

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

G10L 19/00

G10L 19/14



[12] 发明专利说明书

[21] ZL 专利号 00810597.9

[45] 授权公告日 2004 年 8 月 11 日

[11] 授权公告号 CN 1161749C

[22] 申请日 2000.7.19 [21] 申请号 00810597.9

[30] 优先权

[32] 1999. 7. 19 [33] US [31] 09/356,493

[86] 国际申请 PCT/US2000/019670 2000.7.19

[87] 国际公布 WO2001/006490 英 2001.1.25

[85] 进入国家阶段日期 2002.1.18

[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·德加科

审查员 杨 叁

[74] 专利代理机构 上海专利商标事务所

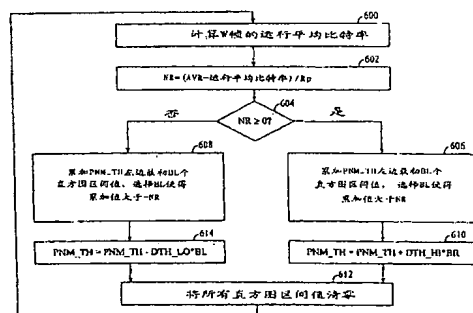
代理人 钱慰民

权利要求书 4 页 说明书 14 页 附图 8 页

[54] 发明名称 保持语言编码中目标比特率的方法和装置

[57] 摘要

保持语言编码器中目标比特率的方法和装置包括语言编码器，用于以预选编码率对帧编码，计算预定个数编码帧的运行平均比特率；从预定目标平均比特率中减去运行平均比特率，并用差值除以预选编码率。如果商值是负的，累加预定个数小于当前性能阈值的语言编码器性能阈值的可能具体值，累加数大于商值的绝对值。如果商值是正的，累加预定个数大于当前性能阈值的语言编码器性能阈值的可能具体值，累加数大于商值。



1. 在以变化的编码率将多个帧编码的语言编码器中, 保持语言编码器目标平均比特率的方法, 其特征在于, 包括以下步骤:

以预选编码率将帧编码;

计算预定个数编码帧的运行平均比特率;

从预定目标平均比特率中减去运行平均比特率, 以获得差值;

用差值除以预选的编码率以获得商值;

如果商值小于零, 累加第一预定个数小于当前性能阈值的语言编码器性能阈值的可能具体值, 以产生第一累加值, 选择语言编码器性能阈值具体值的预定个数, 使得第一累加值大于商值的绝对值;

如果商值小于零, 从当前性能阈值中减去两项的乘积, 其中一项是每个语言编码器性能阈值具体值的减小量, 另一项是语言编码器性能阈值具体值的第一预定个数, 以获得新的性能阈值;

如果商值大于或等于零, 累加第二预定个数大于当前性能阈值的语言编码器性能阈值的可能具体值, 以产生第二累加值, 选择语言编码器性能阈值具体值的预定个数, 使得第二累加值大于商值; 和

如果商值大于或等于零, 将每个语言编码器性能阈值具体值的增加量和语言编码器性能阈值具体值的第二预定个数的乘积与当前性能阈值相加, 以获得新的性能阈值。

2. 如权利要求 1 所述的方法, 其特征在于, 还包括以下步骤, 将语言编码器性能和预定性能测量值比较, 并且如果帧的语言编码器性能低于当前性能阈值, 则调节帧的预选编码率。

3. 如权利要求 2 所述的方法, 其特征在于, 调节步骤包括提高帧的编码率。

4. 如权利要求 2 所述的方法, 其特征在于, 在编码步骤期间, 还包括以下步骤:

对于每个小于当前性能阈值的语言编码器性能阈值的具体值, 从当前性能阈值中减去两项的乘积, 其中一项是每个语言编码器性能阈值具体值的减小量, 另一项是语言编码器性能阈值的具体值和当前性能阈值之间的语言编码器性能阈值具体值数加 1, 并将语言编码器性能阈值的具体值设置成等于该减法的结果;

对于每个大于当前性能阈值的语言编码器性能阈值的具体值, 将两项的乘积

加上当前性能阈值，其中一项是每个语言编码器性能阈值具体值的增加量，另一项是语言编码器性能阈值的具体值和当前性能阈值之间的语言编码器性能阈值具体值数加 1，并将语言编码器性能阈值的具体值设置成等于该加法的结果；和

将对应于当前语言编码器性能的语言编码器性能阈值的具体值增加 1。

5. 如权利要求 1 所述的方法，其特征在于，还包括从帧的开环分类中获得预选编码率的步骤。

6. 如权利要求 1 所述的方法，其特征在于，还包括使当前性能阈值限制于最大值的步骤。

7. 如权利要求 1 所述的方法，其特征在于，还包括使当前性能阈值限制于最小值的步骤。

8. 如权利要求 1 所述的方法，其特征在于，还包括对每个语言编码器性能阈值具体值的减小量和每个语言编码器性能阈值具体值的增加量赋初值的步骤。

9. 如权利要求 1 所述的方法，其特征在于，还包括在执行加或减步骤之后，对所有语言编码器性能阈值的具体值清零的步骤。

10. 如权利要求 1 所述的方法，其特征在于，帧是语言帧。

11. 如权利要求 1 所述的方法，其特征在于，帧是线性预测剩余帧。

12. 如权利要求 1 所述的方法，其特征在于，语言编码器位于无线通信系统的用户单元中。

13. 一种语言编码器，其特征在于，包括：

分析模块，它被构造成分析多个帧；和

量化模块，它耦合到分析模块，并被构造成将分析模块产生的帧参数编码，其中量化模块还被进一步构造成：

以预选编码率将帧编码；

计算预定个数编码帧的运行平均比特率；

从预定目标平均比特率中减去运行平均比特率，以获得差值；

用差值除以预选的编码率以获得商值；

累加第一预定个数可能小于当前性能阈值的语言编码器性能阈值的具体值，以产生第一累加值，选择语言编码器性能阈值具体值的预定个数，使得第一累加值大于商值的绝对值；

从当前性能阈值中减去两项的乘积，其中一项是每个语言编码器性能阈值具体值的减小量，另一项是语言编码器性能阈值具体值的第一预定个数，如果商值

小于零，获得一新的性能阈值；

累加第二预定个数可能大于当前性能阈值的语言编码器性能阈值的具体值，以产生第二累加值，选择语言编码器性能阈值具体值的预定个数，使得第二累加值大于商值；和

将每个语言编码器性能阈值具体值的增加量和语言编码器性能阈值具体值的第二预定个数的乘积与当前性能阈值相加，如果商值大于或等于零，获得一新的性能阈值。

14. 如权利要求 13 所述的语言编码器，其特征在于，量化模块还被构造成将语言编码器性能和预定性能测量值比较，并且如果帧的语言编码器性能低于当前性能阈值，就调节帧的预选编码率。

15. 如权利要求 14 所述的语言编码器，其特征在于，以增加调节编码率。

16. 如权利要求 14 所述的语言编码器，其特征在于，量化模块还被构造成：在帧的编码期间，对于每个小于当前性能阈值的语言编码器性能阈值的具体值，从当前性能阈值中减去两项的乘积，其中一项是每个语言编码器性能阈值具体值的减小量，另一项是语言编码器性能阈值的具体值和当前性能阈值之间的语言编码器性能阈值具体值数加 1，并将语言编码器性能阈值的具体值设置成等于该减法的结果；

在帧的编码期间，对于每个大于当前性能阈值的语言编码器性能阈值的具体值，将两项的乘积加上当前性能阈值，其中一项是每个语言编码器性能阈值具体值的增加量，另一项是语言编码器性能阈值的具体值和当前性能阈值之间的语言编码器性能阈值具体值数加 1，并将语言编码器性能阈值的具体值设置成等于该加法的结果；和

在帧的编码期间，将对应于当前语言编码器性能的语言编码器性能阈值的具体值增加 1。

17. 如权利要求 13 所述的语言编码器，其特征在于，量化模块还被构造成从帧的开环分类中获得预选编码率。

18. 如权利要求 13 所述的语言编码器，其特征在于，量化模块还被构造成使当前性能阈值限制于最大值。

19. 如权利要求 13 所述的语言编码器，其特征在于，量化模块还被构造成使当前性能阈值限制于最小值。

20. 如权利要求 13 所述的语言编码器，其特征在于，量化模块还被构造成

对每个语言编码器性能阈值具体值的减小量和每个语言编码器性能阈值具体值的增加量赋初值。

21. 如权利要求 13 所述的语言编码器, 其特征在于, 量化模块还被构造成在调节当前性能阈值之后, 将所有语言编码器性能阈值的具体值清零。

22. 如权利要求 13 所述的语言编码器, 其特征在于, 帧是语言帧。

23. 如权利要求 13 所述的语言编码器, 其特征在于, 帧是线性预测剩余帧。

24. 如权利要求 13 所述的语言编码器, 其特征在于, 语言编码器位于无线通信系统的用户单元中。

保持语言编码器中目标比特率的方法和装置

发明背景

I. 发明领域

本发明一般涉及语言处理领域，尤其涉及保持语言编码器中目标比特率的方法和装置。

II. 背景技术

用数字技术发送语音已非常普及，尤其在长距离和数字无线电话应用中。这反过来产生了对确定在信道上发送的最小信息量的兴趣，同时又保持重建语言的感觉质量。如果通过简单地采样和数字化发送语言，那么需要大约每秒 64 千比特 (kbps) 的数据率，以获得传统模拟电话的语言质量。然而，通过使用语言分析，接着通过适当的编码、发送、并在接收机端再合成，可以大大降低数据率。

用于压缩语言的装置在电信的众多领域中可找到其应用。典型的领域是无线通信。无线通信领域具有很多应用，包括如无绳电话、寻呼、无线局域环路、如蜂窝和 PCS 电话系统的无线电话，移动因特网协议 (IP) 电话，和卫星通信系统。尤其重要的应用是移动用户的无线电话。

已经发展了各种空中接口用于无线通信系统，包括如频分多址 (FDMA)、时分多址 (TDMA)、和码分多址 (CDMA)。在与之的连接中，建立了各种国内和国际的标准，包括如高级移动电话业务 (AMPS)，全球移动通信系统 (GSM)、和临时标准 (IS-95)。典型的无线电话通信系统是码分多址 (CDMA) 系统。电信工业协会 (TIA) 和其它公知的标准团体颁布了 IS-95 标准及其衍生物 IS-95A、ANSI J-STD-008、IS-95B、提议的第三代标准的 IS-95C 和 IS-2000 等等 (这里统称为 IS-95)，以规定 CDMA 空中接口在蜂窝或 PCS 电话通信系统中的使用。美国专利号 5,103,459 和 4,901,307 中描述了根据使用 IS-95 标准实质性构造的典型无线通信系统，这两个申请都转让给本发明的受让人，并通过引用全面结合于此。

使用通过提取关于人类语言产生模型的参数以压缩语言技术的装置称为语言编码器。语言编码器将输入语言信号分成时间块或分析帧。语言编码器通常包括编码器和解码器。编码器分析输入语言帧，以提取某些相关的参数，然后将参数

量化成二进制表示，也就是一组比特或二进制数据包。在通信信道上将数据包从接收机发送到解码器。解码器处理数据包，将它们去量化以产生参数，并使用去量化参数再合成语言帧。

语言编码器的功能是通过除去语言中的所有固有冗余将数字化的语言信号压缩成低比特率的信号。通过用一组参数表示输入语言帧并使用量化用一组比特表示参数，以实现数字压缩。如果输入语言帧具有比特数 N_i 且语言编码器产生的数据包具有比特数 N_o ，那么语言编码器获得的压缩因数为 $C_r = N_i/N_o$ 。问题是在获得目标压缩因数的同时要保持解码语言的高语音质量。语言编码器的性能依赖于（1）语言模型、或上述分析和合成过程的组合执行得多好，（2）以目标比特率每帧 N_o 比特执行的参数量化过程执行得多好。因此，语言模型的目标是用每个帧一小组参数获取语言信号的本质，或目标语言质量。

也许语言编码器设计中最重要的是搜索一组好的参数（包括向量），以描述语言信号。一组好的参数要求有较低的系统带宽，用于感觉上精确的语言信号的重建。音调、信号功率、频谱包络（或共振峰），振幅和相位谱都是语言编码参数的实例。

语言编码器可以是时域编码器，它通过使用较高的时间分辨处理每次对小片段语言（通常是 5 毫秒（ms）子帧）编码，以试图获取时域语言波形。对于每个子帧，通过本领域所熟知的各种搜索算法从码本空间找出高精度的表示。另外，语言编码器可以用频域编码器实现，它试图用一组参数（分析）获取输入语言帧的短期语言频谱，并使用相应的合成过程从频谱参数中重建语言波形。参数量化器根据 A. Gersho & R. M. Gray, Vector Quantization and Signal Compression (1992) 中描述的熟知的量化技术通过用存储的代码向量表示法表示参数，以保存参数。

公知的时域语言编码器是 L. B. Rabiner & R. W. Schafer, Digital Processing of Speech Signal 396-453 (1978) 中描述的代码受激线性预测编码器（CELP），它通过引用全面结合于此。在 CELP 编码器中，线性预测（LP）分析除去语言信号中的短期相关或冗余，它找出短期共振峰滤波器的系数。将短期预测滤波器施加到输入语言帧产生 LP 剩余信号，用长期预测滤波器参数和随后的随机码本将该信号进一步模型化和量化。因此，CELP 编码将时域语言波形的编码任务分成 LP 短期滤波器系数编码和 LP 剩余编码的分离的任务。可以以固定的速率（也就是对每个帧使用相同的比特数 N_o ）或变化的速率（其中不同比特率用于不同类型的帧内

容) 执行时域编码。变化速率编码器试图只使用将编译码参数编码以达到目标质量水平所需的比特量。美国专利号 5, 414, 796 中描述了典型的变化速率 CELP 编码器, 该申请转让给了本发明的受让人, 并通过引用全面结合于此。

时域编码器如 CELP 编码器通常依赖于每帧高的比特数 N_0 , 以保持时域语言波形的精确度。这种编码器通常传送由相对大的每帧比特数 N_0 (如 8kbps 或更大) 提供的较好语言质量。然而, 在低比特率 (4kbps 或更低) 时, 由于有限的可获得比特数, 时域编码器就不能保持高质量和稳健的性能。在低比特率时, 有限的码本空间箝制了传统时域编码器的波形匹配能力, 编码器在较高速率的工业应用中已成功发展。因此, 尽管随着时间推移而进行了改进, 但是以低比特率工作的很多 CELP 编码系统遭受感觉上严重的失真, 通常表征为噪声。

当前涌现了研究兴趣和较强的工业需要以发展在中到低比特率 (也就是在 2.4 到 4kbps 范围内或更小) 上工作的高质量的语言编码器。应用领域包括无线电话、卫星通信、因特网电话、各种多媒体和话音流应用、话音邮件和其它语音存储系统。驱动力是高容量的需要和在包丢失情况下的稳健性能的要求。各种近来的语言编码标准化努力是另一直接驱动力, 用于推进低速率语言编码算法的研究和发展。低速率语言编码器在每个允许的应用带宽中产生更多的信道或用户, 而且与适当的信道编码附加层耦合的低速率语言编码器能够适应编码器规范的总比特预算, 并能在信道误差条件下传送较稳健的性能。

以低比特率将语言高效编码的有效技术是多模编码。1998 年 12 月 21 日提交的美国申请序号 09/217, 341 题为 “VARIABLE RATE SPEECH CODING” 的申请中描述了典型的多模编码技术, 该申请转让给了本发明的受让人, 并通过引用全面结合于此。传统的多模编码对于不同类型的输入语言帧使用不同的模式或编码-解码算法。定制每种模式或编码-解码过程以最佳地用最有效地方式表示某些类型的语言片段, 如有声语言、无声语言、过渡语言 (例介于有声和无声之间)、和背景噪声 (无语言)。外部、开环模式判定机构测试输入语言帧, 并确定对该帧施加何种模式。通常通过从输入帧中提取多个参数, 估计关于某些时频特征的参数, 并使模式判定基于该估计, 以执行开环模式判定。因此, 进行模式判定无需事先知道输出语言的精确情况, 也就是在语言质量或其它性能测量方面输出语言和输入语言有多接近。

以大约 2.4kbps 速率工作的编码系统一般在本质上是参数式的。也就是说, 这种编码系统通过在规则时间间隔上发送表示语言信号音调周期和频谱包络 (或

共振峰)的参数而进行工作。这些所谓参数编码器的实例是LP声音编码器系统。

LP声音编码器用每个音调周期的单脉冲模型化有声语言信号。可以增广该基本技术,包括还另外发送关于频谱包络的信息。虽然LP声音编码器一般提供合理的性能,但是它们可能传入感觉上严重的失真,典型的表征为蜂音。

近几年来,形成了混合波形编码器和参数编码器的编码器。这些所谓混合编码器的实例是原型波形内插(PWI)语言编码系统。PWI编码系统也熟知为原型音调周期(PPP)语言编码器。PWI编码系统提供将有声语言编码的有效方法。PWI的基本概念是在固定时间间隔提取典型音调周期(原型波形),发送它的描述,并通过在原型波形中内插而重建语言信号。PWI方法可用于对LP剩余信号或语言信号操作。1998年12月21日提交的美国申请序号09/217,494题为“PERIODIC SPEECH CODING”的申请中描述了典型的PWI或PPP语言编码器,该申请转让给了本发明的受让人,并通过引用全面结合于此。美国专利号5,884,253中和W.Bastiaan Kleijn & Wolfgang Granzow, Methods for Waveform Interpolation in Speech Coding, in 1 Digital Signal Processing 215-230(1991)中描述了其它PWI或PPP语言编码器。

传统的低比特率、变化速率语言编码器根据帧的能量使用开环编码模式判定,以确定何时将较低的编码率切换成较高的编码率。这允许语言编码器使用存在的不同分类的语言,并以不同的速率将它们编码。然而,以开环分类确定的速率编码可能导致使特定帧具有较差或普通的质量。因此,改进开环判定的效率是有利的。期望使用质量估计以改变(也就是如果必要就增加)给定帧的编码率。然而,提高帧的编码率将改变(增加)语言编码器的平均编码率。因此提供语言编码器保持固定平均比特率,同时允许一帧一帧的基础上的编码率偏离开环分类所确定的值是更有利的。还期望指定语言编码器的目标平均速率。保持语言编码器的目标总比特率是更有利的。因此,需要语言编码器,用于精细化带有闭环判定过程的编码模式判定,以给出最佳的话音质量,并保持目标编码比特率。

发明内容

本发明是针对语言编码器,用于精细化带有闭环判定过程的编码模式判定,以给出最佳的话音质量,并保持目标编码比特率。因此,本发明的一方面,在被构造成以变化的编码率将多个帧编码的语言编码器中,保持语言编码器目标平均比特率的方法最好包括以下步骤:以预选编码率将帧编码;计算预定个数编码帧

的运行平均比特率；从预定目标平均比特率中减去运行平均比特率，以获得差值；用差值除以预选的编码率以获得商值；如果商值小于零，累加第一预定个数可能小于当前性能阈值的语言编码器性能阈值的具体值，以产生第一累加值，选择语言编码器性能阈值具体值的预定个数，使得第一累加值大于商值的绝对值；如果商值小于零，从当前性能阈值中减去两项的乘积，其中一项是每个语言编码器性能阈值具体值的减小量，另一项是语言编码器性能阈值具体值的第一预定个数，以获得新的性能阈值；如果商值大于或等于零，累加第二预定个数可能大于当前性能阈值的语言编码器性能阈值的具体值，以产生第二累加值，选择语言编码器性能阈值出现值的预定个数，使得第二累加值大于商值；并且如果商值大于或等于零，将每个语言编码器性能阈值具体值的增加量和语言编码器性能阈值的第二预定出现个数的乘积与当前性能阈值相加，以获得新的性能阈值。

本发明的另一方面，语言编码器最好包括分析模块，用于分析多个帧；和量化模块，耦合到分析模块用于将分析模块产生的帧参数编码，其中量化模块还被构造成以预选编码率将帧编码；计算预定个数编码帧的运行平均比特率；从预定目标平均比特率中减去运行平均比特率，以获得差值；用差值除以预选的编码率以获得商值；累加第一预定的可能小于当前性能阈值的语言编码器性能阈值的具体值个数，以产生第一累加值，选择语言编码器性能阈值具体值的预定个数，使得第一累加值大于商值的绝对值；从当前性能阈值中减去两项的乘积，其中一项是每个语言编码器性能阈值具体值的减小量，另一项是语言编码器性能阈值具体值的第一预定个数，如果商值小于零，以获得新的性能阈值；累加第二预定个数可能大于当前性能阈值的语言编码器性能阈值的具体值，以产生第二累加值，选择语言编码器性能阈值具体值的预定个数，使得第二累加值大于商值；并将每个语言编码器性能阈值具体值的增加量和语言编码器性能阈值具体值的第二预定个数的乘积与当前性能阈值相加，如果商值大于或等于零，以获得新的性能阈值。

附图说明

图 1 是无线电话系统的框图。

图 2 是语言编码器在每个终端处终止的通信信道的框图。

图 3 是编码器的框图。

图 4 是解码器的框图。

图 5 是表示语言编码判定过程的流程图。

图 6A 是语言信号振幅相对于时间的图像；图 6B 是线性预测 (LP) 剩余振幅相对于时间的图像。

图 7 是原型音调周期 (PPP) 语言编码器的框图。

图 8 是表示语言编码器执行的算法步骤的流程图，如图 7 的语言编码器，将闭环编码性能测量应用于每个被编码帧，并同时保持语言编码器的目标平均比特率。

图 9 是表示语言编码器执行的算法步骤的流程图，用于在语言帧的编码期间更新直方图区间的值。

较佳实施例的详细描述

这里以下所述的典型实施例存在于被构造成使用 CDMA 空中接口的无线电话通信系统中。然而，本领域熟练的技术人员应该理解使用本发明特征的子采样方法和装置可以存在于使用本领域熟练技术人员所熟知的较宽范围技术的任何各种通信系统中。

如图 1 所示，CDMA 无线电话系统一般包括多个移动用户单元 10、多个基站 12、基站控制器 (BSC) 14，和移动交换中心 (MSC) 16。MSC 16 被构造成与传统公共交换电话网 (PSTN) 18 连接。MSC 16 还被构造成与 BSC 14 连接。BSC 14 还通过回传线路耦合到基站 12。回传线路可以被构造成支持几个已知接口中的任何一个，如 E1/T1、ATM、IP、PPP、Frame Relay、HDSL、ADSL 或 xDSL。应该理解系统中多于两个的 BSC 14。每个基站 12 最好包括至少一个扇区 (未图示)，每个扇区包括一全向天线或从基站 12 指特定径向方向的天线。另一种情况是，每个扇区可以包括两个天线用于分集接收。每个基站 12 最好被设计成支持多个频率分配。扇区的相交和频率分配可以被称为 CDMA 信道。基站 12 也熟知为基站收发机子系统 (BTS) 12。另一种情况是，“基站”可用于工业中统称为 BSC 14 和一个或多个 BTS 12。BTS 12 也可以表示为“区站”12。另一种情况是，给定 BTS 12 的各个扇区可称为区站。移动用户单元 10 通常是蜂窝或 PCS 电话 10。系统最好被构造成根据 IS-95 标准使用。

在蜂窝电话系统的典型工作中，基站 12 从移动单元组 10 接收反向链路信号组。移动单元 10 进行电话呼叫或其它通信。特定基站 12 接收到的每个反向链路信号在该基站 12 中被处理。所得的数据被发送到 BSC 14。BSC 14 提供呼叫源定位和移动管理功能，包括基站 12 之间软越区切换的控制。BSC 14

还将接收到的数据发送到 MSC 16, MSC 16 为与 PSTN 18 接口提供附加的路由服务。类似地, PSTN 18 和 MSC 16 连接, MSC 16 和 BSC 14 连接, BSC 14 依次控制基站 12 发送前向链路信号组到移动单元组 10。

在图 2 中, 第一编码器 100 接收数字化的语言样本 $s(n)$ 并将样本 $s(n)$ 编码, 用于在传输介质 102、或通信信道 102 上发送到第一解码器 104。解码器 104 将编码语言样本解码, 并合成输出语言信号 $S_{\text{SYNTH}}(n)$ 。为了在反方向上发送, 第二编码器 106 将对通信信道 108 上传输的数字化语言样本 $s(n)$ 编码。第二解码器 110 接收并将编码语言样本解码, 产生合成的输出语言信号 $S_{\text{SYNTH}}(n)$ 。

语言样本 $s(n)$ 表示根据本领域中各种熟知的任何方法数字化并量化的语言信号, 这些方法例如包括脉冲编码调制 (PCM)、压扩 μ 律、或 A 律。如本领域中所熟知的, 语言样本 $s(n)$ 组合成输入数据帧, 其中每个帧包括预定个数的数字化语言样本 $s(n)$ 。在典型实施例中, 使用 8kHz 的采样速率, 每个 20ms 的帧包括 160 个样本。在以下描述的实施例, 数据传输速率最好在帧到帧的基础上从 13.2kbps (全速率) 变化到 6.2kbps (半速率) 到 2.6kbps (四分之一速率) 到 1kbps (八分之一速率)。数据传输速率的变化是有利的, 因为对于包含相对较少语言信息的帧可以选用较低的比特率。本领域熟练的技术人员能够理解还可以使用其它采样速率和数据传输速率。

第一编码器 100 和第二解码器 110 一同组成第一语言编码器。语言编码器可用于发送语言信号的任何通信装置, 包括如以上参考图 1 描述的用户单元、BTS 或 BSC。类似地, 第二编码器 106 和第一解码器 104 一同组成第二语言编码器。本领域熟练的技术人员应该理解语言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、或任何传统可编程软件模块和微处理器实现。软件模块可驻留在 RAM 存储器、快闪存储器、寄存器或本领域中所熟知的任何其它形式的可写存储媒介中。另一种情况是, 任何传统的处理器、控制器或状态机械都可以替代微处理器。美国专利号 5,727,123 中中和 1994 年 2 月 16 日提交的美国申请序号 08/197,417 题为 “VOCODER ASIC” 的申请中都描述了专门为语言编码设计的典型 ASIC, 这两个申请都转让给了本发明的受让人, 并通过引用全面结合于此。

在图 3 中, 可用于语言编码器的编码器 200 包括模式判定模块 202、音调估计模块 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。输入语言帧 $s(n)$ 提供给模式判定模块 202、音调估计模块 204、LP 分析模块 206、LP 分析滤波器 208。模式判定模块 202 根据每个输入语言帧 $s(n)$ 的周期、能量信噪比 (SNR) 或过零速率以及其它特征产生模式索引 I_M 和模式 M 。美国专利号 5,911,128 中描述了根据周期将语言帧分类的各种方法,该申请转让给了本发明的受让人,并通过引用全面结合于此。这种方法还结合入电信工业协会工业临时标准 TIA/EIA IS-127 和 TIA/EIA IS-733。上述的美国申请序号 09/217,341 中还描述了典型的模式判定方案。

音调估计模块 204 根据每个输入语言帧 $s(n)$ 产生音调索引 I_P 和滞后值 P_0 。LP 分析模块 206 执行每个输入语言帧 $s(n)$ 的线性预测分析,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 还接收模式 M ,以此执行模式依赖方式的量化过程。LP 量化模块 210 产生 LP 索引 I_{LP} 和量化 LP 参数 $\hat{a}$ 。除了输入语言帧 $s(n)$ 外 LP 分析滤波器 208 还接收量化 LP 参数 $\hat{a}$ 。LP 分析滤波器 208 产生 LP 剩余信号 $R[n]$,它表示输入语言帧 $s(n)$ 和根据量化线性预测参数 $\hat{a}$ 重建的语言之间的误差。LP 剩余 $R[n]$ 、模式 M 和量化 LP 参数 $\hat{a}$ 被提供给剩余量化模块 212。根据这些值,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,可用于语言编码器的解码器 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 接收并将模式索引 I_M 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 将接收到的值解码,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 将接收到的值解码,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 $\hat{a}$ 被提供给 LP 合成滤波器 308,它从中合成解码输出信号 $\hat{s}[n]$ 。

本领域中熟知图 3 中编码器 200 和图 4 中解码器 300 中各种模块的操作和实施,并且在上述美国专利号 5,414,196 和 L. B. Rabiner & R. W. Schafer, Digital Processing of Speech Signals 396-453(1978)中对其进行了描述。

如图 5 中流程图所示,根据一个实施例的语言编码器遵循处理语言样本用于发送的一组步骤。在步骤 400 中,语言编码器接收连续帧中的语言信号的数字样本。通过接收特定的帧,语言编码器进入步骤 402。在步骤 402 中,语言编码器检测帧的能量。能量是帧中语言活动性的衡量。通过将数字化语言样本振幅的平方求和并将所得的能量和阈值比较,执行语言检测。在一个

实施例中，阈值根据背景噪声电平的变化而改变。上述美国专利号 5,414,796 中描述了典型的可变阈值语言活动性检测器。一些无声的语言声很可能是极低能量样本，它可能被当做背景噪声而错误编码。为了防止这种情况的发生，使用低能量样本的频谱倾斜，以区别无声语言和背景噪声，如上述美国专利号 5,414,796 中所描述的。

在检测帧的能量之后，语言编码器执行步骤 404。在步骤 404 中，语言编码器确定检测到的帧能量是否足以将帧分类为包含语言信息。如果检测到的帧能量低于预定的阈值等级，那么语言编码器执行步骤 406。在步骤 406 中，语言编码器将帧作为背景噪声（也就是无语言或静音）编码。在一个实施例中，以 1/8 速率或 1kbps 将背景噪声帧编码。如果在步骤 404 中检测到的帧能量符合或超过预定的阈值等级，那么帧被分类为语言，语言编码器执行步骤 408。

在步骤 408 中，语言编码器确定帧是否为无声语言，也就是语言编码器测试帧的周期性。各种已知的周期确定方法包括如使用零交叉和使用归一化自相关函数（NACF）。尤其，上述美国专利号 5,911,128 中和美国申请序号 09/217,341 中描述了使用零交叉和 NACF 检测周期。此外，以上用于区别有声语言和无声语言的方法被结合入电信工业协会临时标准 TIA/EIA IS-127 和 TIA/EIA IS-733。如果在步骤 408 中确定帧是无声语言，语言编码器执行步骤 410。在步骤 410 中，语言编码器将帧作为无声语言编码。在一个实施例中，以 1/4 速率或 2.6kbps 将无声语言帧编码。如果在步骤 408 中，确定帧不是无声语言，语言编码器执行步骤 412。

在步骤 412 中，语言编码器使用本领域中熟知的周期检测方法确定帧是否是过渡语言，如上述美国专利号 5,911,128 中所描述的方法。如果确定帧是过渡语言，语言编码器执行步骤 414。在步骤 414 中，将帧作为过渡语言（也就是从无声语言过渡到有声语言）编码。在一个实施例中，根据多脉冲内插编码法将过渡语言帧编码，1999 年 5 月 7 日提交的美国申请序号 09/307,294 题为“MULTIPULSE INTERPOLATIVE CODING OF TRANSITION SPEECH FRAMES”的申请中描述了该方法，该申请转让给了本发明的受让人，并通过引用全面结合于此。在另一实施例中，以全速率或 13.2kbps 将过渡语言帧编码。

如果在步骤 412 中语言编码器确定帧不是过渡语言，那么语言编码器执行步骤 416。在步骤 416 中，语言编码器将帧作为有声语言编码。在一个实施

例中,可以以半速率或 6.2kbps 将有声语言帧编码。也可以以全速率或 13.2kbps (或者在 8k CELP 编码器中以全速率 8kbps) 将有声语言帧编码。然而,本领域熟练的技术人员将理解以半速率将有声帧编码,通过开发有声帧的稳态本质,允许编码器节省有用的带宽。此外,不考虑有声语言编码所用的速率,最好使用来自过去帧的信息将有声语言编码,因此称为预测编码。

本领域熟练的技术人员应该理解遵循图 5 的步骤可以将语言信号或相应的 LP 剩余编码。从图 6A 的曲线图中可以看出作为时间函数的噪声、无声、过渡和有声语言的波形特征。从图 6B 的曲线图中可以看出作为时间函数的噪声、无声、过渡和有声 LP 剩余的波形特征。

如图 7 所示,在一个实施例中,原型音调周期 (PPP) 语言编码器 500 包括反向滤波器 502、原型提取器 504、原型量化器 506、原型去量化器 508、内插/合成模块 510、和 LPC 合成模块 512。语言编码器 500 最好作为 DSP 的一部分实现,并存在于如 PCS 或蜂窝电话系统的用户单元或基站中,或者在卫星系统的用户单元或网关中。

在语言编码器 500 中,数字化语言信号 $s(n)$ 被提供给反向 LP 滤波器 502,其中 n 是帧号。在特殊实施例中,帧长为 20ms。根据以下等式计算反向滤波器 $A(z)$ 的传递函数:

$$A(z) = 1 - a_1 z^{-1} - a_2 z^{-2} - \dots - a_p z^{-p},$$

其中系数 a_1 是滤波器抽头,它具有根据已知方法选出的预定值,上述美国专利号 5,414,796 中和美国申请序号 09/217,494 中描述了该方法,这两个申请都通过引用全面结合于此。数字 p 表示反向 LP 滤波器 502 用于预测目的的先前样本数。在特定的实施例中, p 设置为 10。

反向滤波器 502 将 LP 剩余信号 $r(n)$ 提供给原型提取器 504。原型提取器 504 从当前帧中提取原型。原型是当前帧的一部分,内插/合成模块 510 用类似地位于先前帧中的原型线性内插入当前帧,为了在解码器处重建 LP 剩余信号。

原型提取器 504 将原型提供给原型量化器 506,量化器根据本领域中所熟知的各种量化技术中的任何一种将原型量化。可以通过查表(未图示)获得的量化值被装配成用于在信道上发送的数据包,该数据包包括滞后和其它码本参数。该数据包被提供给发射机(未图示),并通过信道被发射到接收机(未图示)。也就是说反向 LP 滤波器 502、原型提取器 504、和原型量化器 506

执行当前帧的 PPP 分析。

接收机接收数据包，并将数据包提供给原型去量化器 508。原型去量化器 508 可以根据各种已知技术中的任何一种将数据包去量化。原型去量化器 508 将去量化的原型提供给内插/合成模块 510。内插/合成模块 510 用类似地位于先前帧中的原型内插原型，以重建当前帧的 LP 剩余信号。最好根据美国专利号 5,884,253 和上述美国申请序号 09/217,494 中描述的已知方法实施内插和帧的合成。

内插/合成模块 510 将重建的 LP 剩余信号 $\hat{r}(n)$ 提供给 LPC 合成模块 512。LPC 合成模块 512 还接收来自发射数据包的线谱对 (LSP) 值，它用于执行重建 LP 剩余信号 $\hat{r}(n)$ 的 LPC 滤波，以产生当前帧的重建语言信号 $\hat{s}(n)$ 。在另一实施例中，在执行当前帧的内插/合成之前，对于原型可以执行语言信号 $\hat{s}(n)$ 的 LPC 合成。也就是说原型去量化器 508、内插/合成模块 510、和 LPC 合成模块 512 执行当前帧的 PPP 合成。

在一个实施例中，如图 7 中 PPP 语言编码器 500 的语言编码器将闭环编码性能测量应用于每个编码帧，并同时保持语言编码器的目标平均比特率。语言编码器可以是 PPP 语言编码器或任何其它类型的可以通过提高以每帧为基础的编码率以改进语音质量的低比特率语言编码器。

在语言帧（在一个实施例，一个帧包括 20ms 的语言片段）的开环分类之后，使用预选速率 R_p 对语言帧编码。然后执行闭环性能参数。用预选速率 R_p 完全或部分编码之后得到编码器性能测量值。相关领域中熟知的典型性能测量值包括如信噪比 (SNR)、编码方案如 PPP 语言编码器中的 SNR 预测、预测误差量化 SNR、相位量化 SNR、振幅量化 SNR、感觉 SNR、和作为稳态测量值的当前和过去帧之间归一化的互相关。如果性能测量值 PNM 低于阈值 PNM_TH，那么编码率变化成期望编码方案给出较佳质量的值。通常，这意味着编码率的变化是增加。1998 年 11 月 13 日提交的美国申请序号 09/191,643 题为“CLOSED-LOOP VARIABLE-RATE MULTIMODE PREDICTIVE SPEECH CODER”的申请中描述了保持可变速率语言编码器质量的典型闭环分类方案，该申请转让给了本发明的受让人，并通过引用全面结合于此。

性能测量值 PNM 最好用于更新当前阈值 PNM_TH 周围的阈值直方图。直方图用于用以下方式实现语言编码器平均比特率的总体控制。语言编码器计算 W 帧窗口上的运行平均比特率，在 W 帧之后将运行平均比特率清零，并再次计

算下一 W 帧的运行平均比特率。在一个 W 帧周期的末端，从目标平均比特率 AVR 中减去平均比特率，将该差值除以初始预选的编码率值 R_p 。

如果除法 AVR/R_p 的商值 NR 是正的，累加 PNM_TH 右边最初 BR 个区间或直方图直条宽度（也就是与高于阈值的编码率关联的最初 BR 个区间）的直方图值。最好选择 BR 的值，使得累加值大于 NR。然后，阈值 PNM_TH 增加等于 $DTH_HI \times BR$ 的乘积的量值，其中 DTH_HI 是每个区间的增加量。应该注意到 DTH_HI 被首次初始化成适当的值。一个这种适当的值是 $(MAX_TH - PNM_TH) / HB$ （在下文中定义的参数）。

如果商值 NR 是负的，累加 PNM_TH 左边最初 BL 个区间的直方图值。最好选择 BL 的值，使得累加值大于 -NR。然后阈值 PNM_TH 减小等于 $DTH_LO \times BL$ 的乘积的量值，其中 DTH_LO 是每个区间的减小量。应该注意到 DTH_LO 被首次初始化成适当的值。一个这种适当的值是 $(PNM_TH - MIN_TH) / HB$ （下文中定义的参数）。

性能阈值 PNM_TH 分别限制在最大值 MAX_TH 和最小值 MIN_TH 之间，如果这种最大值和最小值或它们的估计值已知。有利的是，如果需要每个区间的减小量 DTH_LO 和每个区间的增加量 DTH_HI 可以分别被更新成商值 $(PNM_TH - MIN_TH) / HB$ 和 $(MAX_TH - PNM_TH) / HB$ ，其中 HB 等于直方图中区间数的一半。当语言编码器完成了使 W 帧窗口的平均比特率接近于目标平均比特率 AVR，所有直方图中 2HB 个区间的直方图值最好被清零。

在一个实施例中，在使用预选速率 R_p 编码期间发生直方图值的更新。这通过以下方式实现。首先，更新区间。设置阈值 PNM_TH 左边 HB 个区间中的每一个，对于阈值 PNM_TH（阈值 PNM_TH 位于直方图的中心）左边的第 i 个区间等于 $PNM_TH - DTH_LO \times i$ 的差值。设置阈值 PNM_TH 右边 HB 个区间中的每一个，对于阈值 PNM_TH 右边的第 i 个区间等于 $PNM_TH + DTH_LO \times i$ 的和。其次，将包括当前性能测量值 PNM 的区间的直方图值增加 1。

在一个实施例中，如图 7 中 PPP 语言编码器 500 的语言编码器执行图 8 流程图所示的算法步骤，它将闭环编码性能测量值 PNM 应用于每个帧，并同时保持语言编码器的目标平均比特率。语言编码器可以是 PPP 语言编码器或任何其它类型的能通过增加以每帧为基础的编码率而改进语音质量的低比特率语言编码器。

根据帧内容的开环分类，以速率 R_p 将当前语言帧编码。然后对帧进行闭

环测试,使得如果语言编码性能测量值 PNM 低于性能阈值 PNM_TH ,就提高编码率。然后根据以下方法步骤调节阈值 PNM_TH ,以保持语言编码器的运行平均比特率,或接近于目标平均比特率 AVR 。

在步骤 600 中,语言编码器计算长度为 W 的帧窗口的运行平均比特率。然后语言编码器执行步骤 602。在步骤 602 中语言编码器计算商值 $NR=(AVR-\text{运行平均比特率})/R_p$ 。然后语言编码器执行步骤 604。在步骤 604 中语言编码器确定 NR 是否大于或等于零。如果 NR 大于或等于零,语言编码器执行步骤 606。如果相反, NR 不大于或等于零,语言编码器执行步骤 608。

在步骤 606 中语言编码器累加 PNM_TH (它位于直方图的中心) 右边最初 BR 个直方图区间值,选择 BR 使得累加值大于 NR 。然后语言编码器执行步骤 610。在步骤 610 中语言编码器将 PNM_TH 设置为等于 PNM_TH 和 DTH_HI*BR 之和,其中 DTH_HI 等于每个直方图区间的增加量。然后语言编码器执行步骤 612。

在步骤 608 中语言编码器累加 PNM_TH 左边最初 BL 个直方图区间值,选择 BL 使得累加值大于 $-NR$ 。然后语言编码器执行步骤 614。在步骤 614 中语言编码器将 PNM_TH 设置为等于 PNM_TH 和 DTH_LO*BR 之差,其中 DTH_LO 等于每个直方图区间的减少量。然后语言编码器执行步骤 612。

如果需要可以在步骤 612 之前执行分别将 PNM_TH 限制在最大值 MAX_TH 和最小值 MIN_TH 之间的步骤。此外,如果需要可以在步骤 612 之前,执行将每个区间减小量 DTH_LO 和每个区间增加量 DTH_HI 分别更新为商值 $(PNM_TH-MIN_TH)/HB$ 和 $(MAX_TH-PNM_TH)/HB$ 的步骤,其中 HB 等于直方图中区间数的一半。应该注意到 DTH_HI 和 DTH_LO 应该首次被初始化成适当的值,如分别为 $(MAX_TH-PNM_TH)/HB$ 和 $(PNM_TH-MIN_TH)/HB$ 。

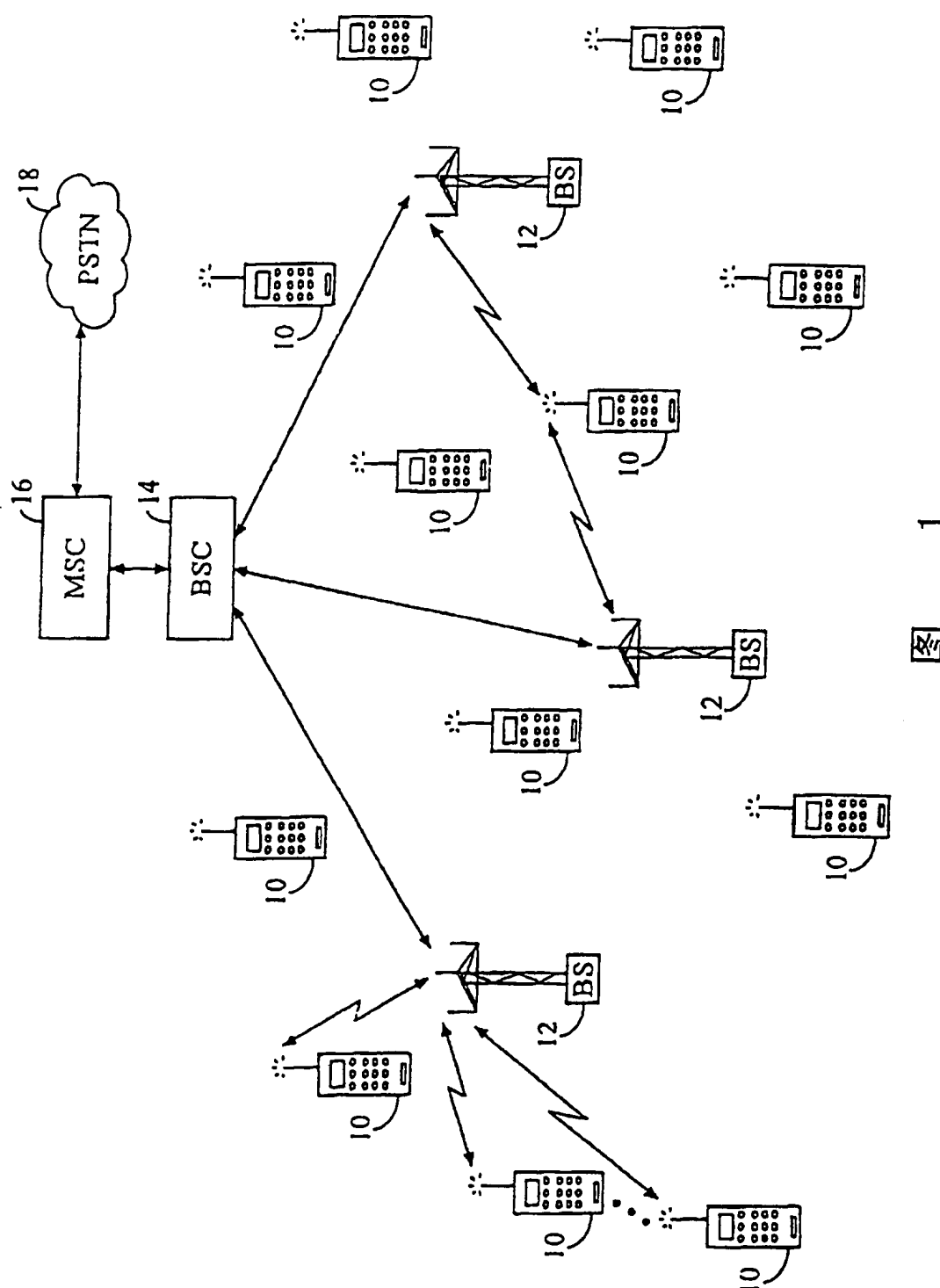
在步骤 612 中,语言编码器将所有 $2HB$ 个直方图区间的直方图值清零。然后语言编码器回到步骤 600 计算下一 W 帧的运行平均比特率。

在一个实施例中,语言编码器执行图 9 流程图所示的算法步骤,在以编码率 R_p 对 W 帧中每个帧的语言帧编码期间更新直方图区间值。在步骤 700 中,语言编码器设置阈值 PNM_TH 左边的所有直方图区间,对于阈值 PNM_TH (阈值 PNM_TH 位于直方图的中心) 左边的第 i 个区间等于 PNM_TH-DTH_LO*i 的差值。然后语言编码器执行步骤 702。步骤 702 中语言编码器设置阈值 PNM_TH 右边的所有直方图区间,对于阈值 PNM_TH 右边的第 i 个区间等于 PNM_TH+DTH_HI*i 的和。然后语言编码器执行步骤 704。在步骤 704 中,语言编码器将包括当前

性能测量值 PNM 的直方图区间值增加 1。

因此，已经描述了语言编码器中保持目标比特率的新颖方法和装置。本领域熟练的技术人员将理解结合这里所揭示实施例描述的各种说明性逻辑模块和算法步骤可以通过以下装置实现或执行：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。处理器最好是微处理器，但是另外处理器也可以是任何传统的处理器、控制器、微控制器、或状态机。软件模块可驻留在 RAM 存储器、快闪存储器、寄存器或本领域中所熟知的任何其它形式的可写存储媒介中。本领域熟练的技术人员还应该理解以上描述中参考的数据、指令、命令、信息、信号、比特、码元和码片最好用电压、电流、电磁波、磁场或粒子、光场或粒子或它们的任何其它组合表示。

因此显示并描述了本发明的较佳实施例。然而对本领域熟练的技术人员显而易见的是不脱离本发明的精神和范围可以对这里揭示的实施例进行各种变化。因此，本发明只受以下权利要求书的限制。



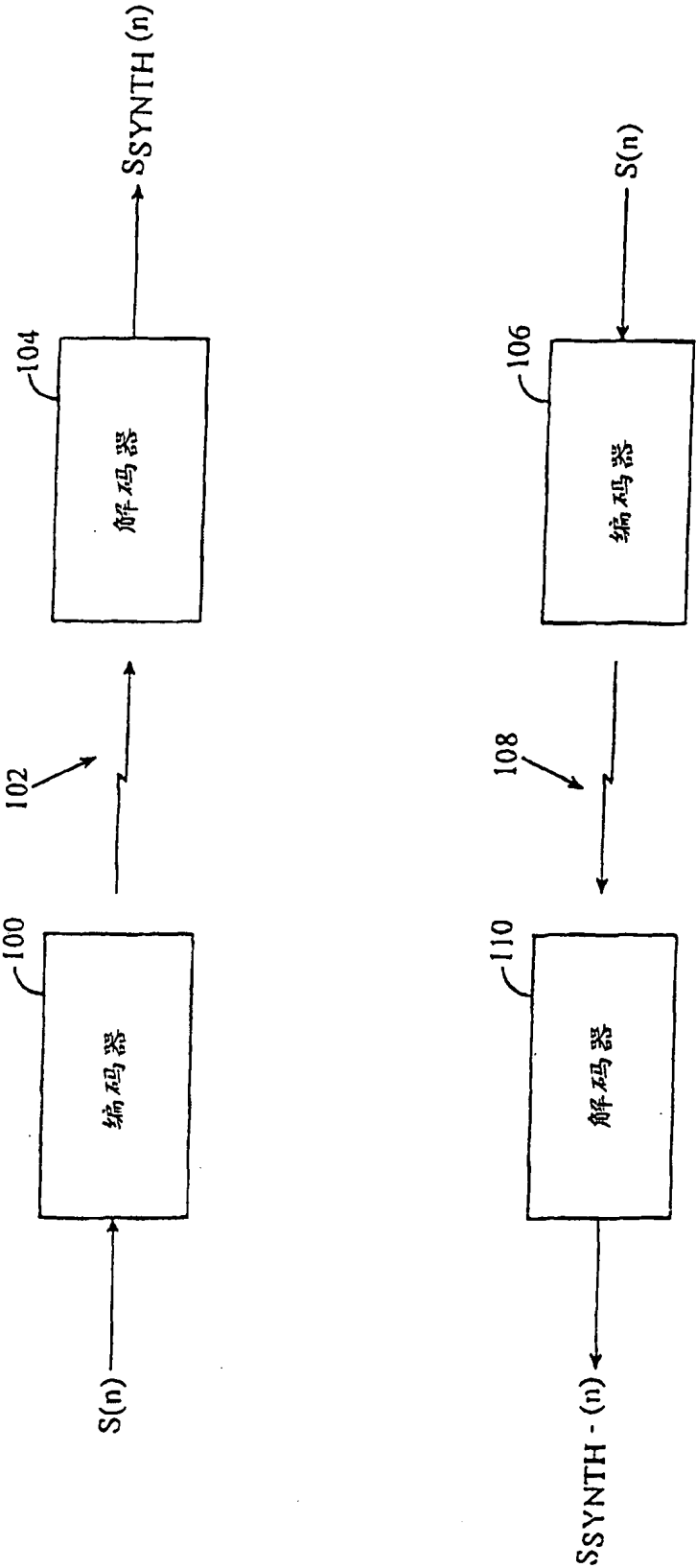


图 2

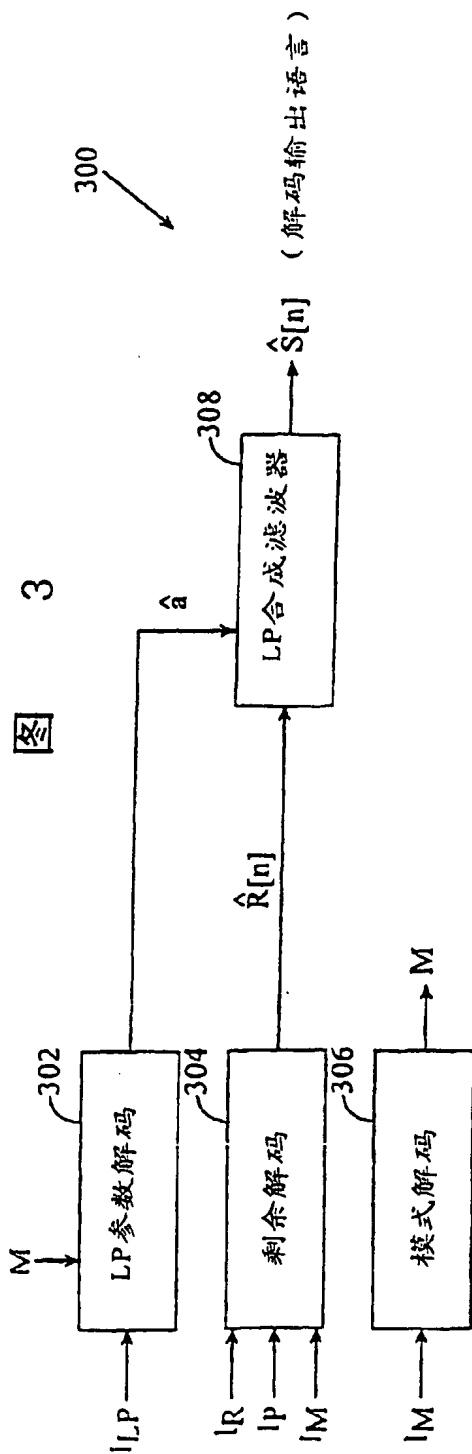
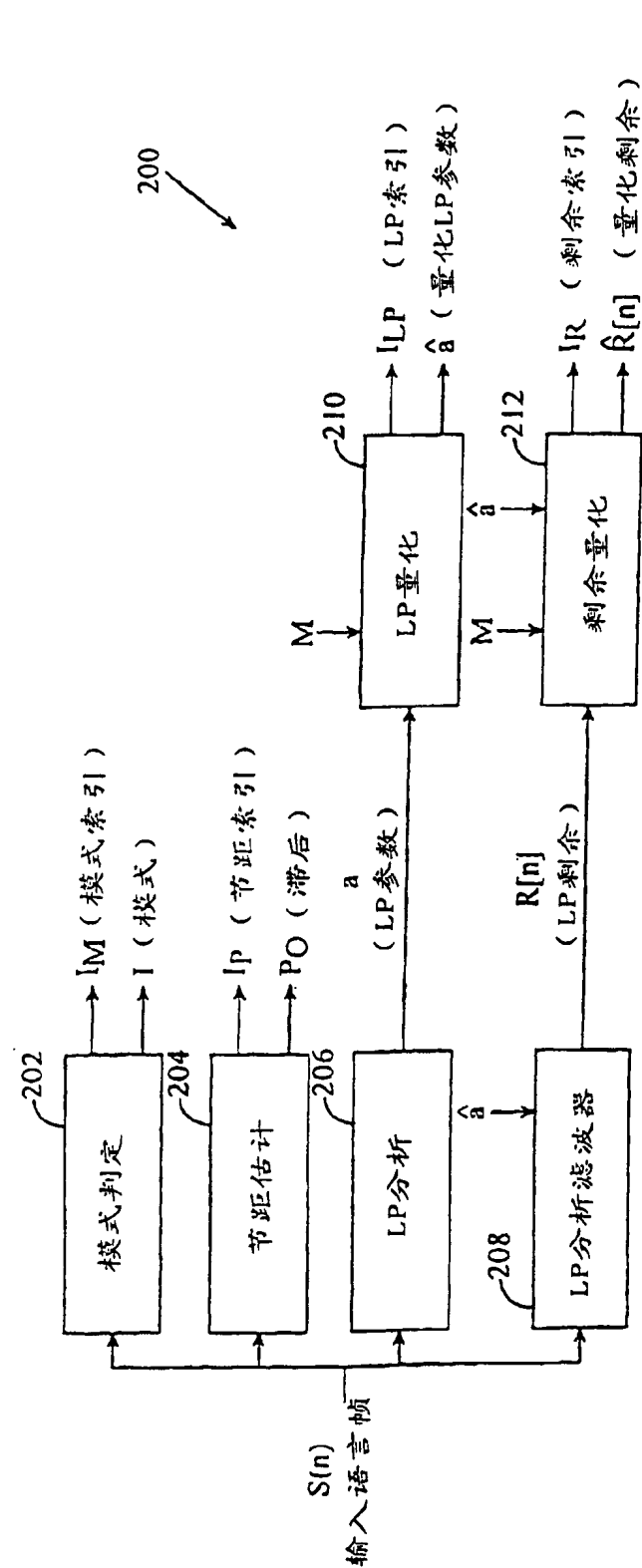


图 3

图 4

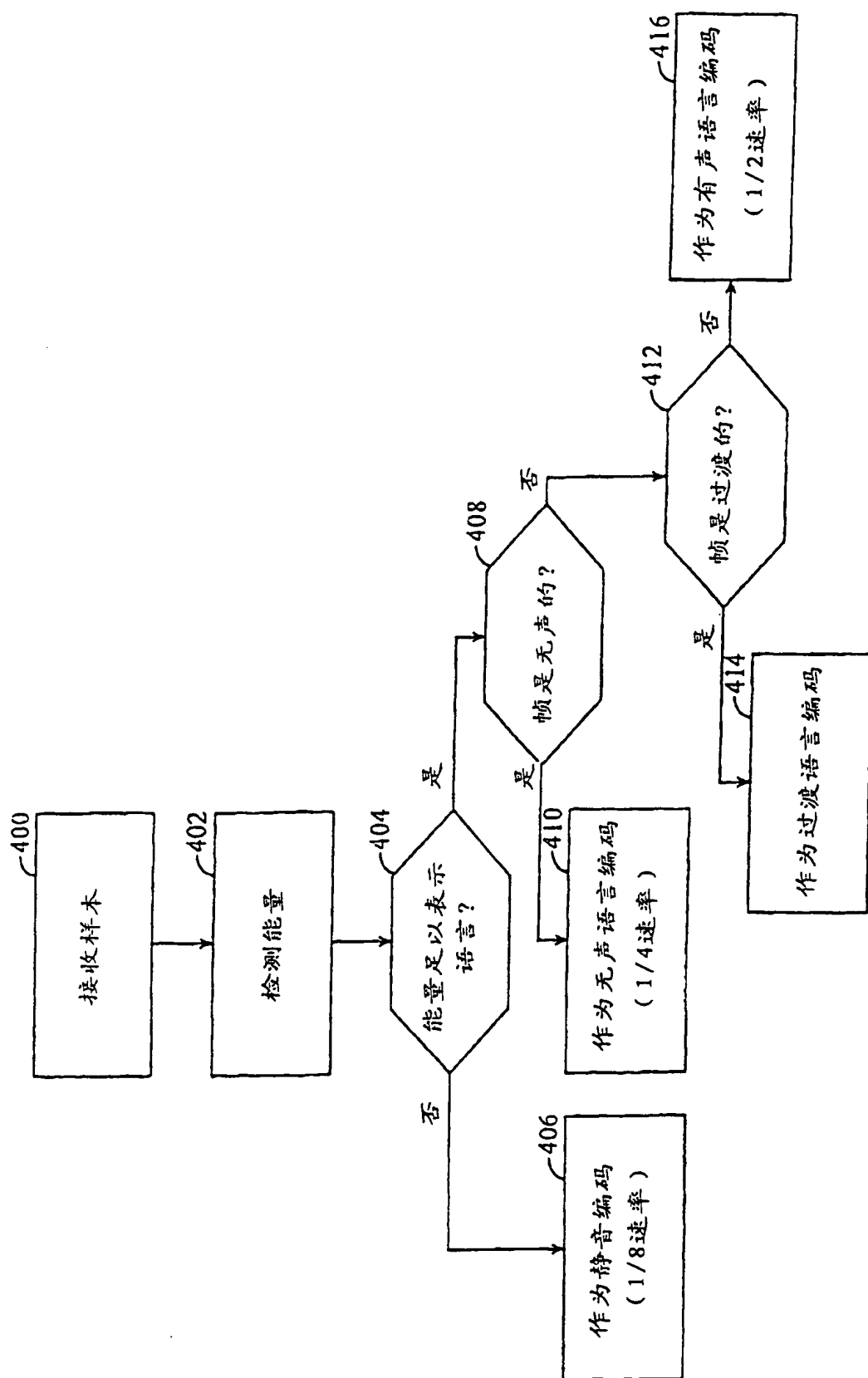


图 5

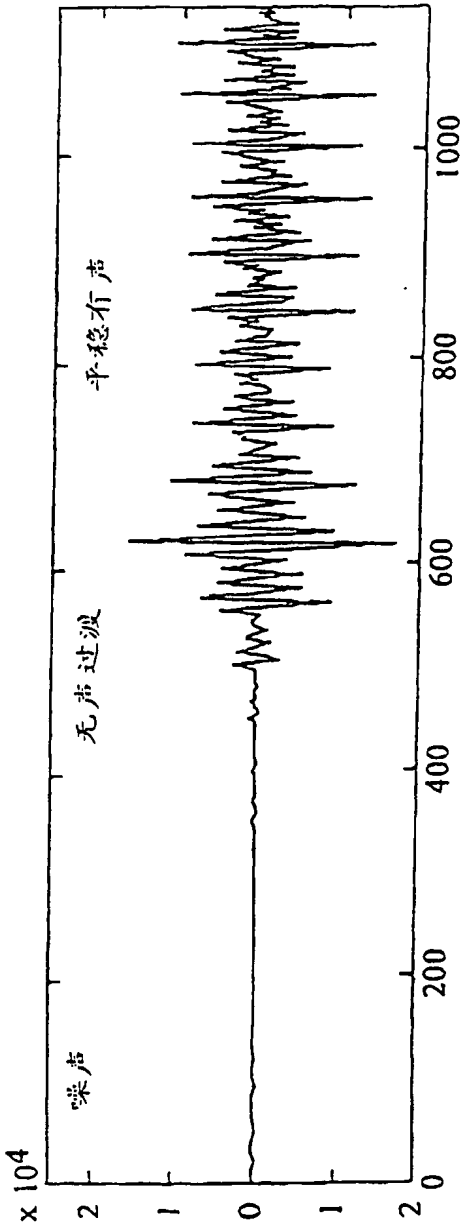


图 6A

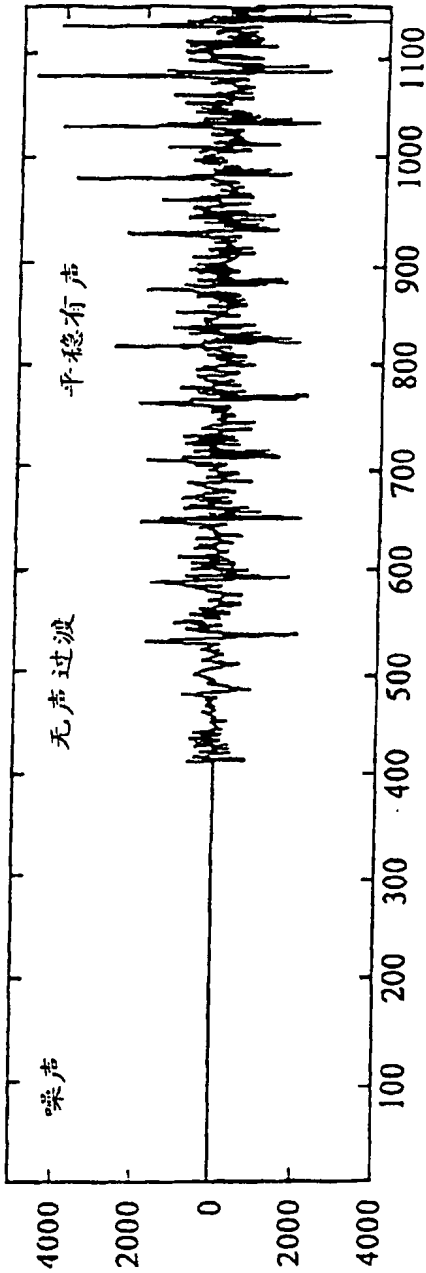


FIG. 6B

图 6B

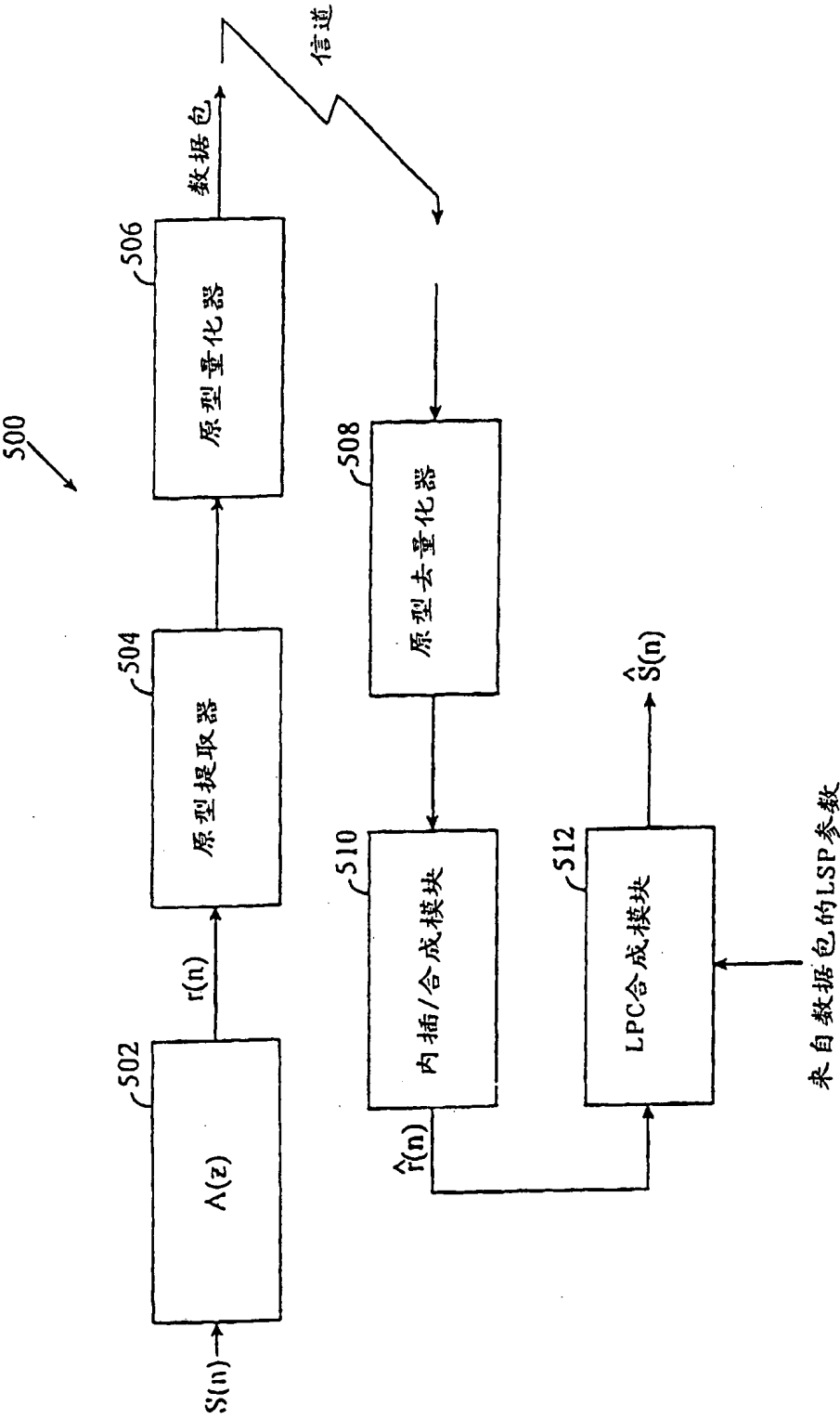


图 7

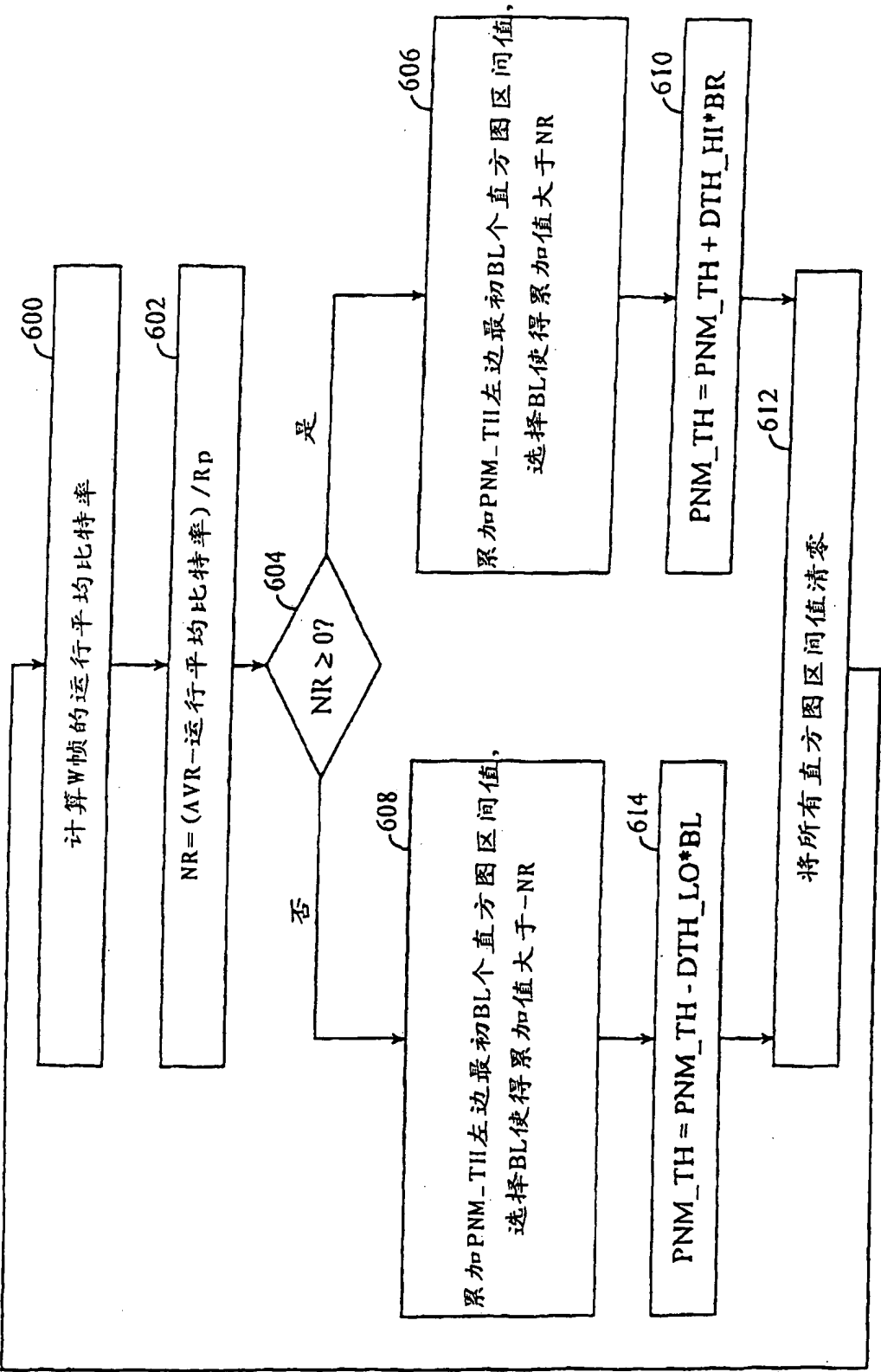


图 8

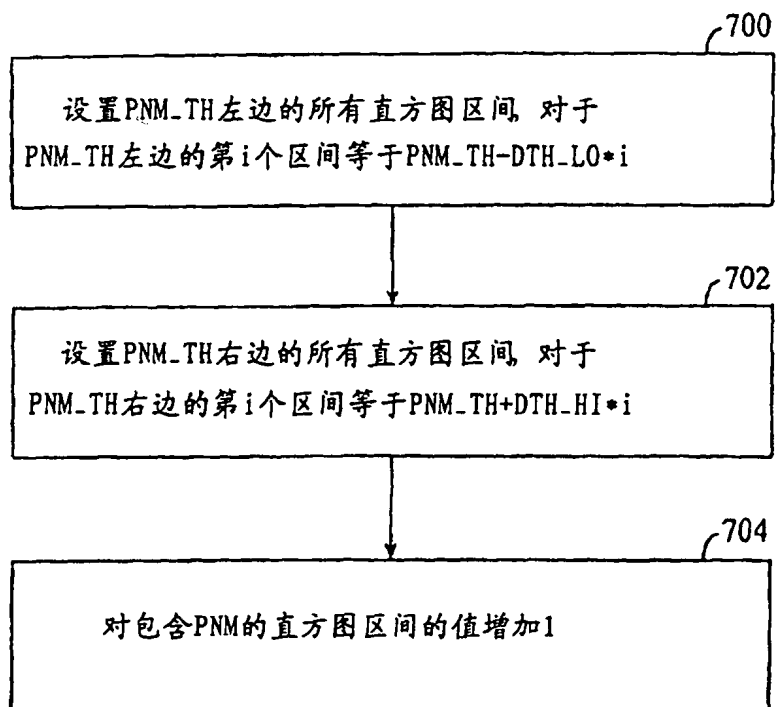


图 9