

[12] 发明专利申请公开说明书

[21] 申请号 00126277.7

[43] 公开日 2001 年 2 月 7 日

[11] 公开号 CN 1282952A

[22] 申请日 2000.6.17 [21] 申请号 00126277.7

[30] 优先权

[32] 1999.6.18 [33] JP [31] 173354/1999

[71] 申请人 索尼公司

地址 日本东京都

[72] 发明人 前田祐児 西口正之

[74] 专利代理机构 中国专利代理(香港)有限公司

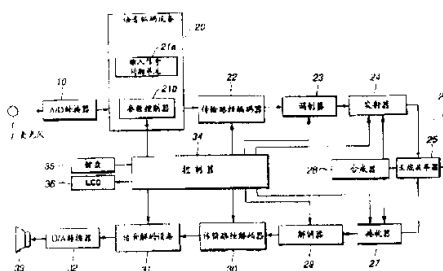
代理人 张志醒

权利要求书 3 页 说明书 22 页 附图页数 17 页

[54] 发明名称 语音编码方法和装置, 输入信号判别方法, 语音解码方法和装置以及程序提供介质

[57] 摘要

在语音编解码器中, 通过给予语音间隔中具有关键意义的有声语音相对大量的比特, 从而通过减少分配给无声声音和背景噪音的比特数, 来减少传输比特的总数从而减少传输比特的平均数。为此, 该系统包括一个用来计算输入端 1 提供的滤波输入语音信号的均方根值(有效值)的 rms 计算单元 2、用来通过 rms 值计算有效值的稳态电平计算单元 3、用来以稳态电平计算单元 3 的输出 min-rms 除 rms 计算单元 2 的输出 rms 值从而得到商 rms_g 的除法器 4、以及用于通过对数振幅差分计算单元 8 的对数振幅差分 wdif 输出一个判别标记 dec-flag 的模糊推理单元 9。



ISSN 1008-4274



权 利 要 求 书

1. 一种以输入语音信号的有声间隔和无声间隔之间的不同比率进行编码的语音编码装置，包括：

5 输入信号检验装置，用来在时间轴上将输入语音信号以预置单位进行划分并且根据信号电平和预置单位的频谱包络的时间变化来检验无声间隔是背景噪音间隔还是语音间隔；

其中

10 编码比特的分配在背景噪音间隔的参数、语音间隔的参数和有声间隔的参数之间是不同的。

2. 根据权利要求 1 的语音编码装置，其中
无声间隔的参数比特率低于有声间隔参数的比特率。

3. 根据权利要求 1 的语音编码装置，其中
背景噪音间隔的参数比特率低于语音间隔的参数比特率。

15 4. 根据权利要求 1 的语音编码装置，其中
在根据背景噪音间隔中的信号电平和频谱包络的时间变化的控制下产生表示所述背景噪音间隔中出现还是消失背景噪音参数更新的信息。

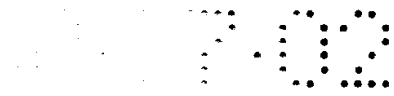
5. 根据权利要求 1 的语音编码装置，其中
20 如果背景噪音间隔中的信号电平和频谱包络的时间变化较小，则发出表示背景噪音间隔的信息和表示背景噪音参数不更新的信息，并且其中

如果背景噪音间隔中的信号电平和频谱包络的时间变化较大，则发出表示背景噪音间隔的信息、更新的背景噪音参数和表示背景噪音参数更新的信息。

6. 根据权利要求 5 的语音编码装置，其中
25 为限制背景噪音间隔中表示背景噪音的参数的持续时间大于一个预置时间，至少在一个预置时间间隔内更新背景噪音参数。

7. 根据权利要求 6 的语音编码装置，其中
所述背景噪音参数是表示 CELP 激发信号的增益参数的频谱包络或者索引的 LPC 系数。

8. 一种以输入语音信号的有声间隔和无声间隔之间的不同比率进行编码
30 的语音编码方法，包括：



输入信号检验步骤，用来在时间轴上将输入语音信号以预置单位进行划分并且根据信号电平和预置单位的频谱包络的时间变化来检验无声间隔是背景噪音间隔还是语音间隔；

其中

5 编码比特的分配在背景噪音间隔的参数、语音间隔的参数和有声间隔的参数之间是不同的。

9. 一种对输入信号进行检验的方法，包括：

以预置单位在时间轴上划分输入语音信号并且以预设单位寻找信号电平的时间改变的步骤，

10 以该单位寻找频谱包络时间改变的步骤，和
根据所述信号电平及所述频谱包络的时间改变检验可能出现背景噪音的步骤。

10. 根据权利要求 9 的对输入信号进行检验的方法，其中
使用模糊推理检验背景噪音出现的可能。

15 11. 一种将不同比特分配的编码比特解码为无声间隔参数和有声间隔参数的解码装置，包括：

检验所述编码比特中的间隔是语音间隔还是背景噪音间隔的检验装置；和
通过使用当前或当前及过去接收到的 LPC 系数、当前或当前及过去接收到的 CELP 增益索引和内部随机产生的 CELP 整形索引在背景噪音间隔解码所
20 述编码比特的解码器件。

12. 根据权利要求 11 的解码装置，其中

所述解码器件通过内插过去和当前接收的 LPS 系数，或者通过内插过去接收的 LPS 系数产生背景噪音间隔信号，其中使用随机数产生 LPC 系数的内插系数。

25 13. 一种将不同比特分配的编码比特解码为无声间隔参数和有声间隔参数的解码方法，包括：

检验所述编码比特中的间隔是语音间隔还是背景噪音间隔的检验步骤；和
通过使用当前或当前及过去接收到的 LPC 系数、当前或当前及过去接收到的 CELP 增益索引和内部随机产生的 CELP 整形索引在背景噪音间隔解码所
30 述编码比特的解码步骤。



14. 一种用于提供在输入语音信号的有声间隔和无声间隔之间以不同比率进行编码的语音编码程序的介质，该程序包括：

输入信号检验步骤，用来在时间轴上将输入语音信号以预置单位进行划分，并且

5 根据信号电平和预置单位的频谱包络的时间变化来检验无声间隔是背景噪音间隔还是语音间隔；

其中

编码比特的分配在背景噪音间隔的参数、语音间隔的参数和有声间隔的参数之间是不同的。

10 15. 一种用于提供将以不同比特分配进行编码的传输比特解码为无声间隔参数和有声间隔参数的语音解码程序的介质，该程序包括：

检测所述编码比特中的间隔是语音间隔还是背景噪音间隔的检测步骤；以及

15 通过使用当前或当前及过去接收到的 LPC 系数、当前或当前及过去接收到的 CELP 增益索引和内部随机产生的 CELP 整形索引在背景噪音间隔解码所述编码比特的解码步骤。

语音编码方法和装置，输入信号判别方法，
语音解码方法和装置以及程序提供介质

5

本发明涉及当无声间隔内的比特率与有声间隔内的比特率不同时对输入语音信号编码的编码方法和装置。本发明还涉及对用该编码方法和装置进行编码和传输的编码数据进行解码的方法和装置，并且涉及用于通过软件相关技术执行编码方法和解码方法的程序提供介质。

10 近来，在需要传输路径的通信领域，为了实现高效地使用传输带宽，希望在传输输入信号前根据输入信号的种类，诸如被分为例如有声声音和无声声音的语音信号间隔，或者背景噪音间隔，来改变传输的输入信号的编码率。

例如，如果给出的间隔被检验为背景噪音间隔，在解码设备不会特意产生背景噪音的情况下，就会希望不要传输编码参数而只是简单地对该间隔进行静
15 噪。

然而这样会导致不自然呼叫，因为背景噪音是叠加在通信一方发出的语音上，并且在没有语音时会突然产生无声状态。

考虑到这一点，常规处理是这样的，即，如果给出的间隔被检验为背景噪音间隔，那么不发出一些编码参数，这时解码设备通过重复使用过去的参数产
20 生背景噪音。

然而，如果一直以重复的方式使用过去的参数，会得到这样的印象，即，噪音自身具有音调，从而产生非自然的噪音。只要线谱对（LSP）参数保持相同，即使电平等发生改变也会发生这样的情况。

因此本发明的一个目的是提出一种语音编码方法和装置，输入信号判别方
25 法，语音解码方法和装置以及程序提供介质，其中，在语音编解码器中，相对大量的传输比特被给予语音间隔的语言语音判据，使非语言语音和背景噪音序列中减少的比特数抑制传输比特的总数并且减少了传输比特的平均量。

一方面，本发明提出了一种在输入语音信号的有声间隔和无声间隔之间以不同比率进行有效编码的语音编码装置，包括输入信号检验装置，用来在时间
30 轴上将语音信号以预置单位进行划分并且根据信号电平和预置单位的频谱包络



的时间变化来检验无声间隔是背景噪音间隔还是语音间隔，其中编码比特的分配在背景噪音间隔的参数、语音间隔的参数和有声间隔的参数之间是不同的。

另一方面，本发明提出了一种在输入语音信号的有声间隔和无声间隔之间以不同比率进行有效编码的语音编码方法，包括输入信号检验步骤，用来在时间轴上将语音信号以预置单位进行划分并且根据信号电平和预置单位的频谱包络的时间变化来检验无声间隔是背景噪音间隔还是语音间隔，其中编码比特的分配在背景噪音间隔的参数、语音间隔的参数和有声间隔的参数之间是不同的。

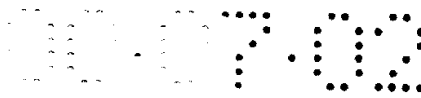
再一方面，本发明提出了一种对输入信号进行检验的方法，包括以预置单位划分输入语音信号并且以预置单位寻找信号电平时间改变的步骤，以该单位寻找频谱包络时间改变的步骤，和根据信号电平时间改变及频谱包络时间改变检验可能出现的背景噪音的步骤。

另一方面，本发明提出了一种将不同比特分配的编码比特解码为无声间隔参数和语音间隔参数的解码装置，包括检验所述编码比特中的间隔是语音间隔还是背景噪音间隔的检验装置，以及通过使用当前或当前及过去接收到的 LPC 系数、当前或当前及过去接收到的 CELP 增益索引信号和如果所述检验装置去掉了表示背景噪音间隔的信息则在内部随机产生的 CELP 整形索引信号而在背景噪音间隔解码编码比特的解码器件。

另一方面，本发明提出了一种将不同比特分配的编码比特解码为无声间隔参数和语音间隔参数的解码方法，包括检验所述编码比特中的间隔是语音间隔还是背景噪音间隔的检测步骤，以及通过使用当前或当前及过去接收到的 LPC 系数、当前或当前及过去接收到的 CELP 增益索引信号和在内部随机产生的 CELP 整形索引信号而在背景噪音间隔解码编码比特的解码步骤。

另一方面，本发明提出了一种用于提供在输入语音信号的有声间隔和无声间隔之间以不同比率进行编码的语音编码程序的介质，其中该程序包括输入信号检验步骤，用来在时间轴上将输入语音信号以预置单位进行划分并且根据信号电平的时间变化和预置单位的频谱包络来检验无声间隔是背景噪音间隔还是语音间隔。编码比特的分配在背景噪音间隔的参数、语音间隔的参数和有声间隔的参数之间是不同的。

另一方面，本发明提出了一种用于提供将以不同比特分配进行编码的传输



比特解码为无声间隔参数和有声间隔参数的语音解码程序的介质，其中该程序包括检验编码比特中的间隔是语音间隔还是背景噪音间隔的检验步骤，以及通过使用当前或当前及过去接收到的LPC系数、当前或当前及过去接收到的CELP增益索引信号和在内部随机产生的CELP整形索引信号而在背景噪音间隔解码编码比特的解码步骤。

通过根据本发明的解码方法和装置，能够保持语音信号的连续性从而解码高质量的语音。

此外，通过根据本发明的程序提供介质，能够使计算机系统保持语音信号的连续性从而解码高质量的语音。

10 图 1 是表示体现本发明的便携电话设备的结构方框图；

图 2 表示便携电话设备除输入信号判别单元和参数控制器之外的语音编码设备的内部详细结构；

图 3 表示输入信号判别单元和参数控制器的详细结构；

图 4 是表示计算稳态电平均方值的处理流程图；

15 图 5 表示模糊推理单元的模糊规则；

图 6 表示模糊规则中关于信号电平的隶属函数；

图 7 表示模糊规则中关于频谱的隶属函数；

图 8 表示模糊规则中关于推理结果的隶属函数；

图 9 表示模糊推理单元中的推理的一个特定的例子；

20 图 10 是表示在参数产生单元中确定传输参数的处理部分的流程图；

图 11 是表示在参数产生单元中确定传输参数的剩余处理部分的流程图；

图 12 表示作为一个例子使用在 MPEG4（运动图象专家组 4）中采用的语音编解码器 HVXC（谐波向量激发编码）编码每一个条件下的比特；

图 13 是表示语音解码装置的详细结构方框图；

25 图 14 是表示语音编码设备的基本和周边部分的结构方框图；

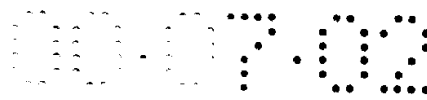
图 15 是 LPC 参数再现控制单元的 LPC 参数再现部分的详细流程图；

图 16 表示标题比特的结构；

图 17 是表示应用本发明的传输系统的方框图；

图 18 是构成传输系统的服务器方框图；

30 图 19 是构成传输系统的客户终端方框图。



参照附图，将详细描述根据本发明的编码方法和装置以及语音解码方法和装置的优选实施例。

基本上，这样的系统可以这样描述，在发射侧分析语音以找到编码参数，发射编码参数并且在接收侧合成语音。特别地，发射侧根据输入语音的特性将编码模式分类，并且改变比特率来消除传输比特率的平均值。

一个特定的例子是便携电话设备，其结构如图 1 所示。该便携电话设备以图 1 所示语音编码设备 20 和语音解码设备 31 的形式使用根据本发明的编码方法和装置以及解码方法和装置。

语音编码设备 20 进行编码以便与有声 (V) 间隔的比特率相比降低输入语音信号的无声 (UV) 间隔的比特率。语音编码设备 20 还在无声间隔将背景噪音间隔 (无语音间隔) 和语音间隔相互判别从而在无声间隔以相当低比特率进行编码。它还从语音间隔判别无语音间隔从而将判别结果发射到语音解码设备 31。

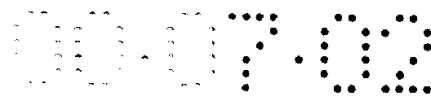
在语音编码设备 20 中，通过输入信号判别单元 21a 在输入语音信号中判别无声间隔和有声间隔或者在无声间隔判别无语音间隔和语音间隔。接着将详细解释该输入信号判别单元 21a。

首先，解释发射侧的结构。从麦克风 1 输入的语音信号通过 A/D 转换器 10 转换为数字信号并且由语音编码设备 20 以不同的比率编码。接着传输路径编码器 22 将编码信号编码从而使语音质量不容易被传输路径的质量所损坏。调制器 23 调制结果信号并且由发射器 24 为发射做处理，以便通过天线共享器 25 经天线 26 发射。

另一方面，接收侧的语音解码器 31 接收一个标记，该标记表示给出的间隔是语音间隔还是无语音间隔。如果间隔是无语音间隔，语音解码器 31 使用当前或当前及过去接收的 LPC 系数、当前或当前及过去接收的 CELP (码激发线性预测) 增益索引信号和在解码器中随机产生的 CELP 的整形索引信号，对所述间隔解码。

下面解释接收侧的结构。接收器 27 通过天线共享器 25 接收天线 26 捕获的电磁波并且由解调器 13 解调，从而接着由传输路径解码器 30 修正传输差错。D/A 转换器 32 将结果信号转换回由扬声器 33 输出的模拟语音信号。

控制器 34 控制上述各个部分，同时合成器 28 将发射/接收频率通知发射



器 24 和接收器 27。键盘 35 和 LCD 显示器 36 作为人机交互界面。

下面参照图 2 和 3 详细解释语音编码设备 20。图 2 给出了语音编码设备 20 内部的编码单元的详细结构，不包括输入信号判别单元 21a 和参数控制单元 21b。图 3 给出了输入信号判别单元 21a 和参数控制单元 21b 的详细结构。

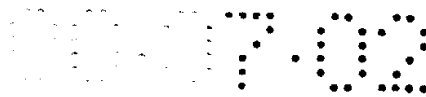
5 将采样频率为 8 kHz 的语音信号提供给输入端 101。在高通滤波器 (HPF) 109 中将不需要波段的信号从输入语音信号中去掉并且接着提供给输入信号判别单元 21a, LPC (线性预测编码) 分析量化单元 113 的 LPC 分析电路 132 和 LPC 反滤波电路 111。

参照图 3, 输入信号判别单元 21a 包括一个均方根计算单元 2, 用来计算
10 滤波输入语音信号的 rms (均方根) 值并提供给输入端 1, 稳态电平计算单元 3, 用来通过有效值的均方根计算有效值的稳态电平, 还包括除法器 4, 用于用稳态电平计算单元 3 的输出 \min_rms 除均方根计算单元 2 的输出均方根以得到商 rms_g , 还包括 LPC 分析单元 5, 用于对来自输入端 1 的输入语音信号进行 LPC 分析以找到 LPC 系数 $\alpha(m)$, 还包括 LPC 对数倒频谱系数计算单元 6, 用来将
15 来自 LPC 分析单元 5 的 LPC 系数 $\alpha(m)$ 转换为 LPC 对数倒频谱系数 $C_L(m)$, 和对数幅值计算单元 7, 用来从 LPC 对数倒频谱系数计算单元 6 的 LPC 对数倒频谱系数 $C_L(m)$ 找到平均对数振幅 $\log Amp(i)$ 。输入信号判别单元 21a 包括对数振幅差分计算单元 8, 用于从对数振幅计算单元 7 的平均对数振幅 $\log Amp(i)$ 中找到对数振幅差分 $wdif$, 以及一个模糊推理单元 9, 用于通过除法器 4 的 rms_g
20 和对数振幅差分计算单元 8 的对数振幅差分 $wdif$ 输出一个判别标记 $decflag$ 。同时, 后面将解释到, 如图 2 所示, 示出包括 V/UV 判断单元 115 的编码单元, 该编码单元用于从输入语音信号输出 $idVUV$ 判断结果并编码各种参数以输出编码参数, 为便于描述, 该编码单元作为语音编码单元 13 在图 3 示出。

参数控制单元 21b 包括一个计数器控制器 11, 用来根据来自 V/VU 判断单元 115 的 $idVUV$ 判断结果和来自模糊推理单元 9 的判断结果 $decflag$ 设置背景
25 噪音计数器 $bgnCnt$, 和参数产生单元 12, 用于确定一个更新标记 $Flag$ 并将该标记在输出端 106 上输出。

现在详细解释输入信号判别单元 21a 以及参数控制单元 21b 的不同部分的操作。首先, 下面是输入信号判别单元 21a 的不同部分的操作。

30 均方根计算单元 2 将以 8kHz 频率采样的输入语音信号分为 20 毫秒基本帧



(160 个样值)。对于语音分析，在重叠的 32 毫秒帧 (256 个样值) 上执行。输入信号 $s(n)$ 被分为 8 个间隔并且由下面的公式 (1) 得到间隔幂 $ene(i)$:

$$ene(i) = \sum_{n=0}^{31} s(32i + n)^2, (i=0, \dots, 7) \quad \dots (1)$$

通过下面的公式 (2) 或者 (3) 由 $ene(i)$ 得到使前一边相对后一边信号间

5 隔部分比率 $ratio$ 最大的边界 m :

$$ratio = \frac{\frac{1}{m} \sum_{i=0}^{m-1} ene(i)}{\frac{1}{8-m} \sum_{i=0}^7 ene(i)} \quad \dots (2)$$

$$ratio = \frac{\frac{1}{8-m} \sum_{i=0}^7 ene(i)}{\frac{1}{m} \sum_{i=0}^{m-1} ene(i)} \quad \dots (3)$$

这里公式 (2) 是当前一部分大于后一部分时的比率，而公式 (3) 是当后一部分大于前一部分时的比率。

10 需要指出的是 m 限制为 $m=2, \dots, 6$ 。

这样根据下面的公式(4)或(5),信号有效值 rms 就能够通过较大的前一部分或后一部分的均方根,以及这样得到的边界值 m 而得到:

$$rms = \sqrt{\frac{1}{32m} \sum_{i=0}^{m-1} ene(i)} \quad \dots (4)$$

$$rms = \sqrt{\frac{1}{32(8-m)} \sum_{i=0}^7 ene(i)} \quad \dots (5)$$

15 需要注意的是公式 (4) 是当前一部分大于后一部分时的有效值 rms ，而公式 (5) 是当后一部分大于前一部分时的有效值 rms 。

通过前面描述的有效值 rms ，稳态电平计算单元 3 按照图 4 的流程图计算有效值的稳态电平。在步骤 S1，根据过去帧的有效值 rms 的稳态检测计数器的状态 st_cnt 是否小于 4。如果在步骤 S1 检测的结果是 YES，则稳态电平计算单元 3 执行步骤 S2 来将过去连续 4 帧的第二大的 rms 值设置为 $near_rms$ 。接着，在步骤 3，从前面的 rms 中找到最小值 $minval$ ，它是 $far_rms(i)$ ($i=0,1$) 和 $near_rms$ 。

如果这样在步骤 S4 中找到的最小值 $minval$ 大于作为稳态 rms 的 min_rms ,

稳态电平计算单元 3 执行步骤 S5 以便如下面的公式 (6) 所示更新 \min_rms :

$$\min_rms = 0.8 \cdot \min_rms + 0.2 \cdot \minval \quad \dots(6)$$

接着, 在步骤 S6, 按下面的公式 (7) 和 (8) 更新 far_rms :

$$\text{far_rms}(0) = \text{far_rms}(1) \quad \dots(7)$$

5 $\text{far_rms}(1) = \text{near_rms} \quad \dots(8)$

接着, 在步骤 S7, rms 和标准电平 STD_LEVEL 中较小的一个被设置为 max_val , 这里 STD_LEVEL 等于信号电平-30dB, 以便设置一个上限电平, 从而使当前 rms 是一个较高信号电平时不会发生故障。在步骤 S8, 将 maxval 与 min_rms 进行比较以便如下更新 min_rms : 即, 如果 maxval 小于 min_rms , 在
10 步骤 S9 只稍微更新 min_rms , 如公式 (9) 所示, 然而, 如果 maxval 不小于 min_rms , 在步骤 S10 只稍微更新 min_rms , 如公式 (10) 所示:

$$\min_rms = \min_rms + 0.001 \cdot \text{maxval} (\text{maxval} \geq \min_rms) \quad \dots(9)$$

$$\min_rms = \min_rms + 0.002 \cdot (\text{maxval} \geq \min_rms) \dots(10)$$

在步骤 S11, 如果 min_rms 小于静音电平 MIN_LEVEL , 设置 $\text{min_rms} =$
15 MIN_LEVEL , 这里 MIN_LEVEL 是信号电平-66dB。

同时, 如果在步骤 S12 前面与后面信号部分的电平比率 ratio 小于 4, 同时 rms 小于 STD_LEVEL , 帧信号是稳定的。因此, 稳态电平计算单元 3 执行步骤 S13 将稳定指示计数器 st_cnt 增加到 1, 如果是其它情况, 稳态电平计算单元 3 执行步骤 S14 设置 $\text{st_cnt} = 0$, 因为这时的稳定性较低。这样实现了目标稳
20 态 rms 。

除法器 4 将 rms 计算单元 2 的输出 rms 除以稳态电平计算单元 3 的输出 min_rms 来计算 rms_g 。即, 该 rms_g 表示当前 rms 相对稳态 rms 的大概电平。

接着 LPC 分析单元 5 从输入语音信号 $s(n)$ 中找到短期预测 (LPC) 系数 $\alpha(m)$ ($m=1, \dots, M$)。同时, 也可以用语音编码单元 13 的内部 LPC 分析得到的 LPC
25 系数 $\alpha(m)$ 。LPC 对数倒频谱系数计算单元 6 将 LPC 系数 $\alpha(m)$ 转换为 LPC 系数 $C_L(m)$ 。

对数振幅计算单元 7 能够按照下面的公式 (11) 从 LPC 系数 $C_L(m)$ 中找到对数平方振幅特性 $\ln|H_L(e^{j\Omega})|^2$:

$$\ln|H_L(e^{j\Omega})|^2 = 2 \sum_{m=0}^{\infty} C_L(m) \cos(\Omega m) \quad \dots(11)$$

然而，这里上面公式右边的求和计算的上限，在无穷大的位置上，设为 16，并且按照下面的公式（12）和（13）找到一个整数以找到间隔平均 $\log\text{Amp}(i)$ 。同时， $C_L(0)=0$ 因此忽略掉。

$$5 \quad \log\text{Amp}(i) = \frac{1}{\omega} \int_{\Omega_i}^{\Omega_{i+1}} \ln|H_L(e^{j\Omega})|^2 d\Omega \quad \dots(12)$$

$$\log\text{Amp}(i) = \frac{1}{\omega} \left[2 \sum_{m=1}^{16} \frac{1}{m} C_L(m) \sin(\Omega m) \right]_{\Omega_i}^{\Omega_{i+1}} \quad \dots(13)$$

其中对于平均间隔 ($\omega = \Omega_{i+1} - \Omega_i$) 设置 ω 为 500Hz($= \pi/8$)。这里，对应于以 500Hz 为间隔 4 等分范围 0 到 2 kHz，计算 $i=0, \dots, 3$ 时的 $\log\text{Amp}(i)$ 。

现在解释对数振幅差分计算单元 8 和模糊推理单元 9。在本发明中，使用模糊理论检测静音和背景噪音。模糊推理单元 9 如下所述使用除法器 4 用 \min_rms 除 rms 得到的 rms_g 值和来自对数振幅差分计算单元 8 的 $wdif$ 输出判断标记 $decflag$ 。

图 5 给出的是模糊推理单元 9 中使用的模糊规则。在图 5 中，上一行(a)，中间行 (b) 和下一行 (c) 是用于背景噪音的规则，分别主要是噪音参数更新规则 15 和语言规则。还有，在图 5 中，左面的一列，中间列和右面的一列分别表示 rms 的隶属函数，频谱包络的隶属函数和推理结果。

模糊推理单元 9 首先使用图 5 左面一列的隶属函数将除法器 4 用 \min_rms 除 rms 得到的 rms_g 值分类。从上面一行，如图 6 所示定义隶属函数 $\mu_{A_{i1}}(x_1)$ ($i=1,2,3$)。同时， $x_1=rms_g$ 。

另一方面，对数振幅差分计算单元 8 保持过去 n (例如，4) 帧的频谱的对数振幅 $\log\text{Amp}(i)$ 并找到平均值 $ave\text{Amp}(i)$ 。对数振幅差分计算单元 8 接着通过下面的公式 (14) 找到 $ave\text{Amp}(i)$ 与当前 $\log\text{Amp}(i)$ 之差的平方和 $wdif$ ：

$$20 \quad wdif = \sqrt{\frac{1}{4} \sum_{i=0}^3 (\log\text{Amp}(i) - ave\text{Amp}(i))^2} \quad \dots(14)$$

模糊推理单元 9 用图 5 中间一行所示的隶属函数对对数振幅差分计算单元 8 如上所述得到的 $wdif$ 进行分类。通过上面一行，如图 7 所示定义隶属函数 $\mu_{A_{j1}}(x_1)$ ($i=1, 2, 3$)，这里 $x_2=wdif$ 。即，图 5 中间一列所示的隶属函数从上面一行



(a), 中间行(b)到下面一行(c)定义为 $\mu_{A12}(x_2)$, $\mu_{A22}(x_2)$ 和 $\mu_{A32}(x_2)$ 。同时, 如果 rms 小于上述常数 MIN_LEVEL (静音电平), 图 7 未给出, 但是这时 $\mu_{A12}(x_2)=1$ 而 $\mu_{A22}(x_2)=\mu_{A32}(x_2)=0$ 。原因是, 如果信号较弱, 频谱变化比一般情况更加剧烈从而影响判断。

5 由于从 $\mu_{Aij}(x_i)$ 得到推理结果, 模糊推理单元 9 如下得到隶属函数 $\mu_{Bi}(y)$: 首先, 在图 5 的上面, 中间和下面行中的每行的 $\mu_{A1i}(x_1)$ 和 $\mu_{A2i}(x_2)$ 中的较小者被设定为该行的 $\mu_{Bi}(y)$, 如下面的公式 (15) 所示:

$$\mu_{Bi}(y)=\min(\mu_{A1i}(x_1), \mu_{A2i}(x_2)) \quad (i=1,2,3) \quad (15)$$

10 注意到这样的配置, 其中, 如果表示语音的隶属函数 $\mu_{A31}(x_1)$ 和 $\mu_{A32}(x_2)$ 中的一个为 1, 输出 $\mu_{B1}(y)=\mu_{B2}(y)=0$ 和 $\mu_{B3}(y)=1$ 。

注意, 从公式 (15) 得到的每一级的 $\mu_{Bi}(y)$ 等于图 5 的右侧列的函数值。如图 8 所示定义隶属函数 $\mu_{Bi}(y)$ 。即, 右列所示的隶属函数以图 8 所示的上面行(a), 中间行(b)和下面行(c)所示的顺序被定义为 $\mu_{B1}(y)$, $\mu_{B2}(y)$ 和 $\mu_{B3}(y)$ 。

15 根据这些值, 模糊推理单元 9 进行推理, 其通过下面公式 (16) 所示的区域法进行判别:

$$y^* = \frac{\sum_{i=1}^3 S_i \cdot y_i^*}{\sum_{i=1}^3 S_i}, \quad S_i = \int \mu_{Bi}(y) dy \quad (16)$$

这里 y^* 和 y_i^* 表示推理结果和每行的隶属函数的重心。在图 5, 以上面, 中间和下面行的顺序分别是 0.1389, 0.5 和 0.8611。Si 表示一个区域。使用隶属函数 $\mu_{Bi}(y)$, 从下面的公式 (17), (18) 和 (19) 可以得到 S_1 到 S_3 :

20 $S_1 = \mu_{B1}(y) \cdot (1 - \mu_{B1}(y)/3)/2 \quad (17)$

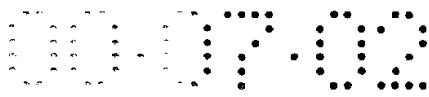
$$S_2 = \mu_{B2}(y) \cdot (2/3 - \mu_{B2}(y)/3) \quad (18)$$

$$S_3 = \mu_{B3}(y) \cdot (1 - \mu_{B3}(y)/3)/2 \quad (19)$$

通过这些值得到的 y^* 的推理结果值, 判定标记 decFlag 的输出值定义如下:

25 $0 \leq y^* \leq 0.34 \quad \rightarrow \text{decFlag} = 0$

$$0.34 < y^* < 0.66 \quad \rightarrow \text{decFlag} = 2$$



$$0.66 \leq y^* \leq 1 \quad \rightarrow \text{decFlag}=1$$

这里 decFlag=0 指出判定结果表示背景噪音，decFlag=2 指出参数需要被更新，以及 decFlag=1 指出语音判别结果。

图 9 示出一个具体的例子。假定 $x_1=1.6$ 和 $x_2=0.35$ ，从这里，如下

5 定义

$\mu_{A_{ij}}(x_j)$, $\mu_{A_{i2}}(x_2)$ 和 $\mu_{B_i}(y)$:

$$\mu_{A_{11}}(x_1)=0.4, \mu_{A_{12}}(x_2)=0, \mu_{B_1}(y)=0$$

$$\mu_{A_{21}}(x_1)=0.4, \mu_{A_{22}}(x_2)=0.5, \mu_{B_2}(y)=0.4$$

$$\mu_{A_{31}}(x_1)=0.6, \mu_{A_{32}}(x_2)=0.5, \mu_{B_3}(y)=0.5$$

10 如果从这里计算一个区域，则 $S_1=0$ ， $S_2=0.2133$ 以及 $S_3=0.2083$ ，因此最终 $y^*=0.6785$ 以及 decFlag =1，这样就指出语音。

前面的是输入信号判别单元 21 a 的操作。下面解释参数控制单元 21b 的各个部分的详细操作。

15 计数器控制器 11 根据来自 V/UV 判定单元 115 的 idVUV 的判定结果和来自模糊推理单元 9 的标记 decflag 设定背景噪音计数器 bgnCnt 以及背景噪音间隔计数器 bgnIntvl。

参数产生单元 12 从计数器控制器 11 的 bgnIntvl 和 idVUV 的判定结果确定 idVUV 参数和更新标记 Flag 以设定从输出端 106 发出的更新标记 Flag。

20 确定传输参数的流程图如图 10 和 11 所示。定义背景噪音计数器 bgnCnt 以及背景噪音间隔计数器 bgnIntvl，二者都具有初始值 0。首先，如果图 10 中的步骤 S21 的输入信号的分析结果指示无声声音(idVUV=0)，并且在步骤 S22 到 S24decFlag=0，则程序进入步骤 S25 将背景噪音间隔计数器 bgnIntvl 增加 1。如果 decFlag=2，则保持 bgnCnt。在步骤 S26，如果 bgnCnt 不小于一个恒定的 BGN_CNT，例如 6，则程序进入步骤 S27 以便将 idVUV 设定到表示背景噪音

25 的值或 1。在步骤 S28，如果 decFlag=0，并且 bgnCnt>BGN_CNT，则在步骤 S29 将 bgnCnt 增加 1。如果在步骤 S31，bgnIntvl 等于一个恒定的 BGN_CNT，例如 16，则程序进入步骤 S32 以便设定 bgnIntvl=0。如果在步骤 S28 dec Flag=2 或者 bgnCnt=BGN=CNT，则程序进入步骤 S30，在这里设定 bgnIntvl=0。

30 在步骤 S21，如果该声音为有声 (idVUV=2,3)，或者如果在步骤 S22，decFlag=1，则程序进入步骤 S23，在这里 bgnCnt =0 并设定 bgnIntvl=0。



参照图 11, 如果在步骤 S33 该声音为无声或者背景噪音 ($idVUV=0,1$), 并且如果在步骤 S35, 如果该声音为无声 ($idVUV=0$), 则在步骤 S36 输出无声参数。

如果在步骤 S35 背景噪音 ($idVUV=1$) 并且如果在步骤 S37, $bgnIntvl=0$, 则在步骤 S38 输出背景噪音参数 ($BGN=$ 背景噪音)。另一方面, 如果在步骤 S37 $bgnIntvl>0$, 则程序进入步骤 S39 以便今发射标题比特。

标题比特的配置如图 16 所示。注意, $idVUV$ 比特被直接设定到上面的两个比特。如果背景噪音周期 ($idVUV=1$) 和帧不是更新帧, 则下一个比特被设定为 0 并且, 如果是其他情况, 则下一个比特被设定为 1。

作为一个示例, 采用 MPEG4 所使用的语音编解码器 HVXC ((谐波向量激发编码), 在图 12 详细示出各种条件下的编码比特。

对于有声, 无声, 背景噪音更新或背景噪音无更新, 使用两个比特编码 $idVUV$ 。作为更新标记, 在背景噪音更新和无更新时每次分配 1 比特。

LSP 参数被分为 LSP 0, LSP2, LSP3, LSP4 和 LSP5。在这些参数中, LSP 0 是第 10 个 LSP 的代码本索引并用作基本的包络参数。对于 20 毫秒帧, 分配 5 比特。LSP2 是第 5 个低频纠错的 LSP 参数的代码本索引并且已将 7 个比特分配在这里。LSP3 是第 5 个高频范围纠错的 LSP 参数的代码本索引并且已将 5 个比特分配在这里。LSP5 是第 10 个全频范围纠错的 LSP 参数的代码本索引并且已将 8 个比特分配在这里。在这些参数中, 当 LSP 0 还不能充分表示包络时, LSP2, LSP3 和 LSP5 是用作补偿前面级的差错并被补充使用的索引。LSP4 是用于选择编码时的编码模式是直接模式还是差模差分模式的 1 比特选择标记。具体地说, 其表示通过量化得到的直接模式的 LSP 和从量化差值得到的 LSP 之间的选择, 无论哪个值都具有与从原始波形分析得到的原始 LSP 参数较小的差值。如果 LSP4 是 0 或 1, 则模式分别为直接模式或差模差分模式。

对于有声声音, 其全部的 LSP 参数为编码比特。对于有声声音和背景噪音更新, 从编码比特中去除 LSP5。在背景噪音的无更新时不发送 LSP 代码比特。特别是, 背景噪音更新时的 LSP 代码比特为量化最新的三个帧的 LSP 参数的平均值时所得到的编码比特。

音调参数 PCH 为仅用于有声声音的 7-比特代码比特。频谱代码本的代码本参数 idS 被分为第零 LPC 剩余频谱代码本索引 $idS0$ 以及第一 LPC 剩余频谱

代码本索引 idS1。对于有声声音，两个索引都为 4 代码比特。对于无声征引在 6 个比特内编码该噪音代码本索引 idSL00,idSL01。

对于有声声音，LPC 剩余频谱增益代码本索引 idG 被设定为 5 比特代码标志 (bot)。对于无声声音，将 4 比特代码比特分配给噪音代码本增益索引 idGL00 和 idGL11 中的每一个。对于背景噪音更新，仅将 4 比特代码比特分配给 idGL00。背景噪音更新内的这 4 比特 idGL00 为量化最新 4 帧（8 个子帧）的 CELP 增益的平均值所得到的代码比特。

对于有声声音，将 7, 10, 9 和 6 比特分别分配给表示为 idS0_4k 的零扩展 LPC 剩余频谱代码本索引，表示为 idS1_4k 的第一扩展 LPC 剩余频谱代码本索引，表示为 idS2_4k 的第二扩展 LPC 剩余频谱代码本索引以及表示为 idS3_4k 的第三扩展 LPC 剩余频谱代码本索引。

分别分配 80 比特用于有声声音，40 比特用于无有声声音，25 比特用于背景噪音更新以及 3 比特用于背景噪音无更新。

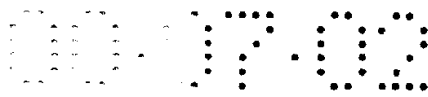
参照图 2，详细解释图 12 所示的用于产生代码比特的语音编码器。

提供给输入端 101 的语音信号由高通滤波器 (HPF) 109 滤波以取出不必要频率范围的信号。如上所述，滤波输出送到输入信号判别单元 21a, LPC (线性预测编码)分析量化单元 113 的 LPC 分析电路 132 以及 LPC 反-滤波电路 111。

LPC 分析量化单元 113 的 LPC 分析电路 132 应用具有大约 256 个样值作为一个模块的输入信号波形的长度的 Hamming 窗口，以便通过自相关方法找到线性预测系数，该系统也称为 α -参数。作为数据输出单元的成帧间隔大约为 160 个样值。采用例如 8kHz 的采样频率 f_s , 帧间隔为 160 个样值或 20 毫秒。

来自 LPC 分析电路 132 的 α -参数被送到用于转换到线性频谱对 (LSP) 参数的 α -LSP 转换电路 133。在这种情况下，所找到的作为直接滤波器系数的 α -参数通过例如 Newton-Rhapson 方法被转换为，例如为 5 对的 10 个 LSP 参数。因为 LSP 参数的插补特性优于 α -参数的描补特性，所以使用到 LSP 参数的转换。

来自 α -LSP 转换电路 133 的 LSP 参数被 LSP 量化器 134 进行矩阵或向量量化。在向量量化之前首先消除帧到帧的差值。作为一种替换，通过矩阵量化可以一起取出并量化数个帧。这里，20 毫秒是一帧并且每 20 毫秒计算的 LSP 参数被一起取出并进行矩阵或向量量化。



作为 LSP 量化索引的 LSP 量化器 134 的量化输出在终端 102 输出，同时量化 LSP 向量被送到 LSP 插补电路 136。

LSP 插补电路 136 插补每 20 毫秒或每 40 毫秒量化的 LSP 向量，以便将比率提高因数 8，从而 LSP 向量每 2.5 毫秒被更新。其原因是，如果剩余波形被谐波编码/解码方法分析-合成，则合成波形的包络极为光滑，因此，如果 LPC 系数变化剧烈，将导致产生极大的声音。即，如果 LPC 系数每 2.5 毫秒逐渐变化，则将阻止产生极大的声音。

对于使用插补基于 2.5 毫秒的 LSP 向量执行输入语音的反-滤波，通过 LSP- α 转换电路 137 将 LSP 参数转换为 α -参数，该 α -参数为具有大约 10 级的直接类型滤波器的系数。LSP- α 转换电路 137 的输出被送到 LPC 反-滤波电路 111，在这里使用每 2.5 毫秒更新的 α -参数完成反-滤波以便实现光滑输出。LPC 反-滤波电路 111 的输出被送到正弦分析编码单元 114，具体为谐波编码电路的诸如离散傅立叶变换电路之类的正交转换电路 145。

来自 LPC 分析量化单元 113 的 LPC 分析电路 132 的 α -参数被送到心理加权滤波器计算电路 139，在这里找到用于心理加权的数据。该加权数据被送到第二编码单元 120 的心理加权向量量化单元 116，心理加权滤波器 125 以及心理加权合成滤波器 122。

正弦分析编码单元 114 可以是诸如谐波编码电路之类的电路，LPC 反-滤波电路 111 的输出通过谐波编码方法进行分析。即，正弦分析编码单元检测音调，计算每个谐波的幅值 A_m 并执行 V/UV 判别。正弦分析编码单元还在幅值上将具有变化音调的幅值 A_m 或谐波的包络转换为恒定数。

在图 2 示出的正弦分析编码单元 114 的一个具体的示例，预先抑制例行谐波编码。特别是，在多-频带激励 (MBE) 编码中，假定在同时发生的每个频率范围或频段，即，在相同的帧块出现一个语言部分和一个无语言部分的情况下进行建模。在其他的谐波编码形式中，一种替换判定是就该帧块内的语音是语言还是无语言进行判定。在下面的解释中，基于帧的 V/UV 含义是在采用 MBE 编码情况下当整个频带为 UV 时的给定帧的 V/UV。就分析方法的合成而言，由本专利受让人提出的日本特开专利 H-5-265487 公开了一种本专利受让人建议的具体示例。

将来自输入端 101 的输入语音信号反馈给图 2 的正弦分析编码单元 114 的



开环音调查找单元 141, 同时将来自高通滤波器 (HPF) 109 的信号反馈给过零计数器 142。将来自 LPC 反-滤波电路 111 的 LPC 剩余或线性预测剩余反馈给正弦分析编码单元 114 的正交转换电路 145。开环音调查找单元 141 取出输入信号的 LPC 剩余以便通过取出取出输入信号的 LPC 剩余执行相对粗糙的音调查找。所提取的粗音调数据被送到高准确性音调查找单元 146, 在这里执行后面所述的闭环查找高准确性音调 (精细音调查找)。从开环音调查找单元 141, 与粗音调数据一起取出标准化 LPC 剩余的最大自相关值时得到的最大标准化自相关值 $r(p)$, 并将其送到 V/UV 判定单元 115。

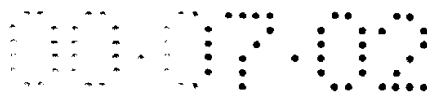
正交转换电路 145 执行诸如离散余弦变换 (DFT) 之类的处理以便将时间轴上的 LPC 剩余转换为频率轴上的频谱幅值数据。正交转换电路 145 的输出被送到高准确性音调查找单元 146 以及用于评价频谱幅值或包络的频谱评价单元 148。

将开环音调查找单元 141 提取的相对较粗音调的粗音调数据和开环音调查找单元 141 提取的频率间隔上的数据反馈给高准确性音调查找单元 146。在该高准确性音调查找单元 146, 以粗音调数据为中心音调数据被偏移正负数个样值以接近具有最佳小数点 (浮点) 的精细音调数据值。作为精细查找技术, 使用由合成方法进行的所谓分析并选择音调以便合成功率频谱将最接近原始语音的功率频谱。通过闭环来自高准确性音调查找单元 146 的音调数据通过开关 118 被送到输出端 104。

在频谱评价单元 148, 根据作为 LPC 剩余的正交转换输出的音调和频谱幅值评价其设定的每个谐波的幅值和频谱包络。评价结果送到高准确性音调查找单元 146, V/UV 判定单元 115 以及心理加权向量量化单元 116。

在 V/UV 判定单元 115, 根据正交转换电路 145 的输出, 高准确性音调查找单元 146 的最佳音调, 频谱评价单元 148 的幅值数据, 开环音调查找单元 141 的最大标准化自相关值 $r(p)$ 以及过零计数器 142 的过零值给出所述帧的 V/UV 判定。MBE 编码情况下的基于频带的 V/UV 判定结果的边界位置也用作所述帧的 V/UV 判定条件。经输出端 105 取出 V/UV 判定单元 115 的判定输出。

频谱评价单元 148 的输出或向量量化单元 116 的输入提供给多个转换单元 119, 该转换单元是一种采样率转换单元。在考虑到音调改变在频率间隔分开的频带数目从而改变数据量的情况下, 该数据转换单元的数目用于将包络的幅



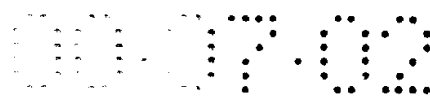
值数据 $|A_m|$ 设定为恒定数。即，如果有效频带高达 3400kHz，则根据音调将该有效频带分为 8 到 63 个频带，以便从频带到频带所得到的幅值数据 $|A_m|$ 的数字 $m_{MX}+1$ 也在 8 到 63 的范围内变化。因此，数据转换单元 119 的数目将变化的数字 $m_{MX}+1$ 幅值数据转换为恒定数 M，例如，44。

5 按照预设定的数据数，例如 44 数据收集上述诸如 44 之类的恒定数，来自频谱评价单元 148 的输出提供的数据转换单元数的振幅数据或包络数据或向量量化单元 116 的输入作为进行加权向量量化的向量。由心理加权滤波器计算电路 139 的输出进行加权。来自向量量化单元 116 的上述包络的索引 idS 在输出端 103 通过开关 117 输出。同时，在加权向量量化之前取出采用适当漏出系数
10 的内帧差值以用于由预设定数据数组成的向量。

下面将描述具有所谓的 CELP（编码激励线性预测）编码配置的编码方法。该编码方法用于编码输入语音信号的无声部分。在用于输入语音信号的无声语音部分的该 CELP 编码配置中，对应于作为噪音代码本各输出的无声语音的 LPC 剩余的噪音输出，或所谓的随机代码本 121 通过增益电路 126 送到心理加权合成滤波器 122。加权合成滤波器 122 通过 LPC 合成方法 LPC 合成输入噪音以便将产生的加权无声语音信号送到减法器 123。将输入端 101 提供的语音信号经高通滤波器（HPF）109 反馈给减法器，该信号已由心理加权滤波器 125 心理加权。这样，减法器从合成滤波器 122 的信号得到差值或误差。注意到在心理加权滤波器 125 的输出开始减去心理加权合成滤波器的零输入响应。该误差
15 被送到距离计算电路 124 以便进行查找最小化噪音代码本 121 的误差的表示值向量的距离计算。也就是通过采用闭环查找，采用合成方法的依次合成所得到的时间间隔波形，该波形是被量化的向量。

就用于体现 CELP 编码配置的编码单元的 UV(无声)部分的数据而言，取出来自噪音代码本 121 的代码本的整形索引 idSI 以及来自增益电路 126 的代码
25 本的增益索引 idGI。作为来自噪音代码本 121 的 UV 数据的整形索引 idSI 经开关 127a 送到输出端 107a，同时作为增益电路 126 的 UV 数据的增益索引 idGI 经开关 127g 送到输出端 107g。

根据 V/UV 判定单元 115 的 V/UV 判定结果接通/断开控制这些开关 127s，127g 以及上述开关 117，118。当与现在发送有关的帧语音信号的 V/UV 判定
30 结果指示有声声音（V）时接通开关 117，118，同时当与现在发送有关的帧语

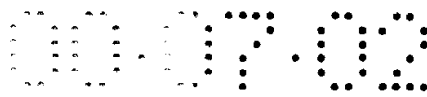


音信号指出无声声音 (UV) 时接通开关 127s, 127g。

由传输路径编码器 22 编码由上述语音编码器以可变比率编码的各个参数, 这些参数为 LSP 参数 LSP, 有声/无声判别参数 idVUV, 音调参数 PCH, 频谱包络的代码本参数 idS 和增益索引 idG, 噪音代码本参数 idS1 以及增益索引 idG1, 以便语音质量不受传输路径质量的影响。所产生的信号由调制器 23 调制并被处理以用于发射器 24 发射, 从而经天线共享器 25 由天线发射。如上所述, 上述参数也送到参数控制单元 21b 的参数产生单元 12。参数产生单元 12 使用来自 V/UV 判定单元 115 的 V/UV 判定结果, 来自计数器控制器 11 的上面参数以及 bgnIntvl 产生 idVUV 以及 0 更新标记。参数控制单元 21b 还管理控制, 因此, 如果从 V/UV 判定单元 115 产生指示背景噪音的 idVUV=1, 则作为 LSP 量化方法的差模差分模式 (LSP4=1) 禁止 LSP 量化器 134 使用, 从而使得仅通过直接模式 (LSP4=0) 执行量化。

解释图 1 所示的便携电话设备的接收侧的语音解码设备 31。将由天线 26 俘获, 由接收器 27 经天线共享器 25 接收, 由解调器 29 解调并由用于传输路径误差的传输路径解码器 30 校正的接收比特反馈给语音解码设备 31。

语音解码设备 31 的结构在图 13 详细示出。具体地说, 语音解码设备包括用于从输入端 200 输入接收比特取出标题比特以便如图 16 所示分开 idVUV 和更新标记并用于输出代码比特的标题比特解释单元 201, 和用于通过 idVUV 和更新标记如后面所解释的控制开关 143 和 248 的切换的切换控制器 241。语音解码设备还包括用于通过后面解释的序列确定 LPC 参数或 LSP 参数的 LPC 参数再现控制器 240, 以及用于从代码比特内的 LSP 索引再现 LPC 参数的 LPC 参数再现单元 213。语音解码设备还包括用于将代码比特分解为各个参数索引的代码比特解释单元 209 以及由切换控制器 241 控制, 在接收到背景噪音更新帧闭合, 反之断开的开关 248。语音解码设备还包括由切换控制器 241 控制, 使其在接收到背景噪音更新帧与 RAM244 连接, 反之断开的开关 243, 以及用于产生 UV 整形索引作为随机数的随机数发生器 208。语音解码设备还包括用于向量量化来自包络索引的包络的向量去量化器 212 和用于合成来自 idVUV, 音调和包络的有声声音的语言语音合成单元 211。语音解码设备还包括 LPC 合成滤波器 214 以及用于在接收到背景噪音更新标记时保留代码比特并在接收到背景噪音无更新标记时提供代码比特的 RAM 244。



首先，标题比特解释单元 201 从输入端 200 提供的接收比特取出标题比特以便从更新标记 Flag 分离 idVUV 以识别所述帧内的帧数。如果存在下一个后面的比特，则标题比特解释单元 201 将其输出作为一个代码比特。如果标题比特配置的上两比特为 00，则该比特被认为是背景噪音(BGN)。因此，如果下一个比特为 0，则该帧为无更新帧，因此该处理接近结束。如果下一个比特为 1，则读出下面的 22 比特以读出背景噪音的更新帧。如果上面两位为 10/11，则该帧被认为是语言，因此读出下面的 78 比特。

切换控制器 241 检查 idVUV 和更新标记。如果 idVUV=1，并且更新标记 Flag=1，则发生更新，因此开关 248 被闭合以便将代码比特发送到 RAM244。同时，开关 243 闭合到标题比特解释单元 201 一侧以便将代码比特送到代码比特解释单元 209。如果相反更新标记 Flag=0，则未发生更新，因此开关 248 被断开。开关 243 闭合到 RAM244 一侧以便在更新开始时提供代码比特。如果 idVUV≠1，则开关 248 断开，同时开关 243 打开到上面一侧。

代码比特解释单元 209 将从标题比特解释单元 201 经开关 243 提供到这里的代码比特分解为相应的参数索引，也就是 LSP 索引，音调，包络索引，UV 增益索引或 UV 整形索引。

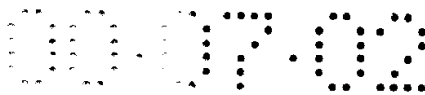
随机数发生器 208 产生 UV 整形索引作为随机数。如果开关 249 接收到具有 idVUV=1 的背景噪音帧，则开关 249 由切换控制器 241 闭合以便将 UV 整形索引发送到无有声声音合成单元 220。如果 idVUV≠1，则 UV 整形索引经开关 249 从代码比特解释单元 209 发送到无有声声音合成单元 220。

LPC 参数再现控制器 240 内部具有一个切换控制器和一个索引判定单元并以后面解释的方式，根据判定结果通过切换控制器检测 idVUV 以控制 LPC 参数再现单元 213 的操作。

LPC 参数再现单元 213，无有声声音合成单元 200，向量去量化器 212，有声声音合成单元 211 以及 LPC 合成滤波器 214 构成语音解码设备 31 的基本部分。图 14 示出这些基本部分和外围部分的结构。

将也被称为代码本索引的 LSP 向量量化输出反馈给输入端 202。

将 LSP 索引发送给 LPC 参数再现单元 213。如上所述，LPC 参数再现单元 213 通过代码比特内的 LSP 索引再现 LPC 参数。LPC 参数再现单元 213 由未示出的 LPC 参数再现控制器 240 内的切换控制器控制。



首先，解释 LPC 参数再现单元 213。LPC 参数再现单元 213 包括一个 LSP 去量化器 231，转换开关 251，LSP 插补电路 232(用于 V)以及 233(用于 UV)，LSP- α 转换电路 234 (用于 V) 以及 235 (用于 UV)，开关 252，RAM 253，帧插补电路 245，LSP 插补电路 246 (用于 BGN) 以及 LSP- α 转换电路 247 (用于 BGN)。

LSP 去量化器 231 去量化来自 LSP 索引的 LSP 参数。解释 LSP 去量化器 231 内的 LSP 参数的生成。这里，引入背景噪音计数器 $bgnIntvl$ (初始值=0)。在有声声音 ($idVUV=2,3$) 或无有声声音 ($idVUV=0$) 的情况下，通过所用的解码处理产生 LSP 参数。

在背景噪音 ($idVUV=1$) 的情况下，如果该帧为更新帧，则设定 $bgnIntvl=0$ ，并且，如果相反，则将 $bgnIntvl$ 增加 1。如果当 $bgnIntvl$ 增加 1 时，其等于后面解释的 BGN_INTVL_RX ，则 $bgnIntvl$ 不增加 1。

接着，如下面公式 (20) 所述产生 LSP 参数。

$$qLSP(t) = \frac{(2 \cdot (BGN_INTVL_RX - bgnIntvl') - 1) \cdot qLSP_{(prev)}(i) + (2 \cdot bgnIntvl' + 1) \cdot qLSP_{(curr)}(i)}{2 \cdot BGN_INTVL_RX} \quad (20)$$

注意到在更新帧为 $qLSP_{(prev)}(1, \dots, 10)$ 之前直接接收 LSP 参数，则在更新帧内接收的 LSP 参数为 $qLSP_{(prev)}(1, \dots, 10)$ 并且通过插补产生的 LSP 参数为 $qLSP_{(prev)}(1, \dots, 10)$ 。

在上面的公式， BGN_INTVL_RX 为恒定值，并且使用 $bgnIntvl$ 和随机数 $rnd(=3, \dots, 3)$ 通过下面公式 (21) 产生 $bgnIntvl'$ ：

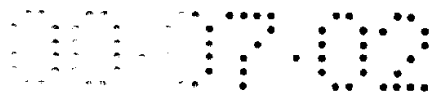
$$bgnIntvl' = bgnIntvl + rnd \quad \dots \quad (21)$$

注意到，如果当 $bgnIntvl' < 0$ ， $bgnIntvl' = bgnIntvl$ 以及 $bgnIntvl' \geq BGN_INTVL_RX$ ，则设定 $bgnIntvl' = bgnIntvl$ 。

LPC 参数再现控制器 240 内的未示出的切换控制器根据 V/UV 参数以及更新标记 Flag 控制 LPC 参数再现单元 213 内的开关 252，262。

对于 $idVUV=0, 2, 3$ 以及对于 $idVUV=1$ ，开关 251 被分别设定到上端和下端。如果更新标记 $Flag=1$ ，即，在背景噪音更新帧的情况下，开关 252 被闭合到将 LSP 参数发送到 RAM253 以便在 $qLSP_{(curr)}$ 更新 $qLSP_{(prev)}$ 之后更新 $qLSP_{(curr)}$ 。RAM253 保留 $qLSP_{(prev)}$ 和 $qLSP_{(curr)}$ 。

帧插补电路 245 使用内部计数器 $bgnIntvl$ 从 $qLSP_{(prev)}$ 和 $qLSP_{(curr)}$ 产生 $qLSP$ 。LSP 插补电路 246 插补 LSPs。LSP- α 转换电路 247 将用于 BGN 的 LSP



转换为 α 。

参照图 15 的流程图将详细解释由 LPC 参数再现控制器 240 进行的 LPC 参数再现单元 213 的控制。

首先，在步骤 S41 LPC 参数再现控制器 240 的切换控制器检测 V/UV 判定参数 $idVUV$ 。如果该参数为 0，则切换控制器进入步骤 S42 以便通过 LSP 插补电路 233 插补 LSPs。接着切换控制器进入步骤 S43，在这里通过 LSP-0 转换电路 235 将 LSPs 转换为 α 。

如果在步骤 S41, $idVUV=1$ ，并且在步骤 S44，更新标记 $Flag=1$ ，则该帧为更新帧，因此在步骤 S45 在帧插补电路 245 设定 $bgnIntvl=0$ 。

如果在步骤 S44, 更新标记 $Flag=0$ ，并且 $bgnIntvl < BGN_INTVL_RX-1$ ，则切换控制器进入步骤 S47 以便将 $bgnIntvl$ 增加 1。

在步骤 S48，产生 $bgnIntvl'$ 作为帧插补电路 245 使用的随机数 rnd 。然而，如果 $bgnIntvl' < 0$ 或者如果 $bgnIntvl' \geq BGN_INTVL_RX$ ，则在步骤 S50 设定 $bgnIntvl' = bgnIntvl$ 。

接着，在步骤 S51, 通过帧插补电路 245 帧-插补 LSPs。在步骤 S52，通过插补电路 246 插补 LSPs 并且在步骤 S53，通过 LSP- α 转换电路 247 将 LSPs 转换为 α 。

如果在步骤 S41, $idVUV=2, 3$ ，则切换控制器进入步骤 S54，在这里通过 LSP 插补电路 232 插补 LSPs。在步骤 S55，通过 LSP- α 转换电路 234 将 LSPs 转换为 α 。

LPC 合成滤波器 214 将用于有声部分的 LPC 合成滤波器 236 与无声部分的 LPC 合成滤波器 237 分开。即，在有声和无声部分独立执行 LPC 系数插补以防止不利影响，该不利影响可能通过插补从有声到无声部分或从无声到有声部分转换时的完全不同特性的 LSPs 引入。

将对应于加权向量量化频谱包络 A_m 的代码索引数据反馈给输入端 203。将音调参数 PCH 的数据和上述 V/UV 判定数据 $id VUV$ 分别反馈给输入端 204，205。

将对应于来自输入端 203 的加权向量量化频谱包络 A_m 的索引数据发送到用于向量去量化的向量去量化器 212。这样，以对应于数据数转换的方式逆-转换数据并检验送到有声声音合成单元 211 的正弦合成电路 215 的频



谱包络数据。

如果在编码内的频谱的向量量化之前得到帧到帧的差值，则在数据数转换所跟随的向量去量化之后执行帧到帧的差值的解码，以产生频谱包络数据。

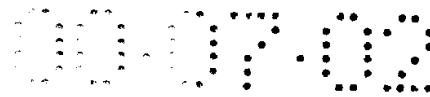
5 将来自输入端 204 的音调以及来自输入端 205 的 V/UV 判定数据 idVUV 反馈给正弦合成电路 215。从正弦合成电路 215，输出对应于图 2 的 LPC 逆-滤波器 111 的输出的 LPC 剩余并将其送到加法器 218。以本专利受让人的姓名申请的日本专利申请 H-4-91422 或日本专利申请 H-6-198451 公开了该正弦合成的特定技术。

10 将来自向量去量化器 212 的包络数据，来自输入端 204, 205 的音调和 V/UV 判定数据以及 V/UV 判定数据 idVUV 路由到适用于增加有声 (V) 部分的噪音的噪音合成电路 216。噪音合成电路 216 的输出经一个加权的加权加法电路 217 被送到加法器 218。这样作的原因是，由于通过正弦合成检验到有声声音的 LPC 滤波器的输入的激励给出诸如男性声音之类的低音调声音的填充感觉并且声音质量在有声 (V) 和无声 (UV) 声音之间突然改变从而产生不自然的感觉，噪
15 音，该噪音考虑从诸如音调，频谱包络幅值，帧的最大幅值之类的编码语言数据得到的参数或者剩余信号电平被增加到 LPC 剩余信号的有声部分。

加法器 218 的和输出被发送到用于 LPC 合成滤波器 214 的有声语音的合成滤波器 236 以进行 LPC 合成处理以便产生一个时间间隔波形信号，该波形信号接着由用于有声语音 238v 的后滤波器滤波并路由到加法器 239。

20 如图 14 所示，作为 UV 数据的整形索引和增益索引被分别路由到输入端 207s 和 207g。接着将增益索引提供给无有声声音合成单元 220。来自终端 207s 的整形索引被送到转换开关 249 的固定端，将随机数发生器 208 的输出反馈给该开关的另一个固定端。如果接收到背景噪音帧，在图 13 所示的切换控制器 241 的控制下，开关 249 闭合到随机数发生器 208 的一侧。将来自随机数发生器 208
25 的整形索引反馈给无有声声音合成单元 220。如果 $idVUV \neq 1$ ，则通过开关 249 从代码比特插补单元 209 提供整形索引。

即，在有声声音 ($idVUV=2, 3$) 或无声声音 ($idVUV=0$) 的情况下通过例行解码处理产生激励信号。在背景噪音 ($idVUV=1$) 的情况下，产生 CELP $idSL00, idSL01$ 的整形索引作为随机数 $md (=0, \dots, N_SHAPE=LO-1$,
30 这里 $N_SHAPE=LO-1$ 为 CELP 整形代码向量数)。将 CELP 增益索引



idGL00,idGL01 提供给更新帧内的两个子帧。

上面已经解释了具有体现本发明的编码方法和设备以及解码方法和设备的便携电话设备。然而，本发明不限于便携电话设备的编码设备和解码设备，而是可以应用于，例如，传输系统。

5 图 17 所示为实施本发明传输系统的实施例的示意性结构。这时，系统是指多个设备的逻辑集成，而没有考虑各个设备是否是处于相同的壳体内。

在该传输系统中，客户终端 63 拥有解码设备，同时服务器 61 拥有编码设备。客户终端 63 和服务器 61 通过网络 62，例如因特网，ISDN（综合服务数字网），LAN（局域网）或 PSTN（公共交换电话网）彼此相连。

10 如果从客户终端 63 通过网络 62 向服务器 61 发出例如音符这样的音频信号请求，对应于请求音符的音频信号的编码参数受到保护以便响应心理声学的比特敏感度，防止在网络 62 上产生传输路径差错，并传输到客户终端 63，它接着解码编码参数，这些编码参数是受到保护的，以便防止对应于解码方法的来自服务器 61 的传输路径差错，并且从象扬声器这样的输出设备中输出解码
15 信号作为语音。

图 18 所示为图 17 的服务器 61 的示意性硬件结构。

这里 ROM（只读存储器）71 存储了例如 IPL（初始程序装入）程序。CPU（中央处理器）72 根据存储在 ROM 71 中的 IPL 程序执行 OS（操作系统）程序。在 OS 控制下，执行存储在外部存储设备 76 中的预先设置的应用程序以便
20 保护音频信号的编码处理并且编码得到的编码以便为客户终端 63 进行编码数据传输处理。RAM（随机存储器）73 存储 CPU 72 的操作所要求的程序或数据。输入设备 74 由例如键盘，鼠标，麦克风或外部接口构成，当需要输入数据或命令时发挥作用。输入设备 74 还用于作为从设置到客户终端 63 的数字音频信号之外接收输入的接口。输出设备 75 由例如显示器，扬声器或打印机构成，
25 并且显示和输出需要的信息。外部存储器 76 包括例如其中存有上述 OS 或预先设置的应用程序的硬盘。通信设备 77 进行在网络 62 上的通信所需要的控制。

存储在外存储器 76 中的预先设置的应用程序是用来使 CPU 72 执行语音编码器 3，传输路径编码器 4 或调制器 7 所发挥的作用的程序。

图 19 所示为图 17 中的客户终端 63 的示意性硬件结构。

30 客户终端 63 由连接到通信设备 87 的 ROM 81 构成并且基本结构与由连接

到通信设备 77 的 ROM 71 构成的服务器 61 的结构类似。

5 需要注意的是，外部存储器 86 中存储了作为应用程序的程序，用来执行本发明的解码方法以解码来自服务器 61 的编码数据，或者存储现在要解释的进行其它处理的程序。通过执行这些应用程序，CPU 82 解码或再现防止出现传输路径差错的编码数据。

具体地，外部存储器 86 中存储一个应用程序，它使 CPU 82 执行解调器 13，传输路径解码器 14 和语音解码器 17 的功能。

这样，客户终端 63 能够通过软件实现存储在外存储器 86 中的解码方法，而不需要图 1 所示硬件结构。

10 客户终端还能够存储从服务器 61 传输到外部存储器 86 的编码数据并在期望的时间读出编码数据并执行编码方法在期望的时间输出语音。编码数据还能够存储在另外的象磁光盘或其它记录介质这样的外部存储器中。

此外，作为服务器 61 的外部存储器 76，可以使用象磁光盘或磁记录介质这样的可记录介质以便在这些记录介质上记录编码数据。

说明书附图

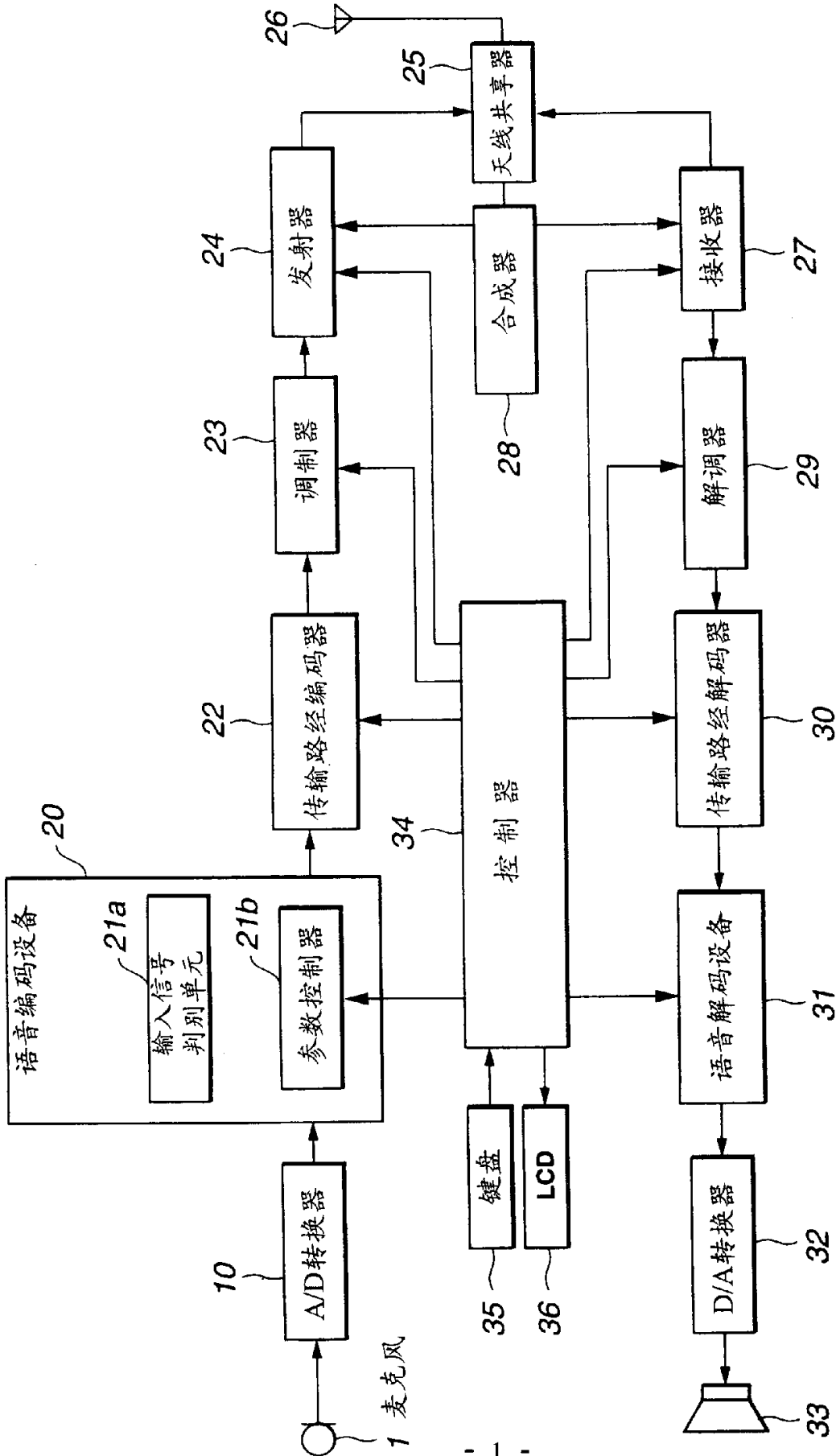


图 1

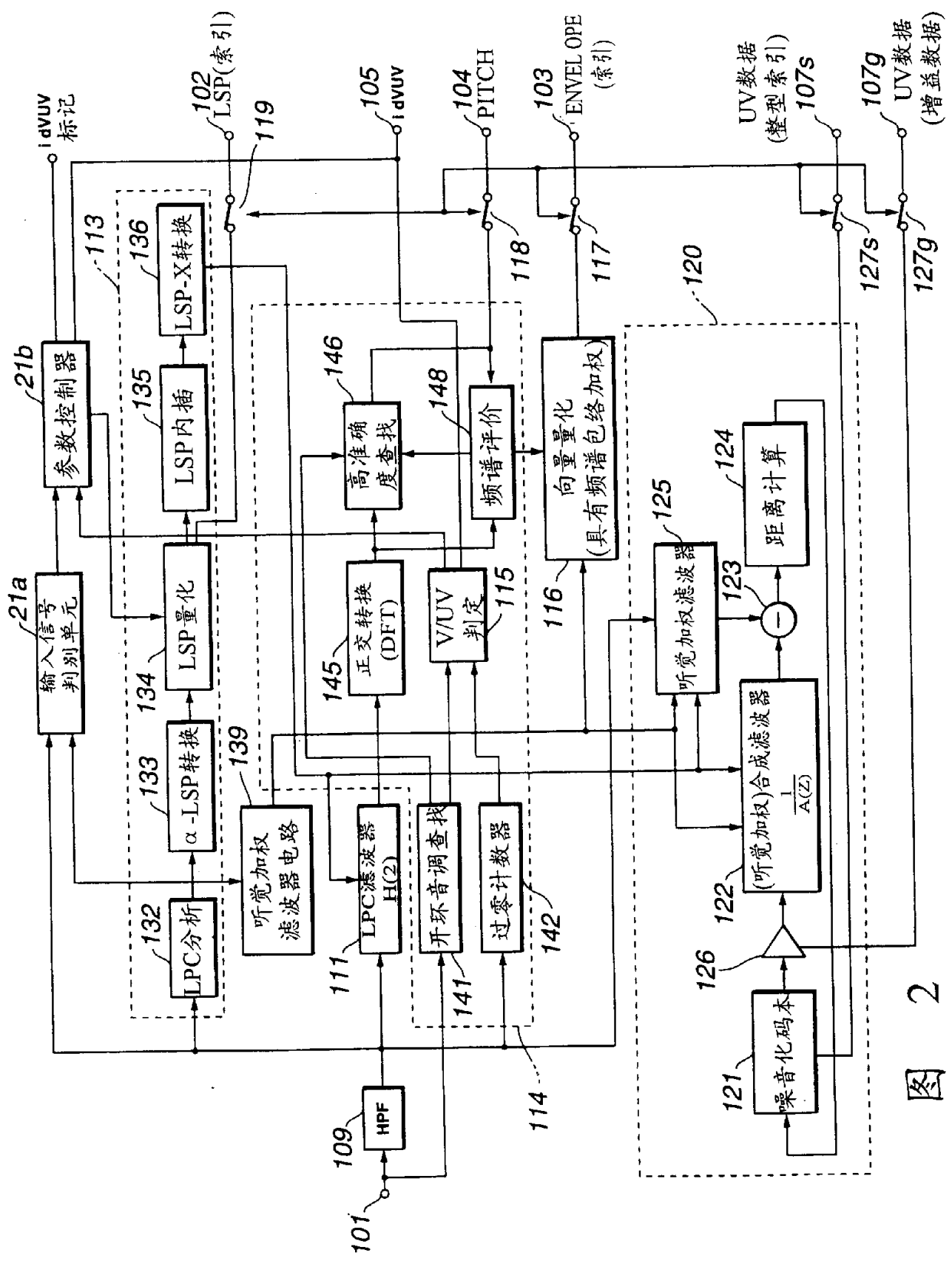


图 2

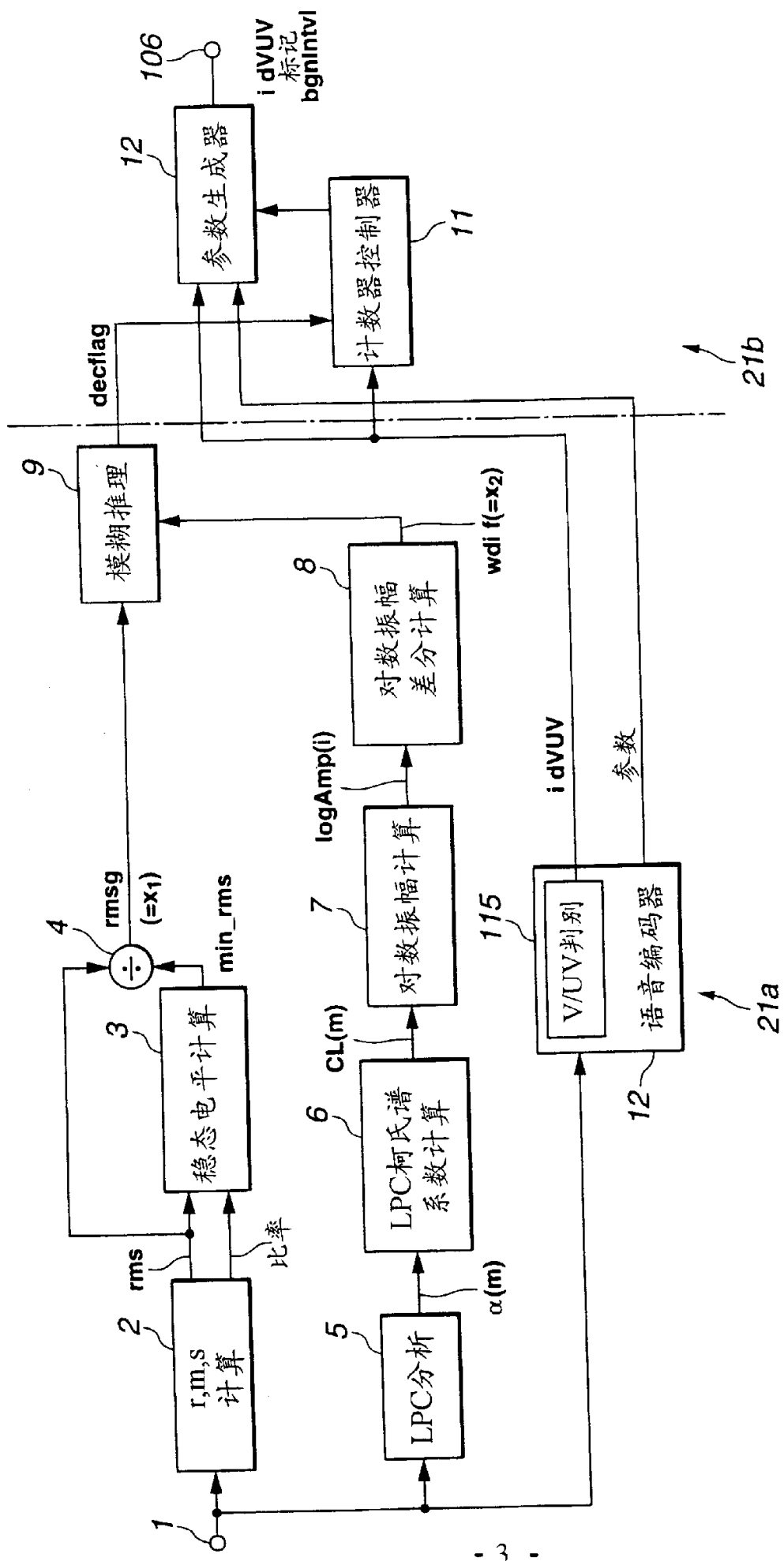


图 3

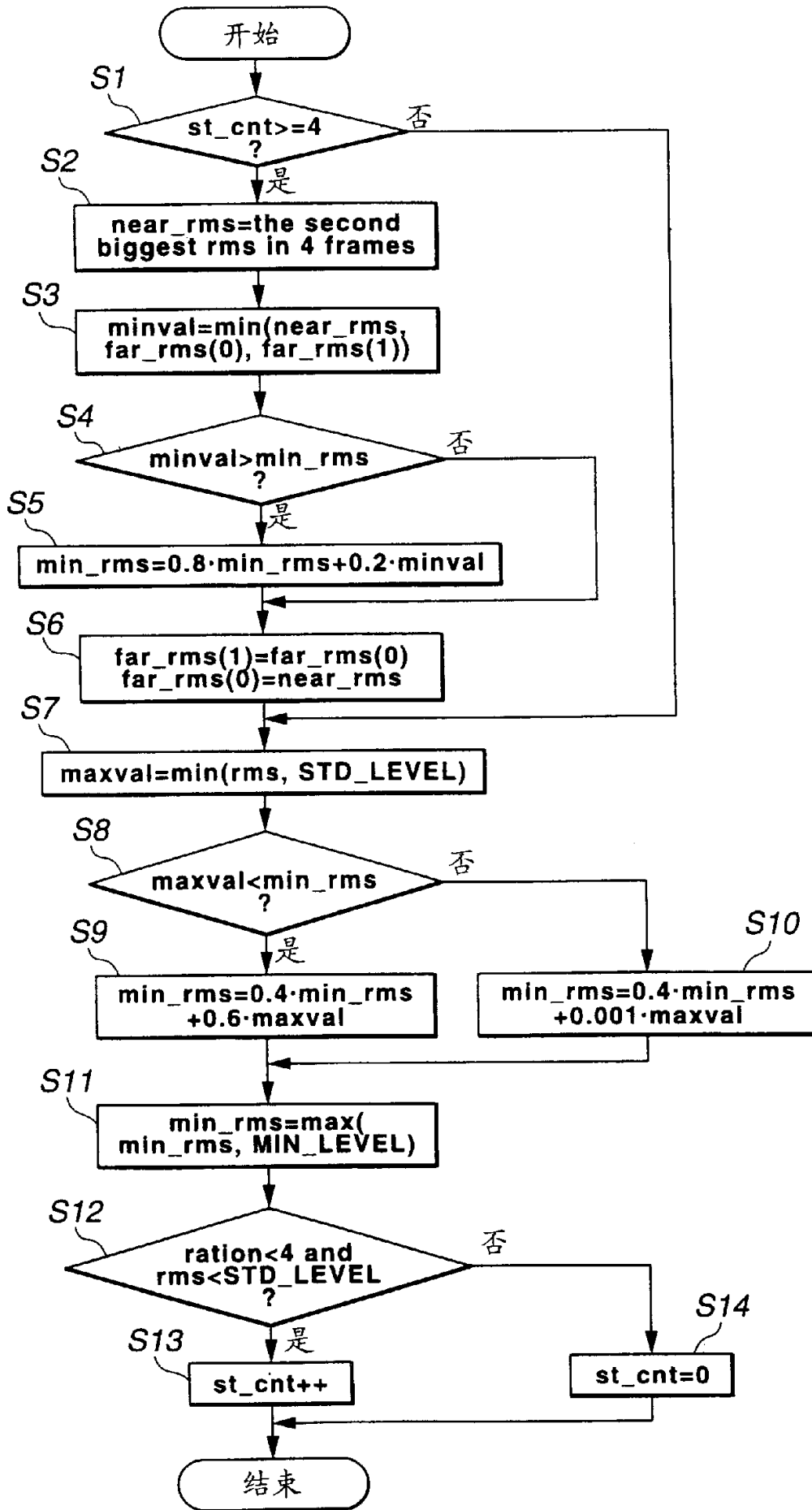


图 4

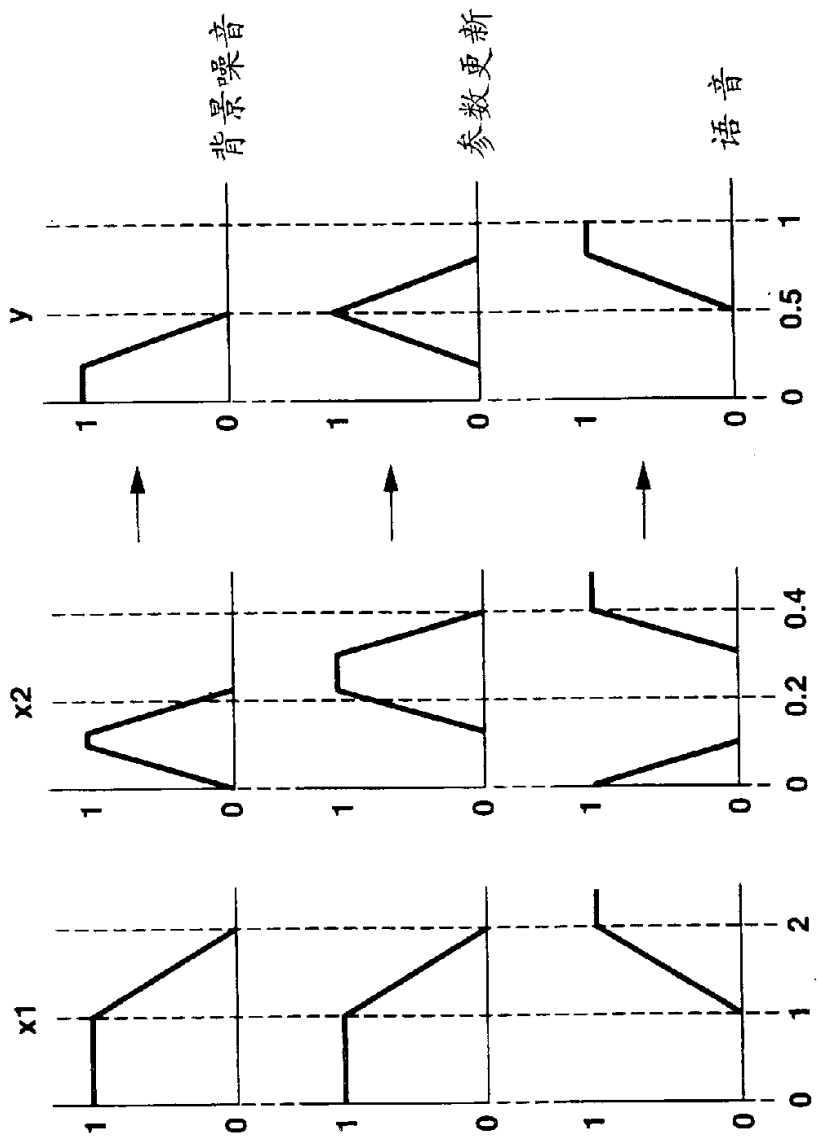


图 5A

图 5B

图 5C

| | $\mu_{A11}(x_1)$ | $\mu_{A21}(x_1)$ | $\mu_{A31}(x_1)$ |
|---------------------|------------------|------------------|------------------|
| $x_1 \leq 1$ | 1 | 1 | 0 |
| $1 \leq x_1 \leq 2$ | $2 - x_1$ | $2 - x_1$ | $x_1 - 1$ |
| $x_1 \geq 2$ | 0 | 0 | 1 |

图 6

| | $\mu_{A12}(x_2)$ | $\mu_{A22}(x_2)$ | $\mu_{A32}(x_2)$ |
|-----------------------------|-----------------------|-----------------------|---------------------|
| $x_2 \leq 0.1$ | $x_2 / 0.1$ | 0 | $1 - x_2 / 0.1$ |
| $0.1 \leq x_2 \leq 0.125$ | 1 | 0 | 0 |
| $0.125 \leq x_2 \leq 0.225$ | $(0.225 - x_2) / 0.1$ | $(x_2 - 0.125) / 0.1$ | 0 |
| $0.225 \leq x_2 \leq 0.3$ | 0 | 1 | 0 |
| $0.3 \leq x_2 \leq 0.4$ | 0 | $(0.4 - x_2) / 0.1$ | $(x_2 - 0.3) / 0.1$ |
| $0.4 \geq x_2$ | 0 | 0 | 1 |

图 7

| | $\mu_{B1}(y)$ | $\mu_{B2}(y)$ | $\mu_{B3}(y)$ |
|-----------------------|---------------------|---------------------|---------------------|
| $0 \leq y \leq 1/6$ | 1 | 0 | 0 |
| $1/6 \leq y \leq 1/2$ | $3 \cdot (1/2 - y)$ | $3 \cdot (y - 1/6)$ | 0 |
| $1/2 \leq y \leq 5/6$ | 0 | $3 \cdot (5/6 - y)$ | $3 \cdot (y - 1/2)$ |
| $5/6 \leq y \leq 1$ | 0 | 0 | 1 |

图 8

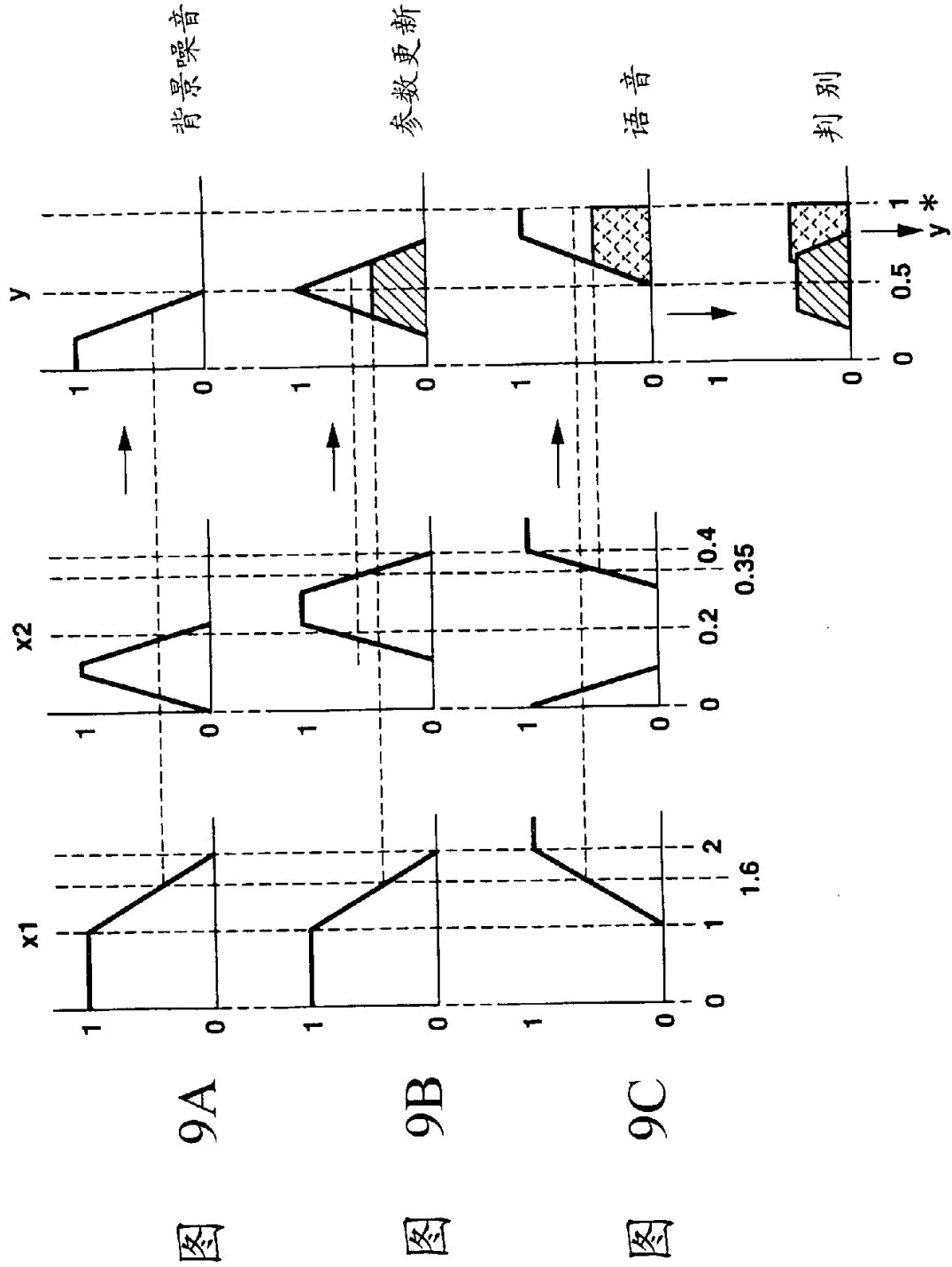


图 9A

图 9B

图 9C

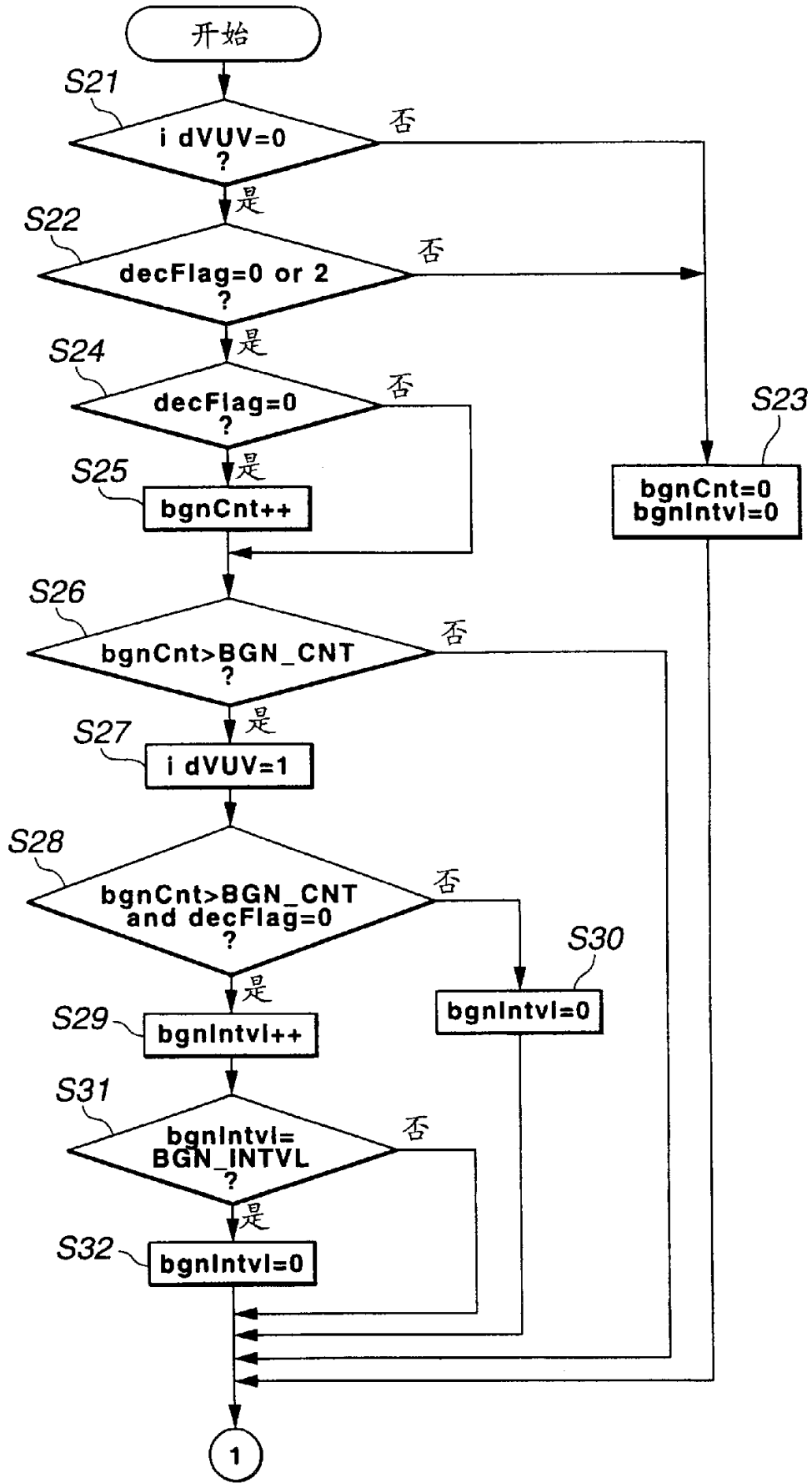


图 10

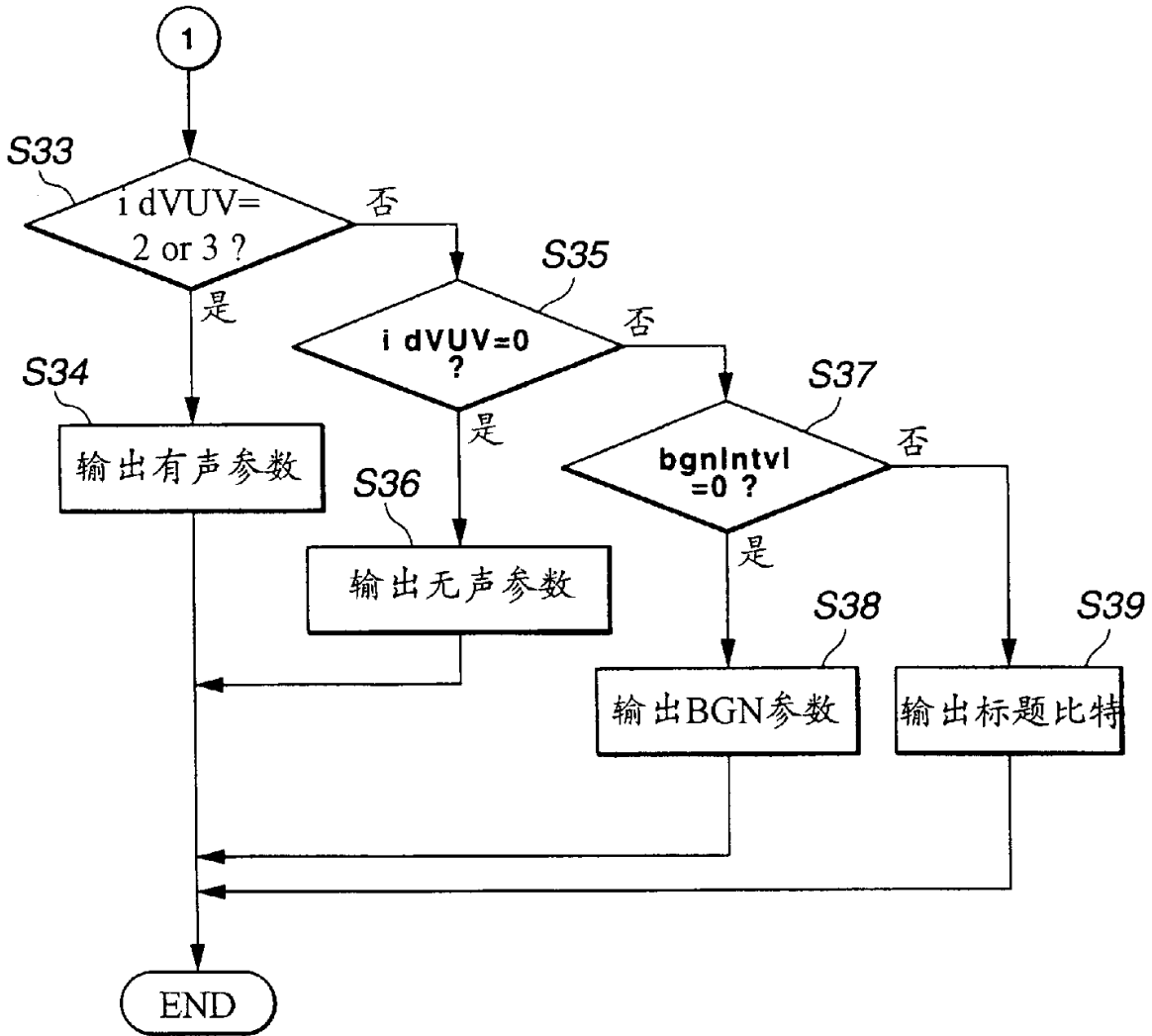
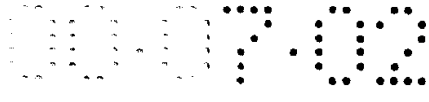


图 11



| 无声语音 | | 有声语音 | | 背景噪音(更新) | | 背景噪音(未更新) | |
|----------|-----|--------|-----|----------|-----|-----------|-----|
| 参数 | 比特数 | 参数 | 比特数 | 参数 | 比特数 | 参数 | 比特数 |
| 标题比特 | | | | | | | |
| idVUV | 2 | idVUV | 2 | idVUV | 2 | idVUV | 2 |
| - | - | - | - | 更新标记 | 1 | 更新标记 | 1 |
| LSP 参数 | | | | | | | |
| LSP0 | 5 | LSP0 | 5 | LSP0 | 5 | - | - |
| LSP2 | 7 | LSP2 | 7 | LSP2 | 7 | - | - |
| LSP3 | 5 | LSP3 | 5 | LSP3 | 5 | - | - |
| LSP4 | 1 | LSP4 | 1 | LSP4 | 1 | - | - |
| LSP5 | 8 | - | - | - | - | - | - |
| PITCH 参数 | | | | | | | |
| PCH | 7 | - | - | - | - | - | - |
| Celp 整形 | | | | | | | |
| Am 整形 | | | | | | | |
| idS0 | 4 | idSL00 | 6 | - | - | - | - |
| idS1 | 4 | idSL01 | 6 | - | - | - | - |
| Celp 增益 | | | | | | | |
| Am 增益 | | | | | | | |
| idG | 5 | idGL00 | 4 | idGL00 | 4 | - | - |
| - | - | idGL01 | 4 | - | - | - | - |
| Am 整形/4K | | | | | | | |
| idS0_4k | 7 | - | - | - | - | - | - |
| idS1_4k | 10 | - | - | - | - | - | - |
| idS2_4k | 9 | - | - | - | - | - | - |
| idS3_4k | 6 | - | - | - | - | - | - |
| 总数 | 80 | 总数 | 40 | 总数 | 25 | 总数 | 3 |

图 12

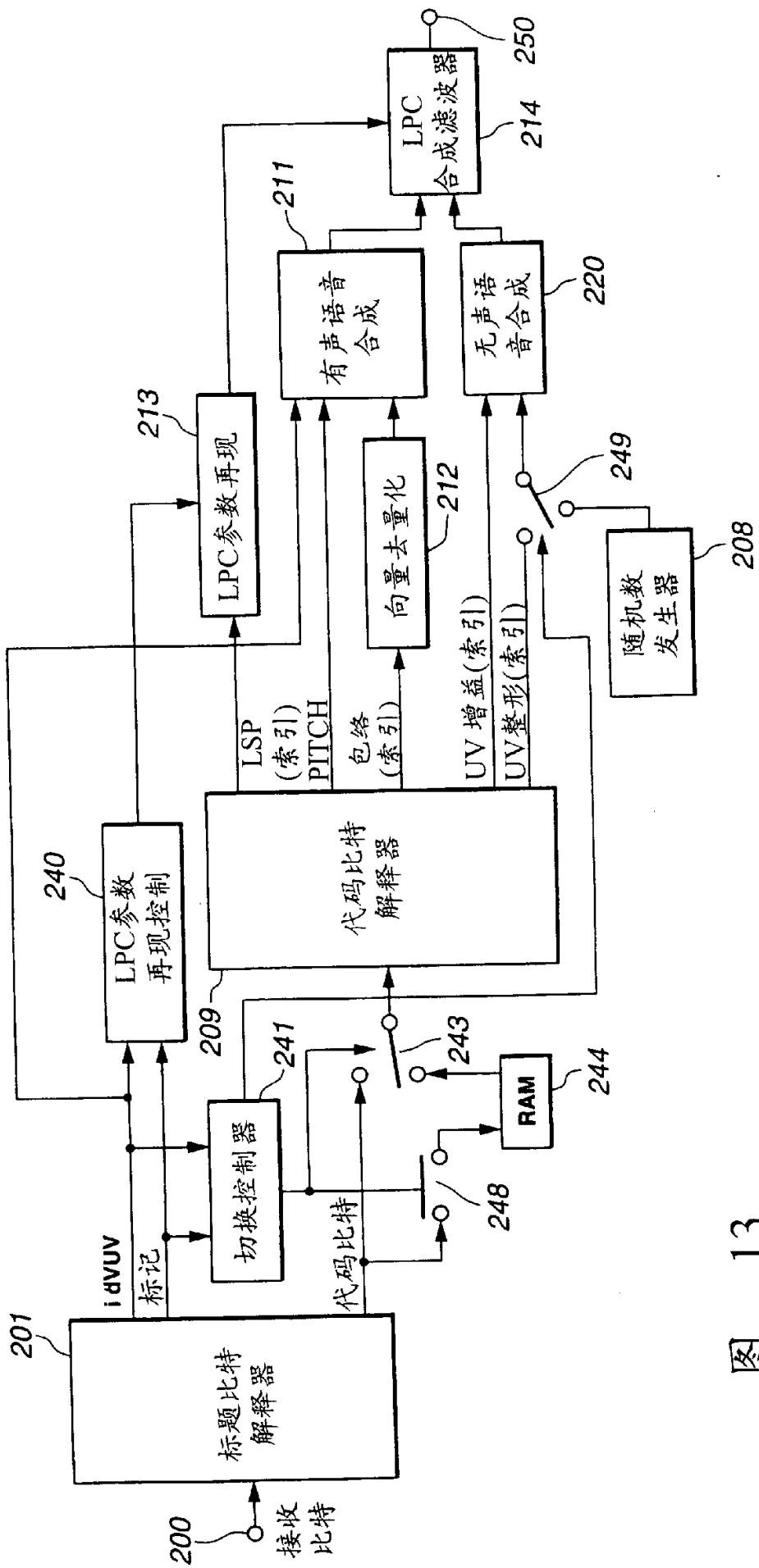


图 13

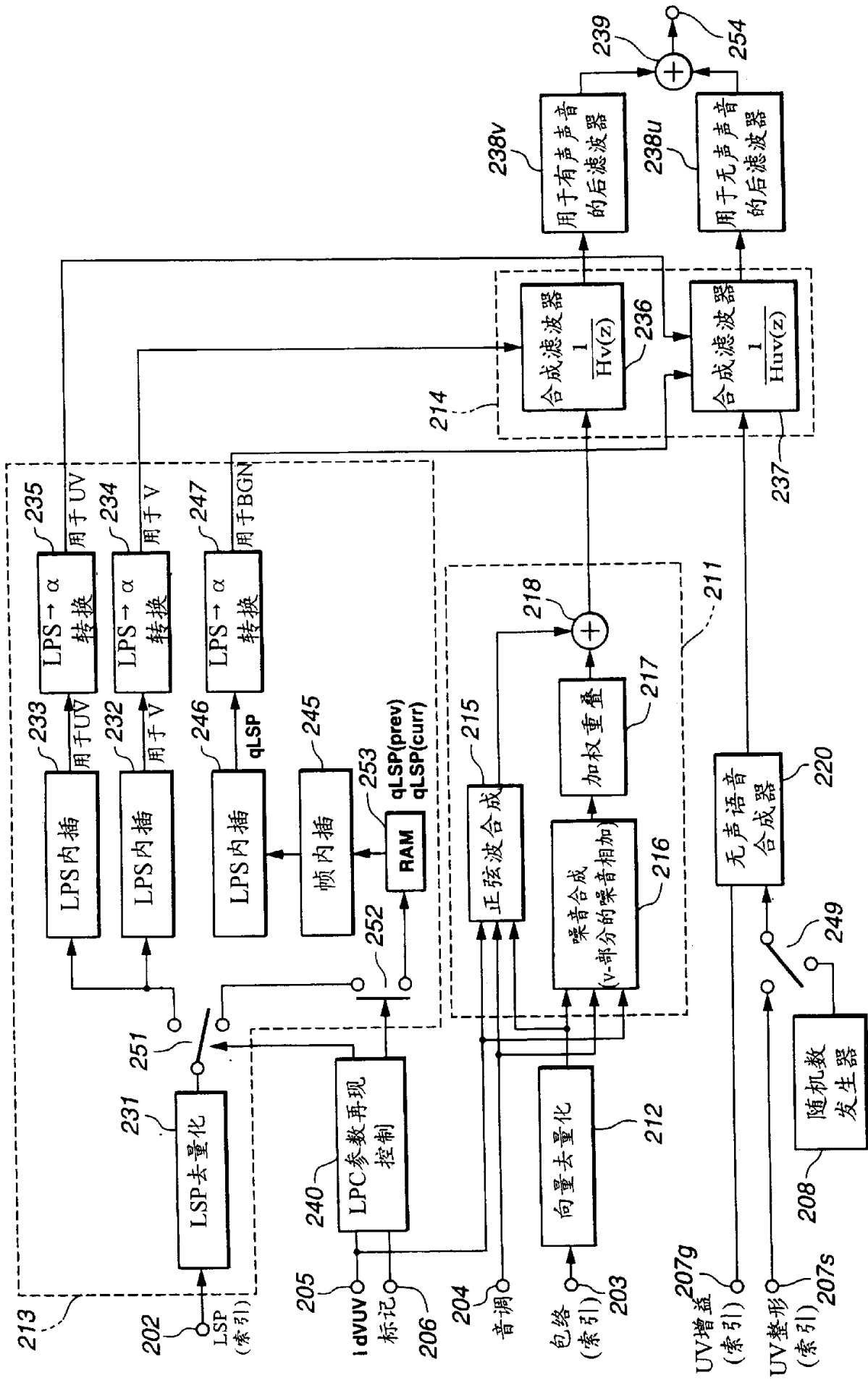


图 14

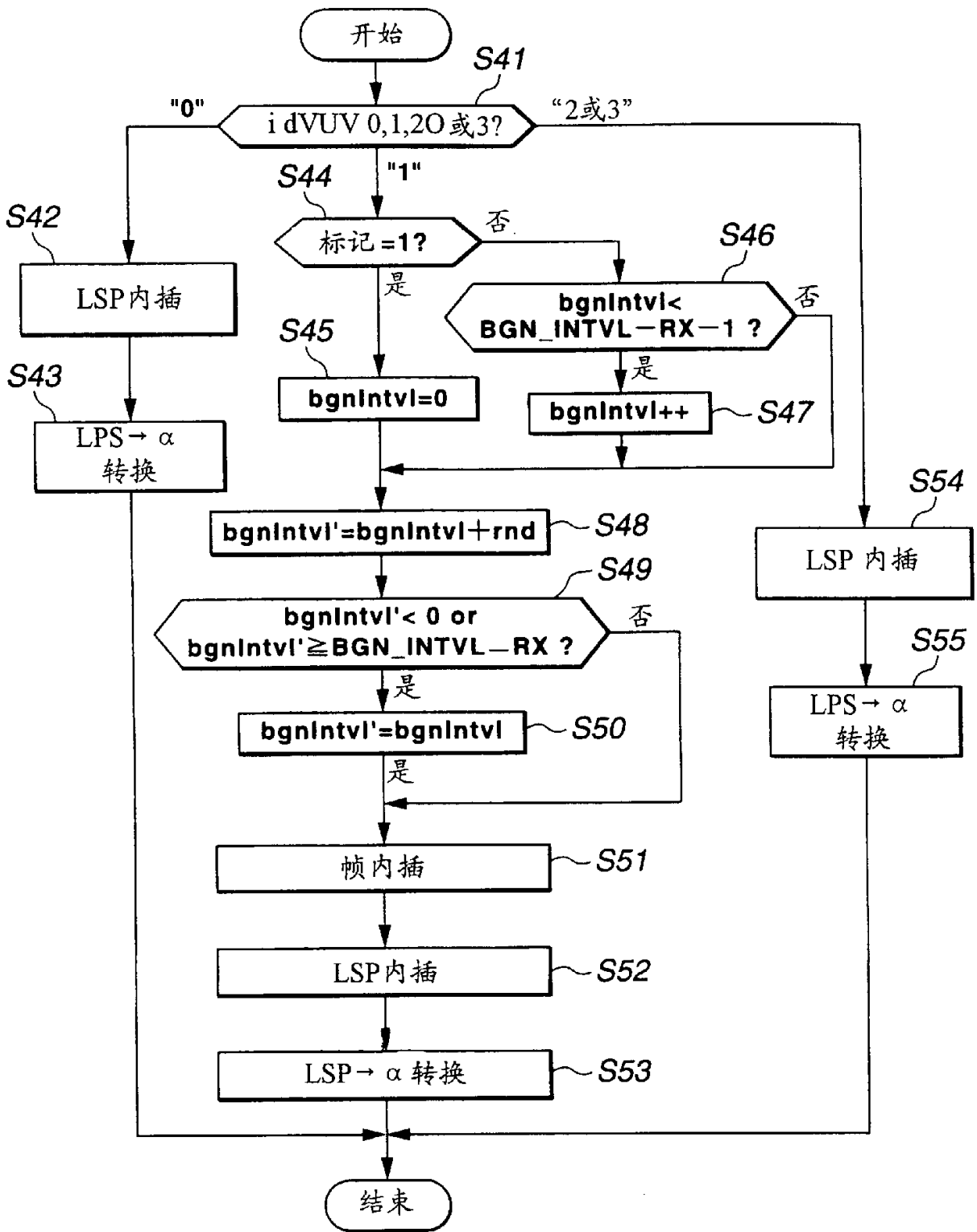


图 15

| | 0 | 1 | 2 |
|---------|---|---|---|
| 无声语音比特 | 0 | 0 | d |
| 无更新 BGN | 0 | 1 | 0 |
| 更新 BGN | 0 | 1 | 1 |
| 有声语音 | 1 | d | d |

(d 不限定)

图 16

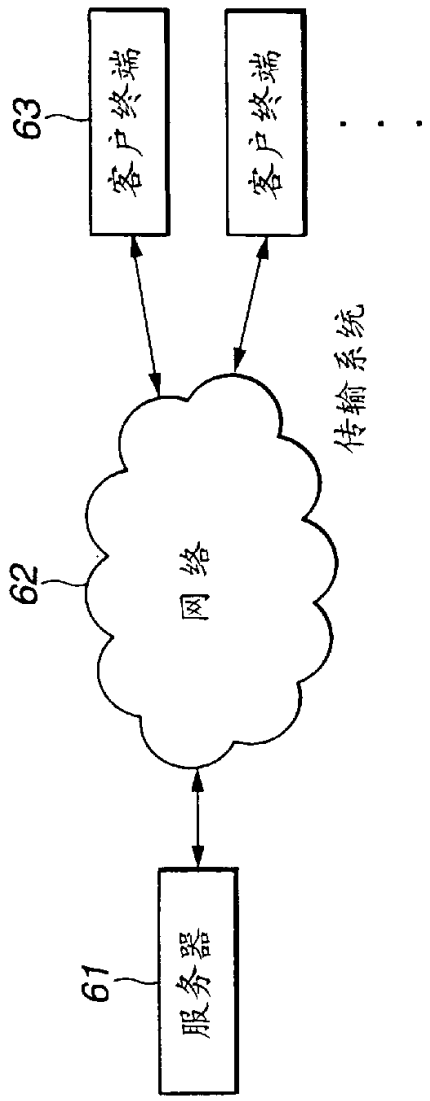


图 17

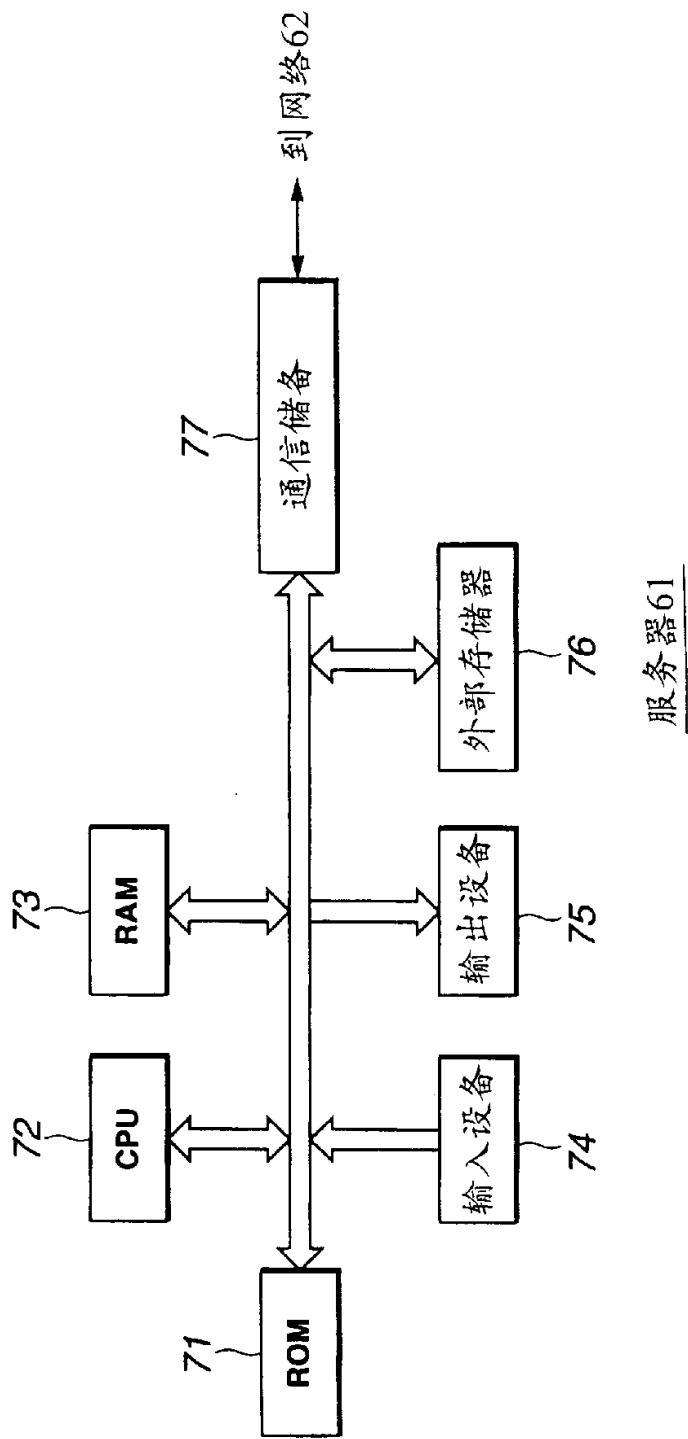


图 18

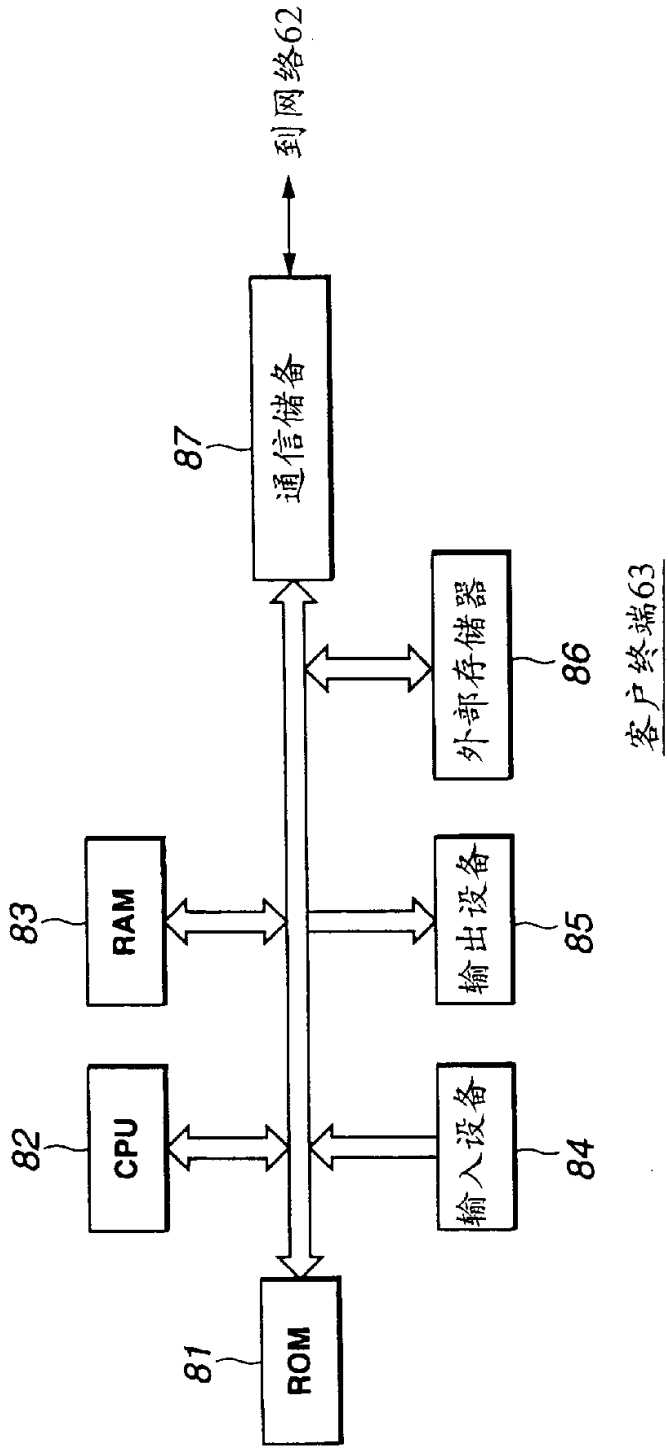


图 19