂窝电话与基站或蜂窝电话与卫星间的无线通信。

[0085] 本领域的技术人员还将明白,通信的每一方通常都作发射与接收,因此每一方都要求有编码器 102 和解码器 104。然而,下面将把信号传输环境 100 描述成在传输媒体 106 的一端包括编码器 102,另一端包括解码器 104。技术人员将容易明白如何将这些设想扩展到双向通信。

[0086] 为了进行描述,假定 $s(n)$ 是在一般交谈中得到的数字语音信号,交谈包括不同的语音发声与静寂周期。语音信号 $s(n)$ 最好分成若干帧,每个帧又分成若干子帧(最好为 4 个)。在作字快处理时,如在本文情况下,一般应用这些任意选择的帧/子帧边界,对帧叙述的操作也适用于子帧,在这方面帧与子帧在这里可互换使用。然而,若是连续处理而不是字块处理, $s(n)$ 就根本无须分为帧/子帧。技术人员很容易明白如何将下述的字块技术扩展到连续处理。

[0087] 在一较佳实施例中, $s(n)$ 以 8kHz 作数字采样。每帧最好含 20ms 数据,即在 8kHz 速率下为 160 个样本,所以各子帧含 40 个数据样本。要着重指出,下面的许多公式都假设了这些值。然而,技术人员将明白,虽然这些参数适合语音编码,但是仅仅为了示例,可以应用其它合适的替代参数。

[0088] II. 发明概述

[0089] 本发明的方法和设备涉及到编码与语音信号 $s(n)$ 。图 2 详细示出了编码器 102 和解码器 104。根据本发明,编码器 102 包括原始参数计算模块 202,分类模块 208 和一种或多种编码器模式 204。解码器 104 包括一种或多种解码器模式 206。解码器模式数 N_d 一般等于编码器模式数 N_e 。如技术人员所知,编码器模式,与解码器模式 1 相联系,其它依次类推。如图所示,编码的语音信号 $s_{\text{enc}}(n)$ 通过传输媒体 106 发送。

[0090] 在一较佳实施例中,根据哪一模式最适合当前帧规定的 $s(n)$ 特性,编码器 102 在各帧的多个编码器模式之间作动态切换,解码器 104 也在各帧的相应解码器模式之间作动态切换。对每一帧选择一具体模式,以获得最低位速率并保持解码器可接受的信号再现。这一过程称为可变速率语音编码,因为编码器的位速率随时间而变化(作为信号变化的特点)。

[0091] 图 3 是流程图 300,描述了本发明的可变速率语音编码法。在步骤 302,原始参数计算模块 202 根据当前帧的数据计算各种参数。在一较佳实施例中,这些参数包括下列参数之一或几个:线性预测编码(LPC)滤波器系数、线路谱信息(LSI)系数、归一化自相关函

数 (MACF)、开环滞后、带能、零交叉速率和元音共振峰分剩余信号。

[0092] 在步骤 304、分类模块 208 把当前帧分为含“有效”或“无效”的语音。如上所述, $s(n)$ 假定对普通谈话包括语音周期与静寂周期。有效语音包括说出的单词, 而无效语音包括其它任何内容, 如背景噪声、静寂、间歇。下面详细描述本发明把语音分为有效 / 无效的方法。

[0093] 如图 3 所示, 步骤 306 研究当前帧在步骤 304 是否被分为有效或无效, 若有效, 控制流程进到步骤 308; 若无效, 控制流程进到步骤 310。

[0094] 被分为有效的帧在步骤 308 再分为话音帧、非话音帧或过渡帧。技术人员应明白, 人类语音可用多种不同的方法分类。两种常用的语音分类是话音声与非话音声。根据本发明, 把非话音语音都归为过渡语音。

[0095] 图 4A 示出一例含话音语音 402 的 $s(n)$ 部分。产生话音声时, 迫使空气通过喉门并调节声带的紧度, 以松弛振荡方式振动, 由此产生激发发音系统的准周期空气脉冲。话音语音测出的一个共同特性是图 4A 所示的音调周期。

[0096] 图 4B 示出一例含非话音语音 404 的 $s(n)$ 部分。产生非话音时, 在发音系统的某一点形成收缩部 (通常朝向嘴端), 迫使空气以足够高的速度通过该收缩部而产生扰动, 得到的非话音语音信号类似于有色噪声。

[0097] 图 4C 示出一例含过渡语音 406 (即既不是话音也不是非话音的语音) 的 $s(n)$ 部分。图 4C 列举的过渡语音 406 可以代表 $s(n)$ 在非话音语音与话音语音的转变。技术人员将明白, 可根据这里描述的技术应用多种不同的语音分类获得可比的结果。

[0098] 在步骤 310, 根据步骤 306 和 308 作出的帧分类, 选择编码器 / 解码器模式。各种编 / 解码器模式平行连接, 如图 2 所示, 一种或多种此类模式可在规定时间工作。但如下所述, 最好在规定时间只有一种模式工作, 并按当前帧分类选择。

[0099] 以下几段描述几种编 / 解码器模式。不同的编 / 解码器模式按不同的编码方案工作。有些模式在语音信号 $s(n)$ 呈现某些特点的编码部分更为有效。

[0100] 在一较佳实施例中, 对分类为过渡语音的代码帧选用“代码受激线性预测” (CELP) 模式, 该模式用量化型线性预测剩余信号激发线性预测发音系统模型。这里描述的所有编 / 解码器模式中, CELP 通常产生最准确的语音再现, 但要求最高的位速率。在一实施例中, CELP 模式实行每秒 8500 位的编码。

[0101] 对分类为话音语音的代码帧, 最好选用“原型音调周期” (PPP) 模式。话音语音包含可被 PPP 模式利用的慢时变周期分量。PPP 模式只对每帧内音调周期的子组编码。语音信号的其余周期通过这些原型周期期间的内插而重建。利用话音语音的周期性, PPP 能实现比 CELP 更低的位速率。且仍能以感性的精确方式再现该语音信号。在一实施例中, PPP 模式实行每秒 3900 位的编码。

[0102] 对分类为非话音语音的代码帧, 可选用“噪声受激线性预测” (NELP) 模式, 它用经滤波的伪随机噪声信号模拟非话音语音。NELP 对编码语音应用最简单的模型, 所以位速率最低。在一实施例中, NELP 模式实行每秒 1500 位的编码。

[0103] 同一种编码技术能以不同的位速率频繁地工作, 性能级别不同。因此, 图 2 中不同的编码器 / 解码器模式可代表不同的编码技术的相同编码技术, 或上述情况相组合。技术人员应明白, 增加编 / 解码器模式数量, 选择模式更灵活, 且能导致更低的平均位速率, 不

过整个系统会更复杂。在指定系统中应用的具体组合,将决定于现有的系统资源与特定的信号环境。

[0104] 在步骤 312,选用的编码器模式 204 对当前帧编码,最好将编码的数据装入数据包传输。在步骤 314,对应的解码器模式 206 打开数据包,对收到的数据解码并重建该语音信号。下面针对合适的编/解码器模式详细描述这些操作。

[0105] III. 原始参数确定

[0106] 图 5 是更详细说明步骤 302 的流程图。各种原始参数按本发明计算。这些参数最好包括如 LPC 系数、线路谱信息 (LSI) 系数、归一化自相关函数 (NACF)、开环滞后、带能、零交叉速率和元音共振峰剩余信号,这些参数在整个系统内按各种方式使用,如下所述。

[0107] 在一较佳实施例中,原始参数计算模块 202 应用“超前 (look ahead)”的 160+40 个样本,这有几个原因。首先,160 样本超前可用下一帧的信息计算音调频率轨迹,明显增强了下述话音编码与音调周期估算技术的耐用性。其次,160 样本超前可对将来一帧计算 LPC 系数、帧能和话音活性,这能有效地多帧量化帧能与 LPC 系数。再次,附加的 40 样本超前可对下述的汉明窗语音计算 LPC 系数。因此,处理当前帧之前缓冲的样本数是 160+160+40,包括当前帧和 160+40 样本超前。

[0108] A. 计算 LPC 系数

[0109] 本发明用 LPC 预测误差滤波器消除语音信号中的短期冗余度。LPC 滤波器的传递函数为:

$$[0110] \quad A(z) = 1 - \sum_{i=1}^{10} a_i z^{-i}$$

[0111] 本发明最好构制一种十阶滤波器,如前述公式所示。解码器中的 LPC 合成滤波器重新插入冗余度,并由 $A(z)$ 的倒数规定:

$$[0112] \quad \frac{1}{A(z)} = \frac{1}{1 - \sum_{i=1}^{10} a_i z^{-i}}$$

[0113] 在步骤 502, LPC 系数 a_i 由 $s(n)$ 计算如下。在对当前帧编码期间,最好对下一帧计算 LPC 参数。

[0114] 对中心位于第 119 与第 120 样本之间的当前帧应用汉明窗 (假定较佳的 160 样本帧有一“超前”)。窗示语音信号 $s_w(n)$ 为:

$$[0115] \quad s_w(n) = s(n+40) \left(0.5 + 0.46 * \cos \left(\pi \frac{n-79.5}{80} \right) \right), \quad 0 \leq n < 160$$

[0116] 40 样本的偏移导致该语音窗的中心位于较佳语音 160 样本帧的第 119 与 120 样本之间。

[0117] 最好将 11 个自相关值计算成:

$$[0118] \quad R(k) = \sum_{m=0}^{159-k} s_w(m) s_w(m+k), \quad 0 \leq k \leq 10$$

[0119] 对自相关值开窗可减少丢失线路谱对 (LSP) 的根的可能性, LSP 对由 LPC 系数得出:

[0120] $R(k) = h(k)R(k), 0 \leq k \leq 10$

[0121] 导致带宽略有扩展,如 25Hz。值 $h(k)$ 最好取自 255 点汉明窗的中心。

[0122] 接着用 Durbin 递归从开窗的自相关值获取 LPC 系数, Durbin 递归是众所周知的高效运算方法,在 Rabiner & Schafer 提出的文本“语音信号数字处理法”中作了讨论。

[0123] B. LSI 计算

[0124] 在步骤 504,把 LPC 系数变换成线路谱信息 (LSI) 系数作量化和内插。LSI 系数按本发明以下述方式计算:

[0125] 如前一样, $A(z)$ 为

$$[0126] \quad A(z) = 1 - a_1 z^{-1} - \dots - a_{10} z^{-10},$$

[0127] 式中 a_i 是 LPC 系数,且 $1 < i < 10$

[0128] $P_A(z)$ 与 $Q_A(z)$ 定义如下:

$$[0129] \quad P_A(z) = A(z) + z^{-11} A(z^{-1}) = p_0 + p_1 z^{-1} + \dots + p_{11} z^{-11},$$

$$[0130] \quad Q_A(z) = A(z) - z^{-11} A(z^{-1}) = q_0 + q_1 z^{-1} + \dots + q_{11} z^{-11},$$

[0131] 其中

$$[0132] \quad p_i = -a_i - a_{11-i}, 1 \leq i \leq 10$$

$$[0133] \quad q_i = -a_i + a_{11-i}, 1 \leq i \leq 10$$

[0134] 和

$$[0135] \quad p_0 = 1 \quad p_{11} = 1$$

$$[0136] \quad q_0 = 1 \quad q_{11} = -1$$

[0137] 线路谱余弦 (LSC) 是下述两函数中 $-0.1 < X < 1.0$ 的 10 个根

$$[0138] \quad P'(x) = p'_0 \cos(5 \cos^{-1}(x)) + p'_1 (4 \cos^{-1}(x)) + \dots + p'_4 + p'_5 / 2$$

$$[0139] \quad Q'(x) = q'_0 \cos(5 \cos^{-1}(x)) + q'_1 (4 \cos^{-1}(x)) + \dots + q'_4 x + q'_5 / 2$$

[0140] 式中

$$[0141] \quad p'_0 = 1$$

$$[0142] \quad q'_0 = 1$$

$$[0143] \quad p'_i = p_i - p'_{i-1} \quad 1 \leq i \leq 5$$

$$[0144] \quad q'_i = q_i + q'_{i-1} \quad 1 \leq i \leq 5$$

[0145] 然而以下式计算 LSI 系数

$$[0146] \quad lsi_i = \begin{cases} 0.5 \sqrt{1 - lsc_i} & lsc_i \geq 0 \\ 1.0 - 0.5 \sqrt{1 + lsc_i} & lsc_i < 0 \end{cases}$$

[0147] LSC 可按下式从 LSI 系数里取回:

$$[0148] \quad lsc_i = \begin{cases} 1.0 - 4 lsi_i^2 & lsi_i \leq 0.5 \\ (4 - 4 lsi_i^2) - 1.0 & lsi_i > 0.5 \end{cases}$$

[0149] LPC 滤波器的稳定性确保这两个函数的根交替,即最小根 lsc_1 就是 $P'(x)$ 的最小根,下一最小根 lsc_2 就是 $Q(X)$ 的最小根,等等。因此, lsc_1 、 lsc_3 、 lsc_5 、 lsc_7 、 lsc_9 都是 $p'(x)$ 的根,而 lsc_2 、 lsc_4 、 lsc_6 、 lsc_8 与 lsc_0 都是 $Q'(x)$ 的根。

[0150] 技术人员将明白,最好应用某种计算 LSI 系数灵敏度的方法来量化。量化处理中可用“灵敏度加权”对每个 LSI 中的量化误差合理地加权。

[0151] LSI 系数用多级矢量量化器 (VQ) 量化, 级数最好取决于所用的具体位速率与代码簿, 而代码簿的选用以当前帧是否为语音为依据。

[0152] 矢量量化将如下定义的加权均方误差 (WMSE) 减至最小 :

$$[0153] \quad E(\bar{x}, \bar{y}) = \sum_{i=0}^{P-1} w_i (x_i - y_i)^2$$

[0154] 式中 \bar{x} 是量化的矢量, \bar{w} 是与其有关的加权, \bar{y} 是代码矢量。在一较佳实施例中, \bar{w} 是灵敏度权和, $p = 10$ 。

[0155] LSI 矢量由 LSI 码重建, 而 LSI 码是量化成 $\bar{q} \bar{1} \bar{s} i = \sum_{i=1}^N C B \bar{i}_{code_i}$ 得到的, 其中 $C B i$ 是语音或非语音帧的第 i 级 VQ 代码簿 (基于指明选择代码簿的代码), $code_i$ 是第 i 级的 LSI 代码。

[0156] 在 LSI 系灵敏变换成 LPC 系数之前, 要作稳定性检查, 确保得到的 LPC 滤波器不因量化噪声或将噪声注入 LSI 系数的语道误差而不稳定。若 LSI 系数保持有序的, 则要确保稳定性。

[0157] 计算原始 LPC 系数时, 使用中心位于帧的第 119 与 120 样本之间的语音窗。该帧其它各点的 LPC 系数可在前一帧的 LSC 与当前帧的 LSC 之间内插近似, 得到的内插 LSC 再到变换回 LPC 系数。各子帧使用的正确内插为 :

$$[0158] \quad ilsc_j = (1-a_i)lscprev_j + ailsccurr_j, \quad 1 \leq j \leq 10$$

[0159] 式中 a_i 是 40 个样本中各四个子帧的内插系数 0.375、0.625、0.875、1.000, $ilsc$ 是内插的 LSC。用内插的 LSC 计算 $\hat{P}_A(z)$ 和 $\hat{Q}_A(z)$ 为 :

$$[0160] \quad \hat{P}_A(z) = (1+z^{-1}) \prod_{j=1}^5 (1 - 2ilsc_{2j-1}z^{-1} + z^{-2})$$

$$[0161] \quad \hat{Q}_A(z) = (1-z^{-1}) \prod_{j=1}^5 (1 - 2ilsc_{2j}z^{-1} + z^{-2})$$

[0162] 所有四个子帧内插的 LPC 系数作为下式的系数来计算 :

$$[0163] \quad \hat{A}(z) = \frac{\hat{P}_A(z) + \hat{Q}_A(z)}{2}$$

[0164] 因此

$$[0165] \quad \hat{a}_i = \begin{cases} -\frac{\hat{p}_i + \hat{q}_i}{2} & 1 \leq i \leq 5 \\ -\frac{\hat{p}_{11-i} - \hat{q}_{11-i}}{2} & 6 \leq i \leq 10 \end{cases}$$

[0166] C. NACF 计算

[0167] 在步骤 506, 归一化自相关函数 (WACF) 按本发明计算。

[0168] 下一帧的元音共振峰余量对 40 个样本子帧计算成

$$[0169] \quad r(n) = s(n) - \sum_{i=1}^{10} \tilde{\alpha}_i s(n-i)$$

[0170] 式中 $\tilde{\alpha}_i$ 是相应子帧第*i*次内插的LPC系数,内插在当前帧的非量化LSC与下一帧的LSC之间进行。下一帧的能量也计算成:

$$[0171] \quad E_N = 0.5 \log_2 \left(\frac{\sum_{n=0}^{159} r^2(n)}{160} \right)$$

[0172] 上述计算的余量经低通滤波和抽取,最好使用一种零相FIR滤波器实施,其长度为15,其系数 df_i ($-7 < i < 7$)为{0.0800,0.1256,0.2532,0.4376,0.6424,0.8268,0.9544,1.000,0.9544,0.8268,0.6424,0.4376,0.2532,0.1256,0.0800}。低通滤波、抽取的余量计算为:

$$[0173] \quad r_d(n) = \sum_{i=-7}^7 df_i r(Fn+i), \quad 0 \leq n < 160/F$$

[0174] 式中 $f=2$ 为抽取系数, $r(Fn+i)$, $-7 \leq Fn+i \leq 6$ 根据非量化LPC系数从当前帧的余量的最后14个值得到。如上所述,这些LPC系数在前一帧计算和存贮。

[0175] 下一帧两子帧(40样本抽取)的WACF的计算如下:

$$[0176] \quad E_{xx_k} = \sum_{i=0}^{39} r_d(40k+i)r_d(40k+i), \quad k=0,1$$

$$[0177] \quad E_{xy_{k,j}} = \sum_{i=0}^{39} r_d(40k+i)r_d(40k+i-j),$$

$$[0178] \quad 12/2 \leq j < 128/2, k=0,1$$

$$[0179] \quad E_{yy_{k,j}} = \sum_{i=0}^{39} r_d(40k+i-j)r_d(40k+i-j),$$

$$[0180] \quad 12/2 \leq j < 128/2, k=0,1$$

$$[0181] \quad n_corr_{k,j-12/2} = \frac{(E_{xy_{k,j}})^2}{E_{xx}E_{yy_{k,j}}},$$

$$[0182] \quad 12/2 \leq j < 128/2, k=0,1$$

[0183] 对*n*为负的 $r_d(n)$,一般使用当前帧的低通滤波和抽取的余量(前一帧存贮的)。当前子帧 c_corr 的NACF也在前一帧计算和存贮。

[0184] D. 音调轨迹与滞后计算

[0185] 在步骤508,按本发明计算音调轨迹音调滞后。最好按下列公式用有反向轨迹的Viterbi类搜索法计算音调滞后:

$$[0186] \quad R1_i = n_corr_{0,j} + \max\{n_corr_{1,j+FAN_{i,0}}\},$$

$$[0187] \quad 0 \leq i < 116/2, 0 \leq j < FAN_{i,j}$$

$$[0188] \quad R2_i = c_corr_{i,j} + \max \{R1_{j+FAN_{i,0}}\},$$

$$[0189] \quad 0 \leq i < 116/2, 0 \leq j < FAN_{i,1}$$

$$[0190] \quad RM_{2i} = R2_i + \max \{c_corr_{0,j+FAN_{i,0}}\},$$

$$[0191] \quad 0 \leq i < 116/2, 0 \leq j < FAN_{i,1}.$$

[0192] 其中 $FAN_{i,j}$ 是 2×58 矩阵, $\{ \{0, 2\}, \{0, 3\}, \{2, 2\}, \{2, 3\}, \{2, 4\}, \{3, 4\}, \{4, 4\}, \{5, 4\}, \{5, 5\}, \{6, 5\}, \{7, 5\}, \{8, 6\}, \{9, 6\}, \{10, 6\}, \{11, 6\}, \{11, 7\}, \{12, 7\}, \{13, 7\}, \{14, 8\}, \{15, 8\}, \{16, 8\}, \{16, 9\}, \{17, 9\}, \{18, 9\}, \{19, 9\}, \{20, 10\}, \{21, 10\}, \{22, 10\}, \{22, 11\}, \{23, 11\}, \{24, 11\}, \{25, 12\}, \{26, 12\}, \{27, 12\}, \{28, 12\}, \{28, 13\}, \{29, 13\}, \{30, 13\}, \{31, 14\}, \{32, 14\}, \{33, 14\}, \{33, 15\}, \{34, 15\}, \{35, 15\}, \{36, 15\}, \{37, 16\}, \{38, 16\}, \{39, 16\}, \{39, 17\}, \{40, 17\}, \{41, 16\}, \{42, 16\}, \{43, 15\}, \{44, 14\}, \{45, 13\}, \{45, 13\}, \{46, 12\}, \{47, 11\} \}$ 。

[0193] 矢量 RM_{2i} 经内插得 R_{2i+1} 值为:

$$[0194] \quad RM_{iF+1} = \sum_{j=0}^4 cf_j RM_{(i-1+j)F}, \quad 1 \leq i < 112/2$$

$$[0195] \quad RM_1 = (RM_0 + RM_2) / 2$$

$$[0196] \quad RM_{2*56+1} = (RM_{2*56} + RM_{2*57}) / 2$$

$$[0197] \quad RM_{2*57+1} = RM_{2*57}$$

[0198] 其中 cf_j 是内插滤波器, 系数为 $\{-0.0625, 0.5625, 0.5625, -0.0625\}$ 。然后选择滞后 L_c , 使 $R_{L_c-12} = \max \{R_i\}, 4 \leq i < 116$, 将当前帧的 NACF 置成 $R_{L_c-12}/4$ 。再搜索对应于大于 $0.9R_{L_c-12}$ 的最大相关的滞后, 消除滞后倍数, 其中

[0199]

$$R_{\max\{\lfloor L_c/M \rfloor - 14, 16\}} \cdots R_{\lfloor L_c/M \rfloor - 10} \text{ for all } 1 \leq M \leq \lfloor L_c / 16 \rfloor$$

[0200] E. 计算带能与零交叉速率

[0201] 在步骤 510, 按本发明计算 0-2kHz 带与 2kHz-4kHz 带内的能量:

$$[0202] \quad E_L = \sum_{i=0}^{159} s_L^2(n)$$

$$[0203] \quad E_H = \sum_{i=0}^{159} s_H^2(n)$$

[0204] 其中

$$[0205] \quad S_L(z) = S(z) \frac{bl_0 + \sum_{i=1}^{15} bl_i z^{-i}}{al_0 + \sum_{i=1}^{15} al_i z^{-i}}$$

$$[0206] \quad S_H(z) = S(z) \frac{bh_0 + \sum_{i=1}^{15} bh_i z^{-i}}{ah_0 + \sum_{i=1}^{15} ah_i z^{-i}}$$

[0207] $S(z)$, $S_L(z)$ 和 $S_H(z)$ 分别是输入语音信号 $s(n)$, 低通信号 $S_L(n)$ 和高通信号 $S_H(n)$ 的 z 变换, $b1 = \{0.0003, 0.0048, 0.0333, 0.1443, 0.4329,$

[0208] $0.9524, 1.5873, 2.0409, 2.0409, 1.5873, 0.9524, 0.4329, 0.1443, 0.0333, 0.0048, 0.0003\}$, $a1 = \{1.0, 0.9155, 2.4074, 1.6511, 2.0597, 1.0584, 0.7976, 0.3020, 0.1465, 0.0394, 0.0122, 0.0021, 0.0004, 0.0, 0.0, 0.0\}$, $bh = \{0.0013, -0.0189, 0.1324, -0.5737, 1.7212, -3.7867, 6.3112, -8.1144, 8.1144, -6.3112, 3.7867, -1.7212, 0.5737, -0.1324, 0.0189, -0.0013\}$ and $andah = \{1.0, -2.8818, 5.7550, -7.7730, 8.2419, -6.8372, 4.6171, -2.5257, 1.1296, -0.4084, 0.1183, -0.0268, 0.0046, -0.0006, 0.0, 0.0\}$.

[0209] 语音信号能量本身为 $E = \sum_{i=0}^{159} S^2(n)$ 。零交叉速率 ECR 计算为：

[0210] $\text{if}(s(n)s(n+1) < 0) \text{ZCR} = \text{ZCR} + 1, \quad 0 \leq n < 159$

[0211] F. 计算元音振峰余量

[0212] 在步骤 512, 对四个子帧计算当前帧的元音共振峰余量：

[0213]
$$r_{curr}(n) = s(n) - \sum_{i=1}^{10} a_i s(n-i)$$

[0214] 其中 a_i , 是相应子帧的第 i 个 LPC 系数。

[0215] IV. 有效 / 无效语音分类

[0216] 再参照图 3, 在步骤 304, 把当前帧分类为有效语音 (如讲出的单词) 或无效语音 (如背景噪声, 静寂)。图 6 的流程图 600 详细列出了步骤 304。在一较佳实施例中, 用基于双能带的取值方法确定有无有效语音。下带 (带 0) 跨越频率为 0.1-2.0kHz, 上带 (带 1) 为 2.0-4.0kHz。在当前帧编码时, 最好以下述方法确定下一帧的话音有效性检测。

[0217] 在步骤 602, 对各带 $i = 0, 1$ 计算带能 $E_b[i]$: 用下列递归公式将 III、A 节中的自相关序列扩展到 19:

[0218]
$$R(k) = \sum_{i=1}^{10} a_i R(k-i), \quad 11 \leq k \leq 19$$

[0219] 利用该公式, 从 $R(1)$ 到 $R(10)$ 中算出 $R(11)$, 从 $R(2)$ - $R(11)$ 中算出 $R(12)$, 依次类推。再用下式从扩展的自相关序列中算出带能:

[0220]
$$E_b(i) = \log_2 \left(R(0)R_h(i)(0) + 2 \sum_{k=1}^{19} R(k)R_h(i)(k) \right), \quad i = 0, 1$$

[0221] 式中 $R(k)$ 是当前帧扩展的自相关序列, $R_h(i)(k)$ 是表 1 中带 i 的带滤波器自相关序列。

[0222] 表 1: 计算带能的滤波器自相关序列

[0223]

k	$R_h(0)(k)$ band 0	$R_h(1)(k)$ band 1
0	4.230889E-01	4.042770E-01
1	2.693014E-01	-2.503076E-01

2	-1.124000E-02	-3.059308E-02
3	-1.301279E-01	1.497124E-01
4	-5.949044E-02	-7.905954E-02
5	1.494007E-02	4.371288E-03
6	-2.087666E-03	-2.088545E-02
7	-3.823536E-02	5.622753E-02

[0224]

8	-2.748034E-02	-4.420598E-02
9	3.015699E-04	1.443167E-02
10	3.722060E-03	-8.462525E-03
11	-6.416949E-03	1.627144E-02
12	-6.551736E-03	-1.476080E-02
13	5.493820E-04	6.187041E-03
14	2.934550E-03	-1.898632E-03
15	8.041829E-04	2.053577E-03
16	-2.857628E-04	-1.860064E-03
17	2.585250E-04	7.729618E-04
18	4.816371E-04	-2.297862E-04
19	1.692738E-04	2.107964E-04

[0225] 在步骤 604,平滑带能估值,并用下式对各帧更新平滑的带能估值 $E_{sm}(i)$:

$$[0226] \quad E_{sm}(i) = 0.6E_{sm}(i) + 0.4E_b(i), i = 0, 1$$

[0227] 在步骤 606,更新信号能与噪声能的估值。信号能估值 $E_s(i)$ 最好用下式更新。

$$[0228] \quad E_s(i) = \max(E_{sm}(i), E_s(i)), i = 0, 1$$

[0229] 噪声能估值 $E_n(i)$ 最好用下式更新

$$[0230] \quad E_n(i) = \min(E_{sm}(i), E_n(i)), i = 0, 1$$

[0231] 在步骤 608,两带的长期信噪比 $SNR(i)$ 计算为

$$[0232] \quad SNR(i) = E_s(i) - E_n(i), i = 0, 1$$

[0233] 在步骤 610,这些 SNR 值最好分成 8 个区 $Reg_{SNR}(i)$, 定义为 :

$$[0234] \quad Reg_{SNR}(i) = \begin{cases} 0 & 0.6SNR(i) - 4 < 0 \\ round(0.6SNR(i) - 4) & \leq 0.6SNR(i) - 4 < 7 \\ 7 & 0.6SNR(i) \geq 7 \end{cases}$$

[0235] 在步骤 612, 以下述方式按本发明判断语音有效性。若 $E_b(0) - E_n(0) > THRESH(Reg_{SNR}(0))$, 或 $E_b(1) - E_n(1) > THRESH(Reg_{SNR}(1))$, 则判定该语音帧有效, 反之为无效。THRESH 值由表 2 规定。

[0236] 表 2 : 阈值系数与 SNR 区的函数关系

[0237]

SNR Region	THRESH
0	2.807
1	2.807
2	3.000
3	3.104
4	3.154
5	3.233
6	3.459
7	3.982

[0238] 信号能估值 $E_s(i)$ 最好用下式更新 :

$$[0239] \quad E_s(i) = E_s(i) - 0.014499, \quad i = 0, 1.$$

[0240] 噪声能估值 $E_n(i)$ 最好用下式更新

$$[0241] \quad E_n(i) = \begin{cases} 4 & E_n(i) + 0.0066 < 4 \\ 23 & 23 < E_n(i) + 0.0066, \quad i = 0, 1 \\ E_n(i) + 0.0066 & otherwise \end{cases}$$

[0242] A. 拖尾帧

[0243] 信噪比很低时, 最好加“拖尾”帧提高重建语音的质量。若三个前帧分为有效而当前帧为无效, 则包括当前帧在内的后 M 帧分类为有效语音。拖尾帧数 M 确定时与表 3 中规定的 $SNR(0)$ 成函数关系。

[0244] 表 3 : 拖尾帧与 $SNR(0)$ 的函数关系

[0245]

SNR(0)	M
0	4

1	3
2	3
3	3
4	3
5	3
6	3
7	3

[0246] V. 有效语音帧的分类

[0247] 再参照图 3, 在步骤 308, 在步骤 304 分为有效的当前帧再按语音信号 $s(n)$ 呈现的特性分类。在一较佳实施例中, 有效语音分为话音, 非话音或过渡。有效语音信号呈现的周期性程度确定了它的分类。话音语音呈现最高度的周期性 (准周期特性)。非话音语音很少或不呈现周期性, 过渡语音的周期性程度在上述二者之间。

[0248] 然而, 这里描述的一般框架不限于该较佳分类方式, 下面描述特定的编 / 解码器模式。有效语音可以不同方式分类, 编码则有不同的编 / 解码器模式。技术人员应明白, 分类与编 / 解码器模式可以有許多组合方式。许多这样的组合可以按这里描述的一般框架降低平均位速率即一般框架即是把语音分成无效或有效, 再对有效语音作分类, 然后用特别适合于每一类范围内语音的编 / 解码器模式编码语音信号。

[0249] 虽然有效语音分类基于周期性程度, 但是分类判断最好不以某种周期性的直接测量为基础, 而是从步骤 302 计算的各种参数为基础, 如上下带中的信噪比和 NACF。较佳的分类可用下列伪码描述。

[0250]

```

if not(previousN ACF < 0.5 and currentN ACF > 0.6)
  if (currentN ACF < 0.75 and ZCR > 60) UNVOICED
    else if (previousN ACF < 0.5 and currentN ACF < 0.55
      and ZCR > 50) UNVOICED
      else if (currentN ACF < 0.4 and ZCR > 40) UNVOICED
if (UNVOICED and currentSNR > 28dB
  and  $E_L > \alpha E_H$ ) TRANSIENT
if (previousN ACF < 0.5 and currentN ACF < 0.5
  and  $E < 5e4 + N$ ) UNVOICED
if (VOICED and low-bandSNR > high-bandSNR
  and previousN ACF < 0.8 and
  0.6 < currentN ACF < 0.75) TRANSIENT

```

[0251] 其中

$$[0252] \quad \alpha = \begin{cases} 1.0, & E > 5e5 + N_{noise} \\ 20.0, & E \leq 5e5 + N_{noise} \end{cases}$$

[0253] N_{noise} 是背景噪声估值, E_{prev} 是前一帧输入能。

[0254] 用该伪码描述的方法可按实施的特定环境提炼。技术人员应明白,上面给出的各种阈值仅作为示例,实践中可根据实施情况要求调节。该方法还可通过增加附加的分类目录予以精炼,如将 TRASIENT 分成两类:一类用于从高能转为低能的信号,另一类用于从低能转为高能的信号。

[0255] 技术人员应明白,其它方法也可以区分语音、非语音与过渡有效语音,还可能其它有效语音的分类方法。

[0256] VI. 编 / 解码器模式选择

[0257] 在步骤 310,根据步骤 304 与 308 分类的当前帧选择编 / 解码器模式。根据一较佳实施例,模式选成如下选择:用 NELP 模式对无效帧和有效非语音帧编码,用 PPP 模式对有效语音帧编码,用 CELP 模式对有效过渡帧编码。下面描述各编 / 译码器模式。

[0258] 在一替代实施例中,无效帧用零速率模式编码。技术人员应明白,有许多要求很低位速率的其它零速率模式。研究过去的模式选择,可改良零速率模式的选择。例如,若前一帧分为有效,就可不对当前帧选择零速率模式。同样地,若下一帧有效,可不对当前帧选择零速率模式。另一方法是不对过多的连续帧(如 9 个连续帧)选用零速率模式。技术人员应明白,可对基本的选模判断作其它许多更改,以改善其在某些环境中的操作。

[0259] 如上所述,在相同一框架内,可交替地应用许多其它分类的组合和编 / 解码器模式。下面详述本发明的几种编 / 解码器模式,先介绍 CELP 模式,然后叙述 PPP 与 NELP 模式。

[0260] VII. 代码受激的线性预测 (CELP) 编码模式

[0261] 如上所述,当当前帧分为有效过渡语音时,可应用 CELP 编 / 解码模式。该模式能最精确地再现信号(与这里描述的其它模式相比),但是位速率最高。

[0262] 图 7 详细示出了 CELP 编码器模式 204 和 CELP 解码器模式 206。如图 7A 图所示,CELP 编码器模式 204 包括音调编码模块 702,编码代码簿 704 和滤波器更新模块 706。模式 204 输出编码的语音信号 $s_{enc}(n)$,最好包括传输给 CELP 编码器模式 206 的代码簿参数与音调滤波器参数。如图 7B 所示,模式 206 包括解码代码簿模块 708,音调滤波器 710 和 LPC 合成滤波器 712。CELP 模式 206 接收编码的语音信号而输出合成的语音信号 $\hat{S}(n)$ 。

[0263] A. 音调编码模块

[0264] 音调编码模块 702 接收语音信号 $s(n)$ 和前一帧量化的余量 $P_c(n)$ (下述)。根据该输入,音调解码模块 702 产生目标信号 $x(n)$ 与一组音调滤波器参数。在一实施例中,这类参数包括最佳音调滞后 L^* 与最佳音调增益 b^* 。这类参数按“分析加合成”法选择,其中解码处理选择的音调滤波器参数,可将输入语音与用这些参数合成的语音之间的加权误差减至最小。

[0265] 图 8 示出了音调编码模块 702,这包括感性加权滤波器 803,加法器 804 与 816,加权的 LPC 合成滤波器 806 与 808,延迟与增益 810 及最小平方和 812。

[0266] 感性加权滤波器 802 用于对原始语音与以感性有意义的方式合成的语音之间的

误差加权。

[0267] 感性加权滤波器的形式为

$$[0268] \quad W(z) = \frac{A(z)}{A(z/\gamma)}$$

[0269] 式中 $A(z)$ 是 LPC 预测误差滤波器, γ 最好等于 0.8。加权的 LPC 分析滤波器 806 接收原始参数计算模块 202 算出的 LPC 系数。滤波器 806 输出的 $a_{zir}(n)$ 是给出 LPC 系数的零输入响应。加法器 804 将负输入 $a_{zir}(n)$ 与滤波器的输入信号相加以形成目标信号 $x(n)$ 。

[0270] 延迟与增益 810 对给定的音调滞后 L 与音调增益 B 输出估算的间调滤波器输出 $bp_L(n)$, 延迟与增益 810 接收前一帧量化的剩余样本 $P_c(n)$ 和估算的音调滤波器将来的输出 $P_0(n)$, 按下式形成 $P(n)$ 。

$$[0271] \quad P(n) = \begin{cases} p_c(n) & -128 < n < 0 \\ p_0(n) & 0 \leq n < L_p \end{cases}$$

[0272] 然后延迟 L 个样本, 用 b 标定, 形成 $bp_L(n)$ 。 L_p 是子帧长度 (最好为 40 样本)。在一较佳实施例中, 音调滞后 L 用 8 位代表, 可以取值 20.0, 20.5, 21.0, 21.5... 126.0, 126.5, 127.0, 127.5。

[0273] 加权的 LPC 分析滤波器 808 用当前 LPC 系数滤波 $bp_L(n)$ 而得出 $by_L(n)$ 。加法器 816 将负输入 $by_L(n)$ 与 $x(n)$ 相加, 其输出被最小平方和 812 接收, 后者选择标为 L^* 的最佳 L 和标为 b^* 的最佳 b , 而 L 和 b 的值按下式将 $E_{pitch}(L)$ 减至最小:

$$[0274] \quad E_{pitch}(L) = \sum_{n=0}^{L_p-1} \{x(n) - by_L(n)\}^2$$

$$[0275] \quad \text{若 } E_{xy}(L) \triangleq \sum_{n=0}^{L_p-1} x(n)y_L(n) \quad \text{且 } E_{yy}(L) \triangleq \sum_{n=0}^{L_p-1} y_L(n)^2 \quad \text{则对规定的 } L \text{ 值将 } E_{pitch} \text{ 减}$$

至最小的 b 值为:

$$[0276] \quad b^* = \frac{E_{xy}(L)}{E_{yy}(L)}$$

[0277] 因此

$$[0278] \quad E_{pitch}(L) = K - \frac{E_{xy}(L)^2}{E_{yy}(L)}$$

[0279] 式中 K 是可以忽略的常数

[0280] 首先确定使 $E_{pitch}(L)$ 最小的 L 值, 再计算 b^* , 求出 L 与 b 的最佳值 (L^* 与 b^*)

[0281] 最好对各子帧算出这些音调滤波器参数, 量化后作有效传输。在一实施例中, 第 j 个子帧的传输代码 $PLAG_j$ 与 $PGAIN_j$ 计算成

[0282]

$$PGAIN_j = \left\lfloor \min\{b^*, 2\} \frac{8}{2} + 0.5 \right\rfloor - 1$$

$$[0283] \quad PLAG_j = \begin{cases} 0, & PGAIN_j = -1 \\ 2L^*, & 0 \leq PGAIN_j < 8 \end{cases}$$

[0284] 若 $PLAG_j$ 置 0, 则将 $PGAIN_j$ 调至 -1。这些传输代码发送给 CELP 解码器模式 206

作为音调滤波器参数,成为编码的语音信号 $s_{\text{enc}}(n)$ 的组成部分。

[0285] B. 编码代码簿

[0286] 编码代码簿 704 接收目标信号 $x(n)$, 并确定一组供 CELP 解码器模式 206 使用的代码簿激励参数, 与音调滤波器参数一起, 以重建量化的剩余信号。

[0287] 编码代码簿 704 首先将 $x(n)$ 更新如下:

$$[0288] \quad x(n) = x(n) - y_{\text{pzir}}(n), 0 \leq n < 40$$

[0289] 式中 $y_{\text{pzir}}(n)$ 是加权的 LPC 合成滤波器 (带有从前一帧结尾保留数据的存储器) 对某一输入的输出, 而该输入是带参数 L^* 与 b^* (和前一子帧处理的存储器) 的音调滤波器的零输入响应。

[0290] 由于 $\vec{d} = H^T \vec{x}$, 用而建立一反滤波目标 $\vec{d} = \{dn\}$, $0 < n < 40$, 其中

$$[0291] \quad H = \begin{bmatrix} h_0 & 0 & 0 & \dots & 0 \\ h_1 & h_0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ h_{39} & h_{38} & h_{37} & \dots & h_0 \end{bmatrix}$$

[0292] 是脉冲响应矩阵, 由脉冲响应 $\{h_n\}$ 和 $\vec{x} = \{x(n)\}$, $0 \leq n < 40$ 形成, 同样产生了两个以上矢量 $\hat{\phi} = \{\phi_n\}$ 和 \vec{s} :

$$[0293] \quad \vec{s} = \text{sign}(\vec{d})$$

$$[0294] \quad \phi_n = \begin{cases} 2 \sum_{i=0}^{39-n} h_i h_{i+n}, & 0 < n < 40 \\ \sum_{i=0}^{39} h_i^2, & n = 0 \end{cases}$$

[0295] 其中

$$[0296] \quad \text{sign}(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$

[0297] 编码代码簿 704 将值 E_{xy}^* 与 E_{yy}^* 初始化为零, 并按以下公式最好用四个 N 值 (0, 1, 2, 3) 搜索最佳激励参数。

$$[0298] \quad \vec{p} = (N + \{0, 1, 2, 3, 4\}) \% 5$$

$$[0299] \quad A = \{p_0, p_0+5, \dots, i' < 40\}$$

$$[0300] \quad B = \{p_1, p_1+5, \dots, k' < 40\}$$

$$[0301] \quad \text{Den}_{i,k} = 2 \phi_0 + s_i s_k \phi_{|k-i|}, i \in A \quad k \in B$$

$$[0302] \quad \{I_0, I_1\} = \arg \max_{\substack{i \in A \\ k \in B}} \left\{ \frac{|d_i| + |d_k|}{\text{Den}_{i,k}} \right\}$$

$$[0303] \quad \{S_0, S_1\} = \{s_{I_0}, s_{I_1}\}$$

$$[0304] \quad E_{xy0} = |d_{I_0}| + |d_{I_1}|$$

$$[0305] \quad E_{yy0} = E_{yy_{I_0, I_1}}$$

$$[0306] \quad A = \{p_2, p_2+5, \dots, i' < 40\}$$

$$[0307] \quad B = \{p_3, p_3+5, \dots, k' < 40\}$$

$$[0308] \quad Den_{i,k} = Eyy0 + 2\phi_0 + s_i \left(S_0 \phi_{|I_0-i|} + S_1 \phi_{|I_1-i|} \right)$$

$$[0309] \quad + s_k \left(S_0 \phi_{|I_0-k|} + S_1 \phi_{|I_1-k|} \right) + s_i s_k \phi_{|k-i|}$$

$$[0310] \quad i \in A, k \in B$$

$$[0311] \quad \{I_2, I_3\} = \arg \max_{\substack{i \in A \\ k \in B}} \left\{ \frac{Exy0 + |d_i| + |d_k|}{Den_{i,k}} \right\}$$

$$[0312] \quad \{S_2, S_3\} = \{s_{I_2}, s_{I_3}\}$$

$$[0313] \quad Exy1 = Exy0 + |d_{I_2}| + |d_{I_3}|$$

$$[0314] \quad Eyy1 = Den_{I_2, I_3}$$

$$[0315] \quad A = \{p_4, p_4+5, \dots, i' < 40\}$$

$$[0316] \quad Den_i = Eyy1 + \phi_0 + s_i \left(S_0 \phi_{|I_0-i|} + S_1 \phi_{|I_1-i|} + S_2 \phi_{|I_2-i|} + S_3 \phi_{|I_3-i|} \right), \quad i \in A$$

$$[0317] \quad I_4 = \arg \max_{i \in A} \left\{ \frac{Exy1 + |d_i|}{Den_i} \right\}$$

$$[0318] \quad S_4 = s_{I_4}$$

$$[0319] \quad Exy2 = Exy1 + |d_{I_4}|$$

$$[0320] \quad Eyy2 = Den_{I_4}$$

[0321] 若

[0322]

$$Exy2^2 Eyy^* > Exy^2 Eyy2 \{$$

$$Exy^* = Exy2$$

$$Eyy^* = Eyy2$$

$$\{ind_{p_0}, ind_{p_1}, ind_{p_2}, ind_{p_3}, ind_{p_4}\} = \{I_0, I_1, I_2, I_3, I_4\}$$

$$\{sgn_{p_0}, sgn_{p_1}, sgn_{p_2}, sgn_{p_3}, sgn_{p_4}\} = \{S_0, S_1, S_2, S_3, S_4\}$$

}

[0323] 编码代码簿 704 把代码簿增益 G^* 计算成 Exy^*/Eyy^* , 然后对第 j 个子帧将该组激励参数量化成下列传输码:

[0324]

$$CBI_{jk} = \left\lfloor \frac{ind_k}{5} \right\rfloor, \quad 0 \leq k < 5$$

$$[0325] \quad \text{SIGN}jk = \begin{cases} 0, & \text{sgn}_k = 1 \\ 1, & \text{sgn}_k = -1 \end{cases}, \quad 0 \leq k < 5$$

[0326]

$$\text{CBG}j = \left\lceil \min\{\log_2(\max\{1, G^*\}), 11.2636\} \frac{31}{11.2636} + 0.5 \right\rceil$$

[0327] 量化的增益 \hat{G}^* 为 $2^{\frac{11.2636}{31} \text{CBG}j}$ 。

[0328] 除去音调解码模块 702, 只作代码簿搜索以对四个子帧都确定指数 I 与增益 G, 就可实现 CELP 编 / 解码器模式的较低位速率实施例。技术人员应明白如何扩充上述想法来实现该较低的位速率实施例。

[0329] C. CELP 解码器

[0330] CELP 解码器模式 206 从 CELP 解码器模式 204 接收解码的语音信号, 最好包括代码簿激励参数与音调滤波器参数, 并根据该数据输出合成的语音 $\hat{S}(n)$ 。解码代码簿模块 708 接收代码簿激励参数, 产生增益为 G 的激励信号 $C_b(n)$ 。j 个子帧的激励信号 $C_b(n)$ 包含大多数零, 但五个位置例外:

$$[0331] \quad I_k = 5\text{CBI}jk+k, \quad 0 \leq k < 5$$

[0332] 它相应地具有脉冲值:

$$[0333] \quad S_k = 1 - 2\text{SIGN}jk, \quad 0 \leq k < 5$$

[0334] 所有值均用计算为 $2^{\frac{11.2636}{31} \text{CBG}j}$ 的增益 G 标定, 以提供 $G_c b(n)$ 。

[0335] 音调滤波器 710 按下列公式对接收传输代码的音调滤波器参数解码:

$$[0336] \quad \hat{L}^* = \frac{\text{PLAG}j}{2}$$

$$[0337] \quad \hat{b}^* = \begin{cases} 0, & \hat{L}^* = 0 \\ \frac{2}{8} \text{PGAIN}j, & \hat{L}^* \neq 0 \end{cases}$$

[0338] 音调滤波器 710 接着滤波 $G_c b(n)$, 滤波器的传递函数为:

$$[0339] \quad \frac{1}{P(z)} = \frac{1}{1 - \hat{b}^* z^{-\hat{L}^*}}$$

[0340] 在一实施例中, 在音调滤波器 710 之后, CELP 解码器模式 706 还加接了额外的滤波操作的音调前置滤波器 (未示出)。音调前置滤波器的滞后与音调滤波器 710 的滞后相同, 但其增益最好是最高达 0.5 的音调增益的一半。

[0341] LPC 合成滤波器 712 接收重建的量化剩余信号 $\hat{r}(n)$, 输出合成的语音信号 $\hat{S}(n)$ 。

[0342] D. 滤波器更新模块

[0343] 滤波器更新模块 706 像前一节描述的那样合成语音, 以便更新滤波器存储器。滤波器更新模块 706 接收代码簿激励参数与音调滤波器参数, 产生激励信号 $c_b(n)$, 对 $G_c b(n)$ 作音调滤波, 再合成 $\hat{S}(n)$ 。在解码器作这一合成, 就更新了音调滤波器与 LPC 合成滤波器中的存储器, 供处理后面的子帧使用。

[0344] VIII. 原型音调周期 (PPP) 编码模式

[0345] 原型音调周期 (PPP) 编码法利用语音信号的周期性实现比 CELP 编码法可得到的更低的位速率。一般而言, PPP 编码法涉及提取一代表性的剩余个周期, 这里称为原型余量, 然后用该原型通过当前帧的原型余量与前一帧的类似音调周期 (如果最后帧是 PPP, 即为原型余量) 之间作内插, 在该帧内建立早期音调周期, PPP 编码法的有效性 (降低位速率) 部分取决于如何使当前与前一原型余量精密地类似于介入的音调周期。为此, 最好将 PPP 编码法应用于呈现出相对高度周期性的语音信号 (如话音语音), 这里指准周期语音信号。

[0346] 图 9 详细示出了 PPP 编码器模式 204 和 PPP 解码器模式 206, 前者包括提取模块 904, 旋转相关器 906, 编码代码簿 908 和滤波器更新模块 910。PPP 编码器模式 204 接收剩余信号 $r(n)$, 输出编码的语音信号 $s_{enc}(n)$, 最好包括代码簿参数和旋转参数。PPP 解码器模式 206 包括代码簿解码器 912、旋转器 914, 加法器 916, 周期内插器 920 和弯曲滤波器 918。

[0347] 图 10 的流程图 1000 示出 PPP 编码的步骤, 包括编码与解码。这些步骤与 PPP 编码器模式 204 和 PPP 解码器模式 206 一起讨论。

[0348] A. 提取模块

[0349] 在步骤 1002, 提取模块 904 从剩余信号 $r(n)$ 中提取原型余量 $r_p(n)$ 。如 III、F、节所述, 初始参数计算模块 202 用 LPC 分析滤波器计算各帧的 $r_p(n)$ 。在一实施例中, 如 VII、A 节所述, 该滤波器的 LPC 系数作感性加权。 $r_p(n)$ 的长度等于原始参数计算模块 202 在当前帧最后一个子帧中算出的音调滞后 L 。

[0350] 图 11 是详细示出步骤 1002 的流程图。PPP 提取模块 904 最好尽量接近帧结束时选择音调周期, 并加下述的某些限制。图 12 示出一例基于准周期语音计算的剩余信号, 包括当前帧与前一帧的最后一个子帧。

[0351] 在步骤 1102, 确定“无切割区”。无切割区限定一组余量中不能是原型余量终点的样本。无切割区保证余量的高能区不出现在原型的开始或结束 (会造成输出中允许出现的断续性)。计算 $r(n)$ 最后 L 个样本每一样本的绝对值。变量 P_s 置成等于最大绝对值 (这里称为“音调尖峰”) 样本的时间指数。例如, 若音调尖峰出现在最后 L 个样本的最后一个样本中, $P_s = L-1$ 。在一实施例中, 无切割区的最小样本 CF_{min} 置成 P_s-6 或 $P_s-0.25L$, 无论哪个更小。无切割区的最大值 CF_{max} 置成 P_s+6 或 $P_s+0.25L$, 无论哪个更大。

[0352] 在步骤 1104, 从余量中切割 L 个样本, 选择原型余量, 在区域终点不能在无切割区内的约束下, 选择的区域尽量接近帧的结束处。用以下列伪码描述的算法确定原型余量的 L 个样本:

[0353]

```

if
    (CFmin < 0) {
        for(i = 0 to L + CFmin-1) rp(i) = r(i+160-L)
        for(i = CFmin to L-1) rp(i) = r(i+160-2L)
    }
else if
    (CFmax ≤ L {
        for(i = 0 to CFmin-1) rp(i) = r(i+160-L)
        for(i = CFmin to L-1) rp(i) = r(i+160-2L)

```

[0354]

```

else {
    for(i = 0 to L-1) rp(i) = r(i+160-L)

```

[0355] B. 旋转相关器

[0356] 再参照图 10, 在步骤 1004, 旋转相关器 906 根据当前原型余量 $r_p(n)$ 和前一帧的原型余量 $r_{prev}(n)$ 计算一组旋转参数。这些参数描述怎样最佳地旋转和标定 r_{prev} 以用作 $r_p(n)$ 的预测器。在一实施例中, 这组旋转参数包括最佳旋转 R^* 与最佳增益 b^* 。图 13 是详细示出步骤 1004 的流程图。

[0357] 在步骤 1302, 对原型音调余量周期 $r_p(n)$ 作循环滤波, 算出感性加权的目标信号 $x(n)$ 。这是按如下方式实现的。由 $r_p(n)$ 产生临时信号 $tmp1(n)$:

$$[0358] \quad tmp1(n) = \begin{cases} r_p(n), & 0 \leq n < L \\ 0, & L \leq n < 2L \end{cases}$$

[0359] 将其用零存储器的加权 LPC 合成滤波器滤波, 以提供输出 $tmp2(n)$ 。在一实施例中, 使用的 LPC 系数是对应于当前帧最后一个子帧的感性加权系数。于是, 目标信号 $x(n)$ 为 :

$$[0360] \quad x(n) = tmp2(n) + tmp2(n+L), 0 \leq n < L$$

[0361] 在步骤 1304, 从前一帧量化的元音共振峰余量 (也存在音调滤波器的存储器中) 中提取前一帧的原型余量 $y_{prev}(n)$ 。该前一原型余量最好定义为前一帧元音共振峰余量的最后 LP 值, 若前一帧不是 PPP 帧, L_p 等于 L , 否则置成前一音调滞后。

[0362] 在步骤 1306, 把 $y_{prev}(n)$ 的长度改为与 $x(n)$ 一样长, 从而正确地计算相关性。这里把改变采样信号长度的这种技术称为弯曲。弯曲的音调激励信号 $y_{w_{prev}}(n)$ 可以描述成 :

$$[0363] \quad rw_{prev} = r_{prev}(n * TWF), 0 \leq n < L$$

[0364] 式中 TWF 是时间弯曲系数 L_p/L 。最好用一套 sinc 函数表计算非整数点 $n * TWF$ 的样本值。选择的 sinc 序列是 $sinc(-3-F : 4-F)$, F 是 $n * TWF$ 的小数部分, 含入最接近的 $1/8$ 倍数。该序列的开头对准 $r_{prev}(N-3) \% L_p$, N 是 $n * TWF$ 在含入最接近第八位后的整数部分。

[0365] 在步骤 1308, 循环滤波弯曲的音调激励信号 $rw_{prev}(n)$, 得出 $y(n)$ 。该操作与上述对步骤 1302 作的操作一样, 但应用于 $rw_{prev}(n)$ 。

[0366] 在步骤 1310, 计算音调旋转搜索范围, 首先计算期望的旋转 E_{rot} :

$$[0367] \quad E_{rot} = L - \text{round} \left(L \text{frac} \left(\frac{(160 - L)(L_p + L)}{2L_p L} \right) \right)$$

[0368] frac(x) 给出 X 的小数部分。若 $L < 80$, 则音调旋转搜索范围定义为 $\{E_{rot}-8, E_{rot}-7.5, \dots, E_{rot}+7.5\}$ 和 $\{E_{rot}-16, E_{rot}-15, \dots, E_{rot}+15\}$, 其中 $L > 80$ 。

[0369] 在步骤 1312, 计算旋转参数, 最佳旋转 R^* 与最佳增益 b^* 。在 $x(n)$ 和 $y(n)$ 之间导致最佳预测的音调旋转与相应的增益 b 一起选择。这些参数最好选成将误差信号 $e(n) = x(n) - y(n)$ 减至最小。最佳旋转 R^* 与最佳增益 b^* 是导致 $E_{xy_R}^2 / E_{yy}$ 最大值的那些旋转 R 与

增益 b 值, 其中 $E_{xy_R} = \sum_{i=0}^{L-1} x((i+R)\%L)y(i)$ 和 $E_{yy} = \sum_{i=0}^{L-1} y(i)y(i)$ 在旋转 R^* 时的最佳增益

b^* 为 $E_{xy_R}^* / E_{yy}$ 。对于旋转的小数值, 通过对在整数旋转值时算出的 E_{xy_R} 值作内插, 求出 E_{xy_R} 的近似值。应用了一种简单的四带内插滤波器, 如

$$[0370] \quad E_{xy_R} = 0.54(E_{xy_{R'}} + E_{xy_{R'+1}}) - 0.04*(E_{xy_{R'-1}} + E_{xy_{R'+2}})$$

[0371] R 是非整数的旋转 (精度 0.5), $R' = |R|$ 。

[0372] 在一实施例中, 旋转参数作量化以有效地传输。最佳增益 \hat{b}^* 最好在 0.0625 和 4.0 之间均匀地量化成:

[0373]

$$PGAIN = \max \left\{ \min \left(63 \left(\frac{b^* - 0.0625}{4 - 0.0625} \right) + 0.5, 63 \right), 0 \right\}$$

[0374] 式中 PGAIN 为传输码, 量化增益 b^* 由 $\max \{0.0625 + (PGAIN(4 - 0.0625) / 63), 0.0625\}$ 给出。将最佳旋转 R^* 量化成传输码 PROT, 若: $L < 80$ 。将其置成 $2(R^* - E_{rot} + 8)$, $L \geq 80$, 则 $R^* - E_{rot} + 16$ 。

[0375] C. 编码代码簿

[0376] 再参照图 10, 在步骤 1006, 编码代码簿 908 根据接收的目标信号 $x(n)$ 产生一组代码簿参数。代码簿 908 设法求出一个或多个代码矢量, 经标定, 相加和滤波后, 加成接近 $x(n)$ 的信号。在一实施例中, 编码代码簿 908 构成多级代码簿, 最好是三级, 每级产生一种标定的代码矢量。因此, 该组代码簿参数包括了对应于三种代码矢量的标引和增益。图 14 是详细示出步骤 1006 的流程图。

[0377] 在步骤 1402, 搜索代码簿之前, 将目标信号 $x(n)$ 更新成

$$[0378] \quad x(n) = x(n) - by((n - R^*) \% L), 0 \leq n < L$$

[0379] 若在上述减法中旋转 R^* 不是整数 (即有小数 0.5), 则

$$[0380] \quad y(i - 0.5) = -0.0073(y(i - 4) + y(i + 3)) + 0.0322(y(i - 3) + y(i + 2))$$

$$[0381] \quad -0.1363(y(i - 2) + y(i + 1)) + 0.6076(y(i - 1) + y(i))$$

[0382] 式中 $i = n - |R^*|$

[0383] 在步骤 1404, 将代码簿值分成多个区域。根据一实施例, 把代码簿确定为:

$$[0384] \quad c(n) = \begin{cases} 1, & n = 0 \\ 0, & 0 < n < L \\ CBP(n - L), & L \leq n < 128 + L \end{cases}$$

[0385] 式中 CBP 是随机或训练的代码簿值。技术人员应知道这些代码簿值是如何产生的。把代码簿划分成多个区域,长度各为 L。第一区是单脉冲,其余各区由随机或训练的代码簿值组成。区数 N 将为 $\lceil 128/L \rceil$ 。

[0386] 在步骤 1406,代码簿的多个区都作循环滤波而产生滤波的代码簿, $y_{reg}(n)$, 其串联是信号 $y(n)$ 。对每一区,按上述步骤 1302 作循环滤波。

[0387] 在步骤 1408,计算各区滤波的代码簿能量 $E_{yy}(reg)$ 并存贮起来:

$$[0388] \quad E_{yy}(reg) = \sum_{i=0}^{L-1} y_{reg}(i), \quad 0 \leq reg < N$$

[0389] 在步骤 1410,计算多级代码簿各级的代码簿参数(即代码矢量标引与增益)。根据一实施例,使 $Region(I) = reg$, 定义为其中有样本 I 的区即,

$$[0390] \quad Region(I) = \begin{cases} 0, & 0 \leq I < L \\ 1, & L \leq I < 2L \\ 2, & 2L \leq I < 3L \\ \dots & \dots \end{cases}$$

[0391] 并假定将 $E_{xy}(I)$ 定义为:

$$[0392] \quad E_{xy}(I) = \sum_{i=0}^{L-1} x(i)y_{Region(I)}((i+I)\%L)$$

[0393] 第 j 代码簿级的代码簿参数 I^* 与 G^* 用下列伪码计算:

[0394]

```

Exy* = 0, Eyy* = 0
for(I = 0 to 127) {
  compute Exy(I)
  if (Exy(I)√Eyy* > Exy*(I)√Eyy(Region(I))) {
    Exy* = Exy(I)
    Eyy* = Eyy(Region(I))
    I* = I
  }
}

```

[0395] 而且 $G^* = E_{xy}^*/E_{yy}^*$ 。

[0396] 根据一实施例,代码簿参数量化后作有效传输。传输代码 CBI_j ($j =$ 级数 $-0, 1$ 或 2) 最好置成 I^* , 而传输代码 CBG_j 与 $SIGN_j$ 通过量化增益 G^* 而设置:

$$[0397] \quad SIGN_j = \begin{cases} 0, & G^* \geq 0 \\ 1, & G^* < 0 \end{cases}$$

[0398]

$$CBG_j = \left\lfloor \min \left\{ \max \left\{ 0, \log_2(|G^*|) \right\}, 11.25 \right\} \frac{4}{3} + 0.5 \right\rfloor$$

[0399] 量化的增益 \hat{G}^* 为

$$[0400] \quad \hat{G}^* = \begin{cases} 2^{0.75CBG_j} & SIGN_j = 0 \\ -2^{0.75CBG_j} & SIGN_j \neq 0 \end{cases}$$

[0401] 然后减量当前级代码簿矢量的贡献,更新目标信号 $x(n)$:

$$[0402] \quad x(n) = x(n) - \hat{G}^* y_{Region(I^*)}((n+I^*)\%L), \quad 0 \leq n < L$$

[0403] 上述从伪码开始的步骤重复进行,对第二和第三级计算 I^* , G^* 和相应的传输代码。

[0404] D. 滤波器更新模块

[0405] 再参照图 10,在步骤 1008,滤波器更新模块 910 更新 PPP 解码器模式 204 所使用的滤波器。图 15A 与 16A 示出两个替代的滤波器更新模块 910 的实施例。如图 15A 的第一替代实施例,滤波器更新模块 910 包括解码代码簿 1502,旋转器 1504,弯曲滤波器 1506,加法器 1510,对准与内插模块 1508,更新音调滤波器模块 1512,和 LPC 合成滤波器 1514。图 16A 的第二实施例包括解码代码簿 1602,旋转器 1604,弯曲滤波器 1606,加法器 1608,更新音调滤波器模块 1610,循环 LPC 合成滤波器 1612 和更新 LPC 滤波器模块 1614,图 17 与 18 是详细示出这两个实施例中步骤 1008 的流程图。

[0406] 在步骤 1702 (和 1802,两实施例的第一步),由代码簿参数与旋转参数重建当前重建的原型余量 $r_{curr}(n)$,长度为 L 样本。在一实施例中,旋转器 1504 (和 1604) 按下式旋转弯曲型的前一原型余量:

$$[0407] \quad r_{curr}((n+R^*)\%L) = brw_{prev}(n), \quad 0 \leq n < L$$

[0408] 式中 r_{curr} 是要建立的当前原型, rw_{prev} 是由音调滤波器存储器中最新 L 个样本获得的弯曲型前一周期 (如 VIII A 节所述, $TWF = L_p/L$),由包传输码获得的音调增益 b 和旋转 R 为:

$$[0409] \quad b = \max \left\{ 0.0625 \left(\frac{PGAIN(4 - 0.0625)}{63} \right), 0.0625 \right\}$$

$$[0410] \quad R = \begin{cases} \frac{PROT}{2} + E_{rot} - 8, & L < 80 \\ PROT + E_{rot} - 16, & L \geq 80 \end{cases}$$

[0411] 其中 E_{rot} 是上述 VIII B 节算出的期望的旋转。

[0412] 解码代码簿 1502 (和 1602) 将三个代码簿级的每级的贡献加到 $r_{curr}(n)$:

$$[0413] \quad r_{curr}((n-i)\%L) = r_{curr}((n-I)\%L) + \begin{cases} G, & I < L, n = 0 \\ GCBP(I-L+n), & I \geq L, 0 \leq n < L \end{cases}$$

[0414] 式中 $I = CBI_j$, G 如上节所述由 CBG_j 和 SIG_j 获得, j 为级数。

[0415] 在这方面,滤波器更新模块 910 的这两个替代实施例有所不同。先参照图 15A 的实施例,在步骤 1704,从当前帧开头到当前原型余量开头,对准与内插模块 1508 填入剩余样本的其余部分 (如图 12 所示)。这里对剩余信号对准和内插。然而,如下所述,还对语音信号作同样的操作。图 19 是详细描述步骤 1704 的流程图。

[0416] 在步骤 1902,确定前一滞后 LP 是否相对于当前滞后 L 为两倍或是一半。在一实施例中,其它倍数不太可能,故不予考虑。若 $L_p > 1.85L$, LP 为一半,只使用前一周期 $r_{prev}(n)$ 的前一半。若 $L_p > 0.54L$,当前滞后 L 可能加倍,因而 LP 也加倍,前一周期 $R_{prev}(n)$ 反复扩展。

[0417] 在步骤 1904,如步骤 1306 所述, $r_{prev}(n)$ 弯成 $rw_{prev}(n)$, $TWF-LP/L$,因而两原型余量的长度现在相同。注意,该操作在步骤 1702 执行,如上所述,做法是弯曲滤波器 1506。技术人员应明白,若弯曲滤波器 1506 对对准与内插模块 1508 有输出,就不需要步骤 1904。

[0418] 在步骤 1906, 计算允许的对准旋转范围。期望的对准旋转 EA 的计算与 VIII B 节所述的 E_{rot} 的计算相同。对准旋转搜索范围定义为 $\{E_A - \delta A, E_A - \delta A + 0.5, E_A - \delta A + 1, \dots, E_A - \delta A - 1.5, E_A - \delta A - 1\}$, $\delta A = \max\{6, 0.15L\}$ 。

[0419] 在步骤 1908, 把整数对准旋转 R 前一与当前原型周期之间的交叉相关性计算成

$$[0420] \quad C(A) = \sum_{i=0}^{L-1} r_{curr}((i+A)\%L) r_{w_{prev}}(i)$$

[0421] 通过在整数旋转处内插相关值, 近似算出非整数旋转 A 的交叉相关性:

$$[0422] \quad C(A) = 0.54(C(A') + C(A'+1)) - 0.04(C(A'-1) + C(A'+2))$$

[0423] 式中 $A' = A - 0.5$ 。

[0424] 在步骤 1910, 将导致 C(A) 最大值的 A 值 (在允许旋转范围内) 选为最佳对准, A^* 。

[0425] 在步骤 1912, 按下述方法算出中间样本的平均滞后或音调周期 L_{av} 。周期数估值 N_{per} 算为

$$[0426] \quad N_{per} = \text{round} \left(\frac{A^*}{L} + \frac{(160-L)(L_p+L)}{2L_pL} \right)$$

[0427] 中间样本的平均滞后为

$$[0428] \quad L_{av} = \frac{(160-L)L}{N_{per}L - A^*}$$

[0429] 在步骤 1914, 根据下述在前一与当前原型余量之间的内插, 算出当前帧中其余的剩余样本:

$$[0430] \quad \hat{r}(n) = \begin{cases} \left(1 - \frac{n}{160-L}\right) r_{w_{prev}}((n\alpha)\%L) \\ + \frac{n}{160-L} r_{curr}((n\alpha + A^*)\%L), & 0 \leq n < 160-L \\ r_{curr}(n+L-160), & 160-L \leq n < 160 \end{cases}$$

[0431] 式中 $x = L/L_{av}$ 。非整数点 \tilde{n} 的样本值 (等于 $n\alpha$ 或 $n\alpha + A^*$) 用一套 sinc 函数表计算。选择的 sinc 序列为 $\text{sinc}(-3-F:4-F)$, 其中 F 是 n 舍入最接近 1/8 倍数的小数部分, 序列开头对准 $r_{prev}((N-3)\%LP)$, N 是 \tilde{n} 舍入最接近 1/8 后的整数部分。

[0432] 注意, 该操作与上述步骤 1306 的弯曲基本上相同。因此, 在一替代实施例中, 步骤 1914 的内插值用弯曲滤波器计算。技术人员应明白, 对于这里描述的各种目的, 重复使用单个弯曲滤波器更经济。

[0433] 参照图 17, 在步骤 1706, 更新音调滤波器模块 1512 从重建的余量 $\hat{r}(n)$ 将值复制到音调滤波器存储器。同样地, 也要更新音调滤波器的存储器。

[0434] 在步骤 1708, LPC 合成滤波器 1514 对重建的余量 $\hat{r}(n)$ 滤波, 作用是更新 LPC 合成滤波器的存储器。

[0435] 现在描述图 16A 的第二个滤波器更新模块 910 实施例。如步骤 1702 所述, 在步骤 1802, 由代码簿与旋转参数重建原型余量, 导致 $r_{curr}(n)$ 。

[0436] 在步骤 1804,按下式从 $r_{curr}(n)$ 复制 L 样本复制件,更新音调滤波器模块 1610 更新音调滤波器存储器。

[0437] $pitch_mem(i) = r_{curr}((L-(131\%L)+i)\%L), 0 \leq i < 131$

[0438] 或者

[0439] $pitch_mem(131-1-i) = r_{curr}(L-1-i\%L), 0 \leq i < 131$

[0440] 其中 131 最好是最大滞后为 127.5 的音调滤波器阶数。在一实施例中,音调前置滤波器的存储器同样用当前周期 $r_{curr}(n)$ 的复制件替换:

[0441] $pitch_prefilt_mem(i) = pitch_mem(i), 0 \leq i < 131$

[0442] 在步骤 1806, $r_{curr}(n)$ 最好应用感性加权的 LPC 系数循环滤波,如 VIII B 节所述,导致 $s_c(n)$ 。

[0443] 在步骤 1808,用 $s_c(n)$ 的值,最好是后 10 个值(对第 10 阶 LPC 滤波器)更新 LPC 合成滤波器的存储器。

[0444] E. PPP 解码器

[0445] 参照图 9 和 10,在步骤 1010,PPP 解码器模式 206 根据收到的代码簿与旋转参数重建原型余量 $r_{curr}(n)$ 。解码代码簿 912,旋转器 914 和弯曲滤波器 918 的工作方式如上节所述。周期内插器 920 接收重建的原型余量 $r_{curr}(n)$ 和前一重建的原型余量 $r_{curr}(n)$,在两个原型之间内插样本,并输出合成的语音信号 $\hat{s}(n)$ 。下节描述周期内插器 920。

[0446] F. 周期内插器

[0447] 在步骤 1012,周期内插器 920 接收 $r_{curr}(n)$,输出合成的语音信号 $\hat{s}(n)$ 。图 15A 与 16b 是两个周期内插器 920 的替代实施例。在图 15B 的第一例中,周期内插器 920 包括对准与内插模块 1516, LPC 合成滤波器 1518 和更新音调滤波器模块 1520。图 16B 的第二例包括循环 LPC 合成滤波器 1616,对准与内插模块 1618,更新音调滤波器模块 1622 和更新 LPC 滤波器模块 1620。图 20 和 21 表示两实施例的步骤 1012 的流程图。

[0448] 参照图 15B,在步骤 2002,对准与内插模块 1516 对当前剩余原型 $r_{curr}(n)$ 与前一剩余原型 $r_{prev}(n)$ 之间的样本重建剩余信号,形成 $\hat{r}(n)$,模块 1516 以步骤 1704 所述的方式(图 19)操作。

[0449] 在步骤 2004,更新音调滤波器模块 1520 根据重建的剩余信号 $\hat{r}(n)$ 更新音调滤波器存储器,如步骤 1706 所述。

[0450] 在步骤 2006, LPC 合成滤波器 1518 根据重建的剩余信号 $\hat{r}(n)$ 合成输出语音信号 $\hat{s}(n)$,操作时, LPC 滤波器存储器自动更新。

[0451] 参照图 16B 和 21,在步骤 2102,更新音高调滤波器模块 1622 根据重建的当前剩余原型 $r_{curr}(n)$ 更新音调滤波器存储器,如步骤 1804 所示。

[0452] 在步骤 2104,循环 LPC 合成滤波器 1616 接收 $r_{curr}(n)$,合成当前语音原型 $s_c(n)$ (长为 L 样本),如 VIII B 节所述。

[0453] 在步骤 2106 更新 LPC 滤波器模块 1620 更新 LPC 滤波器存储器,如步骤 1808 所述。

[0454] 在步骤 2108,对准与内插模块 1618 在前一与当前原型周期之间重建语音样本。前一原型余量 $r_{prev}(n)$ 循环滤波(在 LPC 合成结构中),仅内插可以语音域进行。对准与内插模块 1618 以步骤 1704 的方式操作(见图 19),只是对语音原型而不是对剩余原型操作。对准与内插的结果就是合成的语音信号 $s(n)$ 。

[0455] IX. 噪声激励的线性预测 (NELP) 编码模式

[0456] 噪声激励的线性预测 (NELP) 编码法将语音信号模拟成一个伪随机噪声序列, 由此实现比 CELP 或 PPP 编码法更低的位速率。用信号再现来衡量, NELP 解码的操作最有效, 此时语音信号很少有或没有音调结构, 如非话音或背景噪声。

[0457] 图 22 详细示出了 NELP 编码器模式 204 和 NELP 解码器模式 206, 前者包括能量估算器 2202 和编码代码簿 2204, 后者包括解码代码簿 2206, 随机数发生器 2210, 乘法器 2212 和 LPC 合成滤波器 2208。

[0458] 图 23 是示明 NELP 编码步骤的流程图 2300, 包括编码和解码。这些步骤与 NELP 编 / 解码器模式的各种元件一起讨论。

[0459] 在步骤 2302, 能量估算器 2202 将四个子帧的剩余信号能量都算成:

$$[0460] \quad E_{sf_i} = 0.5 \log_2 \left(\frac{\sum_{n=40i}^{40i+39} s^2(n)}{40} \right), \quad 0 \leq i < 4$$

[0461] 在步骤 2304, 编码代码簿 2204 计算一组代码簿参数, 形成编码的语音信号 $s_{enc}(n)$ 。在一实施例中, 该组代码簿参数包括单个参数, 即标引 I_0 , 它被置成等于 j 值, 并将

$$[0462] \quad \sum_{i=0}^3 (E_{sf_i} - SFEQ(j, i))^2 \text{ 其中 } 0 \leq j < 128$$

[0463] 减至最小。代码簿矢量 SFEQ 用于量化子帧能量 E_{sf_i} , 并包括等于帧内子帧数的元数 (在实施例中为 4)。这些代码簿矢量最好按技术人员已知的普通技术产生, 用于建立随机或训练的代码簿。

[0464] 在步骤 2306, 解码代码簿 2206 对收到的代码簿参数解码。在一实施例中, 按下式解码该组子帧增益 G_i :

$$[0465] \quad G_i = 2^{SFEQ(10, i)}, \text{ 或}$$

$$[0466] \quad G_i = 2^{0.2SFEQ(10, i) + 0.2 \log_2 G_{prev} - 2} \text{ (用零速率编码方案对前一帧编码) 其中 } 0 \leq i < 4,$$

G_{prev} 是代码簿激励增益, 对应于前一帧的最后一个子帧。

[0467] 在步骤 2308, 随机数发生器 2210 产生一单位变化随机矢量 $nz(n)$, 该矢量在步骤 2310 按各子帧内合适的增益 G_i 标定, 建立激励信号 $G_i nz(n)$ 。

[0468] 在步骤 2312, LPC 合成滤波器 2208 对激励信号 $G_i nz(n)$ 滤波, 形成输出语音信号 $\hat{s}(n)$ 。

[0469] 在一实施例中, 也应用了零速率模式, 其中对当前帧的各子帧使用了从最近非零速率 NWLP 子帧获得的增益 G , 与 LPC 参数。技术人员应明白, 在连续出现多个 NELP 帧时, 可有效地应用这种零速率模式。

[0470] X. 结论

[0471] 虽然以上描述了本发明的各种实施例, 但应明白, 这些都是示例, 不用来作限制, 因此, 本发明的范围不受上述任一示例性实施例限制, 仅由所附的权利要求及其等效物限定。

[0472] 上述诸较佳实施例的说明可供任何技术人员用于制作或应用本发明。尽管参照诸较佳实施例具体示出并描述了本发明, 但是技术人员应明白, 在不违背本发明的精神与范围的情况下, 可在形式上和细节上作出各种变化。

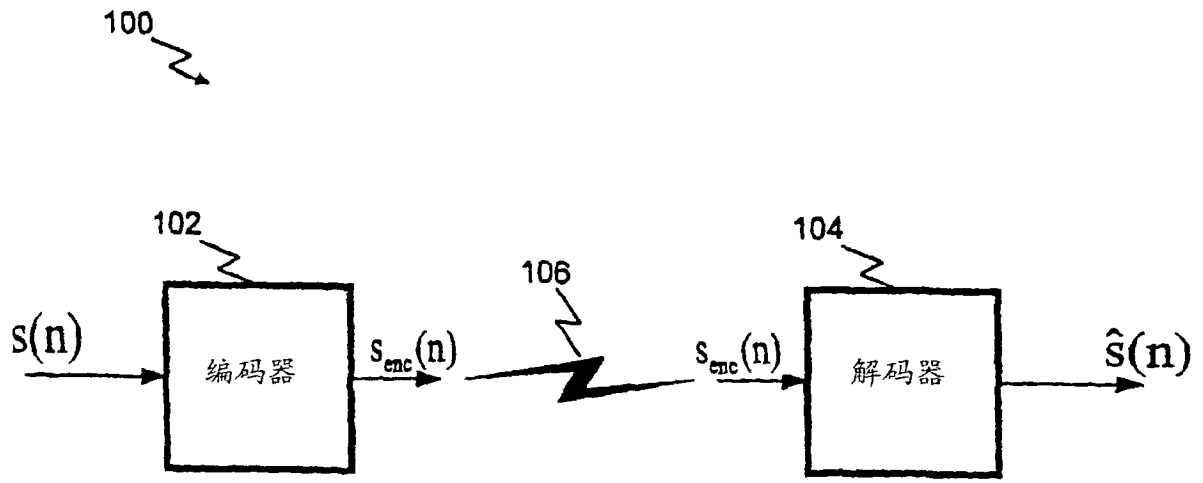


图 1

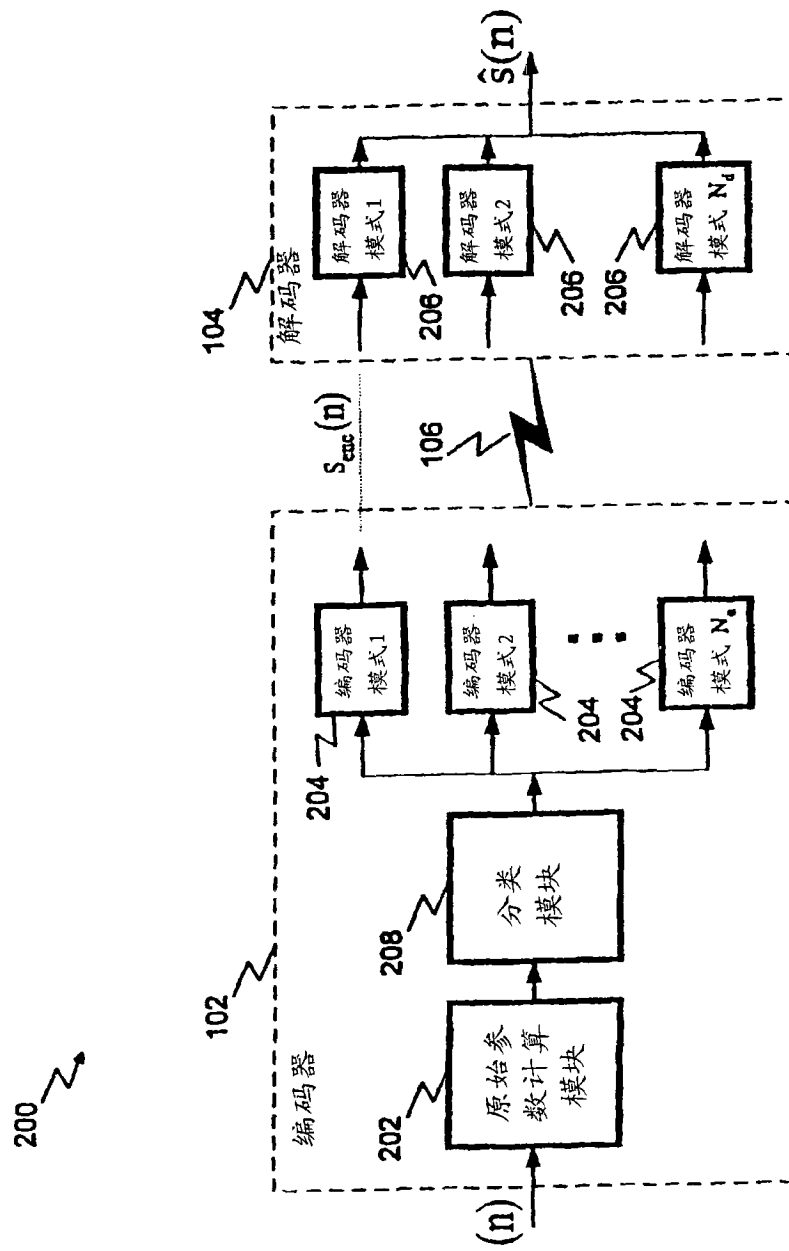


图 2

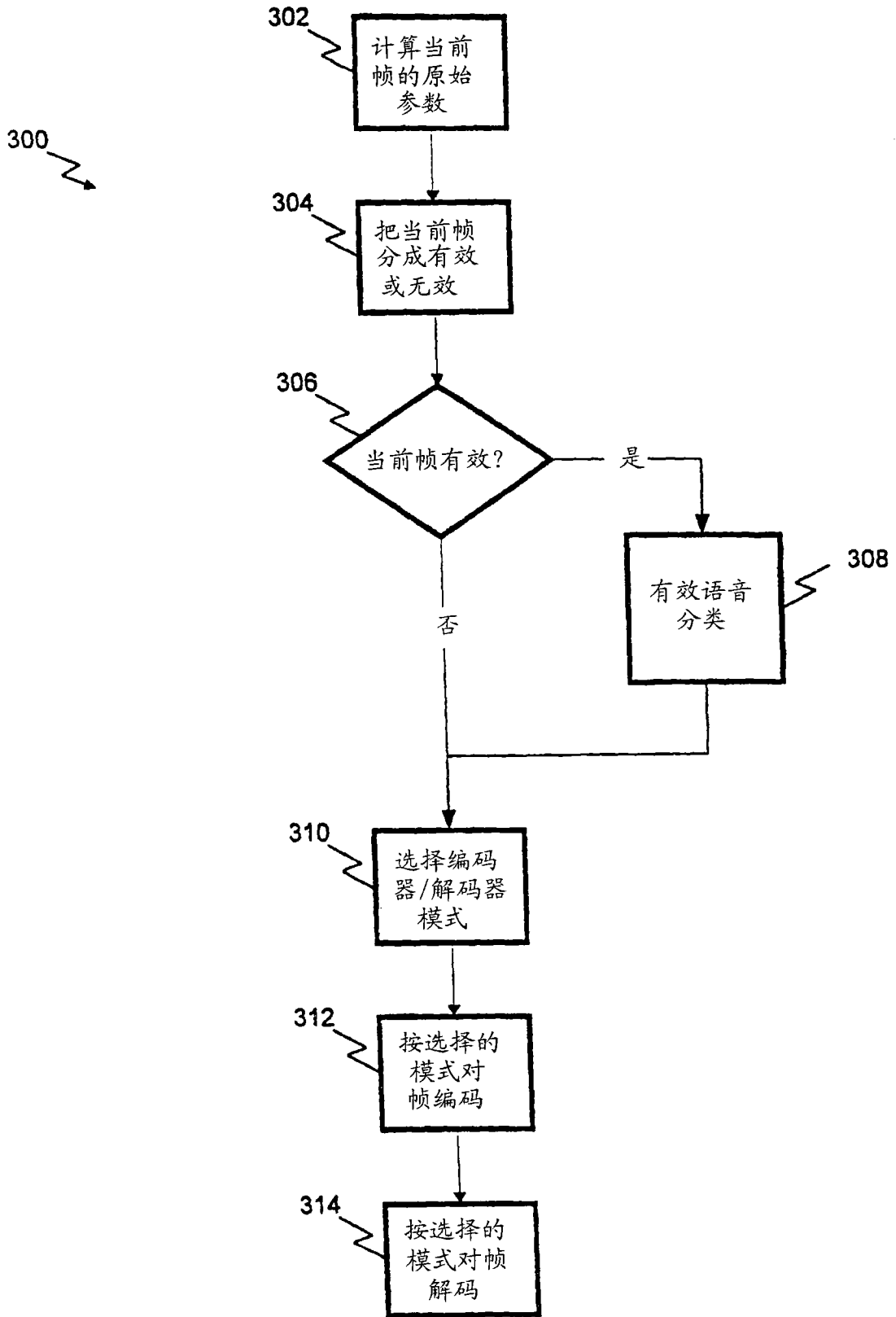


图 3

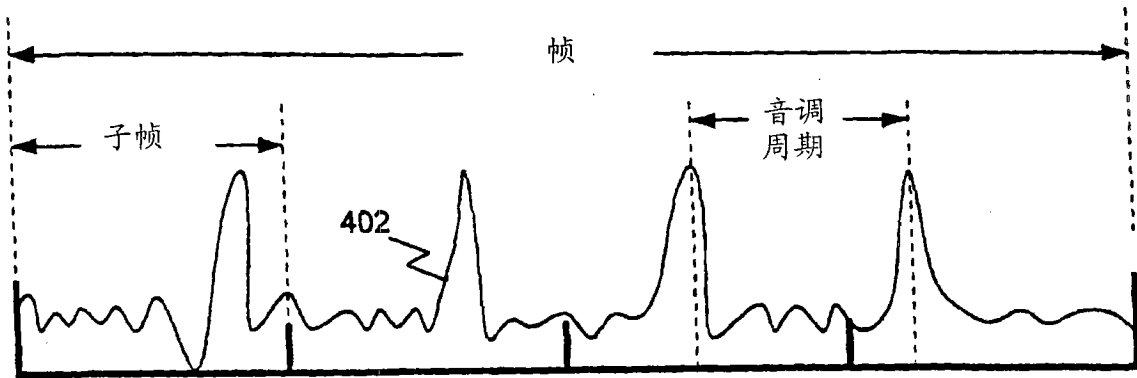


图 4A

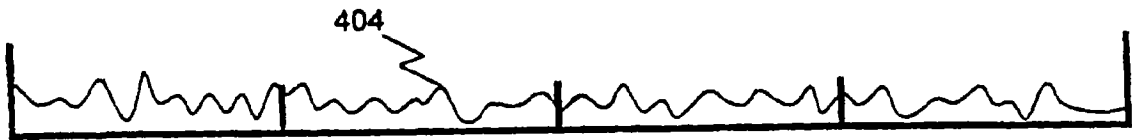


图 4B

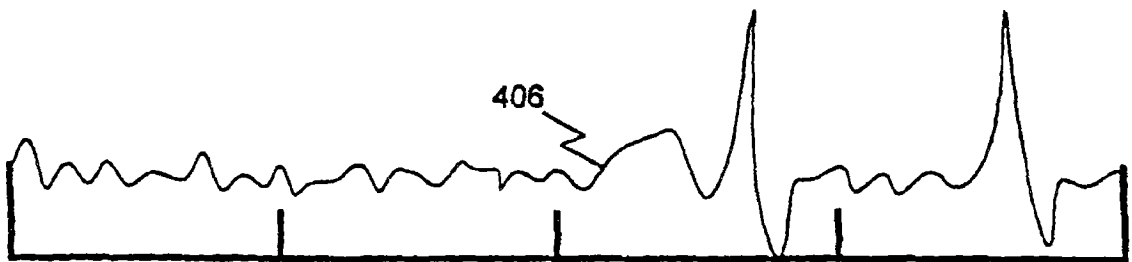


图 4C

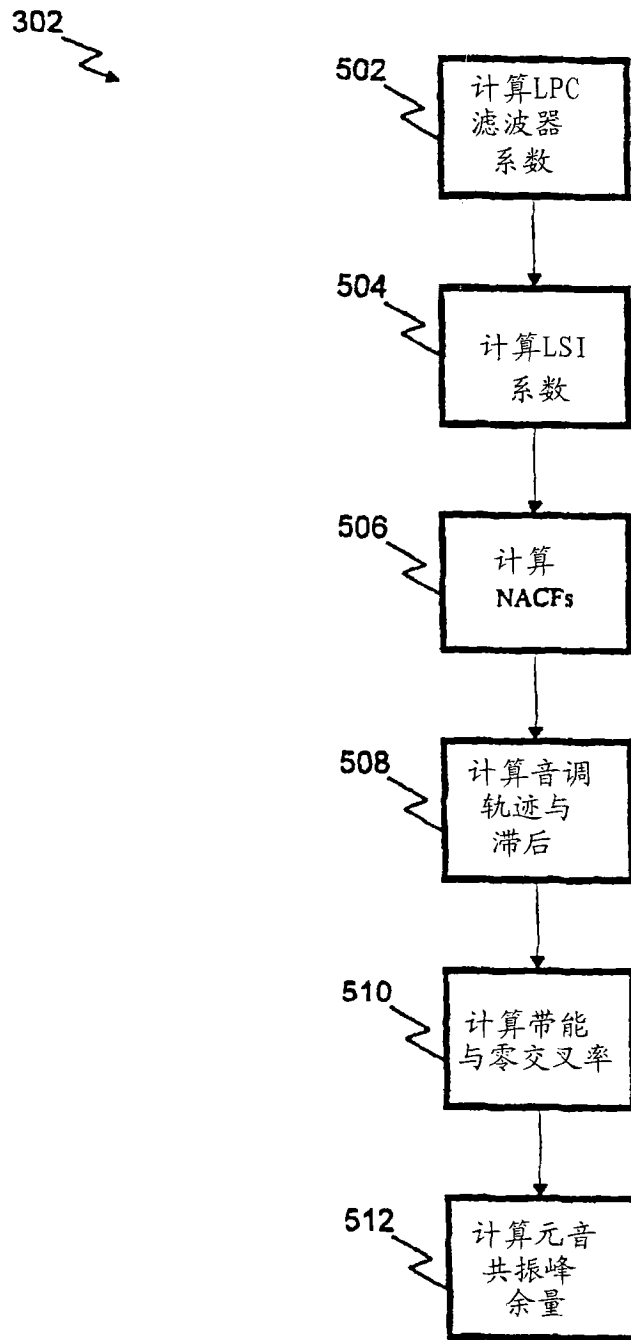


图 5

304

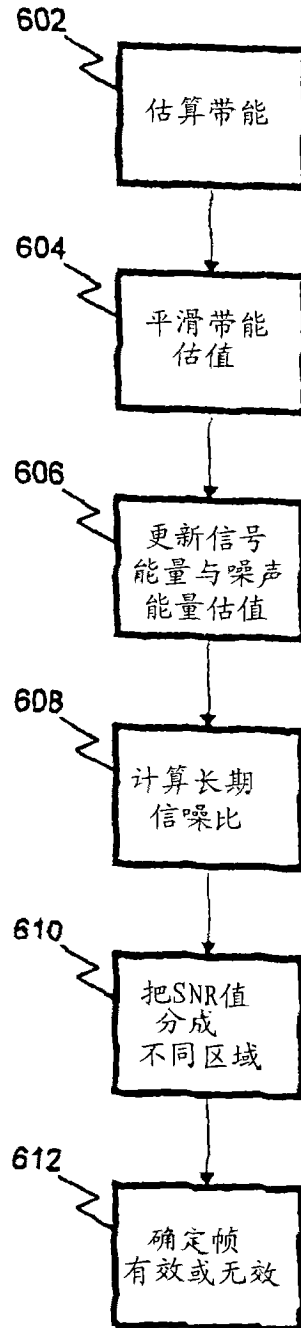


图 6

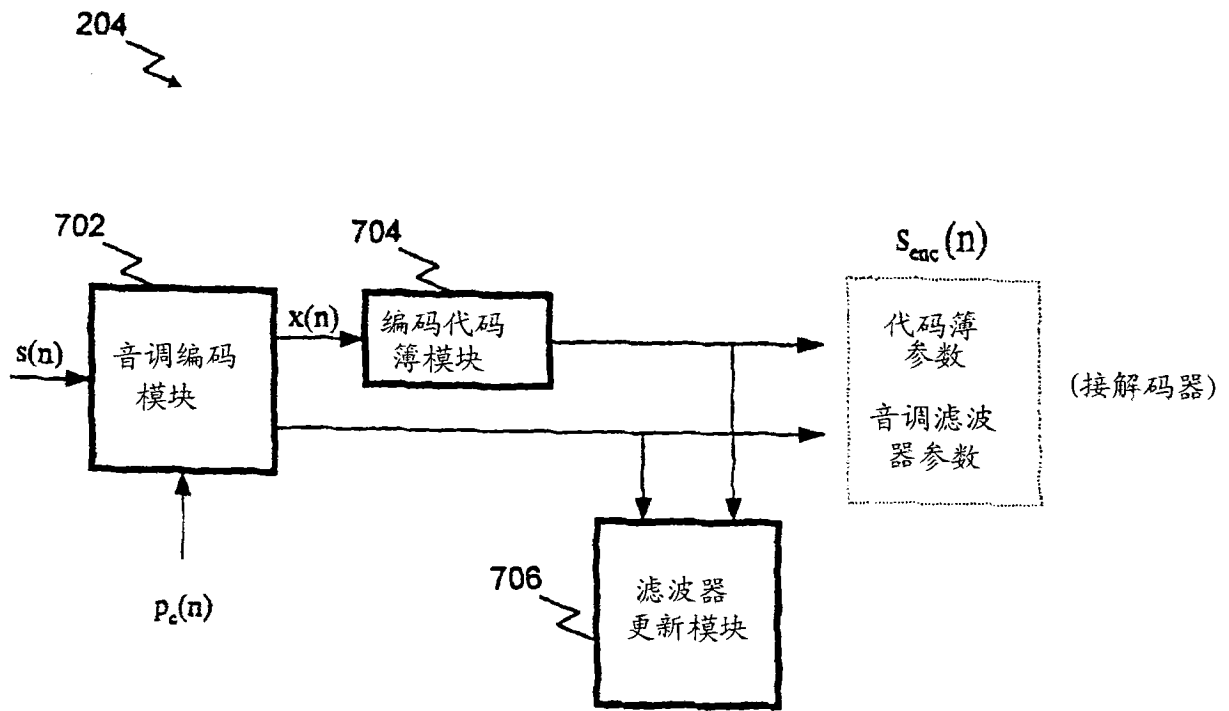


图 7A

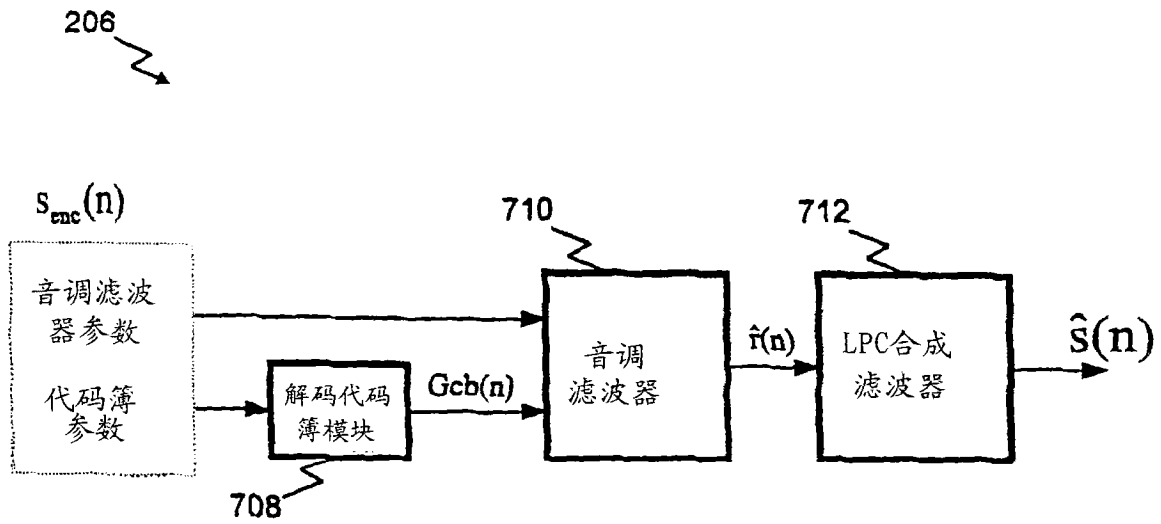


图 7B

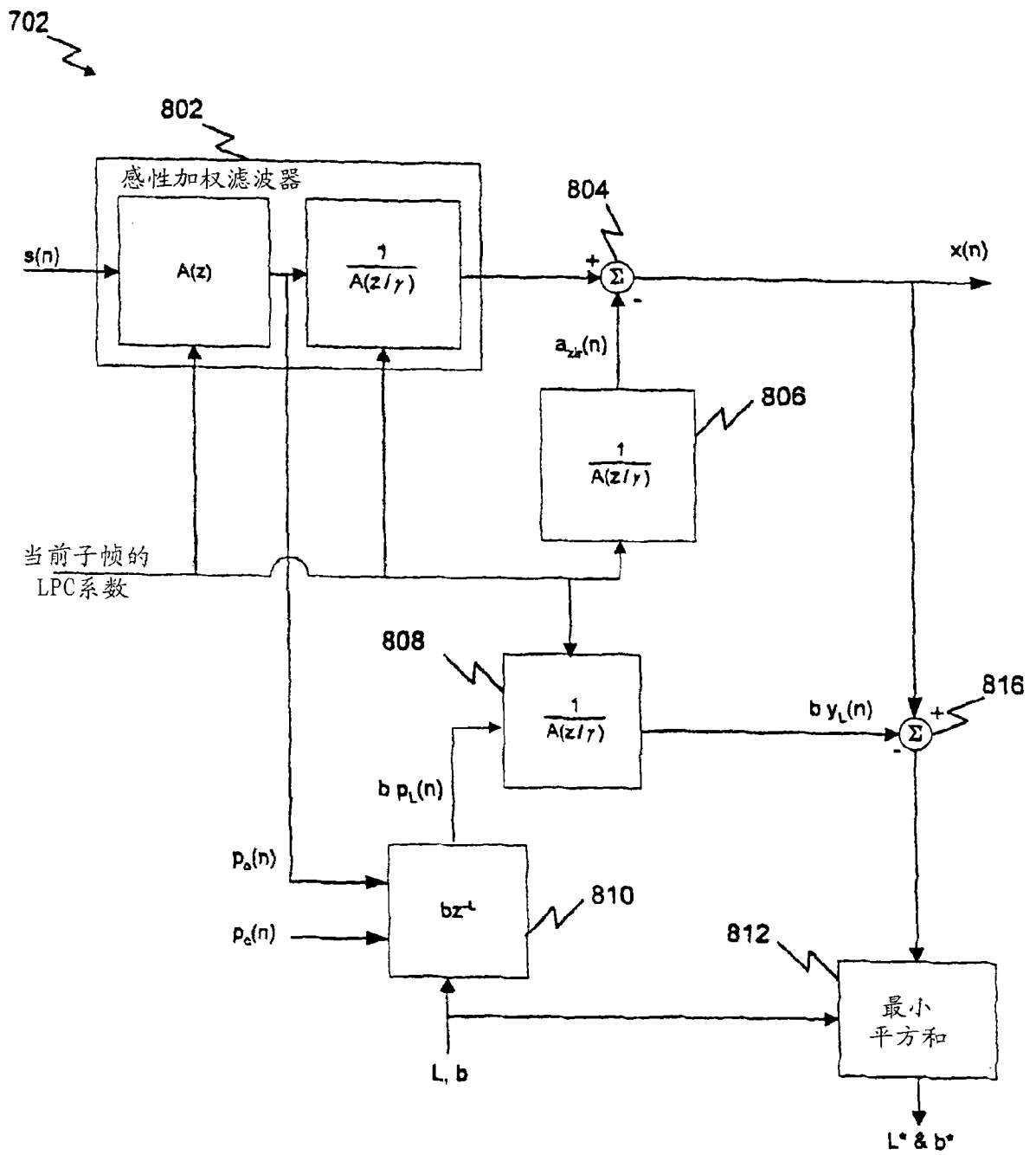


图 8

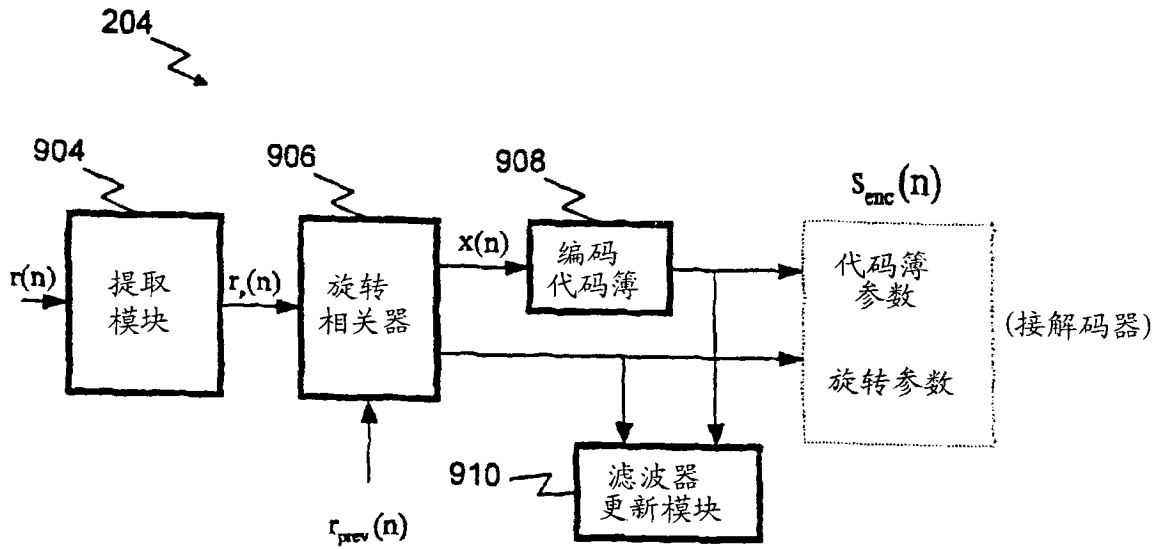


图 9A

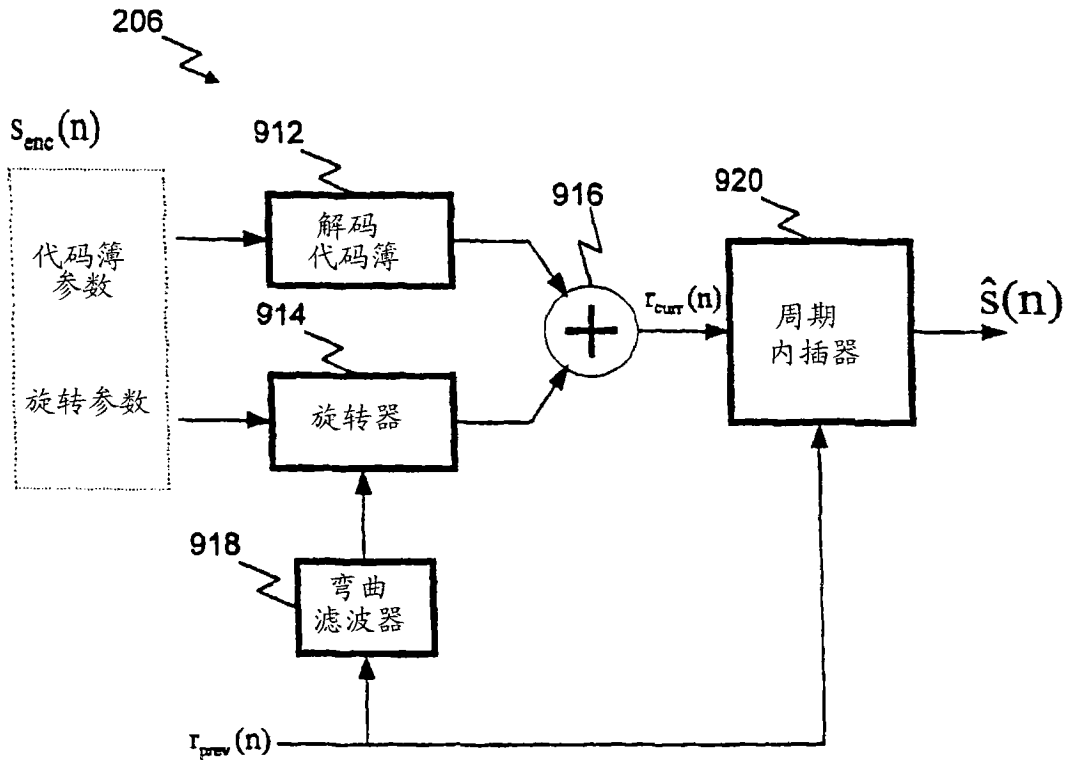


图 9B

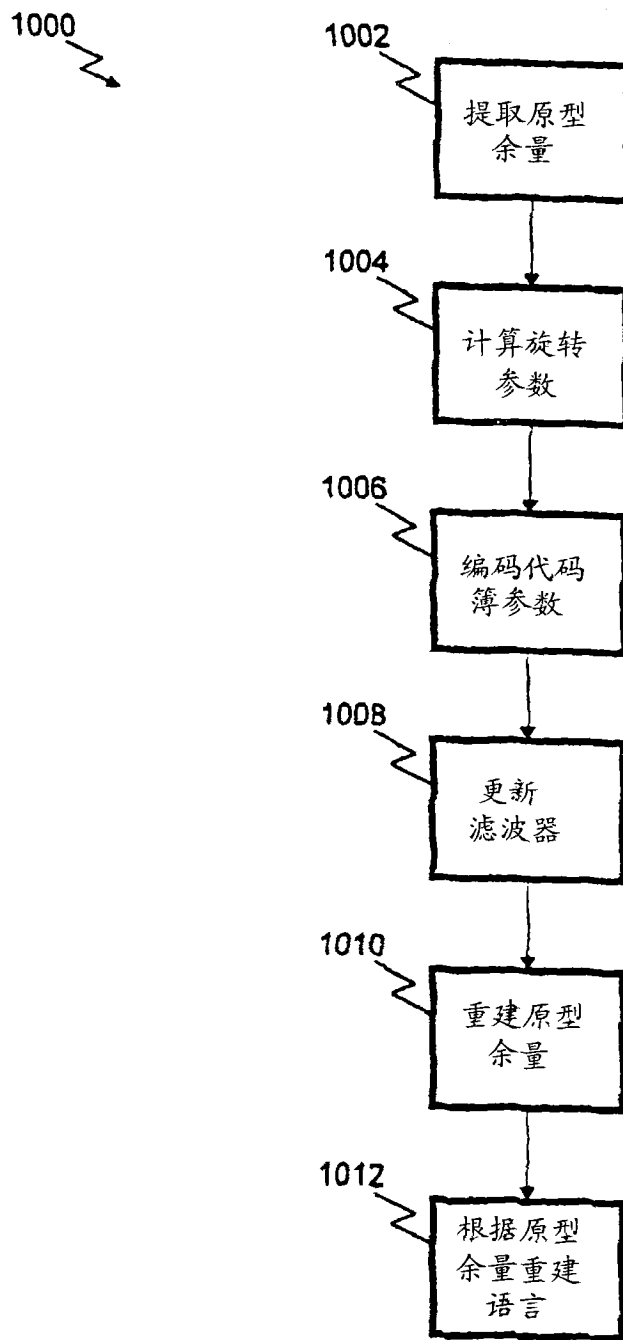


图 10

1002

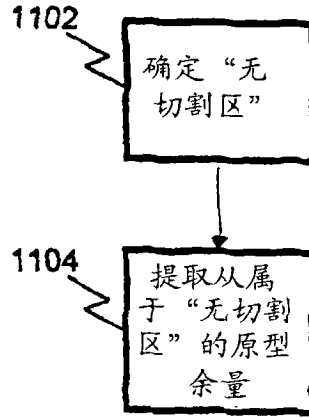


图 11

1200

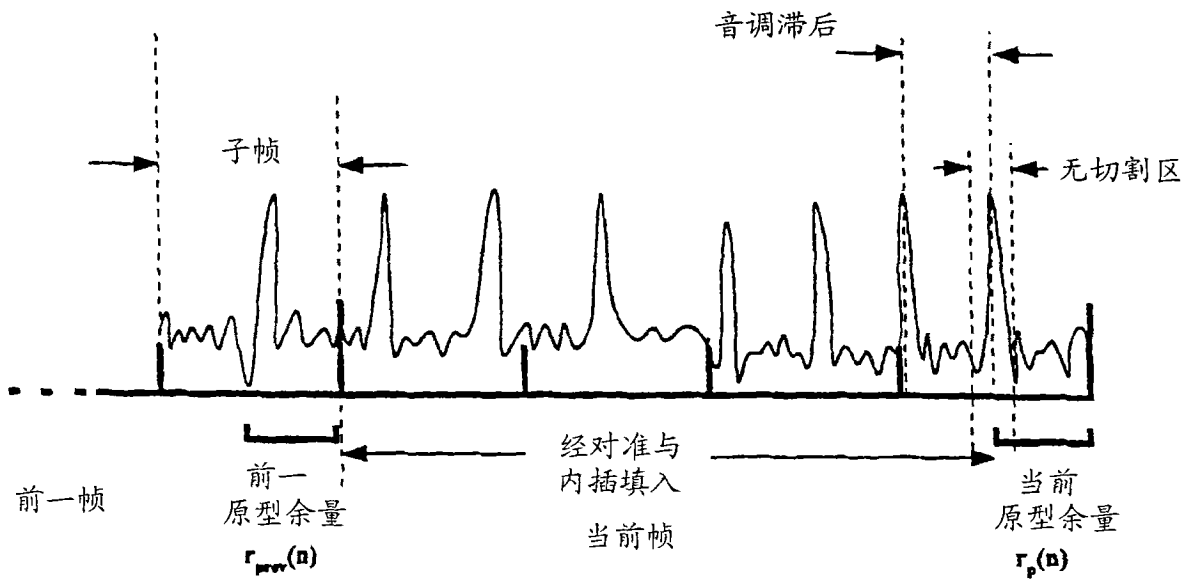


图 12

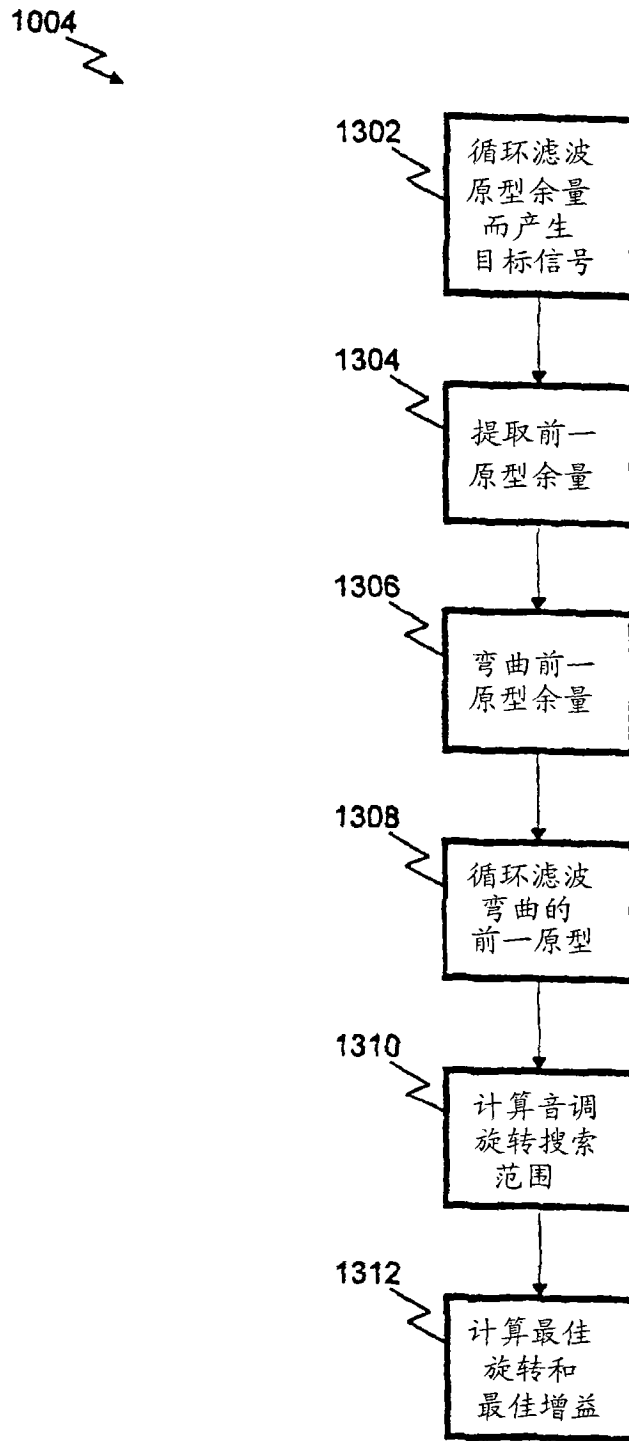


图 13

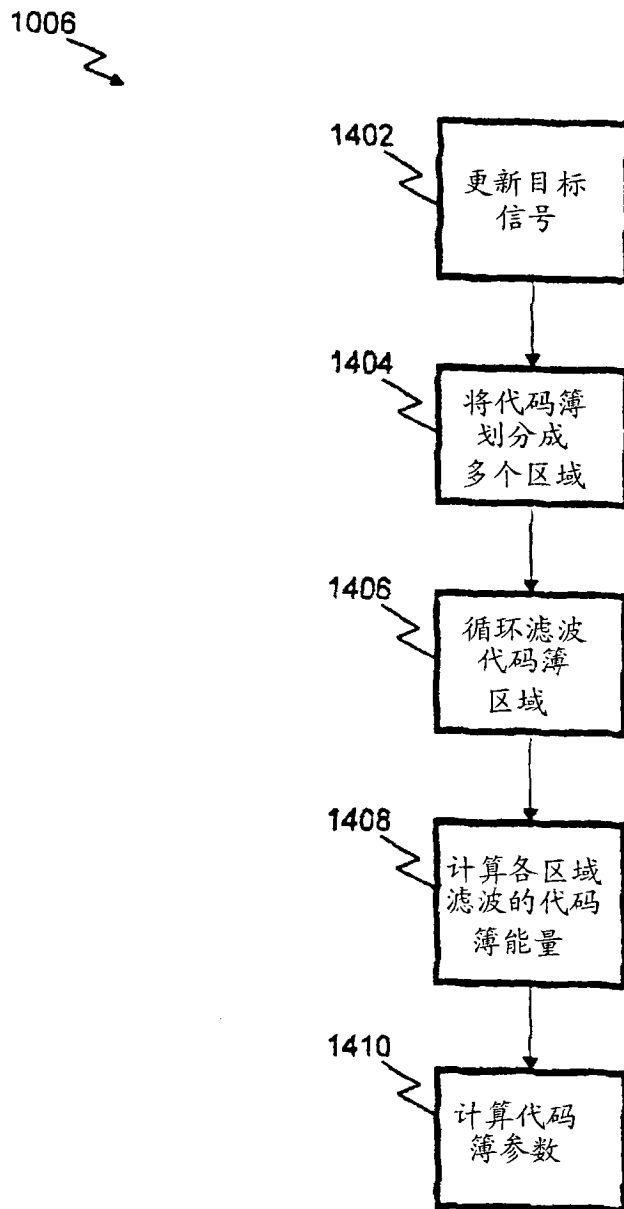


图 14

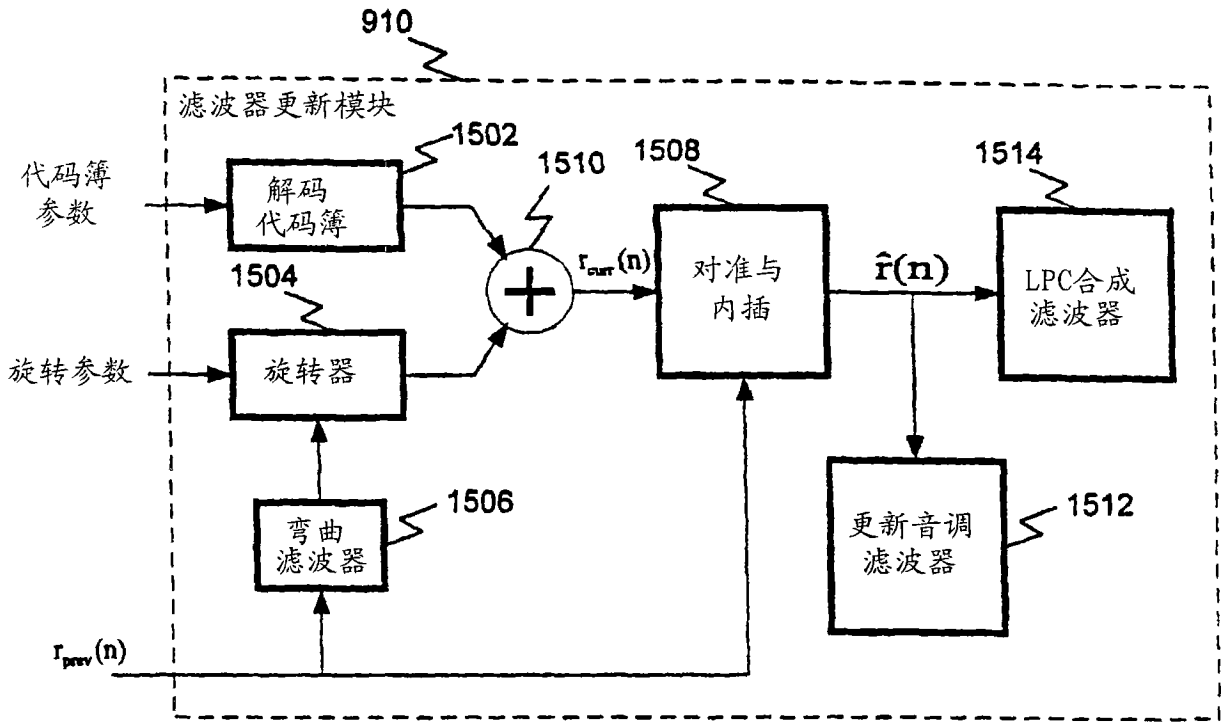


图 15A

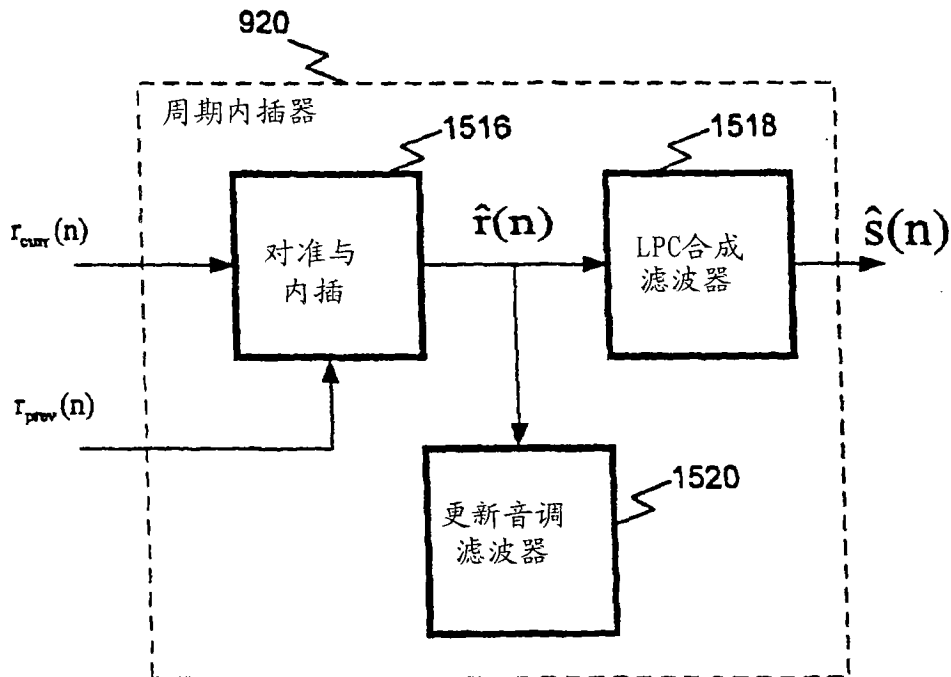


图 15B

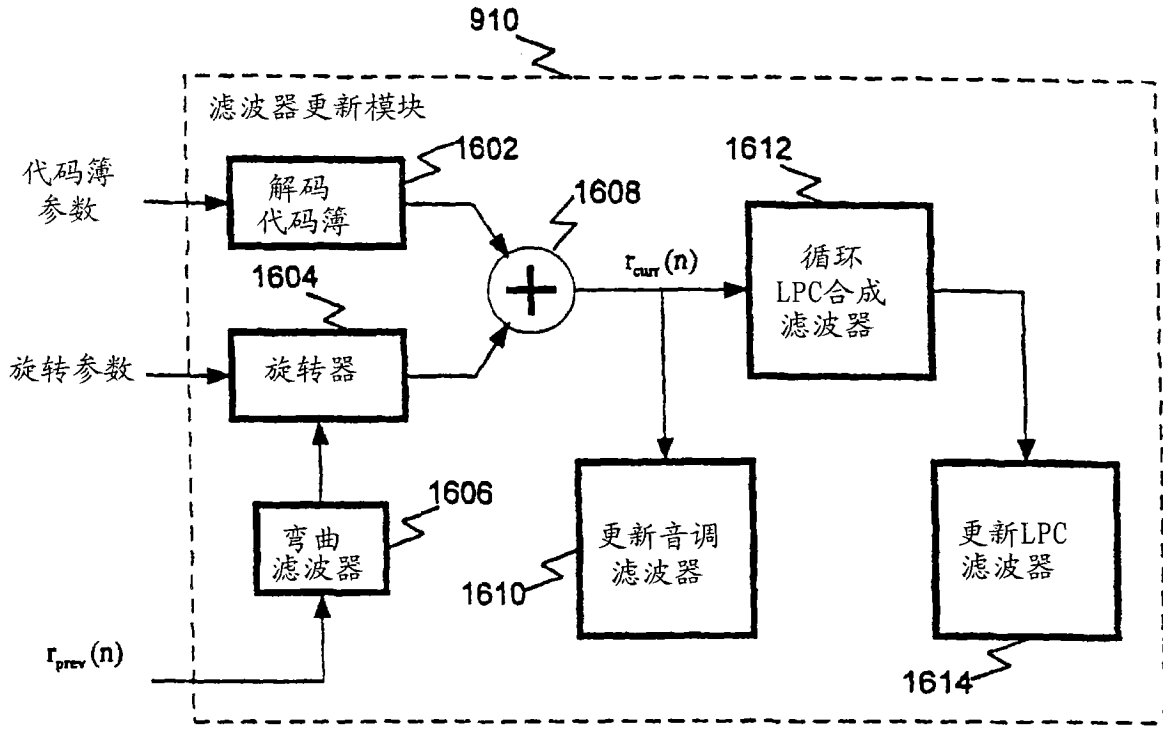


图 16A

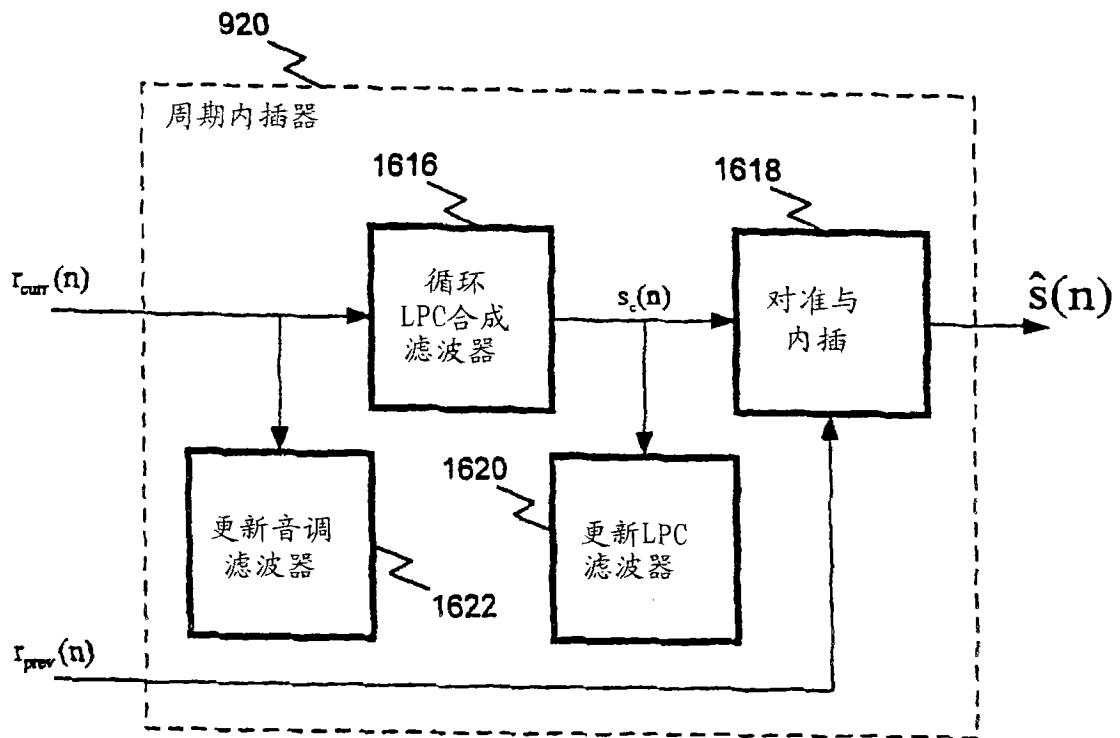


图 16B

1008 ↘

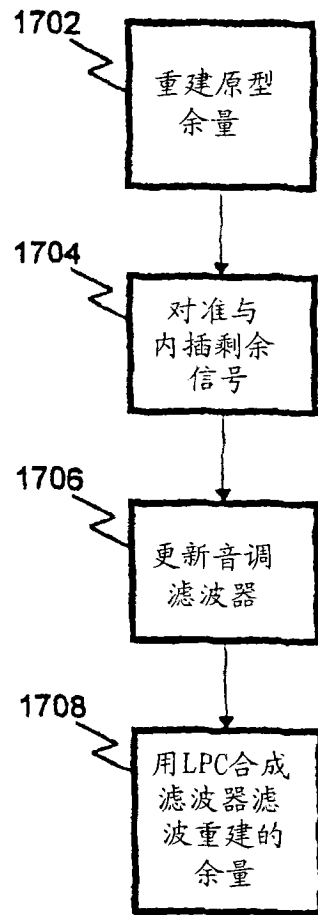


图 17

1008 ↘

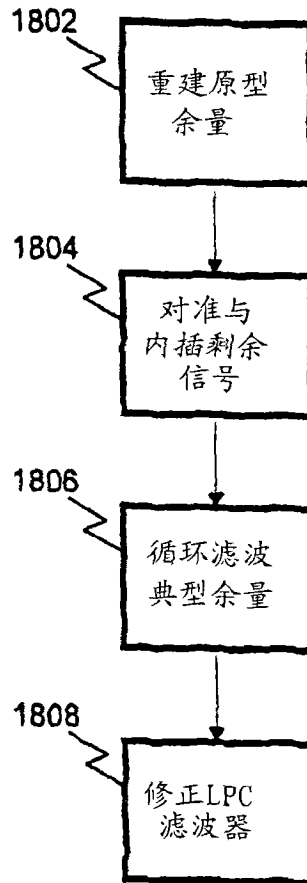


图 18

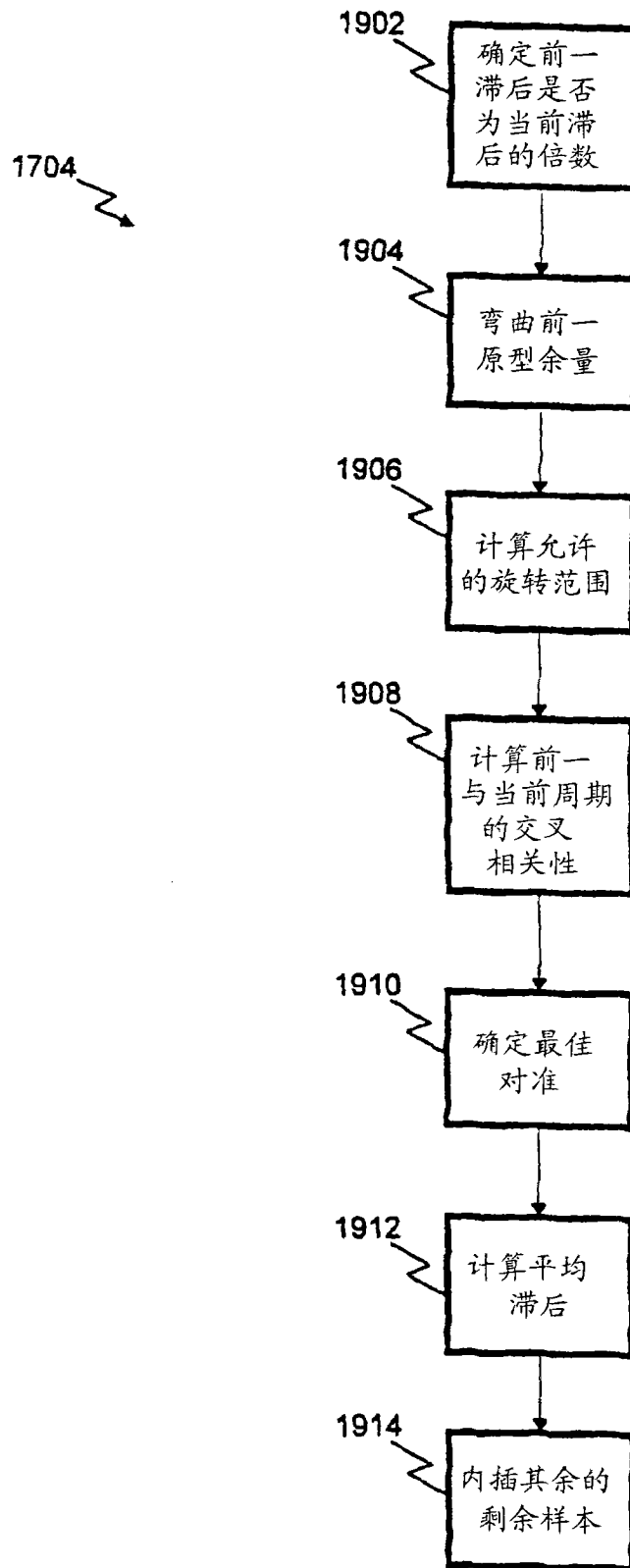


图 19

1012 ↘

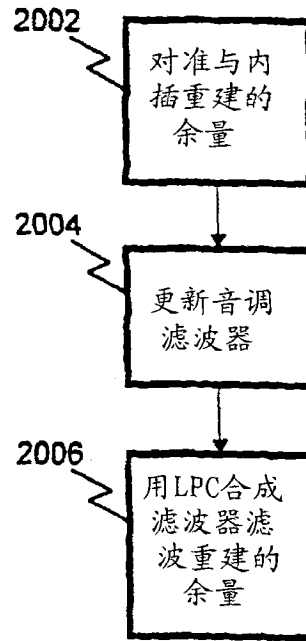


图 20

1012

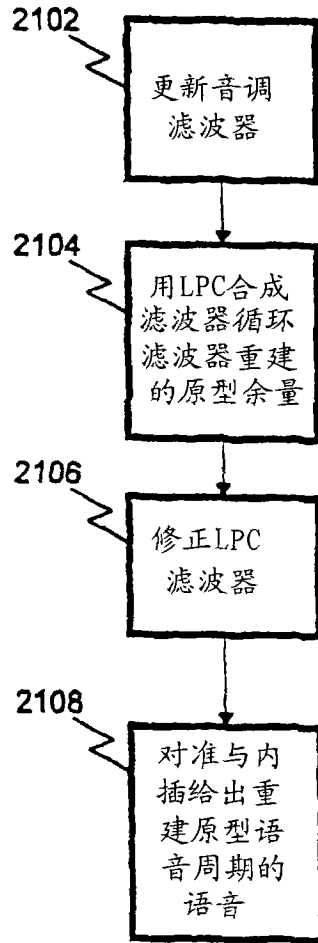


图 21

204

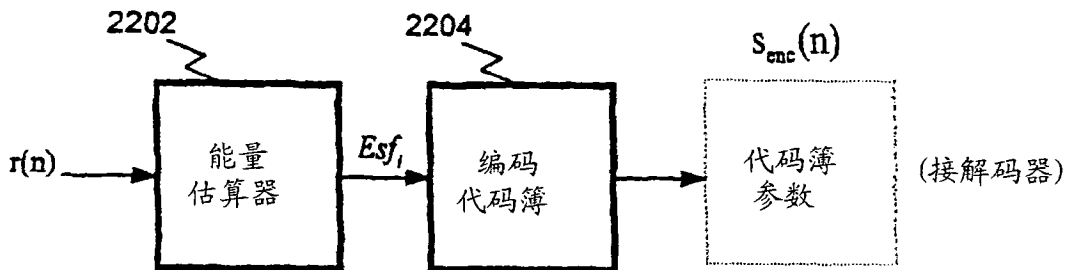


图 22A

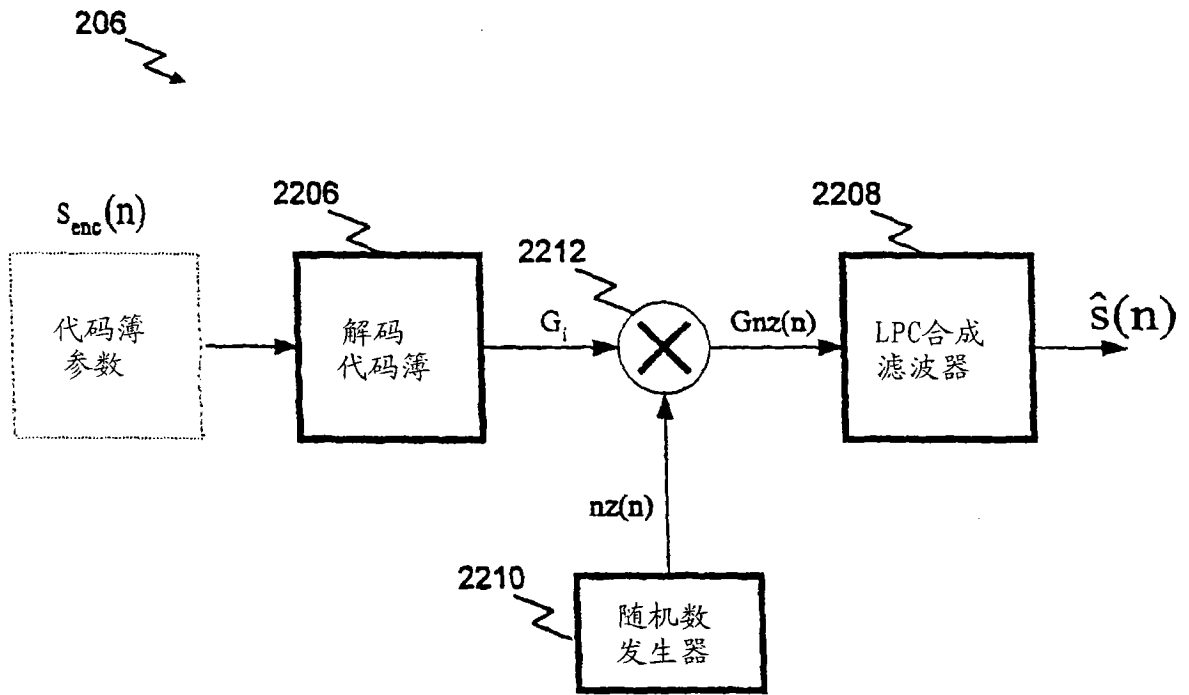


图 22B

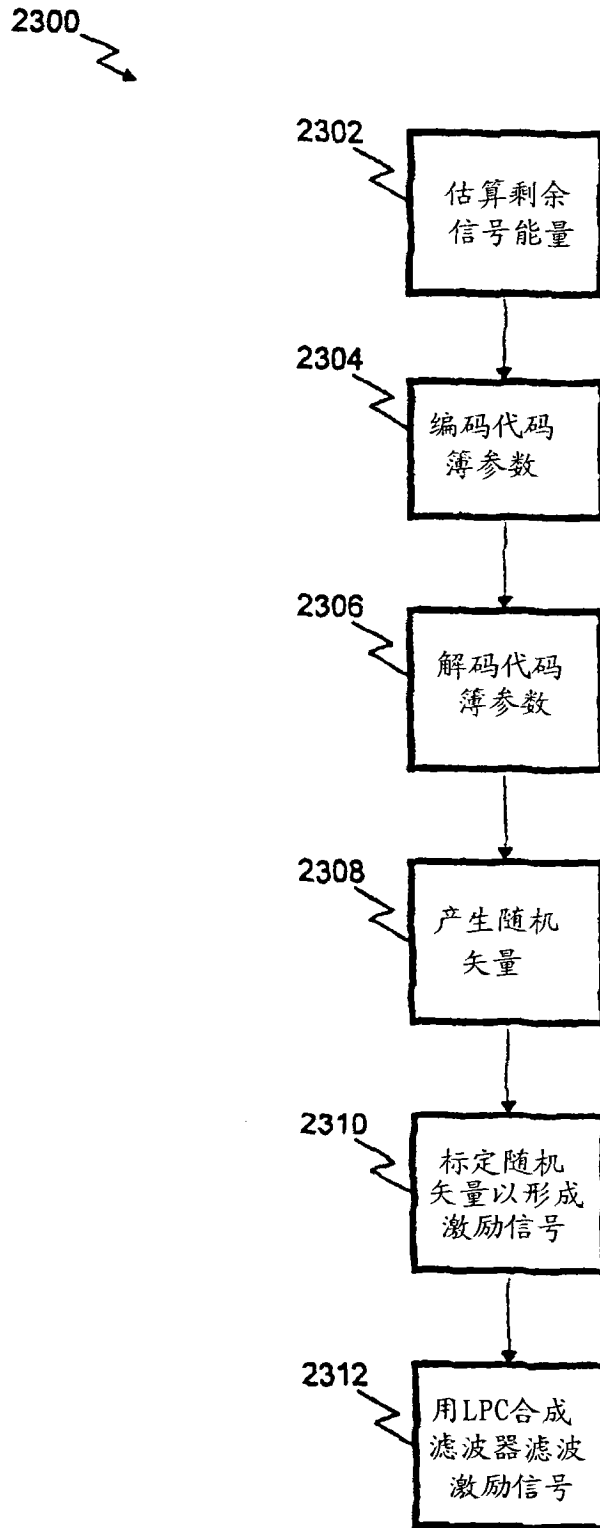


图 23