

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

G10L 21/02

G10L 19/00



[12] 发明专利申请公开说明书

[21] 申请号 01817599.6

[43] 公开日 2004年1月21日

[11] 公开号 CN 1470052A

[22] 申请日 2001.10.17 [21] 申请号 01817599.6

[30] 优先权

[32] 2000.10.18 [33] US [31] 09/691,440

[86] 国际申请 PCT/IB01/01947 2001.10.17

[87] 国际公布 WO02/33697 英 2002.4.25

[85] 进入国家阶段日期 2003.4.18

[71] 申请人 诺基亚有限公司

地址 芬兰埃斯波

[72] 发明人 P·奥亚拉 J·罗托拉-普基拉

J·韦尼奥 H·米科拉

[74] 专利代理机构 中国专利代理(香港)有限公司

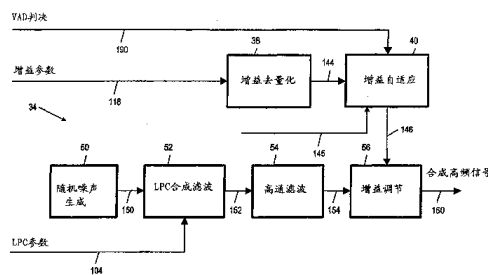
代理人 陈景峻 陈 霁

权利要求书5页 说明书13页 附图8页

[54] 发明名称 宽带语音编解码器中的高频增强层编码

[57] 摘要

用于编码和解码输入信号(100)和提供合成的语音(110)的语音编码方法和设备,其中通过对仿真信号(150)高通滤波和着色获得合成语音(110)的高频部分(160)来提供处理过的仿真信号(154)。处理过的仿真信号(154)在输入信号(100)的激活语音周期通过第一缩放因子(114,144)进行缩放(530,540),在非激活语音周期通过第二缩放因子(114和115,144和145)进行缩放,其中第一缩放因子(114,144)具有输入信号(100)的高频带特性,并且第二缩放因子(114和115,144和145)具有输入信号(100)的低频带特性。特别地,第二缩放因子(114和115,144和145)基于合成语音(110)的低频部分进行估算,并且仿真信号(150)的着色是基于具有输入信号(100)低频特性的线性预测编码系数(104)的。



1. 一种语音编码(500)方法,用于编码和解码具有激活语音周期和非激活语音周期的输入信号(100),并且用于提供一种具有高频部分和低频部分的合成语音信号(110),其中该输入信号在编码和语音合成过程中被分成高频带部分和低频带部分,并且其中具有低频带特性的语音相关参数(104)被用来处理仿真信号(150),用以提供处理过的仿真信号(152),处理过的仿真信号(152)用于进一步提供合成语音的高频部分(160),所述方法包括步骤:

在激活语音周期中,以第一缩放因子(114, 144)缩放(530)处理过的仿真信号(152),以及

在非激活语音周期中,以第二缩放因子(114&115, 144&145)缩放(540)处理过的仿真信号(152),其中第一缩放因子具有输入信号高频带的特性,同时第二缩放因子具有合成信号低频部分的特性。

2. 权利要求1所述的方法,其中处理过的仿真信号(152)被高通滤波,用于在具有合成语音的高频部分的特性的频率范围中提供滤波过的信号(154)。

3. 权利要求2所述的方法,其中,频率范围是在6.4—8.0kHz的范围内。

4. 权利要求1所述的方法,其中输入信号(100)被高通滤波,用于在具有合成语音高频部分特性的频率范围中提供滤波过的信号(112),并且其中第一缩放因子(114, 144)是从滤波过的信号(112)中估算出来的。

5. 权利要求4所述的方法,其中非激活语音周期包括语音释放延迟周期和舒适噪声周期,其中用于在语音释放延迟周期中缩放处理过的仿真信号(152)的第二缩放因子(114&115, 144&145)是从滤波过的信号(112)中估算出来的。

6. 权利要求5所述的方法,其中合成语音的低频部分从输入信号(100)的已编码低频带(106)中再现,并且其中用于在语音释放延迟周期中缩放处理过的仿真信号(152)的第二缩放因子(114&115, 144&145)也是从合成语音信号的低频部分中估算出来的。

7. 权利要求6所述的方法,其中用于在舒适噪声周期中缩放处理过的仿真信号(152)的第二缩放因子(114&115, 144&145)是从合成语音

信号的低频部分中估算出来的。

8. 权利要求 6 所述的方法, 进一步包括向接收端发送已编码比特流, 用于解码的步骤, 其中已编码比特流包括指示第一缩放因子 (114, 144) 的数据。

5 9. 权利要求 8 所述的方法, 其中已编码比特流包括数据 (118), 该数据 (118) 指示用于在语音释放延迟周期中缩放处理过的仿真信号 (152) 的第二缩放因子 (114&115)。

10. 权利要求 8 所述的方法, 其中用于缩放处理过的仿真信号的第二缩放因子 (114&115, 144&145) 在接收端 (34) 中提供。

10 11. 权利要求 6 所述的方法, 其中第二缩放因子 (114&115, 144&145) 指示从合成语音的低频部分中确定的频谱倾斜因子。

12. 权利要求 7 所述的方法, 其中用于在舒适噪声周期中缩放处理过的仿真信号的第二缩放因子 (114&115, 144&145) 指示从合成语音的低频部分中确定的频谱倾斜因子。

15 13. 权利要求 4 所述的方法, 其中第一缩放因子 (114, 144) 进一步从处理过的仿真信号 (152) 中估算出。

14. 权利要求 1 所述的方法, 进一步包括基于输入信号 (100) 提供用于监视激活语音周期和非激活语音周期的语音激活信息 (190) 的步骤。

20 15. 权利要求 1 所述的方法, 其中语音相关参数包括具有输入信号低频带特性的线性预测编码系数。

25 16. 一个语音信号发射机和接收机系统, 用于编码和解码具有激活语音周期和非激活语音周期的输入信号 (100), 并且用于提供一种具有高频部分和低频部分的合成语音信号 (110), 其中该输入信号在编码和语音合成过程中被分成高频带部分和低频带部分, 其中具有输入信号低频部分特性的语音相关参数 (118, 104, 140, 145) 被用来在接收机 (30) 中处理仿真信号 (150) 来提供合成语音信号高频部分 (160) 的, 所述系统包括:

30 发射机中的第一装置 (12, 14), 响应输入信号 (100), 用于提供具有输入信号高频带特性的第一缩放因子 (114, 144);

接收机中的解码器 (34), 用于从发射机接收已编码的比特流, 其中已编码的比特流包括语音相关参数, 该相关参数包括指示第一缩放因子

(114, 144) 的数据; 以及

接收机中的第二装置 (40, 56), 响应语音相关参数 (118, 145), 用于提供第二缩放因子 (144&145), 以及在非激活周期中使用第二缩放因子 (144&145) 缩放处理过的仿真信号 (152), 并且在激活周期中使用
5 第一缩放因子 (114&144) 缩放处理过的仿真信号 (152), 其中第一缩放因子具有输入信号高频带的特性, 同时第二缩放因子具有合成信号低频带的特性。

17. 权利要求 16 所述的系统, 其中第一装置包括一个滤波装置 (12), 用于高通滤波输入信号, 并且提供滤波过的输入信号 (112), 该信号具
10 有相应于合成语音的高频部分的频率范围, 同时其中从滤波过的输入信号 (112) 中估算出第一缩放因子 (114, 144)。

18. 权利要求 17 所述的系统, 其中频率范围是在 6.4—8.0kHz 范围内。

19. 权利要求 17 所述的系统, 进一步包括在发射机中的第三装置
15 (16, 24), 用于在相应于合成信号的频率范围内提供高通滤波的随机噪声 (134), 同时用于基于高通滤波随机噪声改变第一缩放因子 (114, 144)。

20. 权利要求 16 所述的系统, 进一步包括装置 (98), 响应输入信号 (100), 用于监视激活和非激活语音周期。

21. 权利要求 16 所述的系统, 进一步包括装置 (18), 响应第一缩放因子 (114, 144), 用于提供已编码的第一缩放因子 (118), 并且将
20 指示已编码的第一缩放因子的数据包括到用于发送的已编码比特流中。

22. 权利要求 19 所述的系统, 进一步包括装置 (18), 响应第一缩放因子 (114, 144), 用于提供已编码的第一缩放因子 (118), 并且将
25 指示已编码的第一缩放因子的数据包括到用于发送的已编码比特流中。

23. 一个编码器 (10), 用于编码具有激活语音周期和非激活语音周期的输入信号 (100), 并且该输入信号被分为高频带和低频带, 同时用于提供已编码比特流, 该已编码比特流包括具有输入信号低频带特性的语音相关参数, 以便允许解码器 (34) 使用语音相关参数处理仿真信号
30 (150), 用以提供合成语音的高频部分 (160), 并且其中在非激活语音周期中, 使用基于合成语音低频部分的缩放因子 (114&115, 144&145) 缩放处理过的仿真信号 (152), 所述编码器包括:

装置(12), 响应输入信号(100), 用于对输入信号(100)进行高通滤波, 用以在相应于合成语音(110)的高频部分的频率范围中提供高通滤波过的信号(112), 并且基于高通滤波过的信号(112)进一步提供另一个缩放因子(114, 144); 以及

- 5 装置(18), 响应另一个缩放因子(114, 144), 用于在已编码比特流中提供指示另一个缩放因子的已编码信号(118), 以便允许解码器(34)在激活语音周期接收已编码信号, 并使用另一个缩放因子(114, 144)缩放处理过的仿真信号(152)。

24. 一个移动站(200), 其被设置来发送已编码比特流至解码器(34, 10 220), 用以提供具有高频部分和低频部分的合成语音(110), 其中已编码比特流包括指示语音数据输入信号(100)的语音数据, 该输入信号具有激活语音周期和非激活语音周期并且被划分成高频带和低频带, 其中语音数据包括具有输入信号低频带特性的语音相关参数(104), 以便允许解码器(34)基于语音相关参数提供合成语音的低频部分, 并且基于语音相关参数(104)着色仿真信号, 同时基于合成语音的低频部分使用缩放因子(144&145)缩放着色的仿真信号, 用于在非激活语音周期中提供合成语音的高频部分(160), 所述移动站包括:

- 滤波器(12), 响应输入信号(100), 用于高通滤波相应于合成语音高频部分的频率范围的输入信号, 并且用于基于高通滤波过的输入信号(112)提供另一个缩放因子(114, 144); 以及

量化模块(18), 响应另一个缩放因子(114, 144), 用于在已编码比特流中提供指示另一个缩放因子(114, 144)的已编码信号(118), 以便允许解码器(34)在激活语音周期中基于另一个缩放因子(114, 144)缩放着色的仿真信号。

25. 一种电信网络(300)中的元件(34, 320), 其被设置来接收包括指示来自移动站(330)的输入信号的语音数据的已编码的比特流, 用以提供具有高频部分和低频部分的合成语音, 其中输入信号具有激活语音周期和非激活语音周期, 并且输入信号被分为高频带和低频带, 其中语音数据(104, 118, 145, 190)包括具有输入信号低频带特性的语音相关参数(104)和具有输入信号高频带特性的增益参数(118), 并且基于语音相关参数(104)提供合成语音的低频部分, 所述元件包括:

第一机构(38), 响应增益参数(118), 用于提供第一缩放因子(144);

第二机构(52, 54), 响应语音相关参数(104), 用于合成和高通滤波仿真信号(150), 用以提供一个合成和高通滤波过的仿真信号(150);

5 第三机构(40), 响应第一缩放因子(144)和语音数据(145, 190), 用于提供组合的缩放因子(146), 该组合的缩放因子包括具有输入信号高频带特性的第一缩放因子(144), 基于第一缩放因子(144)和具有合成语音低频部分特性的另一个语音相关参数(145)的第二缩放因子(144&145); 以及

10 第四机构, 响应于合成和高通滤波过的仿真信号(154)以及合成缩放因子(146), 用于在激活语音周期和非激活语音周期中, 分别使用第一(144)和第二缩放因子(144&145)缩放合成和高通滤波过的仿真信号(154).

宽带语音编解码器中的高频增强层编码

技术领域

5 本发明通常涉及编码和解码合成语音的领域，尤其是涉及自适应多速率宽带语音编解码器。

背景技术

10 当前很多语音编码方法都是基于线性预测(LP)编码，感性地直接从时间波形而不是从语音信号的频谱(正如所谓的信道声码器或所谓的共振峰声码器一样)中提取语音信号的有效特征。在LP编码中，首先分析语音波形(LP分析)以确定一个随时间变化的产生语音信号的声道激励模型，以及转换函数。解码器(如果通过电信传输编码的语音信号则在接收终端中)然后使用合成器(为了进行LP合成)通过一个用参数表示的模拟声道的系统传递激励以便重新生成原始语音。随着说话者产生语音信
15 号，声道模型参数和模型激励均周期性地更新以适合说话者相应的变化。但是在更新之间，也就是说，在任何特定的间隔之间，激励和系统参数保持不变，因此模型执行的处理是线性的不随时间变化的处理。整个编码和解码(分布式的)系统被称为编解码器。

20 在一个使用LP编码产生语音的编解码器中，解码器需要编码器提供三种输入：如果激励是有声的，则提供音调周期，增益因子和预测系数。(在某些编解码器中，还要提供激励种类，也就是说是有声的还是无声的，但对于代数码激励线性预测(ACELP)编解码器通常并不需要。例如。在前向估计处理中，LP编码是预测型的，因为它使用基于实际输入的应用参数的语音波形片断(在一段特定间隔内)的预测参数。

25 基本的LP编码和解码可用于使用相对低的数据速率以数字方式传输语音，因为它使用非常简单的激励系统，它产生合成的发声的语音。一个所谓的码激励线性预测(CELP)编解码器是一种增强的激励编解码器。它基于"冗余"编码。模拟声道是根据参数被编码成压缩语音的数字滤波器。这些滤波器是由表示原始说话者的声带震动的信号所驱动，即"
30 激励"。音频语音信号的冗余是较少地数字滤波的(原始)音频语音信号。在所谓的"冗余脉冲激励"中，CELP编解码器对冗余编码并将它作为激励

的基础，但是，CELP使用从预设的一套波形模板中选择的波形模板来表示冗余样本块而不是根据样本不同情况分别编码冗余波形。码字是由编码器决定的并提供给解码器，解码器然后使用码字以选择冗余序列表示原始的冗余样本。

5 依据奈奎斯特定理，采样率 F_s 的语音信号可以表示一个从0到 $0.5F_s$ 的频带。当前，大多语音编解码器(编码器-解码器)使用8kHz的采样速率。如果采样速率从8kHz增加，语音的逼真度也会改进因为可以表示更高的频率。现在，语音信号的采样率通常为8kHz，但是开发中的移动电话基站将使用16kHz的采样率。依据奈奎斯特定理，16kHz的采样率在频
10 带0-8kHz表示语音。然后对采样的语音进行编码以使用发射机进行通信，然后被接收机解码。使用16kHz的采样速率采样的语音的语音编码被称为宽带语音编码。

当语音采样率增加时，编码复杂性也增加了。对于某些算法，随着采样率增加，编码复杂性甚至达到指数级增长。因此，编码复杂性经常
15 是确定宽带语音编码算法的一个限制性因素。例如，移动电话基站的功耗，可用的处理功率和内存要求严重影响算法的应用。

在现有技术的宽带编解码器中，如图1所示，预处理阶段用于低通过滤和从原始的16kHz到12.8kHz采样频率下抽样输入语音信号。下抽样信号然后减少十分之一以使得在20ms内320个样本数减少到256。在有效的0
20 到6.4kHz频率带宽内，下抽样和减少了十分之一的信号使用合成分析(A-b-S)循环以提取LPC，音调和激励参数进行编码，并量化成编码的比特流以发送到接收端进行解码。在A-b-S循环中，本地合成信号进一步上抽样并以内插值替换以符合原始样本频率。编码处理之后，6.4 kHz到8.0kHz的频带为空。宽带编解码器在这个空频率范围生成随机噪声并利用
25 如下所述的合成滤波使用LPC参数着色(colors)随机噪声。随机噪声首先根据下式进行缩放

$$e_{scaled} = \text{sqr}[\{ext^T(n)exc(n)exc(n)\}/\{e^T(n)e(n)\}]e(n) \quad (1)$$

其中 $c(n)$ 表示随机噪声 $exc(n)$ 表示LPC激励。上标T表示向量转置。缩放的随机噪声使用着色(coloring)LPC合成滤波器和6.0-7.0kHz带通
30 滤波器进行滤波。这种着色(colored)的高频部分进一步使用关于合成信

号的频谱倾斜的信息进行缩放。频谱倾斜可以通过首先使用下列公式计算自相关系数, r , 来估计:

$$r = \{s^T(i)s(i-1)\} / \{s^T(i)s(i)\} \quad (2)$$

5 其中 $s(i)$ 是合成语音信号。相应地, 估计的增益 f_{est} 由下面决定

$$f_{est} = 1.0 - r \quad (3)$$

并限制 $0.2 \leq f_{est} \leq 1.0$ 。

10 在接收端, 在核心解码处理后, 对合成信号进行进一步的后续处理以通过上抽样信号来满足输入信号采样频率, 以便生成实际的输出。因为高频噪声电平是基于从合成信号的低频带和频谱倾斜获取的LPC参数估算出的, 缩放和着色随机噪声可以在编码器端或解码器端实现。

15 在现有技术的编解码器中, 基于基层信号电平以及频谱倾斜估算高频噪声电平。因而, 合成信号的高频部分被滤波掉。因此, 噪声电平与在6.4-8.0kHz频率范围的实际输入信号特性不相一致。这样, 现有技术的编解码不能提供高质量合成信号。

考虑到在高频范围的实际输入信号的特性, 提供能够提供高质量合成信号的方法和系统是有利和值得的。

20 发明内容

25 本发明的主要目的是在分布式语音处理系统中提高合成语音的质量。这个目的可以通过使用具有在6.0至7.0kHz频率范围的原始语音信号中的高频部分的输入信号特性, 例如, 在激活语音周期, 确定在合成合成语音的高频部分中的着色的 (colored) 高通滤波仿真信号的缩放因子来实现。在非激活语音周期, 可以通过合成语音信号的低频部分确定缩放因子。

因此, 本发明的第一个方面是一种语音编码方法, 用于编码和解码具有激活语音周期和非激活语音周期的输入信号, 并且用于提供一种具有

高频部分和低频部分的合成语音信号，其中该输入信号在编码和语音合成过程中被分成高频带部分和低频带部分，并且其中具有低频部分特性的语音相关参数被用来处理用于提供合成语音信号高频部分的仿真信号。该方法包括步骤：

- 5 在激活语音周期中，以第一缩放因子缩放处理过的仿真信号，以及在非激活语音周期中，以第二缩放因子缩放处理过的仿真信号，其中第一缩放因子是输入信号的高频带特性，并且第二缩放因子是合成信号的低频部分的特性。

10 优选地，输入信号被高通滤波以便在在合成语音的高频部分的频率范围特性中提供滤波过的信号，其中第一缩放因子从滤波过的信号中估算出来的，并且其中当非激活语音周期包括语音释放延迟周期和舒适噪声周期时，从滤波过的信号中估算出在语音释放延迟周期中缩放处理过的仿真信号的第二缩放因子。

15 优选地，用于在语音释放延迟周期中缩放处理过的仿真信号的第二缩放因子也是从合成语音信号的低频部分中估算出来的，并且用于在舒适噪声周期中缩放处理过的仿真信号的第二缩放因子从合成语音信号的低频部分中估算出来。

20 优选地，第一缩放因子在流向接收端的已编码比特流中编码和发送，并且用于语音释放延迟周期的第二缩放因子也包括在已编码的比特流中。

用于语音释放延迟周期的第二缩放因子可以在接收端确定。

优选地，第二缩放因子也可以从频谱倾斜因子 (spectral tilt) 中估算出，该频谱倾斜因子由合成语音的低频部分确定。

优选地，第一缩放因子进一步从处理过的仿真信号中估算出。

25 本发明的第二方面是用于编码和解码具有激活语音周期和非激活语音周期的输入信号并且用于提供一种具有高频部分和低频部分的合成语音信号的语音信号发射机和接收机系统，其中该输入信号在编码和语音合成过程中被分成高频带部分和低频带部分，其中输入信号的低频带的语音相关参数被用来在接收机中处理仿真信号来提供合成语音信号的高频部分。该系统包括：

30 接收机中的解码器，用于从发射机中接收已编码的比特流，其中已编码的比特流包括语音相关参数；

发射机中的第一模块，响应输入信号，提供用于在激活周期中缩放处理过的仿真信号的第一缩放因子，以及

接收机中的第二模块，响应已编码的比特流，提供在非激活周期中缩放处理过的仿真信号的第二缩放因子，其中第一缩放因子是输入信号高频带的特性，并且第二缩放因子是合成信号低频部分的特性。

优选地，本发明的第一模块包括一个滤波器，用于高通滤波输入信号，并且提供滤波过的输入信号，该信号具有相应于合成语音的高频部分的频率范围，以便允许从滤波过的输入信号中估算出第一缩放因子。

优选地，在发射机中使用第三模块来在相应于合成信号的频率范围内提供着色的高通滤波的随机噪声，以便可以基于着色的高通滤波随机噪声修改第一缩放因子。

本发明的第三方面是编码器，用于编码具有激活语音周期和非激活语音周期的输入信号，该输入信号被分为高频带和低频带，用于提供包括具有输入信号低频带特性的语音相关参数的已编码比特流，以便允许解15 码器基于语音相关参数再现合成语音的低频部分并且基于语音相关参数处理仿真信号来提供合成语音高频部分，其中在非激活语音周期中，使用基于合成语音低频部分的缩放因子缩放处理过的仿真信号。该编码器包括：

滤波器，响应输入信号，用于对相应于合成语音的高频部分的频率范围中的输入信号高通滤波，并且提供指示高通滤波过的输入信号的第一20 信号；

装置，响应第一信号，用于基于高通滤波的输入信号以及合成语音的低频部分提供另一个缩放因子，并且提供指示另一个缩放因子的第二信号；以及

25 量化模块，响应第二信号，用于在已编码比特流中提供指示另一个缩放因子的已编码信号，以便允许解码器在激活语音周期基于另一个缩放因子缩放处理过的仿真信号。

本发明的第四方面是一个移动站，其被设置来发送已编码比特流至解码器以便提供具有高频部分和低频部分的合成信号，其中已编码比特流30 包括语音数据，该语音数据指示具有激活语音周期和非激活语音周期的输入信号，并且输入信号被分成高频带和低频带，其中语音数据包括具有输入信号低频带特性的语音相关参数，以便允许解码器基于语音相关

参数提供合成语音的低频部分，并且基于语音相关参数着色仿真信号，同时基于合成语音的低频部分使用缩放因子缩放着色的仿真信号以便在非激活语音周期中提供合成语音的高频部分。移动站包括：

5 滤波器，响应输入信号，用于高通滤波相应于合成语音高频部分的频率范围的输入信号，并且用于基于高通滤波过的输入信号提供另一个缩放因子；以及

量化模块，响应该缩放因子和另一个缩放因子，用于在已编码比特流中提供指示另一个缩放因子的已编码信号，以便允许解码器在激活语音周期中基于另一个缩放因子缩放着色的仿真信号。

10 本发明的第五方面是电信网络中的元件，其被设置来接收用于提供具有高频部分和低频部分的合成语音的已编码比特流，该比特流包括指示来自移动站的输入信号的语音数据，其中具有激活语音周期和非激活语音周期的输入信号被分为高频带和低频带，同时语音数据包括具有输入信号的低频带特性的语音相关参数以及具有输入信号高频带特性的增益参数，其中基于语音相关参数提供合成语音的低频部分，所述元件包括：

15 第一机构，响应增益参数，用于提供第一缩放因子；

第二机构，响应语音相关参数，用于合成和高通滤波的仿真信号用以提供一个合成和高通滤波过的仿真信号；

20 第三机构，响应第一缩放因子和语音数据，用于提供组合的缩放因子，该组合的缩放因子包括具有输入信号高频带特性的第一缩放因子和基于第一缩放因子和具有合成语音低频部分特性的另一个语音相关参数的第二缩放因子；以及

25 第四机构，响应合成和高通仿真信号以及合成缩放因子，用于在激活语音周期和非激活语音周期中，分别使用第一和第二缩放因子缩放合成和高通滤波过的仿真信号。

在结合图 2 至 8 阅读说明书后，本发明将变得更加清楚。

附图说明

图 1 是说明现有技术的宽带语音编解码器的框图。

30 图 2 是说明根据本发明的宽带语音编解码器的框图。

图 3 是说明本发明的宽带语音编解码器的后端处理功能的框图。

图 4 是说明本发明的宽带语音解码器的结构的框图。

图 5 是说明宽带语音编解码器的后端处理功能的框图。

图 6 是说明根据本发明的移动站的框图。

图 7 是说明根据本发明的电信网络的框图。

图 8 是说明根据本发明的语音编码方法的流程图。

5

具体实施方式

如图 2 所示, 根据本发明, 宽带语音编解码器 1 包括用于对输入信号 100 进行预处理的预处理部件 2。如背景技术部分所述, 与现有技术中的编解码器相类似, 预处理部件 2 下抽样并且从输入信号 100 中抽取十分之一, 使其变为有效带宽为 0-6.4kHz 的语音信号 102。为了提取出一套线性预测编码 (LPC) 音调以及激励参数或系数 104, 使用传统 ACELP 技术的合成分析编码部件 (analysis-by-synthesis encoding block) 4 对处理过的语音信号 102 进行编码。可以使用相同的编码参数, 以及高通滤波模块将仿真信号或伪随机噪声处理成着色的高通滤波随机噪声 (134, 图 3; 154, 图 5)。编码部件 4 也可以为后端处理部件 (post-processing block) 6 提供局部合成信号 106。

与现有技术中的宽带编解码器相比, 将后端处理部件 6 的后端处理功能修改成包括增益缩放和增益量化 108, 其相应于具有原始语音信号 100 的高频部分特性的输入信号。更具体地, 可以使用原始语音信号 100 的高频部分, 以及着色的高通滤波随机噪声 134, 154 来确定如图 3 所示的与语音编解码器结合起来进行描述的如方程 4 所示的高带信号缩放因子。后端处理部件 6 的输出内容为后端处理语音信号 110。

图 3 说明了根据本发明的语音编解码器 10 中的后端处理功能的详细结构。如图所示, 使用随机噪声发生器 20 来提供 16kHz 仿真信号 130。LPC 合成滤波器 22 使用 LPC 参数 104 对随机噪声 130 进行着色, 此 LPC 参数 104 基于语音信号 100 的低频带特性由合成分析编码部件 4 (图 2) 中的编码比特流来提供。从着色的随机噪声 132 和高通滤波器 24 抽取出频率为 6.0-7.0kHz 的着色的高频部分 134。在原始语音样本 100 中频率范围为 6.0-7.0kHz 的高频部分 112 也可以通过高通滤波器 12 提取。使用高频部分 112 和 134 的能量来确定增益平衡部件 14 的高带信号缩放因子 g_{scaled} , 根据以下方程:

$$g_{scaled} = \sqrt{\{(s_{hp}^T s_{hp}) / (e_{hp}^T e_{hp})\}}$$

(4)

其中, s_{hp} 为 6.0-7.0kHz 带通滤波原始语音信号 112, e_{hp} 为 LPC 合成(着色)及带通滤波随机噪声 134. 由参考标号 114 所表示的缩放因子 g_{scaled} 可通过增益量化模块 18 进行量化, 并且在编码比特流中进行传输, 从而接收端可以使用缩放因子对随机噪声进行缩放以实现语音信号的再现。

当前的 GSM 语音编解码器中, 非语音周期中的无线电传输过程由不连续传输 (DTX) 函数中止。DTX 函数将会帮助减少不同部分之间的干扰, 同时提高通信系统的容量。DTX 函数依赖于语音激活检测 (VAD) 算法来确定输入信号 100 代表语音还是噪声, 从而防止在激活语音周期内关闭发射机。VAD 算法由参考标号 98 表示。此外, 当发射机在非激活语音周期内被关闭时, 为了消除连接失败的影响, 由接收机提供数量较小的称为“舒适噪声” (CN) 的背景噪声。VAD 算法这样来设计, 以便当监测到非激活语音周期之后, 允许有一个称之为释放延迟或保持延迟的时间段。

根据本发明, 激活语音中的缩放因子 g_{scaled} 可以根据方程 4 进行估算。然而, 完成激活语音到非激活语音的自适应之后, 由于比特速率的限制以及传输系统本身, 增益参数不能够在舒适噪声比特流中进行传输。因此, 同现有技术中的宽带编解码器的实现方式一样, 在非激活语音中, 在接收端不使用原始语音信号来确定缩放因子。因而, 可以从非激活语音中的基层信号中可以隐含地估算出增益值。与之相反, 在基于高频增强层中信号的语音周期中使用显式增益量化。在激活语音转换到非激活语音的过程中, 不同缩放因子之间的转换可能会导致合成信号中的声音瞬变 (audible transients)。为了降低这些声音瞬变, 可以使用增益自适应模块 16 来改变缩放因子。根据本发明, 当语音激活确定 (VAD) 算法的释放延迟周期开始时, 自适应开始启动。为了该目的, 为增益自适应模块 16 提供表示 VAD 判决的信号 190。此外, 不连续传输 (DTX) 的释放延迟周期也将被用来完成增益自适应。DTX 的释放延迟周期之后, 可以使用不通过原始语音信号确定的缩放因子。用来调整缩放因子的整个增益自适应过程可以根据以下方程得以实现:

$$g_{total} = \alpha g_{scaled} + (1.0 - \alpha) f_{est}$$

(5)

其中, f_{est} 由方程 3 来确定并由参考标号 115 表示, α 为自适应参数, 由

以下方程给出：

$$\alpha = (\text{DTXhangovercount})/7$$

(6)

因而，在激活语音中， α 等于 1.0，原因在于 DTX 释放延迟计数等于 7。
5 在从激活到非激活语音的瞬变过程中，DTX 释放延迟计数从 7 降低到 0。
从而，在该瞬变中， $0 < \alpha < 1.0$ 。在非激活语音中，或者是接收到第一舒适
噪声参数之后， $\alpha = 0$ 。

在此情形下，由语音激活监测和源编码比特速率所驱动增强层编码
将依照不同的输入信号周期进行缩放。在激活语音中，增益量化由增强
10 层明显地确定，该增强层包括随机噪声增益参数确定和自适应。在瞬变
周期内，显式确定的增益值将向隐式估算值进行自适应。在非激活语音
中，增益值由基层信号进行隐式估算。因而，高频增益层参数将不会传
输到非激活语音的接收端上。

增益值自适应的好处在于可以获得从激活到非激活语音处理过程完
15 成缩放的高频部分的平滑瞬变。由增益自适应模块 16 所确定且由参考编
号 116 所表示的自适应缩放增益值 g_{total} ，将由增益量化模块 18 作为一套
量化增益参数 118 进行量化。此套增益参数 118 可以被加入到编码比特
流中去，并传输到接收端进行解码。需要注意的是，量化增益参数 118
可以作为查表存储起来，从而可以通过增益索引访问（未示出）。

20 对于自适应后的缩放增益值 g_{total} ，为了降低从激活语音到非激活语音
转换过程中合成信号的瞬变，可以对解码过程中的高频随机噪声进行缩
放。最后，合成的高频部分加入到从编码器的 A-b-S 环路所接收到的上
抽样和内插信号中。在每个 5 毫秒子帧中，彼此独立地实现能量缩放
的后端处理。随着 4 比特电报密码本被用来对高频随机部分增益值进行量
25 化，整个比特率为 0.8 kbit/s。

显式确定的增益值（来自高频增强层上）和隐式估算的增益值（来自
基层，或仅在低频带，信号）之间的增益自适应可以在增益值量化之前
在编码器中完成，如图 3 中所示。在这种情况下，根据方程 5，进行编码
并且传输到接收端的增益值参数为 g_{total} 。可替换地，增益值自适应可仅仅
30 在 VAD 标记显式非语音信号已经开始之后于 DTX 释放延迟周期内的解码
器中实现。在这种情况下，增益参数的量化在编码器中实现，同时在解
码器中实现增益值自适应，传输到接收端上的增益参数可以根据方程 4

简化为 g_{scaled} 。估算的增益值 f_{est} 值可以在解码器中通过使用合成语音信号得以确定。增益值自适应也可以在解码器接收到第一无声描述 (SID first) 之前在舒适噪声周期的初始阶段于解码器中实现。如同前面的情况一样, g_{scaled} 在编码器中量化同时在编码比特流中进行传输。

5 本发明中解码器 30 如图 4 所示。如图所示, 解码器 30 用来合成来自编码参数 140 的语音信号 110, 该编码参数 140 包括 LPC、音调和激励参数 104 以及增益参数 118 (见图 3)。, 解码模块 32 从编码参数 140 提供一套量化 LPC 参数 142。后端处理模块 34 从所接收的语音信号低带部分的 LPC、音调和激励参数 142 产生合成低带语音信号, 如同在现有技术中的解码器一样。后端处理模块 34 由局部产生的随机噪声产生合成高频部分, 它是基于包括语音高频部分的输入信号特性的增益参数之上的。

图 5 给出了解码器 30 的通用后端处理结构。如图 5 所示, 增益参数 118 通过增益去量化 (dequantilization) 部件 38 进行去量化处理。如果增益自适应已经在编码器中完成, 如图 3 中所示, 那么接下来解码器中的相关增益自适应功能将会在舒适噪声周期初期将去量化之后的增益值 144 (g_{total} , $\alpha=1.0$ 以及 $\alpha=0.5$) 自适应为所估算的缩放增益值 $f_{est}(\alpha=0)$, 而无需 VAD 判决信号 190。然而, 如果仅仅在信号 190 提供的 VAD 标记指示非语音信号开始之后, 在 DTX 释放延迟周期内的解码器中进行增益值自适应, 那么增益值自适应部件 40 将根据方程 5 来确定缩放因子 g_{total} 。因此, 当未接收到增益参数 118 时, 在不连续传输过程的初始阶段, 增益值自适应部件 40 将使用估算缩放增益值 f_{est} 消除瞬变, 如参考标号 145 表示。因而, 如增益自适应模式 40 所提供的那样, 根据方程 5 确定缩放因子 146。

25 如图 4 所示的后端处理单元 34 中的随机噪声部分的着色和高通过滤类似于图 3 中所示编码器 10 的后端处理操作。如图所示, 随机噪声发生器 50 用来提供仿真信号 150, 它根据所接收到的 LPC 参数 104 由 LPC 合成滤波器 52 着色。着色的仿真信号 152 由高通滤波器 54 进行滤波操作。然而, 在编码器 10 (图 3) 中提供着色的、高通滤波随机噪声 134 的目的在于产生 e_{hp} (方程 4)。在后端处理模块 34 中, 着色的、高通滤波仿真信号 154 在被基于增益值自适应模块 40 所提供的自适应高带缩放因子 146 上的增益调整模块 56 缩放之后, 被用来产生出合成高频信号

160。最后，高频增强层的输出部分 160 被加入到由基解码器（未示出）所接收到的 16kHz 合成信号上。16kHz 合成信号在本领域是众所周知的。

需要注意到来自解码器的合成信号可以用来实现频谱倾斜 (tilt) 估算。可以使用方程 2 和 3 由解码器后端处理部分估算出参数值 f_{est} 。当出现由于各种原因，如信道带宽限制以及解码器没有接收高带增益值，而导致解码器或传输信道忽略了高带增益参数的情况时，可以缩放着色的、高通滤波随机噪声从而提供合成语音的高频部分。

总之，在宽带语音编解码器中实现高频增强层编码工作的后端处理步骤可以在编码器或者是解码器中来完成。

10 当后端处理步骤在编码器中完成时，高带信号缩放因子 g_{scaled} 从频率范围为 6.0-7.0kHz 的原始语音样本和 LPC 彩色以及带通滤波随机噪声中的高频部分中获得。另外，所估算的增益因子 f_{est} 从编码器中低带合成信号的频谱倾斜值获得。使用 VAD 判定信号来表明输入信号是处于激活语音周期内还是处于非激活语音周期内。针对不同语音周期的所有缩放因子 g_{total} 由缩放因子 g_{scaled} 和估算出的增益因子 f_{est} 运算出。可缩放的高频带信号缩放因子在编码比特流中进行量化和传输。在接收端，全部缩放因子 g_{total} 从所接收到的编码比特流（编码参数）中抽取出来。使用这一全部缩放因子来缩放解码器中所产生的着色的高通滤波随机噪声。

20 当在解码器中完成后端处理步骤时，所估算的增益因子 f_{est} 可以从解码器中的低频带合成语音中获得。这一估算出的增益因子可以用来缩放激活语音内解码器中的着色的高通滤波随机噪声。

图 6 所示为根据本发明的一个实施例所得出的移动台 200 的框图。移动台包括此设备的特有部分，如麦克风 201，数字键盘 207，显示器 206，耳机 214，发送/接收开关 208，天线 209 和控制单元 205。并且，图中给出了此移动台所特有的发送和接收部件 204 和 211。发送部件 204 包括用于编码语音信号的编码器 221。编码器 221 包括图 3 中所示编码器 10 的后端处理功能。发送部件 204 还包括实现信道编码、解密和调制以及 RF 功能的操作，而为了更清楚的表述，这些在图 5 中未给出。接收部件 211 还包括依照本发明的解码部件 220。解码部件 220 包括类似于图 5 中所示解码器 34 的后端处理单元 222。来源于麦克风 201 的信号在放大级上放大，然后在 A/D 转换器中进行数字化处理，然后发送到发送部件 204 上，尤其是发送到发送部件所包括的语音编码设备上。发送部件的发

送, 信号处理、调制以及放大, 通过发送/接收开关 208 传输到天线 209。从天线得到的所要接收的信号通过发送/接收开关 208 传输到接收部件 211, 接收部件 211 能够解调所接收的信号以及解码解密和信道编码。所得到的语音信号将通过 D/A 转换器 212 传输到放大器 213 上, 进一步地
5 传输到耳机 214。控制单元 205 控制移动台 200 的操作, 读取用户通过键盘 207 给出的控制命令, 同时通过显示器 206 向用户发送信息。

根据本发明, 图 3 所示的编码器 10 以及图 5 所示的解码器 34 的后端处理功能也可以用在电信网络 300 上, 如通常的电话网和移动台网络, 如 GSM 网络。图 7 给出了这种电信网络的框图举例。例如, 电信网络 300
10 可以包括电话交换机或相应的交换系统 360, 电信网络中的普通电话 370, 基站 340, 基站控制器 350 以及其它中心设备 355 都可以连接到其上。移动台 330 可以通过基站 340 建立到电信网络的连接。例如, 包括类似于图 5 中所示的后端处理部分 322 的解码部件 320, 可方便地放置于基站 340 中。然而, 解码部件 320 例如也可以置于基站控制器 350 或者示其它中心或交换设备 355 中。例如, 如果移动台系统在基站和基站
15 控制器之间使用的是分开的代码转换器, 为了将由无线电信道接收的编码信号转换成在电信系统中传送的标准的 64 千比特/秒信号并且反之亦然, 解码部件 320 也可以放置在这种代码转换器之中。通常, 包括后端处理部分 322 的解码部件 320 可以放置于能够将编码数据流转换成非
20 编码数据流的电信网络 300 中的任意一个元件中。解码部件 320 对来源于移动台 330 的编码语音信号进行解码和过滤, 然后语音信号可依照通常在电信网络 300 中解压缩的方式进行转换。

图 8 为说明根据本发明所得语音编码方法 500 的流程图。如所示, 由于输入语音信号 100 在步骤 510 上被接收, 语音激活监测算法 98 将在步
25 骤 520 上被使用来确定在当前周期中输入信号 110 是代表语音还是噪声。在语音周期中, 处理完的仿真噪声 152 在步骤 530 上以第一缩放因子 114 进行缩放。在噪声或非语音周期中, 处理完的仿真信号 152 在步骤 540 上以第二缩放因子进行缩放。下一个周期在步骤 520 上重复此操作过程。

30 为了提供合成语音的更高频段部分, 仿真信号或随机噪声在频率范围为 6.0-7.0kHz 上进行过滤。然而, 过滤之后的频率范围例如可以基于编解码器的采样速率而有所不同。

虽然已相对于本发明的优选实施例描述了本发明，本领域的技术人员可以理解在不偏离本发明的精神和范围的情况下，可以在其形式和细节上做出上述的以及不同的变化，省略和偏移。

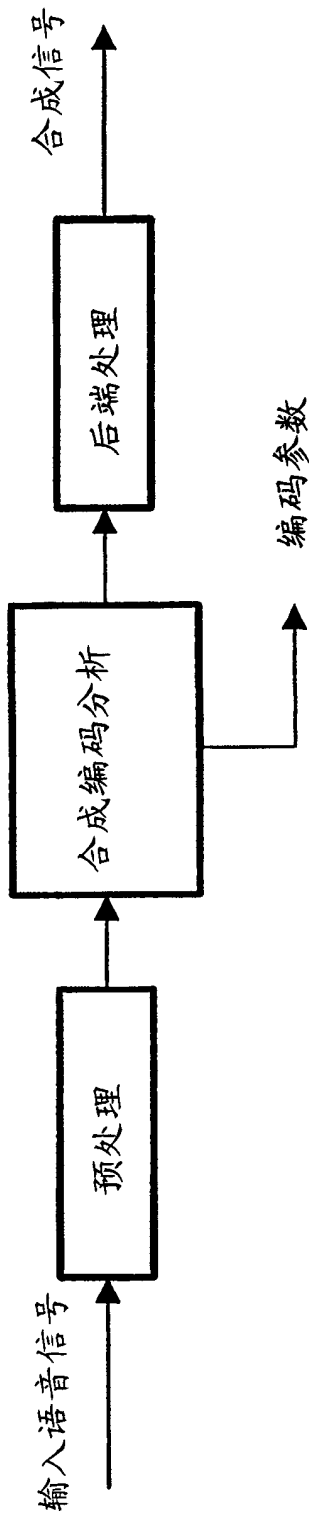


图 1
(现有技术)

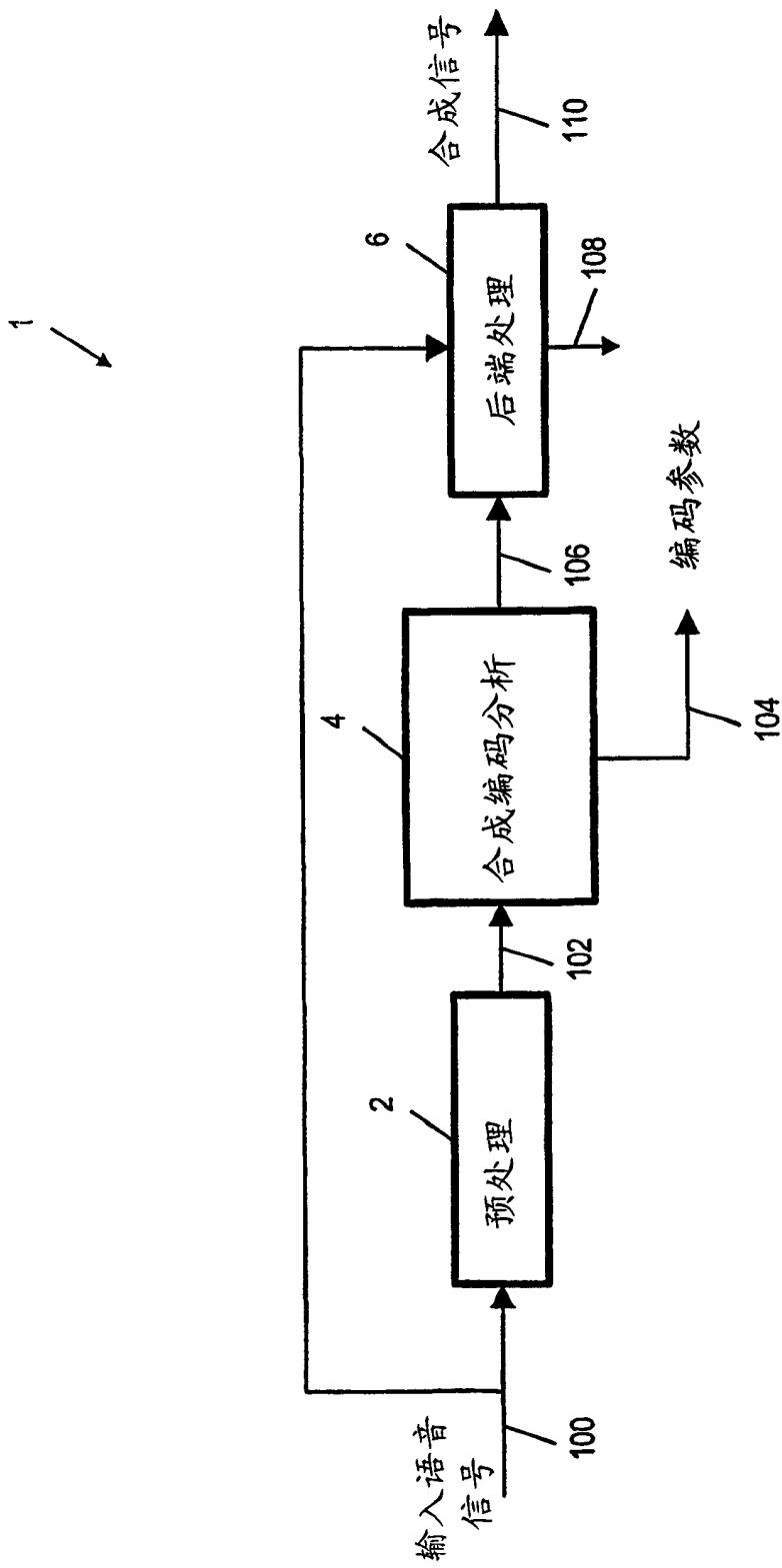


图 2

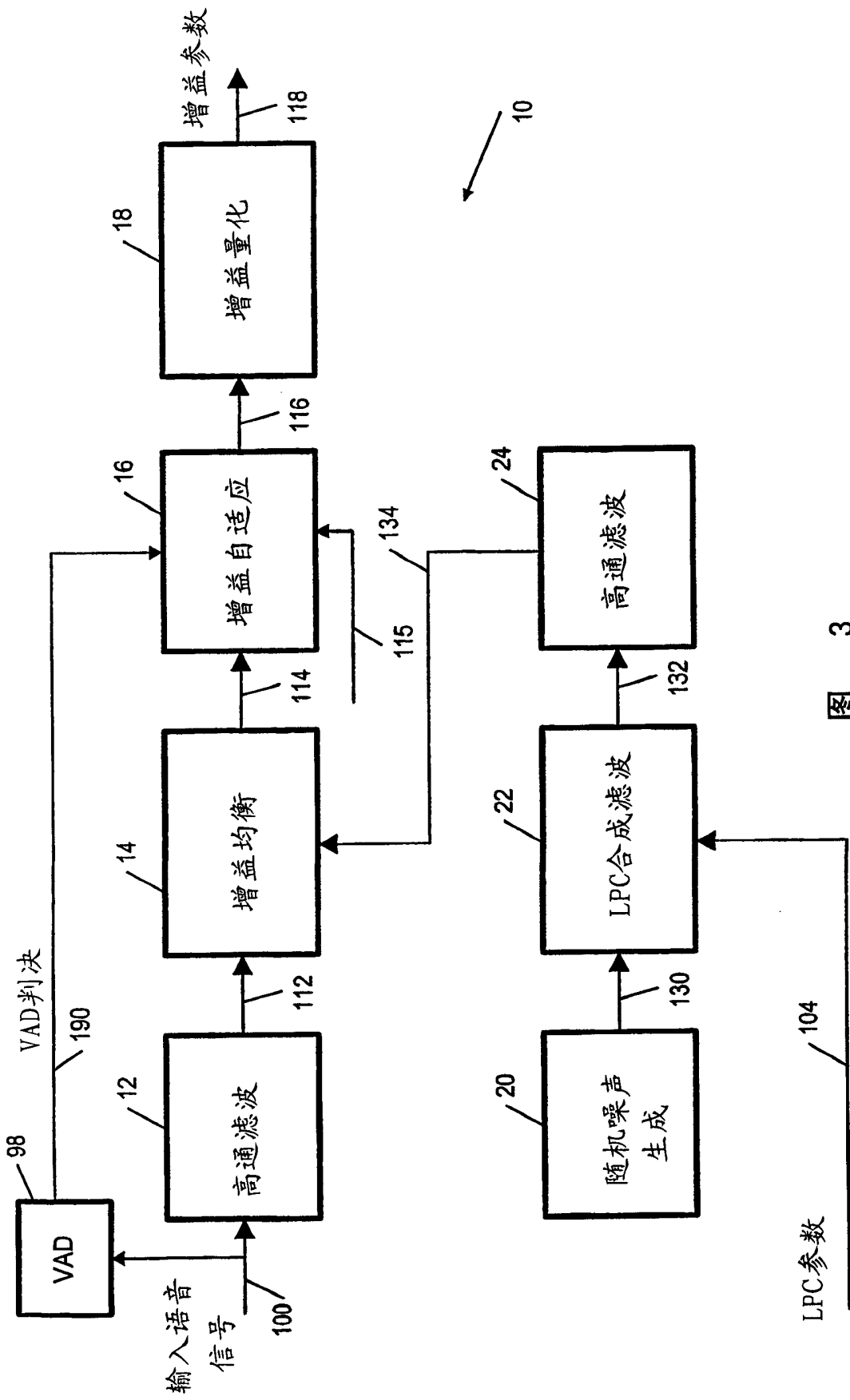


图 3

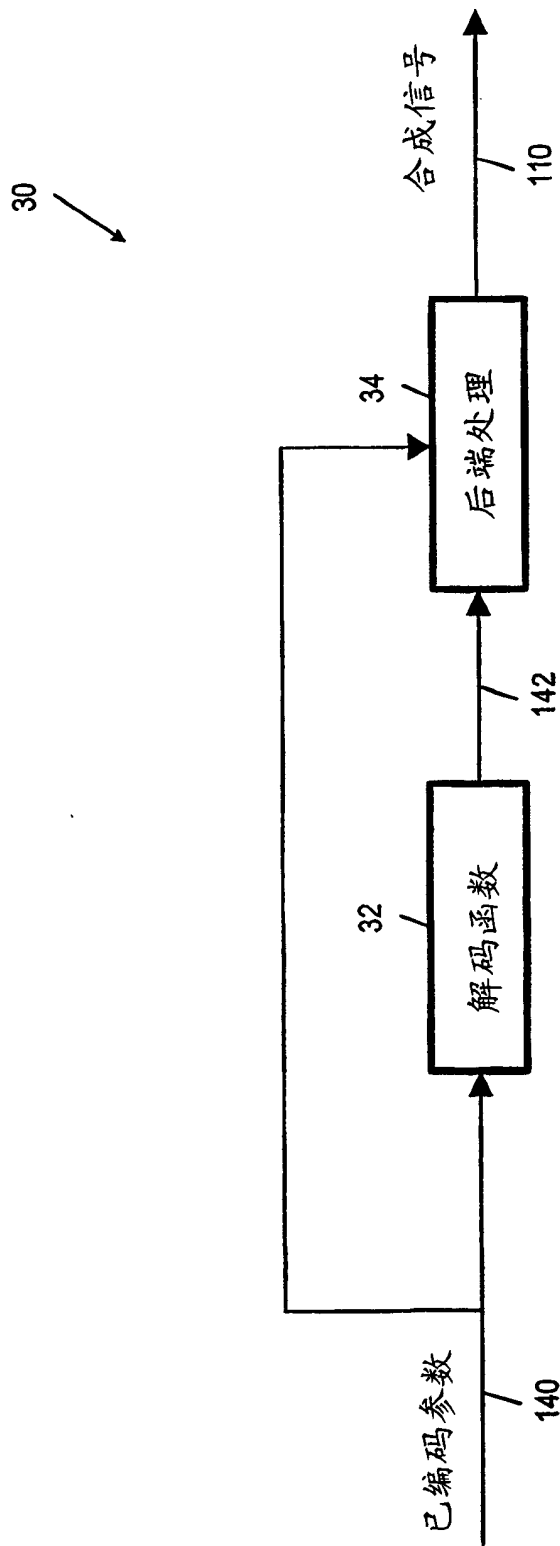


图 4

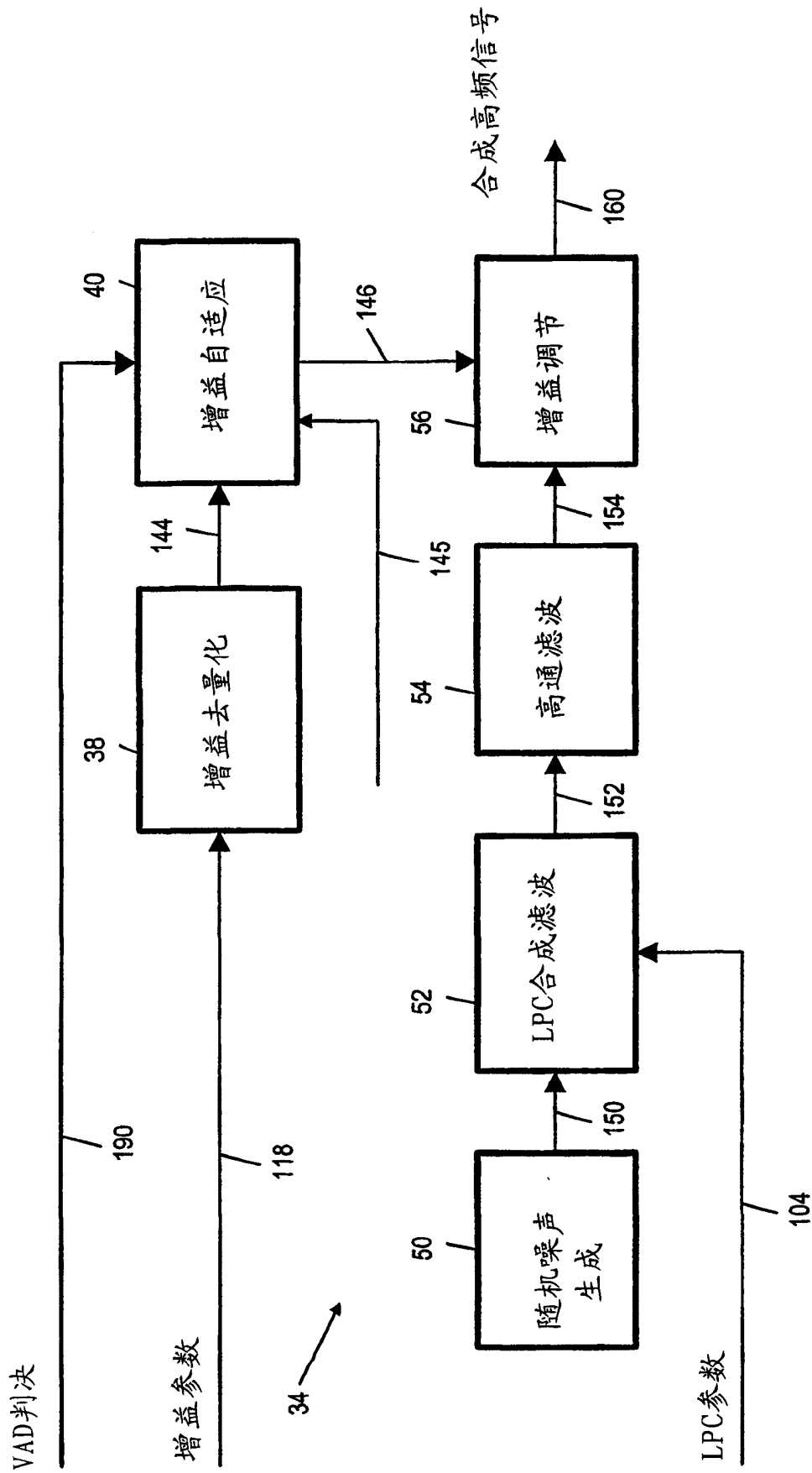
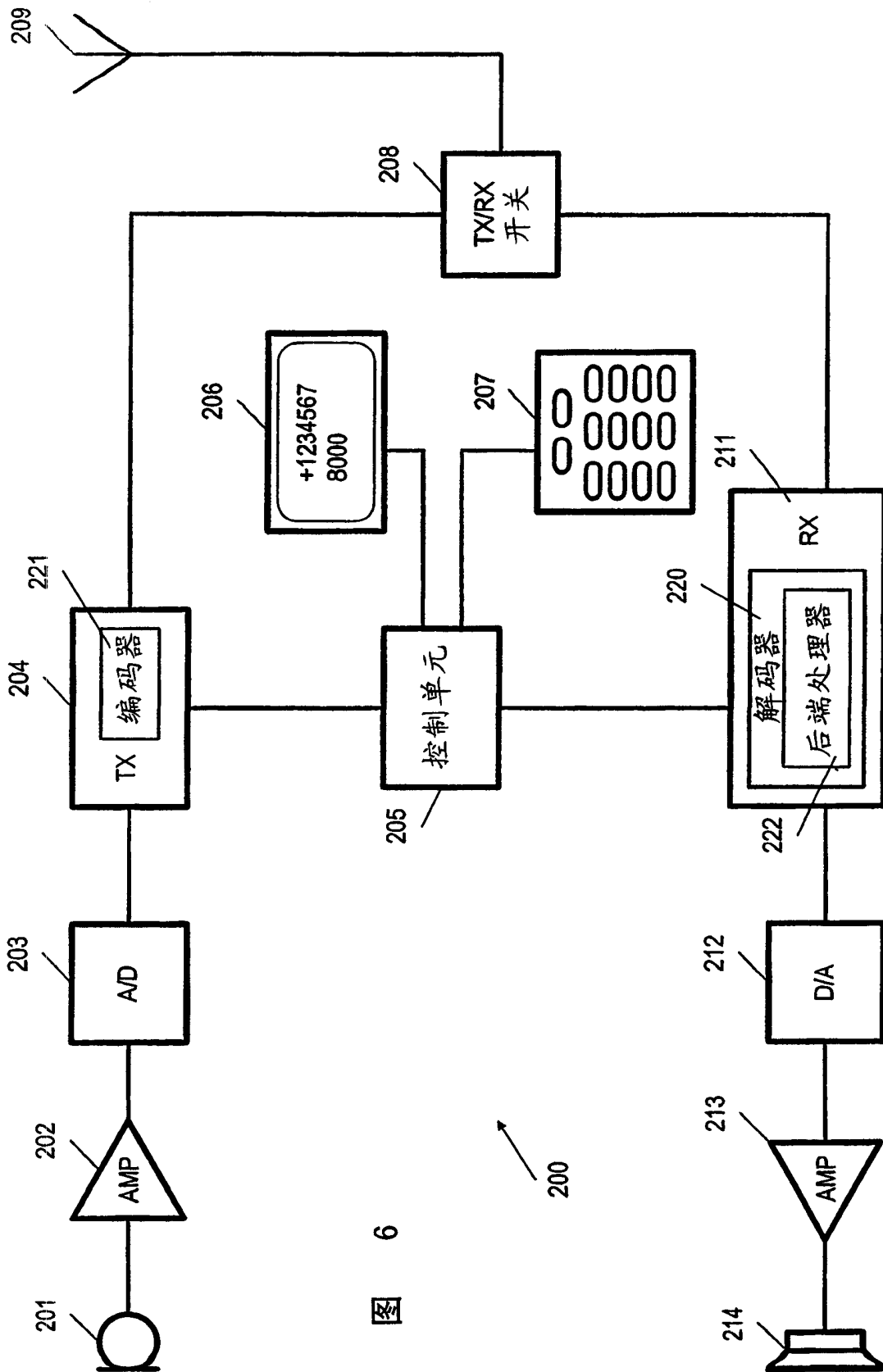


图 5



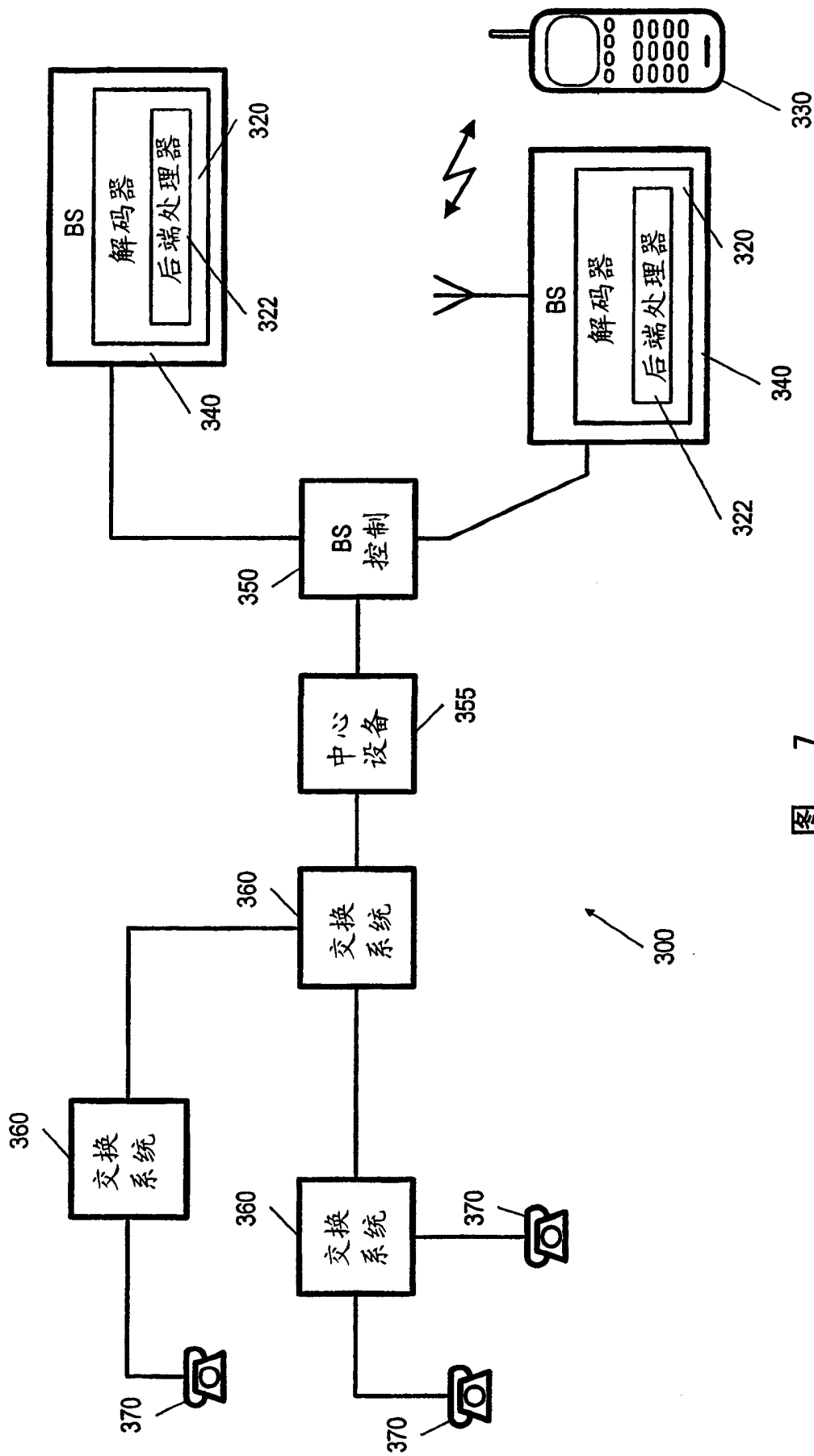


图 7

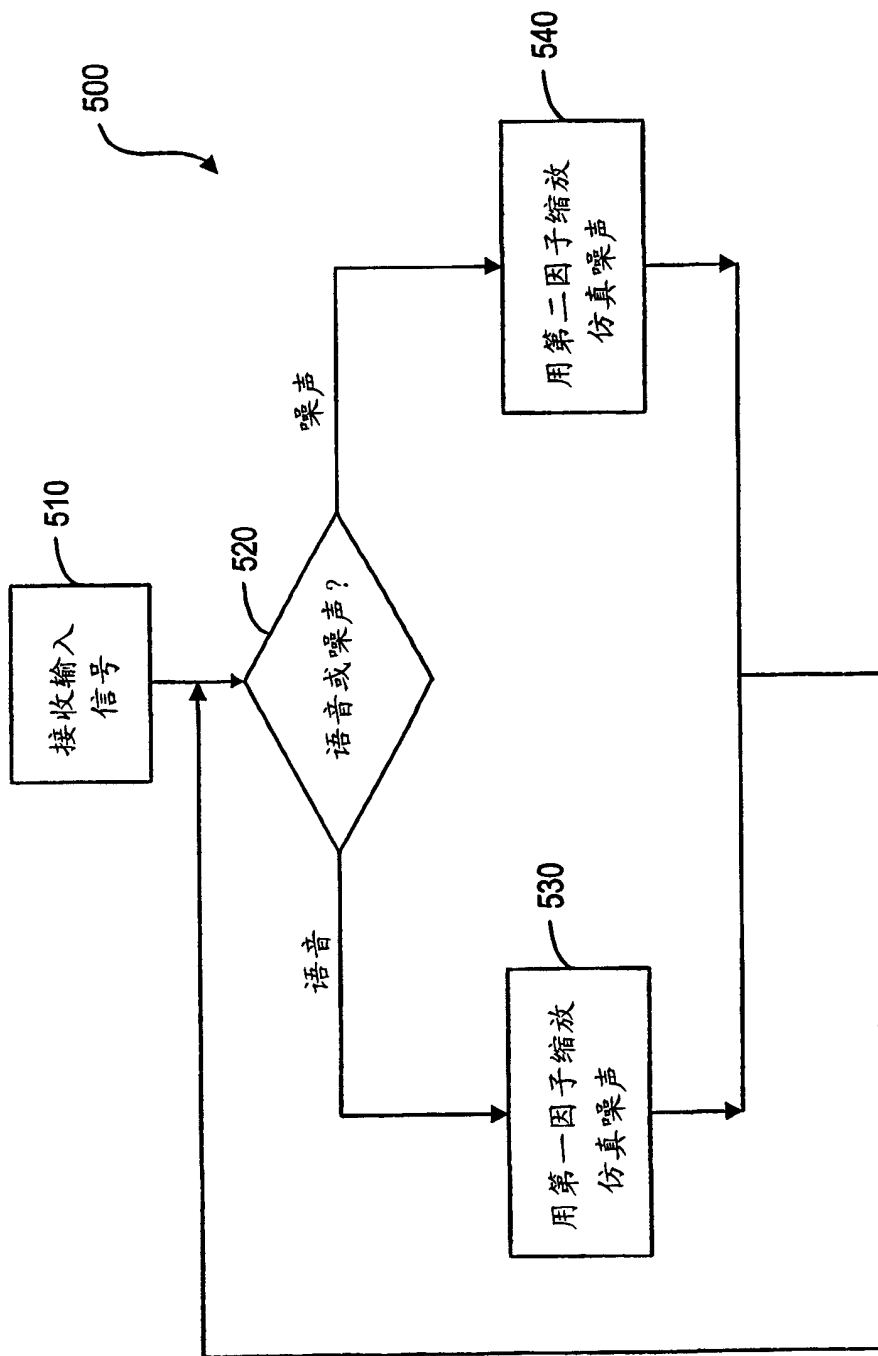


图 8