

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G10L 19/10 (2006.01)

G10L 19/12 (2006.01)



# [12] 发明专利说明书

专利号 ZL 200510006202.8

[45] 授权公告日 2007 年 11 月 14 日

[11] 授权公告号 CN 100349208C

[22] 申请日 1998.10.22

JP 2-282800A 1990.11.20

[21] 申请号 200510006202.8

JP 8-8753A 1996.1.12

分案原申请号 98801556.0

审查员 臧自欣

[30] 优先权

[74] 专利代理机构 上海专利商标事务所有限公司

[32] 1997.10.22 [33] JP [31] 289412/97

代理人 沈昭坤

[32] 1997.10.28 [33] JP [31] 295130/97

[32] 1998.3.31 [33] JP [31] 85717/98

[73] 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

[56] 参考文献

US 5195137A 1993.3.16

JP 6-202699A 1994.7.22

CN 1148449A 1997.4.23

JP 2-280200A 1990.11.16

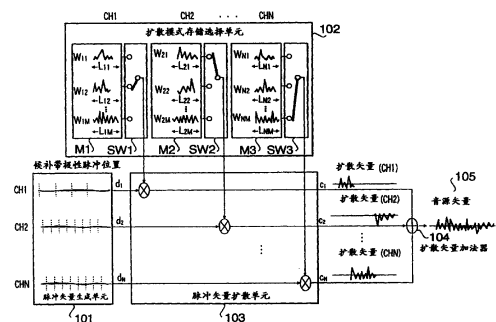
权利要求书 4 页 说明书 29 页 附图 14 页

[54] 发明名称

扩散矢量生成方法及扩散矢量生成装置

[57] 摘要

一种音源矢量生成装置，其中包括：具有 N 个生成脉冲矢量的通道 ( $N \geq 1$ ) 的脉冲矢量生成单元；对应于 N 个通道，每一通道存储 M 种 ( $M \geq 1$ ) 扩散模式的存储单元；从存储单元，每一通道选择性地取出扩散模式的选择单元；每一通道进行所取出的扩散模式和所生成的脉冲矢量的叠加运算，并生成 N 个扩散矢量的扩散单元；根据所生成的 N 个扩散矢量，生成音源矢量的音源矢量生成单元。



1、一种用于话音解码的扩散矢量生成方法，其特征在于，该方法包含下述步骤：

提供在矢量轴的一要素上具有带极性单元脉冲的脉冲矢量；

把自适应码本增益值与预先设定的门限值进行比较；

根据所述比较的结果，从存储在存储器中的多个扩散模式中选择一种扩散模式；

通过对所述脉冲矢量和所述选择的扩散模式进行卷积运算，生成扩散矢量；

其中，在所述选择步骤中，所述自适应码本增益和至少一个表示所述输入话音短时特性的其它参数用作选择判据。

2、一种用于话音解码的扩散矢量生成方法，其特征在于，该方法包含下述步骤：

提供在矢量轴的一要素上具有带极性单元脉冲的脉冲矢量；

把自适应码本增益值与预先设定的门限值进行比较；

根据所述比较的结果，从存储在存储器中的多个扩散模式中选择一种扩散模式；

通过对所述脉冲矢量和所述选择的扩散模式进行卷积运算，生成扩散矢量；

其中，所述选择步骤中选择的所述扩散模式，在所述自适应码本增益值大于所述预先设定的门限值时，为脉冲形扩散模式，在所述自适应码本增益值小于所述预先设定的门限值时，为随机形扩散模式。

3、一种用于话音解码的扩散矢量生成方法，其特征在于，该方法包含下述步骤：

接收与脉冲位置和脉冲极性相应的码号；

相应于所述接收的码号，提供在矢量轴的一要素上具有带极性单元脉冲的脉冲矢量；

把自适应码本增益值与预先设定的门限值进行比较；

根据所述比较的结果，从存储在存储器中的多个扩散模式中选择一种扩散

模式；

通过对所述脉冲矢量和所述选择的扩散模式进行卷积运算，生成扩散矢量；

其中，在所述选择步骤中，所述自适应码本增益和至少一个表示所述输入话音短时特性的其它参数用作选择判据。

4、如权利要求 3 所述的扩散矢量生成方法，其特征在于，所述脉冲矢量由代数码本提供。

5、如权利要求 3 所述的扩散矢量生成方法，其特征在于，所述选择步骤中选择的所述扩散模式，在所述自适应码本增益值大于所述预先设定的门限值时，为脉冲形扩散模式，在所述自适应码本增益值小于所述预先设定的门限值时，为随机形扩散模式。

6、一种 CELP 型话音解码方法，其特征在于，该方法包含下述步骤：

接收通过 CELP 型话音编码产生的代码；

把所述代码分解成线性预测码本索引、自适应码本索引、随机噪声码本索引和加权码本索引；

读取与所述自适应码本索引对应的自适应码矢量；

用扩散矢量生成方法生成随机码矢量；

所述扩散矢量生成方法包含下述步骤：

提供在矢量轴的一要素上具有带极性单元脉冲的脉冲矢量，其中，所述带极性单元脉冲的位置和极性相应于所述随机噪声码本索引；

把自适应码本增益值与预先设定的门限值进行比较；

根据所述比较的结果，从存储在存储器中的多个扩散模式中选择一种扩散模式；

通过对所述脉冲矢量和所述选择的扩散模式进行卷积运算，生成扩散矢量；

该 CELP 型话音解码方法还包含下述步骤：

读取与所述加权码本索引对应的自适应码本增益和随机噪声码本增益；

把所述自适应码矢量与所述自适应码本增益相乘，以生成乘以了增益的自适应码矢量；

把所述随机码矢量与所述随机噪声码本增益相乘，以生成乘以了增益的随机码矢量；

把所述乘以了增益的自适应码矢量和所述乘以了增益的随机码矢量相加，以生成音源矢量；

由所述音源矢量激励合成滤波器，以生成合成语音；

其中，在所述选择步骤中，所述自适应码本增益和至少一个表示所述输入语音短时特性的其它参数用作选择判据。

7、如权利要求 6 所述的 CELP 型语音解码方法，其特征在于，所述脉冲矢量由代数码本提供。

8、一种用于语音解码器以改进语音质量的扩散矢量生成装置，其特征在于，该装置包含：

接收器，接收与脉冲位置和脉冲极性相应的码号；

脉冲矢量提供单元，相应于所述接收的码号，提供在矢量轴的一要素上具有带极性单元脉冲的脉冲矢量；

比较器，把自适应码本增益值与预先设定的门限值进行比较；

选择器，根据所述比较的结果，从存储在存储器中的多个扩散模式中选择一种扩散模式；

扩散矢量生成单元，通过对所述脉冲矢量和所述选择的扩散模式进行卷积运算，生成扩散矢量；

其中，在所述选择中，所述自适应码本增益和至少一个表示所述输入语音短时特性的其它参数用作选择判据。

9、如权利要求 8 所述的扩散矢量生成装置，其特征在于，所述脉冲矢量由代数码本提供。

10、一种 CELP 型语音解码器，其特征在于，该解码器包含：

接收器，接收通过 CELP 型语音编码器产生的代码；

分解单元，把所述代码分解成线性预测码本索引、自适应码本索引、随机噪声码本索引和加权码本索引；

自适应码矢量读取单元，读取与所述自适应码本索引对应的自适应码矢量；

随机码矢量生成单元，用扩散矢量生成装置生成随机码矢量；

所述扩散矢量生成装置包含：

脉冲矢量提供单元，提供在矢量轴的一要素上具有带极性单元脉冲的脉冲矢量，其中，所述带极性单元脉冲的位置和极性相应于所述随机噪声码本索引；

比较器，把自适应码本增益值与预先设定的门限值进行比较；

选择器，根据所述比较的结果，从存储在存储器中的多个扩散模式中选择一种扩散模式；

扩散矢量生成单元，通过对所述脉冲矢量和所述选择的扩散模式进行卷积运算，生成扩散矢量；

该 CELP 型话音解码器还包含：

增益读取单元，读取与所述加权码本索引对应的自适应码本增益和随机噪声码本增益；

自适应码矢量增益乘法器，把所述自适应码矢量与所述自适应码本增益相乘，以生成乘以了增益的自适应码矢量；

随机码矢量增益乘法器，把所述随机码矢量与所述随机噪声码本增益相乘，以生成乘以了增益的随机码矢量；

加法器，把所述乘以了增益的自适应码矢量和所述乘以了增益的随机码矢量相加，以生成音源矢量；

由所述音源矢量激励以生成合成话音的合成滤波器；

其中，在所述选择中，所述自适应码本增益和至少一个表示所述输入话音短时特性的其它参数用作选择判据。

11、如权利要求 10 所述的 CELP 型话音解码器，其特征在于，所述脉冲矢量由代数码本提供。

## 扩散矢量生成方法及扩散矢量生成装置

本申请是申请日为 1998 年 10 月 22 日、申请号为 98801556.0、发明名称为“话音信号编码器和话音信号解码器”的发明专利申请的分案申请。

### 技术领域

本发明涉及高效编码和解码话音信息用的扩散矢量生成方法和扩散矢量生成装置。

### 背景技术

当今，正在开发高效编码和解码话音信息用的话音编码技术。《码激励线性预测：低比特率高质量语音》（Code Excited Linear Prediction: High Quality Speech at Low Bit Rate）（M. R. Schroeder 著；发表于 ICASSP'85, pp. 937 ~ 940）中记载着基于这种话音编码技术的 CELP 型话音信号编码器。此话音信号编码器对用固定时间划分输入话音所得的每一帧进行线性预测，由每帧的线性预测求预测残差（激励信号），并用存放过去驱动音源的自适应码本和存放多个噪声码矢量的噪声码本将该预测残差编码。

图 1 中示出以往的 CELP 型话音信号编码器（下文简称为“话音编码器”）的功能框图。

线性预测分析单元 12 对此 CELP 型话音信号编码器中输入的话音信号 11 进行线性预测分析。利用该线性预测分析，可取得线性预测系数。线性预测系数为表示话音信号 11 的频谱包络特性的参数。线性预测分析单元 12 所得的线性预测系数在线性预测系数编码单元 13 进行量化，并将量化后的线性预测系数送到线性预测系数解码单元 14。又将量化所得量化号作为线性预测码输入到编码输出单元 24。线性预测系数解码单元 14 将线性预测系数编码单元 13 所得量化线性预测系数解码后，取得合成滤波器的系数，线性预测系数解码单元 14 将合成滤波器的系数输出到合成滤波器 15。

自适应码本 17 为输出多种候补自适应码矢量的码本，由存放过去的多帧驱动音源的缓存器构成。自适应码矢量为表现输入话音中周期分量的时间序列矢量。

噪声码本 18 为存放多种候补噪声码矢量的码本，其种类对应于所分配的比特数。噪声码矢量为表现输入话音中非周期分量的时间序列矢量。

自适应码增益加权单元 19 和噪声码增益加权单元 20 分别对自适应码本 17 和噪声码本 18 输出的候补矢量乘以从加权码本 21 读出的自适应增益和噪声码增益后，输出到加法器 22。

加权码本是一种存储器，分别存放多种与候补自适应码矢量相乘的加权数和与候补噪声码矢量相乘的加权数，其种类对应于所分配的比特数。

加法器 22 将分别在自适应码增益加权单元 19、噪声码增益加权单元 20 加权后的候补自适应码矢量和候补噪声码矢量相加，产生候补驱动音源矢量，并输出到合成滤波器 15。

合成滤波器 15 为由线性预测系数解码单元 14 所得的合成滤波器系数构成的全极型滤波器。合成滤波器 15 具有一种功能，输入来自加法器 22 的候补驱动音源矢量时，输出候补合成话音矢量。

失真计算单元 16 计算合成滤波器 15 的输出(即候补合成话音矢量)与输入话音 11 之间的失真，所得失真值输出到码号规定单元 23。码号规定单元 23 分别对三种码本(自适应码本、噪声码本和加权码本)规定使失真计算单元 16 中算出的失真最小的三种码号(自适应码号、噪声码号和加权码号)。然后，将码号规定单元 23 所规定的三种码号输出到编码输出单元 24。编码输出单元 24 汇合线性预测系数编码单元 13 所得的线性预测码号以及码号规定单元 23 所规定的自适应码号、噪声码号和加权码号，并输出到传输线路。

图 2 中示出对上述编码器编码的信号进行解码的 CELP 型话音信号解码器(下文简称为“话音解码器”)的功能框图。此话音信号解码器中，编码输入单元 31 接收话音信号编码器(图 1)送来的编码，将接收到的编码分解为线性预测码号、自适应码号、噪声码号和加权码号，并将分解所得编码分别输出到线性预测系数解码单元 32、自适应码本 33、噪声码本 34 和加权码本 35。

接着，线性预测系数解码单元 32 将编码输入单元 31 所得线性预测码号解码后，取得合成滤波器系数，并输出到合成滤波器 39。然后，从自适应码本中与自适应码号对应的位置读出自适应码矢量，从噪声码本读出与噪声码号对应的噪声码矢量，进而从加权码本读出加权码号对应的自适应码增益和噪声码增益。而且，在自适应码加权单元 36 将自适应码矢量乘以自适应码增益后，送到加法器 38。同样，在噪声码矢量加权单元 37 将噪声码矢量乘以噪声码增益

后，送到加法器 38。

加法器 38 将上述两个编码矢量相加后，产生驱动音源矢量，并将产生的驱动音源送到自适应码本 33，以更新缓存器，该驱动音源也送到合成滤波器 39，以驱动滤波器。合成滤波器 39 由加法器 38 所得驱动音源矢量驱动，并用线性预测系数解码单元 32 的输出再现合成话音。

在 CELP 型话音信号编码器的失真计算单元 16 一般利用下式(式 1)计算所求失真 E:

$$E = \left\| V - (gaHP + gcHC) \right\|^2 \quad (1)$$

V: 输入话音信号(矢量)

H: 合成滤波器脉冲响应卷积矩阵

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

其中，h 为合成滤波器的脉冲响应(矢量)，L 为帧长度，

p: 自适应码矢量

c: 噪声码矢量

ga: 自适应码增益

gc: 噪声码增益

这里，为了使式 (1) 的失真 E 最小，需要对自适应码号、噪声码号和加权码号的全部组合用闭环计算失真，规定各码号。

然而，对式(1)进行闭环检索，则运算处理量过大，因而一般首先用自适应码本由矢量量化规定自适应码号，其次由采用噪声码本的矢量量化规定噪声码号，最后由采用加权码本的矢量量化规定加权码号。现就这种情况下，对采用噪声码本的矢量量化处理作进一步详细说明。

自适应码号和自适应码增益预先确定或暂时确定时，式(1)的失真估算式变为下面的式(2)。

$$E_c = \left\| X - g_c H C \right\| \quad (2)$$

其中，式(2)内的矢量  $X$  为采用预先规定或暂时规定的自适应码号和自适应码增益，并由下列式(3)求得的噪声源信息(规定噪声码号用的目标矢量)。

$$X = V - g_a H P \quad (3)$$

$g_a$ : 自适应码增益

$V$ : 话音信号(矢量)

$H$ : 合成滤波器脉冲响应卷积矩阵

$P$ : 自适应码矢量

在给定噪声码号后规定噪声码增益  $g_c$  的情况下，一般知道可假设式(2)中的  $g_c$  能取任意值，使式(2)最小的规定噪声码矢量号的处理(噪声源信息的矢量化处理)可置换为规定使下式(4)的分数式最大的噪声码矢量号。

$$\frac{(X' H c)^2}{\|H d\|^2} \quad (4)$$

即，在自适应码号和自适应码增益先前已有或暂时规定的情况下，噪声源信息矢量化处理成为规定使失真计算单元 16 算出的式(4)分数式最大的候补噪声码矢量号的处理。

初期的 CELP 型编码器 / 解码器中，将存储器中存储种类与所分配比特数对应的随机数序列所得的数据用作噪声码本。然而，存在以下课题：需要非常大的存储容量，同时对各候补噪声码矢量计算式(4)的失真用的运算处理量庞大。

作为解决此课题的一种方法，可举出《采用 10 毫秒话音帧的 8kb/s ACELP 话音编码：候选 CCITT 标准》(“8KBIT/S ACELP CODING OF SPEECH WITH 10 MS SPEECH-FRAME: A CANDIDATE FOR CCITT STANDARDIZATION”) (R. Salami, C. Laflamme 和 J-P. Adoul 著，刊登于 ICASSP' 94, pp. II-97 ~ II-100, 1994) 等中所记载那样，采用以代数方式产生音源矢量的代数音源矢量生成单元的 CELP 型话音信号编码器 / 解码器。

然而，噪声码本采用上述代数音源生成单元的 CELP 型话音信号编码器 / 解码器中，常以少量脉冲近似表现由式(3)求得的噪声源信息(规定噪声码号的

目标矢量)，因而在谋求改善话音质量方面存在局限。实际察看式(3)中噪声源信息  $X$  的要素，则几乎没有仅用少量脉冲构成该要素的情况。由此可说明存在局限。

### 本发明揭示

本发明的目的在于提供一种新的音源矢量生成装置，该装置能产生与实际分析话音信号时所得音源矢量的形状和统计相似性高的音源矢量。

本发明的又一目的在于提供一种 CELP 话音信号编码器 / 解码器、话音信号通信系统和话音信号记录系统，它们通过将上述音源矢量生成装置用作噪声码本，可获得比代数音源生成单元用作噪声码本时品质高的合成话音。

本发明第 1 形态是一种音源矢量生成装置，其特征在于，包括：具有  $N$  个 ( $N \geq 1$ ) 生成在矢量轴上某一要素建立带极性单元脉冲的脉冲矢量的通道的脉冲矢量生成单元；具有存储所述  $N$  中每一个通道  $M$  ( $M \geq 1$ ) 种扩散模式的功能，同时具有从所存  $M$  种扩散模式选择一种扩散模式的功能的扩散模式存储选择单元；具有每一通道进行所述脉冲矢量生成单元所输出脉冲矢量与所述扩散模式存储选择单元所选择扩散模式的卷积运算，并产生  $N$  个扩散矢量的功能的脉冲矢量扩散单元；具有将所述脉冲矢量扩散单元产生的  $N$  个扩散矢量相加后，产生音源矢量的功能的扩散矢量加法器。使所述脉冲矢量生成单元具有以代数方式产生  $N$  个脉冲矢量 ( $N \geq 1$ ) 的功能，再加上所述扩散模式存储选择单元预先存储通过预先学习实际话音矢量的形状(特性)获得的扩散模式，因而可产生形状比以往的代数音源生成单元更好地接近实际音源矢量形状的音源矢量。

本发明第 2 形态是一种 CELP 话音信号编码器 / 解码器，其特征在于，噪声码本中采用所述音源矢量生成装置。与以往噪声码本中采用代数音源生成单元的话音信号编码器 / 解码器相比，能产生更接近实际形状的音源矢量，因此，能取得可输出质量更高的合成话音的话音信号编码器 / 解码器、话音信号通信系统和话音信号记录系统。

### 附图概述

图 1 为以往的 CELP 型话音信号编码器的功能框图。

图 2 为以往的 CELP 型话音信号解码器的功能框图。

图 3 为本发明第 1 实施形态有关音源矢量生成装置的功能框图。

图 4 为本发明第 2 实施形态有关 CELP 型语音信号编码器的功能框图。  
图 5 为本发明第 2 实施形态有关 CELP 型语音信号解码器的功能框图。  
图 6 为本发明第 3 实施形态有关 CELP 型语音信号编码器的功能框图。  
图 7 为本发明第 4 实施形态有关 CELP 型语音信号编码器的功能框图。  
图 8 为本发明第 5 实施形态有关 CELP 型语音信号编码器的功能框图。  
图 9 为第 5 实施形态中矢量量化功能的框图。  
图 10 为第 5 实施形态中提取目标的算法的说明图。  
图 11 为第 5 实施形态中预测量化的功能框图。  
图 12 为第 6 实施形态中预测量化的功能框图。  
图 13 为第 7 实施形态中 CELP 型语音信号编码器的功能框图。  
图 14 为第 7 实施形态中失真计算单元的功能框图。

### 实施本发明的最佳形态

下面利用附图说明本发明的实施形态。

#### (第 1 实施形态)

图 3 中示出本发明实施形态有关音源矢量生成装置的功能框图。此音源矢量生成装置包括：具有多个通道的脉冲矢量生成单元 101；具有扩散模式存储单元和开关的扩散模式存储选择单元 102；扩散脉冲矢量的脉冲矢量扩散单元 103；将扩散的多个通道脉冲矢量相加的扩散矢量加法器 104。

脉冲矢量生成单元 101 具有 N 个通道(本实施形态中对  $N=3$  的情况进行说明)，这些通道生成在矢量轴上某一要素配置带极性单元脉冲的矢量(下文称为脉冲矢量)。

扩散模式存储选择单元 102 具有存储单元  $M1 \sim M3$  和开关  $SW1 \sim SW3$ ，前者对每一通道存储 M 种扩散模式(本实施形态中对  $M=2$  的情况进行说明)，后者从各存储单元  $M1 \sim M3$  分别选择 M 种扩散模式中的一种扩散模式。

脉冲矢量扩散单元 103 对每一通道进行脉冲矢量生成单元 101 所输出脉冲矢量与扩散模式存储选择单元 102 所输出扩散模式的卷积运算，并产生 N 个扩散矢量。

扩散矢量加法器 104 将脉冲矢量扩散单元 103 生成的 N 个扩散矢量相加后，生成音源矢量 105。

本实施形态中，对脉冲矢量生成单元 101 按照下列表 1 所记载规则以代数

方式产生 N 个脉冲矢量 (N=3) 的情况进行说明。

表 1

通道号	极 性	候补脉冲位置
CH1	$\pm 1$	P1(0, 10, 20, 30, ..., 60, 70)
CH2	$\pm 1$	P2 $\left[ \begin{array}{l} 2, 12, 22, 32, \dots, 62, 72 \\ 6, 16, 26, 36, \dots, 66, 76 \end{array} \right]$
CH3	$\pm 1$	P3 $\left[ \begin{array}{l} 4, 14, 24, 34, \dots, 64, 74 \\ 8, 18, 28, 38, \dots, 68, 78 \end{array} \right]$

说明上文所述那样构成的音源矢量生成装置的运作。扩散模式存储选择单元 102 从每通道分别存储 2 种的扩散模式中选择 1 种, 并输出到脉冲矢量扩散单元 103。但, 对应于选择的扩散模式组合(组合总数  $M^N=8$  种), 专门分配号码。

接着, 脉冲矢量生成单元 101 按照表 1 记载的规则, 以代数方式生成通道数量份额的脉冲矢量(本实施形态中为 3)。

脉冲矢量扩散单元 103 将扩散模式存储选择单元 102 选择的扩散模式和脉冲矢量生成单元 101 生成的脉冲用式(5)作卷积运算, 对每一通道生成扩散矢量。

$$ci(n) = \sum_{k=0}^{L-1} wij(n-k) di(k) \quad (5)$$

其中, n: 0 ~ L-1

L: 扩散矢量长度

i: 通道号

j: 扩散模式号 ( j = 1 ~ M )

ci: 通道 i 的扩散矢量

wij : 通道 i 的第 j 种扩散模式

Wij(m) 的矢量长度为  $2L-1$  (m:  $-(L-1) \sim L-1$ ), 但,  $2L-1$  个要素中能规定值的是  $L$  要素, 其它要素为零

di: 通道 i 的脉冲矢量

$di = \pm \delta ( n-pi )$ , n = 0 ~ L-1

$p_i$  : 通道  $i$  的候补脉冲矢量

扩散矢量加法器 104 利用公式(6)将脉冲矢量扩散单元 103 生成的 3 个扩散矢量相加后, 产生音源矢量 105。

$$c(n) = \sum_{i=1}^N c_i(n) \quad (6)$$

$c$ : 音源矢量

$c_i$ : 扩散矢量

$i$ : 通道号(  $i = 1 \sim N$  )

$n$ : 矢量要素号 (  $n=0 \sim L-1$  , 其中  $L$  为音源矢量长度 )

这样构成的音源矢量生成装置, 通过使扩散模式存储选择单元 102 所选择扩散模式的组合方法以及脉冲矢量生成单元 101 所生成脉冲矢量中的脉冲位置和极性带有变化, 可产生多样的音源矢量。

于是, 以上那样构成的音源矢量生成装置, 对扩散模式存储选择单元 102 所选择扩散模式的组合方法和脉冲矢量生成单元 101 所生成脉冲矢量形状(脉冲位置和脉冲极性)的组合方法等 2 种信息, 可预先分配分别一一对应的号码。扩散模式存储选择单元 102 中, 还可根据实际音源信息进行预先学习, 并预先存储该学习结果所得的扩散模式。

若话音信号编码器 / 解码器的音源信息生成单元中采用上述音源矢量生成装置, 则通过传送扩散模式存储选择单元所选择扩散模式的组合号和脉冲矢量生成单元所生成脉冲矢量的组合号(能规定脉冲位置和脉冲极性)等 2 种号码。能实现噪声源信息的传送。

又, 采用以上那样构成的音源矢量生成单元时, 与采用以代数方式生成的脉冲音源时相比, 可产生形状(特性)和实际音源信息相似的音源矢量。

本实施形态中, 对扩散模式存储选择单元 102 存储每一通道 2 种扩散模式的情况进行说明, 但在对各通道分配 2 种以外的扩散模式时, 也能取得同样的作用和效果。

本实施形态中对脉冲矢量生成单元 101 由 3 通道组成而且以表 1 所记载脉冲生成规则为基础的情况进行说明, 但在信道数不同时, 以及脉冲生成规则采用表 1 所记载以外的规则时, 也能取得同样的作用和效果。

此外, 组成具有上述音源矢量生成装置或话音信号编码器 / 解码器的话音信号通信系统或话音信号记录系统, 可取得上述音源矢量生成装置所具有的

作用和效果。

### (第 2 实施形态)

图 4 中示出本实施形态有关 CELP 型话音信号编码器的功能框图，图 5 中示出本实施形态有关 CELP 型话音信号解码器的功能框图。

有关本实施形态的 CELP 型话音信号编码器，在上述图 1 CELP 型话音信号编码器的噪声码本中应用第 1 实施形态所说明的音源矢量生成装置。与本实施形态有关的 CELP 型话音信号解码器，在上述图 2 CELP 话音信号解码器的噪声码本中，应用上述第 1 实施形态的音源矢量生成装置。因此，除噪声源信息矢量量化处理以外的处理，均与上述图 1、图 2 的装置相同。本实施形态中，以噪声源信息矢量量化处理为中心说明话音信号编码器和话音信号解码器。而且，与第 1 实施形态相同，也设通道数  $N=3$ ，一个通道的扩散模式数  $M=2$ ，脉冲矢量的生成依据图 1。

图 4 话音信号编码器中的噪声源矢量量化处理是规定使式(4)基准值最大的 2 种号码(扩散模式组合号、脉冲位置和脉冲极性组合号)的处理。

将图 3 音源矢量生成装置用作噪声码本时，用闭环规定扩散模式组合号(8 种)和脉冲矢量组合号(考虑极性时为 16384 种)。

因此，扩散模式存储选择单元 215 首先从本身存储的 2 种扩散模式中选择一种扩散模式，并输出到脉冲矢量扩散单元 217。然后，脉冲矢量生成单元 216 按照图 1 的规则，以代数方法产生通道数份额的脉冲矢量(本实施形态中为 3 个)，并输出到脉冲矢量扩散单元 217。

脉冲矢量扩散单元 217 将扩散模式存储选择单元 215 选择的扩散模式和脉冲矢量生成单元 216 产生的脉冲矢量用式(5)的卷积运算对每一通道产生扩散矢量。

扩散矢量加法器 218 将脉冲矢量扩散单元 217 获得的扩散矢量相加后，生成音源矢量(成为候补噪声码矢量)。

然后，失真计算单元 206 算出采用扩散矢量加法器 218 所得候补噪声码矢量的式(4)的值。对按表 1 的规则产生的脉冲矢量的全部组合，进行上述式(4)的值的运算，并将其中式(4)的值最大时的扩散模式组合号、脉冲矢量组合号(脉冲位置及其极性的组合)，以及当时的最大值输出到码号规定单元 213。

接着，扩散模式存储选择单元 215 从所得存储扩散模式选择与先前刚选择

的组合不同的扩散模式。然后,就新选择并更改的扩散模式的组合,与上文所述相同,对按表1的规则在脉冲矢量生成单元216产生的全部脉冲矢量组合算出式(4)的值。再次将其中式(4)最大时的扩散模式组合号、脉冲矢量组合号和最大值再次输出到码号规定单元213。

对能从扩散模式存储选择单元215所存扩散模式选择的全部组合(本实施形态说明中组合总数为8)反复进行上述处理。

码号规定单元213比较失真计算单元206算出的全部8个矢量最大值,选择其中最大的,规定产生该最大值时的2种组合号(扩散模式组合号、脉冲矢量组合号),并作为噪声码号输出到编码输出单元214。

另一方面,图5的话音信号解码器中,编码输入单元301接收话音信号编码器(图4)送来的编码,将接收到的编码分解成对应的线性预测号码、自适应码号、噪声码号(由扩散模式组合号、脉冲矢量组合号等2种组成)和加权码号,并将分解所得的编码分别输出到线性预测系数解码单元302、自适应码本303、噪声码本304和加权码本305。

噪声码号中,扩散模式组合号输出到扩散模式存储选择单元311、脉冲矢量组合号输出到脉冲矢量生成单元312。

然后,线性预测系数解码单元302将线性预测码号解码,取得合成滤波器系数,并输出到合成滤波器309。在自适应码本303,从与自适应码号对应的位置读出自适应码矢量。

噪声码本304中,扩散模式存储选择单元311对每一通道读出与扩散脉冲组合号对应的扩散模式并输出到脉冲矢量扩散单元313;脉冲矢量生成单元312产生通道数份额的与脉冲矢量组合号对应的脉冲矢量并输出到脉冲矢量扩散单元313;脉冲矢量扩散单元313将从扩散模式存储选择单元311接收到的扩散模式和从脉冲矢量生成单元312接收到的脉冲矢量用式(5)的卷积运算产生扩散矢量,并输出到扩散矢量加法器314。扩散矢量加法器314将脉冲矢量扩散单元313产生的各通道的扩散矢量相加后,产生噪声码矢量。

从加权码本305读出与加权码号对应的自适应码增益和噪声码增益,并在自适应码矢量加权单元306将自适应码矢量乘以自适应码增益,同样,在噪声码加权单元307将噪声码矢量乘以噪声码增益后,送到加法器308。

加法器308将已乘上增益的上述2个码矢量相加,生成驱动音源矢量,并将生成的驱动音源矢量输出到自适应码本303,以便更新缓存器,还输出到合

成滤波器 309，以便驱动合成滤波器。

合成滤波器 309 用加法器 308 所得的驱动音源矢量驱动后，再生合成话音 310。又，自适应码本 303 用从加法器 308 接收的驱动音源矢量更新缓存器。

但，图 4 和图 5 中的扩散模式存储选择单元取为将式(2)内的 C 中代入式(6)记载的音源矢量所得式(7)的失真计算基准式当作代价函数，并预先学习，使该代价函数的值较小后，将学习所得的扩散模式按每一通道加以存储。

通过上述操作，可生成形状与实际噪声源信息(式(4)中矢量 X)的形状相似的音源矢量，因而与噪声码本中采用代数音源矢量生成单元的 CELP 话音信号编码器 / 解码器相比，能取得质量高的合成话音。

$$\begin{aligned}
 E_c &= \left\| X - gcH \sum_{i=1}^N C_i \right\|^2 \\
 &= \sum_{n=0}^{L-1} (X(n) - gcH \sum_{i=1}^N C_i(n))^2 \\
 &= \sum_{n=0}^{L-1} (X(n) - gcH \sum_{i=1}^N \sum_{k=0}^{L-1} W_{ij}(n-k) d_i(k))^2 \quad (7)
 \end{aligned}$$

X: 规定噪声码号用的目标矢量

gc: 噪声码增益

H: 合成滤波器脉冲响应卷积矩阵

C: 噪声码矢量

i: 通道号 (i = 1 ~ N)

j: 扩散模式号 (i = 1 ~ M)

c<sub>i</sub>: 通道 i 的扩散矢量

W<sub>ij</sub>: 通道 i 的第 j 种扩散模式

d<sub>i</sub>: 通道 i 的脉冲矢量

L: 音源矢量长度 (n = 0 ~ L-1)

本实施形态中对每一通道预先存储 M 个进行预先学习，使式(7)的代价函数值较小后获得的扩散模式的情况进行说明，但实际上 M 个扩散模式不必全部通过学习取得，如果做成每一通道至少预先存储一种通过学习取得的扩散模式，这种情况下也能取得提高合成话音质量的作用和效果。

本实施形态中说明的情况是，根据扩散模式存储选择单元所存储扩散模式的全部组合和脉冲矢量生成单元 6 所生成候补脉冲矢量位置的全部组合，用闭

环规定使式(4)中基准值最大的组合号。然而，做成根据规定噪声码本号前求得的参数(自适应码矢量的理想增益等)，进行预选，或用开环进行检索等，也能取得同样的作用和效果。

此外，通过构成具有上述话音信号编码器 / 解码器的话音信号通信系统或话音信号记录系统，能取得第 1 实施形态中记载的音源矢量生成装置所具有的作用和效果。

### (第 3 实施形态)

图 6 中示出本实施形态有关 CELP 型话音信号编码器的功能框图。本实施形态在噪声码本内采用上述第 1 实施形态音源矢量生成装置的 CELP 话音编码器中，用检索噪声码本前求得的理想自适应码增益值，进行扩散模式存储选择单元所存扩散模式的预选。除噪声码本外围部分以外，都与图 4 的 CELP 型话音信号编码器相同。因此，本实施形态说明图 6 CELP 型话音信号编码器中的噪声源信息矢量量化处理。

此 CELP 型话音信号编码器具有自适应码本 407、自适应码增益加权单元 409、由实施形态 1 所说明音源矢量生成装置构成的噪声码本 408、噪声码增益加权单元 410、合成滤波器 405、失真计算单元 406、码号规定单元 413、扩散模式存储选择单元 415、脉冲矢量生成单元 416、脉冲矢量扩散单元 417、扩散矢量加法器 418 和自适应增益判定单元 419。

但，本实施形态中，上述扩散模式存储选择单元 415 存储的 M 种扩散模式 ( $M \geq 2$ ) 中至少一种是预先学习，使噪声源信息矢量量化时产生的量化失真较小，并由该学习结果得到的扩散模式。

为了说明简便，本实施形态中设脉冲矢量生成单元的通道数 N 为 3，扩散模式存储选择单元所存储每一通道的扩散脉冲种类数 M 为 2，而且取下列情况进行说明：M 种扩散模式 ( $M=2$ ) 的一种是由上述学习获得的扩散模式，另一种是由随机数矢量生成装置生成的随机矢量串(下文称为随机模式)。附带说一下，上述由学习获得的扩散模式，如图 3 中的 W11 那样，显然为长度比较短的脉冲状扩散模式。

图 6 的 CELP 型话音信号编码器中，在噪声源信息矢量量化前进行规定自适应码本号的处理。因此，在进行噪声源信息矢量量化处理的时刻，可参照自适应码本的矢量号(自适应码号)和理想自适应码增益(暂时确定)。本实施形态

中，使用其中的理想自适应码增益值进行扩散脉冲的预选。

具体而言，首先，自适应码本检索完毕后，立即将码号规定单元 413 所保持的自适应码增益理想值输出到失真计算单元 406。失真计算单元 406 将从码号规定单元 413 接收到的自适应码增益输出到自适应增益判定单元 419。

自适应增益判定单元 419 对从失真计算单元 409 接收到的理想自适应增益值和预先设定的门限值的大小进行比较。接着，自适应增益判定单元 419 根据上述大小比较的结果，将预先用的控制信号送到扩散模式存储选择单元 415。控制信号的内容在上述大小比较中自适应码增益大时，指示选择进行预先学习、使噪声源信息矢量量化时所产生量化失真较小后得到的扩散模式，并在上述大小比较中自适应码增益不大时，指示预选与学习结果所得扩散模式不同的扩散模式。

结果，在扩散模式存储选择单元 415 中，可适应自适应增益的大小，预选各通道存储的  $M$  种扩散模式 ( $M=2$ )，从而能大量减少扩散模式组合数。结果，不需要对扩散模式的全部组合号计算失真，能以少量的运算高效进行噪声源信息的矢量量化处理。

再者，噪声码矢量的形状在自适应增益值大时(有声性强时)为脉冲状，自适应增益值小时(有声性弱时)为随机状。因此，对话音信号的有声区域和无声区域，可分别利用形状适合的噪声码矢量，所以能提高合成语音的质量。

为了说明简便，本实施形态限定于脉冲矢量生成单元的通道数  $N$  为 3，扩散模式存储选择单元所存每一通道扩散脉冲的种类数  $M$  为 2 的情况下进行说明。然而，脉冲矢量生成单元的通道数、扩散模式存储选择单元中每一通道的扩散模式数与上述说明不同时，也能取得同样的效果和作用。

为了说明简便，本实施形态对每一通道所存  $M$  种扩散模式中 ( $M=2$ )，一种为由上述学习获得的扩散模式，另一种为随机模式的情况进行说明。然而，如果每一通道至少预先存储一种由学习取得的扩散模式，即便不是上述那样的情况，也能期望取得同样的效果和作用。

本实施形态对具有将自适应码增益大小信息作为预选扩散模式用的手段的情况进行说明，但如果兼用自适应增益大小信息以外的表示语音信号短时间特征的特征参数，可期望获得更进一步的效果和作用。

此外，通过构成具有上述语音信号编码器的语音信号通信系统或语音信号记录系统，可获得实施形态 1 中记载的音源矢量生成装置所具有的作用和效果。

再者，本实施形态说明了利用在噪声源信息量化的时刻可参照的当前处理帧理想自适应音源增益，预选扩散模式的方法，但不用当前帧理想自适应音源增益，而代之以利用在紧接前一帧求出的解码自适应音源增益时，可取同样的结构，这时也能取得相同的效果。

#### (第4实施形态)

图7为本实施形态有关 CELP 型话音信号编码器的功能框图。本实施形态在噪声码本内采用第1实施形态音源矢量生成装置的 CELP 型话音信号编码器中，使用在噪声源信息矢量量化的时刻可利用的信息，预选扩散模式存储选择单元所存储的多个扩散模式。作为此预选的基准，其特征在于，使用规定自适应码本号时所产生编码失真(用 S/N 比表示)的大小。

除噪声码本外围部分以外，均与图4 CELP 型话音信号编码器相同。因此，本实施形态详细说明噪声源信息的矢量量化处理。

如图7所示，本实施形态的 CELP 型话音信号编码器具有自适应码本 507、自适应码增益加权单元 509、由第1实施形态中说明的音源矢量生成装置构成的噪声码本 508、噪声码增益加权单元 510、合成滤波器 505、失真计算单元 506、码号规定单元 513、扩散模式存储选择单元 515、脉冲矢量生成单元 516、脉冲矢量扩散单元 517、扩散矢量加法器 518 和失真功率判定单元 519。

但，本实施形态中，取上述扩散模式存储选择单元 515 所存 M 种扩散模式中 ( $M \geq 2$ )，至少一种为随机模式。

为了说明简便，本实施形态中，取脉冲矢量生成单元的通道数 N 为 3，扩散模式存储选择单元所存每一通道扩散模式的种类数 M 为 2，而且假设 M 种扩散模式中 ( $M=2$ ) 一种为随机模式，另一种为预先学习，使噪声源信息矢量量化所产生的量化失真较小后，由该学习结果所得的扩散模式。

图7的 CELP 型话音信号编码器中，在噪声源信息矢量量化处理前进行规定自适应码本号的处理。因此，在进行噪声源号矢量量化处理的时刻，可参照自适应码本的矢量号(自适应码号)、理想自适应码增益(暂时确定)和自适应码本检索用的目标矢量。本实施形态中使用可根据上述三种信息算出的自适应码本编码失真(用 S/N 比表示)，进行扩散模式的预选。

具体而言，自适应码本检索完毕后，立即将码号规定单元 513 所保持的自适应码号和自适应码增益(理想增益)的值输出到失真计算单元 506。失真计算

单元 506 利用从码号规定单元 513 接收到的自适应码号和自适应码增益, 以及自适应码本检索用的目标矢量, 算出由规定自适应码本号而产生的编码失真 (S/N 比)。将算出的 S/N 比输出到失真功率判定单元 519。

失真功率判定单元 519 首先进行从失真计算单元 506 接收到的 S/N 比与预先设定的门限值的大小比较。接着, 失真功率判定单元 519 根据上述大小比较的结果, 将预选用的控制信号送到扩散模式存储选择单元 515。控制信号的内容在上述大小比较中 S/N 比大时, 指示选择预先学习, 使噪声码本检索用目标矢量进行编码所产生的编码失真较小后, 其结果所得的扩散模式, 并在上述大小比较中 S/N 比小时, 指示选择随机模式的扩散模式。

结果, 扩散模式存储选择单元 515 中, 从各通道存储的 M 种扩散模式 (M=2) 仅预选一种, 可大量减少扩散模式的组合。因此, 不需要对扩散模式的全部组合号计算失真, 能用少量的运算高效规定噪声码号。再者, 噪声码矢量的形状在 S/N 比大时为脉冲状, S/N 比小时为随机状。因此, 能根据话音信号的短时间特征, 使噪声码矢量的形状变化, 从而可提高合成话音的质量。

为了说明简便, 本实施形态限定于脉冲矢量生成单元的通道数 N 为 3, 扩散模式存储选择单元所存每一通道扩散脉冲的种类数 M 为 2 的情况下进行说明。然而, 脉冲矢量生成单元的通道数、每一通道扩散模式的种类与上述说明不同时, 也能取得同样的效果和作用。

为了说明简便, 本实施形态又对每一通道存储的 M 种扩散模式中 (M=2), 一种为由上述学习取得的扩散模式, 另一种为随机模式的情况进行说明。然而, 如果做成每一通道至少预先存储一种随机模式的扩散模式, 即便不是上述那样的情况也可期望取得同样的效果和作用。

本实施形态中虽然仅用由规定自适应码号而产生的编码失真 (用 S/N 比表示) 的大小信息作为预选扩散模式用的手段, 但如果兼用进一步正确表示话音信号短时间特征的信息, 可期望具有更进一步的效果和作用。

此外, 通过构成具有上述话音信号编码器的话音信号通信系统或话音信号记录系统, 或获得第 1 实施形态中记载的音源矢量生成装置所具有的作用和效果。

## 第 5 实施形态

图 8 中示出本发明第 5 实施形态有关 CELP 型话音信号编码器的功能框图。

此 CELP 型话音信号编码器中，在 LPC 分析单元 600 通过对输入话音数据 601 进行自相关分析和 LPC 分析，获得 LPC 系数。又在对所得 LPC 系数进行编码而取得 LPC 码的同时，将所得 LPC 码解码，取得解码 LPC 系数。

接着，在音源生成单元 602，取出自适应码本 603 和噪声码本 604 所存的音源取样(分别称为自适应码矢量(或自适应音源)和噪声码矢量(或噪声源))，并分别送到 LPC 合成单元 605。

在 LPC 合成单元 605 中，对音源生成单元 602 所获得的 2 个音源，利用 LPC 分析单元 600 所获得的解码 LPC 系数进行滤波，从而获得 2 个合成话音。

在比较单元 606 中，分析 LPC 合成单元 605 所得 2 个合成话音和输入话音 601 的关系，求 2 个合成话音的最佳值(最佳增益)，并将用该最佳增益调整功率后的各合成话音相加，取得总合成话音后，计算该总合成话音与输入话音的距离。

又，对自适应码本 603 和噪声码本 604 的全部音源的取样，计算由驱动音源生成单元 602 和 LPC 合成单元 605 所得的多个合成话音与输入话音 601 的距离，求该结果所得距离中最小时的音源取样索引号。

将所得最佳增益、音源取样索引号以及该索引号对应的 2 个音源送到参数编码单元 607。在参数编码单元 607 通过进行最佳增益编码取得增益码后，将 LPC 码、音源取样索引号汇合在一起送到传输线路 608。

根据与增益码和索引号对应的 2 个音源，生成实际音源信号，将该信号存入自适应码本 603，同时废弃旧的音源取样。

再者，LPC 合成单元 605 中，通常兼用听觉加权滤波器，该滤波器采用线性预测系数、高频强化滤波器和长期预测系数(通过对输入话音进行长期预测分析取得)。一般用将分析区间进一步细分的区间(称为子帧)进行对自适应码本和噪声码本的音源检索。

下面，本实施形态对 LPC 分析单元 600 中的 LPC 系数矢量量化进行详细说明。

图 9 示出实现在 LPC 分析单元 600 执行的矢量量化算法用的功能框图。图 9 中所示矢量量化框包括目标提取单元 702、量化单元 703、失真计算单元 704、比较单元 705、解码矢量存储单元 707 和矢量平滑单元 708。

在目标提取单元 702 中，根据输入矢量 701，算出量化目标。现详细说明提取目标的方法。

本实施形态中的输入矢量由 2 种矢量构成：分析编码对象帧所得的参数矢量；从一个未来帧进行同样分析取得的参数矢量。目标提取单元 702 利用上述输入矢量和解码矢量存储单元 707 所存先前帧解码矢量，算出量化目标。式(8)示出运算方法的例子。

$$X(i) = \{ S_t(i) + p(d(i) + S_{t+1}(i) / 2) \} / (1+p) \quad (8)$$

$X(i)$ ：目标矢量

$i$ ：矢量要素号

$S_t(i)$ 、 $S_{t+1}(i)$ ：输入矢量

$t$ ：时间(帧号)

$p$ ：加权系数(固定)

$d(i)$ ：先前帧解码矢量

下面示出上述目标提取方法的思路。典型的矢量量化中将当前帧的参数矢量  $S_t(i)$  作为目标  $X(i)$ ，并用式(9)进行拟合。

$$E_n = \sum_{i=0}^I (X(i) - C_n(i))^2 \quad (9)$$

$E_n$ ：与  $n$  号码矢量的距离

$X(i)$ ：量化目标

$C_n(i)$ ：码矢量

$n$ ：码矢量号

$i$ ：矢量维数

$I$ ：矢量长度

于是，至此为止的矢量量化中，编码失真仍然与音质的劣化相联系。在即使采取预测矢量量化等对策也不能避免某种程度的编码失真的超低比特率编码中，这成为大问题。

因此，本实施形态中，作为听觉上难以发现差错的方向，着眼于前后解码矢量的中点，在该处导出解码矢量，由此实现听觉方面的改善。这是利用参数矢量内插特性良好时，难以听到时间连续特性在听觉上的劣化这种特性。下面参照示出矢量空间的图 10 说明这种情况。

首先，该前一帧的解码矢量为  $d(i)$ ，未来参数矢量为  $S_{t+1}(i)$  (实际上最好为未来解码矢量，但当前帧中不能编码，所以代用参数矢量)，则码矢量  $C_n(i)$ ：(1) 为码矢量  $C_n(i)$ ：(2) 更接近参数矢量  $S_t(i)$ ，但实际上  $C_n(i)$ ：(2)

接近在  $d(i)$  与  $S_{t+1}(i)$  的连线上, 因而比  $C_n(i)$ : (1) 不易听到劣化。于是, 利用这种特性, 如果将目标  $X(i)$  取为从  $S_t(i)$  以某种程度接近  $d(i)$  与  $S_{t+1}(i)$  的中点位置上的矢量, 则将解码矢量引导到听觉上失真小的方向。

本实施形态中, 通过导入下面式(10)的估算能实现这种目标的移动。

$$X(i) = \{ S_t(i) + p(d(i) + S_{t+1}(i)) / 2 \} / (1+p) \quad (10)$$

$X(i)$ : 量化目标矢量

$i$ : 矢量要素号

$S_t(i)$ 、 $S_{t+1}(i)$ : 输入矢量

$t$ : 时间(帧号)

$p$ : 加权系数(固定)

$d(i)$ : 先前帧解码矢量

式(10)的前半部分为一般矢量量化估算式, 后半部分为听觉加权分量。为了用上述估算式进行量化, 在各  $X(i)$  将估算式进行微分, 并设微分所得结果为 0, 则可得式(8)。

加权系数  $P$  为正的常数, 其值为 0 时和一般矢量量化相同, 无限大时目标完全位于中点。 $P$  非常大, 则目标大为偏离当前帧的参数矢量  $S_t(i)$ , 听觉清晰度下降。根据解码话音信号试听实验, 证实  $0.5 < p < 1.0$  时取得良好的性能。

在量化单元 703 对目标提取单元 702 所得的量化目标进行量化, 求矢量码, 同时求解码矢量, 并和矢量码一起送到失真计算单元 704。

本实施形态中, 采用预测矢量量化作为量化的方法。下面说明预测矢量量化。

图 11 中示出预测矢量量化的功能框图。预测矢量量化是一种利用过去编码和解码所得矢量(合成矢量)进行预测, 并对该预测误差进行矢量量化的算法。

预先生成存储多个预测误差矢量的中心取样(码矢量)的矢量码本 800。通常根据分析多个话音数据所得的多个矢量, 利用 LBG 算法(IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. COM - 28, NO. 1, pp84 - 95, JANUARY 1980), 生成该码本。

在预测单元 802 对量化目标的矢量 801 进行预测。预测利用状态存储单元 803 所存的过去合成矢量进行, 并将所得预测误差矢量送到距离计算单元 804。这里, 作为预测的形态, 举出预测次数为 1 次时利用固定系数进行的预测。下

面的式(11)中示出上此预测时的预测误差矢量计算式。

$$Y(i) = X(i) - \beta D(i) \quad (11)$$

$Y(i)$ : 预测误差矢量

$X(i)$ : 量化目标

$\beta$ : 预测系数(标量)

$D(i)$ : 前 1 帧的合成矢量

$i$ : 矢量维数

上式中预测系数 $\beta$ 的值一般为  $0 < \beta < 1$ 。

在距离计算单元 804 中, 计算预测单元 802 所得预测误差矢量与矢量码本 800 所存码矢量的距离。下面的式(12)示出该距离计算公式。

$$E_n = \sum_{i=0}^I (T(i) - C_n(i))^2 \quad (12)$$

$E_n$ : 与  $n$  号码矢量的距离

$T(i)$ : 预测误差矢量

$C_n(i)$ : 码矢量

$n$ : 码矢量号

$i$ : 矢量维数

$I$ : 矢量长度

在检索单元 805 中比较与各码矢量的距离, 将距离最小的码矢量的号码作为矢量码 806 加以输出。即, 控制矢量码本 800 和距离计算单元 804, 求矢量码本 800 所存全部码矢量中距离最小的码矢量的号码, 并将该矢量号作为矢量码 806。

进而, 根据最终矢量码, 利用从矢量码本 800 所得的码矢量和状态存储单元 803 所存的过去解码矢量, 进行矢量解码, 并用所得的合成矢量更新状态存储单元 803 的内容。因此, 进行下一次解码时可将此处解码的矢量用于预测。

利用下面的式(13)进行上述预测形态例(预测次数为 1 次、固定系数)的解码。

$$Z(i) = CN(i) + \beta D(i) \quad (13)$$

$Z(i)$ : 解码矢量(下一次编码时作为  $D(i)$  使用)

$N$ : 矢量编码

$CN(i)$ : 码矢量

$\beta$ : 预测系数(非向量)

$D(i)$ : 前 1 帧的合成矢量

$I$ : 矢量维数

另一方面, 在解码器中, 通过根据传送来的矢量码求码矢量, 进行解码。解码器中预先备有与编码器相同的矢量码本和状态存储单元, 利用与上述编码算法中的检索单元解码器功能相同的算法, 进行解码。以上是在量化单元 703 执行的矢量量化。

在失真计算单元 704 中根据量化单元 703 所得的解码矢量、输入矢量 701 和解码矢量存储单元 707 所存的先前帧解码矢量, 计算听觉加权编码失真。下面的式(14)示出计算式。

$$E_w = \sum (V(i) - S_t(i))^2 + p \{ V(i) - (d(i) + S_{t+1}(i)) / 2 \}^2 \quad (14)$$

$E_w$ : 加权编码失真

$S_t(i), S_{t+1}(i)$ : 输入矢量

$t$ : 时间(帧号)

$i$ : 矢量要素号

$V(i)$ : 解码矢量

$p$ : 加权系数(固定)

$d(i)$ : 先前帧解码矢量

在式(14)中, 加权系数  $p$  与目标提取单元 702 所用目标计算式的系数相同。将上述加权编码失真值、解码矢量和矢量码送到比较单元 705。

比较单元 705 将失真计算单元 704 送来的矢量码送到传输线路 608, 而且用失真计算单元 704 送来的解码矢量, 更新解码矢量存储单元 707 的内容。

根据上述实施形态, 在目标提取单元 702 将目标矢量修正为从  $S_t(i)$  以某种程度接近  $d(i)$  和  $S_{t+1}(i)$  的中点的位置的矢量, 因而可进行加权检索而不觉得听觉上劣化。

至此, 说明了本发明适应于便携电话等所用低比特率话音信号编码技术的情况, 但本发明不但是话音信号编码, 而且还能用于音乐编码器。图像编码器中内插性较好的参量矢量量化。

上述算法中 LPC 分析单元的 LPC 编码通常是变换为一般的 LSP(线谱对)等便于编码的参数矢量, 利用欧几里德距离和加权欧几里德距离进行矢量量化

(VQ)。

本实施形态中，目标提取单元 702 接受比较单元 705 的控制，将输入矢量 701 送到矢量平滑单元 708，目标提取单元 703 接收矢量平滑单元 708 中修改过的输入矢量，再进行目标的提取。

这时，在比较单元 705 比较失真计算单元 704 送来的加权编码失真值和比较单元内部准备的基准值。根据此比较结果，处理分为二种。

未达到基准值时，将失真计算单元 704 送来的矢量码送到传输线路 606，而且用失真计算单元 704 送来的解码矢量，更新解码矢量存储单元 707 的内容。通过用得到的解码矢量改写解码矢量存储单元 707 的内容，进行此更新。然后，过渡到下一帧参数编码处理。

反之，在基准值以上时，控制矢量平滑单元 708，对输入矢量加以修改，使目标提取单元 702、量化单元 703 和矢量计算单元 704 再次起作用，进行重新编码。

在比较单元 705 中未达到基准值前，反复进行编码处理。然而，有时会反复进行几次也不能变成未达到基准值，因而比较单元 705 内部具有计数器，计算判定为基准值以上的次数，达到一定次数以上时，中止反复编码，并进行未达到基准值时的处理和计数器清零。

矢量平滑单元 708 中，接收比较单元 705 的控制，根据由目标提取单元 702 得到的输入矢量和从解码矢量存储单元 707 得到的先前帧解码矢量，利用下面的式(15)修改作为输入矢量之一的当前帧参数矢量  $S_t(i)$ ，并将修改后的输入矢量送到目标提取单元 702。

$$S_t(i) \leftarrow (1-q) \cdot S_t(i) + q (d(i) + S_{t+1}(i)) / 2 \quad (15)$$

上述  $q$  为平滑系数，表示当前帧参数矢量接近先前帧解码矢量与未来帧参数矢量的中点的程度。根据编码实施，证实  $0.2 < q < 0.4$  且比较单元 705 的内部反复次数上限值为 5—8 次时，可获得良好的性能。

本实施形态虽然在量化单元 703 采用预测矢量量化，但借助上述平滑处理，失真计算单元 704 所得加权编码失真变小的可能性大。其原因在于利用平滑处理使量化目标更接近先前帧解码矢量。因此，利用反复进行比较单元 705 控制的编码，比较单元 705 的失真比较中未达到基准值的可能性提高。

解码器中，预先备有与编码器量化单元对应的解码单元，根据从传输线路送来的矢量码进行解码。

本实施形态也用于 CELP 型编码中出现的 LSP 参数量化(量化单元进行预测 VQ)进行话音信号的编码和解码实验。其结果,证实听觉上的音质当然能提高,而且也能使客观值(S/N 比)提高。这是因为利用具有矢量平滑的反复编码处理,达到即使频谱急剧变化时也能抑制预测 VQ 编码失真的效果。以往的预测 VQ 具有的缺点是:由于根据过去合成矢量进行预测,讲话开始的部分等频谱急剧变化部分的频谱失真反而变大。然而,应用本实施形态,则失真大时进行平滑处理,直到失真变小,因而目标虽然有些偏离实际的参数矢量,但编码失真变小,可取得话音信号解码时总体劣化变小的效果。因此,根据本实施形态,不仅听觉上改善音质,而且也能使客观值提高。

本实施形态中,可利用比较单元和矢量平滑单元的特征,在矢量量化失真大时,将其劣化的方向控制在听觉上比较不会觉察的方向上,而且在量化单元采用预测矢量量化时通过反复进行平滑处理+编码,直到编码失真变小也能使客观值提高。

至此,说明了本发明适应于便携电话等所用低比特率话音编码技术的情况,但本发明不仅是话音信号编码,而且也可用于音乐编码器和图像编码器中内插性较好的参数矢量量化。

#### (第 6 实施形态)

下面说明本发明第 6 实施形态有关的 CELP 型话音信号编码器。本实施形态除量化方法采用多级预测矢量量化的量化单元的量化算法外,其他的结构与上述第 5 实施形态相同。即,噪声码本采用上述第 1 实施形态的音源矢量生成装置。现详细说明量化单元的量化算法。

图 12 中示出量化单元的功能框图。多级矢量量化中,进行目标矢量量化后,以量化所得目标码字利用其码本进行解码,求编码后的矢量与原目标之差(称为编码失真矢量),进而将求得的编码失真矢量加以量化。

预先生成存放多个预测误差矢量中心取样(码矢量)的矢量码本 899、矢量码本 900。借助对多个学习用的预测误差矢量,应用与典型“多级矢量量化”码本生成方法相同的算法,生成这些码本。即,通常根据分析许多话音数据所得的多个矢量,利用 LBG 算法(I EEE TRANSACTIONS ON COMMUNICATIONS, VOL. COM - 28, NO. 1, pp84 - 95, JANUARY 1980)生成上述码本。但,矢量码本 899 的学习总体为许多量化目标的集合,矢量码本 900 的学习总体为对上述许多量

化目标用量化码本 899 进行编码时的编码失真矢量的集合。

首先，在预测单元 902 对量化目标矢量 901 进行预测。预测用状态存储单元 903 所存过去合成矢量进行，并将得到的预测误差矢量送到距离计算单元 904 和距离计算单元 905。

本实施形态中，作为预测形态，举出预测次数为 1 次时利用固定系数进行的预测。下面的式 (16) 示出用这种预测时的预测误差矢量运算式。

$$Y(i) = X(i) - \beta D(i) \quad (16)$$

$Y(i)$ : 预测误差矢量

$X(i)$ : 量化目标

$\beta$ : 预测系数(标量)

$D(i)$ : 前 1 帧的合成矢量

$i$ : 矢量维数

上式中，预测系数 $\beta$ 的值通常为  $0 < \beta < 1$ 。

在距离计算单元 904 中，计算预测单元 902 所得预测误差矢量与矢量码本 899 所存码矢量 A 的距离。下面的式 (17) 示出距离计算式。

$$E_n = \sum_{i=0}^I (X(i) - C_n(i))^2 \quad (17)$$

$E_n$ : 与 n 号码矢量 A 的距离

$X(i)$ : 预测误差矢量

$C_n(i)$ : 码矢量 A

$n$ : 码矢量 A 的号码

$i$ : 矢量维数

$I$ : 矢量长度

在检索单元 906 中，比较与各码矢量 A 的距离，将距离最小的码矢量 A 的号码作为码矢量 A 的编码。即，控制矢量码本 899 和距离计算单元 904 求矢量码本 899 所存全部码矢量中距离最小的码矢量 A 的号码，并将该号码作为码矢量 A 的编码。然后，将码矢量 A 的编码和参照该编码从矢量码本 899 取得的解码矢量 A 送到距离计算单元 905。又将码矢量 A 的编码送到传输线路、检索单元 907。

距离计算单元 905 根据预测误差矢量和从检索单元 906 取得的解码矢量 A，取得编码失真矢量，或者参照从检索单元 906 取得的码矢量 A 的编码，从幅度

存储单元 908 取得幅度，然后计算上述编码失真矢量与矢量码本 900 中存储的码矢量 B 乘以上述幅度所得结果的距离，并将该距离送到检索单元 907。下面的式(18)示出距离计算式。

$$Z(i) = Y(i) - CIN(i)$$

$$Em = \sum_{i=0}^I (Z(i) - aNC2m(i))^2 \quad (18)$$

Z(i): 解码失真矢量

Y(i): 预测误差矢量

C1N(i): 解码矢量 A

N: 码矢量 A 的编码

Em: 与 m 号码矢量 B 的距离

aN 与码矢量 A 的编码对应的幅度

C2m(i): 码矢量 B

m: 码矢量 B 的号码

i: 矢量维数

I: 矢量长度

在检索单元 907 中，比较与各码矢量 B 的距离，将距离最小的码矢量 B 的号码作为码矢量 B 的编码。即，控制矢量码本 900 和距离计算单元 905，求矢量码本 900 所存全部码矢量 B 中距离最小的码矢量 B 的号码，并将该号码作为码矢量 B 的编码。然后，将码矢量 A 和码矢量 B 的编码合在一起，作为矢量 909。

检索单元 907 还根据码矢量 A、B 的编码，用从矢量码本 899 和矢量码本 900 获得的解码矢量 A 和 B、从幅度存储单元 908 获得的幅度，以及状态存储单元 903 存储的过去解码矢量进行矢量的解码，并利用得到的合成矢量更新状态存储单元 903 的内容。(因此，进行下一次编码时，将此处解码的矢量用于预测。)利用下面的式(19)进行本实施形态预测(预测次数为 1 次。固定系数)中的解码。

$$Z(i) = CIN(i) + aN \cdot C2M(i) + \beta D(i) \quad (19)$$

Z(i): 解码矢量(下一次编码时作为 D(i)使用)

N: 码矢量 A 的编码

M: 码矢量 B 的编码

C1M: 解码矢量 A

C2M: 解码矢量 B

aN: 与码矢量 A 的编码对应的幅度

$\beta$ : 预测系数(非向量)

D(i): 前一帧的合成矢量

i: 矢量维数

预先设定幅度存储单元 908 存储的幅度, 下面示出此设定方法。对许多话音数据进行编码, 并对第 1 级码矢量的各编码求下面式(20)的总编码失真后, 进行学习, 使该失真最小, 从而设定幅度。

$$EN = \sum_{i=0}^I (Y_t(i) - C1N(i) - aN C2m_t(i))^2 \quad (20)$$

EN: 码矢量 A 的编码为 N 时的编码失真

N: 码矢量 A 的编码

t: 码矢量 A 的编码为 N 的时间

$Y_t(i)$ : 时间 t 的预测误差矢量

C1N(i): 解码矢量 A

aN: 与码矢量 A 的编码对应的幅度

$C2m_t(i)$ : 码矢量 B

$m_t$ : 码矢量 B 的号码

i: 矢量维数

I: 矢量长度

即, 编码后, 设定并修改上述式(20)的失真, 使在各幅度微分的值为 0, 由此, 进行幅度学习。然后, 反复进行上述编码+学习, 从而求出最佳幅度。

另一方面, 解码器中, 通过根据传递来的矢量码, 求码矢量, 进行解码。解码器具有和编码器相同的矢量码本(对应于码矢量 A、B)、幅度存储单元和状态存储单元, 用与上述编码算法中检索单元(对应于码矢量 B)的解码功能相同的算法进行解码。

因此, 本实施形态中, 利用幅度存储单元和距离计算单元的特征以较少的计算量使第 2 级的码矢量适应第 1 级, 从而能使幅度失真较小。

至此, 说明了本发明适应于便携电话等所用低比特率话音信号编码技术的情况, 但本发明不仅是话音信号编码, 而且还可用于音乐编码器和图像编码器等中内插性较好的参数矢量量化。

### (第7实施形态)

下面说明本发明第7实施形态有关的 CELP 型语音信号编码器。本发明形态是一种编码器的例子，该编码器可减少采用 ACELP 型噪声码本时码检索的运算量。

图13中示出本实施形态有关 CELP 型语音编码器的功能框图。此 CELP 型语音信号编码器中，滤波器系数分析单元1002对输入语音信号1001进行线性预测分析，取得合成滤波器系数，并将所得合成滤波器系数输出到滤波器系数量化单元1003。滤波器系数量化单元1003将输入的合成滤波器系数量化后，输出到合成滤波器1004。

合成滤波器1004是根据滤波器系数量化单元1003所供给的滤波器系数建立的，由激励信号1011驱动。该激励信号1011通过将自适应码本1005输出的自适应矢量1006乘以自适应增益1007所得结果与噪声码本1008输出的噪声矢量1009乘以噪声增益1010所得结果相加而取得。

这里，自适应码本1005是存储每一音调周期取出过去对合成滤波器的激励信号的多个自适应矢量的码本，噪声码本1007是存储多个噪声矢量的码本。噪声码本1007可采用上述第1实施形态的音源矢量生成装置。

失真计算单元1013算出作为激励信号1011所驱动合成滤波器1004的输出合成语音信号1012与输入语音信号1001之间的失真，并进行码检索处理。码检索处理是一种规定使失真计算单元1013所计算失真最小用的自适应矢量1006的号码和噪声矢量1009的号码，同时算出各输出矢量所乘自适应增益1007和噪声增益1010的最佳值的处理。

编码输出单元1014输出的是将分别与从滤波器系数量化单元1003得到的滤波器系数量化值，以及失真计算单元1013中选择的自适应矢量1006的号码和噪声矢量1009的号码相乘的自适应增益1007和噪声增益1009编码后所得的结果。将从编码输出单元1014输出的信息加以传输或存储。

失真计算单元1013中的码检索处理，通常首先对激励信号中的自适应码本分量进行检索，然后对激励信号中的噪声码本分量进行检索。

上述噪声分量的检索使用下面说明的正交检索。

正交检索中，规定使式(21)的检索基准值  $E_{ort} (=N_{ort} / D_{ort})$  最大的噪声矢量  $c$ 。

$$E_{ort} = \frac{N_{ort}}{D_{ort}} = \frac{[(P^t H^t H c) x - (x^t H p) H p] H c}{(c^t H^t H c) (p^t H^t H p) - (p^t H^t H c)^2} \quad (21)$$

Nort : Eort 的分子项

Dort : Eort 的分母项

p: 已规定的自适应矢量

H: 合成滤波器系数矩阵

H': H 的转置矩阵

X: 目标信号(输入话音信号与合成滤波器零输入响应差分所得的结果)

c: 噪声矢量

正交检索是对事先规定自适应矢量为候补的噪声矢量分别正交, 并从正交的多个噪声矢量规定 1 个失真最小的检索方法。这种检索方法与非正交检索相比, 其特征在于可提高规定噪声矢量的精度, 从而能提高合成语音信号的质量。

ACELP 方式中, 只用少数带极性的脉冲构成噪声矢量。利用这点, 将式(21)所示检索基准值的分子项(Nort)变换为下面的式(22), 由此, 可减少分子项的运算。

$$N_{ort} = \{ a_0 \psi(l_0) + a_1 \psi(l_1) + \dots + a_{n-1} \psi(l_{n-1}) \}^2 \quad (22)$$

a<sub>i</sub>: 第 i 个脉冲的极性 (+ 1 / -1)

l<sub>i</sub>: 第 i 个脉冲的位置

N: 脉冲个数

$\psi: \{ (p^t H^t H p) x - (x^t H p) H p \} H$

将式(22)中 $\psi$ 的值作为前处理预先计算, 并在阵列中展开, 则可将阵列 $\psi$ 中的(N-1)个要素带符号进行相加, 并对其结果取平方, 从而计算式(21)的分子项。

下面具体说明可对分母项减少运算量的失真计算单元 1013。

图 14 中示出失真计算单元 1013 的功能框图。本实施形态中的话音信号编码器, 其结构为在图 13 的结构中, 将自适应矢量 1006 和噪声矢量 1009 输入失真计算单元 1013。

在图 14 中, 作为对所输入噪声矢量计算失真时的前处理, 进行以下 3 种处理。

(1)算出第 1 矩阵(N): 计算在合成滤波器合成自适应矢量后所得矢量的功率  $(p^t H^t H p)$  和合成滤波器中滤波器系数的自相关矩阵  $(H^t H)$ , 并将上述功

率与上述自相关矩阵各要素相乘，从而算出矩阵  $N (= (p^t H^t H p) H^t H)$ 。

(2) 算出第 2 矩阵 (M)：将在合成滤波器合成自适应矢量后所得的矢量按反时针顺序合成，并对其结果所得的信号  $(p^t H^t H)$  取矢积后，算出矩阵 M。

(3) 生成第 3 矩阵 (L)：对 (1) 中算出的矩阵 N 和 (2) 中算出的矩阵 M 进行差分，生成矩阵 L。

又，式 (21) 的分母项 (Dort) 可展开为式 (23)。

$$\begin{aligned}
 \text{Dort} &= (c^t H^t H c) (p^t H^t H p) - (p^t H^t H c)^2 & (23) \\
 &= c^t N c - (r^t c)^2 \\
 &= c^t N c - (r^t c)^t (r^t c) \\
 &= c^t N c - (c^t r r^t c) \\
 &= c^t N c - (c^t M c) \\
 &= c^t (N - M) c \\
 &= c^t L c
 \end{aligned}$$

$N: (p^t H^t H p) H^t H \leftarrow$  上述前处理 (1)

$r: p^t H^t H \leftarrow$  上述前处理 (2)

$M: r r^t \leftarrow$  上述前处理 (2)

$L: N - M \leftarrow$  上述前处理 (3)

$c: \text{噪声矢量}$

由此，将计算式 (21) 检索基准值 (Eort) 时的分母项 (Dort) 的计算方法置换为式 (23)，可用较少的运算量规定噪声码本分量。

用上述前处理获得的矩阵 L 和噪声矢量 1009，进行分母项的计算。

这里，为了简便，对输入话音信号取样频率为 8000Hz，代数结构噪声码本检索的单位时间宽度 (帧时间) 为 10ms，噪声矢量用每 10ms 5 个单元脉冲 (+1 / -1) 的规则组合生成的情况，说明基于式 (23) 的分母项计算方法。

又设构成噪声矢量的 5 个单位脉冲由处于从表 2 所示第 0 组到第 4 组所规定位置分别选择 1 个位置的脉冲组成，候补噪声矢量可用下面式 (24) 记述。

$$\begin{aligned}
 C &= a_0 \delta (k - l_0) + a_1 \delta (k - l_1) + \dots + a_4 \delta (k - l_4) & (24) \\
 &(k = 0, 1, \dots, 79)
 \end{aligned}$$

$a_i: \text{第 } i \text{ 组所属脉冲的极性 (+1 / -1)}$

$l_i: \text{第 } i \text{ 组所属脉冲的位置}$

表 2

组号	符号	候补脉冲位置
0	± 1	0, 10, 20, 30, ..., 60, 70
1	± 1	2, 12, 22, 32, ..., 62, 72
2	± 1	4, 14, 24, 34, ..., 64, 74
3	± 1	8, 18, 28, 38, ..., 68, 78
4	± 1	

这时，可用下面的式(25)求式(23)所示的分母项(Dort)。

$$Dort = \sum_{i=0}^4 \sum_{j=0}^4 a_i a_j L(l_i, l_j) \quad (25)$$

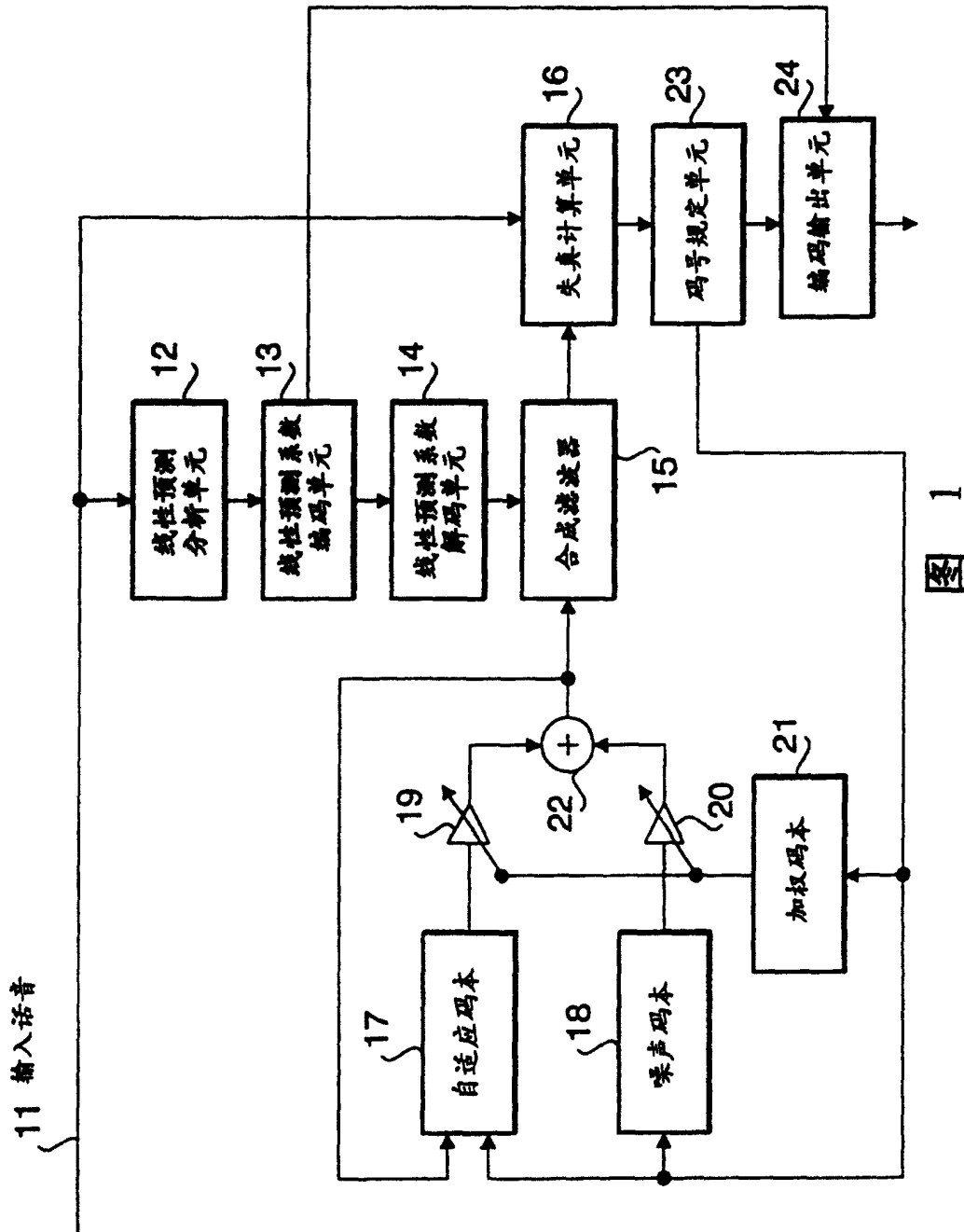
$a_i$ : 第  $i$  组所属脉冲的极性

$l_i$ : 第  $i$  组所属脉冲的位置

$L(l_i, l_j)$ : 矩阵  $L$  中  $l_i$  行、 $l_j$  列的要素

根据以上说明，证明采用 ACELP 型噪声码本时，可用式(22)计算式(21)的码检索基准值的分子项(Nort)，可用式(25)计算其分母项(Dort)。因此，采用 ACELP 型噪声码本时，不是原样计算式(21)的基准值，而是用(22)和式(25)分别计算其分子项和分母项，从而可大幅度削减码检索运算量。

以上说明的本实施形态，说明了不带有预选的噪声码本检索。然而，预选使式(22)的值大的噪声矢量，并对利用预选收敛为多个候补的噪声矢量计算式(21)，选择使该值最大的噪声矢量，这种情况下应用本发明，也能取得相同的效果。



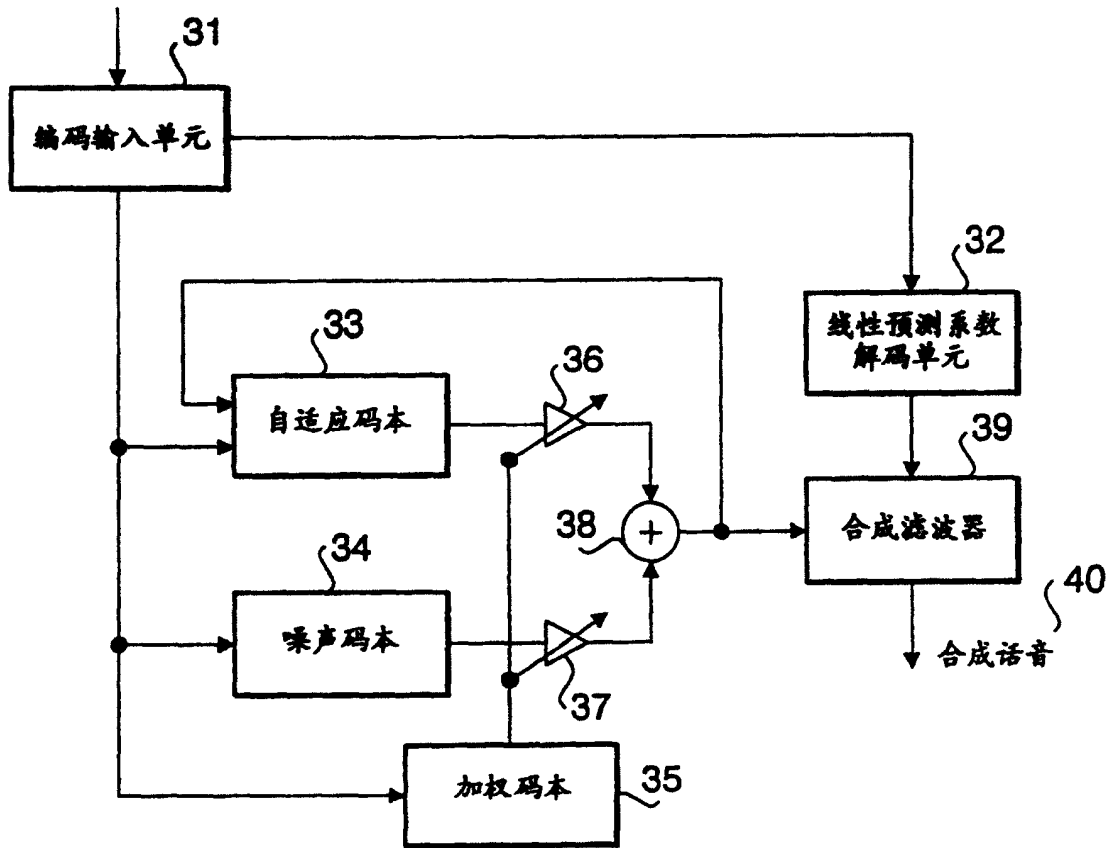
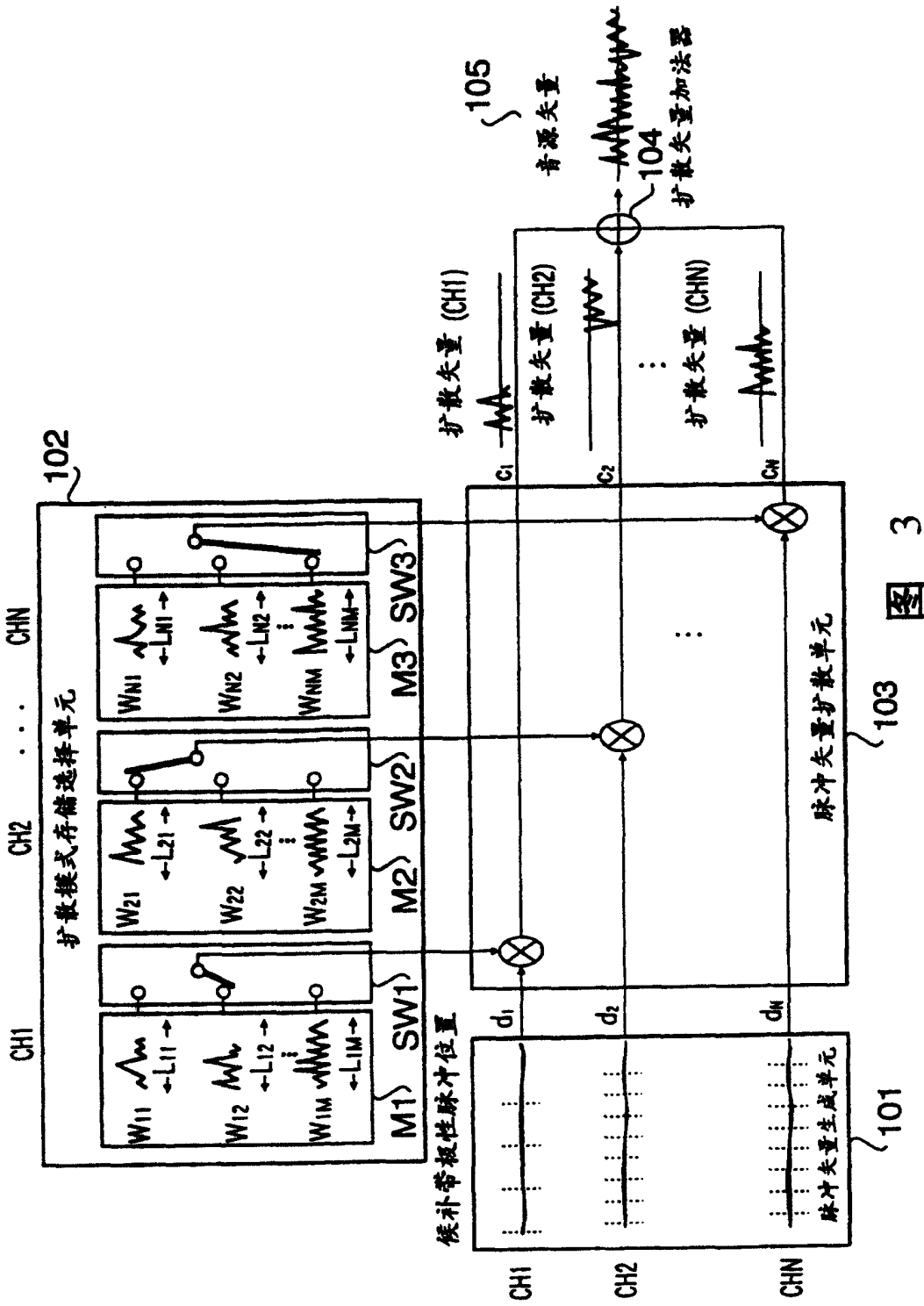
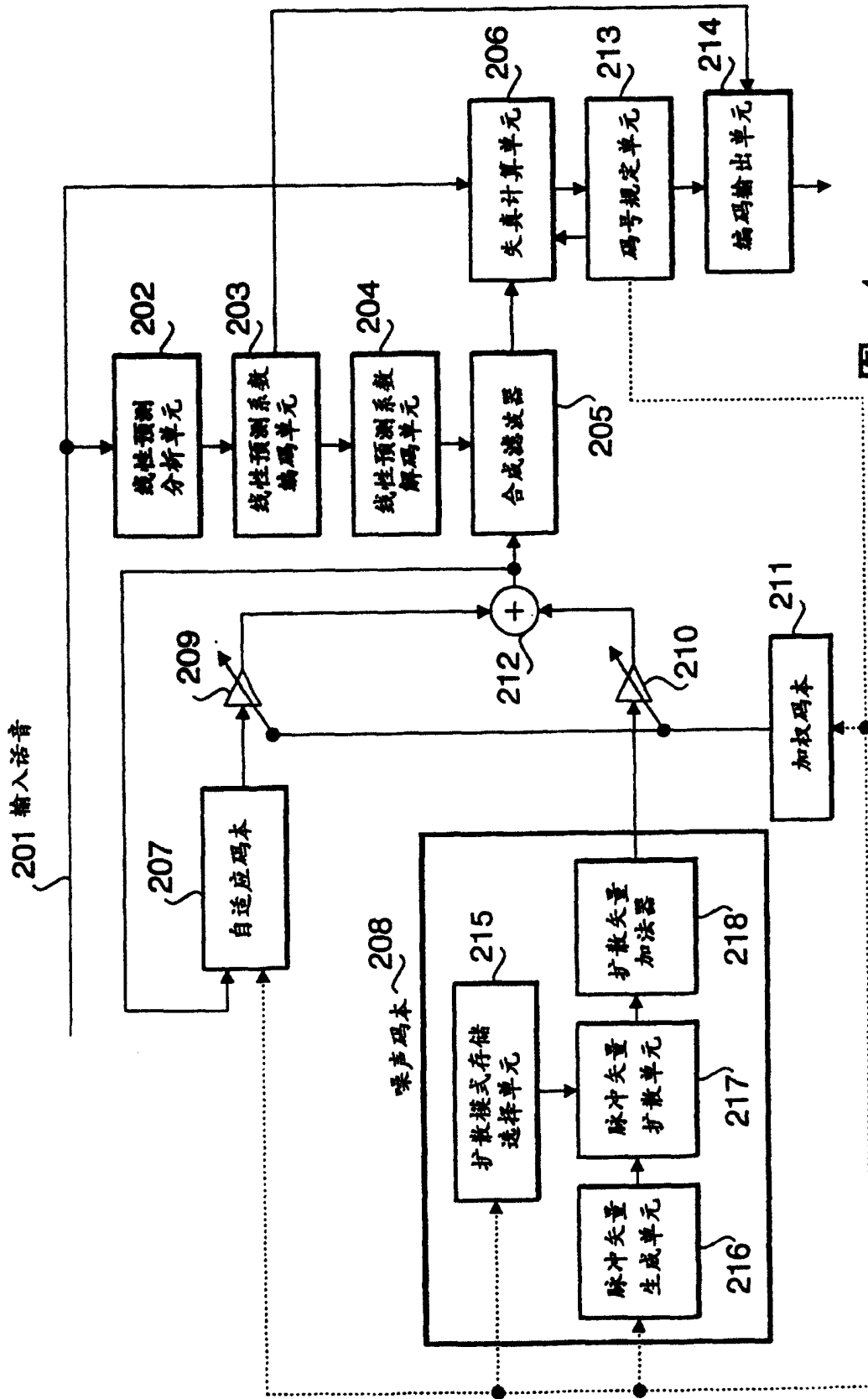


图 2





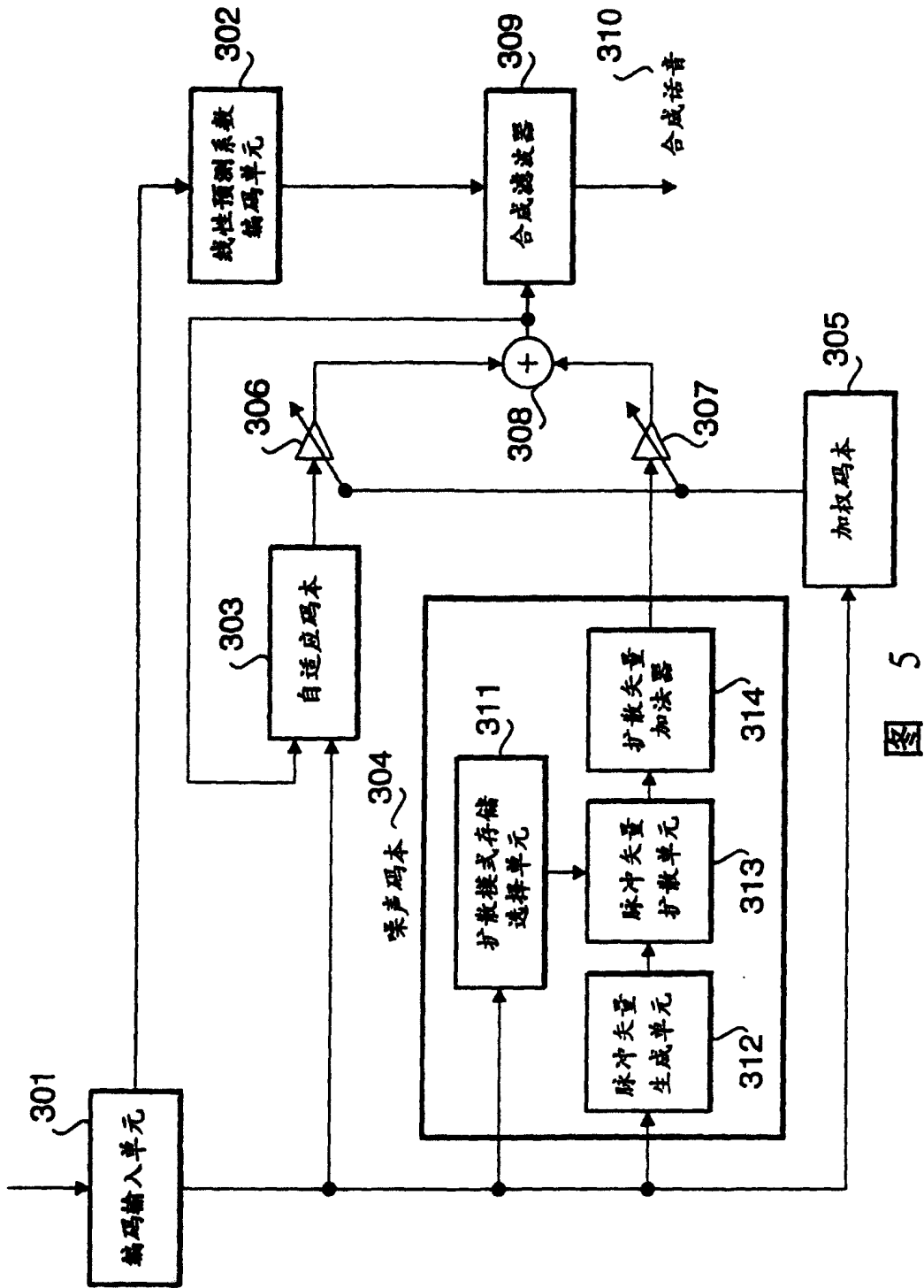


图 5

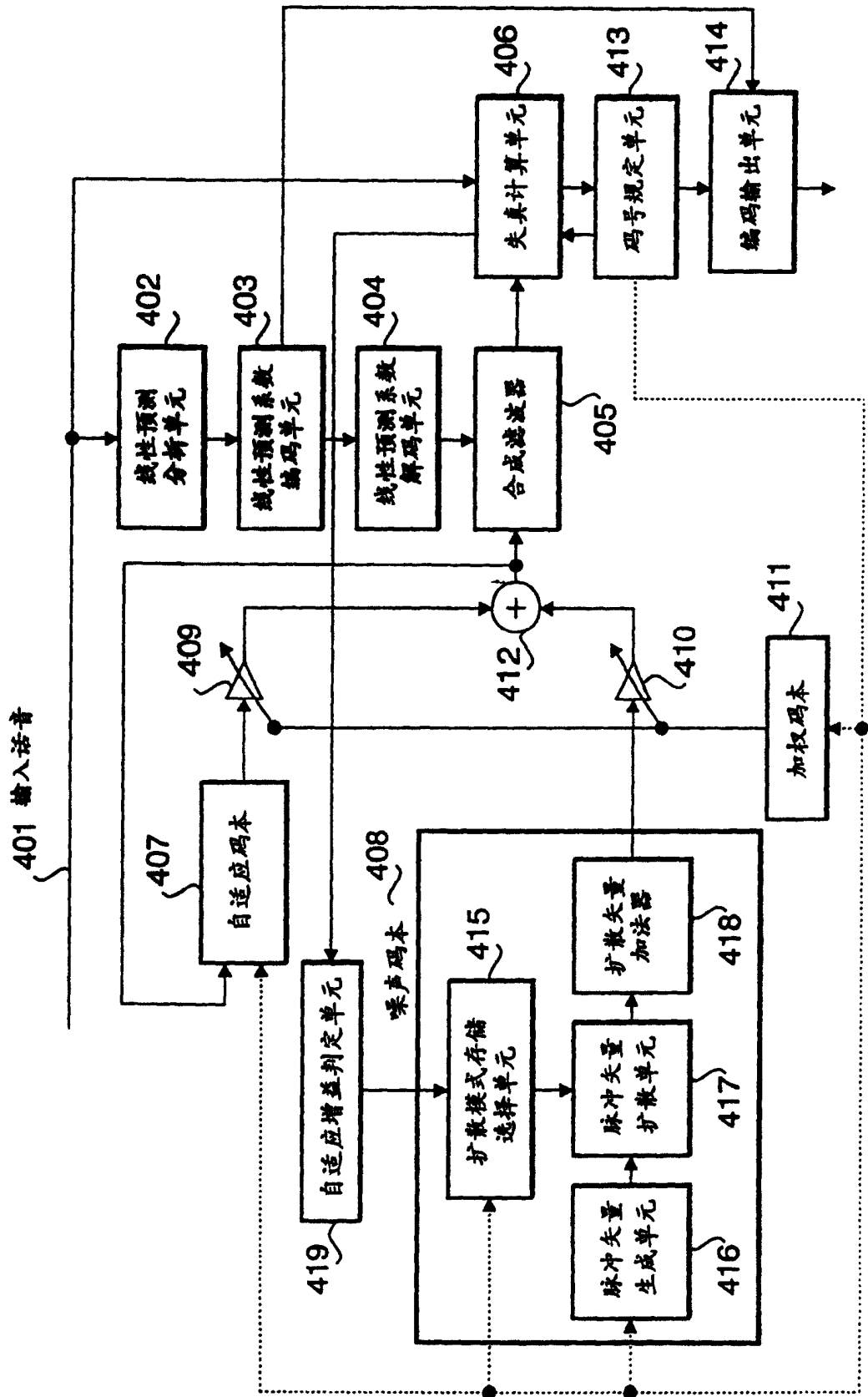


图 6

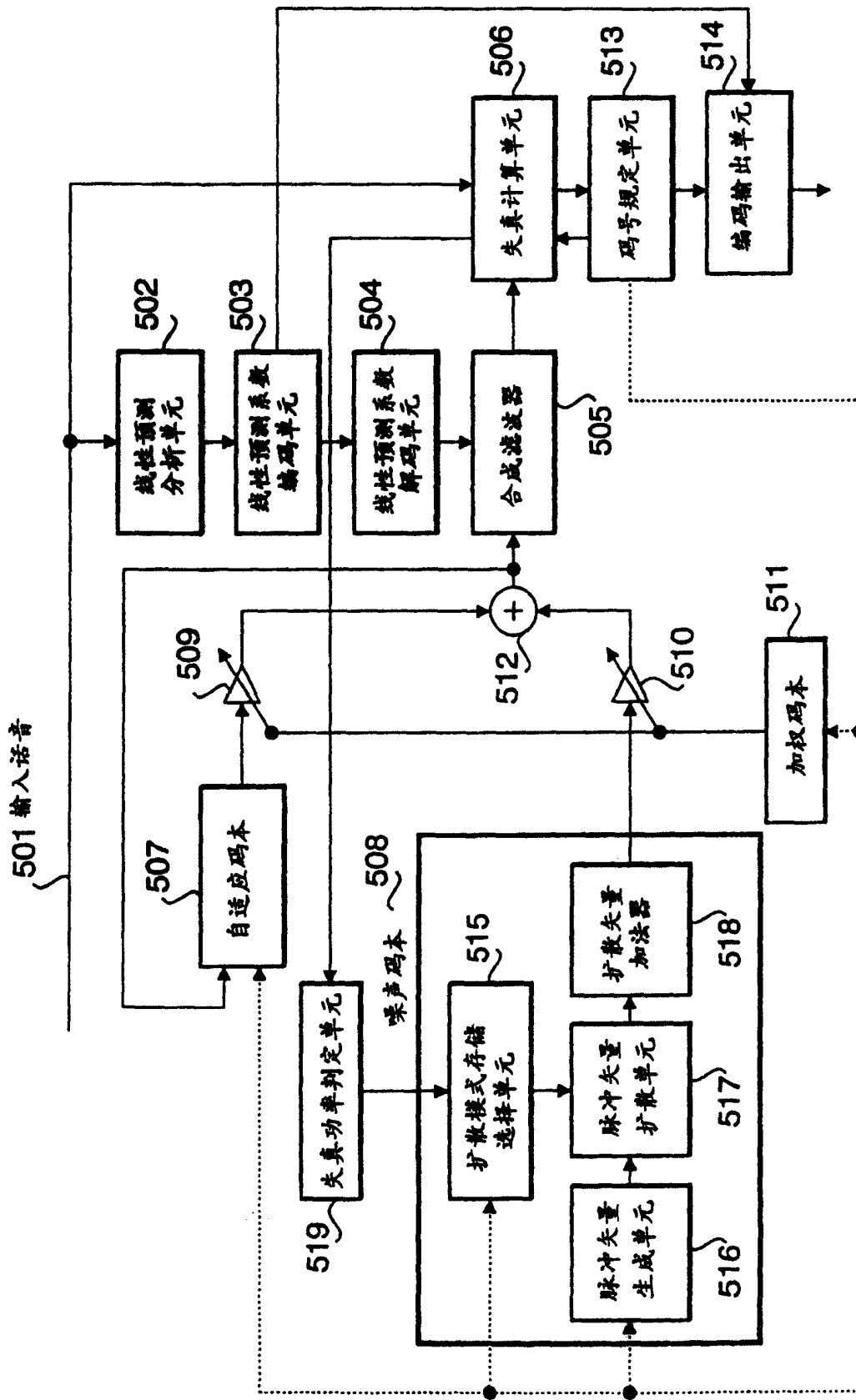


图 7

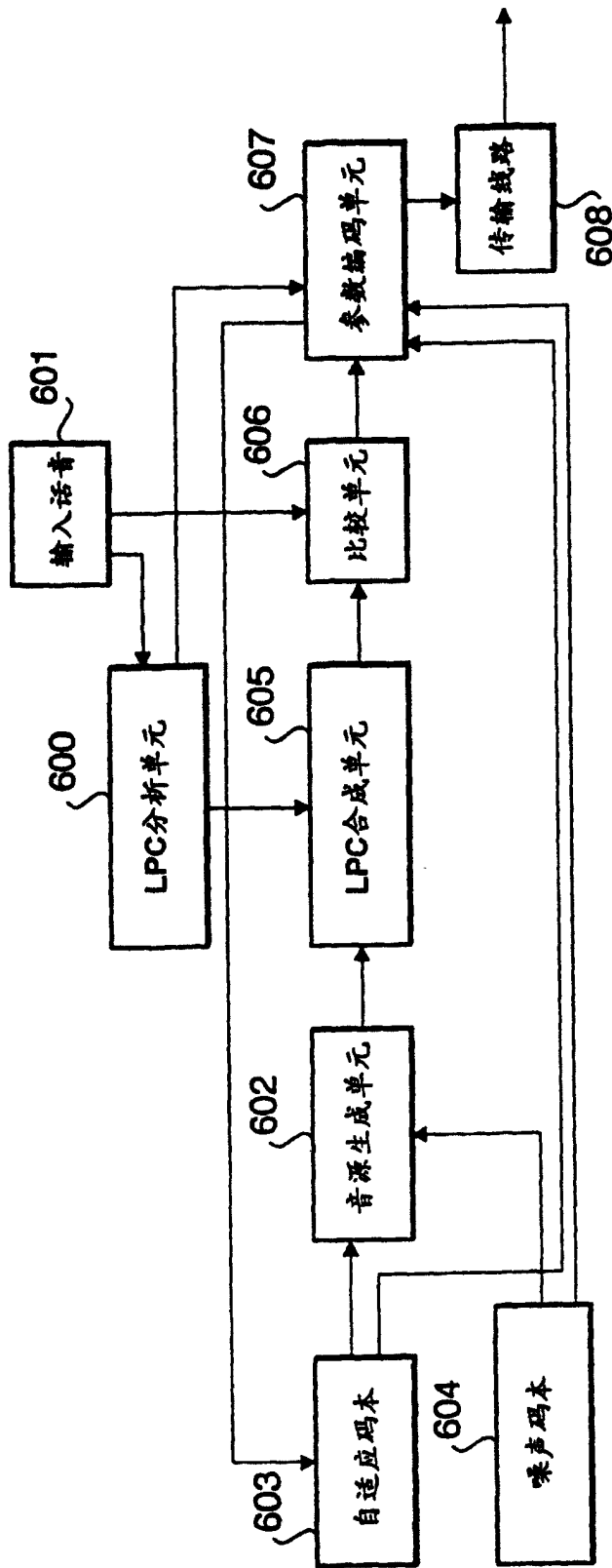


图 8

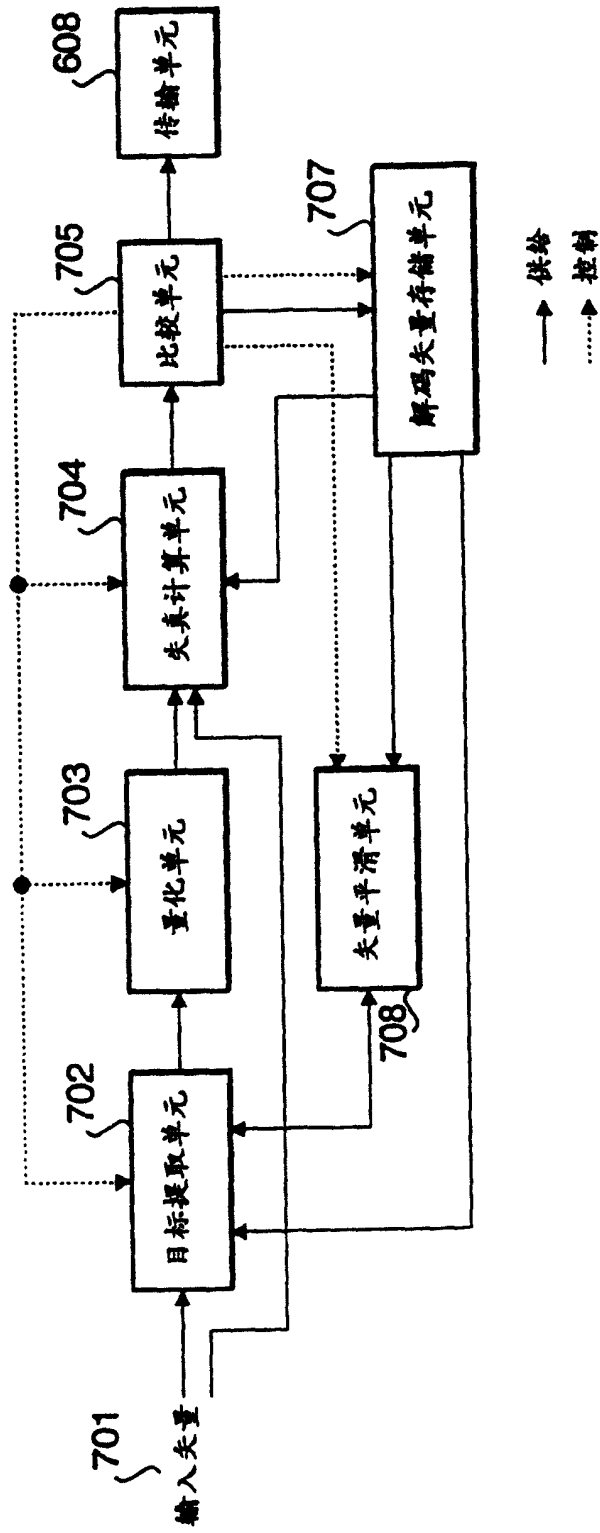


图 9

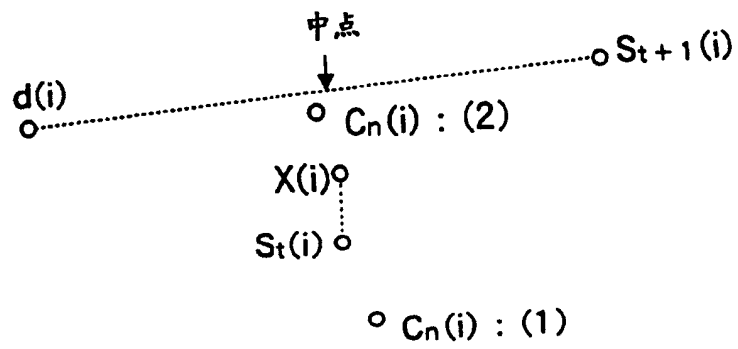


图 10

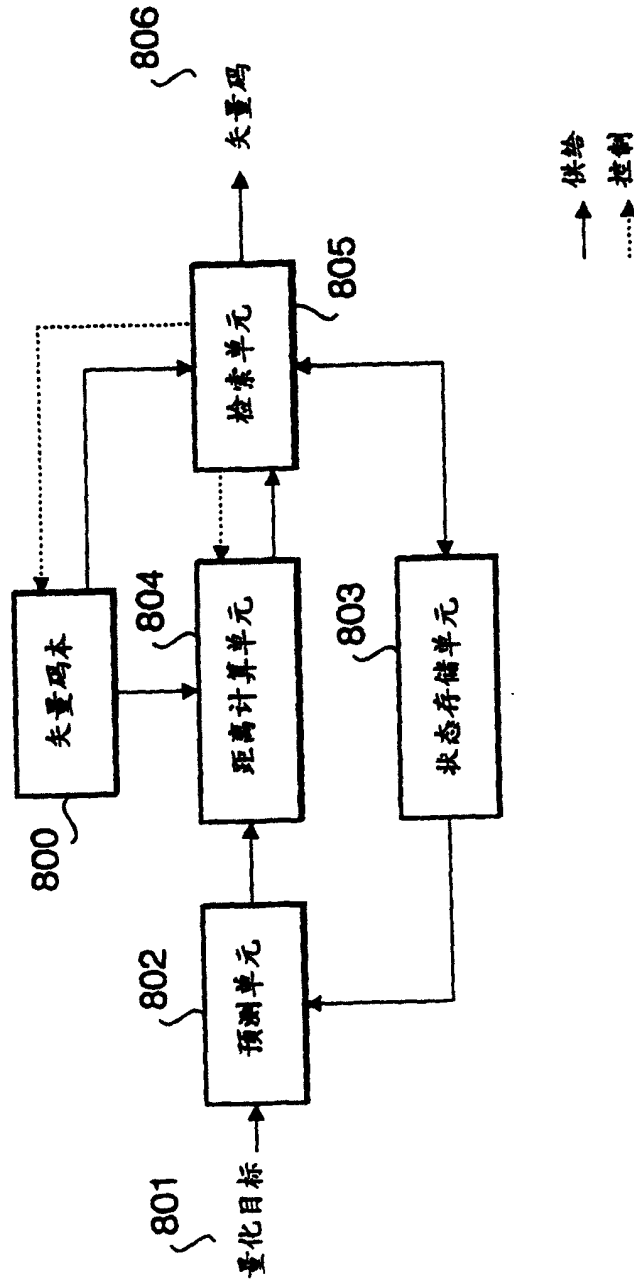


图 11

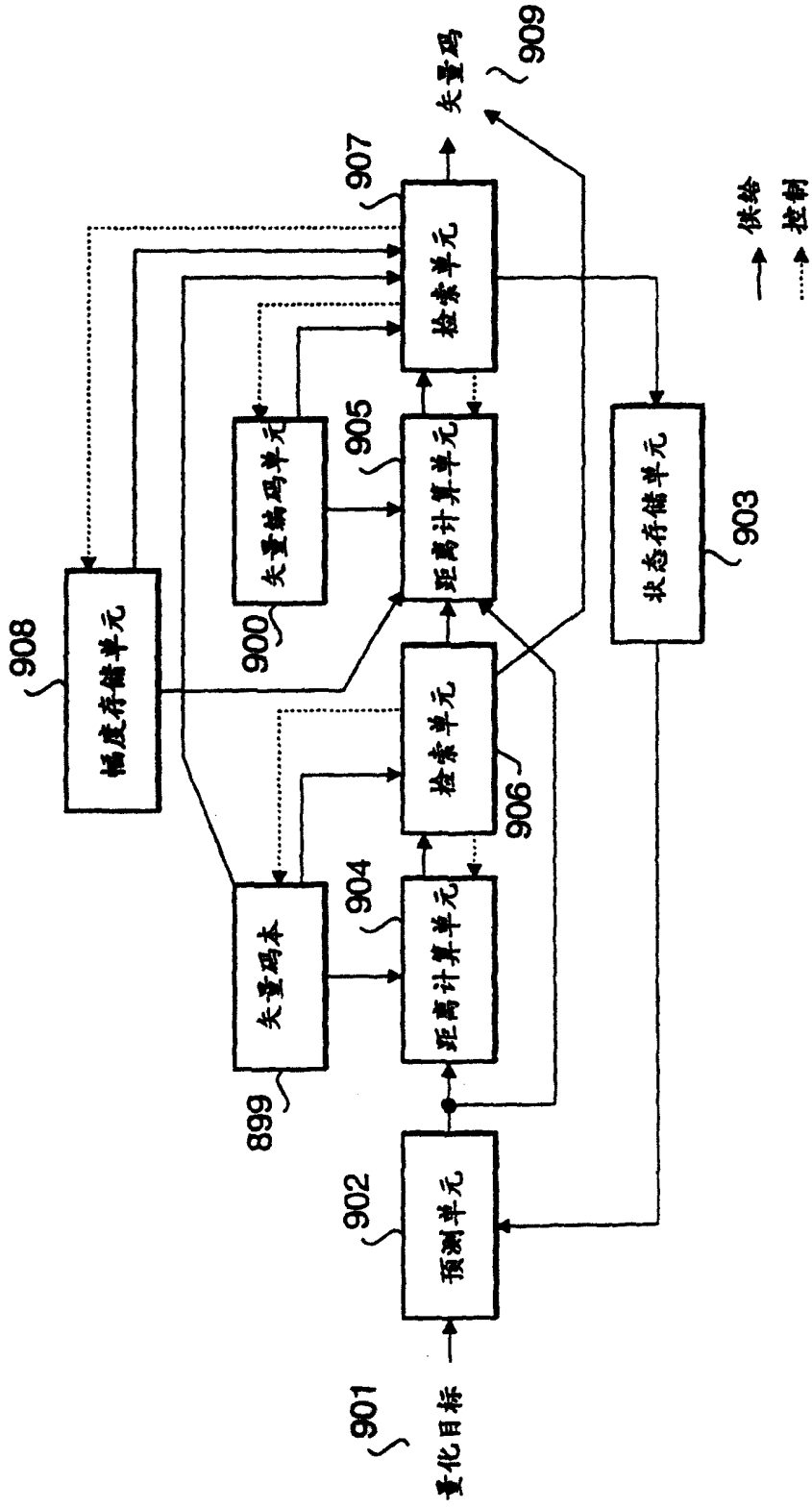


图 12

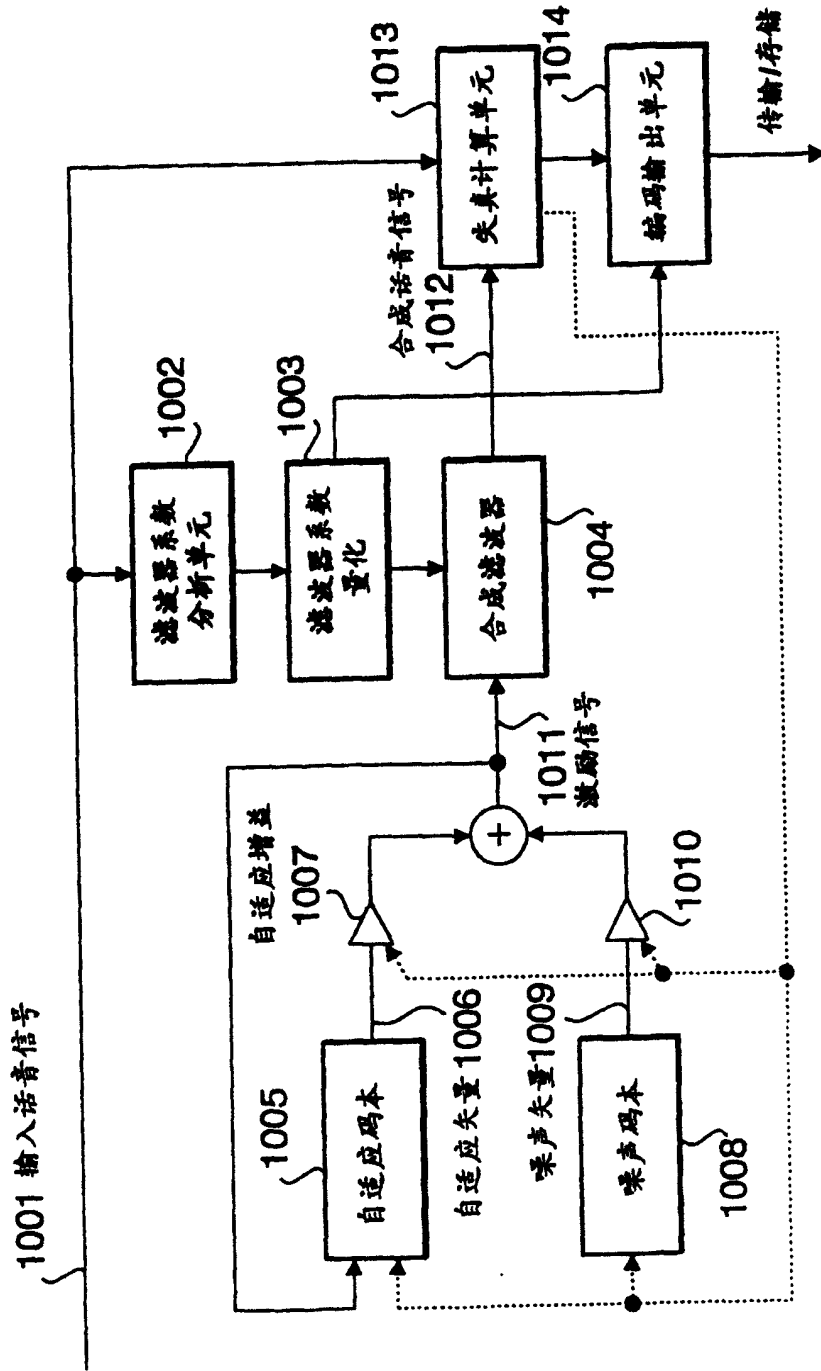


图 13

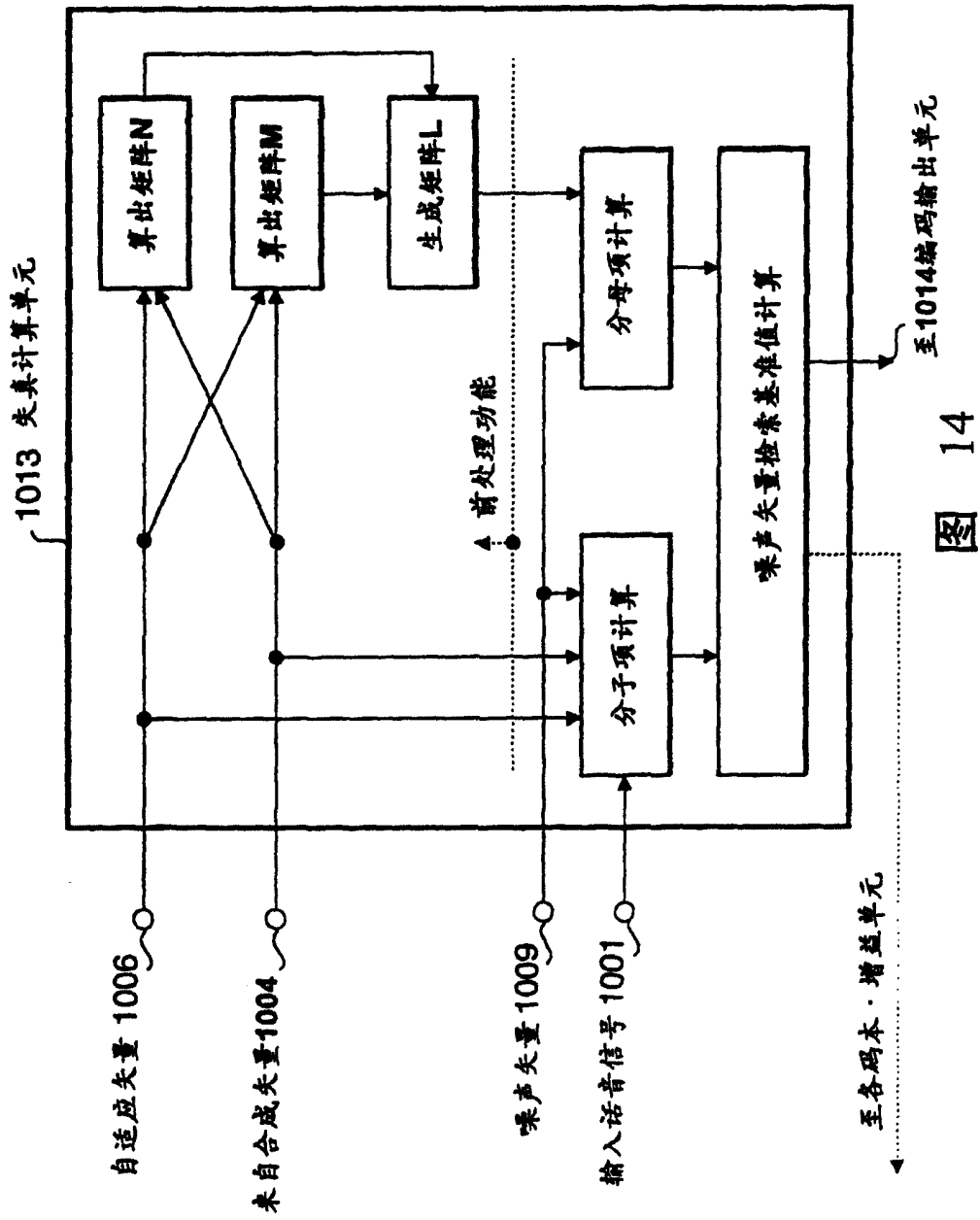


图 14