[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G10L 19/04 (2006.01)

[12] 发明专利说明书

专利号 ZL 96121992.0

[45] 授权公告日 2008 年 8 月 27 日

[11] 授权公告号 CN 100414605C

[22] 申请日 1996.10.26 [21] 申请号 96121992.0

[30] 优先权

[32] 1995.10.26 [33] JP [31] 279412/95

[73] 专利权人 索尼公司

地址 日本东京都

[72] 发明人 西口正之 饭岛和幸 松本淳 大森士郎

[56] 参考文献

CN1071036A 1993.4.14

审查员 马红梅

[74] 专利代理机构 北京市柳沈律师事务所

代理人 周少杰

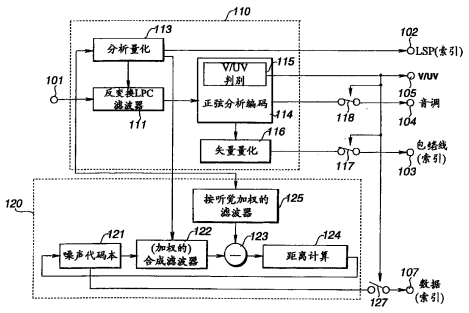
权利要求书 2 页 说明书 45 页 附图 14 页

[54] 发明名称

语音编码方法和装置

[57] 摘要

一种语声编码方法和装置，其中编码数据的比特速率是可变的。 利用第一矢量量化单元 500 的矢量量化单元 502 对输出的矢量$\underline{X}$进行加权的矢量量化。 在输出端 503 输出波形索引，在加法器 505 由源矢量$\underline{X}$减去量化值$\underline{X}_0$'。 形成的量化误差矢量$\underline{Y}$利用第二矢量量化单元 510 按维进行划分。 在输出端 512_1，512_2输出形成的加权的矢量量化的波形索引数据。 量化值$\underline{Y}_1$'和$\underline{Y}_2$'按维组合并利用加法器 513 相加到量化值$\underline{X}_0$'上。 输出形成的量化值$\underline{X}_1$。



1.一种语音编码方法，用于将输入的语音信号沿时间轴分成单位数据块，并对形成的信号进行编码，包含的步骤是:

至少对于输入的语音信号的发声部分求出短期预测的余部;

根据经此求出的短期预测的余部求出正弦分析编码的参数；以及

根据正弦分析编码的参数进行按听觉加权的矢量量化。

2.如权利要求 1 所述的语音编码方法，还包含步骤:

利用波形编码对输入的语音信号的不发声部分编码。

3. 如权利要求 2 所述的语音编码方法，其中还包含下述步骤：判别输入的语音信号是发声还是不发声部分，根据判别的结果，对发现是输入的语音信号的发声部分利用所述的正弦分析编码进行处理；并且对发现是输入的语音信号不发声部分利用合成法分析通过闭环查找最佳矢量进行矢量量化。

4. 如权利要求 2 所述的语音编码方法，其中对代表频谱包络线即用作正弦分析编码参数的数据进行所述按听觉加权的矢量量化。

5. 一种语音编码装置，用于将输入的语音信号沿时间轴分成单位数据块，并且对形成的信号进行编码，包含:

一个装置，用于求出至少输入的语音信号的短期预测的余部;

一个装置，用于根据经此求出的短期预测的余部求出正弦分析编码的参数;

一个装置，用于根据正弦分析编码的参数进行按听觉加权的矢量量化;以及

一个装置，用于通过波形编码对输入的语音信号的不发声部分进行编码。

6. 如权利要求 1 所述的语音编码方法，至少包含:

第一矢量量化步骤；以及

第二矢量量化步骤，对在所述第一矢量量化时产生的量化误差矢量进行量化。

7. 如权利要求 6 所述的语音编码方法，其中对于低的比特速率，取出在第一矢量量化步骤中的输出，并且对于高的比特速率，取出所述第一失量量化步骤的输出和所述第二矢量量化步骤的输出。

8. 一种语音编码装置，用于将输入的语音信号沿时间轴分成单位数据块

并对形成的信号进行编码，包含:

一个装置，用于求出输入的语音信号的短期预测的余部，

一个装置，用于根据经此求出的短期预测的余部求出正弦分析编码的参数；以及

一个装置，根据正弦分析编码的参数，进行按听觉加权的多级矢量量化。

9. 一种便携式无线电终端装置，包含:

放大装置，用于放大输入的语音信号;

A/D 变换装置；用于对所述放大的信号进行 A／D 变换;

语音编码装置，用于对所述 A/D 变换装置的语音输出信号进行编码;

传输通道编码装置，用于对形成的编码信号进行通道编码;

调制装置，用于调制所述传输通道编码装置的输出;

D/A 变换装置，用于对形成的调制信号进行 D/A 变换；以及

放大装置，用于放大来自所述 D/A 变换装置的信号，将所形成的放大信号送到天线;

所述语音编码装置还包含:

一个装置，用于求出所述输入的语音信号的短期预测的余部;

一个装置，用于根据经此求出的短期预测的余部求出正弦分析编码的参数;

一个装置，用于根据所述的正弦分析编码的参数进行按听觉加权的矢量量化；以及

一个装置，用于利用波形编码对所述输入的语音信号进行编码。

语音编码方法和装置

本发明涉及一种语音编码方法和装置，其中输入的语音信号按数据块划分并以所形成的数据块作为一个单位进行编码。

迄今为止已有对于声音信号（包括语音和音响信号）编码的各种编码方法，利用在时域和在频域中的信号的统计特性和人耳的心理声学特性进行压缩。编码方法可以粗分为时域编码、频域编码和分析/合成编码。

对语音信号高效编码的实例包括正弦分析编码，例如谐波编码、多频带激励（MBE）编码、副频带编码（SBT）、线性预测编码（LPC）、离散余弦变换（DCT）、改进的DCT（MDCT）和快速傅里叶变换（FFT）。

利用正弦分析编码，利用采用对语音信号的高效编码的语音信号编码装置，对于短期预测的余部，例如线性预测的编码（LPC）的余部进行编码，并且所形成的频谱包络线的幅值数据进行矢量量化，用以输出代码本索引（index）数据。

利用上述语音信号编码装置，包括矢量量化的代码本索引的编码数据的比特速率维持恒定，而且不能改变。

此外，如果编码数据是，例如M比特的，则用于将编码数据解码的语音信号解码装置必须是M比特的解码装置。就是说，利用这种语音信号解码装置，只能得到与编码数据具有相同比特数的解码数据，同时解码数据的比特数是不能改变的。

因此，本发明的目的是提供一种语音编码方法和装置，利用它们可以改变编码数据的比特速率。

利用本发明的语音编码方法和装置，至少对于输入的语音信号的发声部分求出短期预测的余部，并根据该短期预测的余部求出正弦分析编码参数。利用按听觉加权的矢量量化对这些正弦分析编码参数进行量化。利用具有相位再现性的波形编码对输入的语音信号的不发声部分进行编码。按照按听觉加权的矢量量化，首先进行第一矢量量化，然后通过第二矢量量化，对在第一矢量量化

时产生量化误差矢量进行量化。按照这种方式，输出的已编码的数据的比特数可以根据数据传输通道的容量简便地进行转换，使得可以适应各种数据比特速率，此外，可以产生这样一种编码数据序列，这种序列可以易于适应解码器方，即使在编码器和解码器之间的比特传输速率不同。

图1 是表示语音信号编码方法和用以实施本发明的编码方法的语音信号编码装置（编码器）的基本结构的方块图。

图2 是表示语音信号解码装置（解码器）的基本结构的方块图，该装置是针对图1 所示装置的解码装置（解码器）。

图3 是表示图1 中所示的语音信号编码器的更详细的结构的方块图。

图4 是表示图2 中所示的语音信号编码器的更详细的结构的方块图。

图5 是表示L P C 量化器的基本结构的方块图。

图6 是表示L P C 量化器的更详细的结构的方块图。

图7 是表示矢量量化器的基本结构的方块图。

图8 是表示矢量量化器的更详细的结构的方块图。

图9 是表示本发明的语音信号编码器的C E L P 编码部分（第二编码单元）的更详细的结构的示意电路方块图。

图1 0 是描述在图9 所示配置中的处理程序的流程图。

图1 1 A和1 1 B 表示按照不同的阈值进行限制之后的高斯噪声。

图1 2 是表示通过学习在产生波形（s h a p e ）代码本时的处理程序的流程图。

图1 3 是表示采用按本发明体现的语音信号编码器的便携式终端的发送侧的结构的方块图。

图1 4 是表示采用与图1 3 所示的装置成对的语音信号解码器的便携式终端的接收侧的结构的方块图。

图1 5 是表示在本发明的语音信号编码器中对于不同的比特速率的输出数据。

下面参照附图，详细介绍本发明的各优选实施例。

图1 表示用于实施本发明的语音编码方法的语音信号编码器的基本结构的方块图。该语音信号编码器包括：一反变换L P C 滤波器1 1 ，作为用于求出输入的语音信号的短期预测的余部的装置，以及正弦分析编码器1 1 4 ，作为

用于从短期预测的余部求出正弦分析编码的参数的装置。该语音信号编码器还包括：矢量量化单元116，作为用于根据正弦分析编码参数进行按听觉加权的矢量量化的装置，以及第二编码单元120，作为用于利用相位传输波形编码对输入的语音信号进行编码的装置。

图2 是表示与在图1 中所示的编码装置（编码器）配套的语音解码装置（解码器），图3 是表示图1 中所示的语音信号编码器的更详细的结构的方块图，图4 是表示图2 中所示的语音解码器的更详细的结构的方块图。

下面解释图1 到图4 各方块图的结构。

图1 中所示的语音信号编码器的基本概念在于，该编码器具有第一编码单元110，用于求出 输入的语音信号的短期预测的余部，例如线性预测编码（LPC）的余部，以便进行正弦分析编码、例如谐波编码；以及第二编码单元120，用于利用具有相位再现性的波形编码对输入的语声信号进行编码，以及第一和第二编码单元110、120 分别用于对输入信号的发声部分和不发声部分进行编码。

第一编码单元110 具有利用正弦分析编码，例如谐波编码或多频带编码（MBE）对LPC 的余部进行编码的结构。第二编码单元120 具有利用合成法进行分析，通过闭环查找最佳矢量，采用矢量量化的代码激励线性预测的结构。

在该实施例中，传送到输入端101 上的语音信号被送到第一编码单元110 的反变换LPC 滤波器111 和LPC 分析/量化单元113。由LPC 分析/ 量化单元113 得到的LPC 系数或所谓的α参数送到反变换LPC 滤波器111，用以利用反变换LPC 滤波器111 取出输入的语音信号的线性预测的余部（LPC 余部）。由LPC 合成/ 量化单元113 取出线性频谱对（LSP）的量化输出并送到输出端102。来自反变换LPC 滤波器111 的LPC 的余部送到正弦分析编码单元114。正弦分析编码单元114 进行音调检测，频谱包络线幅值计算和利用发声（V）/ 不发声（UV）鉴别单元115 进行V /UV 鉴别。来自正弦分析编码单元114 的频谱包络线幅值数据送到矢量量化单元116。来自矢量量化单元116 的作为频谱包络线的矢量量化输出的代码本索引经过开关117 送到输出端103，而正弦分析编码单元114 的输出经过开关118 送到输出端104。来自V / UV 鉴别单

元115 的V / UV鉴别输出送到输出端105 和作为转换控制信号送到开关117、118。对于发声（V）信号，选择索引和音调，以便在输出端103、104 取出。

在本实施例中，图1 中的第二编码单元120 具有一种代码激励线性预测（CELP）编码结构，利用合成法进行分析采用闭环查找对时域的波形进行矢量量化。其中噪声代码本121 的输出由加权的合成滤波器122 合成，所形成的加权的语音信号送到减法器123，取出在加权的语音信号和传送到输入端101 并通过按听觉加权滤波器125 的语音信号之间的差，送到距离计算电路124，以便进行距离计算和利用噪声代码本121 查找使误差最小的一个矢量，这种CELP 编码用于对如上所述的不发声部分进行编码。在输出端107 经过开关127 取出来自噪声代码本121 的作为UV数据的代码本索引，开关127 当来自V / UV鉴别单元115 表明为不发声（UV）声音部分的V /UV鉴别结果时接通。

图2 是表示作为图1 中的语音信号编码器的配套装置的语音信号解码器的基本结构的方块图，用于实施本发明的语音解码方法。

参阅图2，来自图1 中的输出端102 的作为线性频谱对（LSP）的量化输出传送到输入端202。图1 中的输出端103、104 和105 的输出即作为包络线量化输出的索引数据、音调和V / UV鉴别输出分别传输到输入端203 到205。作为用于不发声部分的数据的索引数据由图1 中的输出端107 提供到输入端207。

作为输入端203 的量化输出的索引送到用于反矢量量化的反矢量量化单元212，以便求出LPC的余部的频谱包络线，该包络线再送到发声部分的语音合成器211。发声部分的语音合成器211 利用正弦合成对发声语言部分的线性预测编码（LPC）的余部进行合成。还由输入端204、205 向发声部分的语音合成器211 提供音调和V /UV鉴别输出，来自发声部分语音合成单元 211 的发声部分语音的LPC的余部送到LPC合成滤波器214。来自输入端207 的UV数据的索引数据送到不发声部分的声音合成单元220，在其中必须参照一噪声代码本，用于取出不发声部分的LPC的余部。这些 LPC的余部还送到LPC合成滤波器214。在LPC合成滤波器214 中，通过LPC合成对发声部分的LPC的余部和不发声部分的LPC

的余部进行处理。另外，发声部分的LPC余部和不发声部分的LPC余部一起求和，也可以利用LPC合成进行处理。来自输入端202的LSP索引数据送到LPC参数再现单元213，在其中取出LPC的α参数并送到LPC合成滤波器214。在输出端201取出利用LPC合成滤波器214合成的语音信号。参阅图3，下面解释图1中所示的语音信号编码器的更详细的结构。在图3中与图1中所示相似的部分使用相同的标号来标注。

在图3中所示的语音编码器中，提供到输入端101的语音信号利用一用于除去不需要的范围的信号的高通滤波器进行滤波，并由其再提供到LPC分析/量化单元113的LPC分析电路132和提供到反变换LPC滤波器111。LPC分析/量化单元113的LPC分析电路132提供一个输入信号波形长度为按照256个采样量级作为一个数据块的汉明窗口，并利用自相关法求出线性预测系数即所谓的α参数。将作为数据输出单元的帧形成间隔设定接近160个采样。如果采样频率fs为8KHz（千赫），对于160个采样一帧间隔为20ms（毫秒）。

来自LPC分析电路132的α参数送到α-LSP变换电路133，用以变换为线性频谱对（LSP）参数。该数电路将α参数例如由直接型滤波器系数求出的变换为10个即5对LSP参数。例如利用牛顿—拉普松方法进行这种变换。将α参数变换为LSP参数的原因在于，LSP参数在内插特性方面比α参数更优越。

来自α-LSP变换电路133的LSP参数利用LSP量化器134进行矩阵或矢量量化。可以在进行矢量量化之前取得帧与帧的差，或者汇集多个帧，以便进行矩阵量化。在本实例中，取得每20ms计算的LSP参数的2个帧（20ms）并按照矩阵量化和矢量量化进行处理。

在连接端102取出量化器134的量化输出即LSP量化数据，同时将量化的LSP矢量送到LSP内插电路136。

LSP内插电路136内插每20ms或40ms量化的LSP矢量，以便提供8倍的速率。即LSP每2.5ms刷新LSP矢量。原因在于，如果利用谐波编码/解码法通过分析/合成对余部的波形进行处理，则合成波形的包络线表示极为平滑的波形，使得如果每20ms LPC系数急剧地改变，很可能会产生额外的噪声。即如果每2.5ms LPC系数逐渐地改变，就可

防止这样额外的噪声产生。

由于利用每2．5ms产生的内插的LSP矢量对输入的语音进行反变换滤波，利用一个由LSP向α变换电路137将LSP参数变换为α参数，作为例如10阶直接型滤波器的参数。由LSP向α变换电路137的输出送到LPC反变换滤波器电路111，该电路然后利用每2．5ms刷新的α参数进行反变换滤波，用以产生平滑的输出。反变换LPC滤波器111的输出送到例如为谐波编码电路的正弦分析编码单元114的正交变换电路145，例如DCT电路。

来自LPC分析/量化单元113的LPC分析电路132的α参数送到接听觉加权的滤波器计算电路139，在其中求出用于按听觉加权的数据。这些加权的数据送到按听觉加权的矢量量化器116，第二编码单元120的按听觉加权的滤波器125和按听觉加权的合成滤波器122。

谐波编码电路的正弦分析编码单元114利用谐波编码法分析反变换LPC滤波器111的输出。即进行音调检测，各个谐波的幅值Am的计算和发声（V）/不发声（UV）部分的鉴别，并且利用维的变换使随音调变化的各幅值Am的数目或各个谐波包络线的数目是恒定的。

在图3中所示的正弦分析编码单元114的说明性的实例中，采用平常的谐波编码。按照多频带激励（MBE）编码，在模型化时假设，发声部分和不发声部分出现在相同时间点（在相同的数据块或帧中）的频率范围或频带中。按照另一种谐波编码技术，按单值方式判别在一数据块中或一帧中的语音是发声部分还是不发声部分。在如下的介绍中，如果频带的总体是UV部分，则指定的帧判别为UV部分，这是就MBE编码而论的。

分别将来自输入端101的输入的语音信号和来自高通滤波器（HPF）109的信号提供到图3中的正弦分析编码单元114的开环音调查找单元141和过零计数量142。由反变换LPC滤波器111向正弦分析编码单元114的正交变换电路提供LPC的余部或线性预测的余部。开环音调查找单元141取得输入信号的LPC的余部，以便通过开环进行相对粗略音调查找。抽取的粗略音调数据送到下文将要介绍的通过闭环工作的细微音调查找单元146。由开环音调查找单元114取出通过使LPC余部的自相关最大值连同粗略音调数据归一化得到的归一化的自相关的r（p）的最大值以便粗略

音调数据一道送到V /U V 鉴别单元1 1 5 。

正交变换电路1 4 5 进行正交变换，例如离散的傅里叶变换，用以将沿时间轴的L P C 的余部变换成沿频率轴的频谱幅值数据。正交变换电路1 4 5 的输出送到细微音调查找单元1 4 6 和用于计算频谱幅值或包络线的频谱计算单元1 4 8 。

向细微音调查找单元1 4 6 提供有由开环音调查找单元1 4 1 抽取的相对粗略的音调数据和利用正交变换单元1 4 5 通过D F T 得到的频域数据。细微音调查找单元1 4 6 使该音调数据按照0 .2 到0 .5 的速率按中心围绕该粗略音调值数据摆动±几个采样，以便最大到达具有一个最佳十进制点（浮点）的细微音调数据的数值。将合成法分析用作为用于选择一个音调的细微查找技术，以便使该功率谱最接近原有信号的功率谱，来自闭环细微音调查找单元1 4 6 的音调数据经开关1 1 8 送到输出端1 0 4 。

在频谱计算单元1 4 8 ，根据作为L P C 余部的正交变换输出的频谱幅值和音调，计算每个谐波的幅值和作为各谐波之和的频谱包络线，并关到细微音调查找单元（4 6 、V / U V 鉴别单元1 1 5 和按听觉加权的矢量量化单元1 1 6 。

V /U V 鉴别电路1 1 5 根据正交变换电路1 4 5 的输出、来自细微音调查找单元1 4 6 的最佳音调、来自频谱计算单元1 4 8 的频谱幅值数据、来自开环音调查找单元1 4 1 的归一化的自相关的r （p ）和来自过零计数器1 4 2 的过零计数值鉴别一个帧的V /U V 部分。此外，还可以将对于M B E 的以频带为基准的V /U V 鉴别的边界位置用作V /U V 鉴别的条件。在输出端1 0 5 取出V /U V 鉴别单元1 1 5 的鉴别输出。

向频谱计算单元1 4 8 中的输出单元或矢量量化单元1 1 6 的输入单元提供一个数据数目变换单元（进行一种采样速率变换的单元）。数据数目变换单元用于设定包络线的幅值数据| A m | ，在于考虑到这样一个事实，频带的数目沿频率轴划分以及数据的数目与音调不同。即如果有效的频带高达3 4 0 0 千赫，该有效频带可以根据音调划分为8 到6 3 个频带。由逐个频带得到的幅值数据| A m | 的数目m M X + 1 在从8 到6 3 的范围内变化。因此，数据数目变化单元将可变数目m M X + 1 的幅值数据变化为预置数目M的数据，例如4 4 个数据。

按照预置的数目的数据例如 44 个数据作为单位，利用矢量量化单元116 通过进行加权的矢量量化，由设在频谱计算单元148 的输出单元或在矢量量化单元116 的输入单元的数据数目变化单元获得预定数目例如44 个幅值数据或包络线数据。利用按听觉加权的滤波计算电路139 的输出端提供这种加权权重。经过开关117 在输出端103 取出来自矢量量化器116 的包络线索引。在加权的矢量量化之前，利用对于由预置数目的数据构成的矢量的适当的泄漏系数适当取得帧内的差值。

下面解释第二编码单元120 。第二编码单元120 具有所谓的CELP 编码结构并特定用在对输入的语音信号的不发声部分进行编码。在用于输入的语音信号的不发声部分的CELP 编码结构中，与发清辅音的声音部分的LPC 余部相对应的作为噪声代码本或所谓的随机代码本121 的代表性数值输出的噪声输出经过增益控制电路126 送到按听觉加权的合成滤波器122 。加权的合成滤波器122 对输入的噪声进行LPC 合成并将所产生的加权的不发声部分的信号送到减法器123 。由输入端101 经过高通滤波器 (HPF) 109 向减法器123 提供一个信号，该信号经过按听觉加权的滤波器125 已进行按听觉加权，取出该信号和来自合成滤波器122 的信号之间的差或误差。同时由按听觉加权的滤波器125 的输出端预先减去按听觉加权的合成滤波器的零输入响应。这一误差被送到用于计算距离的距离计算电路124 。在噪声代码本121 查找使误差最小的一个代表性的矢量数值。上述就是采用闭环查找接着采用合成法分析对时域波形进行矢量量化的概要。

像对于来自采用CELP 编码结构的第二编码器120 的不发声 (UV) 部分的数据一样，取出来自噪声代码本121 的代码本的波形索引和来自增益电路126 的代码本的增益索引。将作为来自噪声代码本的UV 数据的波形索引和作为增益电路126 的UV 数据的增益索引经过开关127g 送到输出端107g 。

这些开关127s 、127g 和开关117 、118 根据来自V/UV 鉴别单元115 的V/UV 鉴别结果接通或者断开。具体地说，现时传输的帧的语音信号的V/UV 鉴别结果表明为发声 (V) 部分，则开关117 和118 接通，而当现时传输的帧的语音信号如为不发声 (UV) 部分时，开关127s 、127g 接通。

图4 表示图2 中所示语音信号解码器的更详细的结构。在图4 中，使用的数码标注与图2 中相对应的元件。

在图4 中，与图1 和图3 的输出端102 相对应的LSP 的矢量量化输出即代码本索引提供到输入端202 。

LSP 索引送到属于LPC 参数重现单元213 的反矢量量化器231 ，以便对线性频谱对（LPS） 数据进行反矢量量化，再提供到用于内插的LSP 内插电路232 、233 。利用由LSP 向α变换电路将所形成的内插的数据变换成α参数，再送到LPC 合成滤波器214 。LSP 内插电路232 和由LSP 向α变换电路234 设计用于发声部分（V） 的声音，而LSP 内插电路233 和由LPS 向α变换电路235 设计用于不发声部分（UV） 的声音。LPC 合成滤波器214 分为发浊音的语音部分的LPC 合成滤波器236 和发清辅音的语音部分的LPC 合成滤波器237 。即，对于发浊音的语音部分和发清辅音的语音部分独立地进行LPC 系数内插，用于抑制在从发浊音的语音部分到发清辅音的语音部分或者相反的过渡部分由于内插具有完全不特征的LSP 可能产生的不利影响。

向图4 中的输入端203 提供与图1 和图3 中的解码器的连接端103 的输出相应的加权的矢量量化的频谱包络线相对应的代码索引数据。向输入端204 提供来自图1 和图3 中的连接端104 的音调数据，向输入端205 提供来自图1 和3 中的连接端105 的V /UV 鉴别数据。

来自输入端203 的频谱包络线Am 的矢量量化的索引数据送到用于反矢量量化的反矢量量化器212 ，在其中对于数据数目变换进行反变换。形成的频谱包络线数据送到正弦合成电路215 。

如果在对编码过程中的频谱进行矢量量化之前求出帧内的差，在用于产生频谱包络线数据的反矢量量化之后对该帧内差进行解码。

向正弦合成电路215 提供来自输入端204 的音调和来自输入端205 的V /UV 鉴别数据。由正弦合成电路215 取出与在图1 和图3 中所示的LPC 反变换滤波器111 的输出相对应的LPC 余部数据并送到加法器218 。

反矢量量化器212 的包络线数据和来自输出端204 、205 的V /UV 鉴别数据送到用于对发声部分（V） 添加噪声的噪声合成电路216 。噪声合成电路216 的输出经过加权叠加一加法电路217 送到加法器218 。

具体地说，考虑到这样一个事实，如果利用正弦波合成产生作为发声声音部分的LPC合成滤波器的输入的激励信号，则在低音调声音例如男人的声音中产生堵塞的感觉，并且在发声部分的声音和不发声部分的声音之间声音质量急剧变化，因此产生不自然的听觉，故要将噪声添加到LPC余部信号的发声部分。这种噪声考虑到与语声编码数据，例如音调、频谱包络线的幅值、帧内的最大幅值或余部信号电平相关的参数，并与发浊音的语声部分的LPC合成滤波输入即激励信号相结合。

加法器218的加法输出送到用于LPC合成滤波器214的发声声音部分的合成滤波器236，在其中进行LPC合成，以便形成时间或波形数据，再由用于发声语声部分的后置滤波器238V进行滤波并送到加法器239。

来自图3中的输出端107s和107g的作为UV数据的波形索引和增益索引，提供到图4中的输入端207sa和207g，并由其再提供到不发声部分的语音合成单元220。来自连接端207s的波形索引送到不发声部分的语音合成单元220的噪声代码本221，而来自连接端207g的增益索引送到增益电路222。由噪声代码本221读出的代表性的数值输出是与不发声部分的语音的LPC余部相对应的噪声信号分量。这一数值变为在增益电路222中的预置的增益幅值并送到窗口形成电路223，以便形成用于平滑到发声语音部分的结合部的窗口。

窗口形成电路223的输出送到属于不发声（UV）的语音部分的LPC合成滤波器214的合成滤波器237。送到合成滤波器237的数据利用LPC合成进行处理，以便变成不发声部分的时间波形数据。在送到加法器239前，利用一个用于不发声部分的后置滤波器对不发声部分的时间波形数据进行滤波。

在加法器239，来自用于发声部分的后置滤波器238v的时间波形信号和来自用于发清辅音的语音部分的后置滤波器238u的发清辅音的语音部分的时间波形数据彼此相加，在输出端201取出形成的求和数据。

上述语音编码器根据所要求的声音质量可以输出不同比特速率的数据。即输出的数据可以按照可变的比特速率输出。例如，如低的比特速率为2kbps（千比特/秒），以及高的比特速率为6kbps，则输出的数据如图15所示为具有如下比特速率的数据。

表 1

	2kbps	6kbps
UV鉴别输出	1比特/20毫秒	1比特/20毫秒
LSP量化索引	32比特/40毫秒	48比特/40毫秒
对于发声的语音部分(V)	音调数据8比特/20毫秒	音调数据8比特/20毫秒
	索引15比特/20毫秒	音调数据8比特/20毫秒
	波形(对于第一级)5+5比特/20毫秒 增益5比特/20毫秒	波形(对于第一级) 5+5比特/20毫秒 增益 5比特/20毫秒 增益 (对于第二级) 72比特/20毫秒
对于不发声的语声部分(UV)	索引 11比特/10毫秒	索引23比特/5毫秒
	波形 (对于第一级) 7比特/10毫秒 增益,4比特/10毫秒	波形 (对第一级) 9比特/5毫秒 增益 6比特/5毫秒 波形(对于第二级) 5比特/5毫秒 增益3比特/5毫秒
对于发浊音的语声部分	40比特/20毫秒	120比特/20毫秒
对于不发声的语声部分	39比特/20毫秒	117比特/20毫秒

来自输出端104的音调数据对于发声部分的语音按照8比特/20毫秒的比特速率在a1各时间点输出，来自输出端105的V/UV鉴别输出按照1比特在所有的时间点上输出。用于LSP量化、由输出端102输出的索引在32比特/40毫秒和48比特/40毫秒之间转换。另一方面，在由输出端103输出发声语声部分过程中的索引在15比特/20毫秒和87比特/20毫秒之间转换。由输出端107s和107g输出的发清辅音的语声部分的索引在11比特/10毫秒和23比特/5毫秒之间转换。对于发浊音的声音部分（UV）的输出数据对于2kbps和40比特/20毫秒，对于6kbps为120kpbs/20毫秒。另一方面，对于发浊音的声音部分（UV）的输出数据对于2kbps为39比特/20毫秒和对于6kbps为117kbps/20毫秒。

下面结合相关部分的配置解释用于LSP量化的索引、用于发声语声部分（V）的索引和用于不发声语声部分（UV）的索引。

参阅图5和6，详细解释在LSP量化器134中的矩阵量化和矢量量化。

来自LPC分析电路132的α参数送到用于进行由α向LSP参数变换的α-LSP电路133。如在LPC分析电路132中进行P阶LPC分析，则计算P阶的α参数，这P阶的α参数变换为LSP参数，并存储在缓冲器610。

缓冲器610输出2帧的LSP参数，该两帧LSP参数利用由第一矩阵量化器620_1和第二矩阵量化器620_2构成的矩阵量化器620进行矩阵量化。在第一矩阵量化器620_1中对两帧的LSP参数进化矩阵量化，所形成的量化误差在第二矩阵量化器620_2中进一步进行矩阵量化。矩阵量化利用了沿时间轴和在频率轴的变换。

来自矩阵量化器620_2的对于2个帧的量化误差输入到由第一矢量量化器640_1和第二矢量量化器640_2构成的矢量量化单元640。第一矢量量化器640_1由两个矢量量化部分650、660构成，而第二矢量量化器640_2由两个矢量量化部分670、680构成。来自矩阵量化单元620的量化误差利用第一矢量量化器640_1的矢量化器部分650、660按帧进行量化。所形成的量化误差矢量利用第二矢量量化器640_2的矢量量化部分670、680进一步进行矢量量化。上述的矢量量化利用了沿频率轴的变

换。

执行如上所述矩阵量化的矩阵量化单元620至少包括用于实施第一矩阵量化步骤的第一矩阵量化器620$_1$和用于实施对于由第一矩阵量化产生的量化误差进行矩阵量化的第二矩阵量化步骤的第二矩阵量化器620$_2$。执行上述矢量量化的矢量量化单元640至少包括用于实施第一矢量量化步骤的第一矢量量化器640$_1$和用于实施对于由第一矢量量化产生的量化误差进行矩阵量化的第二矩阵量化步骤的第二矢量量化器640$_2$。

下面详细地解释矩阵量化和矢量量化。

存储在缓冲器600中的用于两个帧的LSP参数，即一个10×2的矩阵送到第一矩阵量化器620$_1$。第一矩阵量化器620$_1$将用于2个帧的LSP参数经过LSP参数加法器621送到加权的距离计算单元623，用以求出最小值的加权距离。

利用方程（1）给出利用第一矩阵量化器620$_1$进行代码本查找的过程中的失真的量值dMQ1：

$$d_{MQ1}(X_1,X_1') = \sum_{t=0}^{1}\sum_{i=1}^{P} w(t,i)(x_1(t,i) - x_1'(t,i))^2 \qquad \ldots(1)$$

其中X_1是LSP参数，X_1'是量化值，t和i是P维的数目。

利用方程（2）给出设有考虑沿频率轴和沿时间轴的加权权重限制的加权函数W（i，t）：

$$W(t,i)=\frac{1}{X(t,i+1)-X(t,i)}+\frac{1}{X(t,i)-X(t,i-1)}$$

其中x(t,0)=0,x(t,p+1)=π 与-无关。

方程（2）中的加权函数用于下游侧的矩阵量化和失量量化。

计算的加权距离送到用于矩阵量化的矩阵量化器MQ$_1$ 622。由这一矩阵量化输出的8 比特索引送到信号转换器690。在加法器620 中由用于二个帧的LSP 参数减去通过矩阵量化得到的量化值。加权的距离计算单元623 顺序地计算每2 个帧的加权的距离，以便在矩阵量化单元622 中进行矩阵量化。此外，选择使加权的距离为最小的量化值。加法器621 的输出送到第二矩阵量化器620$_2$ 的加法器631。

与第一矩阵量化器620$_1$ 相似，第二矩阵量化器620$_2$ 进行矩阵量化。加法器621 的输出经过加法器631 送到加权的距离计算单元633，在其中计算最小的加权距离。

利用方程（3）给出由第二矩阵量化器620$_2$ 进行代码本查找的过程中的失真的量值:

$$d_{MQ2}(X_2,X_2') = \sum_{t=0}^{1}\sum_{i=1}^{P} w(t,i)(x_2(t,i) - x_2'(t,i))^2 \qquad \ldots(3)$$

其中X$_2$ 和X$_2$′分别是来自第一矩阵量化器620$_1$ 的量化误差和量化值。

加权的距离送到用于进行矩阵量化的矩阵量化单元（MQ2）632。在加法器631 由两帧的量化误差减去通过矩阵量化输出的8 比特索引。加权的距离计算单元633 顺序地利用加法器631 的输出计算加权的距离。选择使加权的距离为最小的量化值。加法器631 的输出逐帧地送到第一失量量化器640$_1$ 中的加法器651、661。

第一失量量化器640$_1$ 逐帧地进行失量量化。加法器631 的输出逐帧地经过加法器651、661 送到每个加权的距离计算单元653、663，用于计算最小的加权距离。

量化误差X$_2$ 和量化误差X$_2$′之间的差为一个（10×2）的矩阵。如果该距离按照X$_2$ -X$_2$′=[X$_{3-1}$，X$_{3-2}$]来表示，利用方程（4）

和（5）给出利用第一矢量量化器640$_1$的矢量量化单元652、662进行代码本查找的过程中的失真的量值d_{VQ1}，d_{VQ2}：

$$d_{VQ1}(\underline{x}_{3-1},\underline{x}_{3-1}') = \sum_{i=1}^{P} w(0,i)(x_{3-1}(0,i) - x_{3-1}'(0,i))^2$$

$$d_{VQ2}(\underline{x}_{3-2},\underline{x}_{3-2}') = \sum_{i=1}^{P} w(1,i)(x_{3-2}(1,i) - x_{3-2}'(1,i))^2$$

加权的距离送到用于进行矢量量化的矢量量化单元XQ$_1$ 652和矢量量化单元VQ$_2$ 662，由这种矢量量化输出的8比特的索引送到信号转换器690。利用加法器651、661由输入的两帧的量化误差矢量减去量化值。加权的距离计算单元653、663顺序地利用加法器651、661的输出计算加权的距离，用以选择使加权的距离为最小的量化值。加法器651、661的输出送到第二矢量量化器640$_2$的加法器671、681。

利用方程（6）和（7）：

$$d_{VQ3}(\underline{x}_{4-1}, \underline{x}_{4-1}') = \sum_{i=1}^{P} w(0,i)(x_{4-1}(0,i) - x_{4-1}'(0,i))^2 \quad \ldots(6)$$

$$d_{VQ4}(\underline{x}_{4-2}, \underline{x}'_{4-2}) = \sum_{i=1}^{P} w(1,i)(x_{4-2}(1,i) - x'_{4-2}(1,i))^2 \quad \ldots(7)$$

给出对于

$$\underline{X}_{4-1} = X_{3-1} - \underline{X}_{3-1}'$$
$$\underline{X}_{4-2} = X_{3-2} - \underline{X}_{3-2}'$$

利用第二矢量量化器640_2的矢量量化器672、682进行代码本查找的过程中的失真的量值。

这些加权的距离送到用于矢量量化的矢量量化器（VQ_3）672和矢量量化器（VQ_4）682。利用加法器671、681由对于两帧的输入的量化误差矢量减去来自矢量量化器的8比特输出的索引数据。加权的距离计算单元673、683利用加法器671、682的输出顺序地计算加权的距离，用以选择使加权的距离为最小的量化值。

在学习代码本的过程中，根据各个失真的量值利用通用的劳埃德算法进行“学习”。

在代码本查找过程中和学习过程中的失真的量值可能具有不同的数值。

来自矩量化单元622、632和矢量量化单元652、662、672和682的8位的索引数据利用信号转换器690进行转换并在输出端691输出。

具体地说，对于低的比特速率，取出执行第一矩阵量化步骤的第一矩阵量化器620_1的输出、执行第二矩阵量化步骤的第二矩阵量化器620_2的输出以及执行第一矢量量化步骤的第一矢量量化器640_1的输出，而对于高的比特速率，则将按低比特速率的输出加到执行第二矢量量化步骤的第二矢量量化器640_2的输出上，并取出所形成的和。

这样就分别对于2kbps和6kbps输出32比特/40毫秒的索引和48比特/40毫秒的索引。

矩阵量化单元620和矢量量化单元640依照代表LPC系数的参数的特征进行限于沿频率轴和/或时间轴的加权。

首先解释按照LSP参数的特征限于频率轴的加权。如阶数P=10，LSP参数X（i）按低、中和高范围的3个范围分组为:

$$L_1=\{X(i)\mid 1\leq i\leq 2\}$$
$$L_2=\{X(i)\mid 3\leq i\leq 6\}$$
$$L_3=\{X(i)\mid 7\leq i\leq 10\}$$

如果各组L_1、L_2、L_3的权重为1/4、1/2、1/4，利用方程（8）、（9）、（10）输出仅限于沿频率轴的加权函数：

$$w'(i)=\frac{w(i)}{\sum_{j=1}^{2}w(j)}\times\frac{1}{4} \qquad \ldots(8)$$

$$w'(i)=\frac{w(i)}{\sum_{j=3}^{6}w(j)}\times\frac{1}{2} \qquad \ldots(9)$$

$$w'(i)=\frac{w(i)}{\sum_{j=7}^{10}w(j)}\times\frac{1}{4} \qquad \ldots(10)$$

仅按每组进行各个LSP参数的加权，并对每组通过加权限定权重。

分析沿时间轴方向的情况，设各个帧的总和为1，这样沿时间轴方向的限定是以帧为基准的。利用方程（11）给出仅沿时间轴方向限定的权重：

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{10}\sum_{s=0}^{1} w(j,s)} \quad \ldots(11)$$

其中1 ≤i ≤1 0 和0 ≤t ≤1 。

利用这一方程 (1 1)，在帧号为t =0 和t =1 的两帧之间进行不限定沿频率轴方向的加权。在利用矩阵量化处理的两帧之间进行这种仅沿时间轴方向限定的加权。

在学习过程中，用作学习数据的总数为T 的各帧的总体根据方程 (1 2) 进行加权。

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{10}\sum_{s=0}^{T} w(j,s)} \quad \ldots(12)$$

其中1 ≤i ≤1 0 和0 ≤t ≤T 。

下面解释限定在沿频率轴方向和沿时间轴方向的加权。如果阶数P =1 0 、L S P 参数按照低、中和高范围的三个范围分组为:

$$L_1=\{X(i,t)\mid 1\leq i\leq 2,\ 0\leq t\leq 1\}$$
$$L_2=\{X(i,t)\mid 3\leq i\leq 6,\ 0\leq t\leq 1\}$$
$$L_3=\{X(i,t)\mid 7\leq i\leq 10,\ 0\leq t\leq 1\}$$

如果各组的权重分别为1 /4 、1 /2 、1 /4 ，则利用方程 (1 3) 、 (1 4) 和 (1 5) 给出仅沿频率轴方向限定的加权函数:

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{2}\sum_{s=0}^{1} w(j,s)} \times \frac{1}{4} \quad \ldots(13)$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=3}^{6}\sum_{s=0}^{1} w(j,s)} \times \frac{1}{2} \quad \ldots(14)$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=7}^{10}\sum_{s=0}^{1} w(j,s)} \times \frac{1}{4} \quad \ldots(15)$$

利用这些方程（13）到（15），进行每3个帧沿频率轴方向限定的加权和遍布利用矩阵量化处理的两个帧的加权。在代码本查找的过程中和学习的过程中都是有效的。

在学习的过程中，对全体数据的各帧的总体进行加权。对于低、中和高范围，将LSP参数X（i，t）分组为：

$$L_1=\{X(i,t)\mid 1\leq i\leq 2,\ 0\leq t\leq T\}$$
$$L_2=\{X(i,t)\mid 3\leq i\leq 6,\ 0\leq t\leq T\}$$
$$L_3=\{X(i,t)\mid 7\leq i\leq 10,\ 0\leq t\leq T\}$$

如果各组L_1、L_2和L_3的权重为1/4、1/2和1/4，则利用方程（16）、（17）和（18）给出仅沿频率轴限定的对于各组L_1、L_2和L_3的加权函数：

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{2}\sum_{s=0}^{T} w(j,s)} \times \frac{1}{4} \quad \ldots(16)$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=3}^{6}\sum_{s=0}^{T} w(j,s)} \times \frac{1}{2} \quad \ldots(17)$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=7}^{10}\sum_{s=0}^{T} w(j,s)} \times \frac{1}{4} \quad \ldots(18)$$

利用这些方程（16）到（18），可以对沿频率轴的三个范围和遍布沿时间轴的各帧的总体进行加权。

此外，根据LSP参数的变化幅度，矩阵量化单元620和矢量量化单元640进行加权。在从V到UV或从UV到V的过渡部分，在全部语音各帧中它们占少部分的帧，由于在浊音和清辅音之间的频率响应的差别，LSP参数明显变化。因此，由方程（19）表示的加权函数可以乘以加权函数W′（i，t），用以在该过渡部分中重点加权。

$$wd(t) = \sum_{i=1}^{10} |x_1(i,t) - x_1(i,t-1)|^2$$

...(19)

如下的方程（20）：

$$wd(t) = \sum_{i=1}^{10} \sqrt{|x_1(i,t) - x_1(i,t-1)|}$$

...(20)

可以替代方程（19）。

因此，LSP量化单元134实行两级矩阵量化和两级矢量量化，以便使输出索引的比特数是可变的。

在图7中表示矢量量化单元116的基本结构，而在图8中表示图7中所示的矢量量化单元116的更详细的结构。下面解释在矢量量化单元116中对于频谱包络线Am进行加权的矢量量化的说明性的结构。

首先，在图3所示的语音信号编码装置中，解释用于在频谱计算单元的输出侧或在矢量量化单元116的输入侧提供恒定数量的频谱包络线幅值数据的进行数据数目变换的说明性的配置。

对于这种数据数目变换可以考虑采用各种不同的方法。在本实施例中，将从一个数据块中的最后的数据到该数据块中的第一个数据内插一些值的伪数据，或者其它预置数据例如重复在一数据块中的最后的数据或第一个数据附加到沿频率轴的一个有效频带的一个数据块的幅值数据，用以提高到N_F的数据数目，通过例如利用FIR滤波器进行限定带宽方式的O_S-Fold（O_S倍）例如8倍普通采样（oversampling）求出数目为O_S倍例如为8倍的幅值数据。为了扩展到较大的N_M数目例如2048，线性内插（mMx+1）×O_S个幅值数据。对这一N_M数据再采样，用以变换为上述预置数目M个数据，例如44个数据。

实际上，通过普遍采样和线性内插计算为了表示最需要的M数据所需的数据，不必求出上述N_M数据。

用于实现图7中的加权的矢量量化的矢量量化单元116至少包括用于进行第一矢量量化步骤的第一矢量量化单元500和用于进行对由第一矢量量化

单元500进行第一矢量量化的过程中产生的量化误差矢量进行量化的第二矢量量化步骤的第二矢量量化单位510。这一第一矢量量化单元500是所谓的第一级矢量量化单元，而第二矢量量化单元510是所谓的第二级矢量量化单元。

频谱计算单元148的输出矢量X，即预置数目M的包络线数据输入到第一矢量量化单元500的输入端501。这一输出矢量X 由矢量量化单元502通过加权的矢量量化进行量化。因此，在输出端503输出由矢量量化单元502输出的波形索引，而在输出端504输出量化的数值X_0'，并送到加法器505、513。加法器505由源矢量X减去量化值X_0'，提供一多阶的量化误差矢量Y。

量化的矢量误差Y送到在第二矢量量化单元510中的矢量量化单元511。这一第二矢量量化单元511由多个矢量量化单元构成，或如在图7中由两个矢量量化器511_1、511_2构成。量化误差矢量Y按维划分，以便在两个矢量量化器511_1、512_2中利用加权的矢量量化进行量化。在输出端511、511_2输出由这些矢量量化器511_1、511_2输出的波形索引，同时量化的数值Y'、Y'2沿该维的方向接续并送到加法器513。加法器513将量化值Y'1、Y'2加到量化值X'0上，以产生一个在输出端514输出的量化值X'1。

因此，对于低的比特速率，取出利用第一矢量量化单元500执行第一矢量量化步骤产生的输出，而对于高的比特速率，输出利用第二量化单元510执行第二量化步骤产生的输出。

具体地说，矢量量化部分116中的第一矢量量化单元500中的矢量量化器502具有如图8中所示的L阶，例如44阶的两级结构。

即，将代码本容量为32乘以增益g1的44阶矢量量化代码本的输出矢量之和用作44阶频谱包络线矢量X的量化值X'0。因此，如图8所示，两个代码本为CB_0和CB_1、而输出的矢量为S_1i、S_1j，其中$0 \leq i$和$j < 31$。另一方面，增益代码本CB_g的输出为gl，其中$0 \leq \leq 31$，gl是一个标量。极值输出X'0为$ge\ (S_1i + S_1j)$。

通过LPC的余部的上述的MBE分析得到的并变换为预置阶的频谱包络Am为X。关键的是怎样有效地对X进行量化。

量化误差能量（e n e r g y）E 按下式确定:

$$E = \| W\{Hx - Hgl((S_{oi}+S_{ij})\} \|^2$$
$$= \| WH\{X - g_i(S_{oi}+S_{ij})\} \|^2 \qquad (21)$$

其中H 标注沿L P C 合成滤波器的频率轴的特征，和W为对所代表的特征进行加权的矩阵，用以沿频率轴进行按听觉加权。

假如由现时帧的L P C 分析的结果得到的α参数表示为α_i （$1 \leq i \leq p$），根据方程（2 2）

$$H(z) = \frac{1}{1 + \sum_{i=1}^{P} \alpha_i z^{-i}} \qquad ...(22)$$

的频率响应对L 阶例如4 4 阶对应点的数值进行采样。

为了计算，将各个O 接着序列1 、α_1 、α_2 、$\cdots\alpha_p$ 填加，以便提供序列1 、α、α_2 $\cdots\alpha_p$ ，0 ，0 …，0 ，提供2 5 6 点的数据。然后，利用2 5 6 点的F F T，对于与从0 到π的范围相关的点计算（r e ^{2} +i m ^{2} ）$^{1/2}$，并求出各结果值的倒数。这些导数按L 点再取样（例如4 4 点），并形成一个将这些L 点作为对角线基元的矩阵:

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

利用方程（2 3）给出按听觉加权的矩阵W:

$$W(z) = \frac{1 + \sum_{i=1}^{P} \alpha_i \lambda_b^i z^{-i}}{1 + \sum_{i=1}^{P} \alpha_i \lambda_a^i z^{-i}} \qquad \ldots(23)$$

其中α_1是LPC分析的结果，λ_a、λ_b是常数、例λ_a =0．4和λ_b = 0．9。

矩阵W可以由上述方程（2 3）的频率响应计算。例如，按照256点的数据1，α1 λb，α2 λb²，…αp λbp，0，0…0完成FFT，以便对由0到π的域求（re²[i]+Im²[i]）^1/2，其中0≤i≤128。对于从0到π的域，对1，α1 λa，α2 λa²，…αp λa^p，0，0…0，在128点上，利用256点FFT求出分析的频率响应。以便求出（re'²[i]+im'²[i]）^1/2，其中0≤i≤128。

利用 $w_0[i] = \frac{\sqrt{re^2[i] + im^2[i]}}{\sqrt{re'^2[i] + im'^2[i]}}$ 可以求出方程（2 3）的频率响应，其中0≤i≤128。对于例如44阶矢量的每一个相关的点利用如下的方法求出该频率响应。更确切地说，应当采用线性内插。然而，在如下的实例中，代之以使用最接近的点。

即 ω[i]=ω0[nint{128i/L)]，其中1≤i≤L。

在该方程中，nint（X）是一个函数，重现一个最接近X的数值。

正如对H一样，利用相似的方法求出h（1），h（2），…h（L），

即

$$H = \begin{bmatrix} h(1) & & 0 \\ & h(2) & \\ & & \ddots \\ 0 & & h(L) \end{bmatrix} \quad W = \begin{bmatrix} w(1) & & 0 \\ & w(2) & \\ & & \ddots \\ 0 & & w(L) \end{bmatrix}$$

$$WH = \begin{bmatrix} h(1)w(1) & & 0 \\ & h(2)w(2) & \\ & & \ddots \\ 0 & & h(L)w(L) \end{bmatrix}$$

...(24)

作为另一个实例，首先求出H (Z) W (Z)，然后求出频率响应，用以降低F F T 的次数。即方程 (2 5)：

$$H(z)W(z) = \frac{1}{1 + \sum_{i=1}^{P} \alpha_i z^{-i}} \cdot \frac{1 + \sum_{i=1}^{P} \alpha_i \lambda_b^i z^{-i}}{1 + \sum_{i=1}^{P} \alpha_i \lambda_a^i z^{-i}}$$

...(25)

的分母展开为

$$\left(1 + \sum_{i=1}^{P} \alpha_i z^{-i}\right)\left(1 + \sum_{i=1}^{P} \alpha_i \lambda_a^i z^{-i}\right) = 1 + \sum_{i=1}^{2P} \beta_i z^{-i}$$

利用序列1，$\beta 1$，$\beta 2 \cdots \beta_{255}$，0，0，…0，例如产生256点的数据。然后利用幅值为

$$rms[i]=\sqrt{re''^2[i]+im''^2[i]}$$

的频率响应，完成256点FFT，

其中$0 \leq i \leq 128$。由此

$$wh_o[i]=\frac{\sqrt{re^2[i]+im^2[i]}}{\sqrt{re''^2[i]+im''^2[i]}}$$

其中$0 \leq i \leq 128$。对于L维矢量的每一个各对应点求出这一数值。如果FFT的点数少，应当采用线性内插。然而，在这里利用:

$$wh[i]=wh_o[n\ int(\frac{128}{L} \cdot i)]$$

求出最接近的数值，其中$1 \leq i \leq L$。如果以这些数值作为对角线基元的矩阵为W'，则

(26)

$$W' = \begin{bmatrix} wh(1) & & 0 \\ & wh(2) & \\ & & \ddots \\ 0 & & wh(L) \end{bmatrix}$$

方程（26）表示与方程（24）相同的矩阵。

另外，|H（exp（jω）W（exp（jω））|可以按ω=i/L λ直接由方程（25）直接求出，以便用作于wh[i]。再者，按另外一种方式，求出适当长度例如64个点的方程（25）的脉冲响应并进行FFT，以便求出然后可用于wh[i]的幅值频率特性。

利用这一矩阵重新表示方程（21），即加权的合成滤波器的频率响应，我们得到方程（27）：

$$E=\| W'(\underline{X}-g_1((\underline{S}_{0i}+\underline{S}_{1j}))\|^2 \quad \cdots(27)$$

下面解释用于学习波形代码本和增益代码本的方法。

对于所有的帧K使所希望失真值最小，对于这些帧选择用于CB$_0$'的代码矢量$\underline{S}_{0c}$。如有M个这样的帧，若使

$$J=\frac{1}{M}\sum_{k=1}^{M}\|W_k'(\underline{x}-g_k(\underline{s}_{0c}+\underline{s}_{ik}))\|^2 \quad \ldots(28)$$

数值最小，帧数M就足够了。在方程（28）中，W_k'、$\underline{X}_k$，g_k和$\underline{S}_{ik}$分别代表用于第K帧的权重，第k'帧的一个输入，第K'帧的增益和用于第K'帧的代码本CB1的输出。

为使方程（28）的J最小，

$$J=\frac{1}{M}\sum_{k=1}^{M}\{(\underline{x}_k^T-g_k(\underline{s}_{0c}^T+\underline{s}_{1k}^T))W_k'^TW_k'(\underline{x}_k-g_k(\underline{s}_{0c}+\underline{s}_{1k}))\}$$
$$=\frac{1}{M}\sum_{k=1}^{M}\{\underline{x}_k^TW_k'^TW_k'\underline{x}_k-2g_k(\underline{s}_{0c}^T+\underline{s}_{1k}^T)W_k'^TW_k'\underline{x}_k$$
$$+g_k^2(\underline{s}_{0c}^T+\underline{s}_{1k}^T)W_k'^TW_k'(\underline{s}_{0c}+\underline{s}_{1k})\}$$
$$=\frac{1}{M}\sum_{k=1}^{M}\{\underline{x}_k^TW_k'^TW_k'\underline{x}_k-2g_k(\underline{s}_{oc}^T+\underline{s}_{1k}^T)W_k'^TW_k'\underline{x}_k$$
$$+g_k^2\underline{s}_{oc}^TW_k'^TW_k'\underline{s}_{oc}+2g_k^2\underline{s}_{oc}^TW_k'^TW_k'\underline{s}_{1k}+g_k^2\underline{s}_{1k}^TW_k'^TW_k'\underline{s}_{1k}\}$$

...(29)

$$\frac{\partial J}{\partial \underline{s}_{0c}}=\frac{1}{M}\sum_{k=1}^{M}\{-2g_kW_k'^TW_k'\underline{x}_k+2g_k^2W_k'^TW_k'\underline{s}_{0c}+2g_k^2W_k'^TW_k'\underline{s}_{1k}\}=0$$

...(30)

因此,

$$\sum_{k=1}^{M}(g_kW_k'^TW_k'\underline{x}_k-g_k^2W_k'^TW_k'\underline{s}_{lk})=\sum_{k=1}^{M}g_k^2W_k'^TW_k'\underline{s}_{0c}$$

这样

$$s_{0c}=\left\{\sum_{k=1}^{M}g_k^2W_k'^TW_k'\right\}^{-1}\cdot\left\{\sum_{k=1}^{M}g_kW_k'^TW_k'(\underline{x}-g_k\underline{s}_{lk})\right\}$$

...(31)

其中（ ）表示一逆矩阵，$W_k{}'^T$ 表示$W_k{}'$的换位矩阵。

下面讨论增益最优化问题。

利用下式给出选择增益为代码字g c 的与第k ' 帧相关的失真的预期的数值，该式为:

$$J_g = \frac{1}{M}\sum_{k=1}^{N}\|W_k'(\underline{x}_k - g_c(\underline{s}_{0k} + \underline{s}_{1k}))\|^2$$
$$= \frac{1}{M}\sum_{k=1}^{M}\left\{\underline{x}_k^T W_k'^T W_k' \underline{x}_k - 2g_c \underline{x}_k^T W_k'^T W_k'(\underline{s}_{0k} + \underline{s}_{1k}) + g_c^2(\underline{s}_{0k}^T + \underline{s}_{1k}^T) W_k'^T W_k'(\underline{s}_{0k} + \underline{s}_{1k})\right\}$$

解:

$$\frac{\partial J_g}{\partial g_c} = \frac{1}{M}\sum_{k=1}^{M}\{-2\underline{x}_k^T W_k'^T W_k'(\underline{s}_{0k} + \underline{s}_{1k}) + 2g_c(\underline{s}_{0k}^T + \underline{s}_{1k}^T) W_k'^T W_k'(\underline{s}_{0k} + \underline{s}_{1k})\} = 0$$

得到:

$$\sum_{k=1}^{M}\underline{x}_k^T W_k'^T W_k'(\underline{s}_{0k} + \underline{s}_{1k}) = \sum_{k=1}^{M} g_c(\underline{s}_{0k}^T + \underline{s}_{1k}^T) W_k'^T W'(\underline{s}_{0k} + \underline{s}_{1k})$$

以及

$$g_c = \frac{\sum_{k=1}^{M}\underline{x}_k^T W_k'^T W_k'(\underline{s}_{0k} + \underline{s}_{1k})}{\sum_{k=1}^{M}(\underline{s}_{0k}^T + \underline{s}_{1k}^T) W_k'^T W'(\underline{s}_{0k} + \underline{s}_{1k})}$$

...(32)

上述方程（3 1）和（3 2）对于波形S_{0i}，S_{1j} 的最佳的矩心条件，以及对于0 ≤i ≤3 1 的增益g i ，即最佳解码输出。同时，可以按对S_{0i}相同的方式求出S_{1j} 。

下面讨论最佳编码条件，即最接近的邻近条件。

上述方程（2 7）用于求出失真量值，即每次求出使得方程$E = \| W'(X-g_i (S_{0i}+S_{1j})) \|^2$最小的$S_{0i}$和$S_{1j}$，逐帧给出输入$X$和加权矩阵$W'$。

实质上，对于g_i (0 ≤l ≤3 1)，S_{0i} (0 ≤i ≤3 1)，和S_{1j} (0 ≤i ≤3 1)的所有组合根据rownd robin fashion（图形罗宾方式）求出E，即3 2 ×3 2 ×3 2 =3 2 7 6 8，以便求出成组的S_{0i}，S_{1j}，它们会给出E 的最小值。然而，由于这要求进行大量的计算，在本实施例中要顺序地查找波形和增益。同时将r o u n d r o b i n 查找用于S_{0i}和S_{1j} 的组合。对于S_{0i}和S_{1j} 有3 2 ×3 2 =1 0 2 4 个组合。在如下的介绍中，为了简化，将$S_{0i}+S_{1j}$表示为S_m。

上述方程（2 7）变为$E=\| W'(X-g_1 S_m) \|^2$。如果为了进一步简化，$X_w = W'X$ 和$S_w = W' S_m$，得到：

$$E=\| X_w-g_1 S_w \|^2 \qquad \cdots(33)$$

$$E = \|x_w\|^2 + \|s_w\|^2\left(g_1 - \frac{x_w^T \cdot s_w}{\|s_w\|^2}\right)^2 - \frac{\left(x_w^T \cdot s_w\right)^2}{\|s_w\|^2}$$

...(34)

因此，g_1 可以足够精确地得到，可以按这样两步进行查找，这两步是：

（1） 对于S_w进行查找，该S_w将使下式最小：

$$\frac{(\underline{X}^{T}_{w} \cdot \underline{S}_{w})^{2}}{\| \underline{S}_{w} \|^{2}}$$

及（2） 对于g_i进行查找，该g_i最接近：

$$\frac{\underline{X}^{T}_{w} \cdot \underline{S}_{w}}{\| \underline{S}_{w} \|^{2}}$$

如果利用原有的符号重写上式，

（1′） 对于一组$\underline{S}_{oi}$和$\underline{S}_{oj}$进行查找，该组$\underline{S}_{oi}$，$\underline{S}_{1j}$将会使下式达到最大值：

$$\frac{\left(x^{T}W'^{T}W'(\underline{s}_{0i} + \underline{s}_{1j})\right)^{2}}{\|W'(\underline{s}_{0i} + \underline{s}_{1j})\|^{2}}$$

（2′） 对g_i进行查找，该g_i最接近下式：

$$\frac{x^{T}W'^{T}W'(\underline{s}_{0i} + \underline{s}_{1j})}{\|W'(\underline{s}_{0i} + \underline{s}_{1j})\|^{2}}$$

...(35)

上述（35）代表最佳编码条件（最接近的邻近条件）。

利用方程（31）和（32）的条件（矩心条件）（centfroid）和上式（35）的条件，通过使用所谓广义的劳埃德算法（GLA）可以同时将代码本（CB0、CB1和CBg）排成序列。

在本实施例中，将利用输入的X的范数（norm）相除的W′作为W′。即将W′/‖X‖替代在方程（31）、（32）和（35）中的W′。

另外，利用上述方程（26）确定该用于在利用矢量量化器116进行矢量量化时按听觉加权的加权W′。然而，还可以通过求出考虑了刚才的W′的现时的加权W′，来求出考虑了暂进掩蔽的加权W′。

将当在时间n，即在第n帧求出在上述方程（26）中的wh（1），wh（2），…wh（L）的数值分别表示为whn（1），whn（2），…whn（L），

如果将在时间t的考虑了刚才的数值的权重定义为An（i），其中1≤i≤L，则

$$An(i)=\lambda A_{n-1}(i)+(1-\lambda)whn(i),\ (whn(i)\leq A_{n-1}(i))$$
$$=whn(i),\ (whn(i)>A_{n-1}(i))$$

其中λ例如可设定等于0.2。在An（i）中，1≤i≤L，将这样求出的令这样的An（i）作为对角线上的基元的矩阵可以用作上述加权权重。

在输出端520、522分别输出按这样的方式通过加权的矢量量化得到的波形索引数值S_{0i}，S_{1j}，同时在输出端521输出增益索引g_1。此外，在输出端504输出的量化值X_0'送到加法器505。

加法器505由频谱包络线矢量X减去量化值，产生量化误差矢量Y。具体地说，这种量化误差矢量Y送到矢量量化单元511，以便按维划分并利用矢量量化器511_1到511_8通过加权的矢量量化方式进行量化。

第二矢量量化单元510使用的比特数大于第一矢量量化单元500。随之，代码本的存储容量和用于代码本查找的处理工作量（复杂程度）明显增加。因此，变得不可能按照44阶正如与第一矢量量化单元500一样来进行矢量量化。所以，第二矢量量化单元510中的矢量量化单元511由多个矢量量化器构成，并且输入的量化值按维分成为多个低维的矢量，用以进行加权的矢量量化。

在表2中表示了在量化值Y_0到Y_7（用在矢量量化器511_1到511_8中的）、维数和比特数之间的相互关系。

表 2

量化值	维数	比特数
$\underline{Y}_0$	4	10
$\underline{Y}_1$	4	10
$\underline{Y}_2$	4	10
$\underline{Y}_3$	4	10
Y_4	4	9
$\underline{Y}_5$	8	8
$\underline{Y}_6$	8	8
$\underline{Y}_7$	8	7

在输出端523$_1$到523$_8$输出由矢量量化器511$_1$到511$_8$输出的索引值Idvqo到Idvq7。这些索引数据的比特数之和为72。

如果由沿该维方向的矢量量化器511$_1$到511$_8$输出的量化值$\underline{Y}'_0$到$\underline{Y}'_7$得到的一个值为$\underline{Y}'$，则由加法器513对量化值$\underline{Y}'$和$\underline{X}'_0$求和，以提供量化值$\underline{X}'_1$。因此，量化值$\underline{X}'_1$由下式表示：

$$\underline{X}_1'=\underline{X}_0'+\underline{Y}'$$
$$=\underline{X}-\underline{Y}+\underline{Y}'$$

即极值的量化误差矢量为$\underline{Y}' - \underline{Y}$。

假如，来自第二矢量量化器5 1 0 的量化值$\underline{X}_1'$要进行解码，语音信号解码装置不需要来自第一量化单元5 0 0 的量化值$\underline{X}_1'$。然而，它需要来自第一量化单元5 0 0 和第二量化单元5 1 0 的索引数据。

下文解释在矢量量化部分中的学习方法和代码本查找。

像按学习方法一样，利用在表2 中表示的权重W′，量化误差矢量$\underline{Y}$被划分为8 个低阶的矢量$\underline{Y}_0$到$\underline{Y}_7$。如果权重W′是一个令4 4 点再取样的值作为沿对角线上的基元的矩阵：

$$W' = \begin{bmatrix} wh(1) & & & 0 \\ & wh(2) & & \\ & & \ddots & \\ 0 & & & wh(44) \end{bmatrix} \quad \ldots(36)$$

该权重W′被分成如下8 个矩阵：

$$W_1' = \begin{bmatrix} wh(1) & & 0 \\ & \ddots & \\ 0 & & wh(4) \end{bmatrix}$$

$$W_2' = \begin{bmatrix} wh(5) & & 0 \\ & \ddots & \\ 0 & & wh(8) \end{bmatrix}$$

$$W_3' = \begin{bmatrix} wh(9) & & 0 \\ & \ddots & \\ 0 & & wh(12) \end{bmatrix}$$

$$W_4' = \begin{bmatrix} wh(13) & & 0 \\ & \ddots & \\ 0 & & wh(16) \end{bmatrix}$$

$$W_5' = \begin{bmatrix} wh(17) & & 0 \\ & \ddots & \\ 0 & & wh(20) \end{bmatrix}$$

$$W_6' = \begin{bmatrix} wh(21) & 0 & \\ & \ddots & \\ 0 & & wh(28) \end{bmatrix}$$

$$W_7' = \begin{bmatrix} wh(29) & & 0 \\ & \ddots & \\ 0 & & wh(36) \end{bmatrix}$$

$$W_8' = \begin{bmatrix} wh(37) & & 0 \\ & \ddots & \\ 0 & & wh(44) \end{bmatrix}$$

因此分成低维的$\underline{Y}$和W′分别被称为Y i 和Wi ′，其中1 ≤i ≤8 。

失真量值E 定义如下:

$$E=\| W_i' (\underline{Y}_i-\underline{S}) \|^2 \qquad \cdots(37)$$

代码本矢量$\underline{S}$是$\underline{Y\,i}$量化的结果。查找使失真量值E 为最小的代码本的这样的代码矢量。

在代码本学习过程中，还利用通用的劳埃德算法（G L A）完成加权。首先解释用于学习的最佳矩心条件。如果有M个已选择代码矢量$\underline{S}$作为最佳量化结果的输入矢量$\underline{Y}$，并且形成序列的数据为$\underline{Y}_k$，利用方程（3 8）给出失真了的所希望的数值，使对于所有帧K 加权产生的失真的中心值最小:

$$\begin{aligned} J &= \frac{1}{M}\sum_{k-1}^{M} \|W_k'(\underline{y}_k - \underline{s})\|^2 \\ &= \frac{1}{M}\sum_{k-1}^{M} (\underline{y}_k - \underline{s})^T W_k'^T W_k' (\underline{y}_k - \underline{s}) \\ &= \frac{1}{M}\sum_{k-1}^{M} \underline{y}_k^T W_k'^T W_k' \underline{y}_k - 2\underline{y}_k^T W_k'^T W_k' \underline{s} \\ &\quad + \underline{s}^T W_k'^T W_k' \underline{s} \end{aligned} \qquad \ldots(38)$$

$$\frac{\partial J}{\partial \underline{s}} = \frac{1}{M}\sum_{k-1}^{M} (-2\underline{y}_k^T W_k'^T W_k' + 2\underline{s}^T W_k'^T W_k') = 0$$

求解，得到:

$$\sum_{k=1}^{M} \underline{y}_k^T W_k'^T W_k' = \sum_{k=1}^{M} \underline{s}^T W_k'^T W_k'$$

两侧数值换位，得到：

$$\sum_{k=1}^{M} W_k'^T W_k' \underline{y}_k = \sum_{k=1}^{M} W_k'^T W_k' \underline{s}$$

因此，

$$\underline{s} = \left(\sum_{k=1}^{M} W_k'^T W_k' \right)^{-1} \sum_{k=1}^{M} W_k'^T W_k' \underline{y}_k$$

...(39)

在上述方程（39）中，$\underline{S}$是一最佳代表性的矢量并代表最佳矩心条件。

正如对于最佳编码条件一样，它对于查找使$\| W_i' (\underline{Y}_i - \underline{S}) \|^2$的数值为最小的$\underline{S}$足够了。在查找过程中的$W_i'$（不少与在学习过程中的$W_i'$一样）可以是一个未加权的矩阵：

$$\begin{bmatrix} 1 & & 0 \\ & 1 & \\ & \ddots & \\ 0 & & 1 \end{bmatrix} .$$

由于利用两级矢量量化单元构成在语音信号编码器中的矢量量化单元

116，变得能够使输出的索引数目是可变的。

采用本发明的上述CELP编码器结构的第二编码单元120由多级矢量量化处理器构成，如图9所示。这些多级矢量量化处理器在图9所示实施例中是按照两级编码单元120_1、120_2构成的，其中表示了适合用6kbps的传输比特速率的配置，还可以使传输比特速率例如在2kbps和2kbps之间进行转换。此外，波形和增益索引输出可以在23比特/5毫秒和15比特/5毫秒之间转换，图10中表示了按图9中的配置进行的处理流程图。

参照图9，图9中的LPC分析电路302对应于图3中所示的LPC分析电路132，而LSP参数量化电路303对应于图3中的α向LSP变换电路133到从LSP向α变换电路137组成的结构，以及按听觉加权的滤波器304对应于图3中的按听觉加权的滤波器计算电路139和按听觉加权的滤波器125。因此，在图9中，与图3中的第一编码单元113中的由LSP向α变换电路137的输出相同的输出提供到连接端305，与图3中的按听觉加权的滤波器计算电路139的输出相同的输出提供到连接端307和与图3中的按听觉加权的滤波器125的输出相同的输出提供到连接端306。然而，与按听觉加权的滤波器125不同，图9中的按听觉加权的滤波器304利用输入的语声数据和预量化的α参数而不是利用LSP－α变换电路137的输出，产生按听觉加权的语声信号，即为与图3中的按听觉加权的滤波器125的输出相同的信号。

在图9中所示的两级第二编码单元120_1和120_2中，减法器313和323对应于图3中的减法器123，而距离计算电路314、324对应于图3中的距离计算电路124。此外，增益电路311、321对应于图3中的增益电路126，同时，随机的代码本310、320和增益代码本315、325对应于图3中的噪声代码本121。

在图9所示的结构中，LPC分析电路302在图10中的步骤S1将由连接端301提供的输入的语音数据分成如上所述的各帧，以便进行LPC分析，求出α参数。LSP参数量化电路303将来自LPC分析电路302的α参数变换为LSP参数，以便量化LSP参数。对量化的LSP参数内插并变换为α数。LSP参数量化电路303由该从量化的LSP参数变换的α参

数，即经变换的α参数产生LPC合成滤波器函数1/H(Z)，并将所产生的LPC合成滤波器函数1/H(Z)经过连接端305送到第一级第二编码单元120₁的按听觉加权的合成滤波器312。

按听觉加权的滤波器304由来自LPC分析电路302的α参数即预量化的α参数，求出用于按听觉加权的，与用图3中的按听觉加权的滤波器计算电路139产生的参数相同的参数。这些加权的数据经过连接端307提供到第一级第二编码单元120₁的按听觉加权的合成滤波器312。按照在图10中的步骤S₂，按加觉加权的滤波器304由输入的语音数据和预量化的α参数，产生按听觉加权的信号，该信号与图3中的按听觉加权的滤波器125输出的信号相同。即首先由预量化的α参数产生LPC合成滤波器函数W(Z)。经此产生的滤波器函数W(Z)应用于输入的语声数据X，产生XW，再经过连接端306作为按听觉加权的信号提供到第一级第二编码单元120₁中的减法器303。

在第一级第二编码单元120₁中，9比特波形索引输出的随机代码本310的代表性的数值输出送到增益电路311，在其中将来自随机的代码本310的代表性的输出与来自6比特增益索引输出的增益代码表315的增益(标量)相乘。利用增益电路311与增益相乘的该有代表性的数值输出送到按照1/A(Z)=(1/H(Z)*WZ滤波的按听觉加权的合成滤波器312。加权的合成滤波器312将1/A(Z)零输入响应输出送到减法器313，如在图10中的步骤S3所指示的。减法器313对按听觉加权的合成滤波器312的零输入响应输出和来自按听觉加权的滤波器304的按听觉加权的信号XW进行减法运算，并取出所形成的差或误差作为一个参考矢量上。在第一级第二编码单元120₁中进行查找的过程中，这一参考矢量上送到距离计算电路314，在其中计算距离和查找使量化误差能量E为最小值的波形矢量和增益g，如按照图10中的步骤S4所示。这里，1/A(Z)处于零状态。即如果在利用处于零状态的1/A(Z)合成的代码本中的波形矢量为S_{syn}，查找使方程(40)的E为最小的波形矢量S和增益g，方程(40)如下：

$$E=\sum_{n=0}^{N-1}\left(r(n)-gs_{syn}(n)\right)^2 \qquad \ldots(40)$$

虽然，可以全面查找使量化矢量能量E 为最小值的S 和g ，但可采用如下的方法，以便降低计算量。

第一种方法用于查找使E s 最小的波形矢量S ，E s 由如下方程（4 1 ）确定：

$$E_s = \frac{\sum_{n=0}^{N-1} r(n)s_{syn}(n)}{\sqrt{\sum_{n=0}^{N-1} s_{syn}(n)^2}}$$

...(41)

根据利用第一种方法得到的S ，利用方程（4 2 ）表示理想的增益：

$$g_{ref} = \frac{\sum_{n=0}^{N-1} r(n)s_{syn}(n)}{\sum_{n=0}^{N-1} s_{syn}(n)^2}$$

...(42)

因此，按照第二种方法，查找使方程（4 3 ）的E g 为最小值的g ，方程（4 3 ）为：

$$Eg=(g_{ref}-g)^2 \qquad \cdots (4\ 3)$$

由于E 是g 的平方函数，这样的g （使E g 最小）使E 为最小值。

根据利用第一种和第二种方法得到的S 和g ，利用如下方程（4 4 ）可以计算量化误差矢量e ，方程（4 4 ）为：

$$\underline{e}=\underline{r}-g\underline{s}_{syn} \qquad \cdots (4\ 4)$$

按第二级第二编码单元120$_2$的像第一级中一样的一个参考值进行这种量化。

即，提供到连接端305和307上的信号直接由第一级第二编码单元120$_1$的按加觉加权的合成滤波器312 提供到第二级第二编码单元120$_2$的按听觉加权的合成滤波器222。由第一级第二编码单元120$_1$求出的量化误差矢量e 提供到第二级第二编码单元120$_2$中的减法器323。

按照图10中的步骤S5，在第二级第二编码单元120$_2$中进行的处理与在第一级中进行的相似。即，由5 比特波形索引输出的随机的代码本320输出的代表性的数值送到增益电路321。在其中代码本320的代表性的数值输出与来自按3 比特增益索引输出的增益代码本325的增益相乘。加权的合成滤波器322的输出送到减法器323，在其中求出在按听觉加权的合成滤波器322的输出与第一级量化误差矢量之间的差。这个差送到用于距离计算的距离计算电路324，以便查找使量化误差能量E 为最小值的波形矢量S 和增益g。

随机的代码本310的波形索引输出和第一级第二编码单元120$_1$的增益代码本315的索引输出、随机的代码本320的索引输出和第二级第二编码单元120$_2$的增益代码本325的索引输出都送到索引输出转换电路330。如果由第二编码单元120输出23 比特，对第一级和第二级第二编码单元120$_1$、120$_2$的随机的代码本310、320的索引数据和增益代码本315、325的索引数据求和并输出。假如输出15 比特，输出第一级第二编码单元120$_1$的随机的代码本310和增益代码本315的索引数据。

然后，如在步骤S6 所示，刷新滤波器的状态。

在本实施例中，对于波形矢量，第二级第二编码单元120$_2$的索引比特数是小的为5，同对于增益的索引比特数也是小的为3。如果在这种情况下在代码本中没有出现适当的波形和增益，则量化误差很可能增加而不是降低。

虽然为了防止这种缺陷，可以使增益规定为0，对于增益仅有3 比特。如果将其中之一设定为0，则量化器性能明显劣化。考虑到这一点，将已经分配较大比特数的波形矢量设置为全—0 的矢量，除了全—0 矢量以外，进行上述查找，并且如果量化误差已经很大地增加，则选择该全—0 矢量。增益是任选

的。这样就使得能够在第二级第二编码单元120$_2$中防止量化误差增加。

虽然上面已经介绍了两级配置，但级数可以大于2。在这种情况下，如果通过第一级闭环查找矢量量化已接近终止。利用第（N－1）级的量化误差作为参考输入进行了第N级的量化，其中2＜N，并将第N级的量矢误差用作一个参考值输入到第（N＋1）级。

由图9和图10可以看出，对于第二编码单元通过采用多级矢量量化器，与利用具有相同比特数的纯粹的矢量量化或者利用对偶（conjugate）的代码本相比较，计算量被降低。特别是，在进行CELP编码时，其中通过合成法分析采用闭环查找进行沿时间轴的波形的矢量量化，减少查找操作次数是很关键的。此外，通过来自在两级第二编码单元120$_1$，120$_2$的两方的索引输出和仅采用第一级第二编码单元120$_1$的输出而没有采用第二级第二编码单元120$_2$的输出之间进行转换，可以易于转换比特数。如果将第一级和第二级第二编码单元120$_1$，120$_2$的索引输出组合并输出，解码器易于适应选择其中之一的索引输出的结构。即解码器易于适应利用按2kbps操作的解码器对利用6kbps进行编码的数据来解码。此外，如果在第二级第二编码单元120$_2$的波形代码本中包含零矢量，与假如将0添加到增益中相比，就变得能够在使性能较少劣化的情况下，防止量化误差增加。

随机的代码本的代码矢量例如可以通过对所谓的高斯噪声进行限制来产生。具体地说，通过产生高斯噪声、利用适当的阈值限制该高斯噪声以及使经限制的高斯噪声归一化可以形成该代码本。

然而，有不各类型的语音。例如，高斯噪声可以与接近噪声的例如“Sa，shi，Su，Se和So”的辅音的声语相适应，而该噪声不能与急剧上升的辅音的语音例如“Pa，Pi，Pu，Pe和Po”相适应。根据本发明，高斯噪声适用于某些代码矢量，而代码矢量的其余部分通过学习来处理，使可以适应具有明显上升的辅音和接近噪声的音这两者。假如例如阈值增加，得到具有几个较大峰值的这种矢量，而如果阈值降低，代码矢量接近高斯噪声。因此通过增加限制用的阈值的变化，变得能够适应具有明显升高部分的谐音，例如“Pa，Pi，Pu，Pe和Po”或者接近噪声的辅音，例如“Sa，Shi，Su，Se和So”，因此，提高清晰度。图11分别用实线和虚线表示高斯噪声和经限制的噪声的状态特性。图11A和11B表示利用等于1

的限制用的阈值即较大的阈值限制的噪声和利用等于0．4 的限制用的阈值即较小的阈值限制的噪声。由图1 1 A 和1 1 B 可以看出，如果阈值选择得较大，得到具有几个较大峰值的矢量，而如果阈值选择得较小，则噪声接近高斯噪声本身。

为了实现这一点，通过限制高斯噪声和设定适当数目的非学习型的代码矢量来准备起始的代码本。按照递增的方差值的顺序选择非学习型的代码矢量，用以适应于接近噪声的辅音例如“S a ，s h i ，S u ，S e 和S o ”。通过学习得到的矢量将L B G 算法用于学习。按照最接近的邻近条件编码采用了固定的代码矢量和通过学习得到代码矢量。在矩心条件下，仅刷新为了学习而设定的代码矢量。因此，为了学习而设定的代码矢量能够适应于急剧上升的辅音，例如“P a ，P i ，P u ，P e 和P o ”。

利用常规的学习方式，对于这些代码矢量可以得到最佳增益。

图1 2 表示通过限制高斯噪声来构成代码本的处理流程图。

在图1 2 中，为了起始化，在步骤S 1 0 将学习的次数n 设定为n =0 。利用误差$D_0=\infty$，设定学习的最大次数n_{max}，并设定一个确定学习结束条件的阈值ϵ。

在下一个步骤S 1 1 ，通过限制高斯噪声形成起始代码本。在步骤S 1 2 ，将部分的代码矢量固定作为非学习型的代码矢量。

在下一个步骤S 1 3 ，利用上述代码本完成编码。在步骤S 1 4 ，计算误差。在步骤S 1 5 ，判别是 $(D_{n-1}-D_n)/D_n<\epsilon$ 还是$n=n_{max}$。如果结果是Y E S ，过程终止。如果结果是N o ，过程转换到步骤S 1 6 。

在步骤S 1 6 ，处理设有用于编码的代码矢量。在下一个步骤S 1 7 ，刷新代码本，在步骤S 1 8 ，在返回到步骤S 1 3 之前，递增学习的次数。

上述信号编码和信号解码装置可以用作例如用在便携式通信终端或便携式电话装置中的语音代码本，如图1 4 所示。

图1 3 表示采用图1 和图3 中所示结构的语音编码单元1 6 0 的便携式终端的发送侧。由拾音器1 6 1 获得的语音信号利用放大器1 6 2 放大并利用模/数 (A /D) 变换器1 6 3 变换为数字信号，再送到按图1 和图3 所示构成的语音编码单元1 6 0 。来自A / D 变换器1 6 3 的数字信号提供到输入端1 0 1 。语音编码单元1 6 0 按照结合图1 和3 所解释的方式进行编码。图1 和3

中的输出端的输出信号作为语音编码单元160的输出信号送到传输通道编码单元164，然后在其中对所提供的信号进行通道编码。传输通道编码单元164的输出信号送到用于调制的调制电路165，并由此经过数/模（D/A）变换器166和RF放大器167提供到天线168。

图14表示采用按照图4所示构成的语音解码单元260的便携式终端的接收侧。由图14中的天线261接受的语音信号由RF放大器262放大，并经过模/数（D/A）变换器263送到解调电路264，再由此将解调的信号送到传输通道解码单元265。解码单元265的输出信号提供到按图2和4所示构成的语声解码单元260。语声解码单元260按照结合图2和4所解释的方式对信号进行解码。在图2和4的输出端的输出信号作为语声解码单元260的信号送到数/模（D/A）变换器266。来自D/A变换器266的模拟语音信号送到扬声器268。

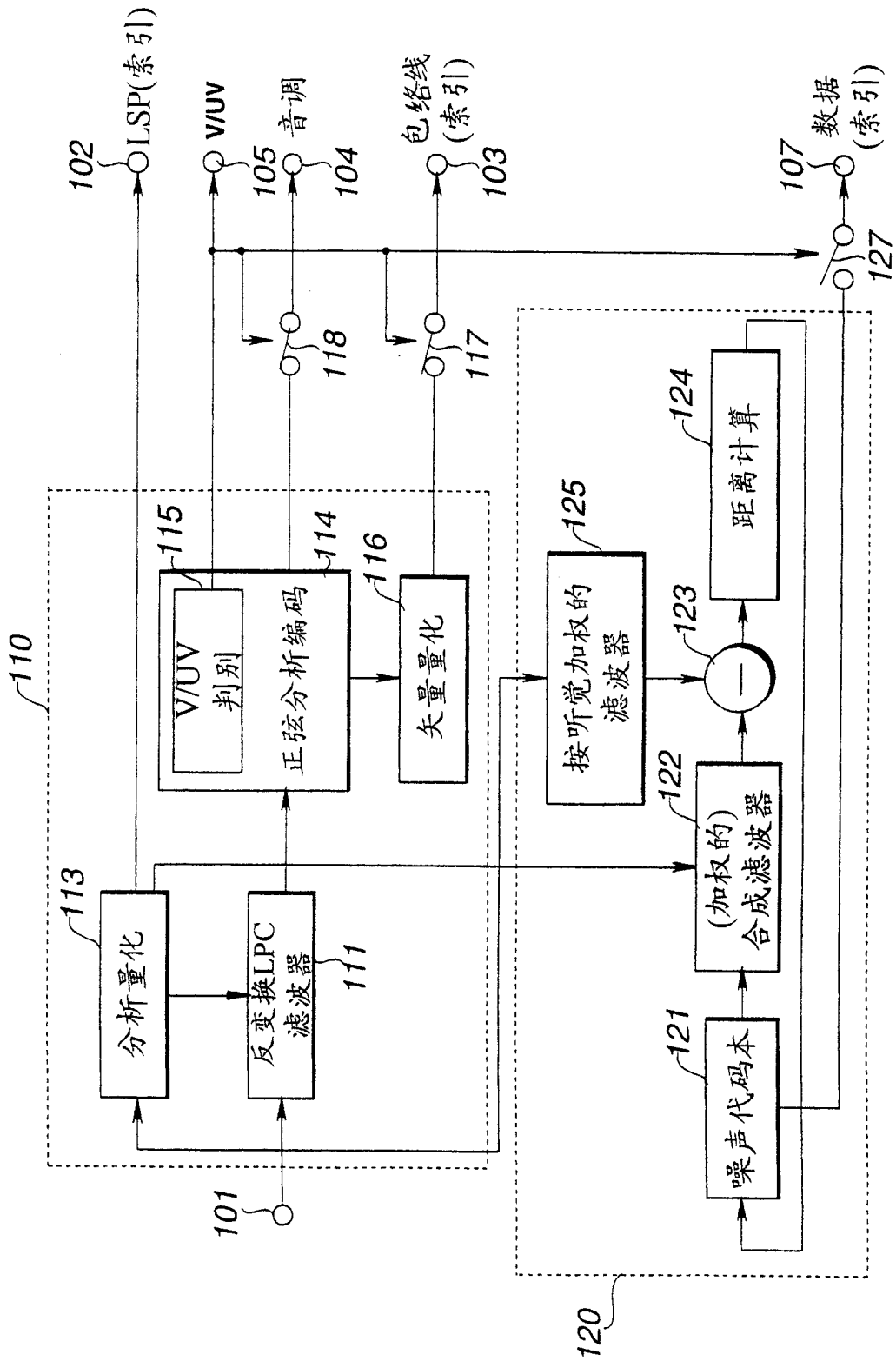

102
LSP(索引)
V/UV
105
音调
104
色络线
(索引)
103
107
数据
(索引)
127
118
117
110
115
V/UV
判别
正弦分析编码
114
116
矢量量化
113
分析量化
反变换LPC
滤波器
111
101
120
125
按听觉加权的
滤波器
123
124
距离计算
122
(加权的)
合成滤波器
121
噪声代码本

图 1

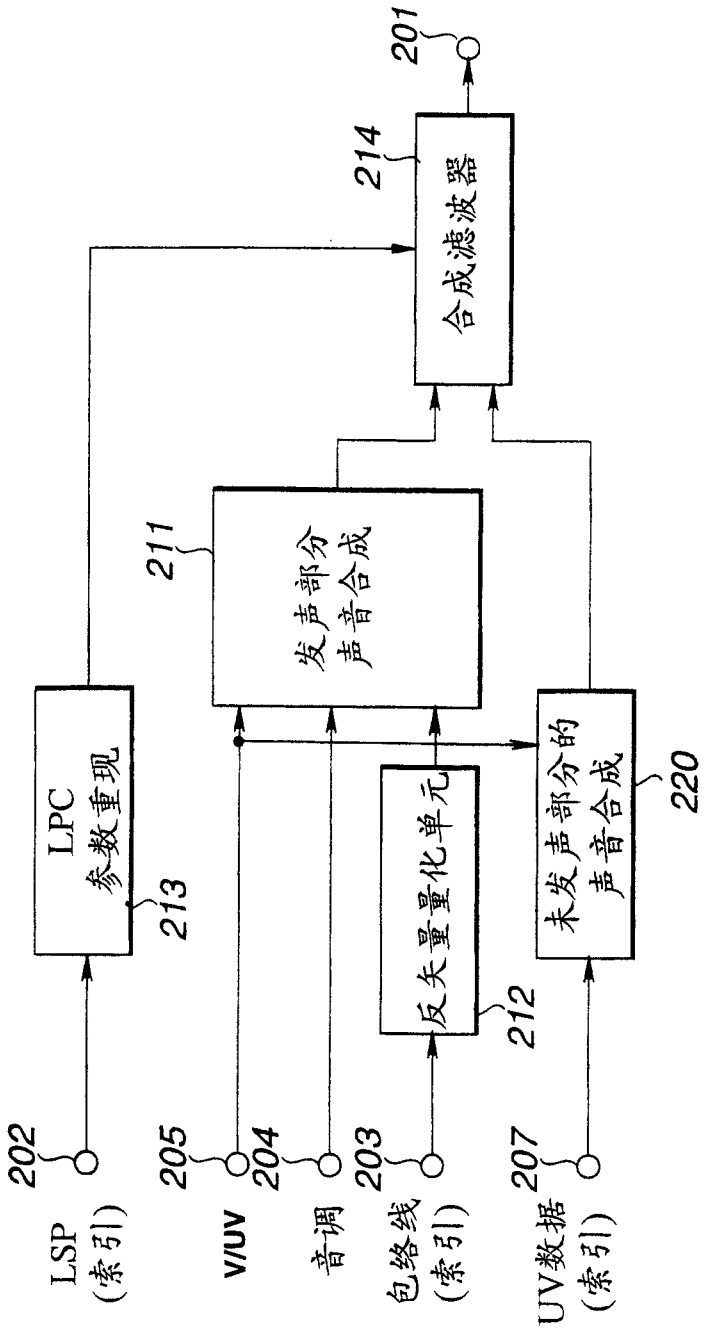

201
214
合成滤波器
211
发声部分
声音合成
LPC
参数重现
213
反矢量量化单元
212
未发声部分的
声音合成
220
202
205
204
203
207
LSP
(索引)
V/UV
音调
包络线
(索引)
UV数据
(索引)

图 2

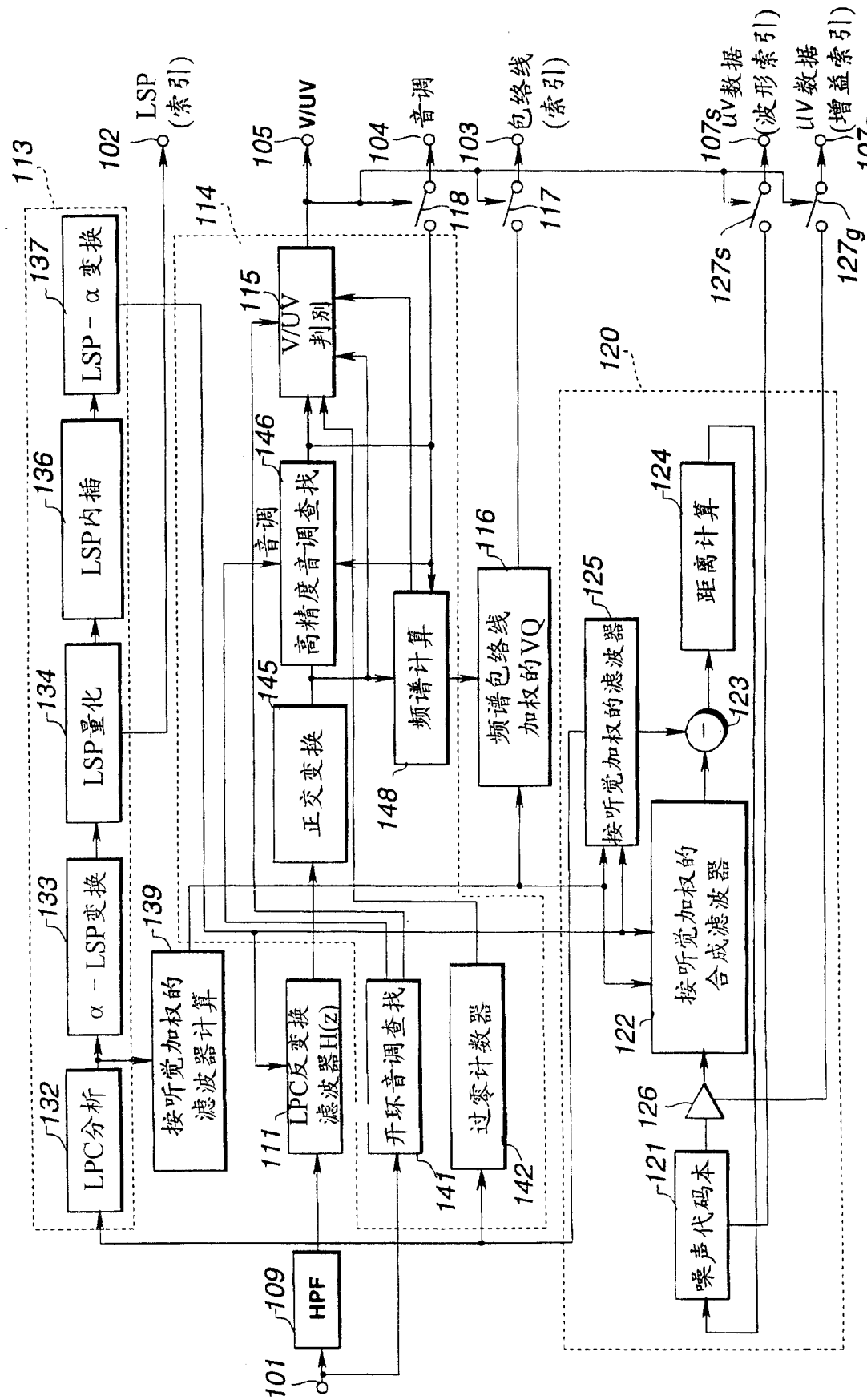

101
109
HPF
111
LPC反变换滤波器H(z)
132
LPC分析
133
α－LSP变换
134
LSP量化
136
LSP内插
137
LSP－α变换
139
按听觉加权的滤波器计算
141
开环音调查找
142
过零计数器
145
正交变换
146
高精度音调查找
148
频谱计算
116
频谱包络线加权的VQ
115
V/UV判别
120
按听觉加权的滤波器
121
噪声代码本
122
按听觉加权的合成滤波器
123
124
距离计算
125
126
113
114
102
LSP（索引）
105
VIUV
104
音调
103
包络线（索引）
118
117
107s
UV数据（波形索引）
127s
107g
UV数据（增益索引）
127g

图 3

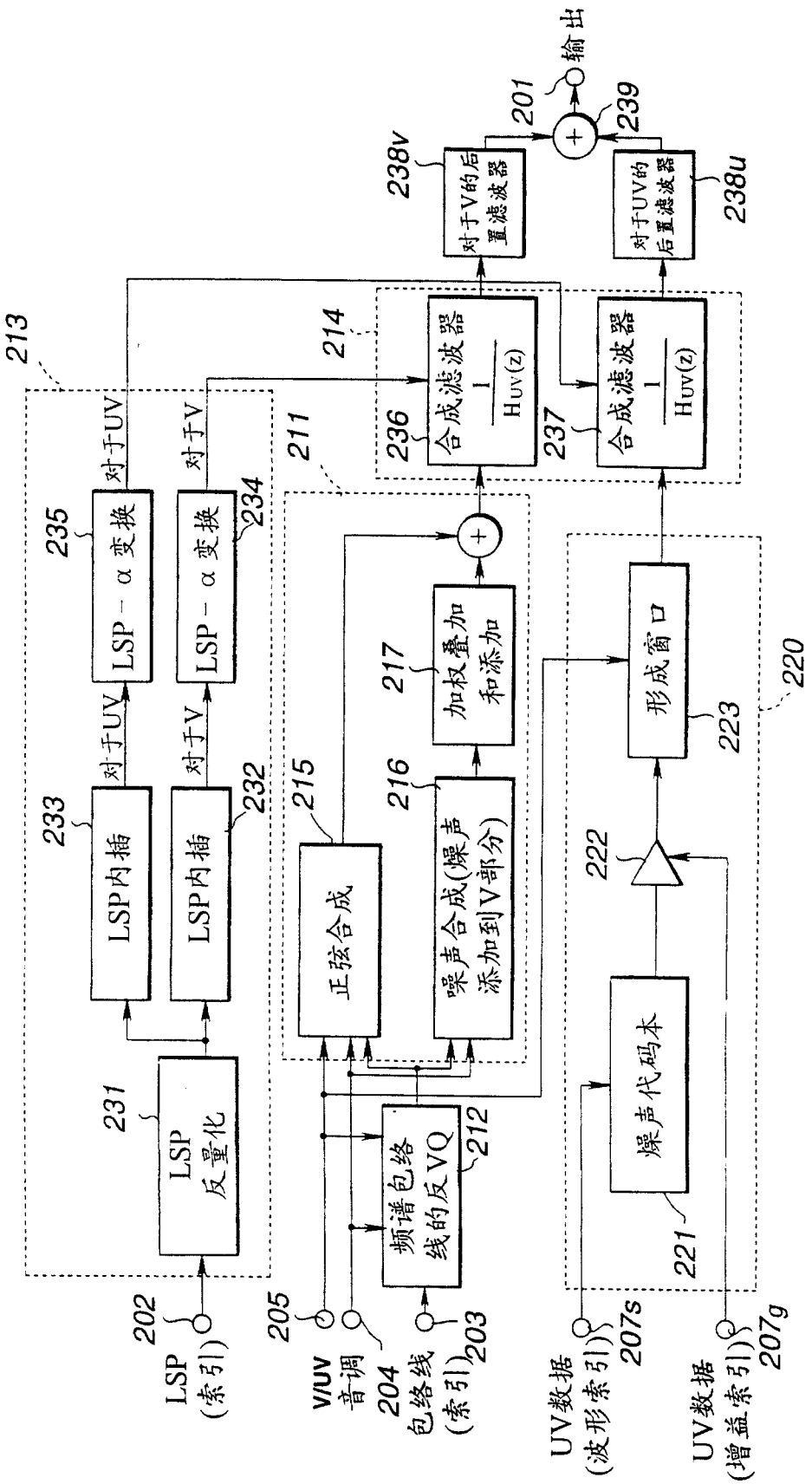

LSP
(索引)
202
LSP
反量化
231
LSP内插
233
LSP内插
232
对于UV
对于V
LSP－α变换
235
LSP－α变换
234
对于UV
对于V
213
V/UV
音调
205
204
包络线
(索引)
203
频谱包络
线的反VQ
212
正弦合成
215
噪声合成(燥声
添加到V部分)
216
加权叠加
和添加
217
211
合成滤波器
$\frac{1}{H_{UV}(z)}$
236
合成滤波器
$\frac{1}{H_{UV}(z)}$
237
214
对于V的后
置滤波器
238v
对于UV的
后置滤波器
238u
201
239
输出
UV数据
(波形索引)
207s
UV数据
(增益索引)
207g
燥声代码本
221
222
形成窗口
223
220

图 4

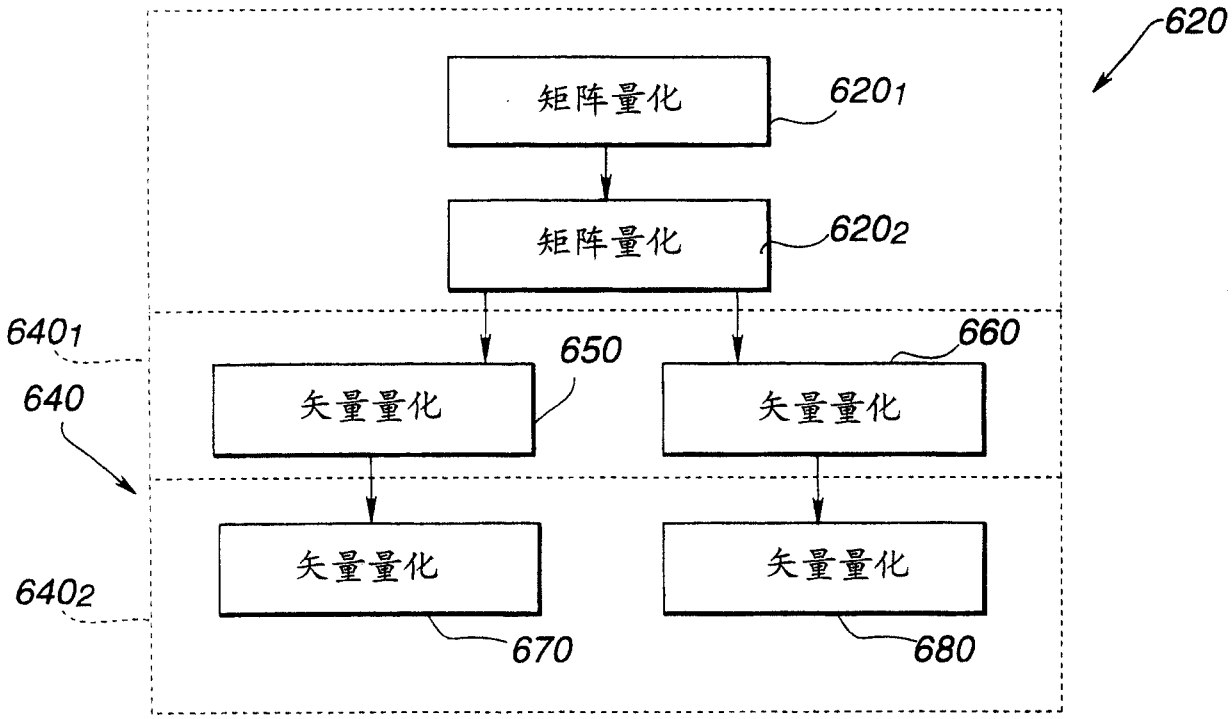

620
矩阵量化
6201
矩阵量化
6202
6401
640
650
660
矢量量化
矢量量化
矢量量化
矢量量化
6402
670
680

图 5

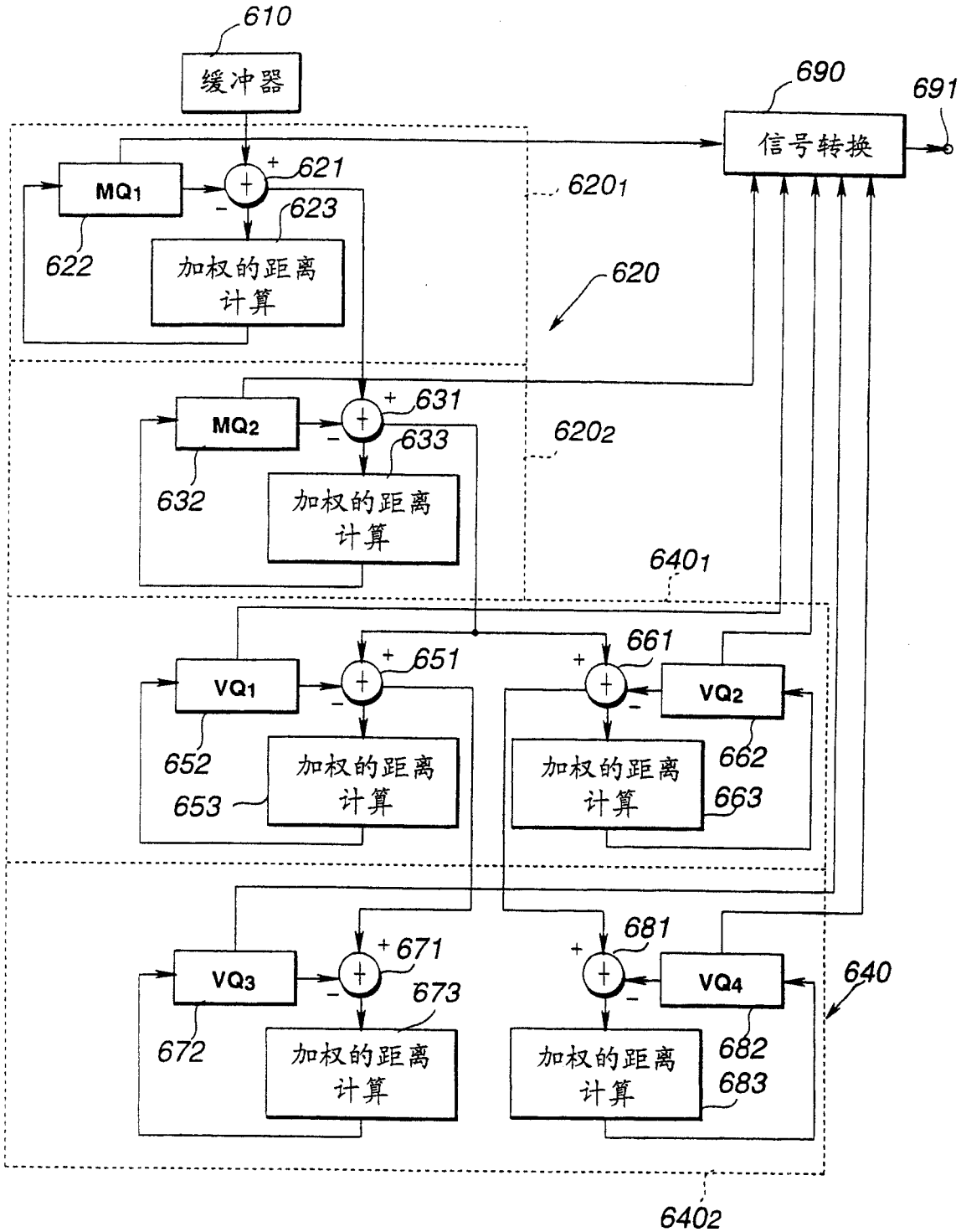

610
缓冲器
690
691
信号转换
621
MQ1
623
622
加权的距离计算
6201
620
631
MQ2
633
632
加权的距离计算
6202
6401
651
661
VQ1
VQ2
652
653
加权的距离计算
662
663
加权的距离计算
671
681
VQ3
VQ4
673
672
加权的距离计算
640
682
683
加权的距离计算
6402

图 6

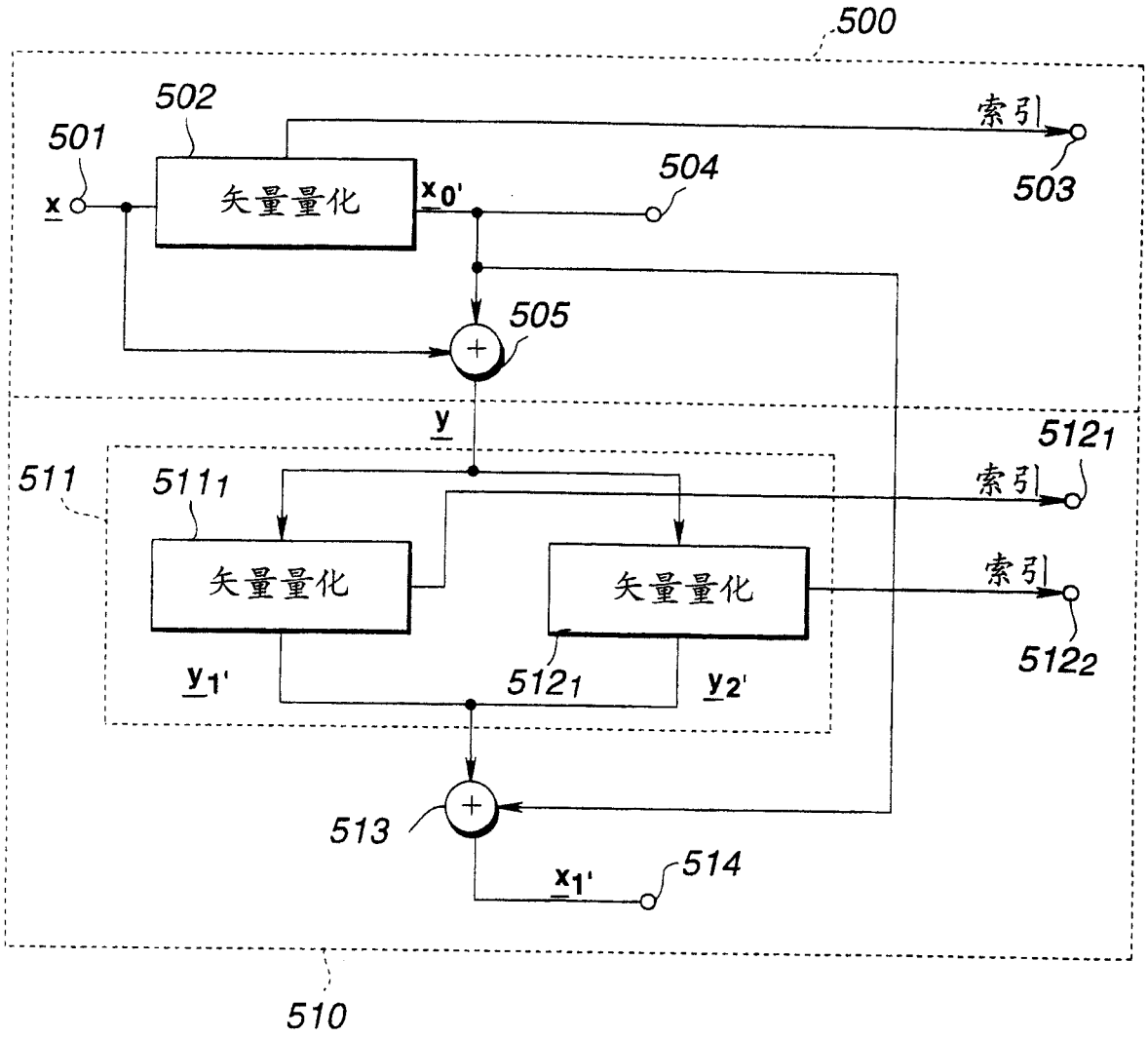

500
502
501
索引
x
矢量量化
x_0'
504
503
505
y
512_1
511
511_1
索引
矢量量化
矢量量化
索引
y_1'
512_1
y_2'
512_2
513
514
x_1'
510

图 7

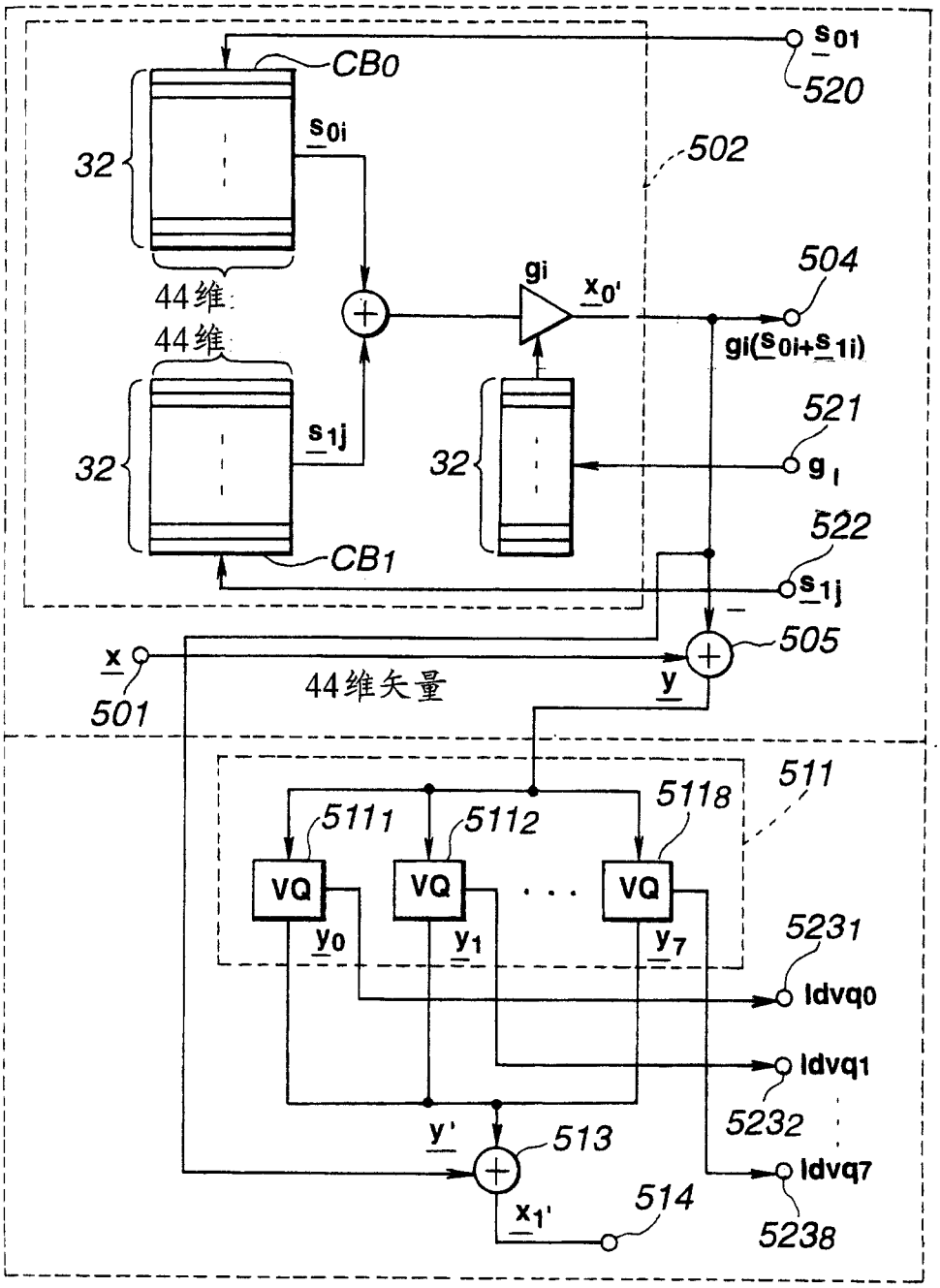

CB0
s0i
32
44维
44维
s1j
32
CB1
gi
x0'
32
s01
520
502
504
gi(s0i+s1i)
521
gi
522
s1j
505
x
501
44维矢量
y
511
5111
5112
5118
VQ
VQ
VQ
y0
y1
y7
5231
Idvq0
Idvq1
5232
Idvq7
5238
y'
513
514
x1'

图 8

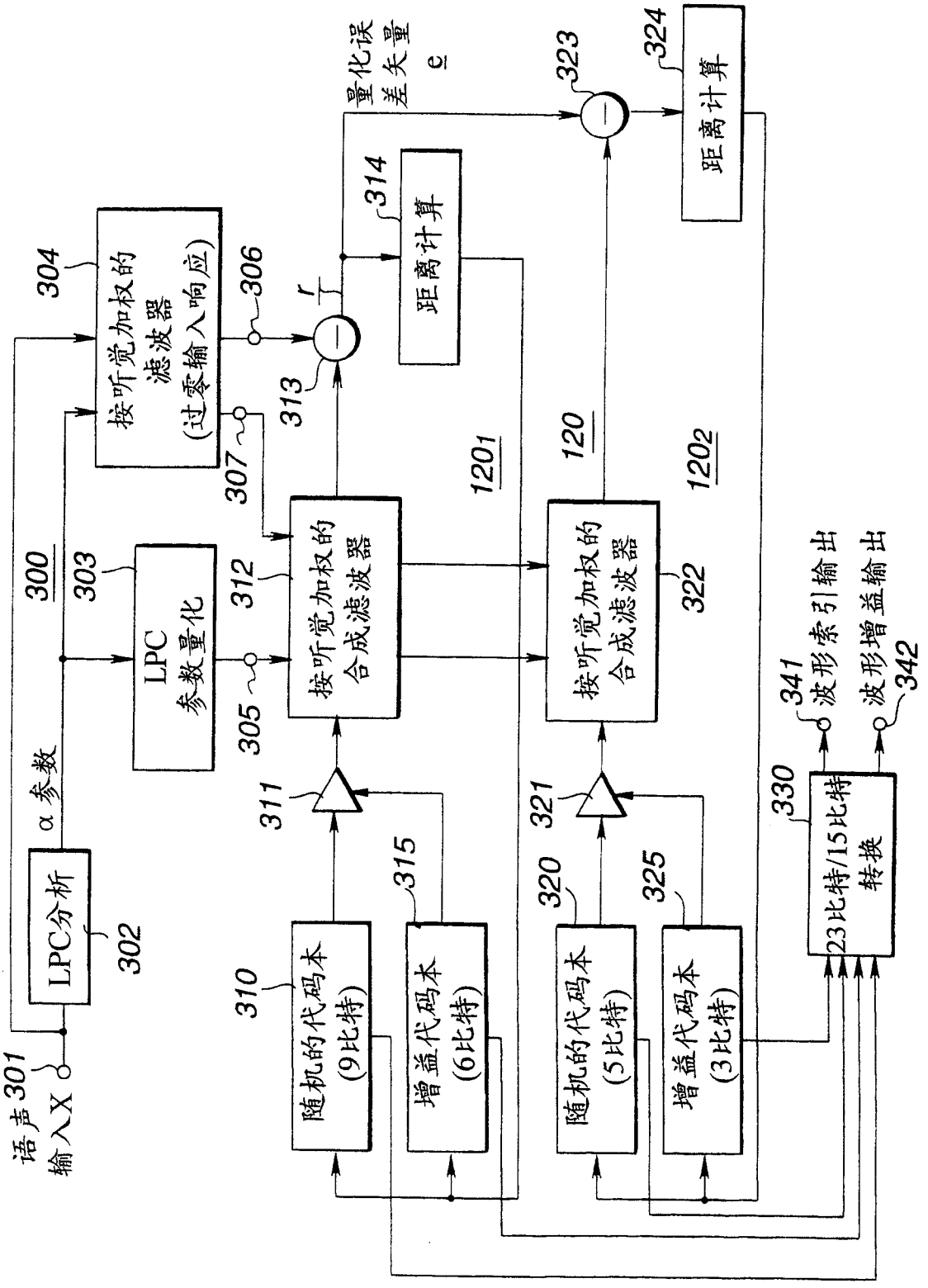

语声输入X
301
LPC分析
302
α参数
300
303
LPC参数量化
304
按听觉加权的滤波器(过零输入响应)
305
306
307
312
按听觉加权的合成滤波器
313
r
314
距离计算
量化误差矢量e
323
324
距离计算
120
120₁
120₂
311
315
310
随机的代码本(9比特)
增益代码本(6比特)
321
320
325
随机的代码本(5比特)
增益代码本(3比特)
322
按听觉加权的合成滤波器
330
23比特/15比特转换
341
波形索引输出
342
波形增益输出

图 9

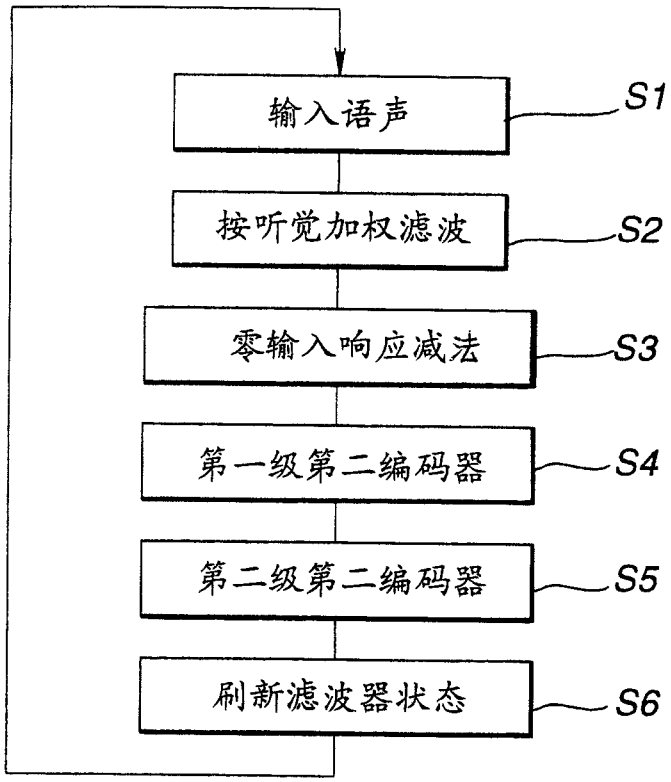

输入语声
S1
按听觉加权滤波
S2
零输入响应减法
S3
第一级第二编码器
S4
第二级第二编码器
S5
刷新滤波器状态
S6

图 10

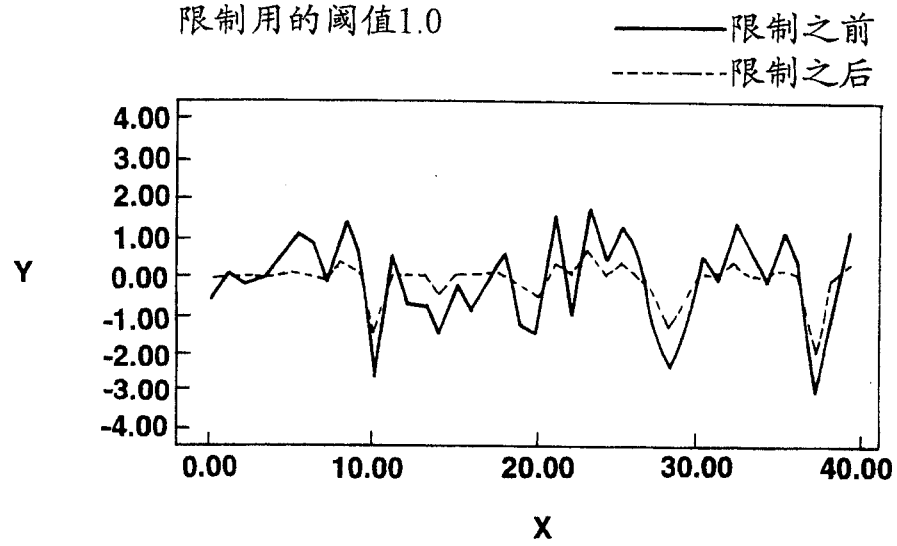

限制用的阈值1.0
限制之前
限制之后
Y
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
0.00
10.00
20.00
30.00
40.00
X

图 11A

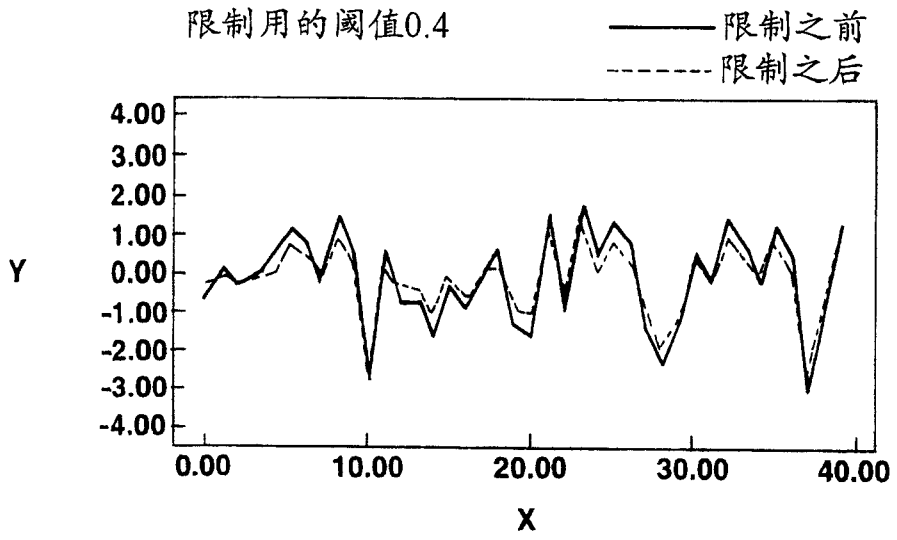

限制用的阈值0.4
限制之前
限制之后
Y
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
0.00
10.00
20.00
30.00
40.00
X

图 11B

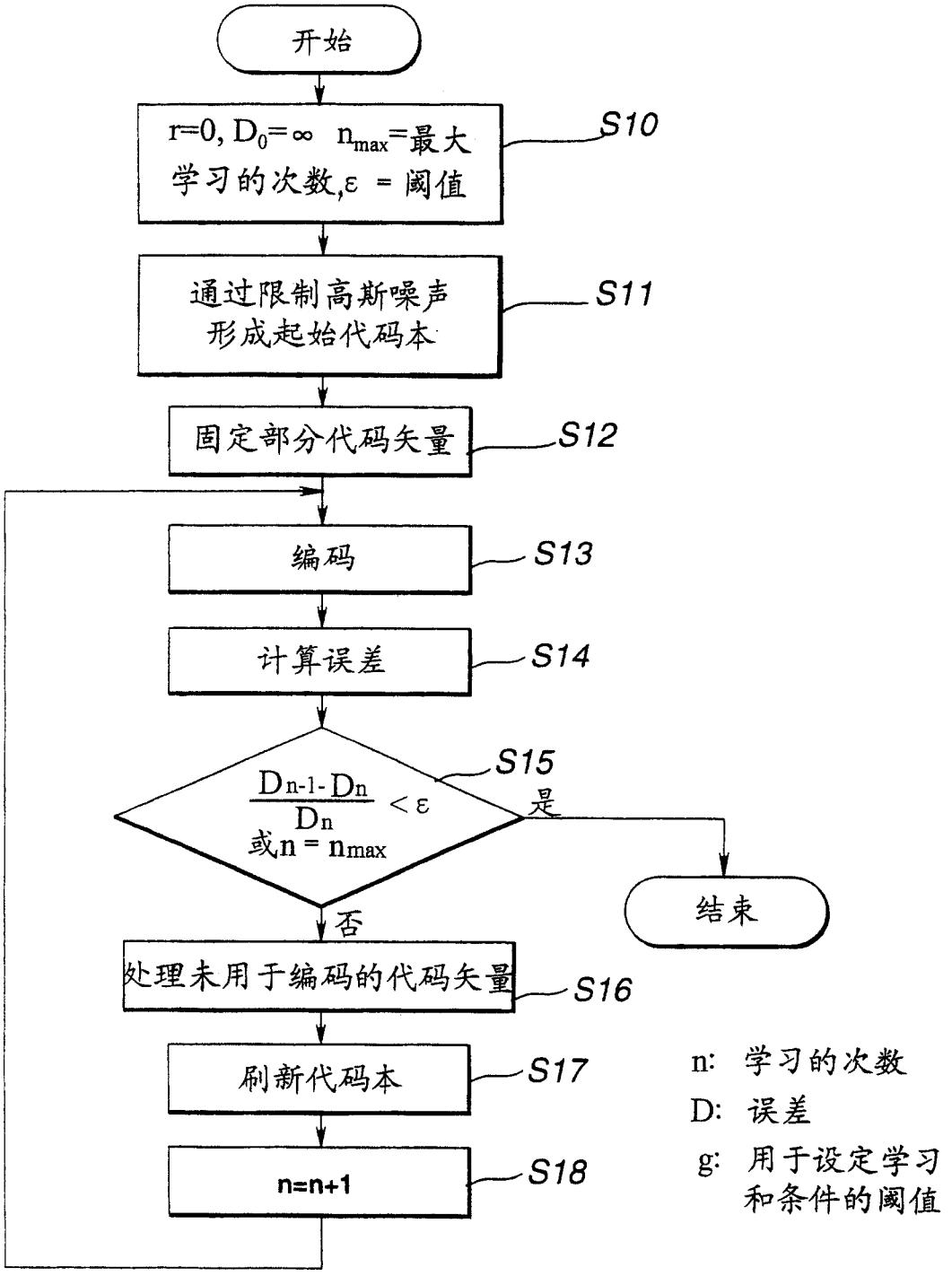

开始
r=0, $D_0=\infty$ n_{max}=最大学习的次数,ε = 阈值
S10
通过限制高斯噪声形成起始代码本
S11
固定部分代码矢量
S12
编码
S13
计算误差
S14
$\frac{D_{n-1}-D_n}{D_n} < \varepsilon$ 或n = n_{max}
S15
是
结束
否
处理未用于编码的代码矢量
S16
刷新代码本
S17
n=n+1
S18
n: 学习的次数
D: 误差
g: 用于设定学习和条件的阈值

图 12

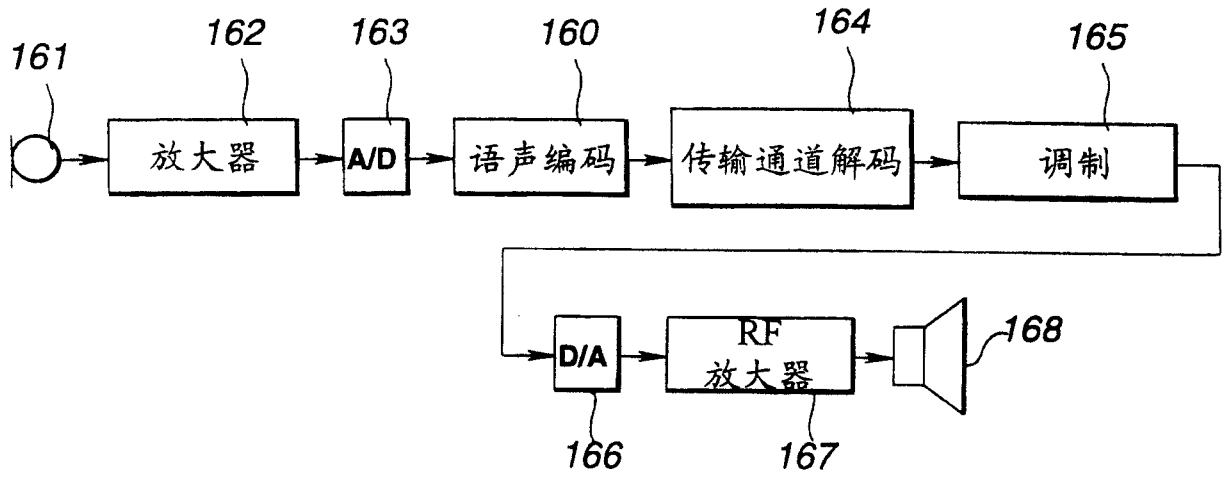

161
162
163
160
164
165
放大器
A/D
语声编码
传输通道解码
调制
D/A
RF
放大器
168
166
167

图 13

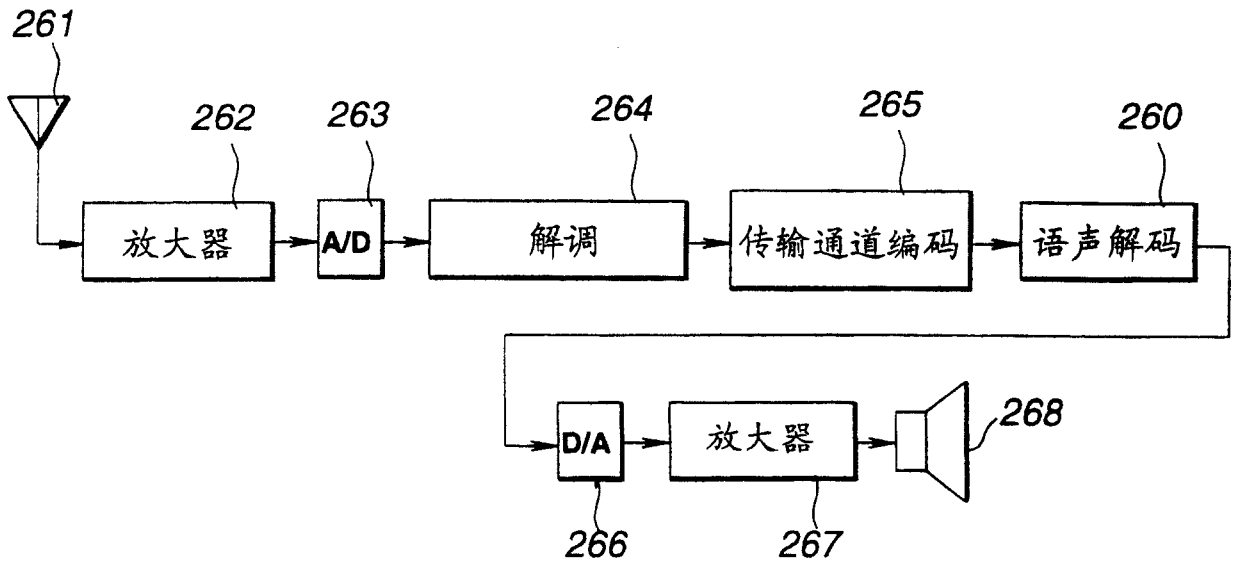

261
262
263
264
265
260
放大器
A/D
解调
传输通道编码
语声解码
D/A
放大器
268
266
267

图 14

	2kbps	6kbps
U/V鉴别输出	1比特/20毫秒	1比特/20毫秒
LSP量化索引	32比特/40毫秒	48比特/40毫秒
对于发声部分的语声 (V)	音调数据8比特/40毫秒	音调数据8比特/20毫秒
	频谱包络线VQ 索引15比特/20毫秒	频谱包络线VQ 索引87比特/20毫秒
	波形(对第一级)，5 + 5 比特/20毫秒 增益，5比特/20毫秒	波形(对第一级)，5 + 5 比特/20毫秒 增益，5比特/20毫秒 索引(对第二级)，72 比特/20毫秒
对于未发声部分的语声 (UV)	时域VQ 索引11比特/10毫秒	时域VQ 索引23比特/5毫秒
	波形(对第一级)，7 比特/10毫秒 增益，4比特/10毫秒	波形(对第一级)，9比特/5毫秒 增益，6比特/5毫秒 波形(对第二 级) 5比特/5毫秒 增益，3比特/5毫秒
对于发声部分的语声 对于未发声部分的语声	40比特/20毫秒 39比特/20毫秒	120比特/20毫秒 117比特/20毫秒

图 15