



# [12] 发明专利说明书

[21] ZL 专利号 01132421.X

[45] 授权公告日 2004 年 10 月 6 日

[11] 授权公告号 CN 1170268C

[22] 申请日 1997. 11. 6 [21] 申请号 01132421.X  
分案原申请号 97191558.X

[30] 优先权

[32] 1996. 11. 7 [33] JP [31] 294738/1996

[32] 1996. 11. 21 [33] JP [31] 310324/1996

[32] 1997. 2. 19 [33] JP [31] 34582/1997

[32] 1997. 2. 19 [33] JP [31] 34583/1997

[71] 专利权人 松下电器产业株式会社

地址 日本国大阪府

[72] 发明人 安永和敏 森井利幸 渡边泰助

江原宏幸

审查员 杨 叁

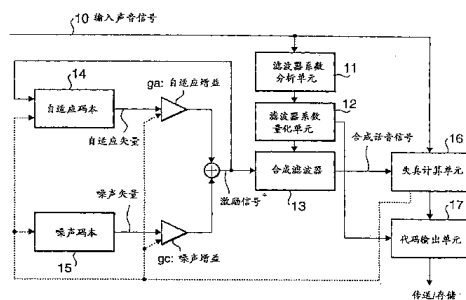
[74] 专利代理机构 上海专利商标事务所  
代理人 孙敬国

权利要求书 3 页 说明书 74 页 附图 22 页

[54] 发明名称 声音编码或解码装置及方法

[57] 摘要

本发明揭示一种声源矢量生成装置以及声音编码装置和声音解码装置。 将以往的 CELP 型声音编码/解码装置的噪声向量读出单元和噪声码本, 分别置换为根据输入振种的值输出不同矢量系列的振荡器和存储多个振种(产生振荡器振荡的“种子”)的振种存储单元。 由此, 不必将固定矢量原样地存储在固定码本(ROM)中, 能大幅度地减小存储器容量。



1. 一种 CELP 型声音编码或解码装置，其特征在于，包括  
用于生成自适应码矢量的自适应码本，  
确定输入声音的浊音/清音特征的装置，  
生成噪声码矢量的噪声码本，以及  
接收由噪声码本生成的噪声码矢量以进行 LPC 合成的合成滤波器，  
其中，所述噪声码本由声源矢量生成装置形成，当输入声音经确定是浊音时，  
所述声源矢量生成装置输出一输入脉冲矢量，当输入声音经确定是清音时，所述  
声源矢量生成装置输出一经过修改的输入脉冲矢量，这种经过修改的输入脉冲矢  
量是按照输入脉冲矢量的位置和极性配置固定波形存储装置中所存储的至少一个  
固定波形来生成的。
2. 如权利要求 1 所述的 CELP 型声音编码或解码装置，其特征在于，  
所述声源矢量生成装置自适应地通过偏置至少一个固定波形来实现所述配置  
步骤。
3. 如权利要求 1 所述的 CELP 型声音编码或解码装置，其特征在于，  
所述声源矢量生成装置自适应地通过将至少一个固定波形与输入脉冲矢量卷  
积来实现所述配置步骤。
4. 如权利要求 1 至 3 中任一项所述的 CELP 型声音编码或解码装置，其特征  
在于，  
根据量化的音调增益值，对输入声音的浊音/清音特征进行确定。
5. 如权利要求 1 至 3 中任一项所述的 CELP 型声音编码或解码装置，其特征  
在于，  
根据自适应码本增益值，对输入声音的浊音/清音特征进行确定。
6. 如权利要求 1 至 3 中任一项所述的 CELP 型声音编码或解码装置，其特征

在于，

根据量化的 LTP 增益值，对输入声音的浊音/清音特征进行确定。

7. 如权利要求 1 至 3 中任一项所述的 CELP 型声音编码或解码装置，其特征  
在于，

声源矢量生成装置通过改变所要配置或偏置或卷积的所述固定波形来生  
成一相应的噪声码矢量。

8. 如权利要求 1 至 3 中任一项所述的 CELP 型声音编码或解码装置，其特征  
在于，

声源矢量生成装置通过改变用于修改所述输入脉冲矢量的所述固定波形  
来生成一不同的噪声码矢量。

9. 如权利要求 1 至 3 中任一项所述的 CELP 型声音编码或解码装置，其特征  
在于，

所述输入脉冲矢量是一由代数码本提供的矢量。

10. 一种 CELP 型声音编码或解码方法，其特征在于，包括下述步骤：

确定输入声音的浊音/清音特征；

当输入声音经确定是浊音时，输出一由输入脉冲矢量组成的噪声码矢量，当  
输入声音经确定是清音时，输出一经过修改的输入脉冲矢量，这种经过修改的输  
入脉冲矢量是按照输入脉冲矢量的位置和极性配置至少一个固定波形来生成的；

用噪声码矢量进行 LPC 合成。

11. 如权利要求 10 所述的 CELP 型声音编码或解码方法，其特征在于，  
所述配置步骤按照输入脉冲矢量的位置和极性偏置所述固定波形。

12. 如权利要求 10 所述的 CELP 型声音编码或解码方法，其特征在于，  
所述配置步骤将所述固定波形与输入脉冲矢量卷积。

13. 如权利要求 10 至 12 中任一项所述的 CELP 型声音编码或解码方法，其

特征在于，

根据量化的音调增益值，对输入声音的浊音/清音特征进行确定。

14. 如权利要求 10 至 12 中任一项所述的 CELP 型声音编码或解码方法，其特征在于，

根据自适应码本增益值，对输入声音的浊音/清音特征进行确定。

15. 如权利要求 10 至 12 中任一项所述的 CELP 型声音编码或解码方法，其特征在于，

根据量化的 LTP 增益值，对输入声音的浊音/清音特征进行确定。

16. 如权利要求 10 至 12 中任一项所述的 CELP 型声音编码或解码方法，其特征在于，

通过改变所要配置或偏置或卷积的所述固定波形来生成噪声码矢量。

17. 如权利要求 10 至 12 中任一项所述的 CELP 型声音编码或解码方法，其特征在于，

通过改变用于修改所述输入脉冲矢量的所述固定波形，来生成一不同的噪声码矢量。

18. 如权利要求 10 至 12 中任一项所述的 CELP 型声音编码或解码方法，其特征在于，

所述输入脉冲矢量是一由代数码本提供的矢量。

### 声音编码或解码装置及方法

本申请是发明名称为“声源矢量生成装置以及声音编码装置和声音解码装置”、申请日为1997年11月6日、申请号为97191558.X的母案的分案申请。

#### 技术领域

本发明涉及能得到高品质合成声音的声源矢量生成装置以及能用低位速率对高品质的声音信号进行编码/解码的声音编码装置和声音解码装置。

#### 背景技术

CELP (Code Excited Linear Prediction: 编码激励线性预测) 型的声音编码装置, 是对每个以一定的时间划分声音的帧进行线性预测, 用存储过去的驱动声源的自适应码本和存储多个噪声矢量的噪声码本, 对每帧线性预测的预测残差(激励信号)进行编码的方式。例如在“低位速率高质量语音”(“High Quality Speech at Low Bit Rate” M. R. Schroeder, Proc. ICASSP' 85, PP937-940) 中公开的 CELP 型声音编码装置。

图1表示 CELP 型的声音编码装置的概略结构。CELP 型的声音编码装置将声音信息分离成声源信息和声道信息并进行编码。对于声道信息, 将输入声音信号 10 输入到滤波器系数分析单元 11 中并进行线性预测, 在滤波器系数量化单元 12 对线性预测系数(LPG)进行编码。借助于对合成滤波器 13 提供线性预测系数, 在合成滤波器 13 能将声道信号掺入声源信息中。对于声源信息, 在每个进一步细分帧的区间(称为子帧)进行自适应码本 14 的检索和噪声码本 15 的声源检索。自适应码本 14 的检索和噪声码本 15 的声源检索, 是决定使式(1)的编码失真最小的自适应码矢量的码号及其增益(音调增益)、和噪声码矢量的码号及其增益(噪声码增益)的处理。

$$\|v - (gaHp + gcHc)\|^2 \quad (1)$$

v: 声音信号(矢量)

H: 合成滤波器的脉冲响应卷积矩阵

$$H = \begin{bmatrix} h(0) & 0 & \Lambda & \Lambda & 0 & 0 \\ h(1) & h(0) & 0 & \Lambda & 0 & 0 \\ h(2) & h(1) & h(0) & 0 & 0 & 0 \\ M & M & M & 0 & 0 & 0 \\ M & M & M & 0 & h(0) & 0 \\ h(L-1) & \Lambda & \Lambda & \Lambda & h(1) & h(0) \end{bmatrix}$$

其中,  $h$ : 合成滤波器的脉冲响应(矢量)

$L$ : 帧长

$p$ : 自适应码矢量

$c$ : 噪声码矢量

$ga$ : 自适应码增益(音调增益)

$gc$ : 噪声码增益

但是, 因当闭环检索使式(1)为最小的前述码时, 码检索中所要的运算量变得膨大, 所以在一般的 CELP 型声音编码装置中, 首先进行自适应码本检索, 规定自适应码矢量的码号, 接着接受其结果, 进行噪声码本检索, 规定噪声码矢量的码号。

这里, 参照图 2A~图 2C 对 CELP 型声音编码装置的噪声码本检索进行说明。

图中, 符号  $x$  是基于式(2)求得的噪声码本检索用的目标矢量。设自适应码本检索已经结束。

$$x = v - gaHp \quad (2)$$

$x$ : 噪声码本检索目标(矢量)

$v$ : 声音信号(矢量)

$H$ : 合成滤波器的脉冲响应卷积矩阵

$p$ : 自适应码矢量

$ga$ : 自适应码增益(音调增益)

如图 2 所示, 噪声码本检索是规定使计算单元 16 中用式(3)定义的编码失真最小的噪声码矢量  $c$  的处理。

$$\|x - gcHc\|^2 \quad (3)$$

$x$ : 噪声码本检索目标(矢量)

$H$ : 合成滤波器的脉冲响应卷积矩阵

c: 噪声码矢量

gc: 噪声码增益

失真计算单元 16 对控制开关 21 进行控制, 切换从噪声码本 15 读出的噪声编码矢量, 直到定出噪声码矢量 c 为止。

为了减少计算的费用, 实际的 CELP 型声音编码装置为图 2B 的结构, 在失真计算单元 16' 中进行规定使式(4)的失真估算值最大的码号的处理。

$$\frac{(x'Hc)^2}{\|Hc\|^2} = \frac{((x'H)c)^2}{\|Hc\|^2} = \frac{(x''c)^2}{\|Hc\|^2} = \frac{(x''c)^2}{c'H'Hc} \quad (4)$$

x: 噪声码本检索目标(矢量)

H: 合成滤波器的脉冲响应卷积矩阵

H': H 的转置矩阵

x': 在 H 对 x 进行时间反转合成倒置所得矢量( $x'' = x'H$ )

c: 噪声码矢量

具体地说, 将噪声码本控制开关 21 连接到噪声码本 15 的 1 端, 从对应于该端的地址读出噪声码矢量 c。由合成滤波器 13, 将读出的噪声码矢量 c 与声道信息合成, 生成合成矢量 Hc。接着, 用对目标 x 进行时间反转、合成、时间反转得到的矢量 x'、以合成滤波器合成噪声码矢量的所得矢量 Hc 和噪声码矢量 c, 失真计算单元 16' 算出式(4)的失真估算值。然后, 切换噪声码本控制开关 21, 对噪声码本内的全部噪声矢量, 算出上述失真估算值。

最后, 将式(4)的失真估算值为最大时连接的噪声码本控制开关 21 的号码, 作为噪声码矢量的码号, 输出到编码输出单元 17 中。

图 2C 表示声音解码转置的部分结构。切换控制噪声码本控制开关 21, 以便读出被传送来的码号的噪声码矢量。在放大电路 23 和合成滤波器 24 中设定被传送来的噪声编码增益 gc 和滤波器系数后, 读出噪声码矢量并复原合成声音。

在前述的声音编码装置和解码装置中, 存储在噪声码本 15 中作为声源信息的噪声码矢量的越多, 越能检索接近实际声音的声源的噪声码矢量。但是, 因噪声码本(ROM)的容量有限制, 所以不能将对应于全部声源的无数的噪声码矢量存储在噪声码本中。因此, 在谋求声音品质的改善方面有其极限。

此外, 提议有能大幅度地降低失真计算单元的编码失真计算, 而且能减小噪声码本(ROM)的代数结构的声源(记载在“8KBIT/S ACELP CODING OF SPEECH WITH

10MS SPEECH-FRAME: A CANDIDATE FOR CCITT STANDARDIZATION” : R. Salami, C. Laflamme, J-P. Adoul, ICASSP' 94, pp. II-97~II-100, 1994 中)。

代数结构的声源预先计算合成滤波器的脉冲响应和时间反转的目标的卷积运算结果以及合成滤波器的自相关,并在存储器中展开,因而能大幅度地减少编码失真计算的费用。借助于代数生成噪声码矢量,能减小存储噪声码矢量的 ROM。在噪声码本中使用前述代数结构声源的 CS-ACELP 和 ACELP 分别被 ITU-T 作为 G. 729 建议和 G. 723. 1 建议提出。

但是,在将前述代数结构声源包括在噪声码本中的 CELP 型的声音编码装置/声音解码装置中,因不断地用脉冲串矢量对噪声码本检索用目标进行编码,所以在谋求声音品质的改善方面有其极限。

## 发明概述

鉴于前述实际情况,本发明的第 1 个目的是提供比原样在噪声码本中存储噪声码矢量的场合能大幅度地减小存储器容量,并能谋得声音品质改善的声源矢量生成装置以及声音编码装置和声音解码装置。

本发明的第 2 个目的是提供在噪声码本中包括代数结构声源,与用脉冲串矢量对噪声码本检索用目标进行编码的场合相比,能生成复杂的噪声码矢量,并能谋得声音品质改善的声源矢量生成装置以及声音编码装置和声音解码装置。

本发明将以往的 CELP 型声音编码/解码装置的固定矢量读出单元和固定码本,分别置换为输出与被输入的振种值相对应的不同的矢量系列的振荡器和存储多个振种(产生振荡器)的振种存储单元中。由此,不必将固定矢量原样存储在固定码本(ROM)中,能大幅度地减小存储器的容量。

本发明将以往的 CELP 型声音编码/解码装置的噪声矢量读出单元和噪声码本置换为振荡器和振种存储单元。由此,不必将噪声矢量原样存储在固定码本(ROM)中,能大幅度地减小存储器的容量。

本发明声源矢量生成装置的结构是:存储多个固定波形,根据起始端候补位置信息将各固定波形配置在各自的起始端位置上,并对这些固定波形进行加法运算,生成声源矢量。因此,能生成接近实际声音的声源矢量。

本发明是噪声码本采用前述声源矢量生成装置而构成的 CELP 型声音编码/解码装置。固定波形配置单元也可以代数生成固定波形的起始端候补位置信息。



本发明的 CELP 型声音编码/解码装置做成存储多个固定波形，生成与每个固定波形起始端候补位置信息相对应的脉冲，对合成滤波器的脉冲响应和各自的固定波形卷积，生成波形别脉冲响应，计算前述波形别脉冲响应的自相关和互相关，并在相关矩阵存储器中展开。由此，能得到与以代数结构声源作为噪声码本使用的场合相同程度的计算费用，同时能改善合成声音的品质的声音编码/解码装置。

本发明的 CELP 型声音编码/解码装置包括多个噪声码本和从前述多个噪声码本中选择一个的切换手段，也可以至少以一个噪声码本作为前述声源矢量生成装置，此外，也可以至少以一个噪声码本作为存储多个随机数序列的向量存储单元或者存储多个脉冲串的脉冲串存储单元，或者至少有两个具有前述声源矢量生成装置的噪声码本，而且各噪声码本存储的固定波形个数不同，还可以使切换手段选择任一噪声码本，使噪声码本检索时编码失真为最小，或者根据声音区间分析结果，自适应选择任一个噪声码本。

#### 附图简要说明

图 1 表示以往的 CELP 型声音编码装置的概略图。

图 2A 是图 1 的声音编码装置的声源矢量生成单元的方框图。

图 2B 谋求减少计算费用的变形的声源矢量生成单元的方框图。

图 2C 是与图 1 的声音编码装置配对使用的声音解码装置中声源矢量生成单元的方框图。

图 3 表示与实施形态 1 相关的声音编码装置的主要部分的方框图。

图 4 表示包括在实施形态 1 的声音编码装置中的声源矢量生成装置的方框图。

图 5 表示实施形态 2 的声音编码装置的主要部分的方框图。

图 6 表示包括在实施形态 2 的声音编码装置中的声源矢量生成装置的方框图。

图 7 表示与实施形态 3 和 4 相关的声音编码装置的主要部分的方框图。

图 8 表示包括在实施形态 3 的声音编码装置中的声源矢量生成装置的方框图。

图 9 表示包括在实施形态 4 的声音编码装置中的非线性数字滤波器的方框

图。

图 10 表示图 9 所示的非线性数字滤波器的加法特性图。

图 11 表示与实施形态 5 相关的声音编码装置的主要部分的方框图。

图 12 表示与实施形态 6 相关的声音编码装置的主要部分的方框图。

图 13A 表示与实施形态 7 相关的声音编码装置的主要部分的方框图。

图 13B 表示与实施形态 7 相关的声音编码装置的主要部分的方框图。

图 14 表示与实施形态 8 相关的声音解码装置的主要部分的方框图。

图 15 表示与实施形态 9 相关的声音编码装置的主要部分的方框图。

图 16 表示包括在实施形态 9 的声音编码装置中的量化对象 LSP 增加部分的方框图。

图 17 表示包括在实施形态 9 的声音编码装置中的 LSP 量化·解码单元的方框图。

图 18 表示与实施形态 10 相关的声音编码装置的主要部分的方框图。

图 19A 表示与实施形态 11 相关的声音编码装置的主要部分的方框图。

图 19B 表示与实施形态 11 相关的声音解码装置的主要部分的方框图。

图 20 表示与实施形态 12 相关的声音编码装置的主要部分的方框图。

图 21 表示与实施形态 13 相关的声音编码装置的主要部分的方框图。

图 22 表示与实施形态 14 相关的声音编码装置的主要部分的方框图。

图 23 表示与实施形态 15 相关的声音编码装置的主要部分的方框图。

图 24 表示与实施形态 16 相关的声音编码装置的主要部分的方框图。

图 25 表示与实施形态 16 相关的矢量量化部分的方框图。

图 26 表示与实施形态 17 相关的声音编码装置的参数编码部分的方框图。

图 27 表示与实施形态 18 相关的降噪装置的方框图。

### 实施发明的最佳方式

下面，参照附图对本发明的实施形态具体地进行说明。

#### 实施形态 1

图 3 表示与实施形态 1 相关的声音编码装置的主要部分的方框图。这种声音编码装置包括具有振种存储单元 31 和振荡器 32 的声源矢量生成装置 30，和 LPC 合成滤波器单元 33。

将从振种存储单元 31 输出的振种(产生振荡的“种子”)34 输入到振荡器 32 中。与输入的振种值相对应, 振荡器 32 输出不同的矢量系列。振荡器 32 用对应于振种(产生振荡的“种子”)34 的值的內容进行振荡, 并输出作为矢量系列的声源矢量 35。LPC 合成滤波器单元 33 用合成滤波器的脉冲响应卷积矩阵的形式, 提供声道信息, 以脉冲响应对声源矢量 35 进行卷积运算后输出合成话音 36。将以脉冲响应对声源矢量 35 进行卷积运算称为 LPC 合成。

图 4 表示声源矢量生成装置 30 的具体的结构。按照由失真计算单元提供的控制信号, 振种存储单元控制开关 41 切换从振种存储单元 31 读出的振种。

这样, 仅将从振荡器 32 输出不同的矢量系列的多个振种预先存储在振种存储单元 31 中, 与将复杂的噪声码向量原样存储在噪声码本中的场合相比, 能用较小的容量发生更多的噪声码矢量。

此外, 虽然在本实施形态中对声音编码装置进行了说明, 但也能将声源矢量生成装置 30 用于声音解码装置中。这种场合, 在声音解码装置中具有与声音编码装置的振种存储单元 31 相同內容的振种存储单元, 并将编码时选择的振种号码提供给振种存储单元控制开关 41。

## 实施形态 2

图 5 表示基于本实施形态的声音编码装置的主要部分的方框图。这种声音编码装置包括具有振种存储单元 51 和非线性振荡器 52 的声源矢量生成装置 50, 和 LPC 合成滤波器单元 53。

将从振种存储单元 51 输出的振种(产生振荡的“种子”)54 输入到非线性振荡器 52 中。从非线性振荡器 52 输出的作为矢量系列的声源矢量 55, 输入到 LPC 合成滤波器单元 53 中。合成滤波器单元 53 的输出是合成话音 56。

非线性振荡器 52 输出对应于输入的振种 54 的值的不同的矢量系列, LPC 合成滤波器单元 53 对输入的声源矢量 55 进行 LPC 合成, 并输出合成话音 56。

图 6 表示声源矢量生成装置 50 的功能的方框图。按照由失真计算单元提供的控制信号, 振种存储单元控制开关 41 切换从振种存储单元 51 读出的振种。

这样, 借助于在声源矢量生成装置 50 的振荡器中使用非线性振荡器 52, 利用遵循非线性特性的振荡, 能抑制发散, 得到实用的声源矢量。

此外, 虽然在本实施形态中对声音编码装置进行了说明, 但也能将声源矢量生成装置 50 用于声音解码装置中。这种场合, 在声音解码装置中包括与声音编

码装置的振种存储单元 51 相同内容的振种存储单元，并将编码时选择的振种号码提供给振种存储单元控制开关 41。

### 实施形态 3

图 7 表示基于本实施形态的声音编码装置的主要部分的方框图。这种声音编码装置包括具有振种存储单元 71 和非线性数字滤波器 72 的声源矢量生成装置 70，和 LPC 合成滤波器单元 73。74 是从振种存储单元 71 输出并输入到非线性数字滤波器 72 中的振种(产生振荡的“种子”)，75 是作为从非线性数字滤波器 72 输出的矢量系列的声源矢量，76 是从 LPC 合成滤波器 73 输出的合成话音。

如图 8 所示，声源矢量生成装置 70 具有利用由失真计算单元供给的控制信号，切换从振种存储单元 71 读出的振种 74 的振种存储单元控制开关 41。

非线性数字滤波器 72 输出对应于输入的振种的值的不同的矢量系列，LPC 合成滤波器单元 73 对输入的声源矢量 75 进行 LPC 合成，并输出合成话音 76。

这样，借助于在声源矢量生成装置 70 的振荡器中使用非线性数字滤波器 72，利用遵循非线性特性的振荡，能抑制发散，得到实用的声源矢量。

此外，虽然在本实施形态中对声音编码装置进行了说明，但也能将声源矢量生成装置 70 用于声音解码装置中。这种场合，在声音解码装置中包括与声音编码装置的振种存储单元 71 相同内容的振种存储单元，并将编码时选择的振种号码提供给振种存储单元控制开关 41。

### 实施形态 4

如图 7 所示，与本实施形态相关的声音编码装置包括具有振种存储单元 71 和非线性数字滤波器 72 的声源矢量生成装置 70，和 LPC 合成滤波器单元 73。

特别指出的是，非线性数字滤波器 72 具有图 9 所示的结构。这种非线性数字滤波器 72 包括具有如图 10 所示的非线性加法特性的加法器 91，具有保存数字滤波器的状态 ( $y(k-1) \sim y(k-N)$  的值) 的作用的状态变量保持单元 92~93，以及并联连接到各状态变量保持单元 92~93 的输出上，将状态变量中乘以增益后，输出到加法器 91 中的乘法器 94~95。根据从振种存储单元 71 读出的振种，状态变量保持单元 92~93 设定状态变量初始值。乘法器 94~95 限定增益的值，使数字滤波器的极点存在于  $Z$  平面的单位圆外。

图 10 是表示包括在非线性数字滤波器 72 中的加法器 91 的非线性加法特性的概念图，表示具有 2 的补数特性的加法器 91 的输入输出关系。加法器 91 首先

求得作为对加法器 91 的输入值总和的加法器输入和，接着使用图 10 所示的非线性特性，以算出对该输入和的加法器输出。

特别是，因非线性数字滤波器 72 采用 2 次全极结构，所以串联连接 2 个状态变量保持单元 92、93，并对状态变量保持单元 92、93 连接乘法器 94、95。采用加法器 91 的非线性加法特性为 2 的补数的数字滤波器。此外，振种存储单元 71，特别存储记载在表 1 中的 32 字的振种矢量。

表 1：噪声矢量生成用的振种矢量

i	$Sy(n-1)[i]$	$Sy(n-2)[i]$	i	$Sy(n-1)[i]$	$Sy(n-2)[i]$
1	0.250000	0.250000	9	0.109521	-0.761210
2	-0.564643	-0.104927	10	-0.202115	0.198718
3	0.173879	-0.978792	11	-0.095041	0.863849
4	0.632652	0.951133	12	-0.634213	0.424549
5	0.920360	-0.113881	13	0.948225	-0.184861
6	0.864873	-0.860368	14	-0.958269	0.969458
7	0.732227	0.497037	15	0.233709	-0.057248
8	0.917543	-0.035103	16	-0.852085	-0.564948

在前述结构的语音编码装置中，将从振种存储单元 71 读出的振种矢量作为初始值，供给非线性数字滤波器 72 的状态变量保持单元 92、93。非线性数字滤波器 72 每从输入矢量(0 系列)将 0 输入到加法器 91 中，就输出 1 个采样( $y(k)$ )，并作为状态变量顺次地传送到状态变量保持单元 92、93 中。这时，对从状态变量保持单元 92、93 输出的状态变量，分别由各乘法器 94、95 乘以增益  $a_1$ 、 $a_2$ 。用加法器 91 对乘法器 94、95 的输出进行相加，求出加法器输入和，并根据图 10 的特性，发生抑制在  $+1 \sim -1$  之间的加法器输出。在输出这种加法器输出( $y(k+1)$ )作为声源矢量的同时，顺次地传送到状态变量保持单元 92、93 中，生成新的采样( $y(k+2)$ )。

在本实施形态中，作为非线性数字滤波器，为了极存在于  $Z$  平面的单位圆外，

特地固定乘法器 94~95 的系数 1~N, 使加法器 91 持有非线性加法特性, 因而即使非线性数字滤波器 72 的输入变大, 也能抑制输出发散, 可连续生成能实用的声源矢量。还能确保生成的声源矢量的随机性。

此外, 虽然在本实施形态中对声音编码装置进行了说明, 但也能将声源矢量生成装置 70 用于声音解码装置中。这种场合, 在声音解码装置中包括与声音编码装置的振种存储单元 71 相同内容的振种存储单元, 并将编码时选择的振种号码提供给振种存储单元控制开关 41。

### 实施形态 5

图 11 表示基于本实施形态的声音编码装置的主要部分的方框图。这种声音编码装置包括具有声源存储单元 111 和声源加法矢量生成单元 112 的声源矢量生成装置 110, 和 LPC 合成滤波器单元 113。

声源存储单元 111 存储过去的声源矢量, 利用接受来自未图示的失真计算单元的控制信号的控制开关, 读出声源矢量。

声源加法矢量生成单元 112, 对从声源存储单元 111 读出的过去的声源矢量, 施行用生成矢量特定号码指示的规定的处理, 生成新的声源矢量。声源加法矢量生成单元 112, 具有对应于生成矢量特定号码, 切换对过去的声源矢量的处理内容的功能。

在如前所述结构的声音编码装置中, 从例如执行声源检索的失真计算单元供给生成矢量特定号码。声源加法矢量生成单元 112, 根据输入生成矢量特定号码的值对过去的声源矢量进行不同的处理, 生成不同的声源加法矢量, 而且 LPC 合成滤波器单元 113 对输入的声源矢量进行 LPC 合成并输出合成话音。

采用本实施形态, 则仅将少数的过去的声源矢量预先存储在声源存储单元 111 中, 并切换在声源加法矢量生成单元 112 的处理内容, 就能生成随机的声源矢量, 因不必预先将噪声矢量原样地存储在噪声码本 (ROM) 中, 所以能大幅度地减小存储器的容量。

此外, 虽然在本实施形态中对声音编码装置进行了说明, 但也能将声源矢量生成装置 110 用于声音解码装置中。这种场合, 在声音解码装置中包括与声音编码装置的声源存储单元 111 相同内容的声源存储单元, 并对声源加法矢量生成单元 112 提供编码时选择的生成矢量特定号码。

### 实施形态 6

图 12 表示与本实施形态相关的声源矢量生成装置的功能的方框图。这种声源矢量生成装置包括声源加法矢量生成单元 120 和存储多个要素矢量 1~N 的声源存储单元 121。

声源加法矢量生成单元 120 包括进行从声源存储单元 121 的不同的位置读出多个不同长度的要素矢量的处理的读出处理单元 122，进行对读出处理后的多个要素矢量作倒置排列变换的处理的倒置处理单元 123，进行对倒置处理后的多个矢量分别乘以不同的增益的处理的乘法处理单元 124，进行缩短乘法处理后的多个矢量的矢量长度的处理的间抽处理单元 125，进行伸长间抽处理后的多个矢量的矢量长度的处理的内插处理单元 126，进行使内插处理后的多个矢量相加的处理的加法处理单元 127，以及同时具有决定对应于所输入生成矢量特定号码值的具体处理方法并对决定各处理单元作指示的功能及保持决定该具体处理内容时参照的号码变换对应映射表 2 的功能的处理判定和指示单元 128。

表 2：号码变换对应映射

位串 (MS...LSB)	6	5	4	3	2	1	0
V1 读出位置 (16 种)				3	2	1	0
V2 读出位置 (32 种)	2	1	0				
V3 读出位置 (32 种)	4	3	2	1	0		
反向处理 (2 种)							0
乘法处理 (4 种)	1	0					
间抽处理 (4 种)				1	0		
内插处理 (2 种)			0				

这里，对声源加法矢量生成单元 120 进一步详细地进行说明。声源加法矢量生成单元 120 将输入生成矢量特定号码(用 7 位的位串取 0 到 127 的整数)与号码变换对应映射表 2 进行比较，以决定读出处理单元 122、倒置处理单元 123、乘法处理单元 124、间抽间距处理单元 125、内插处理单元 126、加法处理单元 127 的各自的具体地处理方法，并向各处理单元输出其具体的处理方法。

首先，着眼于输入生成矢量特定号码的低端的 4 位串 ( $n_1$ : 从 0 到 15 的整数值)，从声源存储单元 121 的一端到  $n_1$  的位置为止，切出长度 100 的要素矢量 1 ( $V_1$ )。接着，着眼于结合输入生成矢量特定号码的低端的 2 位串和高端 3 位串的 5 位串 ( $n_2$ : 从 0 到 31 的整数值)，从声源存储单元 121 的一端到  $n_2+14$  (从 14 到 45 的整数值) 的位置为止，切出长度 78 的要素矢量 2 ( $V_2$ )。进而，着眼于输入生成矢量特定号码的高端 5 位串 ( $n_3$ : 从 0 到 31 的整数值)，从声源存储单元 121 的一端到  $n_3+46$  (从 46 到 77 的整数值) 的位置为止，切出长度  $N_s (=52)$  的要素矢量 3 ( $V_3$ )。读出处理单元 122 进行向倒置处理单元 123 输出  $V_1$ 、 $V_2$ 、 $V_3$  的处理。

如果生成矢量特定号码的最低端的 1 位是“0”，则倒置处理单元 123 进行以倒置排列变换  $V_1$  和  $V_2$  和  $V_3$  的矢量作为新的  $V_1$ 、 $V_2$ 、 $V_3$  并输出到乘法处理单元 124 中的处理，如果生成矢量特定号码的最低端的 1 位是“1”，则进行原样地将  $V_1$  和  $V_2$  和  $V_3$  输出到乘法处理单元 124 中的处理。

乘法处理单元 124 着眼于组合输入生成矢量特定号码的高端第 7 位和高端第 6 位的 2 位串，如果该位串是‘00’，则  $V_2$  的振幅乘-2 倍、如果该位串是‘01’，则以  $V_3$  的振幅的-2，如果该位串是‘10’，则  $V_1$  的振幅乘-2，如果该位串是‘11’，则  $V_2$  的振幅乘 2，所得各矢量分别作为新的  $V_1$ 、 $V_2$ 、 $V_3$ ，输出到间抽单元 125 中。

间抽单元 125 着眼于组合输入生成矢量特定号码的高端第 4 位和高端第 3 位的 2 位串，如果该位串是(a) ‘00’，则从  $V_1$ 、 $V_2$ 、 $V_3$  开始间隔 1 个采样，取出 26 个采样的矢量作为新的  $V_1$ 、 $V_2$ 、 $V_3$ ，输出到内插处理单元 126 中，如果该位串是(b) ‘01’，则从  $V_1$ 、 $V_3$  开始间隔 1 个采样，从  $V_2$  开始间隔 2 个采样，取出 26 个采样的矢量作为新的  $V_1$ 、 $V_2$ 、 $V_3$ ，输出到内插处理单元 126 中，如果该位串是(c) ‘10’，则从  $V_1$  开始间隔 3 个采样，从  $V_2$ 、 $V_3$  开始间隔 1 个采样，取出 26 个采样的矢量作为新的  $V_1$ 、 $V_2$ 、 $V_3$ ，输出到内插处理单元 126 中，如果该位串是(d) ‘11’，则从  $V_1$  开始间隔 3 个采样，从  $V_2$  开始间隔 2 个采样，从  $V_3$  开始间隔 1 个采样，取出 26 个采样的矢量作为新的  $V_1$ 、 $V_2$ 、 $V_3$ ，输出到内插处理单元 77 中，

内部插处理单元 126 着眼于输入生成矢量特定号码的高端第 3 位，如果其值是(a) ‘0’，则以将  $V_1$ 、 $V_2$ 、 $V_3$  分别代入长度  $N_s (=52)$  的 0 矢量的偶数号采样



中的矢量作为新的 V1、V2、V3，输出到加法处理单元 75 中，如果其值是 (b) ‘1’，则以将 V1、V2、V3 分别代入长度  $N_s (=52)$  的 0 矢量的奇数号采样中的矢量作为新的 V1、V2、V3，输出到加法处理单元 75 中。

加法处理单元 127 对由内插处理单元 126 生成的 3 个矢量 (V1、V2、V3) 进行加法运算，生成并输出声源加法矢量。

这样，本实施形态，因对应于生成矢量特定号码随机地组合多个处理，生成随机的声源矢量，所以不必预先将噪声矢量原样地存储在噪声码本 (ROM) 中，能大幅度地减小存储器的容量。

此外，借助于在实施形态 5 的声音编码装置中使用本实施形态的声源矢量生成装置，不必持有大容量的噪声码本，就能生成复杂随机的声源矢量。

#### 实施形态 7

下面，在以作为日本国内的 PDC 数字便携电话的声音编码/解码标准方式的 PSI-CELP 为基础做成的 CELP 型声音编码装置中，使用前述的实施形态 1~实施形态 6 的任一个所示的声源矢量生成装置的例，作为实施形态 7 进行说明。

图 13A 和图 13B 表示与实施形态 7 相关的声音编码装置的方框图。在这种编码装置中，将数字化的输入声音数据 1300 以帧为单位 (帧长  $N_f = 104$ ) 供给到缓存器 1301 中。这时，由供给的新的数据更新缓冲器 1301 中的旧的数据。帧功率量化和解码单元 1302 首先从缓存器 1301 中读出长度  $N_f (=104)$  的处理帧  $s(i)$  ( $0 \leq i \leq N_f - 1$ )，由式 (5) 求出该处理帧内采样的平均功率  $\text{amp}$ 。

$$\text{amp} = \sqrt{\frac{\sum_{i=0}^{N_f} s^2(i)}{N_f}} \quad (5)$$

$\text{amp}$ : 处理帧内采样的平均功率

$i$ : 处理帧内的要素号码 ( $0 \leq i \leq N_f - 1$ )

$s(i)$ : 处理帧内采样

$N_f$ : 处理帧长 ( $=52$ )

利用式 (6)，将求得的处理帧内采样的平均功率  $\text{amp}$  变换成对数变换值  $\text{amplog}$ 。

$$\text{amplog} = \frac{\log_{10}(255 \times \text{amp} + 1)}{\log_{10}(255 + 1)} \quad (6)$$

amplog: 处理帧内采样的平均功率的对数变换值

amp: 处理帧内采样的平均功率

将求得的 amplog 存储在功率量化表存储单元 1303 中, 用表 3 所示的 10 字的标量量化用表 Cpow 进行标量量化, 得到 4 位的功率索引 Ipow, 从得到 4 位的功率索引 Ipow 求出解码帧功率 spow, 并将功率索引 Ipow 和解码帧功率 spow 输出到参数编码单元 133 中。功率量化表存储单元 1303 存储 16 字的功率标量量化表(表 3), 在帧功率量化·解码单元 1302 对处理帧内采样的平均功率的对数变换值进行标量量化时参照该表。

表 3: 功率标量量化用表

i	Cpow(i)	i	Cpow(i)
1	0.00675	9	0.39247
2	0.06217	10	0.42920
3	0.10877	11	0.46252
4	0.16637	12	0.49503
5	0.21876	13	0.52784
6	0.26123	14	0.56484
7	0.30799	15	0.61125
8	0.35228	16	0.67498

LPC 分析单元 1304, 首先从缓存器 1301 读出分析区间长度  $N_w (=256)$  的分析区间数据, 在读出的分析区间数据上乘以窗长  $N_w (=256)$  的汉明窗  $W_h$ , 得到乘以汉明窗后的分析区间数据后, 多次求所得乘以汉明窗后的分析区间数据的自相关函数, 直到次数为预测次数  $N_p (=10)$  为止。。在求得的自相关函数上乘以存储在滞后窗存储单元 1305 中的 10 字的滞后窗表(表 4), 得到乘以滞后窗后的自相关函数, 对于得到的乘以滞后窗后的自相关函数, 进行线性预测分析, 算出 LPC 的参数  $\alpha(i) (1 \leq i \leq N_p)$ , 并输出到音调预选单元 1308 中。

表 4: 滞后窗表

i	Wlag(i)	i	Wlag(i)
0	0.9994438	5	0.9801714
1	0.9977772	6	0.9731081
2	0.9950056	7	0.9650213
3	0.9911382	8	0.9559375
4	0.9861880	9	0.9458861

接着, 将求得的 LPC 参数  $\alpha(i)$  变换成 LSP(线频谱对)  $\omega(i)$  ( $1 \leq i \leq N_p$ ), 并输出到量化/解码单元 1306 中。滞后窗存储单元 1305 存储 LPC 分析单元参照的滞后窗。

LSP 量化/解码单元 1306, 首先参照 LSP 量化表存储单元 1307 中存储的 LSP 的矢量量化用表, 对从 LPC 分析单元 1304 接收到的 LSP 进行矢量量化, 选择最佳索引, 并以选择的索引作为 LSP 码  $I_{lsp}$  输出到参数编码单元 1331 中。接着, 从 LSP 量化表存储单元 1307 读出对应于 LSP 码的形心作为解码 LSP  $\omega_q(i)$  ( $1 \leq i \leq N_p$ ), 并将读出的解码 LSP 输出到 LSP 插入单元 1311 中。此外, 将解码 LSP 变换成 LPC, 得到解码 LSP  $\alpha_q(i)$  ( $1 \leq i \leq N_p$ ), 并将得到的解码 LPC 输出到矢量加权滤波器系数运算单元 1312 和听觉加权 LPC 合成滤波器系数运算单元 1314 中。

LSP 量化表存储单元 1307 存储 LSP 量化/解码单元 1306 对 LSP 进行矢量量化时参照的 LSP 矢量量化表。

音调预选单元 1308, 首先对从缓存器 1301 读出的处理帧数据  $s(i)$  ( $1 \leq i \leq N_f - 1$ ), 施行根据由 LPC 分析单元 1304 接收到的 LSP  $\alpha(i)$  ( $1 \leq i \leq N_p$ ) 构成的线性预测反滤波, 得到线性预测残差信号  $res(i)$  ( $1 \leq i \leq N_f - 1$ ), 计算得到的线性预测残差信号  $res(i)$  的功率, 求得用处理子帧声音采样功率使计算的残差信号功率归一化的值的归一化预测残差功率  $resid$ , 并输出到参数编码单元 1331 中。接着, 在线性预测残差信号  $res(i)$  上乘以长度  $N_w (= 256)$  的汉明窗, 生成乘以汉明

窗后的线性预测残差信号  $resw(i)$  ( $1 \leq i \leq Nw-1$ ), 在  $Lmin-2 \leq i \leq Lmax+2$  (其中,  $Lmin$  为长期预测系数的最短分析区间为 16、 $Lmax$  为长期预测系数的最长分析区间, 分别取为 16 的 128) 的范围内, 求得生成的  $resw(i)$  的自相关函数  $\phi_{int}(i)$ 。在求得的自相关函数  $\phi_{int}(i)$  上叠加存储在多相系数存储单元 1309 上的 28 字的多相滤波器系数  $C_{ppf}$  (表 5), 分别求得整数滞后  $int$  的自相关函数  $\phi_{int}(i)$ 、偏离整数滞后  $int-1/4$  的分数位置的自相关函数  $\phi_{dq}(i)$ 、偏离整数滞后  $int+1/4$  的分数位置的自相关函数  $\phi_{aq}(i)$ 、偏离整数滞后  $int+1/2$  的分数位置的自相关函数  $\phi_{ah}(i)$ 。

表 5: 多相滤波器系数  $C_{ppf}$ 

i	$C_{ppf}(i)$	i	$C_{ppf}(i)$	i	$C_{ppf}(i)$	i	$C_{ppf}(i)$
0	0.100035	7	0.000000	14	-0.128617	21	-0.212207
1	-0.180063	8	0.000000	15	0.300105	22	0.636620
2	0.900316	9	1.000000	16	0.900316	23	0.636620
3	0.300105	10	0.000000	17	-0.180063	24	-0.212207
4	-0.128617	11	0.000000	18	0.100035	25	0.127324
5	0.081847	12	0.000000	19	-0.069255	26	-0.090946
6	-0.060021	13	0.000000	20	0.052960	27	0.070736

此外, 分别对  $Lmin-2 \leq i \leq Lmax+2$  范围内的自变量  $i$ , 将  $\phi_{int}(i)$ 、 $\phi_{dq}(i)$ 、 $\phi_{aq}(i)$ 、 $\phi_{ah}(i)$  中最大的代入到  $\phi_{max}(i)$  中, 进行式 (7) 的处理, 求得  $Lmax-Lmin+1$  个的  $\phi_{max}(i)$ 。

$$\phi_{max}(i) = \text{MAX}(\phi_{int}(i), \phi_{dq}(i), \phi_{aq}(i), \phi_{ah}(i)) \quad (7)$$

$\phi_{max}(i)$ :  $\phi_{int}(i)$ 、 $\phi_{dq}(i)$ 、 $\phi_{aq}(i)$ 、 $\phi_{ah}(i)$  的最大值

I: 长期预测系数的分析区间 ( $Lmin \leq i \leq Lmax$ )

$Lmin$ : 长期预测系数的最短分析区间 (=16)

$Lmax$ : 长期预测系数的最长分析区间 (-128)

$\phi_{int}(i)$ : 预测残差信号整数滞后 ( $int$ ) 的自相关函数

$\phi_{dq}(i)$ : 预测残差信号分数滞后 ( $int-1/4$ ) 的自相关函数

$\phi_{aq}(i)$ : 预测残差信号分数滞后  $(int+1/4)$  的自相关函数

$\phi_{ah}(i)$ : 预测残差信号分数滞后  $(int+1/2)$  的自相关函数

从求得的  $(L_{max}-L_{min}+1)$  个的  $\phi_{max}(i)$  中, 由高端顺次地选出值大的 6 个, 保存作为音调候补  $psel(i)$  ( $0 \leq i \leq 5$ ), 并将线性预测残差信号  $res(i)$  和音调第 1 候补  $psel(0)$  输出到音调增强滤波器系数运算单元 1310, 将  $psel(i)$  ( $0 \leq i \leq 5$ ) 输出到自适应矢量生成单元 1319 中。

多相系数存储单元 1309, 存储音调预选单元 1308 用分数滞后精度求出线性预测残差信号的自相关函数时和自适应矢量生成单元 1319 用分数精度生成自适应矢量时参照的多相滤波器的系数。

音调增强滤波器系数运算单元 1310, 根据音调预选单元 1308 中求得的线性预测残差和  $res(i)$  和从音调第 1 候补  $psel(0)$ , 求 3 次音调预测系数  $cov$  ( $0 \leq i \leq 2$ )。借助使用求得的音调预测系数  $cov$  ( $0 \leq i \leq 2$ ) 的式 (8), 求音调增强滤波器  $Q(z)$  的脉冲响应, 并输出到频谱加权滤波器系数运算单元 1312 和听觉加权滤波器系数运算单元 1313 中。

$$Q(z) = 1 + \sum_{i=0}^{2} cov(i) \times \lambda p_i \times z^{-i} - psel(0) + i - 1 \quad (8)$$

$Q(z)$ : 音调增强滤波器的传递函数

$cov(i)$ : 音调预测系数 ( $0 \leq i \leq 2$ )

$\lambda p_i$ : 音调增强常数 ( $=0.4$ )

$psel(0)$ : 音调第 1 候补

LSP 内插单元 1311, 首先借助使用在 LSP 量化/解码单元 1306 中求得的当前处理帧的解码 LSP  $\omega_q(i)$  和以前求得并保持的前处理帧的解码 LSP  $\omega_{qp}(i)$  的式 (9), 对每个子帧, 求解码 mw 插 LSP  $\omega_{intp}(n, i)$  ( $1 \leq i \leq N_p$ )。

$$\omega_{intp}(n, i) = \begin{cases} 0.4 \times \omega_q(i) + 0.6 \times \omega_{qp}(i) & n = 1 \\ \omega_q(i) & n = 2 \end{cases} \quad (9)$$

$\omega_{intp}(n, i)$ : 第  $n$  子帧的内插 LSP

$n$ : 子帧号码 ( $=1, 2$ )

$\omega_q(i)$ : 处理帧的解码 LSP

$\omega_{qp}(i)$ : 前处理帧的解码 LSP

用将求得的  $\omega_{intp}(n, i)$  变换成 LPC, 求得解码内插 LPC  $a_q(n, i)$  ( $1 \leq i$

$\leq N_p$ ), 并将求得的解码内插 LPC  $\alpha_q(n, i)$  ( $1 \leq i \leq N_p$ ) 输出到频谱加权滤波器系数运算单元 1312 和听觉加权 LPC 合成滤波器系数运算单元 1314 中。

频谱加权滤波器系数运算单元 1312 构成式(10)的 MA 型频谱加权滤波器  $I(z)$ , 将其脉冲响应输出到听觉加权滤波器系数运算单元 1313 中。

$$I(z) = \sum_{i=1}^{N_{fir}} \alpha_{fir}(i) \times z^{-i} \quad (10)$$

$I(z)$ : MA 型频谱加权滤波器的传递函数

$N_{fir}$ :  $I(z)$  的滤波器次数 (=11)

$\alpha_{fir}(i)$ :  $I(z)$  的脉冲响应 ( $1 \leq i \leq N_{fir}$ )

其中, 式(10)的脉冲响应  $\alpha_{fir}(i)$  ( $1 \leq i \leq N_{fir}$ ) 是截短到  $N_{fir}$  (=11) 项为止的(11)供给的 ARMA 型频谱增强滤波器  $G(z)$  的脉冲响应。

$$G(z) = \frac{1 + \sum_{i=1}^{N_p} \alpha(n, i) \times \lambda_{ma}^i \times z^{-i}}{1 + \sum_{i=1}^{N_p} \alpha(n, i) \times \lambda_{ar}^i \times z^{-i}} \quad (11)$$

$G(z)$ : 频谱加权滤波器的传递函数

$n$ : 子帧号码 (=1, 2)

$N_p$ : LPC 分析次数 (=10)

$\alpha(n, i)$ : 第  $n$  子帧的解码内插 LSP

$\lambda_{ma}$ :  $G(z)$  的分子常数 (=0.9)

$\lambda_{ar}$ :  $G(z)$  的分母常数 (=0.4)

听觉加权滤波器系数运算单元 1313, 首先将叠加从频谱加权滤波器系数运算单元 1312 接收到的频谱加权滤波器  $I(z)$  的脉冲响应和从音调增强滤波器系数运算单元 1310 接收到的音调增强滤波器  $Q(z)$  的脉冲响应的结果作为脉冲响应, 构成听觉加权滤波器  $W(z)$ , 并将构成的听觉加权滤波器  $W(z)$  的脉冲响应输出到听觉加权 LPC 合成滤波器系数运算单元 1314 和听觉加权单元 1315 中。

听觉加权 LPC 合成滤波器系数运算单元 1314, 利用从 LSP 内插单元 1311 接收到的解码内插 LPC  $\alpha_q(n, i)$  和从听觉加权滤波器系数运算单元 1313 接收到的听觉加权滤波器  $W(z)$ , 由式(12)构成听觉加权 LPC 合成滤波器  $H(z)$ 。

$$H(z) = \frac{1}{1 + \sum_{i=1}^{N_p} \alpha_q(n, i) \times z^{-i}} W(z) \quad (12)$$

$H(z)$ : 听觉加权合成滤波器的传递函数

$N_p$ : LPC 分析次数

$\alpha_q(n, i)$ : 第  $n$  子帧的解码内插 LSP

$n$ : 子帧号码 (=1, 2)

$W(z)$ : 听觉加权滤波器(级联  $I(z)$  和  $Q(z)$  而成)的传递函数

将构成的听觉加权 LPC 合成滤波器  $H(z)$  的系数, 输出到目标生成单元

A1316、听觉加权 LPC 倒置合成单元 A1317、听觉加权 LPC 合成单元 A1321、听觉加权 LPC 倒置合成单元 B1326 和听觉加权 LPC 合成单元 B1329 中。

听觉加权单元 1315 将从缓冲器 1301 中读出的子帧信号输入到 0 状态的听觉加权 LPC 合成滤波器  $H(z)$  中, 并以其输出作为听觉加权残差  $spw(i)$  ( $0 \leq i \leq N_s-1$ ), 输出到目标生成单元 A1316 中。

目标生成单元 A1316 从听觉加权单元 1315 中求得的听觉加权残差  $spw(i)$  ( $0 \leq i \leq N_s-1$ ), 减去作为在听觉加权 LPC 合成滤波器系数运算单元 1314 中求得的听觉加权 LPC 合成滤波器  $H(z)$  中输入 0 系列时的输出的 0 输入响应  $Zres(i)$  ( $0 \leq i \leq N_s-1$ ) 后, 所得结果输出到 LPC 倒置合成单元 A1317 和目标生成单元 B1325 中, 作为声源选择用的目标向量  $r(i)$  ( $0 \leq i \leq N_s-1$ )。

听觉加权 LPC 倒置合成单元 A1317 时间反转地将目标生成单元 1316 接收到的目标系列  $r(i)$  ( $0 \leq i \leq N_s-1$ ) 变换排列, 并将变换得到的向量输入到初始状态为 0 的听觉加权 LPC 合成滤波器  $H(z)$  中, 将其输出再次时间反转变换排列, 从而得到目标系列的时间反转合成向量  $rh(k)$  ( $0 \leq k \leq N_s-1$ ), 并输出到比较单元 A1322 中。

自适应码本 1318 存储自适应矢量生成单元 1319 生成自适应矢量时参照的过去的驱动声源。自适应矢量生成单元 1319 在生成从音调预选单元 1308 接收到的 6 个音调候补  $psel(j)$  ( $0 \leq j \leq 5$ ) 的同时, 生成  $N_{ac}$  个自适应矢量  $P_{acb}(i, k)$  ( $0 \leq i \leq N_s-1, 0 \leq k \leq N_s-1, 6 \leq N_{ac} \leq 24$ ), 并输出到自适应/固定选择单元 1320 中。具体地说, 如表 6 所示, 在  $16 \leq psel(j) \leq 44$  的场合, 对于相当于一个整数滞后位置的 4 种分数滞后位置, 生成自适应矢量, 在  $45 \leq psel(j) \leq 64$  的场合, 对于

相当于一个整数滞后位置的 2 种分数滞后位置,生成自适应矢量,在  $65 \leq \text{psel}(j) \leq 128$  的场合,对整数滞后位置,生成自适应矢量。由此,根据  $\text{psel}(j)$  ( $0 \leq j \leq 5$ ) 的值,自适应矢量的候补数  $N_{ac}$  最少为 6,最多为 24。

表 6: 自适应矢量和固定矢量的总数

总矢量数	255 个
自适应矢量数	222 个
$16 \leq \text{psel}(i) \leq 44$	116 个 (29 个 $\times$ 分数滞后 4 种)
$45 \leq \text{psel}(i) \leq 64$	142 个 (21 个 $\times$ 分数滞后 2 种)
$65 \leq \text{psel}(i) \leq 128$	64 个 (64 个 $\times$ 分数滞后 1 种)
固定矢量数	32 个 (16 $\times$ 符号 2 种)

此外,生成分数精度的自适应矢量时,利用在以整数精度从自适应码本 1318 读出的过去的声源矢量中,叠加存储在多相系数存储单元 1309 中的多相滤波器系数的内插处理进行。

这里,对应于  $\text{lagf}(i)$  的值的内插,是进行在  $\text{lagf}(i)=0$  的场合对应于整数滞后位置、在  $\text{lagf}(i)=1$  的场合对应于从整数滞后位置偏离  $-1/2$  的分数滞后位置、在  $\text{lagf}(i)=2$  的场合对应于从整数滞后位置偏离  $+1/4$  的分数滞后位置、在  $\text{lagf}(i)=3$  的场合对应于从整数滞后位置偏离  $-1/4$  的分数滞后位置的内插。

自适应/固定选择单元 1320,首先接受自适应矢量生成单元 1319 生成的  $N_{ac}(6 \sim 24)$  一个候补的自适应矢量,并输出到听觉加权 LPC 合成单元 A1321 和比较单元 A1322 中。

比较单元 A1322,首先为了适应矢量生成单元 1319 生成的自适应矢量  $\text{Pacb}(i, k)$  ( $0 \leq i \leq N_s-1, 0 \leq k \leq N_s-1, 6 \leq N_{ac} \leq 24$ ) 从  $N_{ac}(6 \sim 20)$  个候补中预先  $N_{ac}(=4)$  个候补,利用式 (13) 求得由听觉加权 LPC 倒置合成单元 1317 受理的目标矢量的时间反转合成矢量  $\text{rh}(k)$  ( $0 \leq k \leq N_s-1$ ) 和自适应矢量  $\text{Pacb}(i, k)$  的内积  $\text{prac}(i)$ 。



$$\text{prac}(i) = \sum_{k=0}^{N_s-1} \text{Pacb}(i, k) \times \text{rh}(k) \quad (13)$$

$\text{prac}(i)$ : 自适应矢量预选基准值

$N_{ac}(i)$ : 预选后自适应矢量候补数 (=6~24)

$i$ : 自适应矢量的号码 ( $0 \leq i \leq N_{ac}-1$ )

$\text{Pacb}(i, k)$ : 自适应矢量

$\text{rh}(k)$ : 目标矢量  $r(k)$  的时间反转合成矢量

比较求得的内积  $\text{prac}(i)$ ，选择其值变大时的标号和以该标号作为引数时的内积（直到高端第  $N_{acb}(=4)$  个为止，并分别作为自适应矢量预选后标号  $\text{apsel}(j)$  ( $0 \leq j \leq N_{acb}-1$ ) 和自适应矢量预选后基准值  $\text{prac}(\text{apsel}(j))$  进行保存，而且将自适应矢量预选后标号  $\text{apsel}(j)$  ( $0 \leq j \leq N_{acb}-1$ ) 输出到自适应/固定选择单元 1320 中。

听觉加权 LPC 合成单元 A1321 对通过在自适应矢量生成单元 1319 中生成的自适应/固定选择单元 1320 的预选后自适应矢量  $\text{Pacb}(\text{apsel}(j), k)$ ，施行听觉加权 LPC 合成，生成合成自适应矢量  $\text{SYNacb}(\text{apsel}(j), k)$ ，并输出到比较单元 A1322 中。接着，比较单元 A1322 为了对其自身已预选的  $N_{acb}(=4)$  个的预选后适应矢量  $\text{Pacb}(\text{apsel}(j), k)$  进行正式选择，由式(14)求出自适应矢量正式选择基准值  $\text{sacbr}(j)$ 。

$$\text{sacbr}(j) = \frac{\text{prac}^2(\text{apsel}(j))}{\sum_{k=0}^{N_s-1} \text{SYNacb}^2(j, k)} \quad (14)$$

$\text{sacbr}(j)$ : 自适应矢量正大选择基准值

$\text{prac}()$ : 自适应矢量预选后基准值

$\text{apsel}(j)$ : 自适应矢量预选标号

$k$ : 矢量次数 ( $0 \leq k \leq N_s-1$ )

$j$ : 被预选的自适应矢量的标号的号码 ( $0 \leq j \leq N_{acb}-1$ )

$N_s$ : 子帧长 (=52)

$N_{acb}$ : 自适应矢量的预选数

$\text{SYNacb}(J, K)$ : 合成自适应矢量

分别用式(14)的值增大时的标号和以该标号作为引数时的式(14)的值，作为

自适应矢量正大选择后标号 ASEL 和自适应矢量正式选择后基准值 sacbr(ASEL)，并输出到自适应/固定选择单元 1320 中。

固定码本 1323 对固定矢量读出单元 1324 读出的矢量存储 Nfc(=16) 个候补。这里，比较单元 A1322 为了对从固定矢量读出单元 1324 读出的固定矢量 Pfc(i, k) ( $0 \leq i \leq Nfc-1$ ,  $0 \leq k \leq Ns-1$ )，从 Nfc(=16) 个候补中预选 Nfcb(=2) 个候补、利用式(15)求出由听觉加权 LPC 倒置合成单元 A1317 受理的目标矢量的时间反转合成矢量 rh(k) ( $0 \leq k \leq Ns-1$ ) 和固定矢量 Pfc(i, k) 的内积的绝对值 |prfc(i)|。

$$|\text{prfc}(i)| = \sum_{k=0}^{Ns-1} \text{Pfc}(i, k) \times \text{rh}(k) \quad (15)$$

|prfc(i)|: 固定矢量预选基准值

k: 矢量的要素号码 ( $0 \leq k \leq Ns-1$ )

I: 固定矢量的号码 ( $0 \leq i \leq Nfc-1$ )

Nfc: 固定矢量数 (=16)

Pfc(i, k): 固定矢量

rh(k): 目标矢量 r(k) 的时间反转合成矢量

比较式(15)的值 |prac(i)|，选择其值变大时的标号和以该标号作为引数时的内积的绝对值（直到高端第 Nfcb(=2) 为止），并分别作为固定矢量预选后标号 fpsel(j) ( $0 \leq j \leq Nfcb-1$ ) 和固定矢量预选后基准值 |prfc(fpsel(j))| 进行保存，而且将固定矢量预选后标号 fpsel(j) ( $0 \leq j \leq Nfcb-1$ ) 输出到自适应/固定选择单元 1320 中。

听觉加权 LPC 合成单元 A1321，对通过在固定矢量读出单元 1324 中读出的自适应/固定选择单元 1320 的预选后固定矢量 Pfc(fpsel(j), k)，施行听觉加权 LPC 合成，生成合成固定矢量 SYNfcb(fpsel(j), k)，并输出到比较单元 A1322 中。

接着，比较单元 A1322 为了从其自身预选的 Nfcb(=2) 个的预选后固定矢量 Pfc(fpsel(j), k) 中正式选择最佳固定矢量，由式(16)求出固定矢量正式选择基准值 sfcb(j)。

$$\text{sfcbr}(j) = \frac{|\text{prfc}(\text{fpsel}(j))|^2}{\sum_{k=0}^{N_s-1} \text{SYNfcb}^2(j, k)} \quad (16)$$

$\text{sfcbr}(j)$ : 固定矢量正式选择基准值

$|\text{prfc}()|$ : 固定矢量预选后基准值

$\text{fpsel}(j)$ : 固定矢量预选后标号 ( $0 \leq j \leq N_{\text{fcb}}-1$ )

$k$ : 矢量的要素号码 ( $0 \leq k \leq N_s-1$ )

$j$ : 被预选的固定矢量的号码 ( $0 \leq j \leq N_{\text{fcb}}-1$ )

$N_s$ : 子帧长 (=52)

$N_{\text{acb}}$ : 固定矢量的预选数 (=2)

$\text{SYNacb}(J, K)$ : 合成固定矢量

分别用式(16)的值增大时的标号和以该标号作为引数时的式(16)的值, 作为固定矢量正式选择后标号  $\text{FSEL}$  和固定矢量正式选择后基准值  $\text{facbr}(\text{FSEL})$ , 并输出到自适应/固定选择单元 1320 中。

自适应/固定选择单元 1320 利用从比较单元 A1322 收到的  $\text{prac}(\text{ASEL})$ 、 $\text{sacbr}(\text{ASEL})$ 、 $|\text{prfc}(\text{FSEL})|$  和  $\text{sfcbr}(\text{FSEL})$  的大小和正负关系(记载在式(17)中), 选择正式选择后自适应矢量或正式选择后固定矢量, 作为自适应/固定矢量  $\text{AF}(k)$  ( $0 \leq k \leq N_s-1$ )。

$$\text{AF}(k) = \begin{cases} \text{Pacb}(\text{ASEL}, k) & \text{sacbr}(\text{ASEL}) \geq \text{sfcbr}(\text{FSEL}), \text{prac}(\text{ASEL}) > 0 \\ 0 & \text{sacbr}(\text{ASEL}) \geq \text{sfcbr}(\text{FSEL}), \text{prac}(\text{ASEL}) \leq 0 \\ \text{Pfcb}(\text{FSEL}, k) & \text{sacbr}(\text{ASEL}) < \text{sfcbr}(\text{FSEL}), \text{prfc}(\text{FSEL}) \geq 0 \\ -\text{Pfcb}(\text{FSEL}, k) & \text{sacbr}(\text{ASEL}) < \text{sfcbr}(\text{FSEL}), \text{prfc}(\text{FSEL}) < 0 \end{cases} \quad (17)$$

$\text{AF}(k)$ : 自适应/固定矢量

$\text{ASEL}$ : 自适应矢量正式选择后标号

$\text{FSEL}$ : 固定矢量正式选择后标号

$k$ : 矢量的要素号码

$\text{Pacb}(\text{ASEL}, k)$ : 正式选择后自适应矢量

$\text{Pfcb}(\text{FSEL}, k)$ : 正式选择后固定矢量

$\text{sacbr}(\text{ASEL})$ : 自适应矢量正式选择后基准值

$\text{sfcbr}(\text{FSEL})$ : 固定矢量正式选择后基准值

prac(ASEL): 自适应矢量预选后基准值

prfc(FSEL): 固定矢量预选后基准值

将选择的自适应/固定矢量 AF(k) 输出到听觉加权 LPC 合成滤波器单元 A1321 中, 将表示生成选择的自适应/固定矢量 AF(k) 的号码的标号作为自适应/固定标号 AFSEL 输出到参数编码单元 1331 中。此外, 这里因设计成自适应矢量和固定矢量的总矢量数为 255 个(参照表 6), 所以自适应/固定标号 AFSEL 为 8 位编码。

听觉加权 LPC 合成滤波器单元 A1321 对在自适应/固定选择单元 1320 中选择的自适应/固定矢量 AF(k), 施行听觉加权 LPC 合成滤波, 生成合成自适应/固定矢量 SYNaf(k) ( $0 \leq k \leq N_s - 1$ ), 并输出到比较单元 1322 中。

比较单元 1322, 首先利用式(18)求出从听觉加权 LPC 合成滤波器单元 A1321 收到的合成自适应/固定矢量 SYNaf(k) ( $0 \leq k \leq N_s - 1$ ) 的功率 powp。

$$\text{powp} = \sum_{k=0}^{N_s-1} \text{SYNaf}^2(k) \quad (18)$$

powp: 自适应/固定矢量(SYNaf(k))的功率

k: 矢量的要素号码( $0 \leq k \leq N_s - 1$ )

Ns: 子帧长(=52)

SYNaf(k): 自适应/固定矢量

接着, 由式(19)求出从目标生成单元 A1316 收到的目标矢量和合成自适应/固定矢量 SYNaf(k) 的内积 pr。

$$\text{pr} = \sum_{k=0}^{N_s-1} \text{SYNaf}(k) \times r(k) \quad (19)$$

pr: SYNaf(k) 和 r(k) 的内积

Ns: 子帧长(=52)

SYNaf(k): 自适应/固定矢量

r(k): 目标矢量

k: 矢量的要素号码( $0 \leq k \leq N_s - 1$ )

进而, 将由从自适应/固定选择单元 1320 收到的自适应/固定矢量 AF(k) 输出到自适应码本更新单元 1333 中, 计算 AF(k) 的功率 POWaf, 将合成自适应/固定矢量 SYNaf(k) 和 POWaf 输出到参数编码单元 1331 中, 并将 powp 和 pr 以及 rh(k) 输出到比较单元 B1330 中。

目标生成单元 B1325, 从目标生成单元 A1316 收到的声源选择用的目标矢量  $r(i)$  ( $0 \leq k \leq N_s-1$ ) 减去从比较单元 A1322 收到的合成自适应/固定矢量  $SYNaf(k)$  ( $0 \leq k \leq N_s-1$ ), 生成新的目标矢量, 并将生成的新的目标矢量输出到听觉加权 LPC 倒置合成单元 B1326 中。

听觉加权 LPC 倒置合成单元 B1326 对目标生成单元 B1325 中生成的新的目标矢量, 进行时间反转排列变换, 并将该变换后的矢量输入到 0 状态的听觉加权 LPC 合成滤波器中, 再次对该输出向量进行时间反转排列变换, 生成新的目标矢量的时间反转合成矢量  $ph(k)$  ( $0 \leq k \leq N_s-1$ ), 并输出到比较单元 B1330 中。

声源矢量生成装置 1337 使用与例如实施形态 3 中说明了的声源矢量生成装置 70 相同的装置。声源矢量生成装置 70 从振种存储单元 71 读出第 1 个振种, 输入到非线性数字滤波器 72 中, 并生成噪声矢量。将在声源矢量生成装置 70 生成的噪声矢量输出到听觉加权 LPC 合成单元 B1329 和比较单元 B1330 中。接着, 从振种存储单元 71 输入到读出第 2 个振种, 输入到非线性数字滤波器 72 中, 生成噪声矢量, 并输出到听觉加权 LPC 合成单元 B1329 和比较单元 B1330 中。

较单元 B1330 为了对根据第 1 振种生成的噪声矢量, 从  $N_{st}(=64)$  个候补中预选  $N_{stb}(=6)$  个候补, 由式(20)求出第 1 噪声矢量预选基准值  $cr(i1)$  ( $0 \leq i1 \leq N_{stb1}-1$ )。

$$cr(i1) = \sum_{j=0}^{N_s-1} P_{stb1}(i1j) \times rh(j) - \frac{pr}{powp} \sum_{j=0}^{N_s-1} P_{stb1}(i1j) \times ph(j) \quad (20)$$

$cr(i1)$ : 第 1 噪声矢量预选基准值

$N_s$ : 子帧长(=52)

$rh(j)$ : 目标矢量( $rh(j)$ )的时间反转合成矢量

$powp$ : 自适应/固定矢量( $SYNaf(k)$ )的功率

$pr$ :  $SYNaf(k)$ 和  $r(k)$ 的内积

$P_{stb1}(i1, j)$ : 第 1 噪声矢量

$ph(j)$ :  $SYNaf(k)$ 的时间反转合成矢量

$i1$ : 第 1 噪声矢量的号码 ( $0 \leq i1 \leq N_{stb1}-1$ )

$j$ : 矢量的要素号码

比较求得的  $cr(i1)d1$  值, 选择其值变大时的标号和以该标号作为引数时的

式(20)的值(直到高端第  $N_{stb}(=6)$  个为止), 分别作为第 1 噪声矢量预选后标号  $s1p_{sel}(j1)$  ( $0 \leq j1 \leq N_{stb}-1$ ) 和预选后第 1 噪声矢量  $P_{stb1}(s1p_{sel}(j1), k)$  ( $0 \leq j1 \leq N_{stb}-1, 0 \leq k \leq N_s-1$ ) 进行保存。接着, 对于第 2 噪声矢量也进行与第 1 噪声矢量相同的处理, 分别作为第 2 噪声矢量预选后标号  $s2p_{sel}(j2)$  ( $0 \leq j2 \leq N_{stb}-1$ ) 和预选后第 2 噪声矢量  $P_{stb2}(s2p_{sel}(j2), k)$  ( $0 \leq j2 \leq N_{stb}-1, 0 \leq k \leq N_s-1$ ) 进行保存。

听觉加权 LPC 合成单元 B1329, 对预选后第 1 噪声矢量  $P_{stb1}(s1p_{sel}(j1), k)$ , 施行听觉加权 LPC 合成, 生成合成第 1 噪声矢量  $SYN_{stb1}(s1p_{sel}(j1), k)$ , 并输出到比较单元 B1330 中。接着, 对预选后第 2 噪声矢量  $P_{stb2}(s2p_{sel}(j2), k)$ , 施行听觉加权 LPC 合成, 生成合成第 2 噪声矢量  $SYN_{stb2}(s2p_{sel}(j2), k)$ , 并输出到比较单元 B1330 中。

比较单元 B1330 为了对其自身预选的预选后第 1 噪声矢量和预选后第 2 噪声矢量进行正式选择, 对在听觉加权 LPC 合成单元 B1329 中计算的合成第 1 噪声矢量  $SYN_{stb1}(s1p_{sel}(j1), k)$ , 进行式(21)的计算。

$$SYN_{ostb1}(s1p_{sel}(j1), k) = SYN_{stb1}(s1p_{sel}(j1), k) - \frac{SYN_{af}(j1)}{powp} \sum_{k=0}^{N_s-1} P_{stb1}(s1p_{sel}(j1), k) \times ph(k) \quad (21)$$

$SYN_{ostb1}(s1p_{sel}(j1), k) = (21)$

$SYN_{ostb1}(s1p_{sel}(j1), k)$ : 正交化合成第 1 噪声矢量

$SYN_{stb1}(s1p_{sel}(j1), k)$ : 合成第 1 噪声矢量

$P_{stb1}(s1p_{sel}(j1), k)$ : 预选后第 1 噪声矢量

$SYN_{af}(j)$ : 自适应/固定矢量

$powp$ : 自适应/固定矢量( $SYN_{af}(j)$ )的功率

$N_s$ : 子帧长( $=52$ )

$ph(k)$ :  $SYN_{af}(j)$ 的时间反转合成矢量

$j1$ : 预选后第 1 噪声矢量的号码

$k$ : 矢量的要素号码( $0 \leq k \leq N_s-1$ )

求出正交化合成第 1 噪声矢量  $SYN_{ostb1}(s1p_{sel}(j1), k)$  后, 对合成第 2 噪声矢量  $SYN_{ostb2}(s2p_{sel}(j2), k)$  也进行同样的计算, 求出正交化合成第 2 噪声

矢量  $\text{SYNOstb2}(s2\text{psel}(j2), k)$ ，并分别用式(22)和式(23)，对(( $s1\text{psel}(j1)$ ， $s2\text{psel}(j2)$ ))的全部组合(36项)，以闭环方式计算第1噪声矢量本选择基准值  $\text{scr1}$  和第2噪声矢量本选择基准值  $\text{scr2}$ 。

$$\text{scr1} = \frac{\text{cscr1}^2}{\sum_{k=0}^{N_s-1} [\text{SYNOstb1}(s1\text{psel}(j1), k) + \text{SYNOstb2}(s2\text{psel}(j2), k)]^2} \quad (22)$$

$\text{scr1}$ : 第1噪声矢量本选择基准值

$c \text{scr1}$ : 由式(24)事先计算的常数

$\text{SYNOstb1}(s1\text{psel}(j1), k)$ : 正交合成第1噪声矢量

$\text{SYNOstb2}(s2\text{psel}(j2), k)$ : 正交合成第2噪声矢量

$r(k)$ : 目标矢量

$s1\text{psel}(j1), k$ : 第1噪声矢量预选后标号

$s2\text{psel}(j2), k$ : 第2噪声矢量预选后标号

$N_s$ : 子帧长(=52)

$k$ : 矢量的要素号码

$$\text{scr2} = \frac{\text{cscr2}^2}{\sum_{k=0}^{N_s-1} [\text{SYNOstb1}(s1\text{psel}(j1), k) - \text{SYNOstb2}(s2\text{psel}(j2), k)]^2} \quad (23)$$

$\text{scr2}$ : 第2噪声矢量本选择基准值

$c \text{scr2}$ : 由式(25)事先计算的常数

$\text{SYNOstb1}(s1\text{psel}(j1), k)$ : 正交合成第1噪声矢量

$\text{SYNOstb2}(s2\text{psel}(j2), k)$ : 正交合成第2噪声矢量

$r(k)$ : 目标矢量

$s1\text{psel}(j1), k$ : 第1噪声矢量预选后索引

$s2\text{psel}(j2), k$ : 第2噪声矢量预选后标号

$N_s$ : 子帧长(=52)

$k$ : 矢量的要素号码

其中，式(22)中的  $\text{cscr1}$  和式(23) 中的  $\text{cscr2}$ ，分别是由式(24)和式(25)预先计算的常数

$$cscrl = \sum_{k=0}^{Ns-1} \text{SYNOstb1}(s1psel(j1), k) \times r(k) + \sum_{k=0}^{Ns-1} \text{SYNOstb2}(s2psel(j2), k) \times r(k) \quad (24)$$

cscrl: 式(22)用常数

SYNOstb1(s1psel(j1), k): 正交合成第1噪声矢量

SYNOstb2(s2psel(j2), k): 正交合成第2噪声矢量

r(k): 目标矢量

s1psel(j1), k: 第1噪声矢量预选后标号

s2psel(j2), k: 第2噪声矢量预选后标号

Ns: 子帧长(=52)

k: 矢量的要素号码

$$cscrl = \sum_{k=0}^{Ns-1} \text{SYNOstb1}(s1psel(j1), k) \times r(k) - \sum_{k=0}^{Ns-1} \text{SYNOstb2}(s2psel(j2), k) \times r(k) \quad (25)$$

cscrl: 式(23)用常数

SYNOstb1(s1psel(j1), k): 正交合成第1噪声矢量

SYNOstb2(s2psel(j2), k): 正交合成第2噪声矢量

r(k): 目标矢量

s1psel(j1), k: 第1噪声矢量预选后标号

s2psel(j2), k: 第2噪声矢量预选后标号

Ns: 子帧长(=52)

k: 矢量的要素号码

比较单元 B1330 进一步将 s1cr 的最大值代入到 MAXs1cr 中、将 s2cr 的最大值代入到 MAXs2cr 中，并用 MAXs1cr 和 MAXs2cr 中大的一个作为 scr，将求得 scr 时参照的 s1psel(j1) 的值作为第 1 噪声矢量正式选择后标号 SSEL1，输出到参数编码单元 1331 中。保存对应于 SSEL1 的噪声矢量作为正式选择后第 1 噪声矢量 Pstb1(SSEL1, k)，求出对应于 Pstb1(SSEL1, k) 的本选择后合成第 1 噪声矢量 SYNstb1(SSEL1, k) (0 ≤ k ≤ Ns-1)，并输出到参数编码单元 1331 中。

同样，将求得 scr 时参照的 s2psel(j2) 的值作为第 2 噪声矢量正式选择后标号 SSEL2 输出到参数编码单元 1331 中，而且保存对应于 SSEL2 的噪声矢量作为正式选择后第 2 噪声矢量 Pstb2(SSEL2, k)，求出对应于 Pstb2(SSEL2, k) 的正



式选择后合成第 2 噪声矢量  $SYNstb2(SSEL2, k)$  ( $0 \leq k \leq N_s - 1$ ), 并输出到参数编码单元 1331 中。

比较单元 B1330 进一步求出分别乘以  $Pstb1(SSEL1, k)$  和  $Pstb2(SSEL2, k)$  的符号  $S1$  和  $S2$ , 并以求得的  $S1$  和  $S2$  的正负信息作为增益正负标号  $Is1s2$  (2 位信息), 输出到参数编码单元 1331 中。

$$(S1, S2) = \begin{cases} (+1, +1) & scr1 \geq scr2, cscr1 \geq 0 \\ (-1, -1) & scr1 \geq scr2, cscr1 < 0 \\ (+1, -1) & scr1 < scr2, cscr2 \geq 0 \\ (-1, +1) & scr1 < scr2, cscr2 < 0 \end{cases} \quad (26)$$

$S1$ : 正式选择后第 1 噪声矢量的符号

$S2$ : 正式选择后第 2 噪声矢量的符号

$scr1$ : 式(22)的输出

$scr2$ : 式(23)的输出

$cscr1$ : 式(24)的输出

$cscr2$ : 式(25)的输出

在根据式(27)生成噪声矢量  $ST(k)$  ( $0 \leq k \leq N_s - 1$ ), 并输出到自适应码本更新单元 1333 中的同时, 求出其功率  $POWsf$ , 并输出到参数编码单元 1331 中。

$$ST(k) = S1 \times Pstb1(SSEL1, k) \div S2 \times Pstb2(SSEL2, k) \quad (27)$$

$ST(k)$ : 随机矢量

$S1$ : 正式选择后第 1 噪声矢量的符号

$S2$ : 正式选择后第 2 噪声矢量的符号

$Pstb1(SSEL1, k)$ : 正式选择后第 1 级确定的矢量

$Pstb2(SSEL2, k)$ : 正式选择后第 2 级确定的矢量

$SSEL1$ : 第 1 噪声矢量正式选择后标号

$SSEL2$ : 第 2 噪声矢量正式选择后标号

$k$ : 矢量的要素号码 ( $0 \leq k \leq N_s - 1$ )

根据式(28)生成合成噪声矢量  $SYNst(k)$  ( $0 \leq k \leq N_s - 1$ ), 并输出到参数编码单元 1331 中。

$$SYNst(k) = S1 \times SYNstb1(SSEL1, k) \div S2 \times SYNstb2(SSEL2, k) \quad (28)$$

$SYNst(k)$ : 合成随机的矢量

S1: 正式选择后第 1 噪声矢量的符号

S2: 正式选择后第 2 噪声矢量的符号

SYNstb1(SSEL1, k): 正式选择后合成第 1 噪声矢量

SYNstb2(SSEL2, k): 正式选择后合成第 2 噪声矢量

k: 矢量的要素号码 ( $0 \leq k \leq N_s - 1$ )

参数编码单元 1331, 首先根据利用在帧功率量化/解码单元 1302 中求得的解码帧功率  $spow$ 、以及音调预选单元 1308 中求得的归一化预测残差功率  $resid$  的式 (29), 求出子帧推定残差功率  $rs$ 。

$$rs = N_s \times spow \times resid \quad (29)$$

$rs$ : 子帧推定残差功率

$N_s$ : 子帧长 (=52)

$spow$ : 解码帧功率

$resid$ : 归一化预测残差功率

使用求得的子帧推定残差功率  $rs$ 、比较单元 A1322 中计算的自适应/固定矢量的功率  $P0Waf$ 、比较单元 B1330 中求得的噪声矢量的功率  $P0Wst$ 、表 7 所示的增益量化表存储单元 1332 中存储的 256 字的增益量化用表 ( $CGaf[i]$ 、 $CGst[i]$ ) ( $0 \leq i \leq 127$ ) 等, 根据式 (30) 求出量化增益选择基准值  $STDg$ 。

表 7: 增益量化用表

i	CGaf(i)	CGst(i)
1	0.38590	0.23477
2	0.42380	0.50453
3	0.23416	0.24761
126	0.35382	1.68987
127	0.10689	1.02035
128	3.09711	1.75430

$$STDg = \sum_{k=0}^{Ns-1} \left( \sqrt{\frac{rs}{POWaf}} \cdot CGaf(Ig) \times SYNaf(k) + \sqrt{\frac{rs}{POWst}} \cdot CGst(Ig) \times SYNst(k) - r(k) \right)^2 \quad (30)$$

STDg: 量化增益选择基准值

rs: 子帧推定残差功率

POWaf: 自适应/固定矢量的功率

POWst: 噪声矢量的功率

i: 增益量化表的标号 (  $0 \leq i \leq 127$  )

CGaf(i): 增益量化表中自适应/固定矢量栏的组成部分

CGst(i): 增益量化表中噪声矢量栏的组成部分

SYNaf(k): 合成自适应/固定矢量

SYNst(k): 合成噪声矢量

r(k): 目标矢量

Ns: 子帧长 (=52)

k: 矢量的要素号码 (  $0 \leq k \leq Ns-1$  )

借助使用选择 1 个求得的量化增益选择基准值 STDg 为最小时的标号，作为增益量化标号 Ig，以选择的增益量化标号 Ig 为基础从增益量化用表读出的自适应/固定矢量栏的选择后增益 CGaf(Ig)，以及以选择的增益量化标号 Ig 为基础从增益量化用表读出的噪声矢量侧选择后增益 CGst(Ig) 等的式(31)，求出在 AF(k) 中实际用的自适应/固定矢量方面的正式增益 Gaf 和在 ST(k) 中实际用的噪声矢量方面的正式增益 Gst，并输出到自适应码本更新单元 1333 中。

$$(Gaf, Gst) = \left( \sqrt{\frac{rs}{POWaf}} CGaf(Ig), \sqrt{\frac{rs}{POWst}} CGst(IG) \right) \quad (31)$$

Gaf: 自适应/固定矢量侧本增益

Gst: 噪声矢量侧本增益

rs: rs: 子帧推定残差功率

POWaf: 自适应/固定矢量的功率

POWst: 噪声矢量的功率

CGaf(Ig): 固定/适应矢量方面的功率

CGst(Ig): 噪声矢量方面的功率

Ig: 增益量化标号

参数编码单元 1331 收集在帧功率量化和解码单元 1302 中求得的功率标号 Ipow、在 LSP 量化和解码单元 1306 中求得的 LSP 码 I1sp、在自适应/固定选择单元 1320 中求得的自适应/固定标号 AFSEL、在比较单元 B1330 中求得的第 1 噪声矢量正式选择后标号 SSEL1 和第 2 噪声矢量正式选择后标号 SSEL2 以及增益正负标号 Is1s2、在参数编码单元 1331 自身中求得的增益量化标号 Ig, 作为声音码, 并将收集到的声音码输出到传送单元 1334 中。

自适应码本更新单元 1333, 进行对比较单元 A1322 中求得的自适应/固定矢量 AF(k) 和比较单元 B1330 中求得的噪声矢量 ST(k) 分别乘以用参数编码单元 1331 求得的自适应/固定矢量正式增益 Gaf 和噪声矢量正式噪声 Gst 后进行相加的式(32)的处理, 生成驱动声源 ex(k) (0 ≤ k ≤ Ns-1), 并将生成的驱动声源 ex(k) (0 ≤ k ≤ Ns-1) 输出到自适应码本 1318 中。

$$ex(k) = Gaf \times AF(k) + Gst * ST(k) \quad (32)$$

ex(k): 驱动声源

AF(k): thd 适应固定矢量

ST(k): 噪声矢量的增益

k: 矢量的要素号码 (0 ≤ k ≤ Ns-1)

这时, 用由 thd 适应码本更新单元 1333 收到的新驱动声源 ex(k), 更新自适应码本 1318 内旧的驱动声源。

#### 实施形态 8

下面, 对在以作为数字便携电话的声音编码/解码标准方式的 PSI-CELP 开发的声源解码装置中, 用前述实施形态 1~实施形态 6 说明了的声源矢量生成装置的实施形态进行说明。这种解码装置是与前述的实施形态 7 配对的装置。

图 14 表示与实施形态 8 相关的声源解码装置的功能方框图。参数解码单元 1402 通过传送单元 1401 获得从图 13 所述的 CELP 型声音编码装置送来的声音编码(功率标号 Ipow、LSP 码 I1sp、自适应/固定标号 AFSEL、第 1 噪声矢量正式选择后标号 SSEL1、第 2 噪声矢量正式选择后标号 SSEL2、增益量化标号 Ig、增益正负标号 Is1s2)。

接着, 从存储在功率量化表存储单元 1405 中的功率量化用表(参照表 3)读出

功率标号  $I_{pow}$  所示的标量值，并作为解码帧功率  $spow$  输出到功率复原单元 1417 中，从存储在 LSP 量化表存储单元 1404 中的 LSP 量化用表读出 LSP 编码  $I_{lsp}$  的所示的矢量，并作为解码 LSP 输出到 LSP 内插单元 1406 中。将自适应/固定标号 AFSEL 输出到自适应矢量生成单元 1408 固定矢量读出单元 1411 以及自适应/固定选择单元 1412 中，将第 1 噪声矢量正式选择后标号 SSEL1 和第 2 噪声矢量正式选择后标号 SSEL2 输出到声源矢量生成装置 1414 中。从存储在增益量化表存储单元 1403 中的增益量化用表(参照表 7)读出增益量化索引  $I_g$ 。所示的矢量 ( $CAaf(I_g)$ ,  $CGst(I_g)$ )，与编码装置侧相同，根据式(31)求出在  $AF(k)$  中实际用的自适应/固定矢量正式增益  $Gaf$  和在  $ST(k)$  中实际用的噪声矢量正式增益  $Gst$ ，并将求得的自适应/固定矢量正式增益  $Gaf$  和噪声矢量正式增益  $Gst$  与增益正负标号  $Is1s2$  一起输出到驱动声源生成单元 1413 中。

LSP 内插单元 1406 用与编码装置相同的方法，根据从参数编码单元 1402 收到的解码 LSP 对每一子帧求出解码内插 LSP  $\omega_{intp}(n, i)$  ( $0 \leq i \leq N_p$ )，用求得的 LSP  $\omega_{intp}(n, i)$  变换成 LPC，从而得到解码内插 LPC，并将得到的解码内插 LPC 输出到 LPC 合成滤波器单元 1413 中。

自适应矢量生成单元 1408 根据从参数解码单元 1402 收到的自适应/固定标号 AFSEL，在从自适应码本 1407 读出的矢量上叠加存储在多相系数存储单元 1409 中的多相系数(参照表 5)的一部分，生成分数滞后精度的自适应矢量，并输出到自适应/固定选择单元 1412 中。固定矢量读出单元 1411 根据从参数解码单元 1402 收到的自适应/固定标号 AFSEL，从固定码本 1410 读出固定矢量，并输出到自适应/固定选择单元 1412 中。

自适应/固定选择单元 1412 根据从参数解码单元 1402 收到的自适应/固定标号 AFSEL，选择从自适应矢量生成单元 1408 输入的自适应矢量或从固定矢量读出单元 1411 输入的固定矢量作为自适应/固定矢量  $AF(k)$ ，并将被选择的自适应/固定矢量  $AF(k)$  输出到驱动声源生成单元 1413 中。声源矢量生成装置 1414 根据从由参数解码单元 1402 收到的第 1 噪声矢量正式选择后标号 SSEL1 和第 2 噪声矢量正式选择后标号 SSEL2，从振种存储单元 71 取出第 1 振种和第 2 振种，输入到非线性数字滤波器 72 中，分别发生第 1 噪声矢量和第 2 噪声矢量。这样，在重现的第 1 噪声矢量和第 2 噪声矢量上分别乘以增益正负标号的第 1 级信息  $S1$  和第 2 级信息  $S2$ ，生成声源矢量  $ST(k)$ ，并将生成的声源矢量输出到驱动声源生成单元 1413 中。

驱动声源生成单元 1413 在从自适应/固定选择单元 1412 收到的自适应/固定矢量  $AF(k)$  和从声源矢量生成装置 1414 收到的声源矢量  $ST(k)$  上分别乘以在参数编码单元 1402 求出的自适应/固定矢量正式增益  $G_{af}$  和噪声矢量正式增益  $G_{st}$  后, 根据增益正负标号  $Is1s2$  进行相加或者相减, 得到驱动声源  $ex(k)$ , 并将得到驱动器声源输出到 LPC 合成滤波器 14136 和自适应码本 1407 中。在这里, 用从驱动声源生成单元 1413 输入的新的驱动声源更新自适应码本 1407 内的旧的驱动声源。

LPC 合成滤波器 1416 对在驱动声源生成单元 1413 生成的驱动声源, 采用以从 LSP 内插入单元 1406 收到的解码内插 LPC 构成的合成滤波器进行 LPC 合成, 并将滤波器的输出送到功率复原单元 1417 中。功率复原单元 1417 首先求出在 LPC 合成滤波器单元 1413 求得的驱动声源合成矢量的平均功率, 接着用将从参数解码单元 1402 收到的解码功率  $spow$  除以求得的平均功率, 并将所得结果与驱动声源的合成矢量乘, 从而生成合成语音 518。

#### 实施形态 9

图 15 表示与实施形态 9 相关的声音编码装置的主要部分的方框图。这种声音编码装置是在图 13 所示的声音编码装置上增加量化对象 LSP 增加单元 151、LSP 量化/解码单元 152 和 LSP 量化误差比较单元 153, 或者变更一部分功能。

LPC 分析单元 1304 对缓存器 1301 内的处理帧进行线性预测分析并得到 LPC 后, 对得到的 LPC 进行变换生成量化对象 LSP, 并将生成的量化对象 LSP 输出到量化对象 LSP 增加单元 151 中。具体地说, 兼备对缓存器内的首读区间进行线性预测分析, 得到对首读区间的 LPC 后, 对得到的 LPC 进行变换, 生成对先读区间的 LSP, 并输出到量化对象 LSP 增加单元 151 中的功能。

量化对象 LSP 增加单元 151 借助 LPC 分析单元 1304 中变换处理帧的 LPC, 除直接得到的量化对象 LSP 以外, 还生成多个量化对象 LSP。

LSP 量化表存储单元 1307 存储 LSP 量化/解码单元 152 参照的量化表, LSP 量化/解码单元 152 对生成的量化对象 LSP 进行量化和解码, 生成各自的解码 LSP。

LSP 量化误差比较单元 153 对生成的多个解码 LSP 进行比较, 以闭环的方式选择 1 个异常噪声最少的解码 LSP, 并将选择的解码 LSP 作为对于处理帧的解码 LSP 重新采用。

图 16 表示量化对象 LSP 增加部分 151 的方框图。

量化对象 LSP 增加部分 151 由存储 LPC 分析单元 1304 中所求处理帧的量化对象 LSP 的当前帧 LSP 存储单元 161、存储 LPC 分析单元 1304 中求出的首读区间的 LSP 的首读区间 LSP 存储单元 162、存储前处理帧的解码 LSP 的前帧 LSP 存储单元 163 和对于从前述 3 个存储单元读出的 LSP 进行线性内插计算、并增加多个量化对象 LSP 的线性内插单元 164 构成。

对处理帧的量化对象 LSP、首读区间的 LSP 以及前处理帧的解码 LSP，进行线性内插计算，增加多个生成量化对象 LSP，并将生成的量化对象 LSP 输出到全部 LSP 量化/解码单元 152 中。

这里，对量化对象 LSP 增加单元 151 进一步详细地进行说明。LPC 分析单元 1304，对缓存器内的处理帧进行线性预测分析，得到预测次数  $N_p (=10)$  次的 LPC  $\alpha(i)$  ( $0 \leq i \leq N_p$ )，对得到的 LPC 进行变换生成量化对象 LSP  $\omega(i)$  ( $0 \leq i \leq N_p$ )，并将生成的量化对象 LSP  $\omega(i)$  ( $0 \leq i \leq N_p$ ) 存储到量化对象 LSP 增加单元 151 内的当前帧 LSP 存储单元 161 中。此外，对缓存器内的首读区间进行线性预测分析，得到对首读区间的 LPC，变换得到的首读区间的 LPC，生成对首读区间的 LSP  $\omega(i)$  ( $0 \leq i \leq N_p$ )，并将生成的首读区间的 LSP  $\omega(i)$  ( $0 \leq i \leq N_p$ ) 存储在量化对象 LSP 增加单元 151 内的首读区间 LSP 存储单元 162 中。

接着，线性内插单元 164 分别从当前帧 LSP 存储单元 161 读出对应于处理帧的量化对象 LSP  $\omega(i)$  ( $0 \leq i \leq N_p$ )，从首读区间 LSP 存储单元 162 读出对应于首读区间的 LSP  $\omega_f(i)$  ( $0 \leq i \leq N_p$ )，从前帧 LSP 存储单元 163 读出对应于前处理帧的解码 LSP  $\omega_{qp}(i)$  ( $0 \leq i \leq N_p$ )，借助于进行式(33)所示的变换，分别生成量化对象增加第 1LSP  $\omega_1(i)$  ( $0 \leq i \leq N_p$ )，量化对象增加第 2LSP  $\omega_2(i)$  ( $0 \leq i \leq N_p$ )，量化对象增加第 3LSP  $\omega_3(i)$  ( $0 \leq i \leq N_p$ )。

$$\begin{bmatrix} \omega_1(i) \\ \omega_2(i) \\ \omega_3(i) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 & 0.0 \\ 0.5 & 0.3 & 0.2 \\ 0.8 & 0.3 & 0.5 \end{bmatrix} \begin{bmatrix} \omega_{qp}(i) \\ \omega_f(i) \\ \omega(i) \end{bmatrix} \quad (33)$$

$\omega_1(i)$ : 量化对象增加第 1LSP

$\omega_2(i)$ : 量化对象增加第 2LSP

$\omega_3(i)$ : 量化对象增加第 3LSP

$i$ : LPC 次号 ( $0 \leq i \leq N_p$ )

$N_p$ : LPC 分析次数 (=10)

$\omega_q(i)$ : 对应于处理帧的解码 LSP

$\omega_{qp}(i)$ : 对应于前处理帧的复合 LSP

$\omega_f(i)$ : 对应于首读区间的 LSP

将生成的  $\omega_1(i)$ 、 $\omega_2(i)$ 、 $\omega_3(i)$  输出到 LSP 量化/解码单元 152 中。LSP 量化/解码单元 152 在对 4 个量化对象 LSP  $\omega(i)$ 、 $\omega_1(i)$ 、 $\omega_2(i)$ 、 $\omega_3(i)$  全部进行矢量量化/解码后, 分别求出对应于  $\omega(i)$  的量化误差的功率  $E_{pow}(\omega)$ 、对应于  $\omega_1(i)$  的量化误差的功率  $E_{pow}(\omega_1)$ 、对应于  $\omega_2(i)$  的量化误差的功率  $E_{pow}(\omega_2)$ 、对于  $\omega_3(i)$  的量化误差的功率  $E_{pow}(\omega_3)$ , 并对求出的各个量化残差功率施行式(34)的变换, 求出解码 LSP 选择基准值  $STD_{lsp}(\omega)$ 、 $STD_{lsp}(\omega_1)$ 、 $STD_{lsp}(\omega_2)$ 、 $STD_{lsp}(\omega_3)$ 。

$$\begin{bmatrix} STD_{lsp}(\omega) \\ STD_{lsp}(\omega_1) \\ STD_{lsp}(\omega_2) \\ STD_{lsp}(\omega_3) \end{bmatrix} = \begin{bmatrix} E_{pow}(\omega) \\ E_{pow}(\omega_1) \\ E_{pow}(\omega_2) \\ E_{pow}(\omega_3) \end{bmatrix} - \begin{bmatrix} 0.0010 \\ 0.0005 \\ 0.0002 \\ 0.0000 \end{bmatrix} \quad (34)$$

$STD_{lsp}(\omega)$ : 对应于  $\omega(i)$  的的复合 LSP 选择基准值

$STD_{lsp}(\omega_1)$ : 对应于  $\omega_1(i)$  的复合 LSP 选择基准值

$STD_{lsp}(\omega_2)$ : 对应于  $\omega_2(i)$  的复合 LSP 选择基准值

$STD_{lsp}(\omega_3)$ : 对应于  $\omega_3(i)$  的复合 LSP 选择基准值

$E_{pow}(\omega)$ : 对应于  $\omega(i)$  的量化误差的功率

$E_{pow}(\omega_1)$ : 对应于  $\omega_1(i)$  的量化误差的功率

$E_{pow}(\omega_2)$ : 对应于  $\omega_2(i)$  的量化误差的功率

$E_{pow}(\omega_3)$ : 对应于  $\omega_3(i)$  的量化误差的功率

比较求出的解码 LSP 选择基准值, 在选择并输出该基准值最小的量化对象 LSP 所对应的解码 LSP 作为对应于处理帧的解码 LSP  $\omega_q(i)$  ( $0 \leq i \leq N_p$ ), 同时在前帧 LSP 存储单元 163 中存储下一个帧的 LSP, 以便能在矢量量化时参照。

本实施形态有效地利用 LSP 具有的内插特性的优越性(即使用内插后的 LSP 合成也不会发生异常噪声), 能对 LSP 进行向量量化, 即使象话头那样频谱变动大的区间, 也不发生异常噪声, 所以能减小在 LSP 的量化特性不充分的情况下可能发生的合成语音中的异常噪声。

图 17 表示本实施形态的 LSP 量化/解码单元 152 的方框图。LSP 量化/解码单元 152 包括增益信息存储单元 171、自适应增益选择单元 172、乘增益运算单元



173、LSP 量化单元 174 和 LSP 解码单元 175。

增益信息存储单元 171 存储自适应增益选择单元 172 中选择自适应增益时参照的多个增益候补。乘增益运算单元 173 将由 LSP 量化表存储单元 1307 读出的码矢量乘以自适应增益选择单元 172 中选择的自适应增益。LSP 量化单元 174 用乘以自适应增益后的码矢量，对量化对象 LSP 进行矢量量化。LSP 解码单元 175 具有对矢量量化的 LSP 进行解码，生成并输出解码 LSP 的功能，还具有求出作为量化对象 LSP 与解码 LSP 的差分的 LSP 量化误差，输出到自适应增益选择单元 172 中的功能。自适应增益选择单元 172 以矢量量化时在码矢量上乘以前处理帧的 LSP 的自适应增益的大小，和对应于前帧的 LSP 量化误差的大小为基准，以存储在增益存储单元 171 中的增益生成信息为基础进行自适应调节，同时求出对处理帧的量化对象 LSP 进行矢量量化时乘到码矢量上自适应增益，并将求得的自适应增益输出到乘法增益运算单元 173 中。

这样，LSP 量化/解码单元 152 是在自适应码矢量上的自适应增益的同时，对量化对 LSP 进行矢量量化和解码。

这里，对 LSP 量化/解码单元 152 进一步详细地进行说明。增益信息存储单元 171 存储自适应增益选择单元 103 参照的 4 个增益候补 (0.9, 1.0, 1.1, 1.2)，自适应增益选择单元 103，利用在量化前帧的量化对象 LSP 时生成的功率 ERpow 除以矢量量化前处理帧的量化对象 LSP 时选择的自适应增益 Gqlsp 的平方的式 (35)，求出自适应增益选择基准值 Slsp。

$$Slsp = \frac{ERpow}{Gqlsp^2} \quad (35)$$

Slsp: 自适应增益选择基准值

ERpow: 量化前帧的 LSP 时生成的量化误差的功率

Gqlsp: 量化前帧的 LSP 时选择的自适应增益

根据使用求得的自适应增益选择基准值 Slsp 的式 (36)，从由增益信息存储单元 171 读出的 4 个增益候补 (0.9, 1.0, 1.1, 1.2) 中选择 1 个增益。并且，在将被选择的自适应增益 Gqlsp 的值输出到乘增益运算单元 173 中的同时，将用于指定被选择的适应增益是 4 种中的哪一种的信息 (2 位信息) 输出到参数编码单元中。

$$G_{lsp} = \begin{cases} 1.2 & S_{lsp} > 0.0025 \\ 1.1 & S_{lsp} > 0.0015 \\ 1.0 & S_{lsp} > 0.0008 \\ 0.9 & S_{lsp} \leq 0.0008 \end{cases} \quad (36)$$

$G_{lsp}$ : 乘在 LSP 量化用码矢量上的自适应增益

$S_{lsp}$ : 自适应增益选择基准值

在变量  $G_{qlsp}$  和变量  $ER_{pow}$  中, 保持所选择的自适应增益  $G_{lsp}$  和伴随量化产生的误差, 直到矢量量化下一帧的量化对象 LSP 时为止。

乘增益运算单元 173 在由 LSP 量化表存储单元 1307 读出的码矢量上乘以自适应增益选择单元 172 中选择的自适应增益  $G_{lsp}$ , 并输出到 LSP 量化单元 174 中。LSP 量化单元 174, 用乘以自适应增益的码矢量, 对量化对象 LSP 进行矢量量化, 并将其标号输出到参数编码单元中。LSP 解码单元 175 对在 LSP 量化单元 174 量化的 LSP 进行解码, 得到解码 LSP, 输出到得到的解码 LSP, 同时从量化对象 LSP 减去得到的解码 LSP, 求出 LSP 量化误差, 计算求出的 LSP 量化误差的功率  $ER_{pow}$ , 并输出到自适应增益选择单元 172 中。

本实施形态能减小在 LSP 的量化特性不充分的场合可能发生的合成话音中的异常噪声。

#### 实施形态 10

图 18 表示与本实施形态相关的声源矢量生成装置的结果的方框图。这种声源矢量生成装置包括存储通道 CH1、CH2、CH3 的 3 个固定波形 (V1(长度: L1)、V2(长度: L2)、V3(长度: L3)) 的固定波形存储单元 181, 具有各通道的固定波形起始端候补位置信息, 并将从固定波形存储单元 181 读出的固定波形 (V1、V2、V3) 分别配置在 P1、P2、P3 的位置上的固定波配置在单元 182 和对基于固定波形配置单元 182 配置的固定波形相加, 并输出声源矢量的加法运算单元 183。

下面, 对如前所述结构的声源矢量生成装置的动作进行说明。

在固定波形存储单元 181 上预先存储 3 个固定波形 V1、V2、V3。固定波形配置单元 182 根据表 8 所示的其本身具有的固定波形起始端候补位置信息, 在从 CH1 用的起始端候补位置中选择的位置 P1 上配置(移位)从固定波形存储单元 181 读出的固定波形 V1, 同样, 在从 CH2、CH3 用的起始端候补位置中选择的位置 P2、

P3 上分别配置固定波形 V2、V3。

表 8：固定波形起始端候补位置信息

通道号	符号	固定波形起始端候补位置
CH1	$\pm 1$	P1 0, 10, 20, 30, ..., 60, 70
CH2	$\pm 1$	P2 2, 12, 22, 32, ..., 62, 72 6, 16, 26, 36, ..., 66, 76
CH3	$\pm 1$	P3 4, 14, 24, 34, ..., 64, 74 8, 18, 28, 38, ..., 68, 78

加法运算单元 183 对由固定波形配置单元 182 配置的固定波形进行加法运算并生成声源矢量。

其中，对固定波形配置单元 182 具有的固定波形起始端候补位置信息，分配与能被选择的各固定波形的起始端候补位置的组合信息(表示选择哪一个位置作为 P1、选择哪一个位置作为 P2、选择哪一个位置作为 P3 的信息)一一对应的码号。

采用这样结构的声源矢量生成装置，则在能利用传送与固定波形配置单元 182 具有的固定波形起始端候补位置信息有对应关系的码号，行声音信息的传送的同时，借助于码号仅存在于各起始端候补数的积的部分，能不增加计算和必要的存储器，生成接近实际声音的声源矢量。

为了能利用码号的传送进行声音信息的传送，可前述声源矢量生成装置作为噪声码本用在声音编码/解码装置中。

在本实施形态中，虽然对图 18 所示的用 3 个固定波形的场合进行了说明，但固定波形的个数(图 18 和表 8 的通道数一致)为其它的个数的场合，也能得到同样的作用和效果。

此外，在本实施形态中，虽然对固定波形配置单元 182 具有表 8 所示的固定

波形起始端候补位置信息的场合进行了说明，但对于具有表 8 以外的固定波形起始端候补位置信息的场合，也能得到同样的作用·效果。

#### 实施形态 11

图 19A 表示与本实施形态相关的 CELP 型声音编码装置的结构方框图。图 19B 表示与 CELP 型声音编码装置配对的 CELP 型声音解码装置的结构方框图。

与本实施形态相关的 CELP 型声音编码装置包括由固定波形存储单元 181A 和固定波形配置单元 182A 以及加法运算单元 183A 组成的声源矢量生成装置。固定波形存储单元 181A 存储多个固定波形，固定波形配置单元 182A 根据自己具有的固定波形起始端候补位置信息将从固定波形存储单元 181A 读出的固定波形分别配置(移位)在选择的位置上，加法运算单元 183A 对由固定波形配置单元 182A 配置的固定波形进行加法运算、生成声源矢量 C。

这种 CELP 型声音编码装置包括对被输入的噪声码本检索用目标 X 进行时间反转的时间反转单元 191、对时间反转单元 191 的输出进行合成的滤波器 192、对合成滤波器 192 的输出再次进行反转并输出到时间反转合成目标 X' 的时间反转单元 193、对乘以噪声编码矢量增益  $g_c$  的声源矢量 C 进行合成并输出合成声源矢量 S 的合成滤波器 194，以及输入 X'、C、S 并计算失真的失真计算单元 205 和传送单元 196。

在本实施形态中，固定波形存储单元 181A、固定波形配置单元 182A 和加法运算单元 183A，对应于图 18 所示的固定波形存储单元 181、固定波形配置单元 182 和加法运算单元 183，各通道的固定波形起始端候补位置对应于表 8，因而下文中表示通道号、固定波形号及其长度和位置的记号，使用图 18 和表 8 所示的。

另一方面，图 19B 的 CELP 型声音解码装置包括存储多个固定波形的固定波形存储单元 181B、根据基于自己具有的固定波形起始端候补位置信息，将从固定波形存储单元 181B 读出的固定波形分别配置(移位)在选择的位置上的固定波形配置单元 182B、对由固定波形配置单元 182B 配置的固定波形进行加法运算，生成声源矢量 C 的加法运算单元 183B、乘以噪声编码矢量增益  $g_c$  的乘增益运算单元 197 和对声源矢量 C 进行合成并输出合成声源矢量 S 的合成滤波器 198。

声音解码装置的固定波形存储单元 181B 和固定波形配置单元 182B，与声音编码装置的固定波形存储单元 181A 和固定波形配置单元 182A 具有相同的结构，固定波形存储单元 181A 和 181B 存储的固定波形，是具有借助于将以使用噪声码

本检索用目标的式(3)的编码失真计算式作为价值函数的学习,使式(3)的价值函数统计上最小的特性的固定波形。

下面,对如前所述结构的声码装置的动作进行说明。

噪声码本检索用目标  $X$ , 在时间反转单元 191 被倒置后, 在合成滤波器被合成, 并在时间反转单元 193 再次被倒置后, 作为噪声码本检索用的时间反转合成目标  $X'$  输出到失真计算单元 205 中。

接着, 固定波形配置单元 182A 根据表 8 所示的自己具有的固定波形起始端候补位置信息, 将从固定波形存储单元 181A 读出的固定波形  $V1$  配置(移位)在从 CH1 用的起始端候补位置选择的位置  $P1$  上, 同样, 将固定波形  $V2$ 、 $V3$  配置在从 CH2、CH3 用的起始端候补位置选择的位置  $P2$ 、 $P3$  上。被配置的各固定波形输出到加法器 183A 中进行相加, 成为声源矢量  $C$ , 并输入到合成滤波器 194 中。合成滤波器 194 对声源矢量  $C$  进行合成, 生成合成声源矢量  $S$ , 并输出到失真计算单元 205 中。

失真计算单元 205 输入时间反转合成目标  $X'$ 、声源矢量  $C$ 、合成声源矢量  $S$ , 计算式(4)的编码失真。

失真计算单元 205 在计算失真后, 对固定波形配置单元 182A 能选择的起始端候补位置的全部组合, 重复进行从将信号送到固定波形配置单元 182A 中, 从固定波形配置单元 182A 选择分别对应于 3 个通道的起始端候补位置起, 到在失真计算单元 205 计算失真为止的前述处理。

然后, 选择编码失真最小的起始端候补位置的组合, 将与该起始端候补位置的组合一一对应的码号、以及这时的最佳噪声码矢量增益  $g_c$  作为噪声码本的码, 传送到传送单元 196 中。

接着, 对图 19B 的声音解码装置的动作进行说明。

固定波形配置单元 181B 根据从传送单元 196 送来的信息, 从表 8 所示的自己具有的固定波形起始端候补位置信息中选择各通道的固定波形的的位置, 将从固定波形配置单元 181B 读出的固定波形  $V1$  配置(移位)在从 CH1 用的起始端候补位置中选择的位置  $P1$  上, 同样, 将固定波形  $V2$ 、 $V3$  配置在从 CH2、CH3 用的起始端候补位置中选择的位置  $P2$ 、 $P3$  上。被配置的各固定波形输出到加法器 43 中进行相加, 成为声源矢量  $C$ , 并乘以由来自传送单元 196 的信息选择的噪声码矢量增益  $g_c$  后, 输出到合成滤波器 198 中。合成滤波器 198 对乘以  $g_c$  后的声源矢量

C 进行合成，生成并输出合成声源矢量 S。

采用这样结构的语音编码/解码装置，则因由固定波形存储单元、固定波形配置单元和加法器组成的声源矢量生成单元生成声源矢量，所以增加具有实施形态 10 的效果，此外，用合成滤波器合成这种声源矢量而得的合成声源矢量还具有与实际的目标统计上接近的特性，因而能得到高品质的合成声音。

在本实施形态中，虽然示出了将学习得到的固定波形存储在固定波形存储单元 181A 和 181B 中的情况，但在采用其它的统计分析噪声码本检索用目标 X，并根据其分析结果生成的固定波形的情况下，在采用根据实际见识生成的固定波形的情况下，也能同样地得到高品质的合成声音。

在本实施形态中，虽然对固定波形存储单元存储 3 个固定波形的情况进行了说明，但在固定波形的个数为其它的个数的情况下也能得到同样的作用和效果。

此外，在本实施形态中，虽然对固定波形配置单元具有表 8 所示的固定波形起始端候补位置信息的情况进行了说明，但在具有表 8 以外的固定波形起始端候补位置信息的情况下也能得到同样的作用和效果。

#### 实施形态 12

图 20 是表示本实施形态的 CELP 型语音编码装置的结构方框图。

该 CELP 型语音编码装置具有存储多个固定波形（本实施形态中是 CH1: W1、CH2: W2、CH3: W3 个）的固定波形存储器 200，以及有作为对固定波形存储器 200 中存储的固定波形由代数规则生成其起始端位置用的信息的固定波形起始端候补位置信息的固定波形配置单元 201。又，该 CELP 型语音编码装置具备波形别脉冲响应运算单元 202、脉冲发生器 203 及相关矩阵运算器 204，还具备时间反转单元 193 及失真计算单元 205。

波形别脉冲响应运算单元 202 具有对固定波形存储器 200 来的 3 个固定波形和合成滤波器的脉冲响应  $h$ （长度  $L$ =子帧长度）进行卷积，计算出 3 种波形别脉冲响应（CH1:  $h_1$ 、CH2:  $h_2$ 、CH3:  $h_3$ ，长度  $L$ =子帧长度）的功能。

波形别合成滤波器 192' 具有对使输入的噪声码检索目标 X 时间反转的时间反转单元 191 的输出与来自波形别脉冲响应运算单元 202 的各个波形别脉冲响应  $h_1$ 、 $h_2$ 、 $h_3$  进行卷积的功能。

脉冲发生器 203 只在固定波形配置单元 201 选择的起始候补位置 P1、P2、P3 分别使振幅 1（有极性）的脉冲上升，产生不同通道的脉冲（CH1:  $d_1$ 、CH2:

d2、CH3: d3)。

相关矩阵运算器 204 计算来自波形别脉冲响应运算单元 202 的波形别脉冲响应 h1、h2 与 h3 各自的自相关, 以及 h1 与 h2、h1 与 h3、h2 与 h3 的互相关, 将求得的相关值在相关矩阵存储器 RR 中展开。

失真运算单元 205 用 3 个波形别时间反转合成目标 (X' 1、X' 2、X' 3)、相关矩阵存储器 RR、3 个通道别脉冲 (d1、d2、d3), 借助于式(4)的变形式(37)指定使编码失真最小的噪声码矢量。

$$\frac{(\sum_{i=1}^3 x_i^* d_i)^2}{\sum_{i=1}^3 \sum_{j=1}^3 H_i^* H_j d_j} \quad (37)$$

$d_i$ : 通道别脉冲 (矢量)

$d_i = \pm 1 \times \delta(k - p_i), k=0 \sim L-1, p_i$ : 第 i 通道 n 固定波形起始端候补位置

$H_i$ : 波形别脉冲响应卷积矩阵 ( $H_i = HW_i$ )

$W_i$ : 固定波形卷积矩阵

$$W_i = \begin{bmatrix} w_i(0) & 0 & \Lambda & \Lambda & 0 & 0 & 0 & 0 \\ w_i(1) & w_i(0) & 0 & \Lambda & 0 & 0 & 0 & 0 \\ w_i(2) & w_i(1) & w_i(0) & 0 & 0 & 0 & 0 & 0 \\ M & M & M & 0 & 0 & 0 & 0 & 0 \\ w_i(L_i-1) & w_i(L_i-2) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & w_i(L_i-1) & w_i(L_i-2) & 0 & 0 & 0 & \Lambda & 0 \\ M & 0 & w_i(L_i-1) & 0 & 0 & 0 & 0 & 0 \\ M & M & 0 & 0 & 0 & 0 & 0 & 0 \\ M & M & M & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_i(L_i-1) & \Lambda & w_i(1) & w_i(0) \end{bmatrix}$$

其中  $W_i$  为第 i 通道的固定波形 (长度:  $L_i$ )

$x'_i$ : 在  $H_i$  将  $x$  时间反转合成倒置的矢量 ( $x'_i = H_i x$ )

这里对从式(4)变成式(37)的变换, 分别用式(38)和式(39)表示出分母项和分子项的变换。

$$\begin{aligned}
& (x' H c)^2 \\
&= (x' H (W_1 d_1 + W_2 d_2 + W_3 d_3))^2 \\
&= (x' (H_1 d_1 + H_2 d_2 + H_3 d_3))^2 \\
&= ((x' H_1) d_1 + (x' H_2) d_2 + (x' H_3) d_3)^2 \\
&= (x_1' d_1 + x_2' d_2 + x_3' d_3)^2 \\
&= \left( \sum_{i=1}^3 x_i' d_i \right)^2 \tag{38}
\end{aligned}$$

$x$ : 噪声码检索目标 (矢量)

$x'$ :  $x$  的倒易矢量

$H$ : 合成滤波器的脉冲响应卷积矩阵

$c$ : 噪声码矢量 ( $c = W_1 d_1 + W_2 d_2 + W_3 d_3$ )

$W_i$ : 固定波形卷积矩阵

$d_i$ : 通道别脉冲 (矢量)

$H_i$ : 波形别脉冲响应卷积矩阵 ( $H_i = H W_i$ )

$x'_i$ : 在  $H_i$  将  $x$  时间反转合成倒置的矢量 ( $x'_i = x' H_i$ )

$$\begin{aligned}
& \|Hc\|^2 \\
&= \|H(W_1 d_1 + W_2 d_2 + W_3 d_3)\|^2 \\
&= \|H_1 d_1 + H_2 d_2 + H_3 d_3\|^2 \\
&= (H_1 d_1 + H_2 d_2 + H_3 d_3)' (H_1 d_1 + H_2 d_2 + H_3 d_3) \\
&= (d_1' H_1' + d_2' H_2' + d_3' H_3') (H_1 d_1 + H_2 d_2 + H_3 d_3) \\
&= \sum_{i=1}^3 \sum_{j=1}^3 d_i' H_i' d_j H_j \tag{39}
\end{aligned}$$

$H$ : 合成滤波器的脉冲响应卷积矩阵

$c$ : 噪声码矢量 ( $c = W_1 d_1 + W_2 d_2 + W_3 d_3$ )

$W_i$ : 固定波形卷积矩阵

$d_i$ : 通道别脉冲 (矢量)

$H_i$ : 波形别脉冲响应卷积矩阵 ( $H = H W_i$ )

下面对具有如上所述结构的 CELP 型声音编码装置的动作加以说明。

首先, 对波形别脉冲响应运算单元 202 存储的 3 个固定波形  $W_1$ 、 $W_2$ 、 $W_3$



和脉冲响应  $h$  进行卷积，计算出 3 种波形别脉冲响应  $h_1$ 、 $h_2$ 、 $h_3$ ，输出到波形别合成滤波器 192' 及相关矩阵运算器 204。

接着，波形别合成滤波器 192' 对由时间反转单元 191 进行过时间反转的噪声码检索目标  $X$  和输入的 3 种波形别脉冲响应  $h_1$ 、 $h_2$ 、 $h_3$  的各个进行卷积，用时间反转单元 193 再度对来自波形别合成滤波器 192' 的 3 种输出矢量进行时间反转，分别生成 3 个波形别时间反转合成目标  $X' 1$ 、 $X' 2$ 、 $X' 3$  输出到失真计算单元 205。

接着，相关矩阵运算单元 204 计算输入的 3 种波形别脉冲响应  $h_1$ 、 $h_2$ 、 $h_3$  各自的自相关和  $h_1$  与  $h_2$ 、 $h_1$  与  $h_3$ 、 $h_2$  与  $h_3$  的互相关，将求得的相关值在相关矩阵存储器  $RR$  展开后输出到失真运算单元 205。

将上述处理作为前处理实施后，固定波形配置单元 201 在每一个通道各选一个固定波形的起始端候补位置，向脉冲发生器 203 输出该位置信息。

脉冲发生器 203 在从固定波形配置单元 121 得到的选择位置上分别使振幅 1（有极性）的脉冲上升，产生通道别脉冲  $d_1$ 、 $d_2$ 、 $d_3$  并输出到失真计算单元 205。

然后，失真计算单元 205 用 3 个波形别时间反转合成目标  $X' 1$ 、 $X' 2$ 、 $X' 3$ 、相关矩阵  $RR$  及 3 个通道别脉冲  $d_1$ 、 $d_2$ 、 $d_3$ ，计算式(37)的最小编码失真基准值。

固定波形配置单元 201 就该单元能够选择的起始端候补位置的全部组合，反复进行从选择对分别与 3 个通道对应的起始端候补位置起，到失真计算单元 205 计算失真为止的上述处理。然后，在将噪声码矢量增益  $g_c$  指定为噪声码本的代码后，将使式(37)的编码失真检索基准值最小的起始端候补位置的组合编号所对应的码号及那时的最佳增益传送到传输单元。

还有，本实施形态的声音解码装置的结构与实施形态 10 的图 19B 相同，声音编码装置的固定波形存储单元及固定波形配置单元与声音解码装置装置的固定波形存储单元及固定波形配置单元有相同的结构。固定波形存储单元存储的固定波形是具有将使用噪声码本检索目标的式(3)（编码畸变计算式）作为价值函数学习，以在统计上使式(3)的价值函数最小的特性的固定波形。

采用这样构成的声音编码/解码装置，在能够以代数计算算出固定波形配置单元内的固定波形起始端修补位置的情况下，将前处理阶段求得的波形别时间反转合成目标的 3 项相加，取其结果的平方，可以计算式(37)的分子项。又，将前处理阶段求得的波形别脉冲响应的相关矩阵的 9 项相加，可以计算式(37)的分子

项。因此，可以用与将已有的代数结构声源（以振幅 1 的几个脉冲构成声源矢量）用于噪声码本的情况相同的运算量完成检索。

再者用合成滤波器合成的合成声源矢量与实际目标有在统计上相近的特性，因此可以得到高质量的合成话音。

还有，本实施形态示出了将学习得到的固定形状存储于固定波形存储单元的情况，此外，在使用对噪声码本检索用的目标 X 进行统计分析，根据该分析结果作成的固定波形的情况下，以及使用根据实际见识作成的固定波形的情况下，也一样能够得到高质量的合成话音。

又，本实施形态对固定波形存储单元存储 3 个固定波形的情况作出了说明，但是固定波形的个数取其他数值的时也能得到相同的作用与效果。

又，本实施形态对固定波形配置单元具有表 8 所示的固定波形起始端候补位置信息的情况作了说明，但如果是能够以代数方法生成的，则具有表 8 以外的固定波形起始端候补位置信息的情况也能得到同样的作用和效果。

### 实施形态 13

图 21 是本实施形态的 CELP 型声音编码装置的结构方框图。本实施形态的编码装置具备 2 种噪声码本 A211、B212、切换两种噪声码本的开关 213、进行噪声码矢量乘以增益的运算的乘法器 214、将由开关 213 连接的噪声码本输出的噪声码矢量加以合成的合成滤波器 215，以及计算式(2)的编码失真的失真计算单元 216。

噪声码本 A211 具有实施形态 10 的声源矢量生成装置的结构，另一噪声码本 B212 由存储根据随机数序列作出的多个随机矢量的随机数序列存储单元 217 构成。以闭环进行噪声码本的切换。X 是噪声码本检索用的目标。

下成对具有如上所述结构的 CELP 型声音编码装置的动作加以说明。

开始时，开关 213 连接于噪声码本 A211 一侧，固定波形配置单元 182 根据示于表 8 的本身具有的固定波形起始端候补位置信息，将从固定波形存储单元 181 读出的固定波形分别配置（移位）到从起始端候补位置选择出的位置上。所配置的固定波形由加法器 183 进行加法运算，变成噪声码矢量，并乘以噪声码矢量增益后被输入合成滤波器 215。合成滤波器 215 将所输入的噪声码矢量加以合成后，输出到失真计算单元 216。

失真计算单元 216 使用噪声码本的检索用目标 X 和从合成滤波器 215 得到的

合成，进行式(2)的使编码失真最小的处理。

失真计算单元 216 在计算失真之后，向固定波形配置单元 182 传送信号，就固定波形配置单元 182 能选择的起始端候补位置的全部组合，反复进行从固定波形配置单元 182 选择起始端候补位置起，到失真计算单元 216 计算失真为止的上述处理。

然后，选择最小编码失真的起始端候补位置的组合，存储与该起始端候补位置的组合一一对应的噪声码矢量的码号、那时的噪声码矢量增益  $g_c$ ，及编码失真最小值。

接着，开关 213 连接于噪声码本 B212 一侧，从随机数序列存储单元 217 读出的随机数序列成为噪声码矢量，乘以噪声码矢量增益后，输出到合成滤波器 215。合成滤波器 215 将所输入的噪声码矢量加以合成后，输出到失真计算单元 216。

失真计算单元 216 用噪声码本检索用的目标 X 和从合成滤波器 215 得到的合成矢量，计算式(2)的编码失真。

失真计算单元 216 在计算失真之后向随机数序列存储单元 217 传送信号，就随机数序列存储单元 217 能选择的全部噪声码矢量，反复进行从随机数序列存储单元 217 选择噪声码矢量起，到在失真计算单元 216 计算失真为止的上述处理。

然后，选择编码失真最小的噪声码矢量，将该噪声码矢量的码号、那时的噪声码矢量增益  $g_c$ ，以及编码失真最小值存储起来。

接着，失真计算单元 216 将把开关 213 连接于噪声码本 A211 时得到的编码失真最小值与把开关 213 连接于噪声码本 B212 时得到的编码失真最小值加以比较，将得到较小编码失真时的开关连接信息及那时的码号和噪声码矢量增益判定为声音码，传送到未图示的传输单元。

还有，与本实施形状的声音编码装置配对的声音解码装置是将噪声码本 A、噪声码本 B、开关、噪声码矢量增益，及合成滤波器以同图 21 一样的结构配置而成的，根据由传输单元输入的声音码，决定所使用的噪声码本。噪声码矢量及噪声码矢量增益，得到合成声源矢量作为合成滤波器的输出。

采用这样构成的声音编码装置/解码装置，可以从由噪声码本 A 生成的噪声码矢量和由噪声码本 B 生成的噪声码矢量中，以闭环的方式选择使式(2)的编码失真最小的，因此，能够生成更接近实际声音的声源矢量，同时能够得到高音质的

合成语音。

本实施形态示出以作为已有 CELP 型声音编码装置的图 2 所示结构为基础的声音编码/解码装置，但是在图 19A、B 或图 20 的结构为基础的 CELP 型声音编码装置/解码装置中使用本实施形态也能得到同样的作用与效果。

本实施形态设噪声码本 A211 图 18 的结构，但是在固定波形存储单元 181 具有其他结构的情况（例如有 4 种固定波形等）等也能得到同样的作用和效果。

在本实施形态中，对噪声码本 A211 的固定波形配置单元 182 具有表 8 中所示的固定波形起始端候补位置信息的情况作了说明，但是，具有其他固定波形起始端候补位置信息时也能得到同样的作用和效果。

又，本实施形态对噪声码本 B212 由直接在存储器中存储多个随机数序列的随机数序列存储单元 217 构成的情况进行了说明，但是噪声码本 B212 具有其他声源结构的情况（例如由代数结构声源生成信息构成的情况）也能得到同样的作用和效果。

再者，本实施形态对具有 2 种噪声码本的 CELP 型声音编码/解码装置作了说明，但采用具有 3 种以上噪声码本的 CELP 型声音编码/解码装置时，也能取得同样的作用和效果。

#### 实施形态 14

图 22 表示本实施形态的 CELP 型声音编码装置的结构。本实施形态的声音编码装置具有两种噪声码本，一种噪声码本是实施形态 10 的图 18 所示的声源矢量生成装置的结构，另一噪声码本由存储多个脉冲串的脉冲串存储单元构成，利用噪声码本检索以前已经得到的量化音调增益，自适应地换用噪声码本。

噪声码本 A211 由固定波形存储单元 181、固定波形配置单元 182、加法器 183 构成，与图 18 的原矢量生成装置对应。噪声码本 B221 由存储多个脉冲串的脉冲串存储单元 222 构成。开关 213' 对噪声码本 A211 与噪声码本 B211 进行切换。又，乘法器 224 输出自适应码本 223 的输出乘以在噪声码本检索时已经得到的音调增益得出的自适应码矢量。音调增益量化器 225 的输出传送给开关 213'。

下面对具有上述结构的 CELP 型声音编码装置的动作加以说明。

已有的 CELP-型声音代码装置首先进行自适应码本 223 的检索，接着接受其结果，进行噪声码本检索。该自适应码本检索是从自适应码本 223 存储的多个自

适应码矢量（自适应码矢量与噪声码矢量乘以各自的增益后进行相加而得到的矢量）选择最合适的自适应码矢量的处理，结果是生成自适应码矢量的码号及音调增益。

本实施形态的 CELP 型声音编码装置在音调增益量化单元 225 将该音调增益量化，并在生成量化音调增益之后进行噪声码本检索。音调增益量化单元 225 得到的量化音调增益送往切换噪声码本用的开关 213'。

开关 213' 在量化音调增益的值小的时候判断为输入声音清音性强，连接噪声码本 A211，在量化音调增益值大的时候判断为输入声音浊音性强，连接噪声码本 B221。

开关 213' 连接于噪声码本 A211 一侧时，固定波形配置单元 182 根据示于表 8 的本身具有的固定波形起始端候补位置信息，将从固定波形存储单元 181 读出的固定波形分别配置（移位）到从起始端候补位置选择出的位置上。所配置的各固定波形输出到加法器 183 进行加法运算，成为噪声码矢量，乘以噪声码矢量增益后输入合成滤波器 215。合成滤波器 215 将输入的噪声码矢量加以合成，输出到失真计算单元 216。

失真计算单元 216 利用噪声码本检索用目标 X 和从合成滤波器 215 得到的矢量，计算式(2)的编码失真。

失真计算单元 216 在计算失真之后向固定波形配置单元 182 传送信号 182，就固定波形配置单元 182 能够选择的起始端候补位置的全部组合，反复进行从固定波形配置单元 182 选择起始端候补位置起，到失真计算单元 216 计算失真为止的上述处理。

然后，选择编码失真最小的起始端候补位置的组合，将与该起始端候补位置的组合一一对应的噪声码矢量的码号、那时的噪声码矢量增益  $g_c$ ，及量化音调增益作为声音码传送到传输单元。本实施形态在进行声音编码之前事先使固定波形存储单元 181 存储的固定波形图呈现清音性质。

另一方面，开关 213' 连接于噪声码本 B221 一侧时从脉冲串存储单元 222 读出的脉冲串成为噪声码矢量，开关 213' 经噪声码矢量增益的乘法运算过程后，输入合成滤波器 215。合成滤波器 215 将输入的噪声码矢量加以合成，并输出到失真计算单元 216。

失真计算单元 216 用噪声码本检索用目标 X 和从合成滤波器 215 得到的合成

矢量，计算式(2)的编码失真。

失真计算单元 216 在计算失真之后向脉冲串存储单元 222 传送信号，就脉冲串存储单元 222 能够选择的所有的噪声码矢量，反复进行从脉冲串存储单元 222 选择噪声码矢量起，到失真计算单元 216 计算失真为止的上述处理。

然后，选择编码失真最小的噪声码矢量，将该噪声码矢量的码号、那时的噪声码矢量增益  $g_c$ ，以及量化音调增益作为声音码向传输单元传送。

还有，与本实施形态的声音编码装置配对的声音解码装置是具有将噪声码本 A、噪声码本 B、开关、噪声码矢量增益，以及合成滤波器以与图 22 相同的结构配置而成的部分的装置，首先，接收传送来的量化音调增益，根据其大小判断在编码装置一方开关 213' 是连接于噪声码本 A211 一侧，还是连接于噪声码本 B221 一侧。接着，根据码号及噪声码矢量增益的代码，得到合成声源矢量作为合成滤波器的输出。

采用具有这样的结构的声源编码/解码装置，可以相应于输入声音的特征（在本实施形态中，利用量化音调增益的大小作为浊音性/清音性的判断数据）自适应地切换 2 种噪声码本，能够在输入声音的浊音性强的情况下选择脉冲串作为噪声码矢量，在清音性强的情况下，选择呈现清音性质的噪声码矢量，可生成更接近原声的声源矢量，同时可以提高合成话音的音质。在本实施形态中，由于如上所述以开环进行开关切换，可以使传送的信息增加，以提高有关作用和效果。

本实施形态中示出以作为已有的 CELP 型声音编码装置的图 2 所示结构为基础的声源编码/解码装置，但是在以图 19A、B 或图 20 的结构为基础的 CELP 型声音编码/解码装置中使用本实施形态也可以得到同样的效果。

本实施形态中，作为用于切换开关 213' 的参数，使用在音调增益量化器 225 将自适应码矢量的音调增益量化而得到的音调增益，但是也可以代之以使用配备音调周期运算器，从自适应码矢量计算出的音调周期。

本实施形态中，设噪声码本 A211 具有图 18 的结构，但是在固定波形存储单元 181 具有其他结构的情况下（例如有 4 种固定波形的情况等），也能得到同样的作用与效果。

在本实施形态中，对噪声码本 A211 的固定波形配置单元 182 具有表 8 所示的固定波形起始端候补位置信息的情况作了说明，但是具有其他固定波形起始端候补位置信息时也能够得到同样的作用与效果。

在本实施形态中，就噪声码本 B211 由直接将脉冲串存储于存储器中的脉冲串存储单元 222 构成的情况作了说明，但是在噪声码本 B221 具有其他声源结构（例如由代数结构声源生成信息构成的情况下）也能够得到同样的作用与效果。

还有，在本实施例中，对具有 2 种噪声码本的 CELP 型声音编码/解码装置进行了说明，但是采用具有 3 种以上噪声码本的 CELP 型声音编码/解码装置时，也能够得到同样的作用与效果。

#### 实施形态 15

图 23 是本实施形态的 CELP 型声音编码装置的结构方框图。本实施形态的声音编码装置具有两种噪声码本，一种噪声码本是实施形态 10 的图 18 所示的声源矢量生成装置的结构，在固定波形存储单元存储 3 个固定波形，另一噪声码本同样是图 18 所示的声源矢量生成装置的结构，但固定波形存储单元存储的固定波形是 2 个，而且以闭环进行上述两种噪声码本的切换。

噪声码本 A211 由存储 3 个固定波形的固定波形存储单元 A181、固定波形配置单元 A182、加法器 183 构成，与以图 18 的声源矢量生成装置的结构在固定波形存储单元存储 3 个固定波形的情况对应。

噪声码本 B230 由存储 2 个固定波形的固定波形存储单元 B231、具备表 9 所示的固定波形起始端候补位置信息的固定波形配置单元 B232、将由固定波形配置单元 B232 配置的 2 个固定波形相加生成噪声码矢量的加法器 233 构成，与以图 18 的声源矢量生成装置的结构在固定波形存储单元存储 2 个固定波形的情况对应。

表 9

通道编号	符号	固定波形起始端候补位置
CH1	±	p1 0, 4, 8, 12, 16, ..., 72, 76
		2, 6, 10, 14, 18, ..., 74, 78
CH2	±	p2 1, 5, 9, 13, 17, ..., 73, 77
		3, 7, 11, 15, 19, ..., 75, 79

其他结构也与上述实施形态 13 相同。

下面对具有如上所述的结构的 CELP 型声音编码装置的动作加以说明。

开始时，开关 213 连接于噪声码本 A211 一侧，固定波形存储单元 A181 根据表 8 所示的本身具有的固定波形起始端候补位置信息，将从固定波形存储单元 A181 读出的 3 个固定波形分别配置（移位）到从起始端候补位置选择出的位置上。所配置的 3 个固定波形输出到加法器 183，经过加法运算，成为噪声码矢量，经过开关 213、乘以噪声码矢量增益的乘法器 213，输入合成滤波器 215。合成滤波器 215 将输入的噪声代码量加以合成，并输出到失真计算单元 216。

失真计算单元用噪声码本检索用的目标 X 和从合成滤波器 215 得到的合成矢量计算式(2)的编码失真。

失真计算单元 216 在计算失真之后向固定波形配置单元 A182 传送信号，就固定波形配置单元 A182 能选择的起始端候补位置的全部组合，反复进行从固定波形配置单元 A182 选择起始端候补位置起，到失真计算单元 216 计算失真为止的上述处理。

然后，选择编码失真最小的起始端候补位置的组合，存储与该起始端候补位置的组合一一对应的噪声码矢量的码号、那时的噪声码矢量增益  $g_c$ ，以及编码失真最小值。

本实施形态中，在进行声音编码之前，存储于固定波形存储单元 A181 的固定波形图使用学习得到的，该学习在固定波形有 3 个的条件下使失真最小。

接着，开关 213 连接于噪声码本 B230 一侧，固定波形存储单元 B231 根据表 9 所示的本身具有的固定波形起始端候补位置信息，将从固定波形存储单元 B231 读出的 2 个固定波形分别配置（移位）到从起始端候补位置选择出的位置上。所配置的 2 个固定波形输出到加法器 233，经过加法运算后，成为噪声码矢量，经过开关 213、将乘以噪声码矢量增益的乘法器 214，输入合成滤波器 215。合成滤波器 215 将输入的噪声码矢量合成，并输出到失真计算单元 216。

失真计算单元 216 用噪声码本检索用的目标 X 和从合成滤波器 215 得到的合成矢量，计算式(2)的编码失真。

失真计算单元 216 在计算失真之后，将信号传送到固定波形配置单元 B232，就固定波形配置单元 B232 能够选择的起始端候补位置的全部组合，反复进行从



固定波形配置单元 B232 选择起始端候补位置，到失真计算单元 216 计算失真为止的上述处理。

然后，选择编码失真最小的起始端候补位置的组合，存储与该起始端候补位置的组合一一对应的噪声码矢量的码号、那时的噪声码矢量增益  $g_c$ ，以及编码失真最小值。本实施形态在进行声音编码之前，存储于固定波形存储单元 B231 的固定波形图使用学习得到的，该学习在固定波形有 2 个的条件下使失真最小。

接着，失真计算单元 216 将开关 213 连接于噪声码本 A211 时得到的编码失真最小值与开关 213 连接于噪声码本 B230 时得到的编码失真最小值加以比较，将得到较小编码失真时的开关连接信息、那时的码号及噪声码矢量增益判定为声音码，传送到传输单元。

还有，在本实施形态中的声音解码装置是具有将噪声码本 A、噪声码本 B、开关、噪声码矢量增益及合成滤波器以与图 23 一样的结构配置而成的部分的装置，根据从传输单元输入的声音码，决定所使用的噪声码本、噪声码矢量及噪声码矢量增益，从而得到合成声源矢量作为合成滤波器的输出。

采用这样构成的声音编码/解码装置，可用闭环从由噪声码本 A 生成的噪声码矢量与噪声码本 B 生成的噪声码矢量中选择使式(2)的编码失真最小的噪声码矢量，因此可以生成更接近原声的声源矢量，同时可以得到高音质的合成话音。

在本实施形态中，示出以作为已有的 CELP 型声音编码装置的图 2 所示结构为基础的声音编码/解码装置，但是，在以图 19A、B 或图 20 的结构为基础的 CELP 型声音编码/解码装置中使用本实施形态也能够得到同样的效果。

在本实施形态中，对噪声码本 A211 的固定波形存储单元 A181 存储 3 个固定波形的情况进行了说明，但是，在固定波形存储单元 A181 具有其他数目的固定波形的情况下（例如有 4 个固定波形的情况等）也能得到同样的作用与效果。对于噪声码本 B230 也相同。

又，在本实施形态中，对噪声码本 A211 的固定波形配置单元 A182 具有表 8 所示的固定波形起始端候补位置信息的情况作了说明，但是，具有其他固定波形起始端候补位置信息时也能够得到同样的作用与效果。对于噪声码本 B230 也相同。

还有，本实施形态对具有 2 种噪声码本的 CELP 型声音编码/解码装置进行了说明，但是采用有 3 种以上噪声码本的 CELP 型声音编码/解码装置时，也能得

到相同的作用与效果。

#### 实施形态 16

图 24 表示本实施形态的 CELP 型声音编码装置的功能方框图。该声音编码装置在 LPC 分析单元 242 对输入的声音数据 241 进行自相关分析与 LPC 分析，以此得到 LPC 系数，又对所得到的 LPC 系数进行编码，得到 LPC 代码，又将得到的 LPC 代码加以编码，得到解码 LPC 系数。

接着，在声源生成单元 245，从自适应码本 243 与声源矢量生成装置 244 取出自适应码矢量与噪声码矢量，分别送往 LPC 合成单元 246。声源矢量生成装置 244 使用上述实施形态 1~4、10 中的任一个声源矢量生成装置。并且，在 LPC 合成单元 246，根据 LPC 分析单元 242 得到的解码 LPC 系数对声源生成单元 245 得到的 2 个声源进行滤波，从而得到两个合成话音。

还在比较单元 247 分析在 LPC 合成单元 246 得到的 2 种合成话音与输入的声音的关系，求两种合成话音的最佳值（最佳增益），把根据该最佳增益进行过功率调整的各合成话音相加，得到总合成话音，计算该总合成话音与输入的声音的距离。

又，对自适应码本 243 与声源矢量生成装置 244 产生的全部声源样本，计算由于使声源生成单元 245、LPC 合成单元 246 起作用而得到的多个合成话音与输入的声音的距离，求得在该结果所得到的距离中为最小的时候的声源样本的标号，再把与该标号对应的两个声源传送到参数编码单元 248。

参数编码单元 248 进行最佳增益的编码，得到增益代码，将 LPC 代码、声源样本标号汇集在一起传送到传输路径 249。又根据增益代码和对应于标号的两个声源生成实际声源信号，将其存储于自适应码本 243，同时废弃旧声源样本。

图 25 是与参数编码单元 248 中增益矢量量化有关的部分的功能方框图。

参数编码单元 248 具备：变换为输入最佳增益 2501 的组成部分的和以及对该和的比率求量化对象矢量的参数变换单元 2502、用解码矢量存储单元存储的过去已解码代码矢量和预测系数存储单元存储的预测系数求目标矢量的目标提取单元 2503，存储过去已解码代码矢量的解码矢量存储单元 2504、存储预测系数的预测系数存储单元 2505、用预测系数存储单元存储的预测系数，计算矢量码本存储的多个代码矢量与目标提取单元得到的目标矢量之间的距离的距离计算单元 2506、存储多个代码矢量的矢量码本 2507、以及控制矢量码本与距离计算单

元，根据对从距离计算单元得到的距离的比较，求出最佳代码矢量的编号，并根据求得的编号取出矢量存储单元存储的代码矢量，用该代码矢量更新解码矢量存储单元的内容的比较单元 2508。

下面对具有如上所述结构的参数编码单元 248 的动作做详细说明。预先生成存储多个量化对象矢量的代表性样本（代码矢量）的矢量码本 2507、这通常以分析多个声音数据得到的多个矢量为基础，用 LBG 算法(IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. COM-28, NO.1, pp84-95, JANUARY 1980) 生成。

又，在预测系数存储单元 2505 存储着用于进行预测编码的系数。关于该预测系数的算法将在后面进行说明。又在解码矢量存储单元 2504 中预先存储表示清音状态的数值作为初始值。例如功率最小的代码矢量。

首先，在参数变换单元 2502 将输入的最佳增益 2501（自适应声源的增益与噪声声源的增益）变换成和与比率要素的矢量（输入）。变换方法示于式(40):

$$\begin{aligned} P &= \log (Ga + Gs) \\ R &= Ga / (Ga + Gs) \end{aligned} \quad \dots\dots(40)$$

$(Ga + Gs)$ : 最佳增益

$Ga$ : 自适应声源增益

$Gs$ : 随机声源增益

$(P, R)$ : 输入矢量

$P$ : 和

$R$ : 比率

上述各量中， $Ga$  不必一定是正值，因而  $R$  也有负值的情况。而且，在  $Ga + Gs$  为负值的情况下代入预先准备的固定值。

接着，在目标提取单元 2503 以在参数变换单元 2052 得到的矢量为基础，利用解码矢量存储单元 2504 存储的过去的解码矢量和预测系数存储单元 2505 存储的预测系数，得到目标矢量。将目标矢量的计算式示于式(41):

$$\begin{aligned}
 T_p &= P - \left( \sum_{i=1}^l U_{pi} \times p_i + \sum_{i=1}^l V_{pi} \times r_i \right) \\
 T_r &= R - \left( \sum_{i=1}^l U_{ri} \times p_i + \sum_{i=1}^l V_{ri} \times r_i \right)
 \end{aligned}
 \tag{4.1}$$

- $(T_p, T_r)$ : 目标矢量  
 $(P, R)$ : 输入矢量  
 $(p_i, r_i)$ : 过去的解码矢量  
 $U_{pi}, V_{pi}, U_{ri}, V_{ri}$ : 预测系数 (固定值)  
 $i$ : 前面第几个解码矢量的标号  
 $l$ : 预测次数

接着在距离计算单元 2506 用预测系数存储单元 2505 存储的预测系数计算在目标提取单元 2503 得到的目标矢量与矢量码本 2507 存储的代码矢量的距离。

距离的计算式示于式(42):

$$\begin{aligned}
 D_n &= W_p \times (T_p - U_{pO} \times C_{pn} - V_{pO} \times Crn)^2 \\
 &\quad + W_r \times (T_r - U_{rO} \times C_{pn} - V_{rO} \times Crn)^2
 \end{aligned}
 \tag{4.2}$$

- $D_n$ : 目标矢量与代码矢量的距离  
 $(T_p, T_r)$ : 目标矢量  
 $U_{pO}, V_{pO}, U_{rO}, V_{rO}$ : 预测系数 (固定值)  
 $(C_{pn}, Crn)$ : 代码矢量  
 $n$ : 代码矢量的编号  
 $W_p, W_r$ : 调节对失真的灵敏度的加权系数 (固定)

接着, 比较单元 2508 控制矢量码本 2507 与距离计算单元 2506, 在矢量码本 2507 中存储的多个代码矢量中求距离计算单元 2506 计算出的距离为最小的代码矢量的编号, 以此作为增益的代码 2509。又以得到的增益代码 2509 为基础求解码矢量, 并利用该矢量更新解码矢量存储单元 2504 的内容。求解码矢量的方法示于式(43):

$$\begin{aligned}
 p &= \left( \sum_{i=1}^l U_{pi} \times p_i + \sum_{i=1}^l V_{pi} \times r_i \right) + U_{pO} \times C_{pn} + V_{pO} \times C_{rn} \\
 R &= \left( \sum_{i=1}^l U_{ri} \times p_i + \sum_{i=1}^l V_{ri} \times r_i \right) + U_{rO} \times C_{pn} + V_{rO} \times C_{rn}
 \end{aligned} \tag{43}$$

$(C_{pn}, C_{rn})$ : 代码矢量

$(p, r)$ : 解码矢量

$(p_i, r_i)$ : 过去的解码矢量

$U_{pi}, V_{pi}, U_{ri}, V_{ri}$ : 预测系数 (固定值)

$i$ : 前面第几个解码矢量的标号

$l$ : 预测次数

$n$ : 代码矢量的编号

又, 进行更新的方法示于式(44)。

处理的顺序:

$$\begin{aligned}
 pO &= C_{pN} \\
 rO &= C_{rN} \\
 p_i &= p_{i-1} \quad (i = 1 \sim l) \\
 r_i &= r_{i-1} \quad (i = 1 \sim l)
 \end{aligned} \tag{44}$$

$N$ : 增益的代码

另一方面, 解码装置(解码器)备有与编码装置相同的矢量码本、预测系数存储单元以及解码矢量存储单元, 根据编码装置传送来的增益的代码, 借助于编码装置中比较单元的编码矢量生成功能和解码矢量存储单元的更新功能进行解码。

这里对预测系数存储单元 2505 存储的预测系数的设定方法加以说明。

首先对许多学习用的声音数据进行量化, 收集从其最佳增益求出的输入矢量和量化时的解码矢量编成组, 然后通过使下面的式(45)所示的总失真最小, 对该组求预测系数。具体地说, 以各  $U_{pi}$ 、 $U_{ri}$  对总失真式进行偏微分, 解所得到的联立方程, 从而求出  $U_{pi}$ 、 $U_{ri}$  的值。

$$Total = \sum_{t=0}^T \left\{ W_p \times (P_t - \sum_{i=0}^l U_{pi} \times p_{t,i})^2 + \right. \\ \left. W_r \times (R_t - \sum_{i=0}^l U_{ri} \times r_{t,i})^2 \right\}$$

$$p_t, O = C_p$$

$$r_t, O = C_{rn} \quad \dots\dots(45)$$

*Total*: 总失真

*t*: 时间 (帧编号)

*T*: 矢量组的数据数目

$(P_t, R_t)$ : 时间 *t* 中的最佳增益

$(p_{ti}, r_{t,i})$ : 时间 *t* 中的解码矢量

*U<sub>pi</sub>*、*V<sub>pi</sub>*、*U<sub>ri</sub>*、*V<sub>ri</sub>*: 预测系数 (固定值)

*i*: 表示前面第几个解码矢量的标号

*l*: 预测次数

$(C_{pn}(t), C_{rn}(t))$ : 时间 *t* 中的代码矢量

*W<sub>p</sub>*、*W<sub>r</sub>*: 调节对失真的灵敏度的权重系数 (固定)

采取这样的矢量量化方法, 可以把最佳增益原样矢量量化, 能借助于参数变换单元的特征, 利用功率与各增益的相对大小的相关性, 因而可实现借助于解码矢量存储单元、预测系数存储单元、目标提取单元及距离计算单元的特征, 利用功率与 2 个增益的相对关系间的相关性的增益预测编码, 并且借助于这些特征, 可以充分利用参数之间的相关性。

#### 实施形态 17

图 26 是表示本实施形态的声音编码装置的参数编码单元的功能的方框图。在本实施形态中, 一边根据与声源的标号对应的两个合成话音和听觉加权输入声音估算增益量化引起的失真, 一边进行矢量量化。

如图 26 所示, 该参数编码单元具备: 根据输入的听觉输入声音、听觉加权 LPC 合成自适应声源、作为听觉加权 LPC 合成噪声声源 2601 的输入数据、解码矢量存储单元存储的解码矢量, 以及预测系数存储单元储存的预测系数计算进行距离计算所需的参数的参数计算单元 2602、存储过去解码的代码矢量的解码矢量存储单元 2603、存储预测系数的预测系数存储单元 2604、使用存储于预测系数

存储单元的预测系数，计算以矢量码本中存储的多个代码矢量解码时的编码失真的距离计算单元 2605、存储多个代码矢量的矢量码本 2606，以及控制矢量码本和距离计算单元，根据从距离计算单元得到的编码失真的比较，求出最佳代码矢量的编号，并根据求得的编号取出矢量存储单元所存的代码矢量，用该代码矢量更新解码矢量存储单元的内容的比较单元 2607。

下面对具有如上所述结构的参数编码单元的矢量量化动作加以说明。预先生成存储多个量化对象矢量的代表性样本（代码矢量）的矢量码本 2606。通常是根据 LBG 算法（IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. COM-28, NO.1, PP84-95, JANUARY 1980）等生成的。又在预测系数存储单元 2604 预先存储用于进行预测编码的系数。该系数使用与实施形态 16 中说明的预测系数存储单元 2505 存储的预测系数相同的系数。又在解码矢量存储单元 2603 存储表示清音状态的数值作为初始值。

首先，在参数计算单元 2602，根据听觉加权输入声音、听觉加权 LPC 合成自适应声源、听觉加权 LPC 合成噪声声源 2601，以及解码矢量存储单元 2603 存储的解码矢量、预测系数存储单元 2604 存储的预测系数，对距离计算所需的参数进行计算。距离计算单元的距离根据下式(46)计算：

$$\begin{aligned}
 E_n &= \sum_{i=0}^I (X_i - G_{an} \times A_i - G_{sn} \times S_i)^2 \\
 G_{an} &= O_{rn} \times e \times p(O_{pn}) \\
 G_{sn} &= (1 - O_{rn}) \times e \times p(O_{pn}) \\
 O_{pn} &= Y_p + U_{pO} \times C_{pn} + V_{pO} \times C_{rn} \\
 Y_p &= \sum_{j=1}^J U_{pj} \times p_j + \sum_{j=1}^J V_{pj} \times r_j \\
 Y_r &= \sum_{j=1}^J U_{rj} \times p_j + \sum_{j=1}^J V_{rj} \times r_j
 \end{aligned} \tag{46}$$

$G_{an}, G_{sn}$ : 解码增益

$(O_{pn}, O_{rn})$ : 解码矢量

$(Y_p, Y_r)$ : 预测矢量

$E_n$ : 使用第  $n$  号增益代码矢量时的编码失真

$X_i$ : 听觉加权输入声音

$A_i$ : 听觉加权 LPC 合成自适应声源

$S_i$ : 听觉加权 LPC 合成随机声源

$n$ : 代码矢量的编号

$i$ : 声源数据标号

$I$ : 子帧长度 (输入声音的编码单位)

$(C_{pn}, C_{rn})$ : 代码矢量

$(p_j, r_j)$ : 过去的解码矢量

$U_{pj}, V_{pj}, U_{rj}, V_{rj}$ : 预测系数 (固定值)

$j$ : 表示前面第几个解码矢量的标号

$J$ : 预测次数

因而, 在参数计算单元 2602 对与代码矢量的编号无关的部分进行计算。预先计算的是上述预测矢量和 3 个合成话音之间的相关性及功率。计算式示于式 (47):

$$\begin{aligned}
 Y_p &= \sum_{j=1}^J U_{pj} \times p_j + \sum_{j=1}^J V_{pj} \times r_j \\
 Y_r &= \sum_{j=1}^J U_{rj} \times p_j + \sum_{j=1}^J V_{rj} \times r_j \\
 D_{xx} &= \sum_{i=0}^I X_i \times X_i \\
 D_{xa} &= \sum_{i=0}^I X_i \times A_i \times 2 \\
 D_{xs} &= \sum_{i=0}^I X_i \times S_i \times 2 \\
 D_{aa} &= \sum_{i=0}^I A_i \times A_i \\
 D_{as} &= \sum_{i=0}^I A_i \times S_i \times 2 \\
 D_{ss} &= \sum_{i=0}^I S_i \times S_i
 \end{aligned} \tag{47}$$

$(Y_p, Y_r)$ : 预测矢量

$D_{xx}, D_{xa}, D_{xs}, D_{aa}, D_{as}, D_{ss}$ : 合成话音间的相关值、功率

$X_i$ : 听觉加权输入声音

$A_i$ : 听觉加权 LPC 合成自适应声源

$S_i$ : 听觉加权 LPC 合成随机声源



- $i$ : 声源数据标号  
 $I$ : 子帧长度 (输入声音的编码单位)  
 $(p_j, r_j)$ : 过去的解码矢量  
 $Up_j, Vp_j, Ur_j, Vr_j$ : 预测系数 (固定值)  
 $j$ : 表示前面第几个解码矢量的标号  
 $J$ : 预测次数

接着, 在距离计算单元 2605, 根据参数运算单元 2602 计算的各参数、预测系数存储单元 2604 存储的预测系数、矢量码本 2606 存储的代码矢量算出编码失真。计算式示于式(48):

$$\begin{aligned}
 E_n &= D_{xx} + (G_{an})^2 \times D_{aa} + (G_{sn})^2 \times D_{ss} \\
 &\quad - G_{an} \times D_{xa} - G_{sn} \times D_{xs} + G_{an} \times G_{sn} \times D_{as} \\
 G_{an} &= O_{rn} \times \exp(O_{pn}) \\
 G_{sn} &= (1 - O_{rn}) \times \exp(O_{pn}) \\
 O_{pn} &= Y_p + U_{pO} \times C_{pn} + V_{pO} \times C_{rn} \\
 O_{rn} &= Y_r + U_{rO} \times C_{pn} + V_{rO} \times C_{rn}
 \end{aligned} \tag{48}$$

- $E_n$ : 使用第  $n$  号增益代码矢量时的编号失真  
 $D_{xx}, D_{xa}, D_{xs}, D_{aa}, D_{as}, D_{ss}$ : 合成话音间的相关值、功率  
 $G_{an}, G_{sn}$ : 解码增益  
 $(O_{pn}, O_{rn})$ : 解码矢量  
 $(Y_p, Y_r)$ : 预测矢量  
 $U_{pO}, V_{pO}, U_{rO}, V_{rO}$ : 预测系数 (固定值)  
 $(C_{pn}, C_{rn})$ : 代码矢量  
 $n$ : 代码矢量的编号

还有, 实际上  $D_{xx}$  与代码矢量的编号  $n$  无关, 因此可以省略其加法运算。

接着, 比较单元 2607 对矢量码本 2606 和距离运算单元 2605 进行控制, 在矢量码本 2606 存储的多个代码矢量中, 求距离运算单元 2605 计算出的距离达到最小的代码矢量的编号, 以此作为增益的代码 2608。又以得到的增益代码 2608

为基础求解码矢量，用它来更新解码矢量存储单元 2603 的内容。解码矢量根据式(43)求得。

又，使用更新方法式(44)。

另一方面，声音解码装置预先备有与声音编码装置相同的矢量码本、预测系数存储单元、解码矢量存储单元，根据从编码器传送来的增益代码，利用编码器比较单元生成解码矢量的功能和解码矢量存储单元的更新功能进行解码。

采用具有这样的结构的实施例形态，可以一边根据与声源的标号对应的两种合成话音和输入声音估算增益量化引起的失真，一边进行矢量量化，借助于参数变换单元的特征，利用功率与各增益的相对大小的相关性，因而能实现借助于解码矢量存储单元、预测系数存储单元、目标提取单元、距离计算单元的特征，利用功率与 2 个增益的相对关系之间的相关性的增益预测编码，以此可以充分利用参数之间的相关性。

#### 实施形态 18

图 27 是本实施形态的降噪装置的主要功能方框图。该降噪装置装备于上述声音编码装置。例如，在图 13 所示的声音编码装置中设置在缓冲器 1301 的前级。

图 27 所示的降噪装置具备：A/D 变换器 272、降噪系数存储单元 273、降噪系数调整单元 274、输入波形设定单元 275、LPC 分析单元 276、傅利叶变换单元 277、降噪/频谱补偿单元 278、频谱稳定单元 279、反傅利叶变换单元 280、频谱增强单元 281、波形匹配单元 282、噪声推定单元 284、噪声频谱存储单元 285、前频谱存储单元 286、随机相位存储单元 287、前波形存储单元 288、最大功率存储单元 289。

首先对初始设定加以说明。表 10 表示固定参数的名称和设定例。

表 10

固定参数	设定例
帧长度	160 (在 8Khz 取样数据中为 20msec)
首读数据长度	80 (上述数据中为 10msec)
FFT 次数	256
LPC 预测次数	10
噪声频谱基准持读数	30
指定最小功率	20.0
AR 增强系数 0	0.5
MA 增强系数 0	0.8
高频增强系数 0	0.4
AR 增强系数 1-0	0.66
MA 增强系数 1-0	0.64
AR 增强系数 1-1	0.7
MA 增强系数 1-1	0.6
高频增强系数 1	0.3
功率增强系数	1.2

噪声基准功率	20000.0
无声功率减少系数	0.3
补偿功率上升系数	2.0
噪声基准持续数	5
降噪系数学习系数	0.8
无声检测系数	0.05
指定降噪系数	1.5

又，随机相位存储单元 287 预先存储用于调整相位的相位数据。这些数据在频谱稳定化单元 279 用于使相位转动。相位数据有 8 种的例子示于表 11。

表 11

相位数据
(-0.51, 0.86), ( 0.98, -0.17 )
( 0.30, 0.95 ), ( -0.53, -0.84 )
( -0.94, -0.34 ), ( 0.70, 0.71 )
(-0.22, 0.97), ( 0.38, -0.92 )

以使用上述相位数据为目的的计数器（随机相位计数器）也在随机相位存储单元 287 存储着。该数值预先初始化为 0 存储着。

接着，设定静态的 RAM 区域。亦即对降噪系数存储单元 273、噪声频谱存储单元 285、前频谱存储单元 286、前波形存储单元 288、最大功率存储单元 289 清零。下面叙述对各存储单元的说明和设定例。

降噪系数存储单元 273 是存储降噪系数的区域，作为初始值存储着 20.0。噪声频谱存储单元 285 是对各频率存储表示平均噪声功率、平均噪声频率，以及 1 级候补的补偿用噪声频谱与 2 级候补的补偿用噪声频谱各自的频谱值在几帧以前有过变化的帧数（持续数）的区域，而且作为初始值对平均噪声功率存储足够大的值，对平均噪声频谱存储指定的最小功率，对补偿用噪声频谱和持续数分别

存储足够大的数。

前频谱存储单元 286 是存储补偿用噪声功率、以前的帧的功率（全频带、中频带）（前帧功率）、以前的帧的平滑功率（全频带、中频带）（前帧平滑功率），以及噪声持续数的区域，作为补偿用噪声功率，存储足够大的值，作为前帧功率、全帧平滑功率都存储 0.0，而作为噪声持续数存储噪声基准持续数。

前波形存储单元 288 是存储用于使输出信号匹配的先前帧输出信号末尾首读数据长度份额的数据的区域，作为初始值全部存储 0。频谱增强单元 281 进行 ARMA 及高频增强滤波，而且将以此为目的的各滤波器的状态都清 0。最大功率存储单元 289 是存储输入的信号的功率的最大值的区域，作为最大功率存储 0。

下面用图 27 在每个方框图中对降噪算法加以说明。

首先，用 A/D 变换器 272 对含有声音的模拟输入信号进行 A/D 变换，输入 1 帧长度 + 首读数据长度（上述设定例中为  $160 + 80 = 240$  点）份额。降噪系数调节单元 274 根据降噪系数存储单元 273 存储的降噪系数、指定降噪系数、降噪系数学习系数及补偿功率上升系数，利用式(49)计算出降噪系数及补偿系数。然后，将得到的降噪系数存储于降噪系数存储单元 273，同时将 A/D 变换器 272 得到的输入信号传送到输入波形设定单元 275，再将补偿系数与降噪系数传送到噪声推定单元 284 与降噪频谱补偿单元 278。

$$\begin{aligned} q &= q * C + Q * (1 - C) \\ r &= Q / q * D \end{aligned} \quad \dots\dots(49)$$

$q$ : 降噪系数

$Q$ : 指定的降噪系数

$C$ : 降噪系数学习系数

$r$ : 补偿系数

$D$ : 补偿功率上升系数

还有，降噪系数是表示噪声降低的比例的系数，指定降噪系数是指预先指定的固定降噪系数、降噪系数学习系数是表示降噪系数接近指定降噪系数的比例的系数，补偿系数是调节频谱补偿的补偿功率的系数，补偿功率上升系数是调节补偿系数的系数。

在输入波形设定单元 275，为了能够进行 FFT（快速傅利叶变换），将来自 A/D 变换器 272 的输入信号从后面开始写入具有 2 的乘方的长度的存储器阵列。

前面的部分填上 0。在上述设定例中，在长度为 256 的阵列中 0~15 写入 0，16~255 写入输入信号。这一数组在进行 8 阶快速傅利叶变换（FFT）时用作实数部分。又，虚数部分准备与实数部分相同长度的阵列，全部写着 0。

在 LPC 分析单元 276，对输入波形设定单元 275 设定的实数区域加上汉明窗，并对加汉明窗后的波形进行自相关分析，求自相关函数，进行基于自相关法的 LPC 分析，得到线性预测系数。再把得到的线性预测系数传送到频谱增强单元 281。

傅利叶变换单元 277 有在输入波形设定单元 275 得到的实数部分、虚数部分的存储器阵列进行采用高速傅利叶变换的离散傅利叶变换。计算得到的复数频谱的实数部分与虚数部分的绝对值之和，以此求输入信号的模拟振幅频谱（下称输入频谱）。又求出各频率的输入频谱值的总和（下称输入功率），传送到噪声推定单元 284。又将复数频谱本身传送到频谱稳定单元 279。

下面对噪声推定单元 284 的处理加以说明。

噪声推定单元 284 将傅利叶变换单元 277 得到的输入功率与最大功率存储单元 289 存储的最大功率数值加以比较，在最大功率较小的情况下，以输入功率数值作为最大功率数值，将该数值存储于最大功率存储单元 289，然后，在符合下面三个条件中的至少一个时进行噪声推定，在完全不满足时不进行噪声推定。

- (1) 输入功率比最大功率乘以无声检测系数的积小。
- (2) 降噪系数比指定降噪系数加 0.2 的和大。
- (3) 输入比从噪声频谱存储单元 285 得到的平均噪声功率乘以 1.6 的积小。

这里地噪声推定单元 284 的噪声推定算法加以叙述。

首先，对噪声频谱存储单元 285 存储的 1 级候补、2 级候补的全部频率的持续数进行更新（加 1）。然后，调查 1 级候补的各频率的持续数，在比预先设定的噪声频谱基准持续数大时，以 2 级候补的补偿用频谱与持续数作为 1 级候补，以 2 级候补的补偿用频谱作为 3 级候补的补偿用频谱，取持续数为 0。但是，在调换该 2 级候补的补偿用频谱时不存储 3 级候补，而以 2 级候补经若干放大代用，以此可以节省存储器。在本实施形态中，以 2 级候补的补偿用频谱放大 1.4 倍代用。

在持续数更新后，对各频率进行补偿用噪声频谱与输入频谱的比较。首先，将各频率的输入频谱与 1 级候补的补偿用噪声频谱用比较，如果输入频谱较小，

就取 1 级候补的补偿用噪声频谱与持续数为 2 级候补，以输入频谱作为 1 级候补的补偿用频谱，并将 1 级候补的持续数取 0。在上述条件以外的情况下，进行输入频谱与 2 级候补的补偿用噪声谱的比较，如果是输入频谱较小，取输入频谱为 2 级候补的补偿用频谱，并将 2 级候补的持续数取 0。然后，将得到的 1、2 级候补的补偿用频率与持续数存储于补偿用噪声频谱存储单元 285。同时，对平均噪声频谱也按照下面的式(50)进行更新。

$$s_i = s_i * g + S_i * (1 - g) \quad \dots\dots(50)$$

$s$ : 平均噪声频谱                       $S$ : 输入频谱

$g$ : 0.9 (输入功率比平均噪声功率的一半大的情况下)

0.5 (输入功率比平均噪声功率的一半小的情况下)

$i$ : 频率编号

还有，平均噪声频谱是用模拟的方式求得平均噪声频谱，式(50)中的系数  $g$  是调节平均噪声频谱学习的快慢的系数。亦即，是具有在输入功率与噪声功率相比较小的情况下，判断为是只有噪声的区间的可能性大，提高学习速度，在并非较小的情况下判断为有可能是在声音区间中，降低学习速度的效果的系数。

然后，求平均噪声频谱各频率值的总和，以此作为平均噪声功率。补偿用噪声谱、平均噪声谱、平均噪声功率存储于噪声频谱存储单元 285。

又，在上述噪声推定处理中，如果使 1 个频率的噪声频谱与多个频率的输入频谱对应，则可以节省构成噪声频谱存储单元 285 用的 RAM 容量。下面举出使用本实施形态的 256 点的 FFT 的情况下，根据 4 个频率的输入频谱推定 1 个频率的噪声频谱时的、噪声频谱存储单元 285 的 RAM 容量为例。考虑（模拟）振幅频谱以频率轴左右对称，在所有频率进行推定的情况下，由于存储 128 个频率的频谱和持续数，需要 128（频率）×2（频谱与持续数）×3（补偿用的 1、2 级候补、平均），即共计 768W 的 RAM 容量。

与此相反，在使 1 个频率的噪声频谱与 4 个频率的输入频谱对应的情况下，需要 32（频率）×2（频谱与持续数）×3（补偿用的 1、2 级候补、平均），即共计 192W 的 RAM 容量即可。实验证实，虽然在这种情况下，噪声频谱频率的分辨率降低，但是在上述 1 对 4 的情况下性能几乎没有变坏。而且由于这一做法不是以 1 个频率的频谱推定噪声频谱，在稳态声（正弦波、元音等）长时间持续的情况下，也有防止把这种频谱错误推定为噪声频谱的效果。

下面对降噪/频谱补偿单元 276 进行的处理加以说明。

从输入的频谱中减去噪声频谱存储单元 285 存储的平均噪声频谱与降噪系数调节单元 274 得到的降噪系数的乘积（以下称差频谱）。在节约上述噪声推定单元 284 的说明中所示的、噪声频谱存储单元 285 的 RAM 容量的情况下，减去与输入频谱对应的频率的平均噪声频谱与降噪系数的乘积。然后，在差额谱为负的情况下，将噪声频谱存储单元 285 存储的补偿用噪声频谱的 1 级候补与降噪系数调节单元 274 求出的补偿系数的乘积代入以进行补偿。这一点对所有频率进行。又，对每一频率生成标志数据，以便判明补偿差频谱的频率。例如，每一频率有一个区域，在不补偿时代入 0，在补偿时代入 1。这一标志数据与差频谱一起被送到频谱稳定单元 279。又，调查标志数据的值以求出补偿的总数（补偿数），也将其送往频谱稳定单元 279。

接着，对频谱稳定单元 279 的处理加以说明。这一处理主要是为了起减小对不含声音的区间的异常感觉的作用。

首先，计算降噪/频谱补偿单元 278 得到的各频率的差频谱之和求当前帧的功率。当前帧功率求全频带与中频带两种。全频带是对全部频率（所谓全频带，在本实施形态是 0~128）求得的，中频带是对听觉重要的中间附近的频带（所谓中频带，在本实施形态是 16~79）求得的。

同样，求关于噪声频谱存储单元 285 存储的补偿用噪声频谱的 1 级候补的和，以此作为当前帧噪声功率（全频带、中频带）。在这里，调查降噪/频谱补偿单元 278 得到的补偿数值，在足够大的情况下，并且又是满足下述 3 个条件中的至少 1 个的情况下，判断当前帧是只有噪声的区间，进行频谱的稳定处理。

- (1) 输入功率比最大功率乘以无声检测系数的积小。
- (2) 当前帧功率（中频带）比当前帧噪声功率（中频带）乘以 5.0 的积小。
- (3) 输入功率比噪声基准功率小。

不进行稳定处理时，前频谱存储单元 286 存储的噪声持续数为正时减小，又以当前帧噪声功率（全频带、中频带）为前帧功率（全频带、中频带），分别存储于前频谱存储单元 286，并进入相位扩散处理。

在这里对频谱稳定处理加以说明。这一处理的目的是实现无声区间（没有声音只有噪声的区间）的频谱的稳定和减小功率。处理有两种，在噪声持续数比噪声基准持续数小的情况下实施处理 1，在前者超过后者的情况下实施处理 2。下



面对两种处理进行说明。

#### 处理 1

对前频谱存储单元 286 存储的噪声持续数加 1，又将当前帧噪声功率（全这、中频带）当作前帧功率（全频带、中频带），分别存储于前频谱存储单元 286，并进入相位调整处理。

#### 处理 2

参照前频谱存储单元 286 存储的前帧功率、前帧平滑功率、还有作为固定系数的无声功率降低系数，按照式(51)分别使其变更。

$$\begin{aligned} Dd80 &= Dd80 * 0.8 + A80 * 0.2 * P \\ D80 &= D80 * 0.5 + Dd80 * 0.5 \\ Dd129 &= Dd129 * 0.8 + A129 * 0.2 * P \\ D129 &= D129 * 0.5 + Dd129 * 0.5 \end{aligned} \quad (51)$$

*Dd80*: 前帧平滑功率（中频带）

*D80*: 前帧功率（中频带）

*Dd129*: 前帧平滑功率（全频带）

*D129*: 前帧功率（全频带）

*A80*: 当前帧噪声功率（中频带）

*A129*: 当前帧噪声功率（全频带）

接着，使这些功率反映于差频谱中。为此，计算中频带所乘的系数（以下称系数 1）与全频带所乘的系数（以下称系数 2）等两个系数。首先，以下式（式 (52)）计算系数 1。

$$\begin{aligned} r1 &= D80 / A80 \quad (A80 > 0 \text{ 时}) \\ &1.0 \quad (A80 = 0 \text{ 时}) \end{aligned} \quad (52)$$

*r1*: 系数 1

*D80*: 前帧功率（中频带）

*A80*: 当前帧噪声功率（中频带）

系数 2 受系数 1 的影响，因此，求取的手段有些复杂。其步骤如下。

(1)在前帧平滑功率（全频带）比前帧功率（中频带）小的情况下，或当前帧噪声功率（全频带）比当前帧噪声功率（中频带）小的情况下，转入步骤 (2)，其他情况下转入步骤(3)。

(2) 系数 2 取 0.0，以前帧功率（全频带）作为前帧功率（中频带），转入步骤(6)。

(3) 在当前帧噪声功率（全频带）与当前帧噪声功率（中频带）相等时转入步骤(4)，在不相等时转入(5)。

(4) 系数取 1.0，并转入(6)。

(5) 利用下述式(53)求系数 2，并转入(6)。

$$r2 = (D129 - D80) / (A129 - A80) \quad (53)$$

$r2$ : 系数 2

$D129$ : 前帧功率（全频带）

$D80$ : 前帧功率（中频带）

$A129$ : 当前帧噪声功率（全频带）

$A80$ : 当前帧噪声功率（中频带）

(6) 系数 2 计算处理结束。

利用上述算法得到的系数 1、2 都把上限箝于 1.0，把下限箝于无声功率降低系数。然后，把中频带的频率（本例中为 16~79）的差频谱乘以系数 1 得到的积作为差频谱，再把该差频谱的全频带中去除中频带后的频率（本例中为 0~15，80~128）的差频谱乘以系数 2 得到的积作为差频谱。与此同时，利用下面的式(54)变换前帧功率（全频带、中频带）。

$$D80 = A80 * r1$$

$$D129 = D80 + (A129 - A80) * r2 \quad (54)$$

$r1$ : 系数 1

$r2$ : 系数 2

$D80$ : 前帧功率（中频带）

$A80$ : 当前帧噪声功率（中频带）

$D129$ : 前帧功率（全频带）

$A129$ : 当前帧噪声功率（全频带）

将这样得到的各种功率数据全部存储于前频谱存储单元 286，结束处理(2)。

根据上述要领在频谱稳定单元 279 实现频谱稳定。

下面对相位调整处理加以说明。在已往的频谱相减中，相位原则上不变，但是本实施形态中，在该频率的频谱在削减时得到补偿的情况下，进行随机修改相

位的处理。由于这一处理，余下的噪声的随机性加强，因此有在听觉上不大会给人以不良印象的效果。

首先，得到随机相位存储单元 287 存储的随机相位计数器。然后，参照全部频率的标志数据（表示有否补偿的数据），正在补偿时，利用下面的式(55)，使在傅利叶变换单元 277 得到的复数频谱的相位旋转。

$$\begin{aligned}
 B_s &= S_i * R_c - T_i * R_c + 1 \\
 B_t &= S_i * R_c + 1 + T_i * R_c \\
 S_i &= B_s \\
 T_i &= B_t
 \end{aligned} \tag{55}$$

$S_i$ 、 $T_i$ : 复数频谱、 $i$ : 表示频率的标号

$R$ : 随机相位数据、 $c$ : 随机相位计数器

$B_s$ 、 $B_t$ : 计算基数寄存器

在式(55)中，成对使用两个随机相位数据。因而，每进行一次上述处理，使随机相位计数器增加 2，在达到上限（在本实施形态中为 16）的情况下取 0。还有，随机相位计数器存储于随机相位存储单元 287，所得到的复数频谱传送到反傅利叶变换单元 280。求出差频谱的总和（以下称差频谱功率），将其传送到频率增强单元 281。

反傅利叶变换单元 280，根据频谱稳定单元 279 得到的差频谱的幅和复数频谱的相位，构成新的复数频谱，用 FFT 进行反傅利叶变换。（把所得到的信号称为第 1 次输出信号）。然后，将所得到的第 1 次输出信号传送到频谱增强单元 281。

下面对频谱增强单元 281 的处理加以说明。

首先，参照噪声频谱存储单元 285 存储的平均噪声功率、频谱稳定单元 279 得到的差频谱功率、作为常数的噪声基准功率，选择 MA 增强系数与 AR 增强系数。选择根据对下述两个条件进行的评价进行。

条件 1

差频谱功率比噪声频谱存储单元 285 存储的平均噪声功率乘以 0.6 得到的积大，并且平均噪声功率比噪声基准功率大。

条件 2

差频谱功率比平均噪声功率大。

满足条件(1)时，以此作为“浊音区间”，取 MA 增强系数为 MA 增强系数 1

-1, 取 AR 增强系数为 AR 增强系数 1-1, 取高频增强系数为高频增强系数 1。而在不满足条件(1), 而满足条件(2)的情况下, 将其当作“清音区间”, 取 MA 增强系数为 MA 增强系数 1-0, 取 AR 增强系数为 AR 增强系数 1-0, 取高频增强系数为 0。又, 在不满足条件(1), 又不满足条件(2)的情况下, 以此作为“无声区间(只有噪声的区间)”, 取 MA 增强系数为 MA 增强系数 0, 取 AR 增强系数为 AR 增强系数 0, 取高频增强系数为高频增强系数 0。

然后, 使用 LPC 分析单元 276 得到的线性预测系数、上述 MA 增强系数、AR 增强系数, 根据下述式(56), 计算出极点增强滤波器的 MA 系数与 AR 系数。

$$\begin{aligned}\alpha(ma)_i &= \alpha_i * \beta^i \\ \alpha(ar)_i &= \alpha_i * \gamma^i\end{aligned}\quad (56)$$

$\alpha(ma)_i$ : MA 系数

$\alpha(ar)_i$ : AR 系数

$\alpha_i$ : 线性预测系数

$\beta$ : MA 增强系数

$\gamma$ : AR 增强系数

$i$ : 编号

然后, 对在反傅利叶变换单元 280 得到的第 1 次输出信号, 用上述 MA 系数与 AR 系数乘极点增强滤波器。此滤波器的传递函数示于下面的式(57)。

$$\frac{1 + \alpha(ma)_1 \times Z^{-1} + \alpha(ma)_2 \times Z^{-2} + \Lambda + \alpha(ma)_j \times Z^{-j}}{1 + \alpha(ar)_1 \times Z^{-1} + \alpha(ar)_2 \times Z^{-2} + \Lambda + \alpha(ar)_j \times Z^{-j}} \quad (57)$$

$\alpha(ma)_i$ : MA 系数

$\alpha(ar)_i$ : AR 系数

$j$ : 次数

进而, 为了增强高频成分, 用上述高频增强系数乘高频增强滤波器。此滤波器的传递函数示于下述式(58)。

$$1 - \delta Z^{-1} \quad \dots\dots(58)$$

$\delta$ : 为高频增强系数

上述处理得到的信号称为第 2 次输出信号。还有, 滤波器的状态保持于频谱增强单元 281 的内部。

最后，在波形匹配单元 282，利用三角窗使频谱增强单元 281 得到的第 2 次输出信号和前波形存储单元 288 存储的信号重迭，得到输出信号。还把该输出信号的末尾首读数据长度份额的数据存储于前波形存储单元 288。这时的匹配方法示于下面的式(59).

$$\begin{aligned} O_j &= (j \times D_j + (L - j) \times Z_j) / L & (j = 0 \sim L - 1) \\ O_j &= D_j & (j = L \sim L + M - 1) \\ Z_j &= O_{M+1} & (j = 0 \sim L - 1) \end{aligned} \quad (59)$$

$O_j$ : 输出信号

$D_j$ : 第 2 次输出信号

$Z_j$ : 输出信号

$L$ : 首读数据长度

$M$ : 帧长度

这里需要注意的是，作为输出信号，输出首读数据长度+帧长度份额的数据，但是，其中能够作为信号处理的只有从数据的始端起，长度等于帧长度的区间。这是因为，后面的首读数据长度的数据在输出下一输出信号时被改写。但是，在输出信号的全部区间内连续性受到补偿，因此能够使用于 LPC 分析和滤波器分析等频率分析中。

采用这样的实施形态，在声音区间中和声音区间外都能够进行噪声频谱推定，即使是在搞不清楚声音在哪一个时间存在于全部数据的情况下，也能够推定噪声频谱。

此外，可以用线性预测系数增强输入的频谱包络的特征，即使是在噪声电平高的情况下也能防止音质劣化。

还可以从平均和最低两个方向推定噪声的频谱，因而能够进行更恰当的降噪处理。

又，将噪声的平均频谱用于降噪处理，可以在更大程度上削减噪声频谱，还可以另外推定补偿用频谱，以更恰当地进行补偿。

而且，可以使不含声音、只有噪声的区间的频谱平滑，因而能够防止同区间的频谱由于噪声的减小而由极端的频谱变动引起异常感觉。

还可以使补偿的频率成分具有随机性，将不削去而留下的噪声变换成听觉上异常感觉小的噪声。

又，在声音区间，可以实施在听觉上更恰当的加权，在无声音的区间和清辅音区间，可以抑制由听觉加权引起的异常感觉。

### 工业应用性

如上所述，本发明的声源矢量生成装置，声音编码装置和声音解码装置对于声源矢量检索是有用的，适合于提高音质。

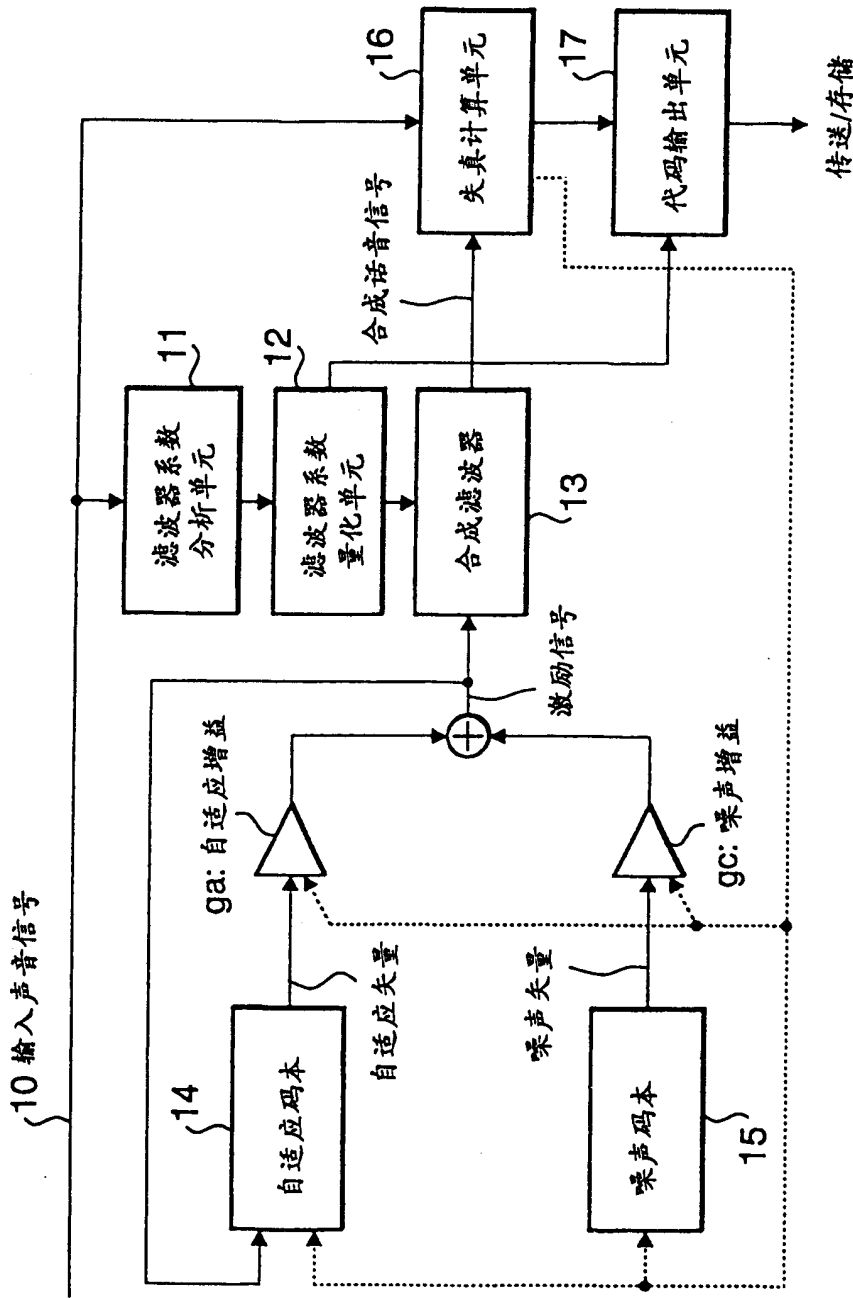


图 1

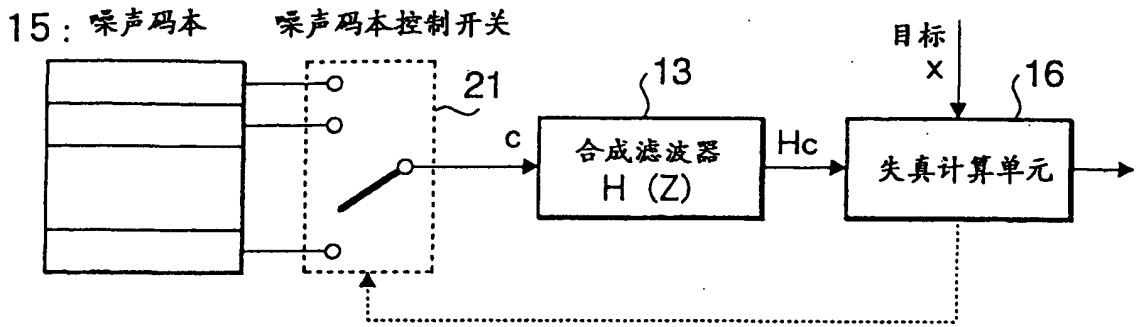


图 2A

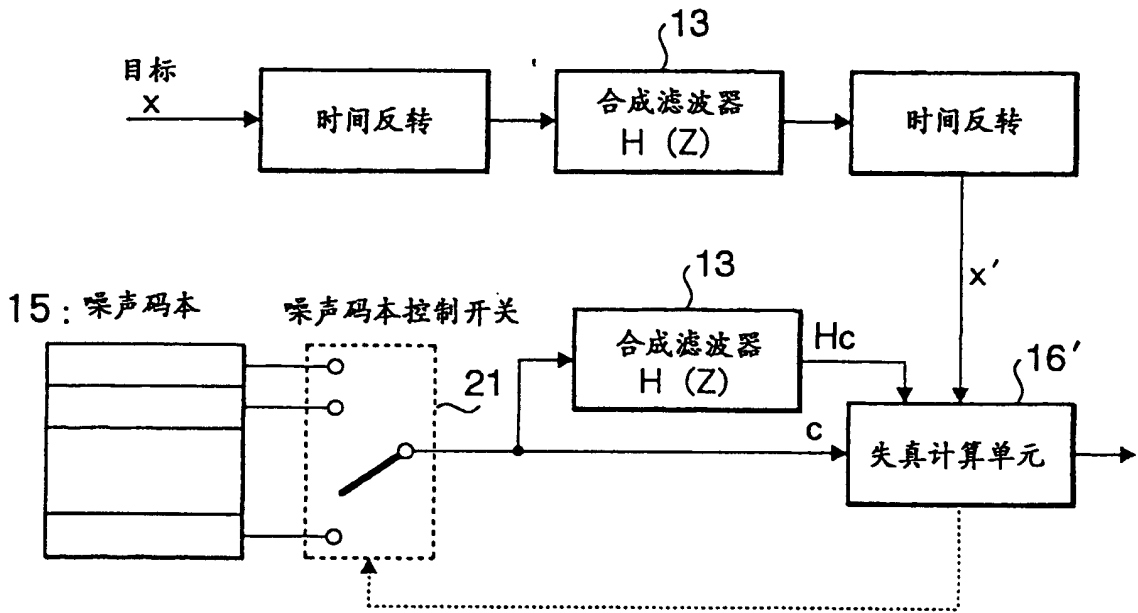


图 2B

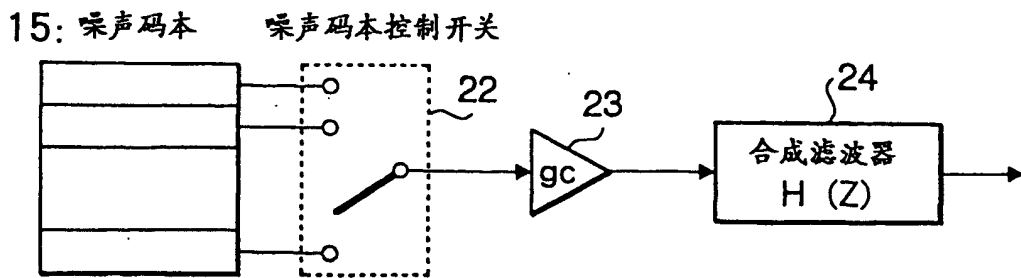


图 2C



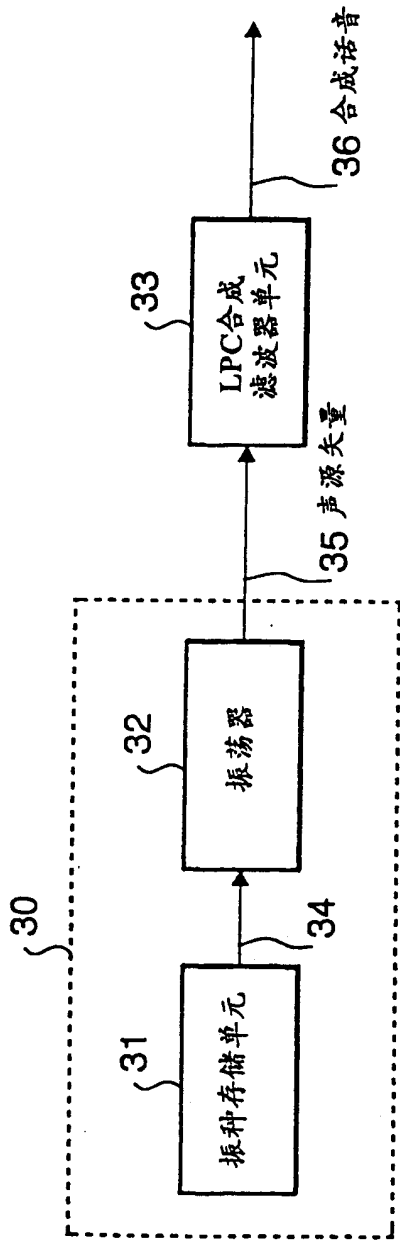


图 3

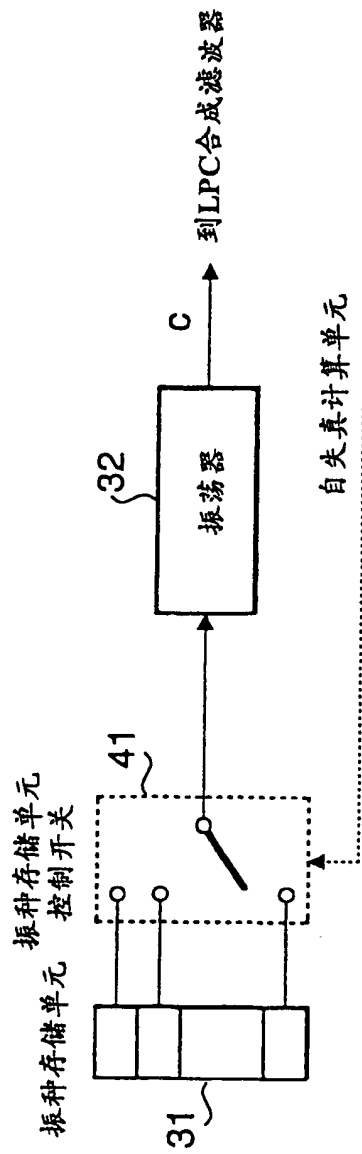


图 4

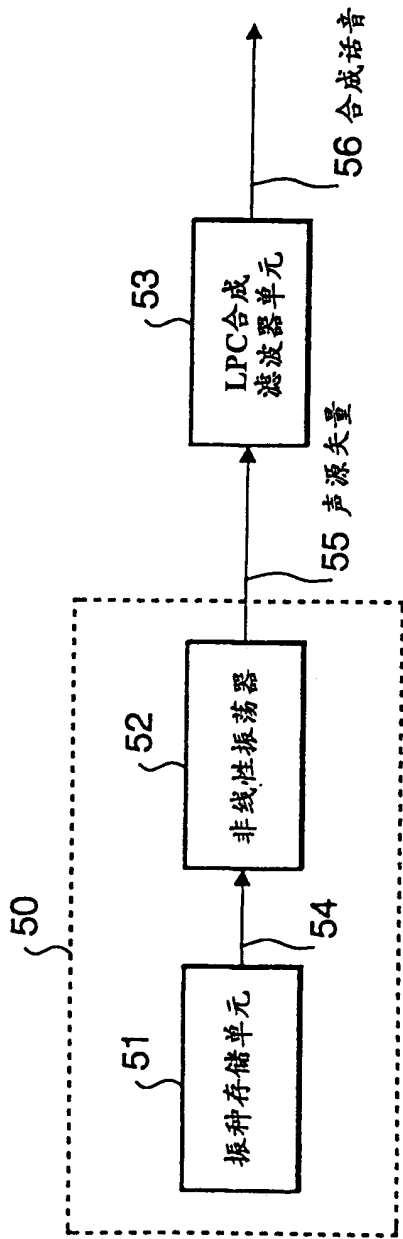


图 5

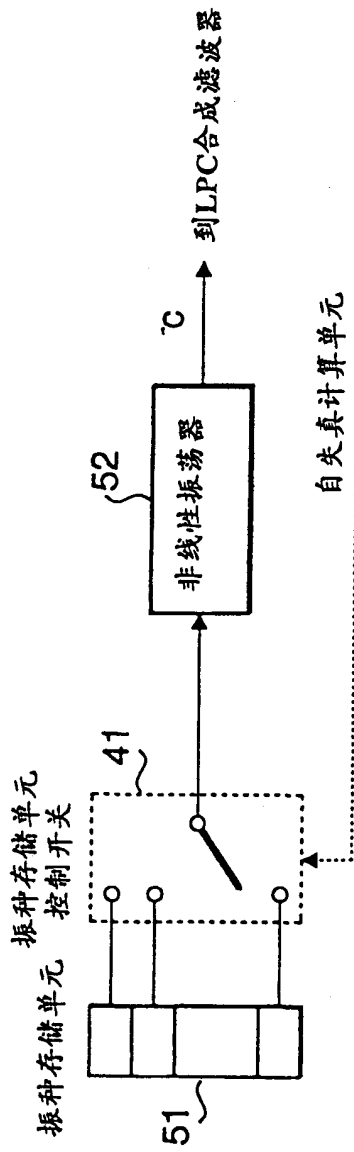


图 6

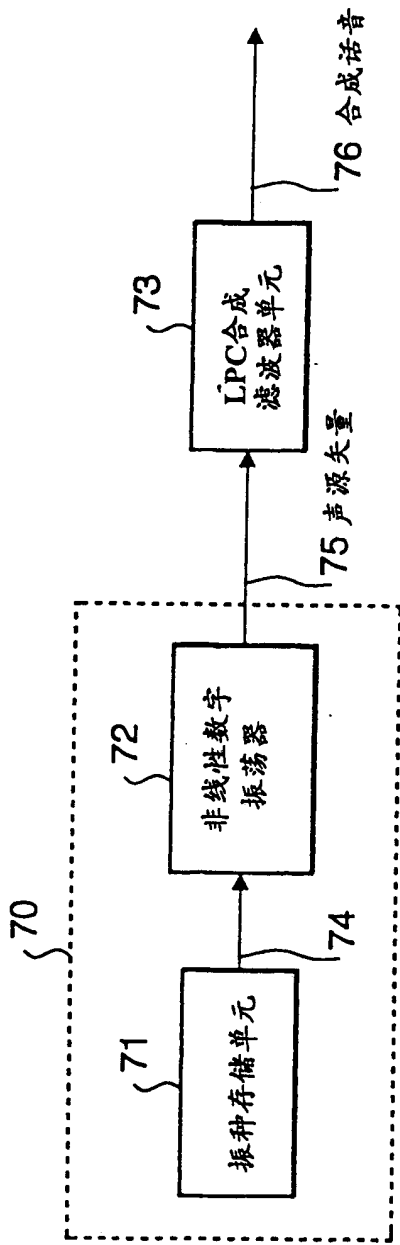


图 7

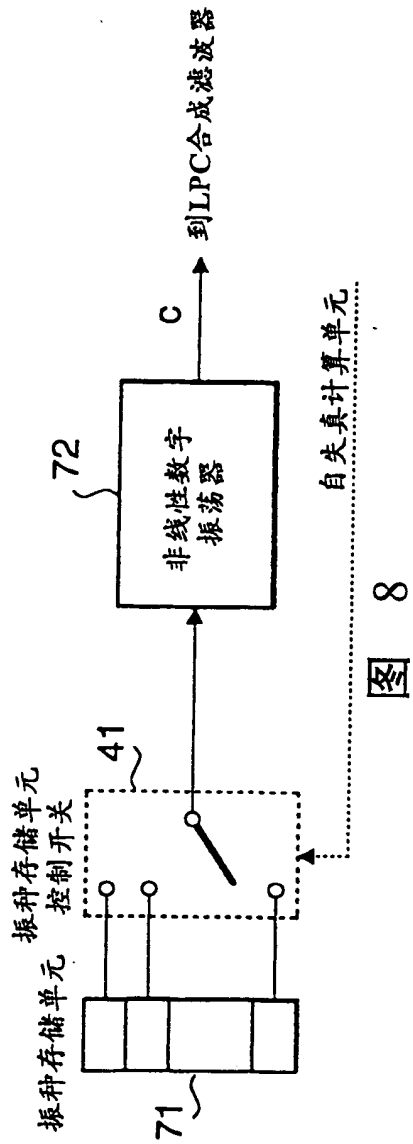


图 8

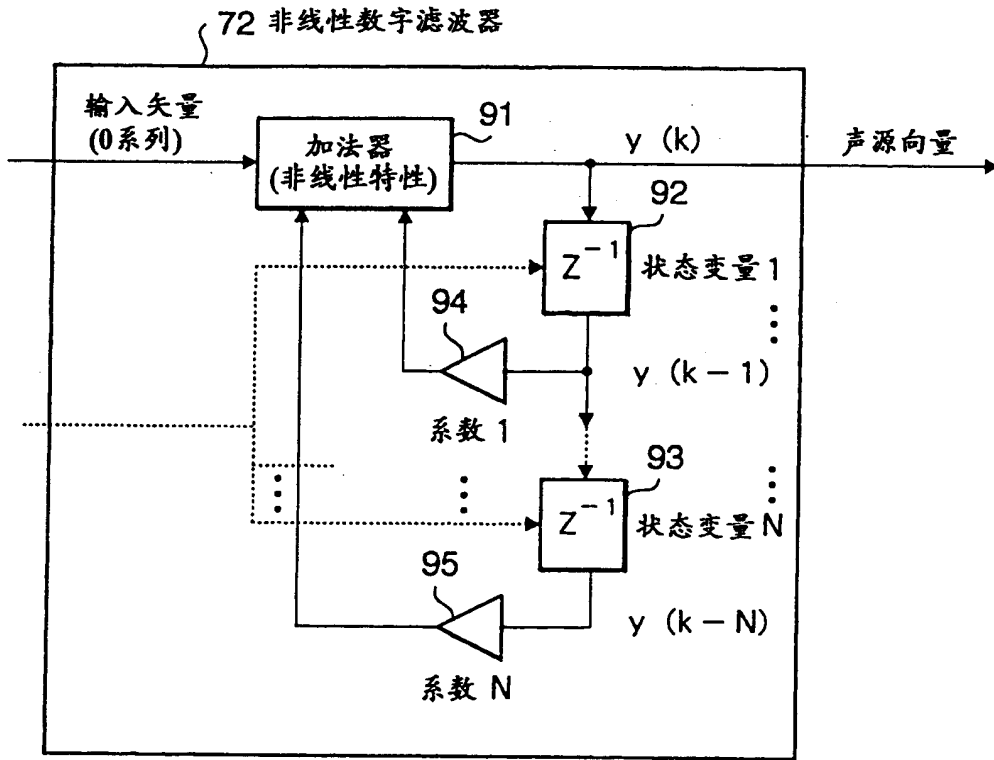


图 9

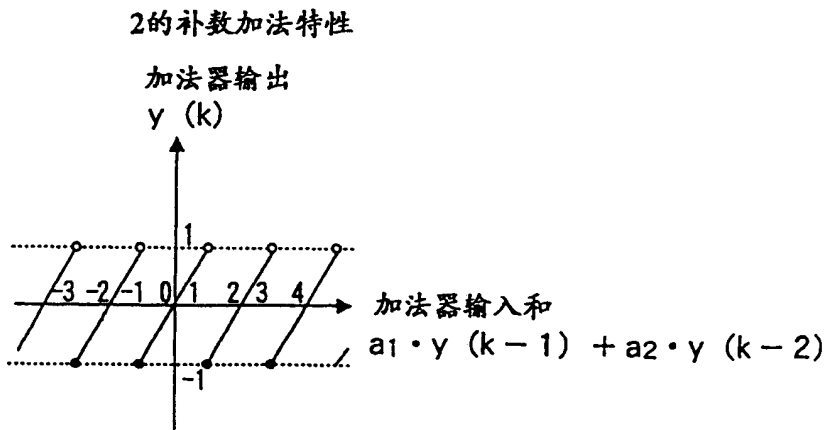


图 10

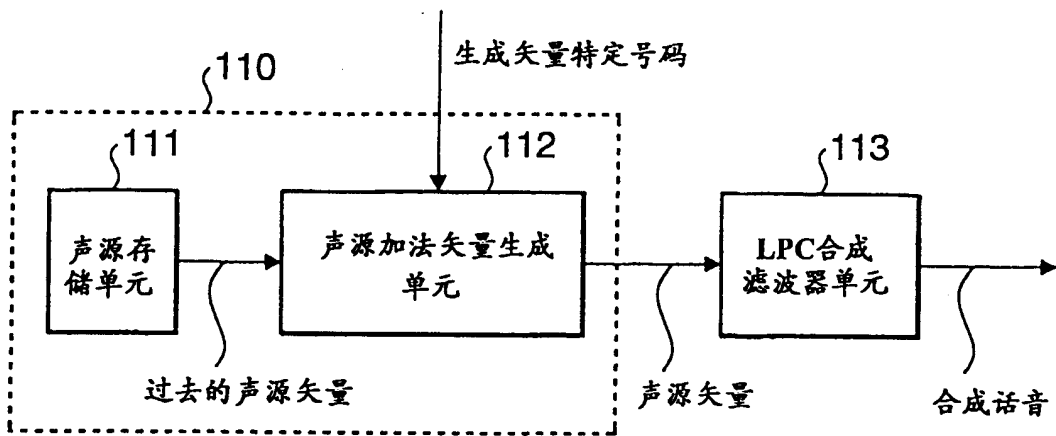


图 11

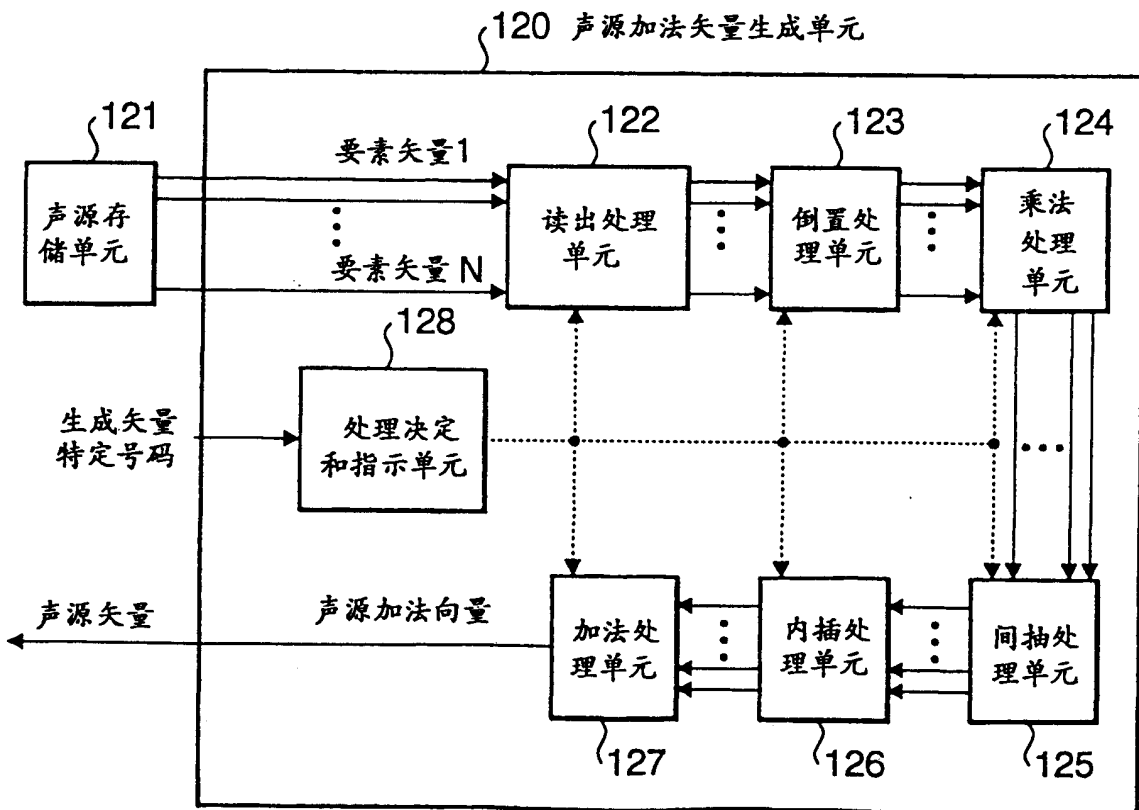


图 12

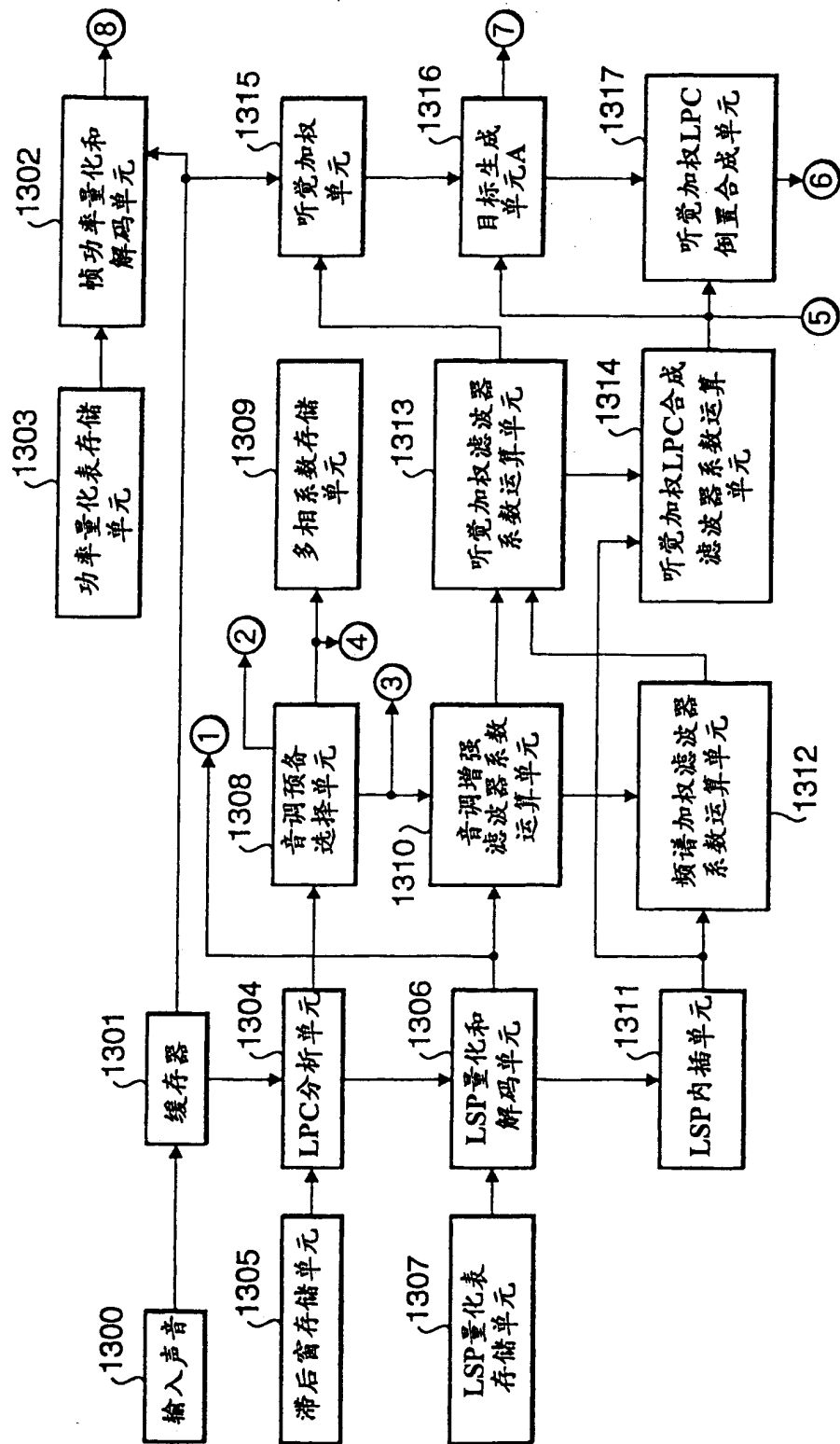


图 13A

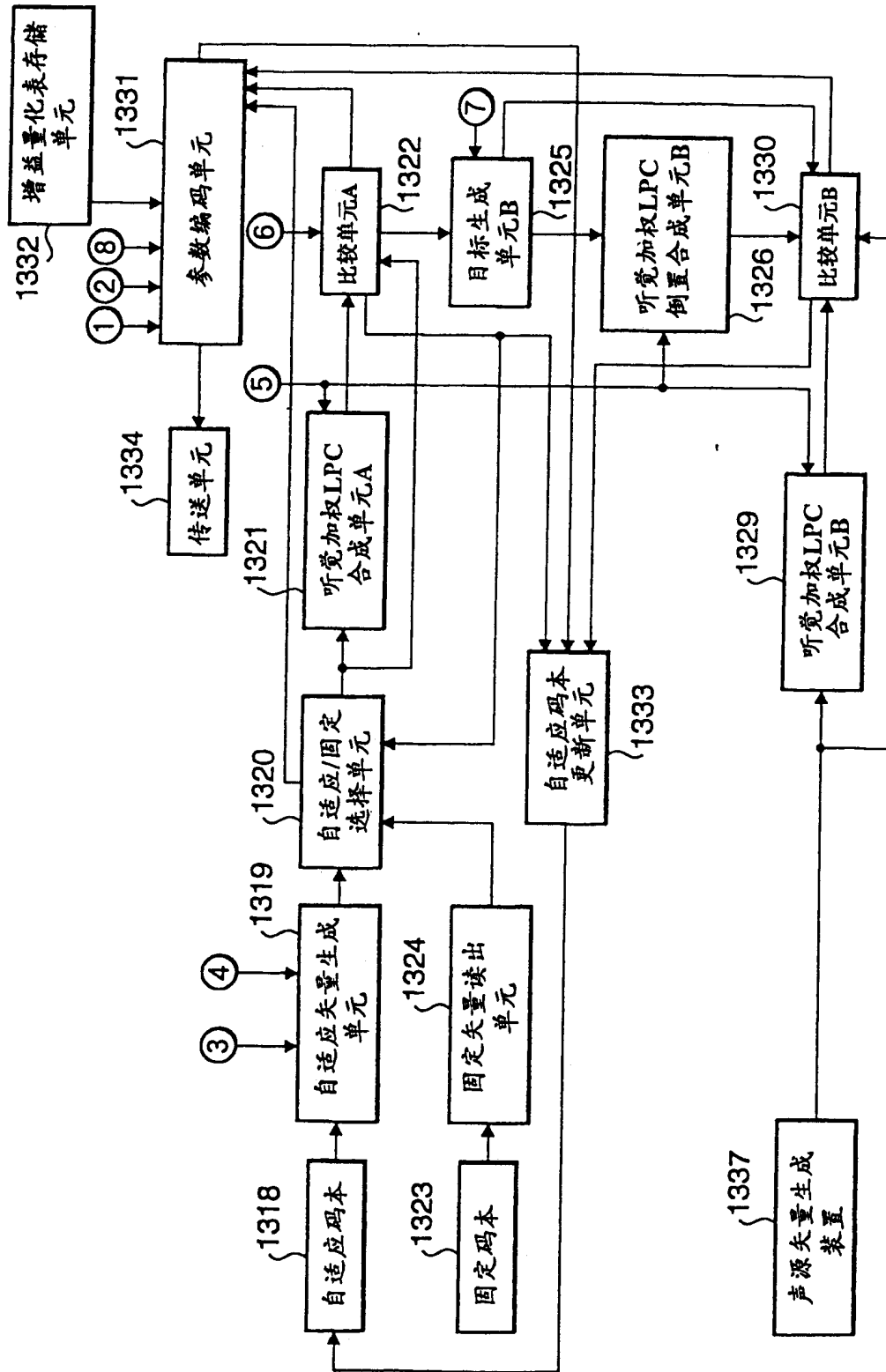


图 13B

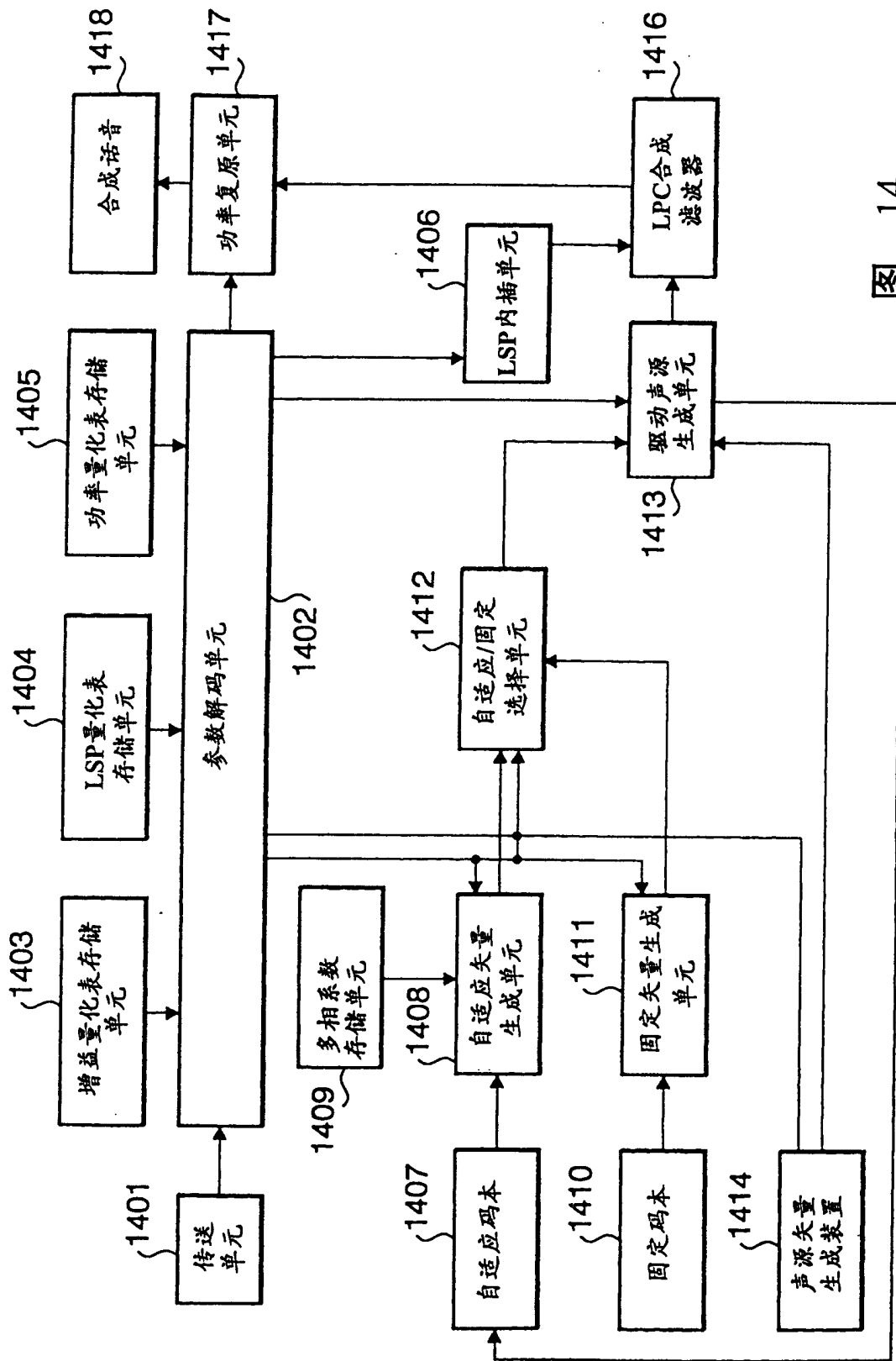


图 14



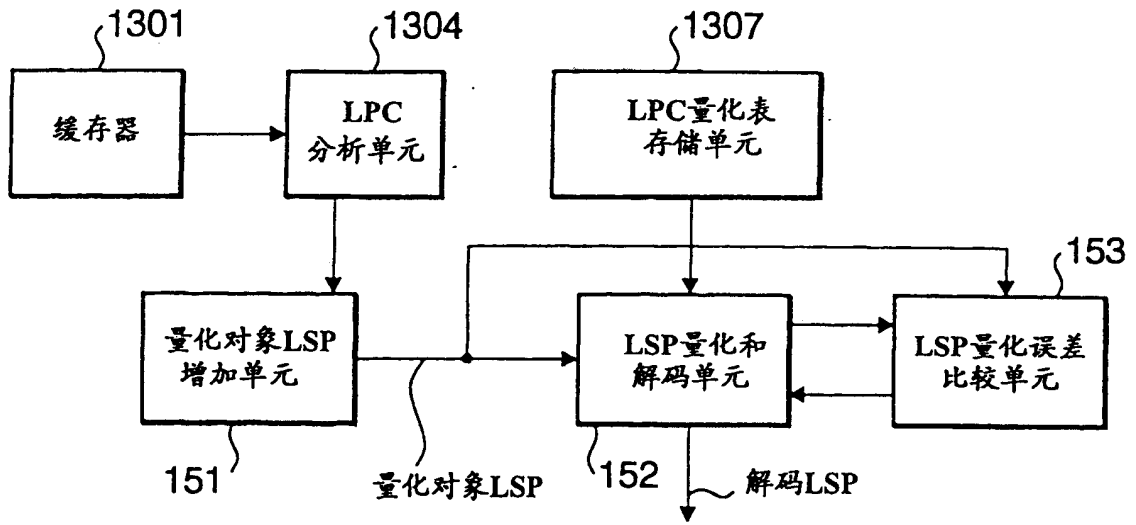


图 15

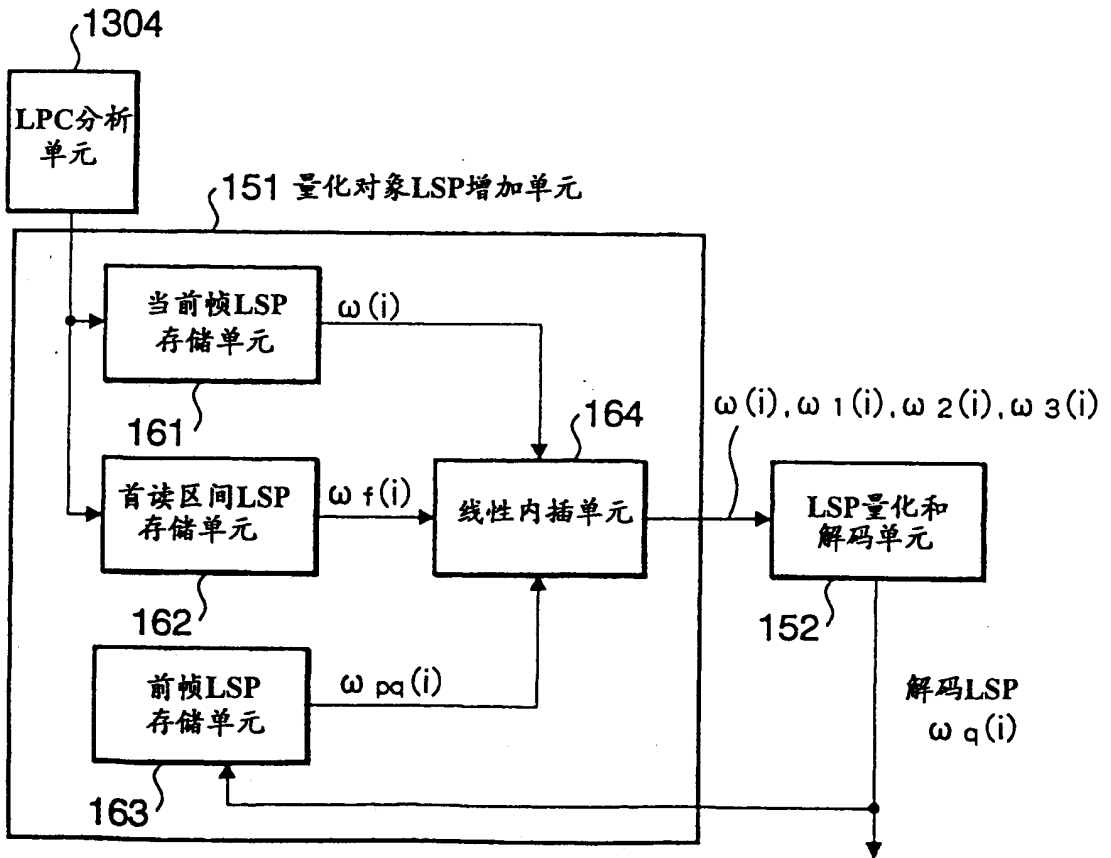


图 16

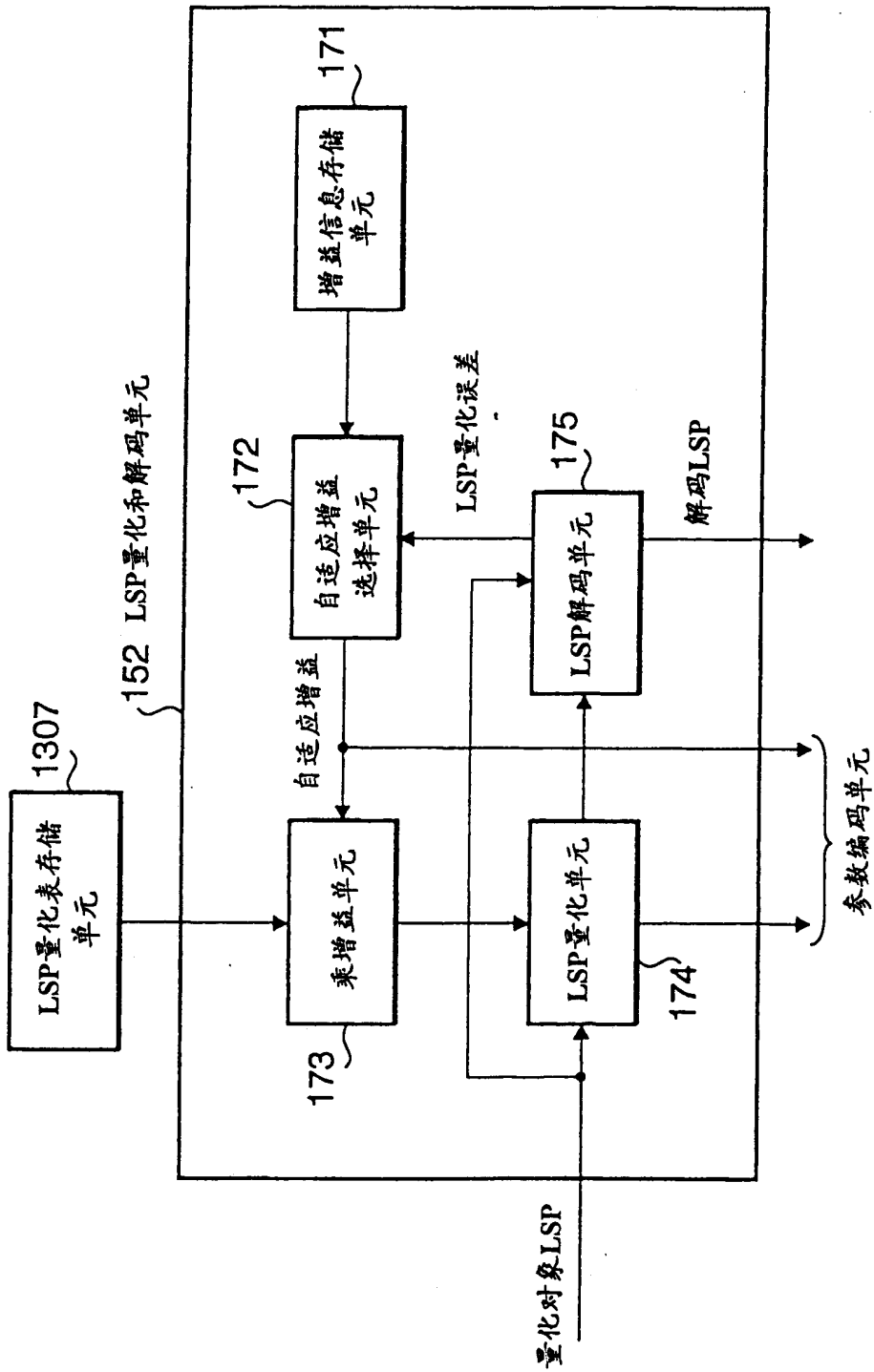


图 17

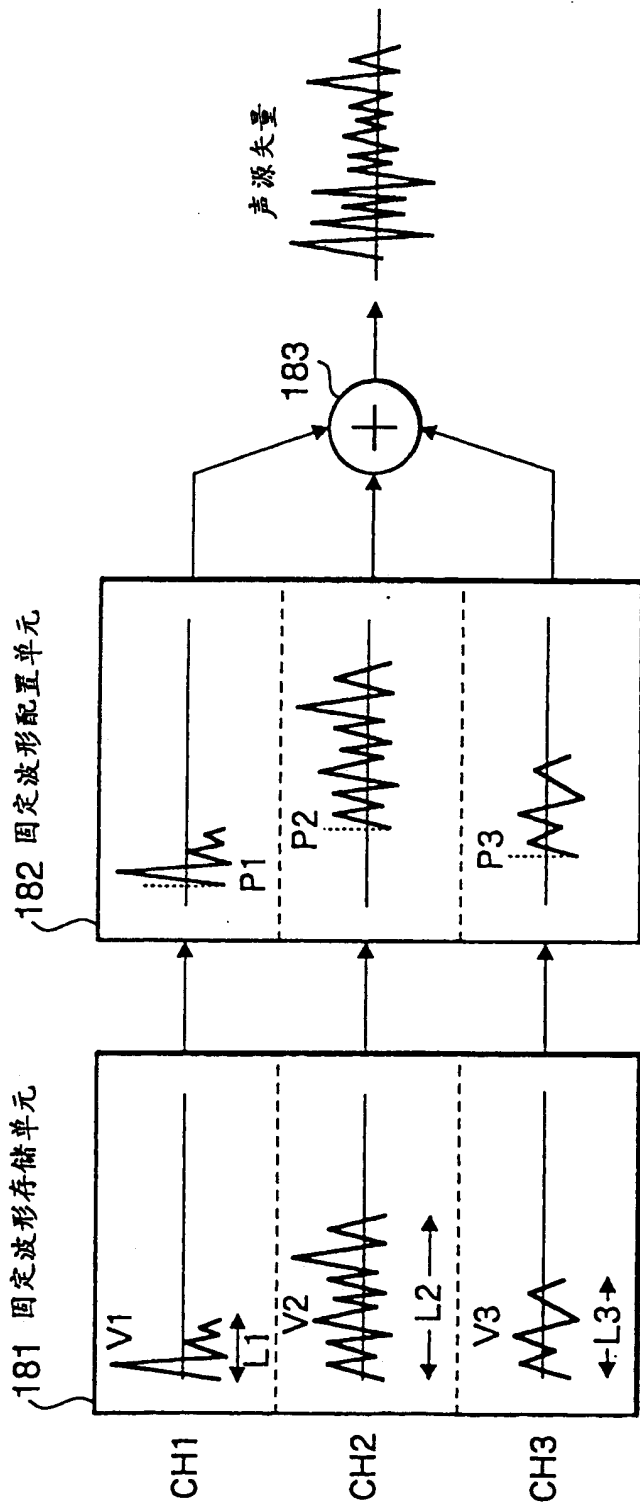


图 18

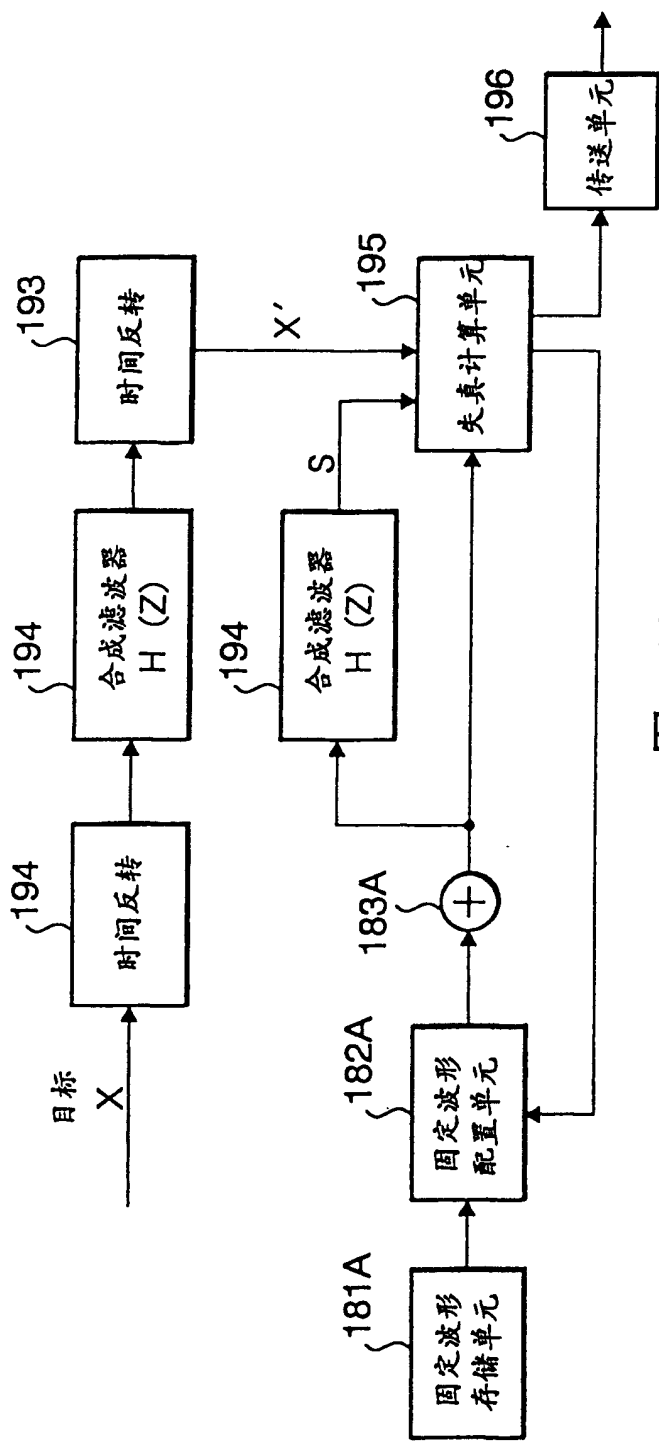


图 19A

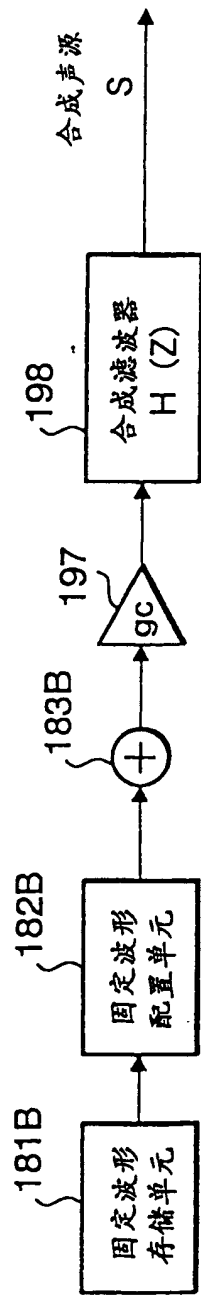


图 19B

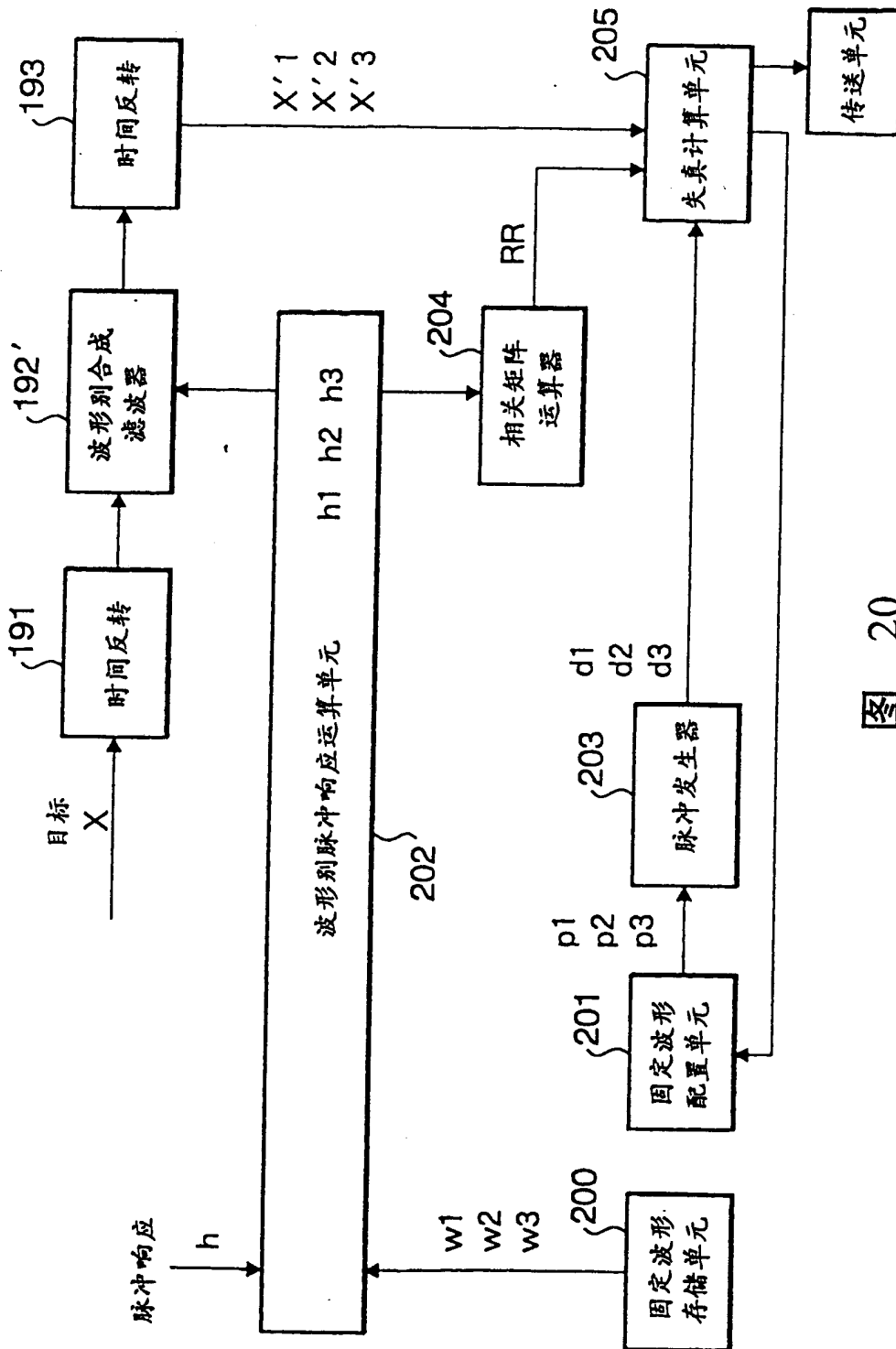


图 20

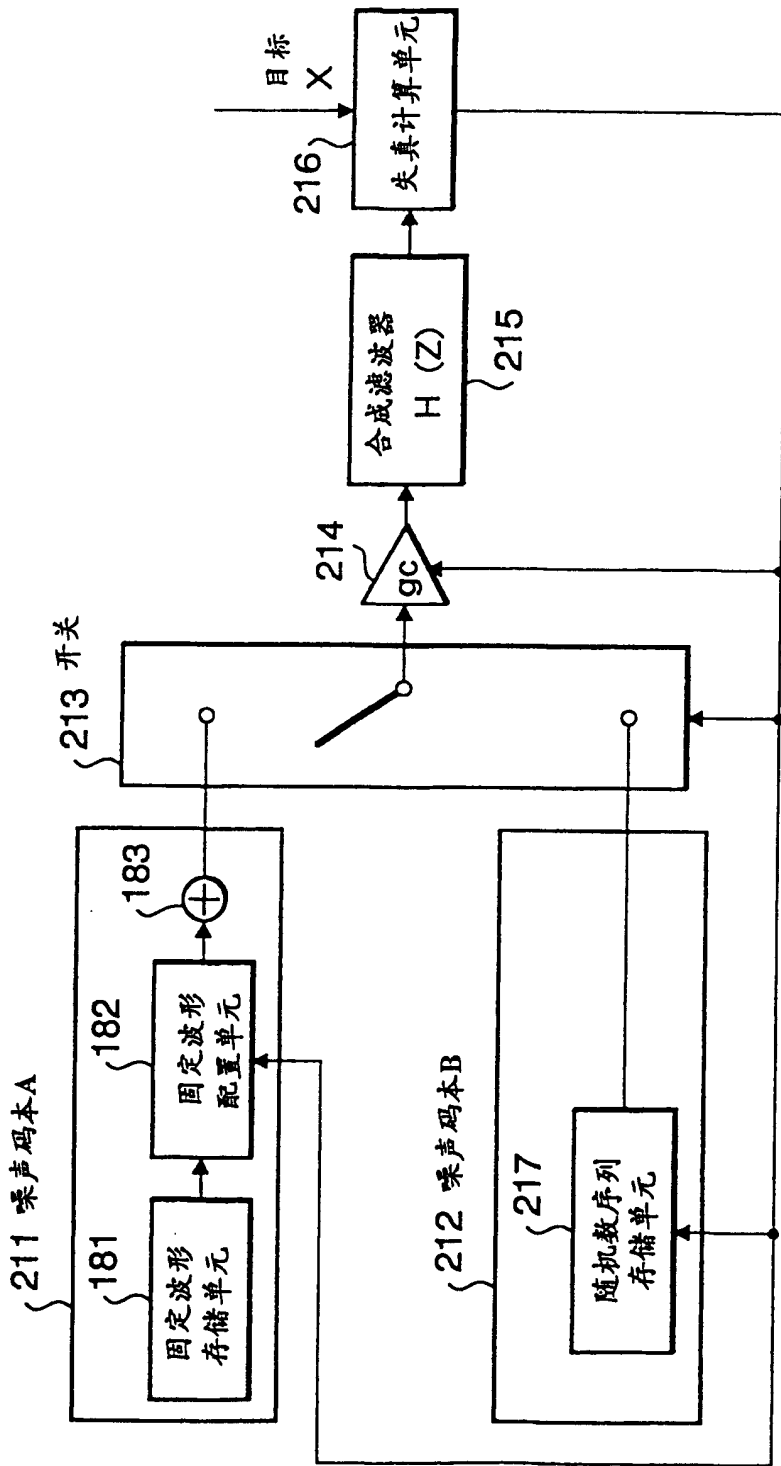


图 21

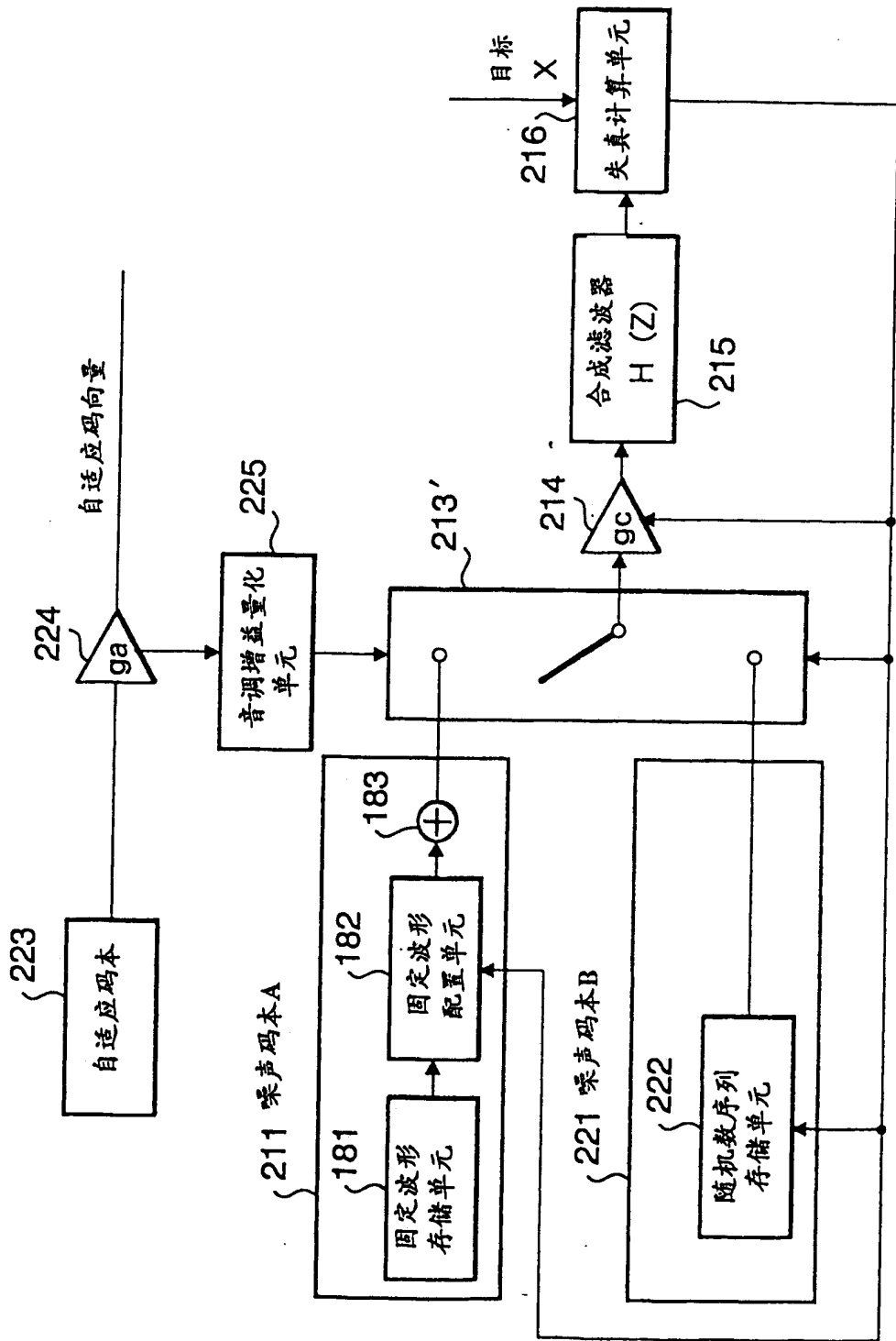


图 22

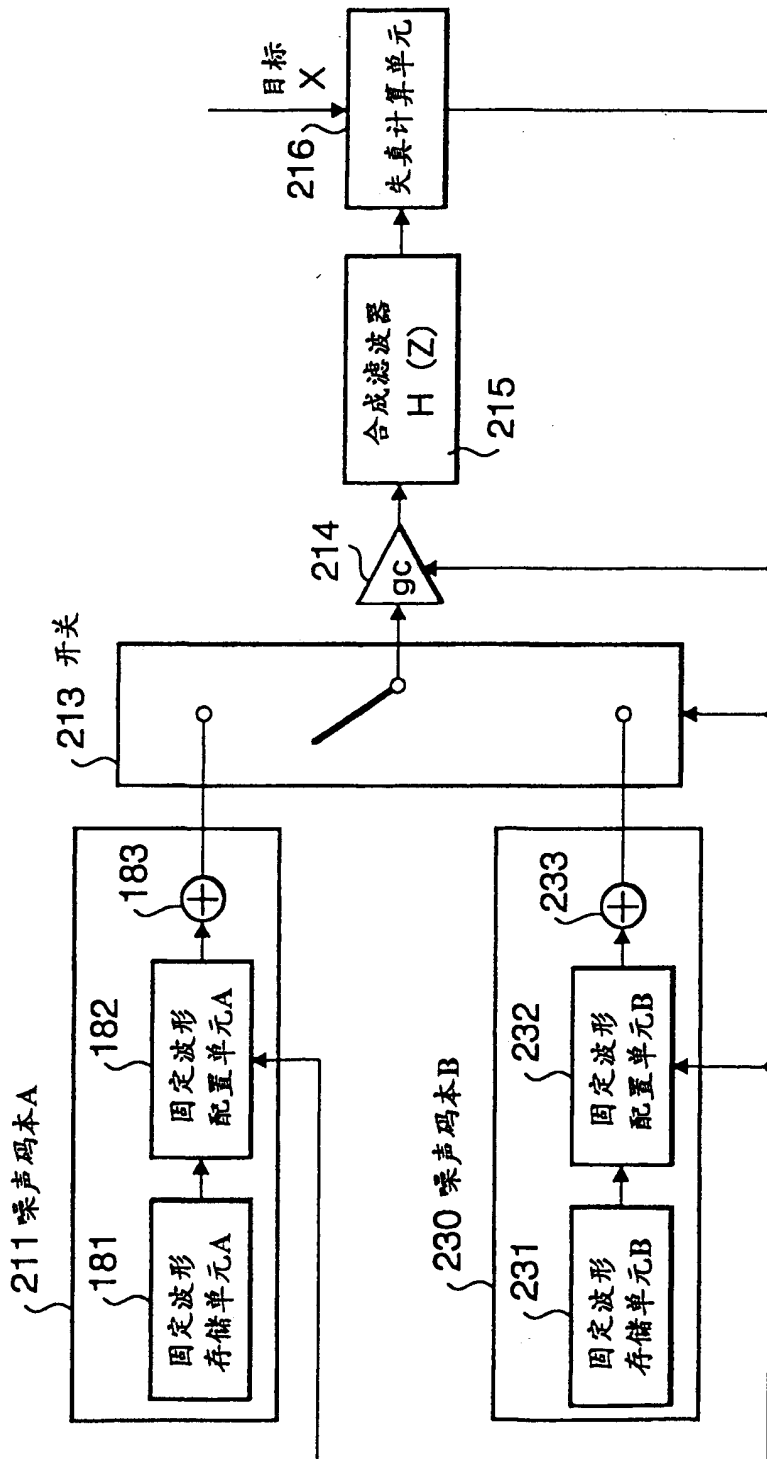


图 23



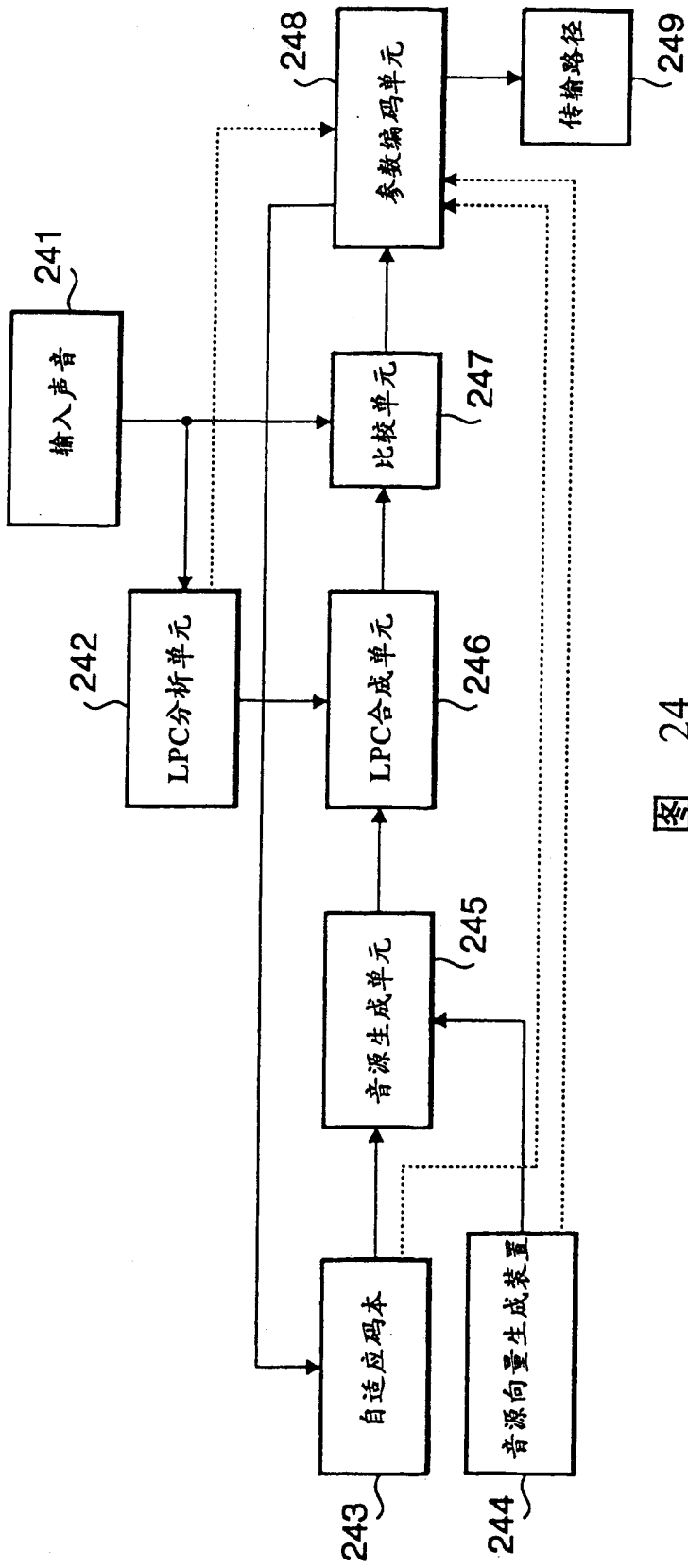


图 24

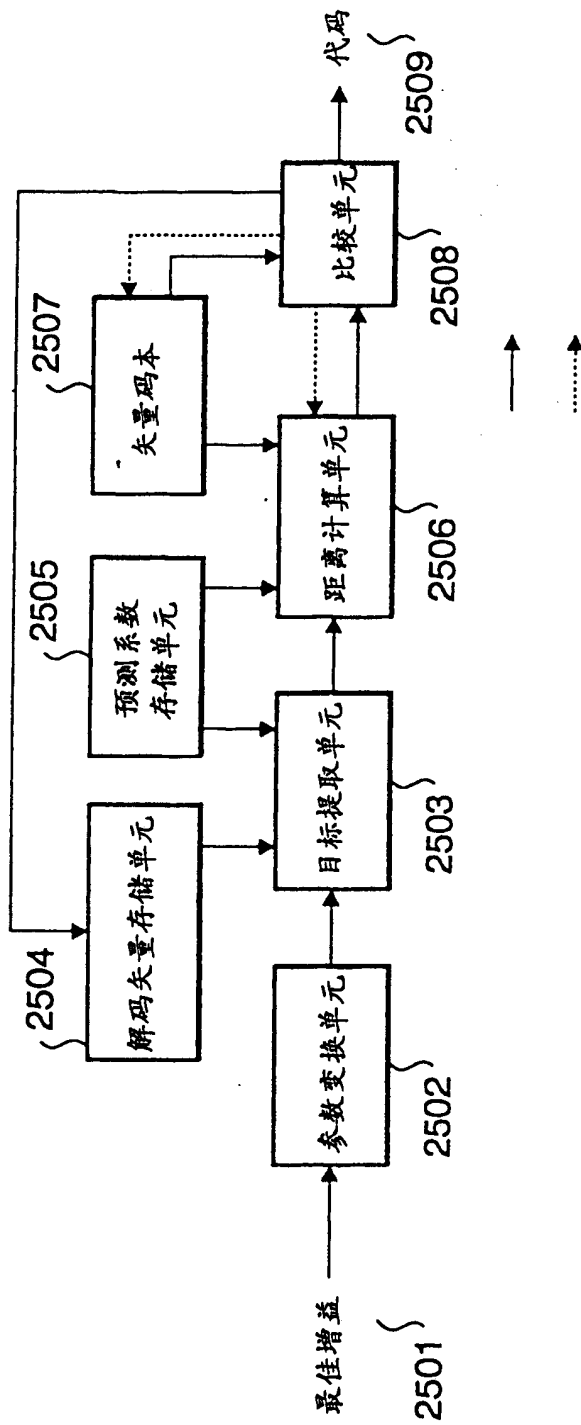


图 25

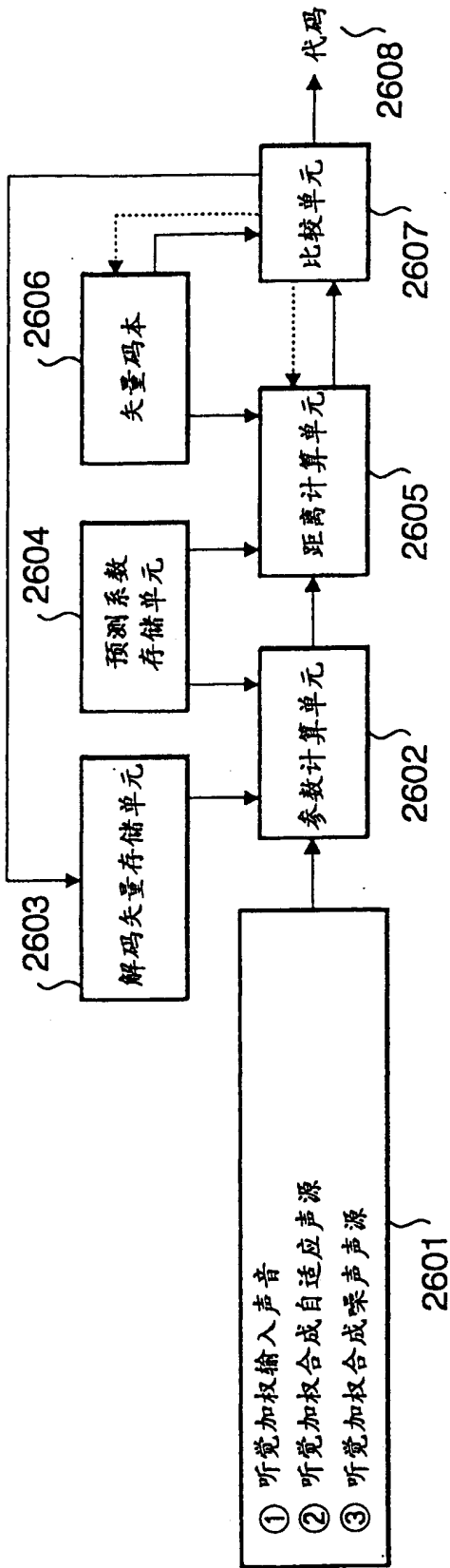


图 26

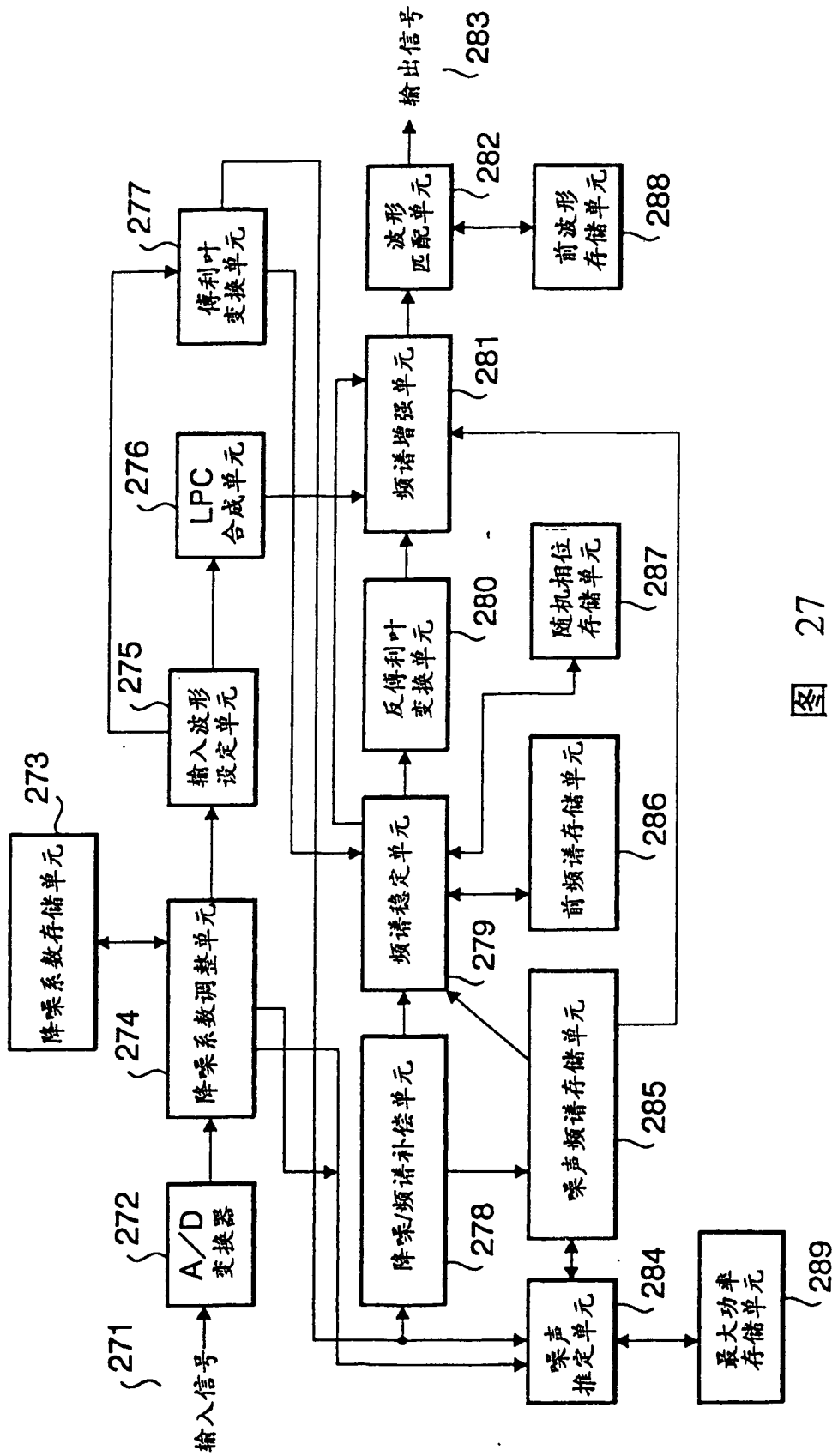


图 27