



# [12] 发明专利说明书

专利号 ZL 00814971.2

[45] 授权公告日 2005 年 7 月 27 日

[11] 授权公告号 CN 1212607C

[22] 申请日 2000.10.26 [21] 申请号 00814971.2

[30] 优先权

[32] 1999.10.28 [33] US [31] 09/429,754

[86] 国际申请 PCT/US2000/029710 2000.10.26

[87] 国际公布 WO2001/031639 英 2001.5.3

[85] 进入国家阶段日期 2002.4.27

[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 S·曼祖那什 A·P·德佳科

A·K·阿南萨帕德曼那伯汉

E·L·T·乔依

审查员 杨 叁

[74] 专利代理机构 上海专利商标事务所有限公司

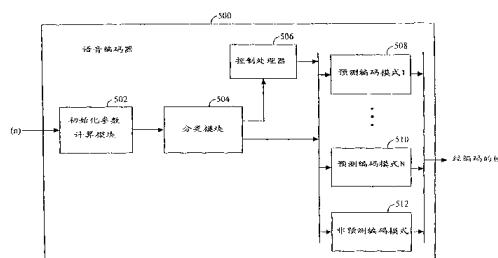
代理人 沈昭坤

权利要求书 3 页 说明书 12 页 附图 7 页

[54] 发明名称 使用编码方案选择模型以减少对帧  
差错敏感性的预测语音编码器

[57] 摘要

一种在预测语音编码器中使用编码方案选择模型 (pattern) 以减少对帧差错状态敏感性的方法和设备包括配置用于从多种预测编码模式中进行选择的语音编码器。在预定数量的语音帧已经预测编码之后, 语音编码器用非预测编码模式或轻度预测编码模式对帧进行编码。预定数量的帧可以提前由听众的主观立场确定。预定数量的帧可以进行周期性变化。通过对每个连续模型或组或预测编码语音帧保持平均编码位率, 就能保持语音编码器的平均编码位率, 所述语音帧包括至少一种非预测编码或轻度预测编码的语音帧。



- 1、一种语音编码器，其特征在于，包括：  
至少一种预测编码模式；  
至少一种较少预测编码模式；以及  
与上述至少一种预测编码模式和至少一种较少预测编码模式配合的处理器，  
所述处理器配置用于通过根据至少一种模型所选的编码模式启动对连续语音帧进行编码，所述至少一种模型包括用所述较少预测编码模式进行编码的至少一个语音帧。
- 2、如权利要求 1 所述的语音编码器，其特征在于，所述至少一种较少预测编码模式是一种轻度预测编码模式。
- 3、如权利要求 1 所述的语音编码器，其特征在于，所述至少一种较少预测编码模式是一种完全的非预测编码模式。
- 4、如权利要求 1 所述的语音编码器，其特征在于，所述处理器进一步配置用于实现经编码语音帧的预定模型以维持平均编码率。
- 5、如权利要求 4 所述的语音编码器，其特征在于，所述经编码语音帧预定模型的长度从听众的主观立场来看是事先确定的。
- 6、如权利要求 1 所述的语音编码器，其特征在于，所述至少一个模型周期性地重现。
- 7、如权利要求 1 所述的语音编码器，其特征在于，所述至少一个模型包括多个随机模型。
- 8、一种对语音帧进行编码的方法，其特征在于，包括下述步骤：  
用预测编码模式对预定数量的连续语音帧进行编码；  
在执行用预测编码模式对预定数量的连续语音帧进行编码的步骤之后，用较少预测编码模式对至少一个语音帧进行编码；以及  
重复所述两个编码步骤，以便产生多个依据模型编码的语音帧。
- 9、如权利要求 8 所述的方法，其特征在于，所述模型周期性地重现。
- 10、如权利要求 8 所述的方法，其特征在于，所述模型是随机的。
- 11、如权利要求 8 所述的方法，其特征在于，所述较少预测编码模式是轻度预测编码模式。
- 12、如权利要求 8 所述的方法，其特征在于，所述较少预测编码模式是完全

的非预测编码模式。

13、如权利要求 8 所述的方法，其特征在于，进一步包括选择经编码语音帧模型以维持平均编码率的步骤。

14、如权利要求 8 所述的方法，其特征在于，所述连续语音帧的预定数量从听众的主观立场看是事先确定的。

15、如权利要求 8 所述的方法，其特征在于，进一步包括在重复所述两个编码步骤的步骤之前，改变所述连续语音帧预定数量的步骤。

16、如权利要求 15 所述的方法，其特征在于，所述改变连续语音帧的预定数量的步骤包括以周期性方式改变所述连续语音帧的预定数量。

17、如权利要求 15 所述的方法，其特征在于，所述改变连续语音帧的预定数量的步骤包括以随机方式改变所述连续语音帧的预定数量。

18、一种语音编码器，其特征在于，包括：

用于以预测编码模式对预定数量的连续语音帧进行编码的装置；

在用所述预测编码模式对所述预定数量的连续语音帧进行编码之后，用于用较少预测编码模式对至少一个语音帧进行编码的装置；以及

用于依据用预测编码模式编码的语音帧和用较少预测编码模式进行编码的语音帧的模型来产生多个语音帧的装置。

19、如权利要求 18 所述的语音编码器，其特征在于，所述模型周期性地重现。

20、如权利要求 18 所述的语音编码器，其特征在于，所述模型是随机的。

21、如权利要求 18 所述的语音编码器，其特征在于，所述较少预测编码模式是轻度预测编码模式。

22、如权利要求 18 所述的语音编码器，其特征在于，所述较少预测编码模式是完全的非预测编码模式。

23、如权利要求 18 所述的语音编码器，其特征在于，进一步包括选择经编码语音帧模型以维持平均编码率的装置。

24、如权利要求 18 所述的语音编码器，其特征在于，所述连续语音帧的预定数量从听众的主观立场看是事先确定的。

25、如权利要求 18 所述的语音编码器，其特征在于，所述产生多个语音帧的装置进一步能用于改变所述连续语音帧的预定数量。

26、如权利要求 25 所述的语音编码器，其特征在于，所述产生多个语音帧

的装置包括用于以周期性方式改变所述连续语音帧的预定数量的装置。

27、如权利要求 25 所述的语音编码器，其特征在于，所述产生多个语音帧的装置包括用于以随机方式改变所述连续语音帧的预定数量的装置。

28、一种对语音帧进行编码的方法，其特征在于，包括以一种模型对多个语音帧进行编码的步骤，所述模型包括至少一个预测编码语音帧和至少一个较少预测编码语音帧。

29、如权利要求 28 所述的方法，其特征在于，所述模型周期性重现。

30、如权利要求 28 所述的方法，其特征在于，所述模型是随机的。

31、一种对语音帧进行编码的方法，其特征在于，包括以一种模型对多个语音帧进行编码的步骤，所述模型包括至少一个深度预测编码语音帧和至少一个轻度预测编码语音帧。

32、如权利要求 31 所述的方法，其特征在于，所述模型周期性重现。

33、如权利要求 31 所述的方法，其特征在于，所述模型是随机的。

## 使用编码方案选择模型以减少对帧差错敏感性的预测语音编码器

### 技术领域

本发明通常涉及语音处理领域，具体说涉及用于减少预测语音编码器中对帧差错情况的敏感性的方法和设备。

### 背景技术

通过数字技术进行语音传输已经被广为应用，特别是在长距离和数字无线电话应用中。因此，这就使得人们对在维持重构语音可感知质量的情况下，如何判定可以在信道上发送的最少信息量产生兴趣。如果语音仅是通过简单的采样和数字化就能发送，那么，就需要大约 64 千位/秒 (kbps) 的数据率来达到传统模拟电话的语音质量。然而，通过使用语音分析，继之以进行合适的编码、发送以及在接收机端的再合成，就可以实现数据率的明显减少。

使用通过提取有关人声生成模型的参数以压缩语音的技术的装置称为语音编码器。语音编码器将输入的语音信号分成时间块或分析帧。语音编码器通常包括编码器和解码器。编码器对输入的语音帧进行分析以提取某些相关参数，并随后将参数量化成二进制表示，即一组位或二进制数据包。数据包在通信信道上发送给接收机和解码器。解码器对数据包进行处理，将它们去量化以产生参数，并且使用这些去量化参数进行语音帧的再合成。

语音编码器的一个功能在于通过去除语音中固有的所有自然冗余以便将数字化语音信号压缩为低位率信号。数字压缩通过用一组参数代表输入语音帧，并使用量化以使用一组位来表示这些参数来实现。如果输入的语音帧具有位数  $N_i$ ，并且由语音编码器产生的数据包具有位数  $N_o$ ，则语音编码器达到的压缩系数为  $C_r = N_i/N_o$ 。其关键是在达到目标压缩系数的情况下，保持经解码语音的高品质语音。一种语音编码器的性能取决于 (1) 上述语音模型或分析和合成处理组合执行效果如何，以及 (2) 在目标  $N_o$  位/帧的位率下，执行参数量化处理的效果如何。语音模型的目的是对每个帧用一小组参数捕获语音信号的本质或目标语音品质。

语音编码器设计中最重要方面可能是对能描述语音信号的一组较好参数 (包括矢量) 的检索。一组较好参数需要一较低系统带宽用于重构精确语音信号

的再合成。间距 (pitch)、信号功率、谱包络 (或共振峰)、振幅和相频谱是语音编码参数的一些示例。

语音编码器可以作为时域编码器实现。该编码器试图通过使用高时间分辨率处理以每次对语音的小片断进行编码 (通常是 5 微秒 (ms) 子帧) 来捕获时域语音波形。对于每个子帧, 依据本领域各种已知算法可以从代码本空间找到高精度的代表。或者, 语音编码器可以作为频域编码器实现, 该编码器试图用一组参数捕获输入语音帧的短期语音频谱 (分析), 并且使用相应的合成处理以便从频谱参数中重新创建语音波形。参数量化器通过根据在 A. Gersho & R. M. Gray 的 *Vector Quantization and Signal Compression* (1992) 中所述的已有量化技术用所存储的代码矢量代表来表示这些参数, 从而保存了这些参数。

一种熟知的时域语音编码器是在 L. B. Rabiner & R. W. Schafer 的 *Digital Processing of Speech Signals* 396-453 (1978) (在此通过参考引入) 中所述的代码激励线性预测 (CELP) 编码器。在 CELP 编码器中, 通过线性预测 (LP) 分析消除了语音信号中的短期相关或冗余, 该分析查明短期共振峰滤波器的系数。将短期预测滤波器应用于输入语音帧产生 LP 残余信号, 该信号被长期预测滤波器参数和后续随机代码本进行进一步的模型化和量化。这样, CELP 编码将对时域语音波形的编码任务分成对 LP 短期滤波器系数的编码和对 LP 残余编码的单独任务。时域编码可以以固定速率 (即对每个帧都使用相同的位数  $N_0$ ) 或以可变速率 (对不同帧内容类型使用不同的位率) 来执行。可变速率编码器试图仅使用能获得目标品质水平的编译码器参数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, 并通过参考完全引入) 中有描述。

时域编码器例如 CELP 编码器通常依靠高位数  $N_0$  每帧以保留时域语音波形的精确性。这种编码器通常通过相对较大的位数  $N_0$  每帧 (例如 8 kbps 或以上) 来提供优异的语音品质。然而, 在低位率 (4 kbps 和以下), 由于有限的可用位数, 时域编码器就不能保持高品质以及稳定的性能。在低位率时, 有限的代码本空间限制了传统时域编码器的波形匹配能力, 该能力在更高速率商业应用中得到很成功的使用。因此, 尽管随着时间进行不断地改进, 许多以低位率工作的 CELP 编码系统遭受到明显的能感知失真的困扰, 该失真通常以噪声来表征。

现在人们对开发以中等到低位率 (即在 2.4 到 4kbps 及以下) 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、各种多媒体以及语音流应用、语音信箱以及其他语音存储系统。其驱动力来自对高性能的需要和对在包丢失情况下稳定性能的需求。近来各种语音编码标准化计划也是推动低速率语音编码算法研究和开发的另一直接驱动力。低速率语音编码器可以在每个容许的应用带宽上创建更多的信道或用户，并且，与适合的信道编码附加层结合的低速率语音编码器能够符合编码器规范的整体位预算，并在信道差错情况下，提供稳固的性能。一种示范低速率编码器是在美国申请序列号 09/217,341 名为 VARIABLE SPEECH CODING, 1998. 12. 21 申请(已转让给本发明的受让人，并通过参考完全引入)中所述的原型间距周期 (PPP) 语音编码器。

在传统的预测语音编码器中，例如 CELP 编码器、PPP 编码器以及波形插值 (WI) 编码器，编码方案主要依赖过去的输出。因此，如果解码器接收到帧差错或帧消除，解码器必须自身创建被怀疑帧的最佳替代。解码器通常使用重复先前输出的智能帧。因为解码器必须创建自己的替代，解码器和编码器就丧失了彼此的同步性。因而，当下一帧到达解码器时，如果该帧经过预先编码，解码器就查阅不同于编码器所使用的先前输出。这就在语音品质和语音编码器性能方面引起下降。语音编码器越是依赖预测编码技术(即语音编码器预先对越多的帧进行编码)，性能方面就会有越大的下降。这样，就需要一种能减少预测语音编码器中对帧差错情况的敏感性方法。

#### 发明内容

本发明针对一种减少预测语音编码器中对帧差错情况的敏感性的方法。因此，在本发明的一个方面中，提供了一种语音编码器。该语音编码器最好包括至少一种预测编码模式；至少一种较少预测编码模式；以及与至少一种预测编码模式和至少一种较少预测编码模式配合的处理器，该处理器配置用于通过根据编码语音帧的模型所选的编码模式启动对连续语音帧的编码，该模型包括至少一个用较少预测编码模式进行编码的语音帧。

在本发明的另一方面，提供了一种对语音帧进行编码的方法。该方法最好包括用预测编码模式对预定数量的连续语音帧进行编码的步骤；在执行用预测编码模式对预定数量的连续语音帧进行编码的步骤之后，用较少预测编码模式对至少一个语音帧进行编码；以及为了产生多个依据模型编码的语音帧重复两个编码步骤。

在本发明的另一方面，提供了一种语音编码器。该语音编码器最好包括用于用预测编码模式对预定数量的连续语音帧进行编码的装置；在用预测编码模式对预定数量的连续语音帧进行编码之后，用于用较少预测编码模式对至少一个语音帧进行编码的装置；以及用于产生多个依据模型编码的语音帧的装置，该模型包括至少一个用较少预测编码模式编码的语音帧。

在本发明的另一方面，提供了一种对语音帧进行编码的方法。该方法最好包括多模型中的多个语音帧进行编码的步骤，该模型包括至少一个经预测编码的语音帧和至少一个经较少预测编码的语音帧。

在本发明的另一方面，提供了一种对语音帧进行编码的方法。该方法最好包括多模型中的多个语音帧进行编码的步骤，该模型包括至少一个经深度预测编码的语音帧和至少一个经轻度预测编码的语音帧。

#### 附图说明

图 1 是在每个端由语音编码器中止的通信信道的框图。

图 2 是可以在图 1 的语音编码器中使用的编码器框图。

图 3 是可以在图 1 的语音编码器中使用的解码器框图。

图 4 是说明语音编码决策处理的流程图。

图 5A 是语音信号振幅与时间比较的图例，而图 5B 是线性预测 (LP) 残余振幅与时间比较的图例。

图 6 是配置使用编码模式选择模型的语音编码器框图。

图 7 是由语音编码器例如图 6 的语音编码器所执行的用来采用编码模式选择模型的方法步骤流程图。

#### 具体实施方式

在图 1 中，第一编码器 100 接收经数字化的语音采样  $s(n)$ ，并且对采样  $s(n)$  进行编码用于在传输媒体 102 或通信信道 102 上发送给第一解码器 104。传输媒体 102 可以是，例如陆基通信线路、在基站和卫星之间的链路、在蜂窝或 PCS 电话和基站之间的无线通信信道或在蜂窝或 PCS 电话和卫星之间的无线通信信道。如下所述，语音采样  $s(n)$  最好以各种代码本索引和量化噪声的形式进行编码。解码器 104 对经编码的语音采样进行解码，并且合成输出语音信号  $s_{\text{SYNTH}}(n)$ 。如下所述，解码处理最好包括使用发送的代码本索引对



各种代码本进行检索以判定用于合成输出语音信号  $s_{\text{SYNTH}}(n)$  的合适值。对于相反方向的传输，第二编码器 106 对在通信信道 108 上发送的数字化语音采样  $s(n)$  进行编码。第二解码器 110 对经编码的语音采样进行接收并解码，产生合成的输出语音信号  $s_{\text{SYNTH}}(n)$ 。

语音采样  $s(n)$  代表了根据本领域的各种已知方法中的任意方法，包括脉冲编码调制 (PCM)、压扩  $\mu$ -律或 A-律，进行数字化和量化的语音信号。如本领域所熟知，语音采样  $s(n)$  组成了输入数据帧，其中每个帧由预定数目的数字化语音采样  $s(n)$  组成。这些帧还可以进一步细分为子帧。在示范实施例中，每个帧包括 4 个子帧。在示范实施例中，使用了 8 kHz 的采样率，每 20ms 的帧包括有 160 个采样。在下述实施例中，数据传输率最好可以基于帧间进行变化。例如，数据传输率可以从全速到半速到 1/4 速到 1/8 速进行变化。因为对于含有相对较少语音信息的帧可以有选择性地使用较低的位率，所以，变化的数据传输率具有优势。如本领域的熟练技术人员所知，可以使用各种采样率、帧大小以及数据传输率。

第一编码器 100 和第二解码器 110 一起组成第一语音编码器或语音编译器。语音编码器可以用于任意发送语音信号的通信装置中包括，例如蜂窝或 PCS 电话、基站和/或基站控制器。同样，第二编码器 106 和第一解码器 104 一起组成第二语音编码器。本领域的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、固件或任意传统的可编程软件模块以及微处理器来实现。软件模块可以驻留在 RAM 存储器、闪存、寄存器或本领域已知的任何其他形式的可写存储媒体中。或者，可以用任意传统的处理器、控制器或状态机来替代微处理器。特别为语音编码设计的示范 ASIC 在美国专利号 5,727,123 (已转让给本发明的受让人，并通过参考完全引入) 以及美国专利号 5,784,532 (已转让给本发明的受让人，并通过参考完全引入) 中有描述。

在图 2 中，可以在语音编码器中使用的编码器 200 包括模式判定模块 202、间距估计模块 204、LP 分析模块 206、LP 分析滤波器 208、LP 量化模块 210 以及残余量化模块 212。输入语音帧  $s(n)$  提供给模式判定模块 202，间距估计模块 204、LP 分析模块 206、LP 分析滤波器 208。模式判定模块 202 依据每个输入语音帧  $s(n)$  的周期性、能量、信噪比 (SNR) 或零交叉率和其他特征，产生模式索引  $I_M$  和模式  $M$ 。根据周期性对语音帧进行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,并通过参考完全引入) 中有描述。这些方法也包括在电信工业协会临时标准 TIA/EIA IS-127 以及 TIA/EIA IS-733。在上述美国申请序列号 09/217,341 中也描述了一种示范模式判定方法。

间距估计模块 204 依据每个输入语音帧  $s(n)$  产生间距索引  $I_p$  和延迟值  $P_0$ 。LP 分析模块 206 对每个输入语音帧  $s(n)$  执行线性预测分析以产生 LP 参数  $a$ , LP 参数  $a$  提供给 LP 量化模块 210。LP 量化模块 210 还接收了模式  $M$ , 因而以独立模式方式执行量化处理。LP 量化模块 210 产生 LP 索引  $I_{LP}$  和经量化的 LP 参数  $\hat{a}$ 。除输入语音帧  $s(n)$  之外, LP 分析滤波器 208 还接收经量化的 LP 参数  $\hat{a}$ 。LP 分析滤波器 208 产生 LP 残余信号  $R[n]$ , 它代表在输入语音帧  $s(n)$  和依据经量化的线性预测参数  $\hat{a}$  所重构的语音之间的误差。LP 残余  $R[n]$ 、模式  $M$  以及经量化的 LP 参数  $\hat{a}$  提供给残余量化模块 212。依据这些值, 残余量化模块 212 产生残余索引  $I_R$  以及经量化的残余信号  $\hat{R}[n]$ 。

在图 3 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、残余解码模块 304、模式解码模块 306 以及 LP 合成滤波器 308。模式解码模块 306 接收并解码模式索引  $I_M$ , 从中产生模式  $M$ 。LP 参数解码模块 302 接收模式  $M$  以及 LP 索引  $I_{LP}$ 。LP 参数解码模块 302 对接收的值进行解码以产生经量化的 LP 参数  $\hat{a}$ 。残余解码模块 304 接收残余索引  $I_R$ , 间距索引  $I_p$  以及模式索引  $I_M$ 。残余解码模块 304 对所接收的值进行解码以产生经量化的残余信号  $\hat{R}[n]$ 。经量化的残余信号  $\hat{R}[n]$  以及经量化的 LP 参数  $\hat{a}$  提供给 LP 合成滤波器 308, 它从中合成出经解码的输出语音信号  $\hat{s}[n]$ 。

在上述美国专利号 5,414,796 和美国申请序列号 09/217,341 中描述了用于图 2 编码器 200 以及图 3 解码器 300 的模块的各种操作和实现技术。

如图 4 流程图所说明, 根据一实施例的语音编码器遵循一组处理传输语音采样的步骤。在步骤 400 中, 语音编码器接收连续帧中的语音信号数字采样。依据接收的给定帧, 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。该能量是帧语音活动的量度。语音检测是通过对数字化语音采样的振幅平方求和, 并将所得能量与阈值比较来执行。在一个实施例中, 依据背景噪声的变化水平来修改阈值。在上述美国专利号 5,414,796 中描述了一种示范可变阈值语音活动检测器。某些无声语音声音可以是非常低能量的采样, 它们很容易被错误作为背景噪声进行编码。为了防止这种情况的发生, 如在上述美国专利号 5,414,796 中所述, 可以使用低能量采样频谱倾斜

以便将无声语音从背景噪声中区分开来。

在检测帧能量之后，语音编码器进入步骤 404。在步骤 404 中，语音编码器判定所检测的帧能量是否足以作为含有语音信息的帧进行分类。如果所检测的帧能量在预定阈值水平之下，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（即非语音或静默）。在一个实施例中，背景噪声以 1/8 速进行编码。如果在步骤 404 中，所检测的帧能量符合或超过预定阈值水平，该帧就分类为语音，并且语音编码器进入步骤 408。

在步骤 408 中，语音编码器判定帧是否是无声帧，即语音编码器检验帧的周期性。周期性判定的各种已知方法包括，例如使用零交叉和使用归一自动纠错函数（NACFs）。在上述美国专利号 5,911,128 以及美国申请序列号 09/217,341 中特别对使用零交叉和使用 NACFs 来检测周期性进行了描述。另外，上述用于将有声语音从无声语音中区分开来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。如果在步骤 408 中判定帧是无声语音，语音编码器进入步骤 410。在步骤 410 中，语音编码器将该帧作为无声帧进行编码。在一个实施例中，无声语音帧以 1/4 速进行编码。如果在步骤 408 中判定帧不是无声语音，语音编码器就进入步骤 412。

在步骤 412 中，如在上述美国专利号 5,911,128 中所述，语音编码器使用在本领域已知的周期检测方法来判断该帧是否是过渡语音。如果判定该帧是过渡语音，语音编码器就进入步骤 414。在步骤 414 中，该帧就作为过渡语音（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，根据在美国专利号 6,260,017, 1999.5.7 公布，（已转让给本发明的受让人，并通过参考完全引入）中所述的多脉冲内插编码方法对过渡语音帧进行编码。在另一实施例中，过渡语音帧以全速进行编码。

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，有声语音帧可以以半速进行编码。也可以以全速对有声语音帧进行编码。然而，本领域的熟练技术人员可以理解通过利用有声帧的稳态性质，以半速对有声帧进行编码允许编码器节约有用的带宽。而且，不考虑用于编码有声语音的速率，有声语音优势是可以使用从过去的帧得到的信息进行编码，并且因此，称为周期性编码。

熟练技术人员可以理解语音信号或对应的 LP 残余可以通过遵循在图 4 中

所述的步骤进行编码。噪声、无声、过渡和有声语音的波形特征可以视作为在图 5A 中图例的时间函数。噪声、无声、过渡和有声 LP 残余的波形特征可以视作是图 5B 图例的时间函数。

在一个实施例中，如图 6 所示，周期性对一定比例的帧进行编码的语音编码器 500 配置用于通过使用确定性的编码方案选择模型以减少对帧差错情况的敏感性。语音编码器 500 包括初始参数计算模块 502、分类模块 504、控制处理器 506、多种（N 种）预测编码模式 508、510（为了简洁，仅示出两种预测编码模式 508、510，剩余的预测编码模式由虚线表示），以及至少一种较少预测编码模式 512。初始参数计算模块 502 与分类模块 504 耦合。分类模块 504 与控制处理器 506 以及各种编码模式 508、510、512 耦合。控制处理器也与各种编码模式 508、510、512 耦合。

经数字化的语音采样  $s(n)$  由语音编码器 500 接收，并输入给初始参数计算模块 502。初始参数计算模块 502 从语音采样  $s(n)$  中获得各种初始参数，包括例如线性预测系数（LPC 系数）、线频谱配对（LSP）系数、归一自动纠正函数（NACFs）、开环延迟参数、带能量、零交叉率以及共振峰残余信号。各种初始参数的计算和使用为本领域所熟知，并在上述美国专利号 5,414,796 和美国申请序列号 09/217,341 中有描述。

这些初始参数提供给分类模块 504。依据这些初始参数值，分类模块 504 依据上面参照图 4 描述的分类步骤对语音帧进行分类。帧分类提供给控制处理器 506，并且语音帧提供给各种编码模式 508、510、512。

控制处理器 506 最好配置用于依据哪个模式最适用于当前帧语音属性，而在多种编码模式 508、510、512 之间进行动态逐帧转换。为每个帧选择特定的编码模式 508、510、512，以便在解码器（未示出）处保持可接受信号再现情况下，达到最低可用位率。这样，当语音信号  $s(n)$  的属性变化时，语音编码器 500 的位率也随时间改变，这种处理称为可变速率语音编码。

在本发明的一个实施例中，控制处理器 506 依据当前语音帧的分类指示使用一个特定的预测编码模式 508、510。预测编码模式 508、510 之一是一种 CELP 编码模式，这种模式在上述美国专利号 5,414,796 中有描述。预测编码模式 508、510 中的另一模式是 PPP 编码模式，这种模式在上述美国申请序列号 0/217,341 中有描述。还是预测编码模式 508、510 的另一模式可以是 WI 编码模式。

在一个实施例中，较少预测编码模式 512 是轻度预测或低存储编码方案。预测编码模式 508、510 最好是深度编码模式。在一个替代实施例中，较少预测编码模式 512 是一种完全非预测或无需存储编码方案。完全非预测编码方案 512 可以是，例如语音采样  $s(n)$  的 PCM 编码，语音采样  $s(n)$  的压扩  $\mu$ -律或语音采样  $s(n)$  的 A-律。

而在结合图 6 所述的实施例中示出一种较少预测编码模式 512，本领域的熟练技术人员可以理解可以使用超过一种的较少预测编码模块。如果使用了超过一种的较少预测编码模块，该较少预测编码模块的类型可以改变。而且，在使用超过一种的较少预测编码模块的替代实施例中，较少预测编码模块中的一部分或全部都是轻度预测编码模块。并且在其他实施例中，较少预测编码模块中的一部分或全部都是完全的非预测编码模块。

在一个实施例中，在确定期间，最好由控制处理器 506 插入较少预测编码模式 512。控制处理器 506 在帧中创建具有长度为  $F$  的模型。在一个实施例中，长度  $F$  取决于最长的帧差错影响容许持续时间。最长容许时间最好提前从听众的主观立场来确定。在另一实施例中，长度  $F$  由控制处理器 506 进行周期性变化。在其他实施例中，长度  $F$  由控制处理器 506 进行随机或伪随机变化。一种示范再现模型是 PPPN，其中  $P$  代表预测编码模式 508、510，而  $N$  表示非预测或轻度预测编码模式 512。在替代实施例中，插入了多种较少预测编码模式。一种示范模型是 PPNPPN。在模型长度  $F$  改变的实施例中，模型 PPPN 可以遵循模型 PPN，该模型 PPN 可以遵循模型 PPPNPN 等。

在一种实施例中，语音编码器，例如图 6 的语音编码器 500，执行在图 7 流程图中所述的算法步骤以智能化地在确定间隔中插入低存储或无需存储编码方案。在步骤 600 中，控制处理器（未示出）设定计数变量  $i$  等于 0。随后，控制处理器进入步骤 602。在步骤 602 中，控制处理器依据当前帧的语音内容分类为当前语音帧选择预测编码模式。随后，控制处理器进入步骤 604。在步骤 604 中，控制处理器用所选的预测编码模式对当前帧进行编码。随后，控制处理器进入步骤 606。在步骤 606 中，控制处理器递增计数变量  $i$ 。随后，控制处理器进入步骤 608。

在步骤 608 中，控制处理器判定计数变量  $i$  是否大于预定阈值  $T$ 。预定阈值  $T$  可以依据帧差错影响的最长容许持续时间，如最长容许时间提前从目标听众的立场来确定。在特定实施例中，预定阈值  $T$  在流程图的预定重复次数

期间保持不变，并随后由控制处理器改变为不同的预定值。如果计数变量  $i$  不大于预定阈值  $T$ ，控制处理器返回步骤 602 以便为下一语音帧选择预测编码模式。另一方面，如果计数变量  $i$  大于预定阈值  $T$ ，控制处理器进入步骤 610。在步骤 610 中，控制处理器用非预测或轻度预测编码模式对下一语音帧进行编码。随后，控制处理器返回步骤 600，再次设定计数变量  $i$  等于 0。

本领域的熟练技术人员可以认识到图 7 的流程图可以进行修改以结合不同的预测编码和非预测或轻度预测编码语音帧的再现模型。例如，计数变量  $i$  可以随流程图的每次重复或在预定次数重复流程图之后，或伪随机，或随机进行变化。或者，例如在步骤 610 中可以用非预测编码模式或轻度预测编码模式对下两个帧进行编码。或者，例如，在步骤 610 中，可以用非预测编码模式或轻度预测编码模式对任意预定数目的帧，或随机选定数目的帧，或伪随机选定数目的帧，或随流程图的每次重复以预定方式变化的许多帧进行编码。

在一个实施例中，图 6 的语音编码器 500 是一种可变速率语音编码器 500，并且最好保持语音编码器 500 的平均位率。在特定实施例中，在模型中所使用的每个预测编码模式 508、510 以彼此不同的速率进行编码，并且较少预测编码模式 512 以不同于任意预测编码模式 508、510 所使用的速率进行编码。在另一特定实施例中，预测编码模式 508、510 以相对较低的位率进行编码，而较少预测编码模式 512 以相对较高的位率进行编码。因此，每隔  $F$  个帧就插入一次高质量、低存储或无需存储的编码方案，而在连续的高位率帧之间使用中-高质量、深度预测、低位率的编码方案，产生减少的平均编码率。而与任意预测语音编码器相比的优势在于，这种技术特别适用于低位率语音编码器，在其中可以仅通过使用深度预测编码方案来达到良好的语音品质。由于它们的预测性质，这样的低位率语音编码器更加容易受到由帧差错引起的不良影响。通过周期性插入高位率、较少预测编码模式 512，而同时保持各种低位率的预测编码模式 508、510，就可以同时达到所需的良好语音品质和低平均编码率。

在一个实施例中，通过将语音片断中的所有帧以重复、确定模型进行编码，使得平均编码率最好在预定平均率保持恒定或接近恒定，这样平均率就等于  $R$ 。一种示范模型是 PPN，其中  $P$  代表预测编码的帧，而  $N$  代表非预测或轻度预测编码的帧。在这种模型中，第一帧以  $R/2$  进行预测编码，第二帧以  $R/2$

速率进行预期编码，而第三帧以  $2R$  速率进行非预测或轻度预测编码。随后，重复模型等。这样，平均编码率为  $R$ 。

另一示范模型是 PPPN。在这种模型中，第一帧以  $R/2$  速率进行预测编码，第二帧以  $R$  速率进行预期编码，第三帧以  $R/2$  速率进行预测编码，而第四帧以  $2R$  速率进行非预测或轻度预测编码。随后，重复模型等。这样，平均编码率为  $R$ 。

另一示范模型是 PPNPPN。在这种模型中，第一帧以  $R/2$  速率进行编码，第二帧以  $R/2$  速率进行编码，第三帧以  $2R$  速率进行编码，第四帧以  $R/3$  速率进行编码，第五帧以  $R/3$  速率进行编码，而第六帧以  $7R/3$  速率进行编码。随后，重复模型等。这样，平均编码率为  $R$ 。

另一示范模型是 PPNPN。在这种模型中，第一帧以  $R/3$  速率进行编码，第二帧以  $R/3$  速率进行编码，第三帧以  $R/3$  速率进行编码，第四帧以  $3R$  速率进行编码，第五帧以  $R/2$  速率进行编码，而第六帧以  $3R/2$  速率进行编码。随后，重复模型等。这样，平均编码率为  $R$ 。

另一示范模型是 PPNPPN。在这种模型中，第一帧以  $R/3$  速率进行编码，第二帧以  $R/3$  速率进行编码，第三帧以  $2R$  速率进行编码，第四帧以  $2R$  速率进行编码，第五帧以  $R/2$  速率进行编码，第六帧以  $R/2$  速率进行编码，而第七帧以  $4R/3$  速率进行编码。随后，重复模型等。这样，平均编码率为  $R$ 。

本领域的熟练技术人员可以理解也能使用任意上述模型的任意循环轮转。本领域的熟练技术人员也可以认识到上述模型以及其他模型可以以所选的随机或伪随机，或实际周期的任意顺序结合在一起。本领域的熟练技术人员可以进一步理解可以使用任意组编码率，使得这些编码率平均为在模型持续时间上（ $F$  帧）所需的平均编码率  $R$ 。

强制要进行非预测或轻度编码的帧以高速率进行编码会使得帧差错的影响仅持续与模型一样长的时间，而保持了语音片断的所需平均编码率  $R$ 。实际上，如果语音片断没有包括确切的  $F$  个帧（模型长度）的倍数，控制处理器可以配置用于对模型进行智能轮换以达到或多或少更低的平均速率。如果语音片断的所需有效平均编码率  $R$  改为通过以固定速率  $R$  对片断中的所有帧进行编码来实现，并且速率  $R$  对于使用预测来说相对较低，语音编码器将非常易受持续的帧差错影响。

本领域的熟练技术人员可以理解，虽然，上述实施例是基于可变速率语

音编码器，但基于模型的方案（例如上述这些方案）也可以适用于固定速率的预测语音编码器。如果固定速率的预测语音编码器是低位率语音编码器，帧差错情况就会对语音编码器有不利影响。非预测编码或轻度预测编码帧可能具有比以相同低速率编码的预测编码帧更低的品质。不过每隔 F 个帧就引入一非预测编码或轻度预测编码帧就会消除每隔 F 个帧的帧差错影响。

这样，就已经描述了一种用于在预测语音编码器中使用编码方案选择模型以减少对帧差错情况敏感性的新颖方法和设备。本领域的熟练技术人员可以理解结合在此揭示的实施例所描述的各种说明逻辑框和算法步骤可以通过电子硬件、计算机软件或两者的结合来实现。各种说明部件、框和步骤通常依据其功能性进行描述这些功能性是由硬件或软件来实现是取决于特定的应用和施加在整个系统上的设计约束。熟练技术人员认识到在情况下的硬件和软件的交换性，以及如何能针对每个特定应用最佳地实现所述功能性。作为示例，结合在此揭示的实施例所描述的各种说明逻辑框和算法步骤可以由数字信号处理器（DSP）、专用集成电路（ASIC）、分立门或晶体管逻辑、分立的硬件部件，例如寄存器和 FIFO、执行一组固件指令的处理器或任意传统可编程软件模块以及处理器来实现或执行。处理器最好是微处理器，但作为替代，处理器可以是任意传统的处理器、控制器、微控制器或状态机。软件模块可以驻留在 RAM 存储器、闪存、寄存器或不利于已知的任意其他形式的可写存储媒体。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所参照的数据、指令、命令、信息、信号、位、符号和芯片最好由电压、电流、电磁波、磁场或粒子、光场或粒子、或其任意结合来表示。

本发明的较佳实施例已经示出，并进行了描述。在不背离本发明的精神和范畴的情况下，对于本领域的熟练技术人员来说，对这些实施例各种修改将是显而易见的。因而，除下述权利要求外，本发明并不受其他限制。



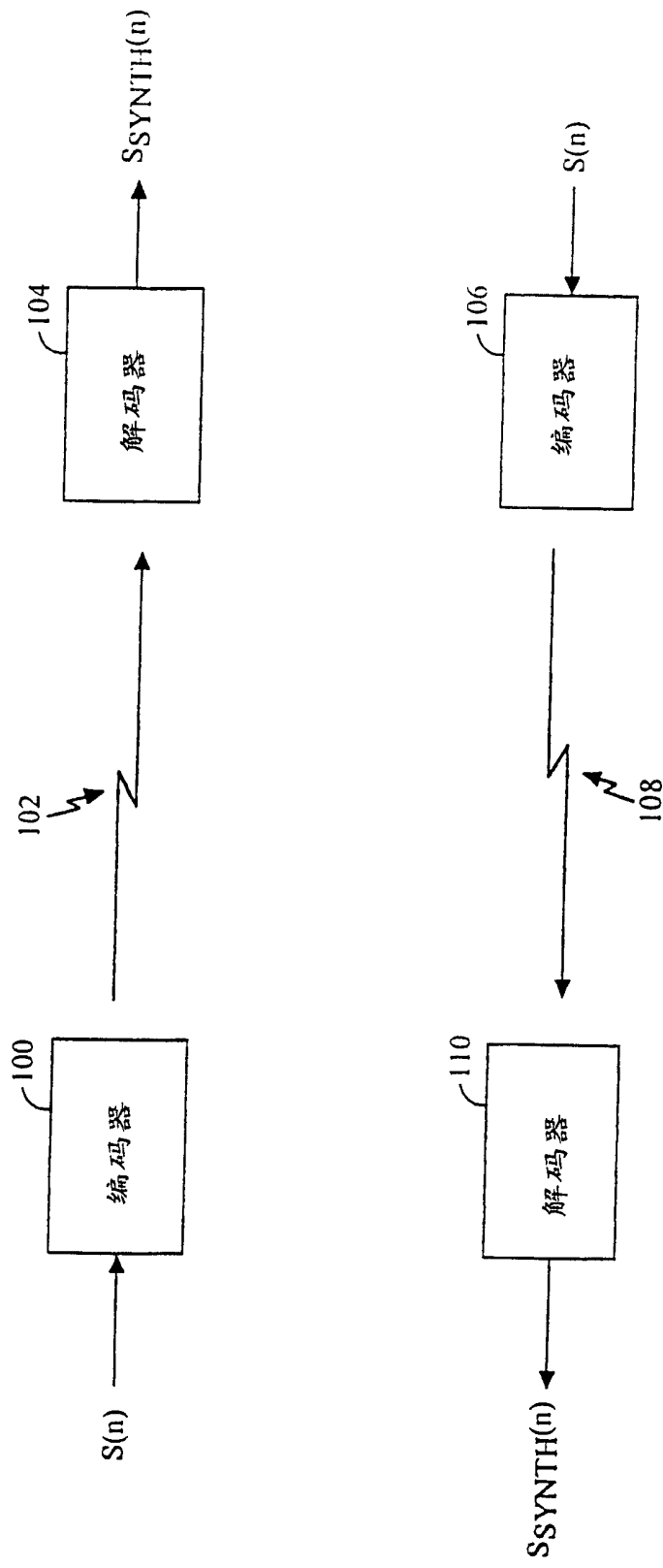


图 1

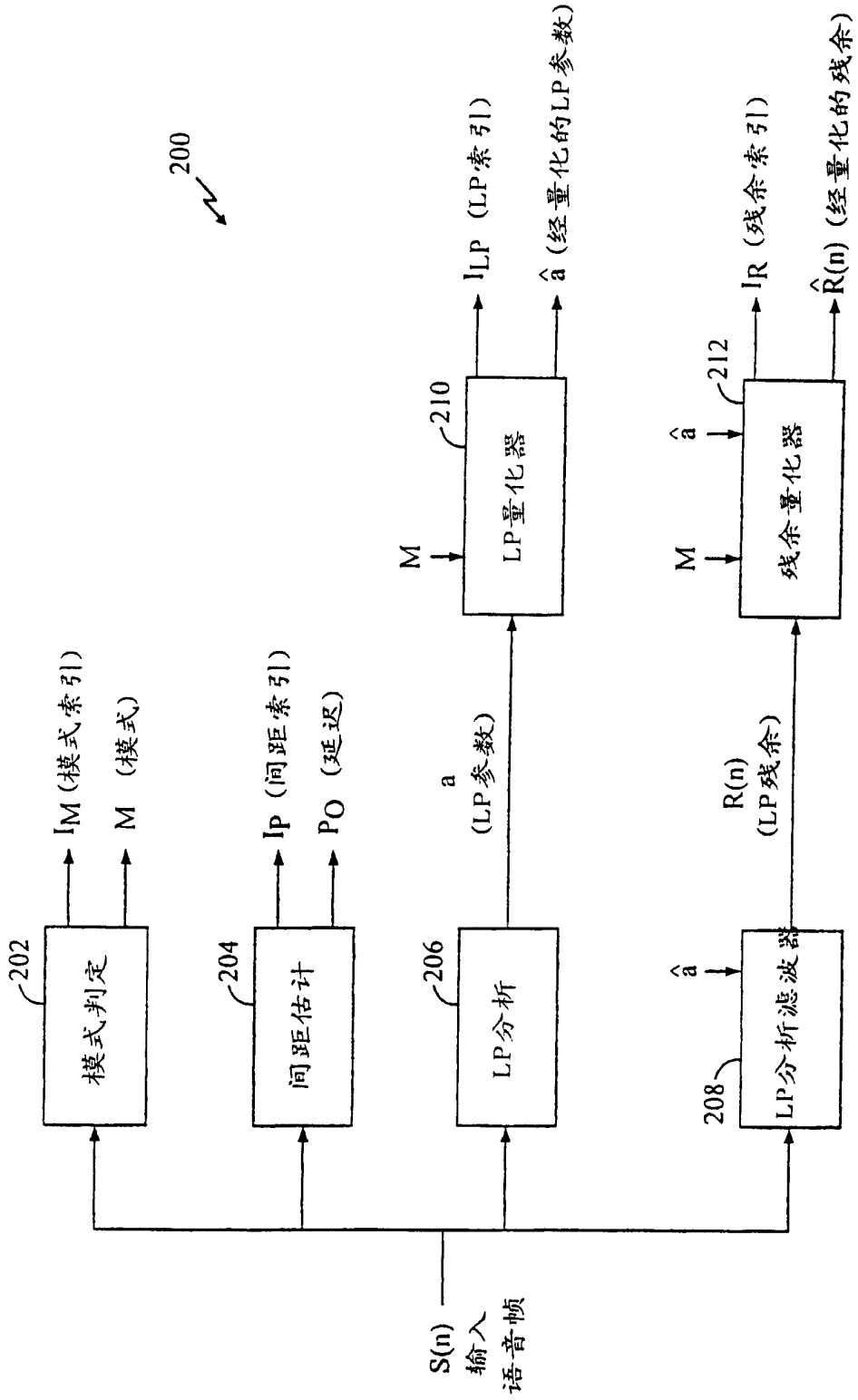


图 2

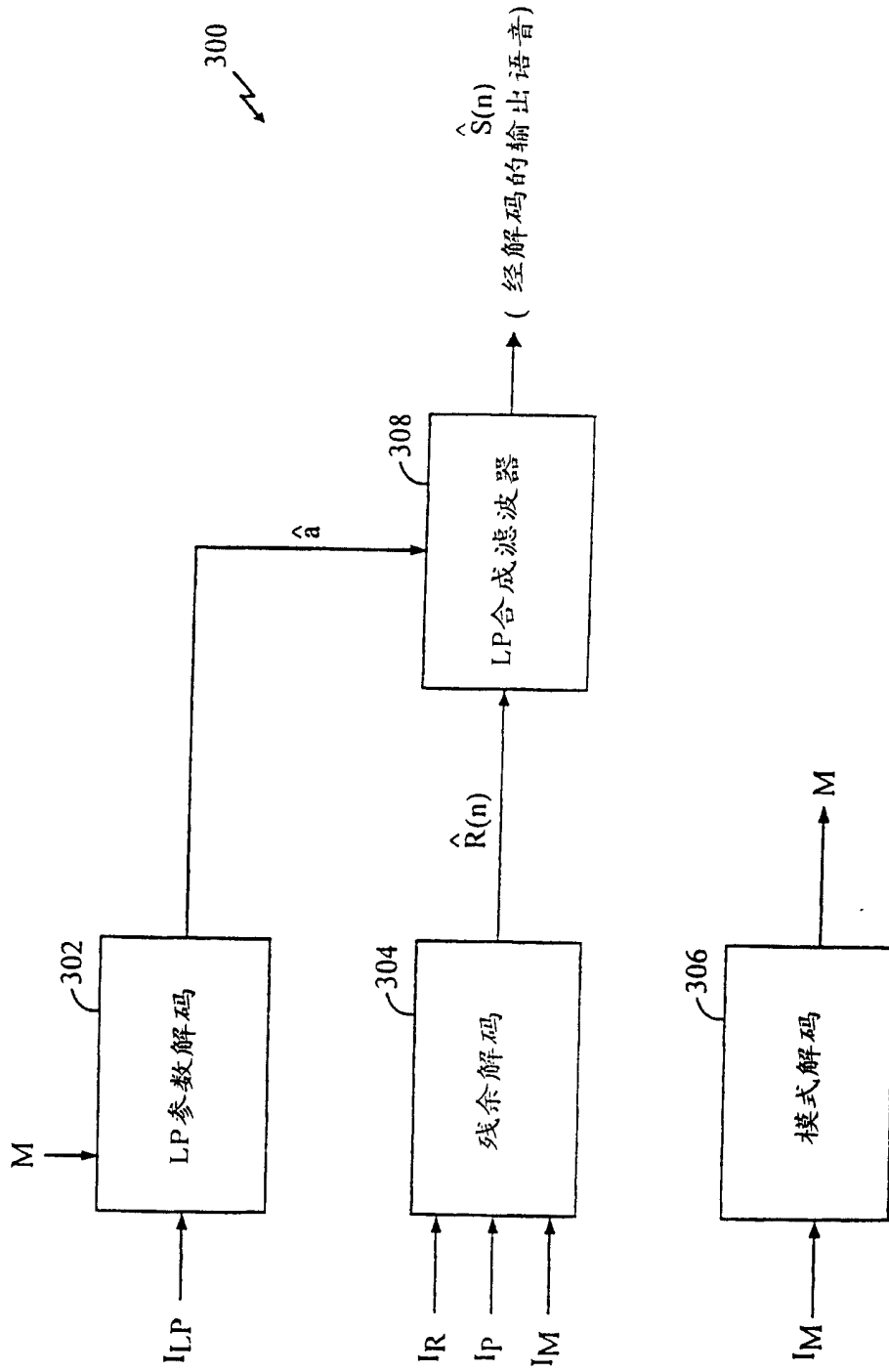
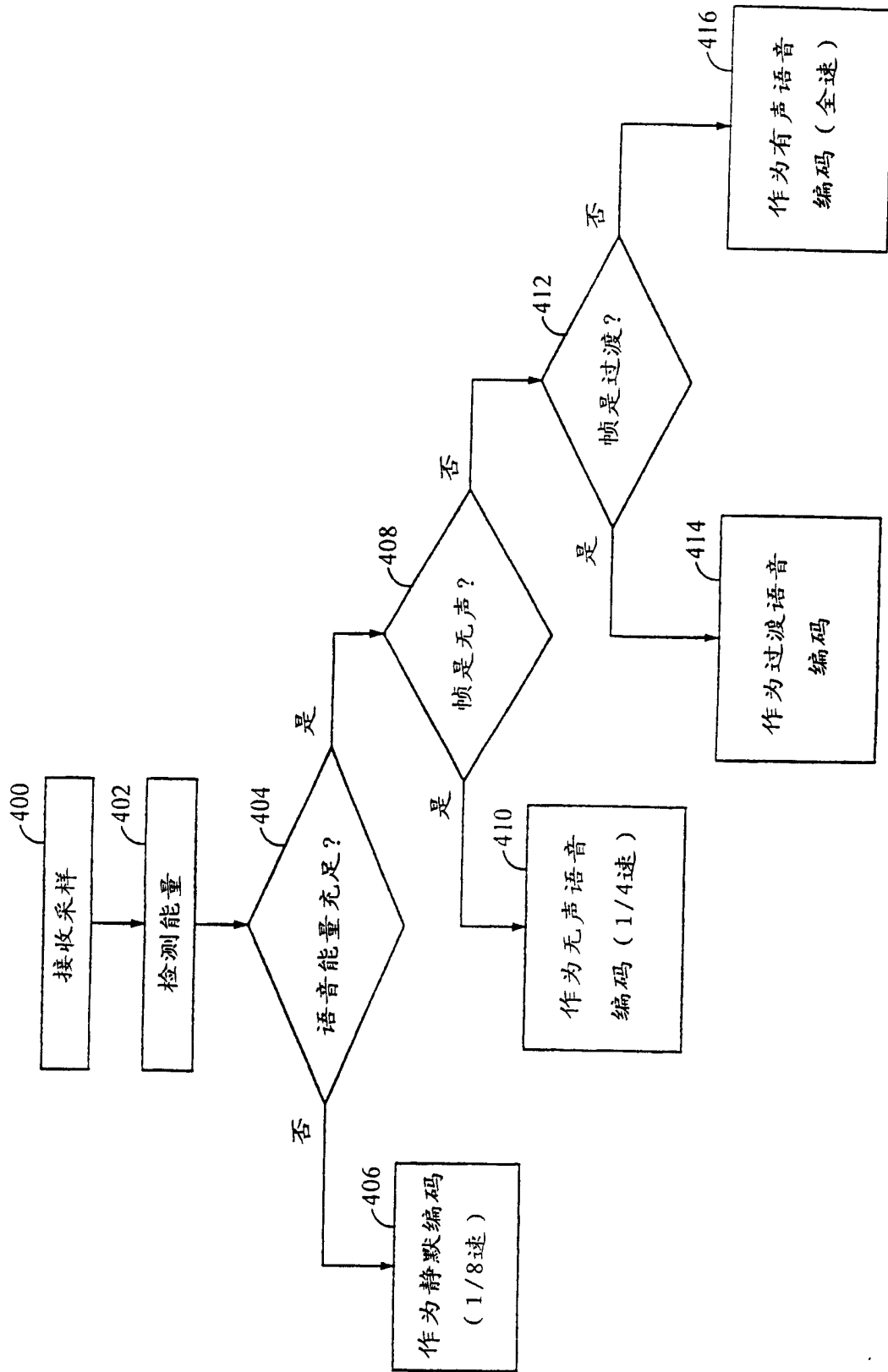


图 3



图

4

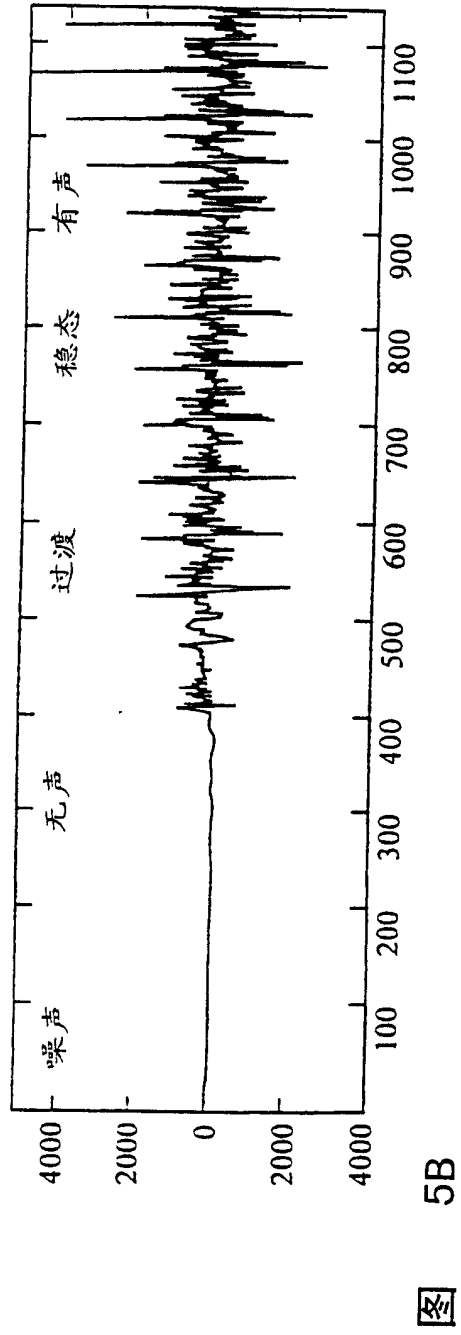
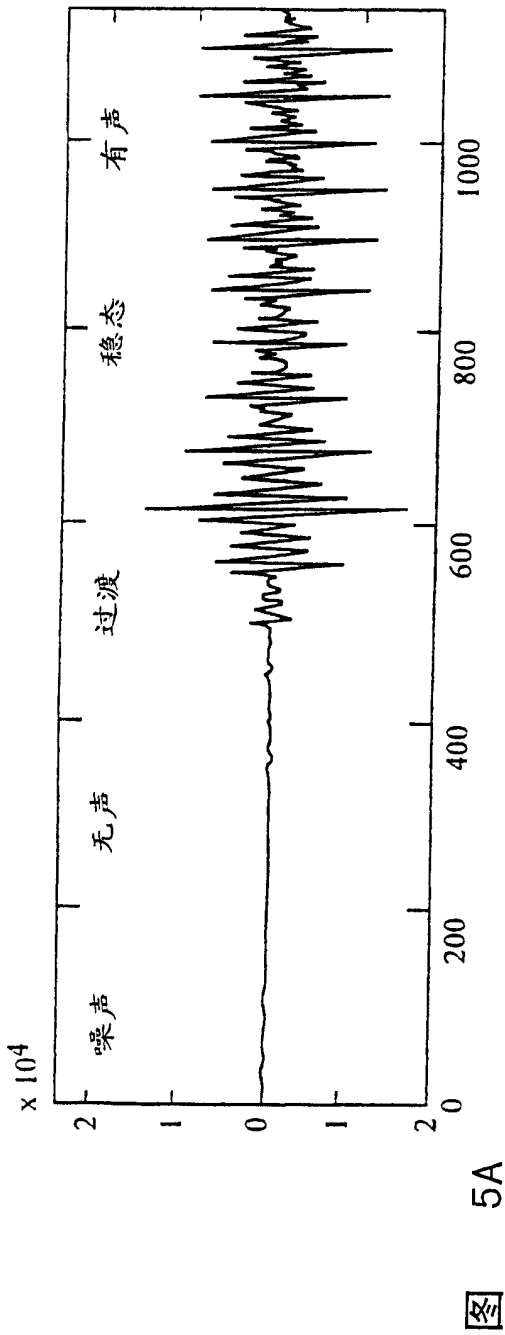


图 5

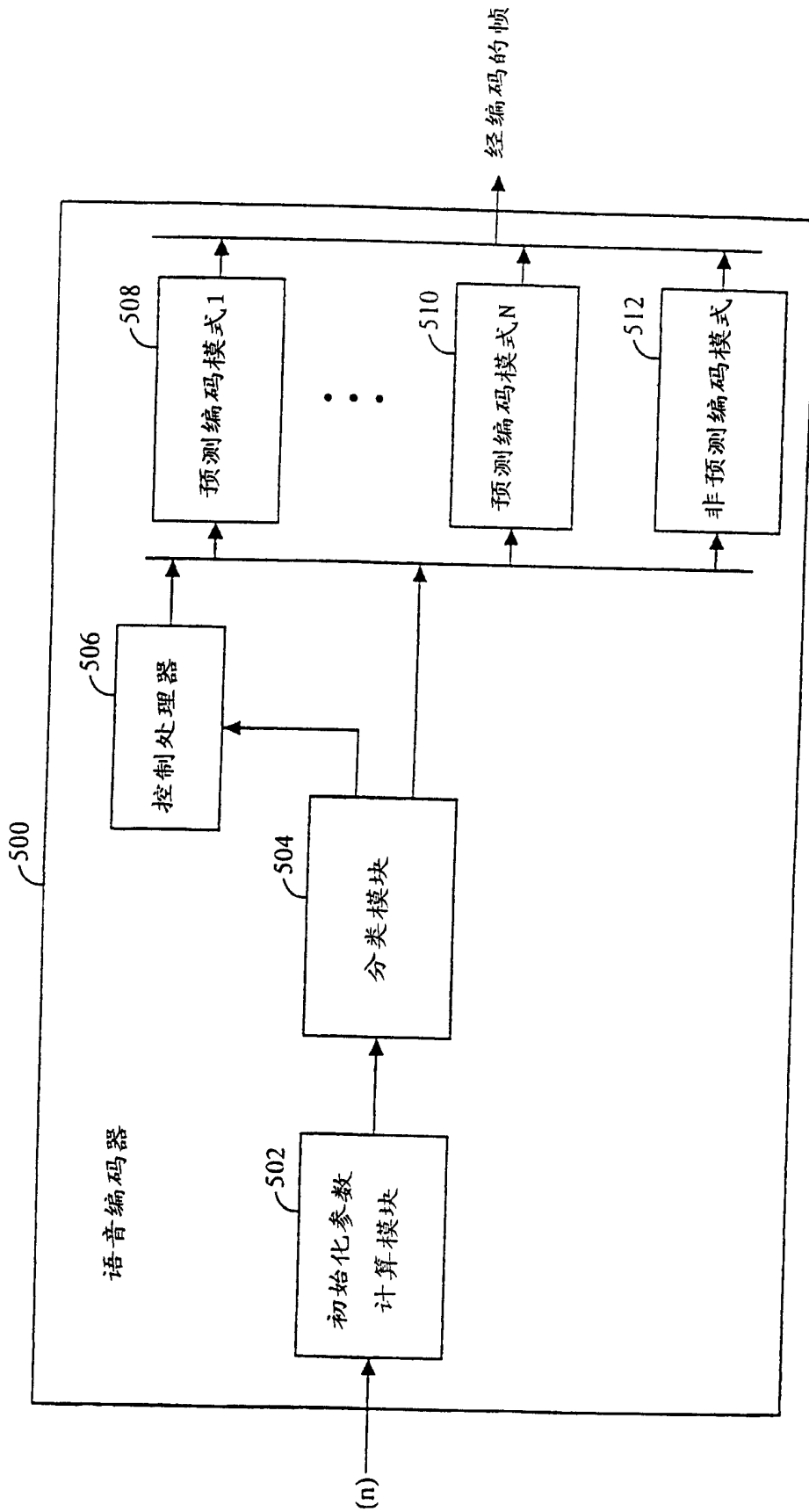


图 6

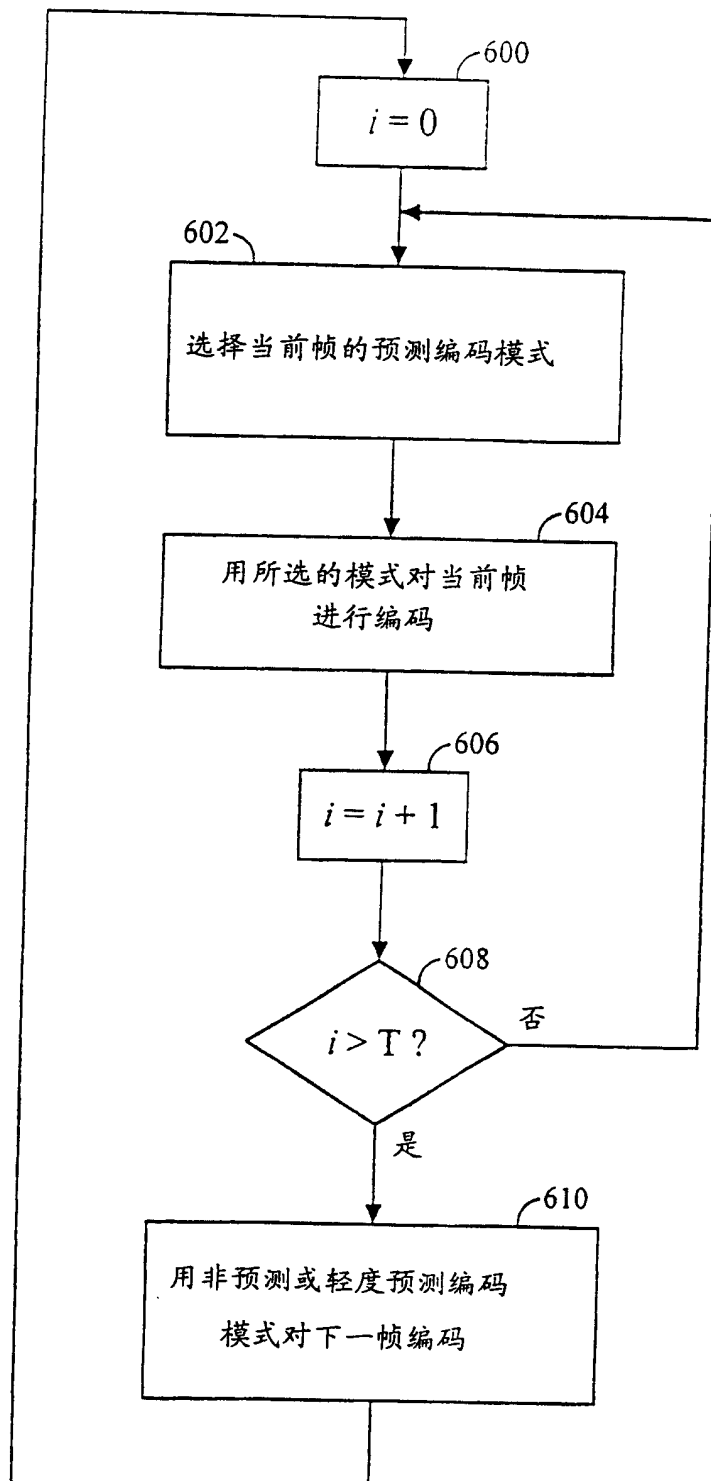


图 7