



## [12] 发 明 专 利 说 明 书

[21] ZL 专利号 00810352.6

[45] 授权公告日 2004 年 4 月 14 日

[11] 授权公告号 CN 1145930C

[22] 申请日 2000.7.19 [21] 申请号 00810352.6

[30] 优先权

[32] 1999.7.19 [33] US [31] 09/356,755

[86] 国际申请 PCT/US2000/019672 2000.7.19

[87] 国际公布 WO01/06495 英 2001.1.25

[85] 进入国家阶段日期 2002.1.14

[71] 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 A·K·阿南塔帕德玛那伯汉

S·曼朱那什

审查员 杨 叁

[74] 专利代理机构 上海专利商标事务所

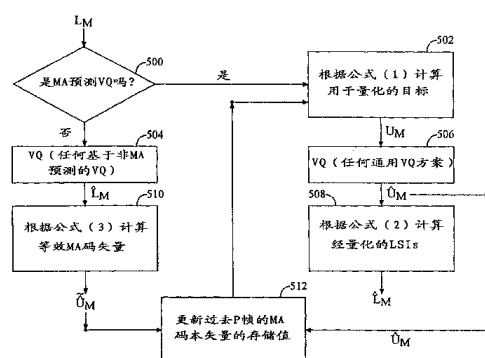
代理人 沈昭坤

权利要求书 4 页 说明书 12 页 附图 6 页

[54] 发明名称 用于交织语音编码器中线性谱信息  
量化方法的方法和设备

## [57] 摘要

一种用于交织语音编码器中线性谱信息量化方法的方法和设备包括用两种矢量量化技术对线性谱信息进行量化, 第一技术是基于非移动平均预测的技术, 而第二技术是基于移动平均预测的技术。用第一技术对线性谱信息矢量进行矢量量化。计算用于第一技术的等效移动平均码矢量。用等效移动平均码矢量以经语音编码器预先处理的预定帧数更新码矢量移动平均码本的存储值。依据已更新的移动平均码本存储值来计算用于第二技术的目标量化矢量。用第二技术对目标量化矢量进行矢量量化来产生量化的目标码矢量。用已量化的目标码矢量来更新移动平均码本的存储值。从已量化的目标码矢量中导出量化线性谱信息矢量。



1、一种语音编码器，包括：

线性预测滤波器，配置为用于分析帧并依据分析生成线性谱信息码矢量；和  
与所述线性预测滤波器耦合的量化器，配置为用于通过使用基于非移动平均预测矢量量化方案的第一矢量量化技术来对线性谱信息矢量进行矢量量化，

其特征在于，所述量化器进一步配置为用来计算用于第一技术的等效移动平均码矢量，用所述等效移动平均码矢量对经语音编码器预先处理的预定帧数的码矢量移动平均码本存储值进行更新，依据已更新的所述移动平均码本存储值计算用于第二技术的目标量化矢量，通过所述第二矢量量化技术对目标量化矢量进行量化来生成经量化的目标码矢量，所述第二矢量量化技术是使用基于移动平均预测的方案，用所述经量化的目标码矢量对所述移动平均码本存储值进行更新，并从所述经量化的目标码矢量中计算经量化的线性谱信息矢量。

2、如权利要求 1 所述的语音编码器，其特征在于，所述帧是语音帧。

3、如权利要求 1 所述的语音编码器，其特征在于，所述帧是线性预测剩余帧。

4、如权利要求 1 所述的语音编码器，其特征在于，所述目标量化矢量是根据下述公式进行计算：

$$\mathbf{U}_M \equiv \left\{ \mathbf{U}_M^n = \frac{(\mathbf{L}_M^n - \alpha_1^n \hat{\mathbf{U}}_{M-1}^n - \alpha_2^n \hat{\mathbf{U}}_{M-2}^n - \dots - \alpha_P^n \hat{\mathbf{U}}_{M-P}^n)}{\alpha_0^n}; n = 0, 1, \dots, N-1 \right\},$$

其中  $\{\hat{\mathbf{U}}_{M-1}^n, \hat{\mathbf{U}}_{M-2}^n, \dots, \hat{\mathbf{U}}_{M-P}^n; n=0, 1, \dots, N-1\}$  是对应于紧接在帧之前已处理的预定数目帧的线性谱信息参量的码本表项，而  $\{\alpha_1^n, \alpha_2^n, \dots, \alpha_P^n; n=0, 1, \dots, N-1\}$  是各参数权重，这样  $\{\alpha_0^n + \alpha_1^n + \dots + \alpha_P^n = 1; n=0, 1, \dots, N-1\}$ 。

5、如权利要求 1 所述的语音编码器，其特征在于，所述经量化线性谱信息矢量是根据下述公式进行计算：

$$\hat{\mathbf{L}}_M \equiv \{\hat{\mathbf{L}}_M^n = \alpha_0^n \mathbf{U}_M^n + \alpha_1^n \hat{\mathbf{U}}_{M-1}^n + \dots + \alpha_P^n \hat{\mathbf{U}}_{M-P}^n; n = 0, 1, \dots, N-1\},$$

其中  $\{\hat{\mathbf{U}}_{M-1}^n, \hat{\mathbf{U}}_{M-2}^n, \dots, \hat{\mathbf{U}}_{M-P}^n; n=0, 1, \dots, N-1\}$  是对应于紧接在帧之前已处理的预定数目帧的线性谱信息参量的码本表项，而  $\{\alpha_1^n, \alpha_2^n, \dots, \alpha_P^n; n=0, 1, \dots, N-1\}$  是各参量权重，这样  $\{\alpha_0^n + \alpha_1^n + \dots + \alpha_P^n = 1; n=0, 1, \dots, N-1\}$ 。

$N-1\}$ 。

6、如权利要求 1 所述的语音编码器，其特征在于，所述等效移动平均码矢量是根据下述公式进行计算：

$$\tilde{U}_{M-K} \equiv \left\{ \tilde{U}_{M-K}^n = \frac{(\hat{L}_{M-K}^n - \beta_1^n \hat{U}_{M-K-1}^n - \beta_2^n \hat{U}_{M-K-2}^n - \dots - \beta_P^n \hat{U}_{M-K-P}^n)}{\beta_0^n}; n = 0, 1, \dots, N-1 \right\}$$

其中  $\{\beta_1^n, \beta_2^n, \dots, \beta_P^n; n=0, 1, \dots, N-1\}$  是各等效移动平均码矢量单元权重使得  $\{\beta_0^n + \beta_1^n + \dots + \beta_P^n = 1; n=0, 1, \dots, N-1\}$ ，并且其中初始条件  $\{\tilde{U}_{-1}, \tilde{U}_{-2}, \dots, \tilde{U}_{-P}\}$  已确立。

7、如权利要求 1 所述的语音编码器，其特征在于，所述语音编码器驻留在无线通信系统用户单元中。

8、一种对帧的线性谱信息矢量进行矢量量化的方法，使用第一和第二量化矢量量化技术，第一技术使用基于非移动平均预测矢量量化方案，第二技术使用基于移动平均预测矢量量化方案，其特征在于，该方法包括下述步骤：

用所述第一矢量量化技术对线性谱信息矢量进行矢量量化；

计算用于所述第一技术的等效移动平均码矢量；

用所述等效移动平均码矢量更新经语音编码器预先处理的预定帧数的码矢量移动平均码本的存储值；

依据所述已更新的移动平均码本的存储值来计算用于所述第二技术的目标量化矢量；

用所述第二矢量量化技术对目标量化矢量进行矢量量化来产生量化的目标码矢量；

用所述已量化的目标码矢量来更新所述移动平均码本的存储值；和

从所述已量化的目标码矢量中导出量化线性谱信息矢量。

9、如权利要求 8 所述的方法，其特征在于，所述帧是语音帧。

10、如权利要求 8 所述的方法，其特征在于，所述帧是线性预测剩余帧。

11、如权利要求 8 所述的方法，其特征在于，所述计算步骤包括根据下述公式计算所述目标量化矢量：

$$U_M \equiv \left\{ U_M^n = \frac{(\hat{L}_M^n - \alpha_1^n \hat{U}_{M-1}^n - \alpha_2^n \hat{U}_{M-2}^n - \dots - \alpha_P^n \hat{U}_{M-P}^n)}{\alpha_0^n}; n = 0, 1, \dots, N-1 \right\},$$

其中  $\{\hat{U}_{M-1}^n, \hat{U}_{M-2}^n, \dots, \hat{U}_{M-P}^n; n=0, 1, \dots, N-1\}$  是对应于紧接在帧之前已处理的预定数目帧的线性谱信息参数的码本表项，而  $\{\alpha_1^n, \alpha_2^n, \dots, \alpha_P^n;$

$n=0, 1, \dots, N-1$  是各参数的权重, 使得  $\{\alpha_0^n + \alpha_1^n + \dots + \alpha_p^n = 1; n=0, 1, \dots, N-1\}$ 。

12、如权利要求 8 所述的方法, 其特征在于, 所述导出步骤包括根据下述公式导出所述经量化线性谱信息矢量:

$$\hat{L}_M = \{\hat{L}_M^n = \alpha_0^n \hat{U}_M^n + \alpha_1^n \hat{U}_{M-1}^n + \dots + \alpha_p^n \hat{U}_{M-p}^n; n = 0, 1, \dots, N-1\},$$

其中  $\{\hat{U}_{M-1}^n, \hat{U}_{M-2}^n, \dots, \hat{U}_{M-p}^n; n=0, 1, \dots, N-1\}$  是对应于紧接在帧之前已处理的预定数目帧的线性谱信息参量的码本表项, 而  $\{\alpha_1^n, \alpha_2^n, \dots, \alpha_p^n; n=0, 1, \dots, N-1\}$  是各参数权重, 这样  $\{\alpha_0^n + \alpha_1^n + \dots + \alpha_p^n = 1; n=0, 1, \dots, N-1\}$ 。

13、如权利要求 8 所述的方法, 其特征在于, 所述计算步骤包括根据下述公式计算所述等效移动平均码矢量:

$$\tilde{U}_{M-K} = \left\{ \tilde{U}_{M-K}^n = \frac{(\hat{L}_{M-K}^n - \beta_1^n \hat{U}_{M-K-1}^n - \beta_2^n \hat{U}_{M-K-2}^n - \dots - \beta_p^n \hat{U}_{M-K-p}^n)}{\beta_0^n}; n = 0, 1, \dots, N-1 \right\}$$

其中  $\{\beta_1^n, \beta_2^n, \dots, \beta_p^n; n=0, 1, \dots, N-1\}$  是各等效移动平均码矢量单元权重使得  $\{\beta_0^n + \beta_1^n + \dots + \beta_p^n = 1; n=0, 1, \dots, N-1\}$ , 并且其中初始条件  $\{\tilde{U}_{-1}^n, \tilde{U}_{-2}^n, \dots, \tilde{U}_{-p}^n\}$  已确立。

14、一种语音编码器, 其特征在于, 包括:

用于通过用第一矢量量化技术对线性谱信息矢量进行矢量量化的装置, 所述技术使用基于非移动平均预测矢量量化方案;

用于计算用于所述第一技术的等效移动平均码矢量的装置;

用于用所述等效移动平均码矢量更新经语音编码器预先处理的预定帧数的码矢量移动平均码本存储值的装置;

用于依据所述已更新的移动平均码本存储值来计算用于第二技术的目标量化矢量的装置;

用于用所述第二矢量量化技术对所述目标量化矢量进行矢量量化来产生量化的目标码矢量的装置;

用于用所述已量化的目标码矢量来更新所述移动平均码本的存储值的装置;  
和

用于从所述已量化的目标码矢量中导出量化线性谱信息矢量的装置。

15、如权利要求 14 所述的语音编码器, 其特征在于, 所述帧是语音帧。

16、如权利要求 14 所述的语音编码器，其特征在于，所述帧是线性预测剩余帧。

17、如权利要求 14 所述的语音编码器，其特征在于，所述目标量化矢量是根据下述公式进行计算：

$$U_M \equiv \left\{ U_M^n = \frac{(L_M^n - \alpha_1^n \hat{U}_{M-1}^n - \alpha_2^n \hat{U}_{M-2}^n - \dots - \alpha_P^n \hat{U}_{M-P}^n)}{\alpha_0^n}; n = 0, 1, \dots, N-1 \right\}$$

其中  $\{\hat{U}_{M-1}^n, \hat{U}_{M-2}^n, \dots, \hat{U}_{M-P}^n; n=0, 1, \dots, N-1\}$  是对应于紧接在帧之前已处理的预定数目帧的线性谱信息参数的码本表项，而  $\{\alpha_1^n, \alpha_2^n, \dots, \alpha_P^n; n=0, 1, \dots, N-1\}$  是各参数的权重，使得  $\{\alpha_0^n + \alpha_1^n + \dots + \alpha_P^n = 1; n=0, 1, \dots, N-1\}$ 。

18、如权利要求 14 所述的语音编码器，其特征在于，所述经量化线性谱信息矢量是根据下述公式导出：

$$\hat{L}_M \equiv \{\hat{L}_M^n = \alpha_0^n \hat{U}_M^n + \alpha_1^n \hat{U}_{M-1}^n + \dots + \alpha_P^n \hat{U}_{M-P}^n; n = 0, 1, \dots, N-1\},$$

其中  $\{\hat{U}_{M-1}^n, \hat{U}_{M-2}^n, \dots, \hat{U}_{M-P}^n; n=0, 1, \dots, N-1\}$  是对应于紧接在帧之前已处理的预定数目帧的线性谱信息参数的码本表项，而  $\{\alpha_1^n, \alpha_2^n, \dots, \alpha_P^n; n=0, 1, \dots, N-1\}$  是各参数权重，这样  $\{\alpha_0^n + \alpha_1^n + \dots + \alpha_P^n = 1; n=0, 1, \dots, N-1\}$ 。

19、如权利要求 14 所述的语音编码器，其特征在于，所述等效移动平均码矢量是根据下述公式计算进行计算：

$$\tilde{U}_{M-K} \equiv \left\{ \tilde{U}_{M-K}^n = \frac{(\hat{L}_{M-K}^n - \beta_1^n \hat{U}_{M-K-1}^n - \beta_2^n \hat{U}_{M-K-2}^n - \dots - \beta_P^n \hat{U}_{M-K-P}^n)}{\beta_0^n}; n = 0, 1, \dots, N-1 \right\}$$

其中  $\{\beta_1^n, \beta_2^n, \dots, \beta_P^n; n=0, 1, \dots, N-1\}$  是各等效移动平均码矢量单元权重使得  $\{\beta_0^n + \beta_1^n + \dots + \beta_P^n = 1; n=0, 1, \dots, N-1\}$ ，并且其中初始条件  $\{\tilde{U}_{-1}, \tilde{U}_{-2}, \dots, \tilde{U}_{-P}\}$  已确立。

20、如权利要求 14 所述的语音编码器，其特征在于，所述语音编码器驻留在无线通信系统用户单元中。

## 用于交织语音编码器中线性谱信息量化方法的方法和设备

### 技术领域

本发明通常涉及语音处理领域，并且特别针对用于对语音编码器中的线性谱信息进行量化的方法和设备。

### 背景技术

通过数字技术进行语音传输已经变得很普遍，特别是在长距离和数字无线电话应用中。这反过来又使人们对在信道上所发送的能保持重构语音感知质量的信息最小量的确定产生了兴趣。如果语音是以简单的采样和数字化进行传输，那么就需要大约 64 千位每秒(kbps)的数据率才能达到传统模拟电话的语音质量。然而，通过语音分析的使用，后随合适的编码、传输和在接收器的再合成，可以使数据率明显下降。

用于压缩语音的设备在许多电信领域中都能找到。一个示例领域就是无线通信。无线通信领域具有很多应用包括例如无绳电话、无线电寻呼、无线本地环路、无线电话例如蜂窝或 PCS 电话系统、移动网际协议(IP)电话和卫星通信系统。一种特别重要的应用就是用于移动用户的无线电话。

针对无线通信系统包括例如频分多路访问(FDMA)、时分多路访问(TDMA)和码分多路访问(CDMA)已经开发出各种空中接口。与之连接中，建立了各种国内和国际标准包括例如高级移动电话服务(AMPS)、全球移动通信系统(GSM)和临时标准 95(IS-95)。一种示范无线电话通信系统是码分多路访问(CDMA)系统。IS-95 标准和其衍生物 IS-95A、ANSI J-STD-008, IS-95B、提议的第三代标准 IS-95C 和 IS-2000 等(在此共同归类为 IS-95)是由电信工业协会(TIA)和其他知名标准团体公布来说明用于蜂窝或 PCS 电话通信系统的 CDMA 空中接口的使用。大致根据使用的 IS-95 标准配置的示范无线通信系统在美国专利号 5,103,459 和 4,901,307(已转让给本发明的受让人并在此作为合作参考)中有所描述。

采用以提取与人类语音生成模型有关的参量来压缩语音的技术的设备被称为语音编码器。语音编码器将输入语音信号分为时间块或分析帧。语音编码器通常由编码器和译码器组成。编码器对输入语音帧进行分析来提取某些相关参量，并

且随后将参量量化为二进制码表示，即量化为一组位或二进制数据包。数据包在通信信道上向接收器和解码器传输。解码器对这些数据包进行处理，把它们去量化来产生参量，并且使用去量化参量来再合成语音帧。

语音编码器的功能是通过去除语音中固有的所有自然冗余来将数字化的语音信号压缩为低比特率信号。通过用一组参量代表输入语音帧并对参量进行量化来用一组位表示参量就可以实现数字压缩。如果输入语音帧具有位数为  $N_i$  并且语音编码器产生的数据包具有位数  $N_o$ ，语音编码器所达到的压缩系数为  $C_r = N_i / N_o$ 。在压缩技术中所面临的挑战是在达到目标压缩系数的情况下还要保持解码语音的高语音质量。评价语音编码器的性能的依据是(1)上述语音模型或分析和合成的混合处理完成的效果有多好，以及(2)以目标比特率每帧  $N_o$  位进行参量量化处理所执行的效果如何。语音模型的目标就是对于每帧用较小一组参量来获得语音信号的实质或目标语音质量。

在语音编码器的设计中最重要的是寻找一组好的参数(包括向量)来描述语音信号。一组好的参数需要较低的系统带宽用于感觉上准确的语音信号重构。音调、信号功率、谱包络(或共振峰)、振幅谱和相谱都是语音编码参数的实例。

语音编码器可以作为时域编码器实现，时域编码器是试图通过每次使用高时间分辨率处理对较小的语音段(通常是 5 毫秒(ms)子帧)进行编码来捕获时域语音波形。对于每个子帧，依靠本领域中已知的各种搜索算法从码本空间中寻找高精度的代表。或者，语音编码器可以作为频域编码器来实现，频域编码器是试图用一组参量(分析)来捕获输入语音帧的短期语音频谱，并且使用相应的合成处理来从谱参量中重建语音波形。参量量化器根据 A. Gersho & R. M. Gray, 的矢量量化和信号压缩(Vector Quantization and Signal Compression)(1992)中描述的已有量化技术通过用已存储的码矢量代表表示这些参量来保存它们。

一种著名的时域编码器是在 L. B. Rabiner & R. W. Schaffer, 的语音信号数字处理(Digital Processing of Speech Signals)396-453(1978, 在此作为合作参考)中所描述的代码激发线性预测(CELP)编码器。在 CELP 编码器中，通过线性预测(LP)分析去除了短期相关或冗余，该分析是找出短期共振峰滤波器的系数。对输入语音帧使用短期预测滤波器就产生 LP 剩余信号，该信号将进一步用长期预测滤波器参数和后续随机码本进行模拟和量化。这样，CELP 编码将对时域语音波形的编码任务分为对 LP 短期滤波器系数编码和对 LP 剩余编码的独立任务。时域编码能以固定速率(即对每个帧使用相同的位数， $N_o$ )或可变速率(对不同类型的帧

内容使用不同的速率)执行。可变速率编码器试图仅使用足够获得目标质量水平而对编解码器参量进行编码所需的位数。一种示范可变速率 CELP 编码器在美国专利号 5,414,796(已转让给本发明的受让人,并在此作为合作参考)中有描述。

时域编码器例如 CELP 编码器通常依靠较高的每帧位数  $N_0$  来保持时域语音波形的精确度。这样的编码器通常以相对较大的每帧位数  $N_0$  (例如 8 kbps 或以上)所提供的极好的语音质量进行传输。然而,在较低比特率(4kbps 和以下),时域编码器由于有限的可用位数而不能保持高质量传输和稳健的性能。在低比特率时,有限的码本空间削减了传统时域编码器的波形匹配能力,该编码器在更高比特率的商业应用中使用得非常成功。因此,虽然随时间进行了很多改进,但是,许多在低比特率上工作的 CELP 编码系统还是受到通常用噪声表征的明显感觉上失真的困扰。

当前人们对开发在中到低比特率(即 2.4 到 4kbps 和以下的范围)工作的高质量语音编码器有着浓厚的研究兴趣和强烈的商业需求。其应用领域包括无线电话、卫星通信、因特网电话、各种多媒体和语音流应用程序、语音邮件和其他语音存储系统。其驱动力是人们对高容量的需求和在包丢失情况下对稳健性能的要求。各种新近的语音编码标准化工作是另一种推动低比特率语音编码算法研究和发展的直接驱动力。低比特率语音编码器在每个允许的应用带宽上创建更多的信道或用户,并且结合有适合信道编码的附加层的低比特率语音编码器能符合编码器规范的总体位预算,并能在信道错误的条件下提供稳健的性能。

一种在低比特率下能有效对语音编码的有用技术是多模编码。一种示范多模编码技术在美国申请序列号 09/217,341 在 1998.12.21 申请的名称为可变比特率语音编码(VARIABLE RATE SPEECH CODING,已转让给本发明的受让人并在此作为合作参考)中有描述。传统的多模编码器对不同类型的输入语音帧采用不同的模式或编码-解码算法。每种模式或编码-解码处理是为以最有效的方式最佳表示某种类型语音段而定制的,例如即有声语音、无声语音、过渡语音(例如有声和无声之间)和背景噪声(无语音)。一种外部开环模式判定机制对输入语音帧进行检验,并做出有关对帧采用什么模式的判定。开环模式判定通常是通过从输入帧中提取许多参量,对有关某些时间和频谱特性的参数进行评估,并以评估值作为模式判定的基础。

在许多传统语音编码器中,通过未充分减少码率而对有声语音帧进行编码,在未利用有声语音的稳态特性情况下,传输线性谱信息例如线性谱对或线性谱余



弦。因此，浪费了宝贵的带宽。在另一些传统语音编码器、多模式语音编码器或低比特率语音编码器中，对每帧都利用有声语音的稳态特性。因此，非稳态帧性能退化，并影响了语音质量。提供一种能反应每帧语音内容特性的自适应编码方法是很有益的。另外，因为有益信号通常是非稳态或非平稳的，在语音编码中使用的线性谱信息(LSI)参数的量化效率可以通过使用对每帧语音的LSI参数可选择性地使用基于移动平均(moving-average) (MA)预测矢量量化(VQ)或其他标准VQ方法进行编码的方案得到改进。这种方案适合发挥上述两种VQ方法的优势。因此，需要提供一种语音编码器，该编码器在从一种方法过渡到另一种方法的边界处通过适当地混合两种方案来交织两种VQ方法。这样，就需要一种使用多种矢量量化方法来适应在周期帧和非周期帧之间变化的语音编码器。

### 发明内容

本发明针对一种使用多种矢量量化方法来适应在周期帧和非周期帧之间变化的语音编码器。因此，在本发明的一个方面中，语音编码器最好包括配置来分析帧并依据上述分析生成线性谱信息码矢量的线性预测滤波器；和与线性预测滤波器耦合并配置用于使用基于非移动平均预测矢量量化方案的第一矢量量化技术对线性谱信息矢量进行矢量量化的量化器，其中该量化器进一步配置来计算用于第一技术的等效移动平均的码矢量，用等效移动平均码矢量来更新经语音编码器预先处理的预定帧数的码矢量移动平均码本的存储值，依据已更新的移动平均码本存储值来计算用于第二技术的目标量化矢量，用第二矢量量化技术对目标量化矢量进行矢量量化来产生量化的目标码矢量，第二矢量量化技术使用基于移动平均预测方案，用已量化的目标码矢量来更新移动平均码本的存储值，并从已量化的目标码矢量中计算量化线性谱信息矢量。

在本发明的另一方面中，对帧的线性谱信息矢量进行矢量量化的方法，使用第一和第二量化矢量量化技术，第一技术使用基于非移动平均预测矢量量化方案，第二技术使用基于移动平均预测矢量量化方案，最好包括用第一矢量量化技术对线性谱信息矢量进行矢量量化的步骤；计算用于第一技术的等效移动平均码矢量的步骤；用等效移动平均码矢量更新经语音编码器预先处理的预定帧数的码矢量移动平均码本存储值的步骤；依据已更新的移动平均码本存储值来计算用于第二技术的目标量化矢量的步骤；用第二矢量量化技术对目标量化矢量进行矢量量化来产生量化的目标码矢量的步骤；用已量化的目标码矢量来更新移动平均码本的

存储的步骤；以及从已量化的目标码矢量中导出量化线性谱信息矢量的步骤。

在本发明的另一方面中，语音编码器最好包括用第一矢量量化技术对线性谱信息矢量进行矢量量化的装置，该技术使用基于非移动平均预测矢量量化方案；用于计算用于第一技术的等效移动平均码矢量的装置；用于用等效移动平均码矢量更新经语音编码器预先处理的预定帧数的码矢量移动平均码本存储值的装置；用于依据已更新的移动平均码本存储值来计算用于第二技术的目标量化矢量的装置；用于用第二矢量量化技术对目标量化矢量进行矢量量化来产生量化的目标码矢量的装置；用于用已量化的目标码矢量来更新移动平均码本的存储的装置；以及用于从已量化的目标码矢量中导出量化线性谱信息矢量的装置。

#### 附图说明

图 1 是无线电话系统的框图。

图 2 是由语音编码器在每个端点终止的通信信道框图。

图 3 是编码器框图。

图 4 是解码器框图

图 5 是说明语音编码判决过程的流程图。

图 6A 是语音信号放大与时间的相对图

#### 具体实施方式

下述示范实施例是驻留在使用 CDMA 空中接口配置的无线电话通信系统中。然而，对于本领域的熟练技术人员来说应该理解使用本发明特征的子抽样方法和设备可以安置在为本领域熟练技术人员所熟知的广阔技术领域中所使用的各种通信系统中的任意系统中。

如图 1 所示，CDMA 无线电话系统通常包括多个移动用户单元 10、多个基站 12、基站控制器(BSCs)14 和移动交换中心(MSC)16。MSC16 配置来与传统的公用电话交换网(PSTN)18 对接。MSC 也配置来与 BSCs 14 对接。BSCs 14 通过回传线与基站 12 连接。回传线可以配置来支持任意几种已知接口包括例如 E1/T1、ATM、IP、PPP、帧中继、HDSL、ADSL 或 xDSL。应该明白在系统中可能有多于 2 个的 BSCs 14。每个基站 12 最好包括至少一个扇区(未示出)，每个扇区由全向天线或沿径向从基站 12 离开指向特定方向的天线组成。或者，每个扇区可能包括两个用于分集接收的天线。每个基站 12 最好能设计成支持

多个频率分配。扇区的相交和频率分配可以称为 CDMA 信道。基站 12 也可以通称为基站收发器子系统(BTSs)12。或者,“基站”在工业界可以用统称为 BSC 14 和一个或多个 BTSs 12。BTSs 12 也能表示为“蜂窝站”12。或者,给定的 BTS 12 的单独扇区可以称为蜂窝站。移动用户单元 10 通常是蜂窝或 PCS 电话 10。根据 IS-95 标准对该系统的使用进行了有利的配置。

在蜂窝电话系统的典型工作期间,基站 12 从移动单元 10 组中接收到反向链路信号集。移动单元 10 处理电话呼叫或其他通信。由给定基站 12 接收的每个反向链路信号在该基站 12 中进行处理。结果数据提交给 BSCs 14。BSCs 14 提供呼叫资源分配和移动性管理的功能包括在基站 12 之间的软切换控制。BSCs 14 也将接收的数据发送给 MSC 16, MSC 16 提供了与 PSTN 18 对接的附加路由服务。同样, PSTN 18 与 MSC 16 对接,并且 MSC 16 与 BSCs 14 对接, BSCs 14 依次控制基站 12 向移动单元 10 组发送前向链路信号集。

在图 2 中,第一编码器 100 接收数字化语音采样  $s(n)$ ,并对采样  $s(n)$  编码用于在传输介质 102 或通信信道 102 上向第一解码器 104 传输。解码器 104 对编码的语音采样进行解码,并合成为输出语音信号  $s_{\text{SYNTH}}(n)$ 。为了能在反向传输,第二编码器 106 对在通信信道 108 上传输的数字化语音采样  $s(n)$  进行编码。第二解码器 110 接收编码的语音采样并对其进行解码,生成经合成的输出语音信号  $s_{\text{SYNTH}}(n)$ 。

语音采样  $s(n)$  代表根据本领域已知各种方法,包括例如脉冲编码调制(PCM)、压扩  $\mu$ -律(companded  $\mu$ -law)或 A-律,中的任何方法经数字化和量化的语音信号。如本领域中所知,语音采样  $s(n)$  是以输入数据帧的形式编制,其中每个帧由预定数量的数字化语音采样  $s(n)$  组成。在示范实施例中,使用 8kHz 的采样率,就是 20ms 的帧由 160 个采样组成。在下述实施例中,数据传输率在帧与帧的基础上从 13.2 kbps(全速)到 6.2 kbps(半速)到 2.6 kbps(1/4 速)到 1 kbps(1/8 速)进行有利地变化。变化的数据传输率具有优势是因为对于含有相对较少语音信息的帧可选择使用低比特率。如本领域熟练技术人员所知,可以使用其他采样率、帧大小和数据传输率。

第一编码器 100 和第二解码器 110 都由第一语音编码器或语音编译码器组成。语音编码器可以用在用于传输语音信号的任意通信设备中,包括例如如图 1 中所述的用户单元、BTSs 或 BSCs。同样,第二编码器 106 和第一解码器 104 都由第二语音编码器组成。本领域熟练技术人员可以了解语音编码器

可以用数字信号处理器(DSP)、专用集成电路(ASIC)、离散门逻辑、固件或任何传统的可编程软件模块和微处理器来实现。软件模块可以驻留在RAM存储器、快闪存储器、寄存器或任何本领域已知的可写入存储媒体的其他形式中。或者,可以用任何传统的处理器、控制器或状态机来替代微处理器。特别设计用于语音编码的示范例ASICs在美国专利号5,727,123(已转让给本发明的受让人,并在此作为合作参考)以及美国申请号08/197,417名为声码器ASIC(VOCODER ASIC, 1994.2.16申请,已转让给本发明的受让人,并在此作为合作参考)中有描述。

在图3中,可以用在语音编码器中的编码器200包括模式判决模块202、音调估值模块204、LP分析模块206、LP分析滤波器208、LP量化模块210和剩余量化模块212。输入语音帧 $s(n)$ 提供给模式判决模块202、音调估值模块204、LP分析模块206和LP分析滤波器208。模式决策模块202依据每个输入语音帧 $s(n)$ 的周期、能量、信噪比(SNR)或过零率和其他特征来产生模式索引 $I_M$ 和模式 $M$ 。根据周期对语音帧分类的各种方法在美国专利号5,911,128(已转让给本发明的受让人,并在此作为合作参考)中有描述。在电信工业协会临时标准TIA/EIA IS-127和TIA/EIA IS-733也包括有这样的方法。一种示范模式判决方案在上述美国申请号09/217,341中也有描述。

音调估计模块204依据每个输入语音帧 $s(n)$ 产生音调索引 $I_p$ 和滞后值 $P_0$ 。LP分析模块206对每个输入语音帧 $s(n)$ 执行线性预测分析来产生LP参量 $a$ 。LP参量 $a$ 提供给了LP量化模块210。LP量化模块210也接收模式 $M$ ,因此,就以与模式有关的方式执行量化处理。LP量化模块210产生LP索引 $I_{LP}$ 和已量化的LP参数。LP分析滤波器208除输入语音帧 $s(n)$ 之外还接收已量化的LP参数 $\hat{a}$ 。LP分析滤波器208生成LP剩余信号 $R[n]$ ,该信号依据量化线性预测参数 $\hat{a}$ 表示了输入语音帧 $s(n)$ 和重构语音之间的错误。LP剩余 $R[n]$ 、模式 $M$ 和量化LP参数 $\hat{a}$ 提供给剩余量化模块212。依据这些值,剩余量化模块212产生剩余索引 $I_R$ 和量化剩余信号 $\hat{R}[n]$ 。

在图4中,可以在语音编码器中使用的解码器300包括LP参数解码模块302、剩余解码模块304、模式解码模块306和LP合成滤波器308。模式解码模块306接收模式索引 $I_M$ 并对其解码,从中产生模式 $M$ 。LP参数解码模块302接收模式 $M$ 和LP索引 $I_{LP}$ 。LP参数解码模块302对接收的值进行解码来产生量化LP参数 $\hat{a}$ 。剩余解码模块304接收剩余索引 $I_R$ 、音调索引 $I_p$ 和模式索引

$I_M$ 。剩余解码模块 304 对接收的值进行解码来产生量化剩余信号  $\hat{r}[n]$ 。量化剩余信号  $\hat{r}[n]$  和量化 LP 参数  $\hat{a}$  提供给 LP 合成滤波器 308, 滤波器 308 将其合成为经解码的输出语音信号  $\hat{s}[n]$ 。

图 3 的编码器 200 以及图 4 的解码器 300 的各种模块的运作和实现为本领域的熟练技术人员所熟知, 并且在上述美国专利号 5,414,796 和 L.B. Rabiner & R.W.Schafer, 的语音信号数字处理(Digital Processing of Speech Signals)396-453(1978)中有描述。

如图 5 中流程图所示, 根据一个实施例的语音编码器按照一组步骤来处理用于传输的语音采样。在步骤 400, 语音编码器接收连续帧中的语音信号数字采样。一当接收到的给定帧, 语音编码器进入步骤 402。在步骤 402 中, 语音编码器检测帧的能量。该能量是测量帧语音活动的一种度量。通过将数字化语音采样振幅的平方求和, 并将结果能量和阈值进行比较就能执行语音检测。在一个实施例中, 阈值依据背景噪声的变化水平进行适应改变。一种示范可变阈值活动检测器在上述美国专利号 5,414,796 中有描述。某些无声语音声音可以是非常低能量采样, 该采样可能被误认为基底噪声编码。为了避免这样的情况发生, 可能用低能量采样的光谱倾斜来从基底噪声中分辨无声语音, 如上述美国专利号 5,414,796 所述。

在检测帧能量之后, 语音编码器进到步骤 404。在步骤 404 中, 语音编码器对检测到的帧能量是否足够将帧分类为含有语音信息的帧进行判定。如果检测到的帧能量降到预定阈值之下, 语音编码器就进入步骤 406。在步骤 406 中, 语音编码器将帧作为背景噪声(即非语音或静音)进行编码。在一个实施例中, 背景噪声以 1/8 速或 1 kbps 速率进行编码。如果在步骤 404 中, 检测到的帧能量达到或超过预定阈值, 帧就分类为语音, 并且语音编码器进到步骤 408。

在步骤 408 中, 语音编码器对帧是否是无声语音进行判定, 即语音编码器检验帧的周期。各种已知周期判定方法包括例如通过使用过零和通过使用标准自相关函数(NACFs)的方法。特别是使用过零和 NACFs 来检测周期在上述美国专利号 5,911,128 和美国申请序列号 09/217,341 中有描述。另外, 上述用于从无声语音中分辨有声语音的方法包括在了电信工业协会临时标准 TIA/EIA IS-127 和 TIA/EIA IS-733 中。如果该帧在步骤 408 中判定为无声语音, 语音编码器就进行步骤 410。在步骤 410, 语音编码器将帧作为无声语音

编码。在一个实施例中，无声语音帧以 1/4 速率或 2.6 kbps 进行编码。如果在步骤 408 中，没有判定该帧为无声语音，语音编码器就进到步骤 412。

在步骤 412 中，语音编码器使用本领域已知的周期检测方法对该帧是否是过渡语音，如例如上述美国专利号 5,911,128 中所述。如果该帧确定为过渡语音，语音编码器就进到步骤 414。在步骤 414，该帧作为过渡语音(即从无声语音到有声语音的过渡)进行编码。在一个实施例中，转换语音帧根据在美国申请序列号 09/307,294 名为过渡语音帧的多脉冲内插编码(MULTIPULSE INTERPOLATIVE CODING OF TRANSITION SPEECH FRAMES)1999.5.7 申请(已转让给本发明的受让人并在此作为合作参考)中所述的多脉冲内插编码方法进行编码。在另一实施例中，过渡语音帧以全速或 13.2kbps 进行编码。

如果在步骤 412 中，语音编码器判定该帧不是过渡语音，语音编码器就进入步骤 416。在步骤 416 中，语音编码器将该帧作为有声语音进行编码。在一个实施例中，有声语音帧能以半速率或 6.2kbps 进行编码。也可以以全速率或 13.2kbps(或在 8k CELP 编码器中以全速率，8kbps)对有声语音帧进行编码。本领域的熟练技术人员可以理解以半速率进行有声帧编码允许编码器通过利用有声帧的稳态特性来节省宝贵的带宽。进一步，不管用于对有声语音编码的速率是多少，有声语音可以使用过去帧的信息方便地进行编码，因此可以说是通过预测进行编码。

本领域的熟练技术人员可以理解语音信号或相应的 LP 剩余可以通过如图 5 中所示的步骤进行编码。噪声、无声、过渡和有声语音的波形特征可以看作是图 6A 中的时间函数。噪声、无声、过渡和有声 LP 剩余的波形特征可以看作是图 6B 中的时间函数。

在一个实施例中，语音编码器执行如图 7 所示的流程图中的步骤来交织两种线性谱信息(LSI)矢量量化(VQ)的方法。语音编码器最好计算用于基于非 MA 预测 LSI VQ 的等效移动平均(MA)码本矢量的估值，该非 MA 预测 LSI VQ 能使语音编码器交织两种 LSI VQ 方法。在基于 MA 预测的方案中，计算 MA 用于先前处理的帧数，P，如下所述，MA 是通过将各矢量码本表项乘以参量权重来计算。如下所述，从 LSI 参量的输入矢量中减去 MA 来产生目标量化矢量。本领域的熟练技术人员能很容易地理解基于非 MA 预测 VQ 的方法可以是不使用基于 MA 预测 VQ 的任何已知 VQ 方案。

通常通过使用具有帧间 MA 预测的 VQ 或通过使用任何其他标准基于非 MA

预测 VQ 方法例如分割 VQ、多级 VQ(MSVQ)、交换预测 VQ(SPVQ)或这些方法中的一些或全部方法的混合来将 LSI 参量量化。在结合图 7 所述的实施例中,使用一种方案来对任何具有基于 MA 预测 VQ 方法的上述 VQ 方法混合。这是因为基于 MA 预测 VQ 的方法最适合用于本质上是稳态或平衡的语音帧(该帧所示出信号例如图 6A-B 中所示的平衡有声帧所示的信号),基于非 MA 预测 VQ 的方法最适用于本质上是非稳态或非平衡的语音帧(该帧所示出信号例如图 6A-B 中所示的无声帧和过渡帧所示的信号)。

在用于量化 N 维 LSI 参数的基于非 MA 预测 VQ 的方案中,对于第 M 帧的输入矢量,  $L_M = \{L_M^n; n=0, 1, \dots, N-1\}$ , 是直接作为目标量化矢量使用,并且使用任何上述标准 VQ 技术将其量化为矢量  $\hat{L}_M = \{\hat{L}_M^n; n=0, 1, \dots, N-1\}$ 。

在示范帧间 MA 预测方案中,目标量化矢量如下计算

$$U_M = \left\{ U_M^n = \frac{(L_M^n - \alpha_1^n \hat{U}_{M-1}^n - \alpha_2^n \hat{U}_{M-2}^n - \dots - \alpha_P^n \hat{U}_{M-P}^n)}{\alpha_0^n}; n=0, 1, \dots, N-1 \right\} \quad (1)$$

其中  $\{\hat{U}_{M-1}^n, \hat{U}_{M-2}^n, \dots, \hat{U}_{M-P}^n; n=0, 1, \dots, N-1\}$  是对应于紧接在帧 M 之前的 P 帧 LSI 参量的码本记录,而  $\{\alpha_1^n, \alpha_2^n, \dots, \alpha_P^n; n=0, 1, \dots, N-1\}$  是各权重,这样  $\{\alpha_0^n + \alpha_1^n + \dots + \alpha_P^n = 1; n=0, 1, \dots, N-1\}$ 。随后,使用任何上述 VQ 技术将目标量化矢量  $U_M$  量化为  $\hat{U}_M$ 。经量化的 LSI 矢量如下计算

$$\hat{L}_M = \{\hat{L}_M^n = \alpha_0^n \hat{U}_M^n + \alpha_1^n \hat{U}_{M-1}^n + \dots + \alpha_P^n \hat{U}_{M-P}^n; n=0, 1, \dots, N-1\} \quad (2)$$

MA 预测方案需要过去 P 帧的码本表项,  $\{\hat{U}_{M-1}, \hat{U}_{M-2}, \dots, \hat{U}_{M-P}\}$ , 的过去值的存在。而码本表项对于那些使用 MA 方案进行自身量化的帧(在过去 P 帧中)是自动可供使用的,过去 P 帧的剩余帧可以使用基于非 MA 预测 VQ 方法来进行量化,并且其相应的码本表项( $\hat{U}$ )对于这些帧是不能直接使用的。这就使得混合或交织上述两种 VQ 方法变得很困难。

在结合图 7 所述的实施例中,下述公式最适用于计算在  $K \in \{1, 2, \dots, P\}$  其中码本表项  $\hat{U}_{M-K}$  没有明示可用的情况下的码本表项  $\hat{U}_{M-K}$  的估值  $\tilde{\hat{U}}_{M-K}$ :

$$\tilde{\hat{U}}_{M-K} = \left\{ \tilde{\hat{U}}_{M-K}^n = \frac{(\hat{L}_{M-K}^n - \beta_1^n \hat{U}_{M-K-1}^n - \beta_2^n \hat{U}_{M-K-2}^n - \dots - \beta_R^n \hat{U}_{M-K-P}^n)}{\beta_0^n}; n=0, 1, \dots, N-1 \right\} \quad (3)$$

其中  $\{\beta_1^n, \beta_2^n, \dots, \beta_P^n; n=0, 1, \dots, N-1\}$  是各权重使得  $\{\beta_0^n + \beta_1^n + \dots + \beta_P^n = 1; n=0, 1, \dots, N-1\}$ , 并且具有初始条件  $\{\tilde{\hat{U}}_{-1}, \tilde{\hat{U}}_{-2}, \dots, \tilde{\hat{U}}_{-P}\}$ 。一种示范初始条件为  $\{\tilde{\hat{U}}_{-1} = \tilde{\hat{U}}_{-2} = \dots = \tilde{\hat{U}}_{-P} = L^B\}$ , 其中  $L^B$  是 LSI 参量的

偏差值。下述是权重的示范集合：

$$\left\{ \begin{array}{l} \beta_0^n = 1; \\ \beta_1^n = \dots = \beta_p^n = 0; \end{array} \right\} \quad n = 0, \dots, N-1 \quad (4)$$

在图 7 流程图的步骤 500，语音编码器判定是否用基于 MA 预测 VQ 的技术来量化输入 LSI 矢量  $L_M$ 。该判决最好依据帧的语音内容。例如，用于平稳有声帧的 LSI 参量量化为最有利于基于 MA 预测 VQ 的方法，而用于无声帧和过渡帧的 LSI 参量量化为最有利于基于非 MA 预测 VQ 的方法。如果语音编码器确定用基于 MA 预测 VQ 的技术来量化输入 LSI 矢量  $L_M$ ，语音编码器就进入步骤 502。另一方面，如果语音编码器确定不用基于 MA 预测 VQ 的技术来量化输入 LSI 矢量  $L_M$ ，语音编码器就进入步骤 504。

在步骤 502 中，语音编码器根据上述公式(1)计算用于量化的目标  $U_M$ 。随后，语音编码器进入步骤 506。在步骤 506 中，语音编码器根据任何各种通常为本领域所知的 VQ 技术来对目标  $U_M$  量化。随后，语音编码器进入步骤 508。在步骤 508 中，语音编码器根据上述公式(2)从经量化的目标  $\hat{U}_M$  中计算经量化的 LSI 参数的矢量  $\hat{L}_M$ 。

在步骤 504 中，语音编码器根据任何各种通常为本领域所知的基于非 MA 预测 VQ 技术来对目标  $U_M$  量化。(如本领域熟练技术人员所知，在基于非 MA 预测 VQ 技术中用于量化的目标矢量为  $L_M$ ，而不是  $U_M$ 。)随后语音编码器进入步骤 510。在步骤 510 中，语音编码器根据上述公式(3)从经量化的 LSI 参数的矢量  $\hat{L}_M$  中计算等效的 MA 码矢量  $\tilde{U}_M$ 。

在步骤 512 中，语音编码器使用在步骤 506 中获得的已量化目标  $\hat{U}_M$  以及在步骤 510 获得的等效 MA 码矢量  $\tilde{U}_M$  来更新过去 P 帧 MA 码本矢量的存储值。随后，将已更新的过去 P 帧 MA 码本矢量的存储值用于步骤 502 来计算用于后继帧输入 LSI 矢量  $L_{M+1}$  量化的目标  $U_M$ 。

这样，就揭示了一种用于交织语音编码器中线性谱信息量化方法的新颖方法和设备。本领域的熟练技术人员应该理解，此处所揭示的与实施例有关的各种说明逻辑块和算法步骤可以由数字信号处理器(DSP)、专用集成电路(ASIC)、离散门或晶体管逻辑、离散硬件部件例如寄存器和 FIFO、执行一组固件指令的处理器或任何传统可编程软件模块和处理器，来实现或执行。该处理器最好是微处理器，但作为替代，该处理器也可以是任何传统处理器、控制器、微控制器或状态机。软件模块可以驻留在 RAM 存储器、快闪存储器、寄存器或任何本领域已知的可



写入存储媒体的其他形式中。本领域的熟练技术人员可以进一步理解，在上述整个描述中提到的数据、指令、命令、信息、信号、位、字符和码片最好由电压、电流、电磁波、磁场或粒子、光场或粒子或其任意组合来表示。

本发明的较佳实施例已经示出并讨论。对于本领域普通技术人员来说，在不背离本发明的精神和范畴的情况下，很明显可以对此处揭示的实施例做出许多改动。因而，本发明仅局限于下述权利要求。

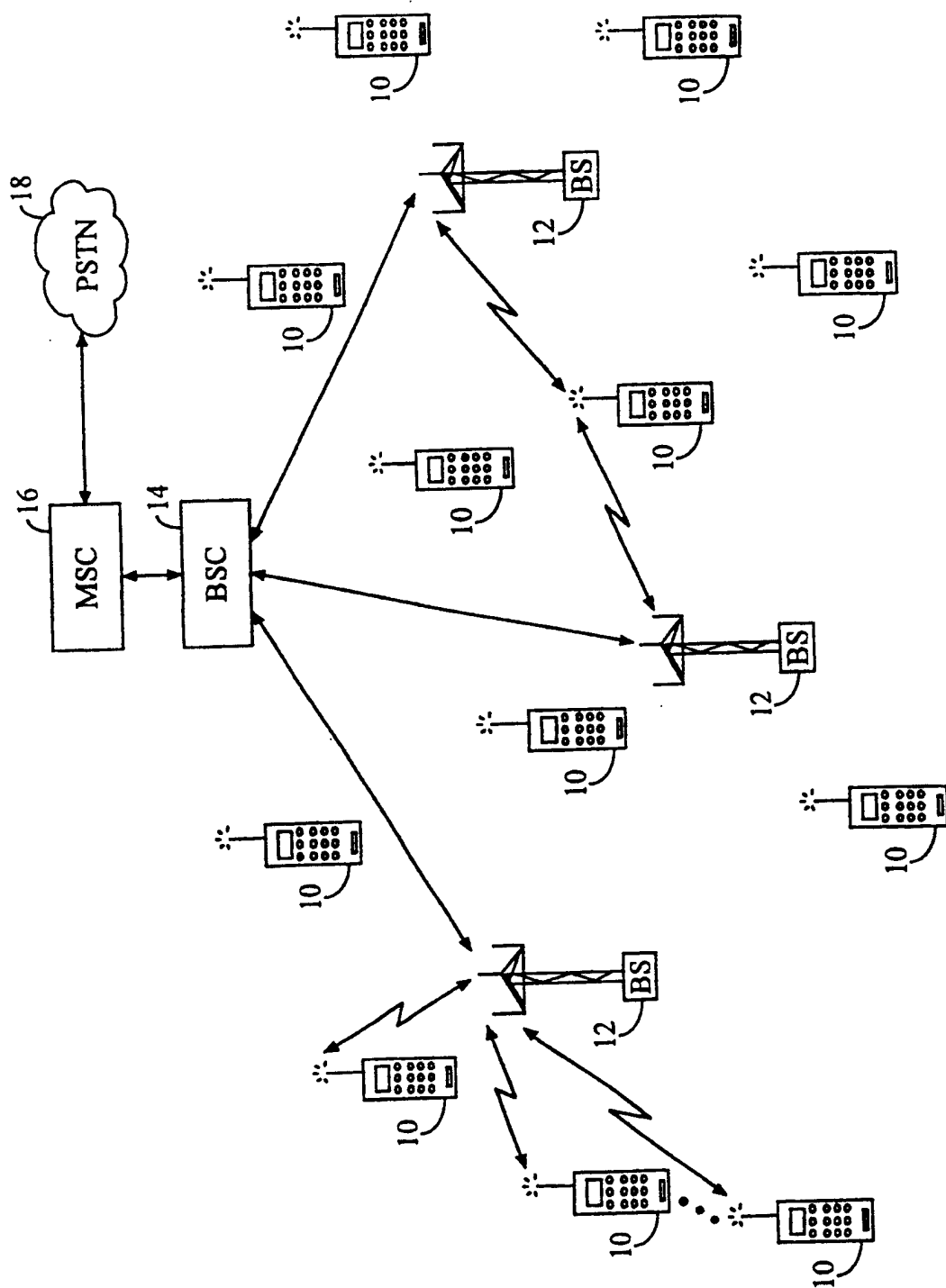


图 1

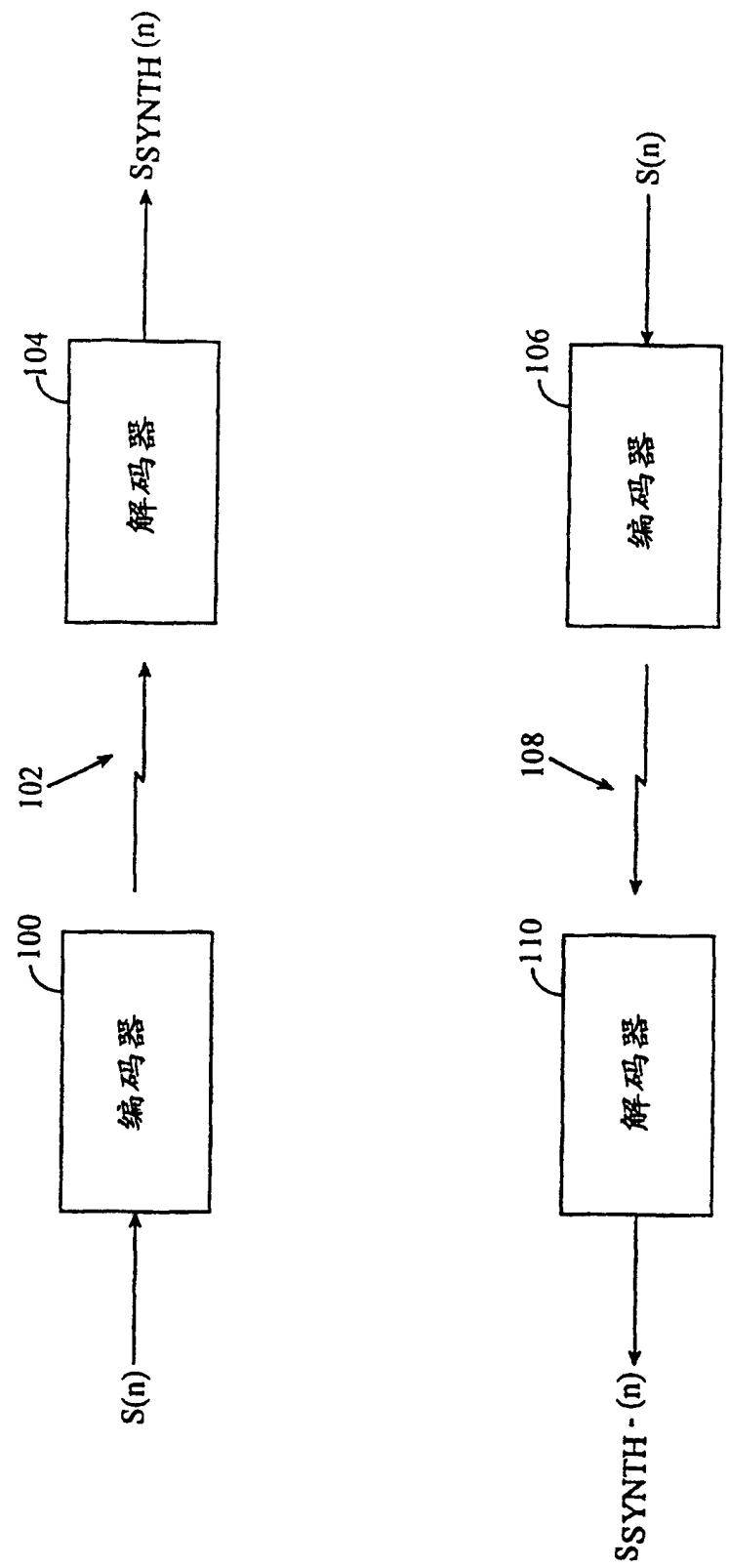
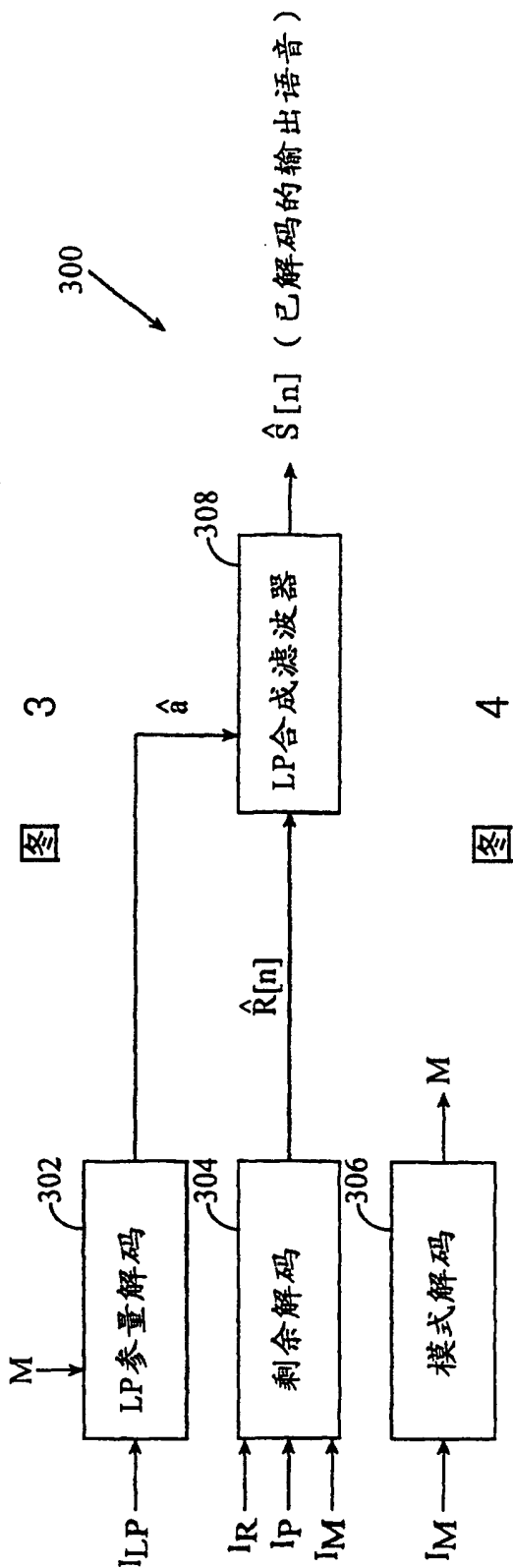
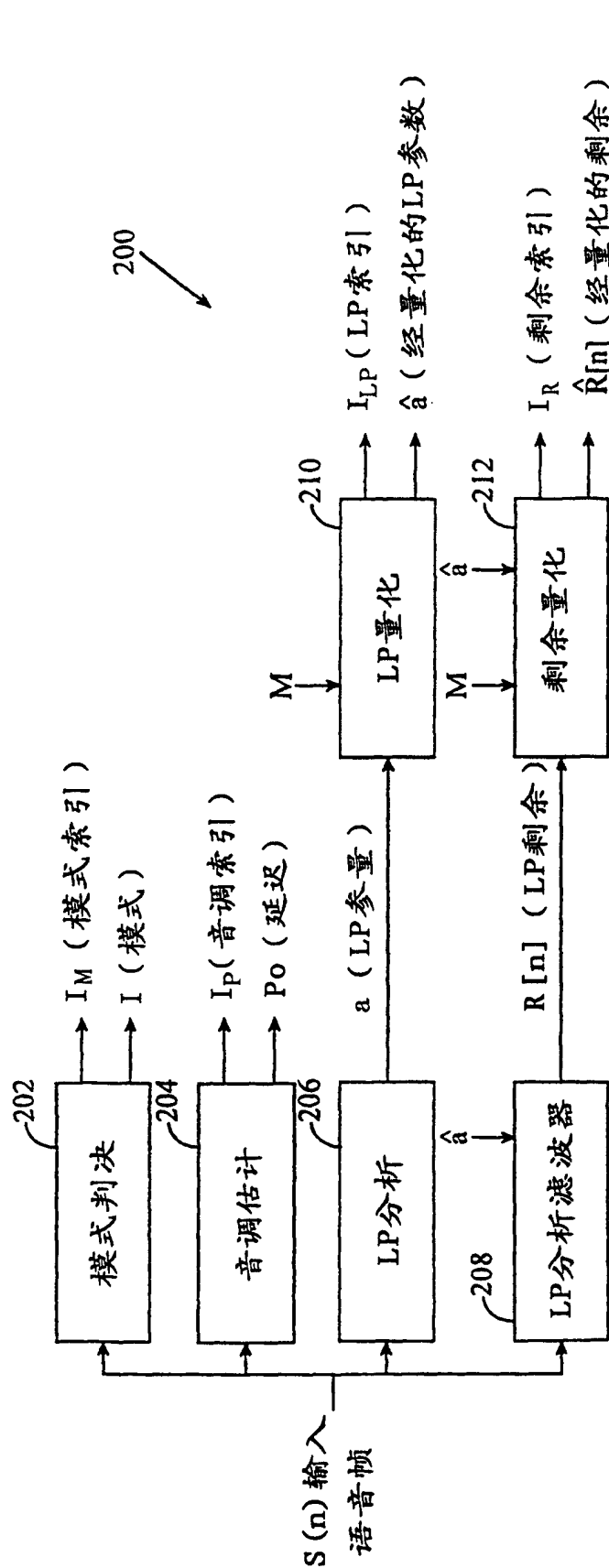


图 2



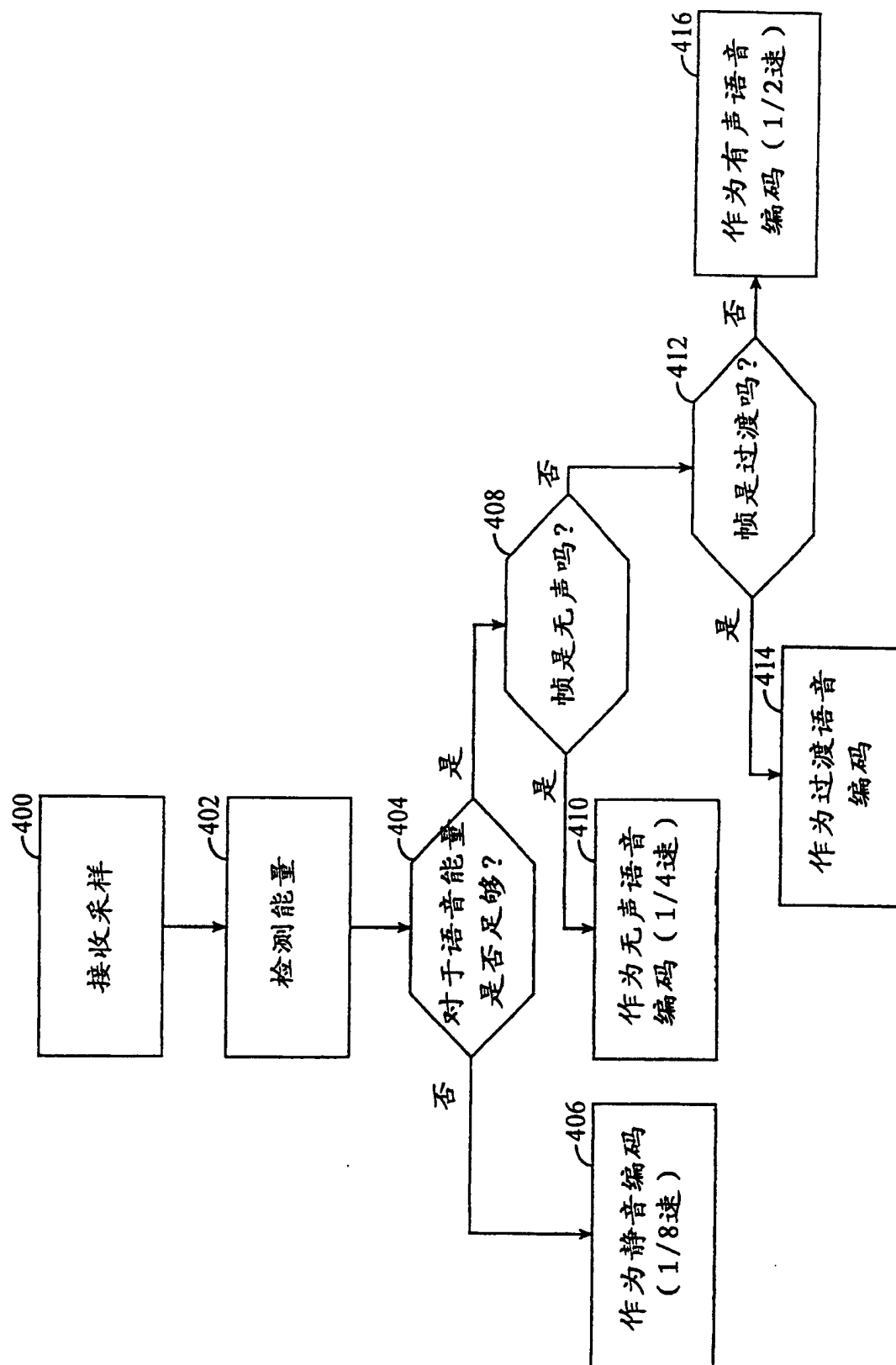


图 5

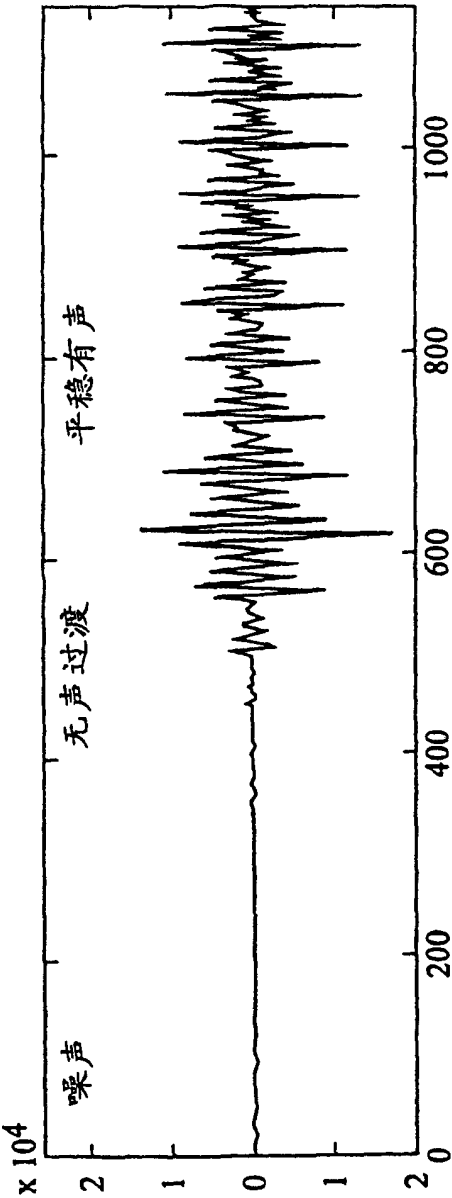


图 6A

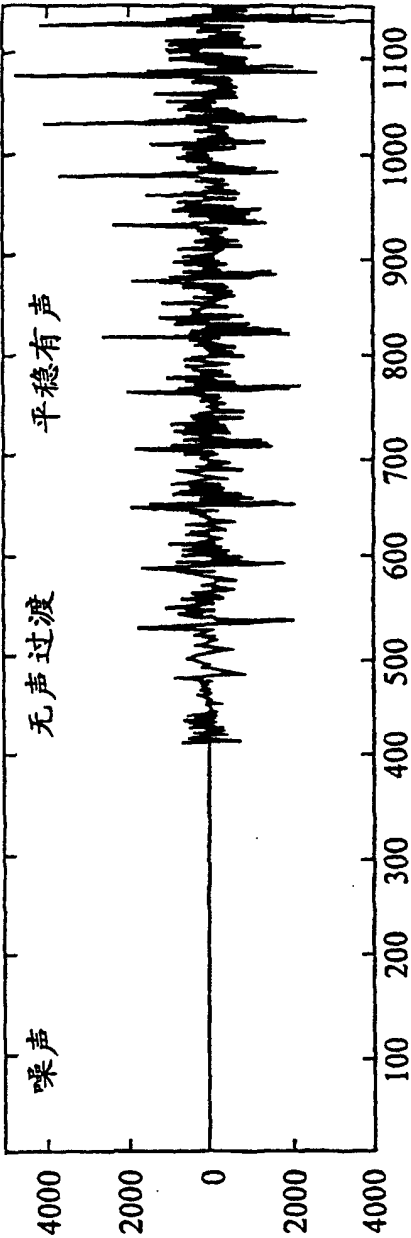


图 6B

