

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G10L 19/04 (2006.01)

G10L 19/12 (2006.01)

H03M 7/30 (2006.01)



# [12] 发明专利申请公开说明书

[21] 申请号 200480014149.X

[43] 公开日 2006年6月28日

[11] 公开号 CN 1795495A

[22] 申请日 2004.4.30

[21] 申请号 200480014149.X

[30] 优先权

[32] 2003.4.30 [33] JP [31] 125665/2003

[86] 国际申请 PCT/JP2004/006294 2004.4.30

[87] 国际公布 WO2004/097796 日 2004.11.11

[85] 进入国家阶段日期 2005.11.22

[71] 申请人 松下电器产业株式会社

地址 日本大阪府

[72] 发明人 佐藤薰 森井利幸

[74] 专利代理机构 北京市柳沈律师事务所

代理人 郭定辉 黄小临

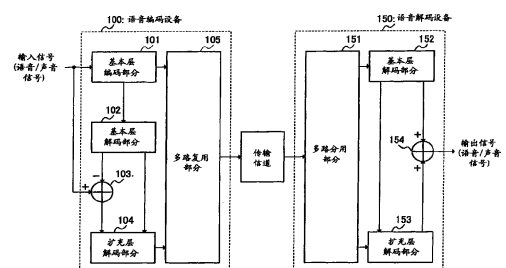
权利要求书 3 页 说明书 16 页 附图 9 页

## [54] 发明名称

音频编码设备、音频解码设备、音频编码方法和音频解码方法

## [57] 摘要

基本层编码部分(101)编码输入信号并获取基本层编码信息。基本层解码部分(102)解码基本层编码信息并获取基本层解码信号和长期预测信息(基音延迟)。加法器(103)颠倒基本层解码信号的极性并将它与输入信号相加以便获取残差信号。扩充层编码部分(104)编码利用长期预测信息和残差信号计算出的长期预测系数并获取扩充层编码信息。基本层解码部分(152)解码基本层编码信息并获取基本层解码信息和长期预测信息。扩充层解码部分(153)利用长期预测信息解码扩充层编码信息并获取扩充层解码信号。加法器(154)将基本层解码信号和扩充层解码信号相加,以便获取音频/音乐复合信号。因此,可以以少量计算和少量编码信息实现可伸缩编码。



1. 一种语音编码设备, 包括:  
基本层编码器, 用于编码输入信号并生成第一编码信息;  
5 基本层解码器, 用于解码第一编码信息并生成第一解码信号, 同时生成包括代表语音或声音的长期相关性的信息的长期预测信息;  
加法器, 用于获取代表输入信号与第一解码信息之间的差值的残差信号;  
和  
扩充层编码器, 用于利用长期预测信息和残差信号计算长期预测系数,  
10 和编码长期预测系数并生成第二编码信息。
  2. 根据权利要求1所述的语音编码设备, 其中, 基本层解码器将指定从激励矢量信号样本中取出自适应激励矢量的取出位置的信息用作长期预测信息。
  3. 根据权利要求1所述的语音编码设备, 其中, 扩充层编码器包括:  
15 根据长期预测信息获取扩充层的长期预测延迟的部分;  
从存储在缓冲器中的前长期预测信号序列中取出倒退长期预测延迟的长期预测信号的部分;  
利用残差信号和长期预测信号计算长期预测系数的部分;  
编码长期预测系数并生成扩充层编码信息的部分;  
20 解码扩充层编码信息并生成解码长期预测系数的部分; 和  
利用解码长期预测系数和长期预测信号计算新长期预测信号, 以及利用新长期预测信号更新缓冲器的部分。  - 4. 根据权利要求3所述的语音编码设备, 其中, 扩充层编码器进一步包括:  
25 获取代表残差信号与长期预测信号之间的差值的长期预测残差信号的部分;  
编码长期预测残差信号并生成长期预测残差编码信息的部分;  
解码长期预测残差编码信息并计算解码长期预测残差信号的部分; 和  
相加新长期预测信号并解码长期预测残差信号, 以及利用相加结果更新  
30 缓冲器的部分。  - 5. 一种从根据权利要求1所述的语音编码设备接收第一编码信息和第二

编码信息并解码语音的语音解码设备，所述语音解码设备包括：

基本层解码器，用于解码第一编码信息以生成第一解码信号，同时生成包括代表语音或声音的长期相关性的信息的长期预测信息；

5 扩充层解码器，用于利用长期预测信息解码第二编码信息并生成第二解码信号；和

加法器，用于相加第一解码信号和第二解码信号，并输出作为相加结果的语音或声音信号。

6. 根据权利要求5所述的语音解码设备，其中，基本层解码器将指定从激励矢量信号样本中取出自适应激励矢量的取出位置的信息用作长期预测信息。

7. 根据权利要求5所述的语音解码设备，其中，扩充层解码器包括：  
根据长期预测信息获取扩充层的长期预测延迟的部分；  
从存储在缓冲器中的前长期预测信号序列中取出倒退长期预测延迟的长期预测信号的部分；

15 解码扩充层编码信息并获取解码长期预测系数的部分；和  
利用解码长期预测系数和长期预测信号计算长期预测信号，以及利用长期预测信号更新缓冲器的部分，

其中，扩充层解码器将长期预测信号用作扩充层解码信号。

8. 根据权利要求7所述的语音解码设备，其中，扩充层解码器进一步包括：

20 解码长期预测残差编码信息并获取解码长期预测残差信号的部分；和  
相加长期预测信号并解码长期预测残差信号的部分，  
其中，扩充层解码器将相加结果用作扩充层解码信号。

9. 一种配有语音编码设备的语音信号发送设备，其中，该语音编码设备

25 包括：  
基本层编码器，用于编码输入信号并生成第一编码信息；

基本层解码器，用于解码第一编码信息并生成第一解码信号，同时生成包括代表语音或声音的长期相关性的信息的长期预测信息；

加法器，用于获取代表输入信号与第一解码信息之间的差值的残差信号；

30 和

扩充层编码器，用于利用长期预测信息和残差信号计算长期预测系数，

编码长期预测系数，并生成第二编码信息。

10. 一种配有从根据权利要求 1 所述的语音编码设备接收第一编码信息和第二编码信息并解码语音的语音解码设备的语音信号接收设备，所述语音信号接收设备包括：

5 基本层解码器，用于解码第一编码信息以生成第一解码信号，同时生成包括代表语音或声音的长期相关性的信息的长期预测信息；

扩充层解码器，用于利用长期预测信息解码第二编码信息并生成第二解码信号；和

10 加法器，用于相加第一解码信号和第二解码信号，并输出作为相加结果的语音或声音信号。

11. 一种语音编码方法，包括：

编码输入信号并生成第一编码信息；

解码第一编码信息并生成第一解码信号，同时生成包括代表语音或声音的长期相关性的信息的长期预测信息；

15 获取代表输入信号与第一解码信息之间的差值的残差信号；和

利用长期预测信息和残差信号计算长期预测系数，编码长期预测系数，并生成第二编码信息。

12. 一种利用在根据权利要求 11 所述的语音编码方法中生成的第一编码信息和第二编码信息解码语音的语音解码方法，所述语音解码方法包括：

20 解码第一编码信息以生成第一解码信号，同时生成包括代表语音或声音的长期相关性的信息的长期预测信息；

利用长期预测信息解码第二编码信息并生成第二解码信号；和

相加第一解码信号和第二解码信号，并输出作为相加结果的语音或声音信号。

## 音频编码设备、音频解码设备、音频编码方法和音频解码方法

## 5 技术领域

本发明涉及用在编码和发送语音和/或声音信号的通信系统中的语音编码设备、语音解码设备及其方法。

## 背景技术

10 在数字无线通信、以因特网为代表的分组通信和语音存储等领域中，为了有效地利用无线电信号的传输信道容量和存储媒体，编码/解码语音信号的技术是必不可少的，并且，人们已经开发出许多语音编码/解码方案。在这些系统当中，CELP（码激励线性预测）语音编码/解码方案已经实际用作主流技术。

15 CELP型语音编码设备根据事先存储的语音模型编码输入语音。更具体地说，CELP型语音编码设备将数字化语音信号划分成大约20 ms的帧，一帧一帧地对语音信号进行线性预测分析，获取线性预测系数和线性预测残差矢量，和分开编码线性预测系数和线性预测残差矢量。

20 为了进行低位速率通信，由于存储的语音模型量是有限的，在传统CELP型语音编码/解码方案中主要存储发声语音模型。

在诸如因特网通信之类发送分组的通信系统中，取决于网络状态，会发生分组丢失，最好，即使部分编码信息丢失了，也可以从其余编码信息部分中解码出语音和声音。类似地，在根据通信容量改变位速率的可变速率通信系统中，当通信容量降低时，最好，通过只发送部分编码信息就可以轻易地减轻通信容量的负担。因此，作为允许利用全部编码信息或部分编码信息解

25 码语音和声音的技术，最近人们把注意力放在可伸缩编码技术上。一些可伸缩编码方案已经按惯例公开。

可伸缩编码系统一般包括基本层和扩充层，这些层构成基本层是最低层的分层结构。在每个层中，编码等于较低层中的输入信号和输出信号之间的

30 差值的残差信号。根据这种构造，可以利用所有层的编码信息或只利用较低层的编码信息解码语音和/或声音信号。

但是，在传统可伸缩编码系统中，CELP 型语音编码/解码系统用作基本层和扩充层的编码方案，从而在计算和编码信息两个方面都需要相当大的数量。

## 5 发明内容

因此，本发明的目的是提供利用少量计算和编码信息就能够实现可伸缩编码的语音编码设备、语音解码设备及其方法。

上述的目的通过如下的步骤达到：提供进行长期预测的扩充层，利用语音或声音的长期相关性特性对扩充层中的残差信号进行长期预测以便提高解码信号的质量，利用基本层的长期预测信息获取长期预测日志，从而减少计算量。

### 附图简述

图 1 是例示根据本发明第一实施例的语音编码设备和语音解码设备的配置 15 的方块图；

图 2 是例示根据上面实施例的基本层编码部分的内部配置的方块图；

图 3 是说明根据上面实施例的基本层编码部分中的参数确定部分确定从自适应激励码本中生成的信号的处理的图形；

图 4 是例示根据上面实施例的基本层解码部分的内部配置的方块图；

20 图 5 是例示根据上面实施例的扩充层（enhancement layer）编码部分的内部配置的方块图；

图 6 是例示根据上面实施例的扩充层解码部分的内部配置的方块图；

图 7 是例示根据本发明第二实施例的扩充层编码部分的内部配置的方块图；

25 图 8 是例示根据上面实施例的扩充层解码部分的内部配置的方块图；和

图 9 是例示根据本发明第三实施例的语音信号发送设备和语音信号接收设备的配置的方块图。

### 具体实施方式

30 下面将参照附图具体描述本发明的实施例。在每个实施例中描述一种情况，在其中，在包括基本层和扩充层的双层语音编码/解码方法的扩充层中

进行长期预测的情况。但是，本发明不局限于这样的层结构，本发明可应用于在存在三层或更多层的分层语音编码/解码方法中利用较低层的长期预测信息在较高层中进行长期预测的任何情况。分层语音编码方法指的是在较高层中存在通过长期预测编码残差信号（较低层的输入信号与较低层的解码信号之间的差值）以便输出编码信息的数种语音编码方法和这些语音编码方法构成分层结构。并且，分层语音解码方法指的是在较高层中存在解码残差信号的数种语音解码方法和这些语音解码方法构成分层结构。这里，存在于最低层中的语音/声音编码/解码方法被称作基本层。存在于比基本层高的层中的语音/声音编码/解码方法被称作扩充层。

10 在本发明的每个实施例中，举例描述基本层进行 CELP 型语音编码/解码的情况。

（第一实施例）

图 1 是例示根据本发明第一实施例的语音编码设备和语音解码设备的配置的方块图。

15 在图 1 中，语音编码设备 100 主要包括基本层编码部分 101、基本层解码部分 102、相加部分 103、扩充层编码部分 104 和多路复用部分 105。语音解码设备 150 主要包括多路分用部分 151、基本层解码部分 152、扩充层解码部分 153 和相加部分 154。

基本层编码部分 101 接收语音或声音信号，利用 CELP 型语音编码方法编  
20 码输入信号，并将通过编码获得的基本层编码信息输出到基本层解码部分 102 和多路复用部分 105。

基本层解码部分 102 利用 CELP 型语音解码方法解码基本层编码信息，并将通过解码获得的基本层解码信号输出到相加部分 103。并且，基本层解码部分 102 将基音延迟（pitch lag）输出到扩充层编码部分 104，作为基本层  
25 的长期预测信息。

“长期预测信息”是指示语音或声音信号的长期相关性的信息。“基音延迟”涉及由基本层指定的位置信息，以后会作更详细描述。

相加部分 103 颠倒从基本层解码部分 102 输出的基本层解码信号的极性以便与输入信号相加，并且将作为相加结果的残差信号输出到扩充层编码部  
30 分 104。

扩充层编码部分 104 利用从基本层解码部分 102 输出的长期预测信息并

从相加部分 103 输出的残差信号计算长期预测系数，编码长期预测系数，并将通过编码获得的扩充层编码信息输出到多路复用部分 105。

5 多路复用部分 105 多路复用从基本层编码部分 101 输出的基本层编码信息和从扩充层编码部分 104 输出的扩充层编码信息，以便通过传输信道输出到多路分用部分 151 作为多路复用信息。

多路分用部分 151 将从语音编码设备 100 发送的多路复用信息多路分用成基本层编码信息和扩充层编码信息，并将多路分用基本层编码信息输出到基本层解码部分 152，同时将多路分用扩充层编码信息输出到扩充层解码部分 153。

10 基本层解码部分 152 利用 CELP 型语音解码方法解码基本层编码信息，并将通过解码获得的基本层解码信号输出到相加部分 154。并且，基本层解码部分 152 将基音延迟输出到扩充层解码部分 153，作为基本层的长期预测信息。扩充层解码部分 153 利用长期预测信息解码扩充层编码信息，并将通过解码获得的扩充层解码信号输出到相加部分 154。

15 相加部分 154 相加从基本层解码部分 152 输出的基本层解码信号和从扩充层解码部分 153 输出的扩充层解码信号，并将作为相加结果的语音或声音信号输出到供随后处理用的设备。

下面参照图 2 的方块图描述图 1 的基本层编码部分 101 的内部配置。

20 将基本层编码部分 101 的输入信号输入预处理部分 200。预处理部分 200 进行除去 DC（直流）成分的高通滤波处理、整形处理以及用以改善后续编码处理的性能的预加强处理，并将经过处理的信号（ $X_{in}$ ）输出到 LPC（线性预测系数）分析部分 201 和加法器 204。

25 LPC 分析部分 201 利用  $X_{in}$  进行线性预测分析，并将线性预测的结果（线性预测系数）输出到 LPC 量化部分 202。LPC 量化部分 202 对从 LPC 分析部分 201 输出的线性预测系数进行量化处理，并将量化 LPC 输出到合成滤波器 203，同时将代表量化 LPC 的代码（L）输出到多路复用部分 213。

合成滤波器 203 通过利用基于量化 LPC 的滤波系数对从如后所述的相加部分 210 输出的激励矢量进行滤波合成，生成合成信号，并将合成信号输出到加法器 204。

30 加法器 204 颠倒合成信号的极性，将所得信号与  $X_{in}$  相加，计算误差信号，并将误差信号输出到听觉加权部分 211。



自适应激励码本 205 含有存储在缓冲器中的较早从加法器 210 输出的激励矢量信号，并且从参数确定部分 212 输出的信号指定的较早激励矢量信号样本中取出与一个帧相对应的样本，输出到乘法器 208。

5 量化增益生成部分 206 将参数确定部分 212 输出的信号指定的自适应激励增益和固定激励增益分别输出到乘法器 208 和 209。

固定激励码本 207 将具有参数确定部分 212 输出的信号指定的形状的脉冲激励矢量乘以扩展矢量，并将获得的固定激励矢量输出到乘法器 209。

乘法器 208 将量化增益生成部分 206 输出的量化自适应激励增益乘以自适应激励码本 205 输出的自适应激励矢量，并将结果输出到加法器 210。乘法器 209 将量化增益生成部分 206 输出的量化固定激励增益乘以固定激励码本 207 输出的固定激励矢量，并将结果输出到加法器 210。

加法器 210 接收分别从乘法器 208 和 209 输入的两者的都已乘以增益的自适应激励矢量和固定激励矢量，并且以矢量形式将它们相加，并将作为相加结果的激励矢量输出到合成滤波器 203 和自适应激励码本 205。另外，将输入自适应激励码本 205 的激励矢量存储在缓冲器中。

听觉加权部分 211 对从加法器 204 输出的信号进行听觉加权，在听觉加权区中计算  $X_{in}$  与合成信号之间的失真，并将结果输出到参数确定部分 212。

参数确定部分 212 选择分别来自自适应激励码本 205、固定激励码本 207 和量化增益生成部分 206 的使听觉加权部分 211 输出的编码失真最小的自适应激励矢量、固定激励矢量和量化增益，并将代表选择结果的自适应激励矢量代码 (A)、量化增益代码 (G) 和固定激励矢量代码 (F) 输出到多路复用部分 213。另外，自适应激励矢量代码 (A) 是与基音延迟相对应的代码。

多路复用部分 213 从 LPC 量化部分 202 接收代表量化 LPC 的代码 (L)，进一步从参照确定部分 212 接收代表自适应激励矢量的代码 (A)、代表固定激励矢量的代码 (F) 和代表量化增益的代码 (G)，并多路复用这几个信息以便作为基本层编码信息输出。

上文是图 1 的基本编码部分 101 的内部配置的说明。

下面参照图 3，主要描述参数确定部分 212 确定要从自适应激励码本 205 中生成的信号的处理。在图 3 中，缓冲器 301 是配备在自适应激励码本 205 中的缓冲器，位置 302 是自适应激励矢量的取出位置，而矢量 303 是取出的自适应激励矢量。数值“41”和“296”分别对应于移动取出位置 302 的范围

的下限和上限。

假设指定给代表自适应激励矢量的代码(A)的位数是“8”，移动取出位置302的范围被设置在长度为“256”(例如，从“41”到“296”)的范围上。移动取出位置302的范围可以任意设置。

- 5 参数确定部分212在设置的范围内移动取出位置302，和按帧长从每个位置中取出自适应激励矢量303。然后，参数确定部分212获取使听觉加权部分211输出的编码失真最小的取出位置302。

参数确定部分212如此获得的缓冲器中的取出位置302是“基音延迟”。

下面参照图4描述图1的基本层解码部分102(152)的内部配置。

- 10 在图4中，输入基本层解码部分102(152)的基本层编码信息由多路分用部分401多路分用，分解成代码(L、G和F)。多路分用LPC代码(L)输出到LPC解码部分402，多路分用自适应激励矢量代码(A)输出到自适应激励码本405，多路分用量化增益代码(G)输出到量化增益生成部分406，而多路分用固定激励矢量代码(F)输出到固定激励码本407。

- 15 LPC解码部分402从多路分用部分401输出的代码(L)中解码LPC，并将结果输出到合成滤波器403。

- 自适应激励码本405从多路分用部分401输出的代码(A)指定的过去激励矢量信号中取出与一个帧相对应的样本作为激励矢量，并将激励矢量输出到乘法器408。并且，自适应激励码本405还将作为长期预测信息的基音延迟输出到扩充层编码部分104(扩充层然解码部分153)。
- 20

量化增益生成部分406分别解码多路分用部分401输出的量化增益代码(G)指定的自适应激励矢量增益和固定激励矢量增益，并将结果输出到乘法器408和409。

- 固定激励码本407生成多路分用部分401输出的代码(F)指定的固定激励矢量，并将结果输出到加法器409。
- 25

乘法器408将自适应激励矢量乘以自适应激励矢量增益，并将结果输出到加法器410。乘法器409将固定激励矢量乘以固定激励矢量增益，并将结果输出到加法器410。

- 加法器410相加分别从乘法器408和409输出的两者都乘以增益的自适应激励矢量和固定激励矢量，生成激励矢量，并将这个激励矢量输出到合成滤波器403和自适应激励码本405。
- 30

合成滤波器 403 利用作为激励信号从加法器 410 输出的激励矢量和进一步利用在 LPC 解码部分 402 中解码的滤波系数进行滤波合成，并将合成信号输出到后处理部分 404。

后处理部分 404 对合成滤波器 403 输出的信号进行诸如共振峰加强和基音加强之类提高语音的主观质量的处理，以及提高静态噪声的主观质量的其他处理，以便作为基本层解码信号输出。

上文是图 1 的基本层解码部分 102 的内部配置的说明。

下面参照图 5 描述图 1 的扩充层编码部分 104 的内部配置。

扩充层编码部分 104 将残差信号划分成  $N$  个样本 ( $N$  是自然数) 的片段，并在将  $N$  个样本作为一个帧的假设下，对每个帧进行编码。在下文中，残差信号用  $e(0) \sim e(X-1)$  表示，而经过编码的帧用  $e(n) \sim e(n+N-1)$  表示。这里， $X$  是残差信号的长度，而  $N$  对应于帧的长度。 $n$  是位于每个帧开头的样本，对应于  $N$  的整数倍。另外，从以前生成的信号中预测一些帧的信号的方法叫作长期预测。进行长期预测的滤波器叫作基音滤波器、梳状滤波器等。

在图 5 中，长期预测延迟指示部分 501 接收在基本层解码部分 102 中获得的长期预测信息  $t$ ，并根据该信息获取扩充层的长期预测延迟  $T$ ，以便输出到长期预测信号存储器 502。另外，当在基本层和扩充层之间出现取样频率差时，从如下方程 (1) 中获取长期预测延迟  $T$ 。另外，在方程 (1) 中， $D$  是扩充层的取样频率，而  $d$  是基本层的取样频率。

$$T = D \times t / d \quad \dots \text{方程 (1)}$$

长期预测信号存储器 502 配有存储较早生成的长期预测信号的缓冲器。当假设缓冲器的长度是  $M$  时，缓冲器包括以前生成长期预测信号的序列  $s(n-M-1) \sim s(n-1)$ 。一旦从长期预测延迟指示部分 501 接收到长期预测延迟  $T$ ，长期预测信号存储器 502 就从存储在缓冲器中的前长期预测信号序列中取出倒退长期预测延迟  $T$  的长期预测信号  $s(n-T) \sim s(n-T+N-1)$ ，并将结果输出到长期预测系数计算部分 503 和长期预测信号生成部分 506。并且，长期预测信号存储器 502 从长期预测信号生成部分 506 接收长期预测信号  $s(n) \sim s(n+N-1)$ ，并通过如下方程 (2) 更新缓冲器。

$$\begin{aligned} \hat{s}(i) &= s(i+N) \quad (i = n-M-1, \dots, n-1) \\ s(i) &= \hat{s}(i) \quad (i = n-M-1, \dots, n-1) \quad \dots \text{方程 (2)} \end{aligned}$$

另外，当长期预测延迟  $T$  短于帧长  $N$  和长期预测信号存储器 502 不能取

出长期预测信号时，将长期预测延迟  $T$  乘以整数，直到  $T$  长于帧长  $N$ ，以便能够取出长期预测信号。否则，重复倒退长期预测延迟  $T$  的长期预测信号  $s(n-T) \sim s(n-T+N-1)$  直到要取出的帧长  $N$ 。

长期预测系数计算部分 503 接收残差信号  $e(n) \sim e(n+N-1)$  和长期预测信号  $s(n-T) \sim s(n-T+N-1)$ ，并将这些信号用在如下方程 (3) 中，计算长期预测系数  $\beta$ ，以便输出到长期预测系数编码部分 504。

$$\beta = \frac{\sum_{i=0}^{N-1} e(n+i)s(n-T+i)}{\sum_{i=0}^{N-1} s(n-T+i)^2} \quad \dots \text{方程 (3)}$$

长期预测系数编码部分 504 编码长期预测系数  $\beta$ ，并将通过编码获得的扩充层编码信息输出到长期预测系数解码部分 505，同时，通过传输信道进一步将信息输出到扩充层解码部分 153。另外，作为编码长期预测系数  $\beta$  的方法，通过可伸缩量化等的方法是已知的。

长期预测系数解码部分 505 解码扩充层编码信息，并将通过解码获得的解码长期预测系数  $\beta_q$  输出到长期预测信号生成部分 506。

长期预测信号生成部分 506 接收解码长期预测系数  $\beta_q$  和长期预测信号  $s(n-T) \sim s(n-T+N-1)$  作为输入，利用该输入，通过如下方程 (4) 计算长期预测信号  $s(n) \sim s(n+N-1)$ ，并将结果输出到长期预测信号存储器 502。

$$s(n+i) = \beta_q \times s(n-T+i) \quad (i=0, \dots, N-1) \quad \dots \text{方程 (4)}$$

上文是图 1 的扩充层编码部分 104 的内部配置的说明。

下面参照图 6 的方块图描述图 1 的扩充层解码部分 153 的内部配置。

在图 6 中，长期预测延迟指示部分 601 利用基本层解码部分 152 输出的长期预测信息获取扩充层的长期预测延迟  $T$ ，以输出到长期预测信号存储器 602。

长期预测信号存储器 602 配有存储较早生成的长期预测信号的缓冲器。当缓冲器的长度是  $M$  时，缓冲器包括较早生成长期预测信号的序列  $s(n-M-1) \sim s(n-1)$ 。一旦从长期预测延迟指示部分 601 接收到长期预测延迟  $T$ ，长期预测信号存储器 602 从存储在缓冲器中的前长期预测信号序列中取出倒退长期预测延迟  $T$  的长期预测信号  $s(n-T) \sim s(n-T+N-1)$ ，以便输出到长期预测信号生成部分 604。并且，长期预测信号存储器 602 从长期预测信号生成部分 604 接收长期预测信号  $s(n) \sim s(n+N-1)$ ，并通过如上所述的方程 (2)

更新缓冲器。

长期预测系数解码部分 603 解码扩充层编码信息，并将通过解码获得的解码长期预测系数 $\beta_q$  输出到长期预测信号生成部分 604。

长期预测信号生成部分 604 接收解码长期预测系数 $\beta_q$  和长期预测信号  
5  $s(n-T) \sim s(n-T+N-1)$  作为输入，利用该输入，通过如上所述方程 (4) 计算长期预测信号  $s(n) \sim s(n+N-1)$ ，并将结果输出到长期预测信号存储器 602 和相加部分 153，作为扩充层解码信号。

上文是图 1 的扩充层解码部分 153 的内部配置的说明。

因此，通过提供进行长期预测的扩充层和利用语音或声音信号的长期相  
10 关性特性在扩充层中对残差信号进行长期预测，可以利用较少编码信息编码/解码频率范围宽的语音/声音信号和减少计算量。

此刻，取代编码/解码长期预测延迟，利用基本层的长期预测信息获取长期预测延迟可以减少编码信息。

并且，通过解码基本层编码信息，可以只获取基本层的解码信号，和可  
15 以实现用 CELP 型语音编码/解码方法（可伸缩编码）从部分编码信息中解码语音或声音的功能。

此外，在长期预测中，利用语音或声音的长期相关性从缓冲器中取出与当前帧有最高相关性的帧，以及利用取出帧的信号表达当前帧的信号。但是，在从缓冲器中取出与当前帧有最高相关性的帧的手段中，当不存在诸如基音延迟之类代表语音或声音的长期相关性的信息时，有必要改变从缓冲器中取出帧的取出位置，同时计算取出帧和当前帧的自相关函数，以便搜索有最高  
20 相关性的帧，并且，用于搜索的计算量变得相当大。

但是，通过确定唯一使用在基本层编码部分 101 中获得的基音延迟的取出位置，可以大量地减少一般长期预测所需的计算量。

另外，上面在本实施例中说明的扩充层长期预测方法中已经描述了从基本层解码部分输出的长期预测信息是基音延迟的情况，但是，本发明不局限于这种情况，只要信息代表语音或声音的长期相关性，就可以将任何信息用作长期预测信息。

并且，在本实施例中描述了长期预测信号存储器 502 从缓冲器中取出长  
30 期预测信号的位置是长期预测延迟  $T$  的情况，但是，本发明可应用于这样的位置是长期预测延迟  $T$  附近的位置  $T + \alpha$  ( $\alpha$  是一个微小数字并可任意设置)

的情况，即使在长期预测延迟  $T$  存在微小误差的情况下，也可以获得与本实施例相同的效果和优点。

例如，长期预测信号存储器 502 从长期预测延迟指示部分 501 接收长期预测延迟  $T$ ，从存储在缓冲器的前长期预测信号序列中取出倒退  $T + \alpha$  的长期预测信号  $s(n-T-\alpha) \sim s(n-T-\alpha+N-1)$ ，利用如下方程 (5) 计算确定值  $C$ ，获取使确定值  $C$  最小的  $\alpha$ ，并编码这个  $\alpha$ 。并且，在解码的情况下，长期预测信号存储器 602 解码  $\alpha$  的编码信息，以及利用长期预测延迟  $T$ ，取出长期预测信号  $s(n-T-\alpha) \sim s(n-T-\alpha+N-1)$ 。

$$C = \frac{\left[ \sum_{i=0}^{N-1} e(n+i)s(n-T-\alpha+i) \right]^2}{\sum_{i=0}^{N-1} s(n-T-\alpha+i)^2} \quad \dots \text{方程 (5)}$$

并且，虽然上面在本实施例中已经描述了利用语音/声音信号进行长期预测的情况，但是，本发明最终可应用于利用诸如 MDCT 和 QMF 之类的正交变换将语音/声音信号从时域变换到频域，以及利用变换信号（频率参数）进行长期预测的情况，并且，仍然可以获得与本实施例相同的效果和优点。例如，在利用语音/声音信号的频率参数进行扩充层长期预测的情况下，在图 5 中，长期预测系数计算部分 503 重新配备了将长期预测信号  $s(n-T) \sim s(n-T+N-1)$  从时域变换到频域的函数，并重新配备了将残差信号变换成频率参数的另一种函数，并且，长期预测系数生成部分 506 重新配备了将长期预测信号  $s(n) \sim s(n+N-1)$  从频域逆变换到时域的函数。并且，在图 6 中，长期预测系数生成部分 604 重新配备了将长期预测信号  $s(n) \sim s(n+N-1)$  从频域逆变换到时域的功能。

在一般语音/声音编码/解码方法中将用在检错或纠错中的冗余位加入编码信息中并在传输信道上发送包含冗余位的编码信息是常见的。在本发明中可以加权指定给基本层编码部分 101 输出的编码信息 (A) 和扩充层编码部分 104 输出的编码信息 (B) 的冗余位对编码信息 (A) 的位指定，以便进行指定。

### (第二实施例)

下面参照编码并解码残差信号与长期预测信号之间的差值（长期预测残差信号）的情况描述第二实施例。

除了扩充层编码部分 104 和扩充层解码部分 153 的内部配置之外，本实

施例的语音编码设备和语音解码设备的配置与图 1 中的那些相同。

图 7 是例示根据本实施例的扩充层编码部分 104 的内部配置的方块图。另外，在图 7 中，将与图 5 中相同的标号指定给与图 5 共同的结构单元，以便省略对它们的描述。

5 与图 5 相比，图 7 中的扩充层编码部分 104 进一步配有相加部分 701、长期预测残差信号编码部分 702、编码信息多路复用部分 703、长期预测残差信号解码部分 704 和相加部分 705。

长期预测信号生成部分 506 将计算出的长期预测信号  $s(n) \sim s(n+N-1)$  输出到相加部分 701 和 702。

10 正如在如下方程 (6) 中所表达的那样，相加部分 701 颠倒长期预测信号  $s(n) \sim s(n+N-1)$  的极性，将结果与残差信号  $e(n) \sim e(n+N-1)$  相加，并将作为相加结果的长期预测残差信号  $p(n) \sim p(n+N-1)$  输出到长期预测残差信号编码部分 702。

$$p(n+i) = e(n+i) - s(n+i) \quad (i = 0, \dots, N-1) \quad \dots \text{方程 (6)}$$

15 长期预测残差信号编码部分 702 编码长期预测残差信号  $p(n) \sim p(n+N-1)$ ，并将通过编码获得的编码信息（下文称为“长期预测残差编码信息”）输出到编码信息多路复用部分 703 和长期预测残差信号解码部分 704。另外，长期预测残差信号的编码一般通过矢量量化进行。

下面将进行 8 位矢量量化的情况用作一个例子描述编码长期预测残差信号  $p(n) \sim p(n+N-1)$  的方法。在这种情况下，在长期预测残差信号编码部分 702 中预备存储事先生成 256 种代码矢量的码本。代码矢量  $CODE(k)(0) \sim CODE(k)(N-1)$  是长度为  $N$  的矢量。 $k$  是代码矢量的索引和取范围从 0 到 255 的值。长期预测残差信号编码部分 702 利用如下方程 (7) 获取长期预测残差信号  $p(n) \sim p(n+N-1)$  与代码矢量  $CODE(k)(0) \sim CODE(k)(N-1)$  之间的平方误差  $er$ 。

$$er = \sum_{i=0}^{N-1} (p(n+i) - CODE^{(k)}(i))^2 \quad \dots \text{方程 (7)}$$

然后，长期预测残差信号编码部分 702 确定使平方误差  $er$  最小的  $k$  值，作为长期预测残差编码信息。

30 编码信息多路复用部分 703 多路复用从长期预测系数编码部分 504 输入的扩充层编码信息和从长期预测残差信号编码部分 702 输入的长期预测残差

编码信息，和通过传输信道将多路复用信息输出到扩充层解码部分 153。

长期预测残差信号解码部分 704 解码长期预测残差编码信息，并将解码长期预测残差信号  $pq(n) \sim pq(n+N-1)$  输出到相加部分 705。

相加部分 705 相加从长期预测信号生成部分 506 输入的长期预测信号  $s(n) \sim s(n+N-1)$  和从长期预测残差信号解码部分 704 输入的解码长期预测残差信号  $pq(n) \sim pq(n+N-1)$ ，并将相加结果输出到长期预测信号存储器 502。其结果是，长期预测信号存储器 502 利用如下方程 (8) 更新缓冲器。

$$\left. \begin{aligned} \hat{s}(i) &= s(i+N) & (i = n-M-1, \dots, n-N-1) \\ \hat{s}(i) &= s(i+N) + p(i-N) & (i = n-N, \dots, n-1) \\ \hat{s}(i) &= s(i) & (i = n-M-1, \dots, n-1) \end{aligned} \right\} \dots \text{方程 (8)}$$

10 上文是根据本实施例的扩充层编码部分 104 的内部配置的说明。

下面参照图 8 中的方块图描述根据本发明的扩充层解码部分 153 的内部配置。另外，在图 8 中，将与图 6 中相同的标号指定给与图 6 共同的结构单元，以便省略对它们的描述。

与图 6 相比，图 8 中的扩充层解码部分 153 进一步配有编码信息多路分用部分 801、长期预测残差信号解码部分 802 和相加部分 803。

编码信息多路分用部分 801 将通过传输信道接收的多路复用编码信息多路分用成扩充层编码信息和长期预测残差编码信息，并将扩充层编码信息输出到长期预测系数解码部分 603，将长期预测残差编码信息输出到长期预测残差信号解码部分 802。

20 长期预测残差信号解码部分 802 解码长期预测残差编码信息，获取解码长期预测残差信号  $pq(n) \sim pq(n+N-1)$ ，并将该信号输出到相加部分 803。

相加部分 803 相加从长期预测信号生成部分 604 输入的长期预测信号  $s(n) \sim s(n+N-1)$  和从长期预测残差信号解码部分 802 输入的解码长期预测残差信号  $pq(n) \sim pq(n+N-1)$ ，并将相加结果输出到长期预测信号存储器 602，同时

25 同时将结果作为扩充层解码信号输出。

上文是根据本实施例的扩充层解码部分 153 的内部配置的说明。

通过如此编码并解码残差信号与长期预测信号之间的差值（长期预测残差信号），可以获得质量比前面在第一实施例中所述高的解码信号。

另外，上面在本实施例中已经描述了通过矢量量化编码长期预测残差信号的情况。但是，本发明不局限于这种编码方法，可以利用，例如，形状 -

30



增益 VQ、分割 VQ、变换 VQ 或多相 VQ 进行编码。

下面将描述通过在形状方面 8 个位和在增益方面 5 个位的 13 位形状 - 增益 VQ 进行编码的情况。在这种情况下，提供了两种码本，形状码本和增益码本。形状码本包括 256 种形状代码矢量，和形状代码矢量 SCODE(k1)(0) ~ SCODE(k1)(N-1) 是长度为 N 的矢量。k1 是形状代码矢量的索引和取范围从 0 到 255 的值。增益码本包括 32 种增益代码，和增益代码 GCODE(k2) 取标量值。k2 是增益代码的索引和取范围从 0 到 31 的值。长期预测残差信号编码部分 702 利用如下方程 (9) 获取长期预测残差信号  $p(n) \sim p(n+N-1)$  的增益和形状矢量  $shape(0) \sim shape(N-1)$ ，和进一步获取增益与增益代码 GCODE(k2) 之间的增益误差  $gainer$  和形状矢量  $shape(0) \sim shape(N-1)$  与形状代码矢量 SCODE(k1)(0) ~ SCODE(k1)(N-1) 之间的平方误差  $shaper$ 。

$$gain = \sqrt{\sum_{i=0}^{N-1} p(n+i)^2}$$

$$shape(i) = \frac{p(n+i)}{gain} \quad (i=0, \dots, N-1) \quad \dots \text{方程 (9)}$$

$$gainer = |gain - GCODE^{(k2)}|$$

$$shaper = \sum_{i=0}^{N-1} (shape(i) - GCODE^{(k2)}(i))^2 \quad \dots \text{方程 (10)}$$

然后，长期预测残差信号编码部分 702 获取使增益误差  $gainer$  最小的  $k2$  值和使平方误差  $shaper$  最小的  $k1$  值，并将获得的值确定为长期预测残差编码信息。

下面描述通过 8 位分割 VQ 进行编码的情况。在这种情况下，预备了两种码本，第一分割码本和第二分割码本。

第一分割码本包括 16 种第一分割代码矢量 SPCODE(k3)(0) ~ SPCODE(k3)(N/2-1)，第二分割码本包括 16 种第二分割代码矢量 SPCODE(k4)(0) ~ SPCODE(k4)(N/2-1)，和每个代码矢量具有 N/2 的长度。k3 是第一分割代码矢量的索引和取范围从 0 到 15 的值。k4 是第二分割代码矢量的索引和取范围从 0 到 15 的值。长期预测残差信号编码部分 702 利用如下方程 (11) 将长期预测残差信号  $p(n) \sim p(n+N-1)$  划分成第一分割矢量  $sp1(0) \sim sp1(N/2-1)$  和第二分割矢量  $sp2(0) \sim sp2(N/2-1)$ ，并获取第一分割矢量  $sp1(0) \sim sp1(N/2-1)$  与第一分割代码矢量 SPCODE(k3)(0) ~ SPCODE(k3)(N/2-1) 之间的平方误差  $splitter1$  和第二分割矢量  $sp2(0) \sim sp2(N/2-1)$  与第二分割代码矢

量 SPCODE (k4) (0) ~ SPCODE (k4) (N/2-1) 之间的平方误差 splitter2。

$$sp_1(i) = p(n+i) \quad (i=0, \dots, N/2-1)$$

$$sp_2(i) = p(n+N/2+i) \quad (i=0, \dots, N/2-1) \quad \dots \text{方程 (11)}$$

$$splitter_1 = \sum_{i=0}^{N/2-1} (sp_1(i) - SPCODE_1^{(k3)}(i))^2$$

$$5 \quad splitter_2 = \sum_{i=0}^{N/2-1} (sp_2(i) - SPCODE_2^{(k4)}(i))^2 \quad \dots \text{方程 (12)}$$

然后，长期预测残差信号编码部分 702 获取使平方误差 splitter1 最小的 k3 值并使平方误差 splitter2 最小的 k4 值，并将获得的值确定为长期预测残差编码信息。

下面描述通过利用离散付里叶变换的 8 位变换 VQ 进行编码的情况。在这种情况下，预备了包括 256 种变换代码矢量的变换码本，和变换代码矢量 TCODE (k5) (0) ~ TCODE (k5) (N/2-1) 是长度为 N/2 的矢量。k5 是变换代码矢量的索引和取范围从 0 到 255 的值。长期预测残差信号编码部分 702 利用如下方程 (13) 对长期预测残差信号 p(n) ~ p(n+N-1) 进行离散付里叶变换以获取变换矢量 tp(0) ~ tp(N-1)，以及利用如下方程 (14) 获取变换矢量 tp(0) ~ tp(N-1) 与变换代码矢量 TCODE (k5) (0) ~ TCODE (k5) (N/2-1) 之间的平方误差 transfer。

$$10 \quad \hat{tp}(i) = \sum_{n=0}^{N-1} p(n+i) e^{-j \frac{2\pi n i}{N}} \quad (\hat{i}=0, \dots, N-1) \quad \dots \text{方程 (13)}$$

$$15 \quad transfer = \sum_{i=0}^{N-1} (\hat{tp}(i) - TCODE^{(k5)}(i))^2 \quad \dots \text{方程 (14)}$$

然后，长期预测残差信号编码部分 702 获取使平方误差 transfer 最小的 k5 值，并将获得的值确定为长期预测残差编码信息。

下面描述 5 个位用于第一阶段和 8 个位用于第二阶段的 13 位二相 VQ 进行编码的情况。在这种情况下，预备了第一阶段码本和第二阶段码本两种码本。第一阶段码本包括 32 种第一阶段代码矢量 PHCODE1 (k6) (0) ~ PHCODE1 (k6) (N-1)。第二阶段码本包括 256 种第二阶段代码矢量 PHCODE2 (k7) (0) ~ PHCODE2 (k7) (N-1)，并且每个代码矢量具有 N/2 的长度。k6 是第一阶段代码矢量的索引并且取范围从 0 到 31 的值。k7 是第二阶段代码矢量的索引并且取范围从 0 到 255 的值。长期预测残差信号编码部分 702 利用如下方程 (15) 获取长期预测残差信号 p(n) ~ p(n+N-1) 与第一阶段代码矢量 PHCODE1 (k6) (0) ~ PHCODE1 (k6) (N-1) 之间的平方误差 phaseer1，进一步获取

使平方误差  $phaseer_1$  最小的  $k_6$  值, 并将该值确定为  $K_{max}$ 。

$$phaseer_1 = \sum_{i=0}^{N-1} (p(i) - PHCODE_{E_1}^{(k_6)}(i))^2 \quad \dots \text{方程 (15)}$$

然后, 长期预测残差信号编码部分 702 利用如下方程 (16) 获取误差矢量  $ep(0) \sim ep(N-1)$ , 获取误差矢量  $ep(0) \sim ep(N-1)$  与第二阶段代码矢量 PHCODE2 ( $k_7$ ) (0)  $\sim$  PHCODE2 ( $k_7$ ) (N-1) 之间的平方误差  $phaseer_2$ , 进一步获取使平方误差  $phaseer_2$  最小的  $k_7$  值, 并将该值和  $K_{max}$  确定为长期预测残差编码信息。

$$ep(i) = p(n+i) - PHCODE_{E_1}^{(k_{max})}(i) \quad (i=0, \dots, N-1) \quad \dots \text{方程 (16)}$$

$$phaseer_2 = \sum_{i=0}^{N-1} (ep(i) - PHCODE_{E_2}^{(k_7)}(i))^2 \quad \dots \text{方程 (17)}$$

### 10 (第三实施例)

图 9 是例示分别含有在第一和第二实施例中描述的语音编码设备和语音解码设备的语音信号发送设备和语音信号接收设备的配置的方块图。

在图 9 中, 通过输入设备 902 将语音信号 901 转换成电子信号和输出到 A/D 转换设备 903。A/D 转换设备 903 将从输入设备 902 输出的 (模拟) 信号转换成数字信号并将结果输出到语音编码设备 904。语音编码设备 904 装有如图 1 所示的语音编码设备 100, 编码从 A/D 转换设备 903 输出的数字语音信号, 并将编码信息输出到 RF 调制设备 905。RF 调制设备 905 将从语音编码设备 904 输出的编码信息转换成诸如无线电信号之类传播媒体的信号以便加以发送, 并将信号输出到发送天线 906。发送天线 906 发送从 RF 调制设备 905 输出的输出信号作为无线电信号 (RF 信号)。另外, 图 9 中的 RF 信号 907 代表从发送天线 906 发送的无线电信号 (RF 信号)。语音信号发送设备的配置和操作就是像上述那样。

RF 信号 908 由接收天线 909 接收, 然后, 输出到 RF 解调设备 910。另外, 图 9 中的 RF 信号 908 代表如果在传播路径上没有出现信号衰减和/或噪声多路复用就与 RF 信号 907 相同、由接收天线 909 接收的无线电信号。

RF 解调设备 910 从接收天线 909 输出的 RF 信号中解调语音编码信息, 并将结果输出到语音解码设备 911。语音解码设备 911 装有如图 1 所示的语音解码设备 150, 从 RF 解调设备 910 输出的语音编码信息中解码语音信号, 并将结果输出到 D/A 转换设备 912。D/A 转换设备 912 将语音解码设备 911 输出的数字语音信号转换成模拟电子信号, 并将结果输出到输出设备 913。

输出设备 913 将电子信号转换成空气振动并将结果输出成人耳听得见的声音信号。另外，在该图中，标号 914 表示输出声音信号。语音信号接收设备的配置和操作就是像上述那样。

通过在带有上述语音信号发送设备和语音信号接收设备的无线通信系统中配备基站设备和通信终端设备，可以获得高质量的解码信号。

如上所述，根据本发明，可以利用较少编码信息编码并解码具有宽带的语音和声音信号，并可以减少计算量。并且，通过利用基本层的长期预测信息获取长期预测延迟，可以减少编码信息。此外，通过解码基本层编码信息，可以只获取基本层的解码信号，并且在 CELP 型语音编码/解码方法中，可以实现从部分编码信息（可伸缩编码）中解码语音和声音的功能。

本申请基于 2003 年 4 月 30 日提出的日本专利申请第 2003-125665 号，特此全文引用，以供参考。

#### 工业应用性

本发明适用于用在编码和发送语音和/或声音信号的通信系统中的语音编码设备和语音解码设备。

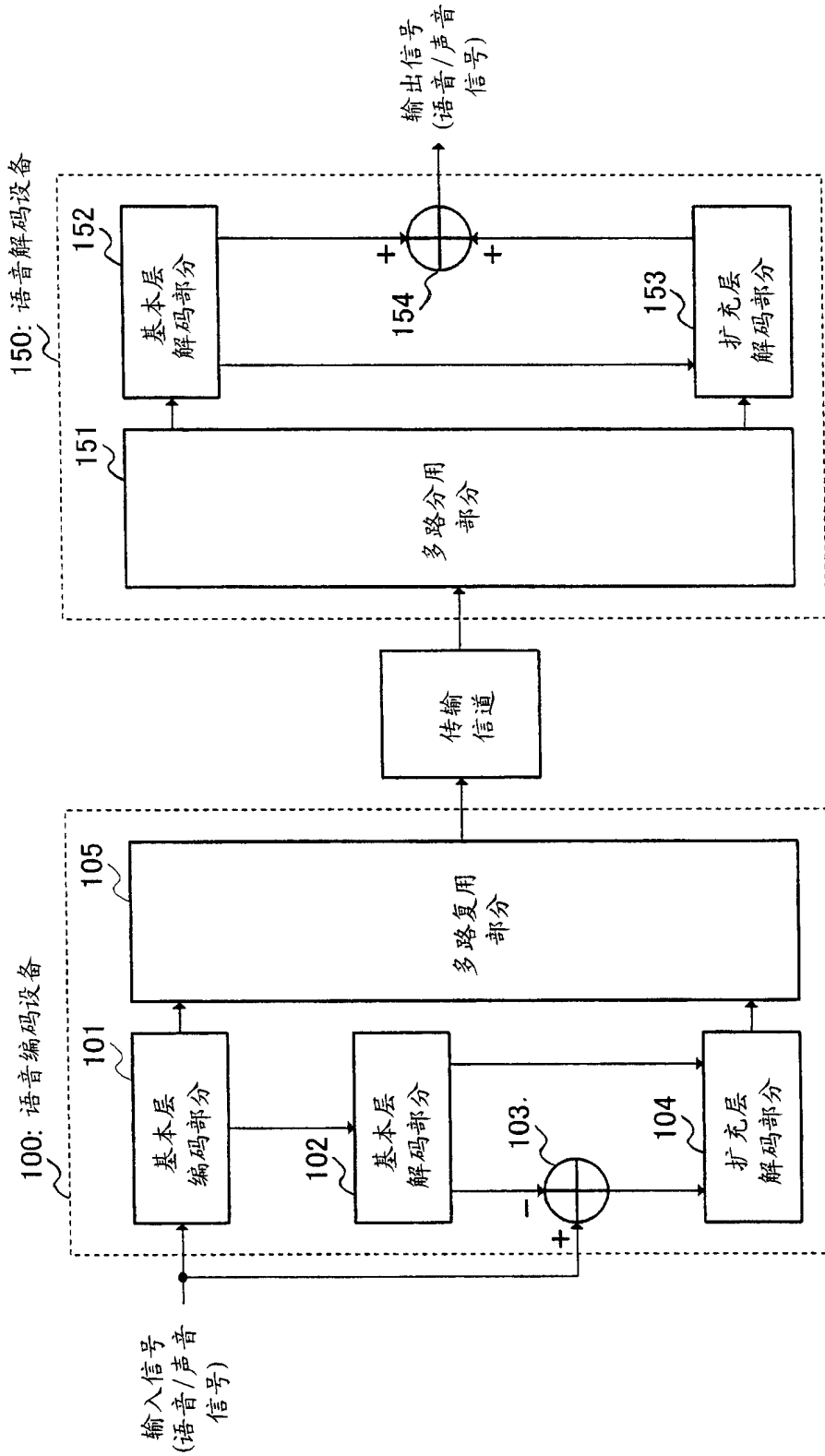


图 1

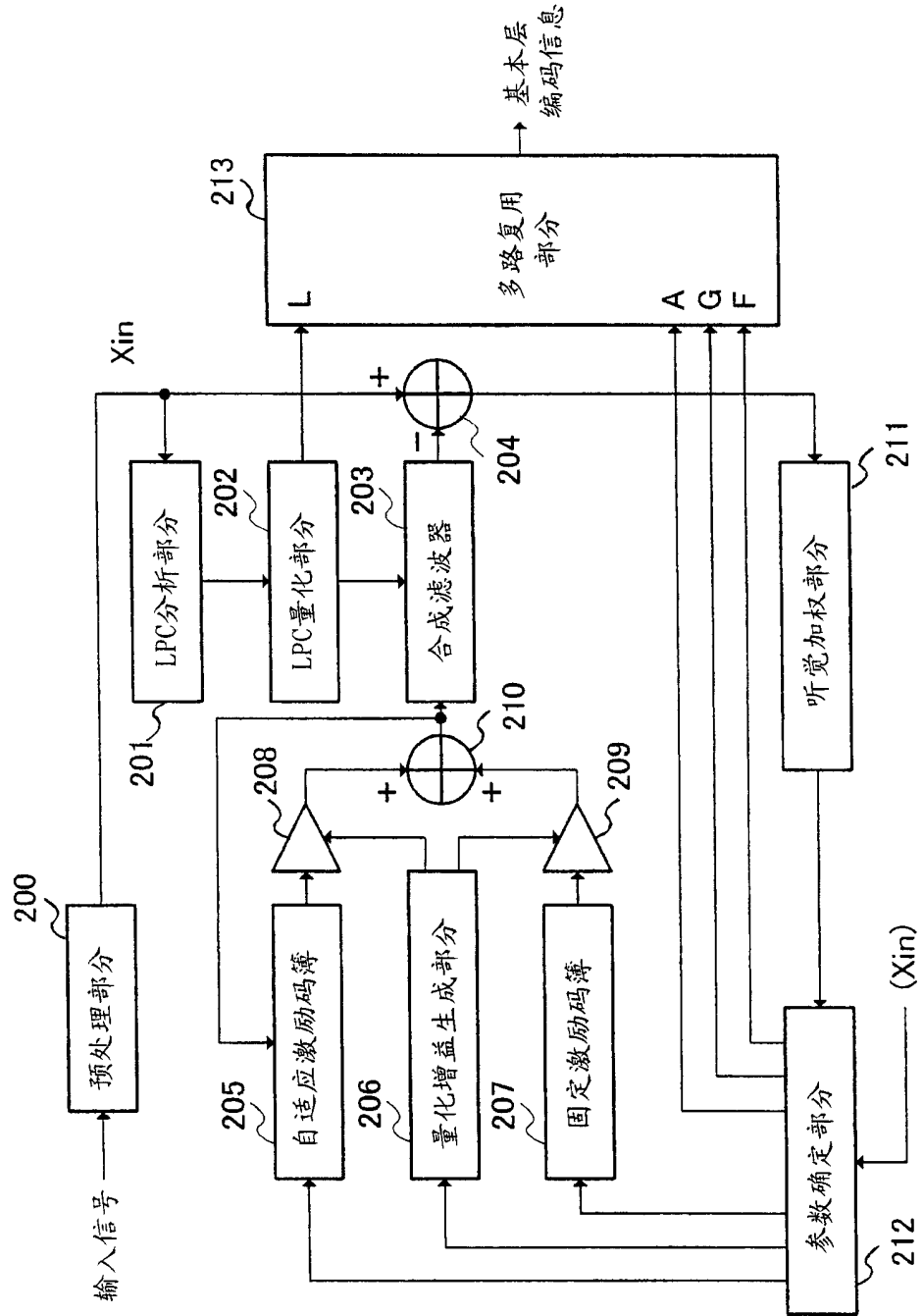


图 2

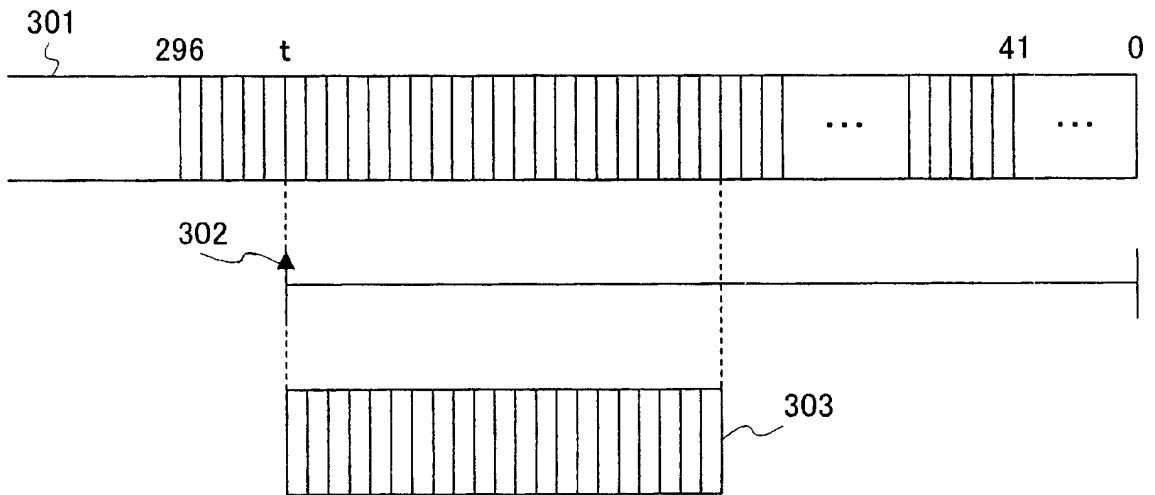


图 3

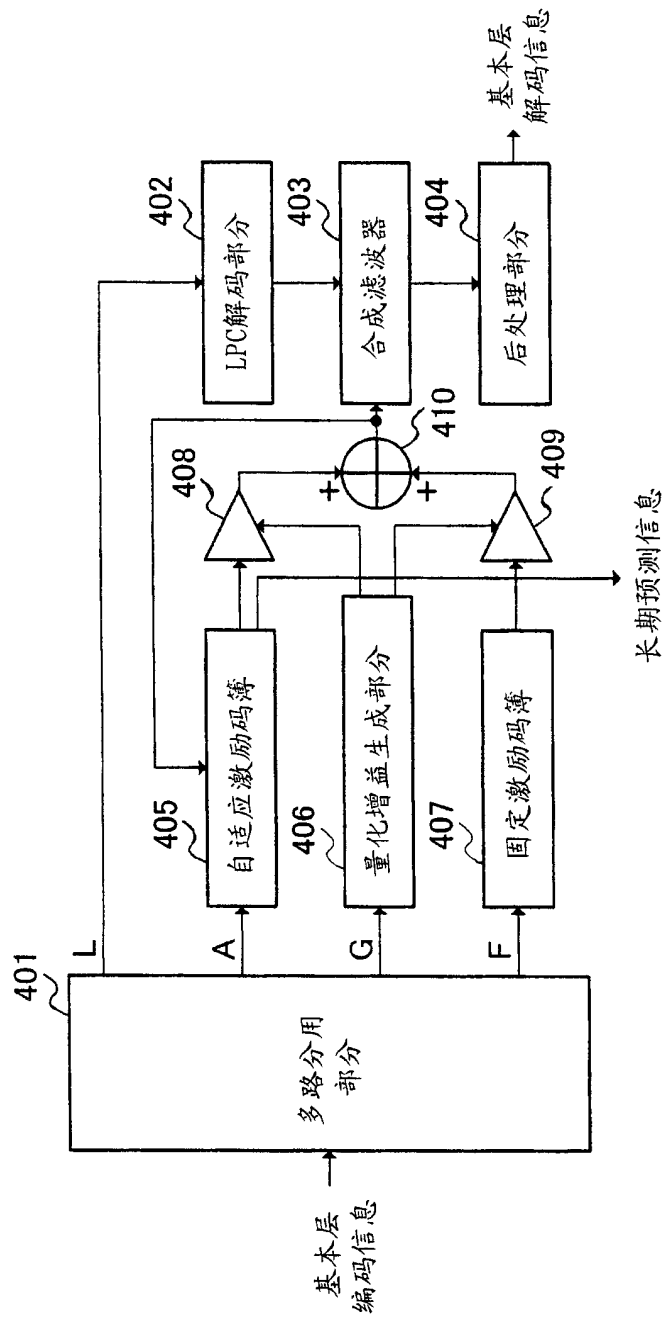


图 4



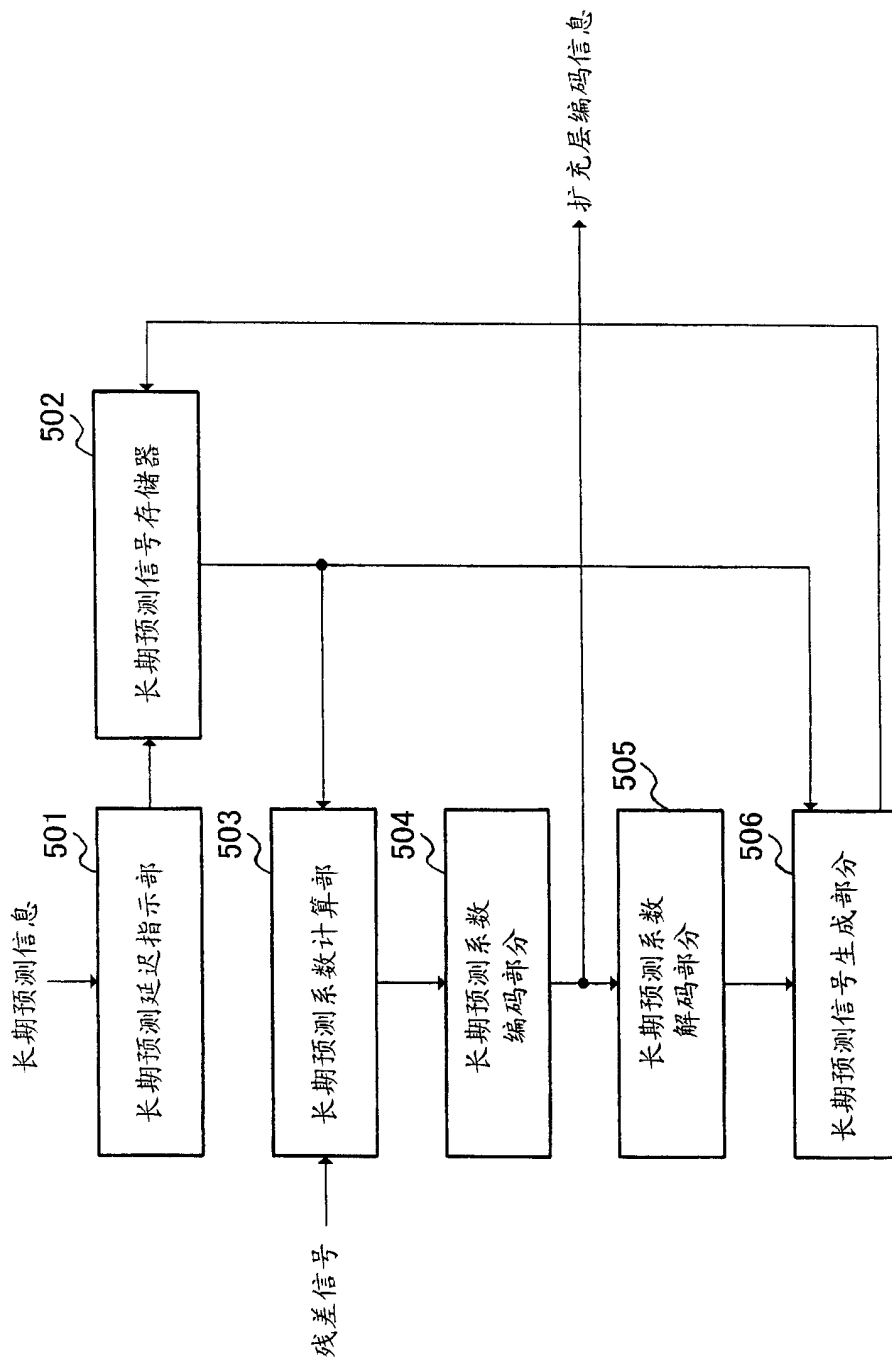


图 5

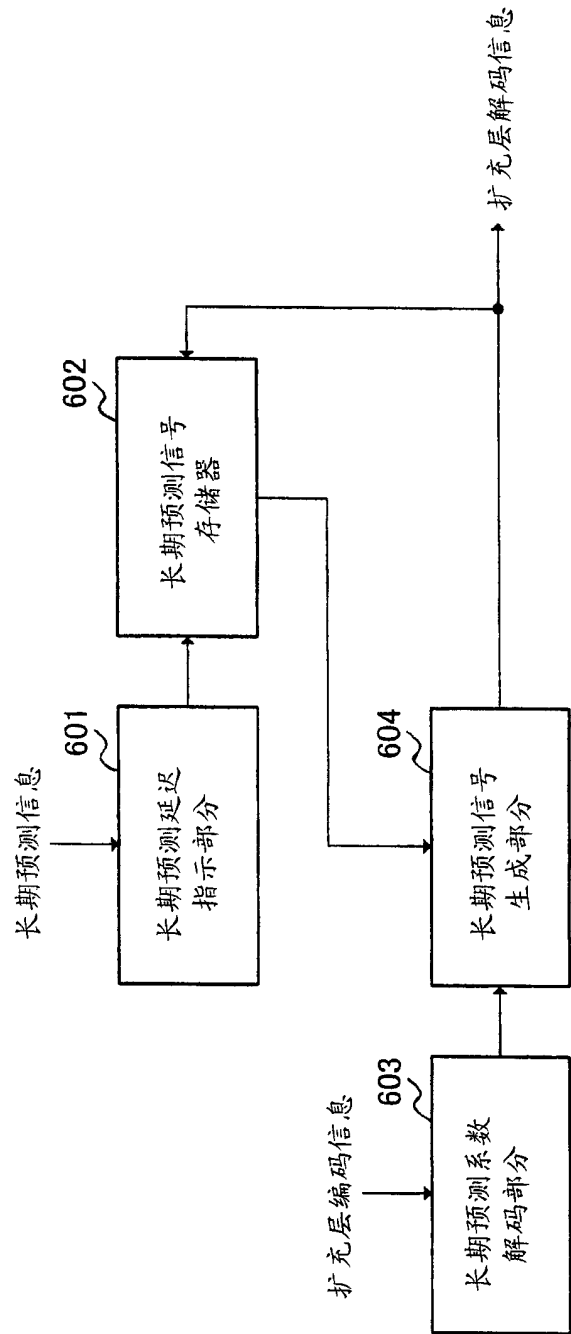


图 6

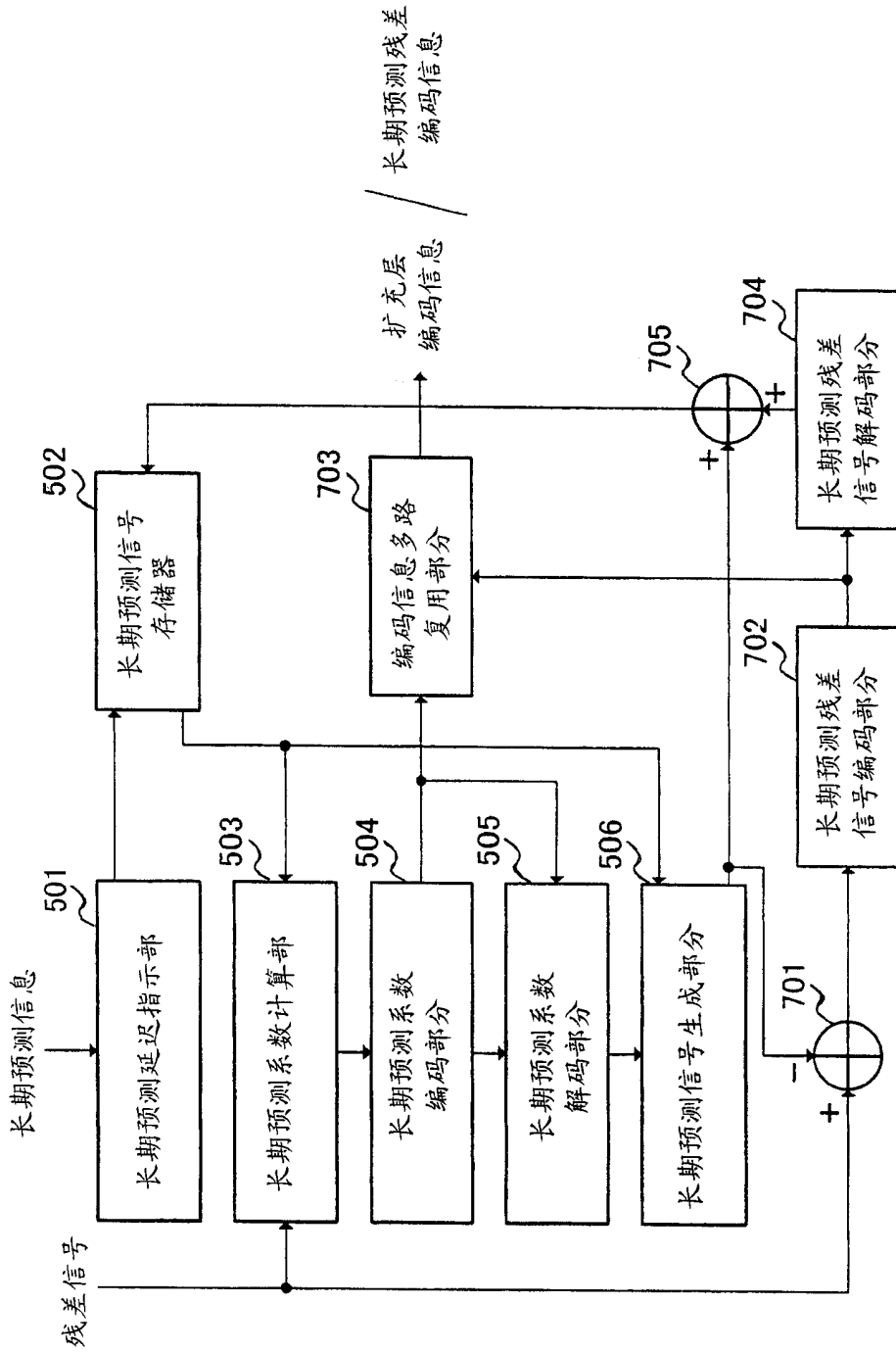


图 7

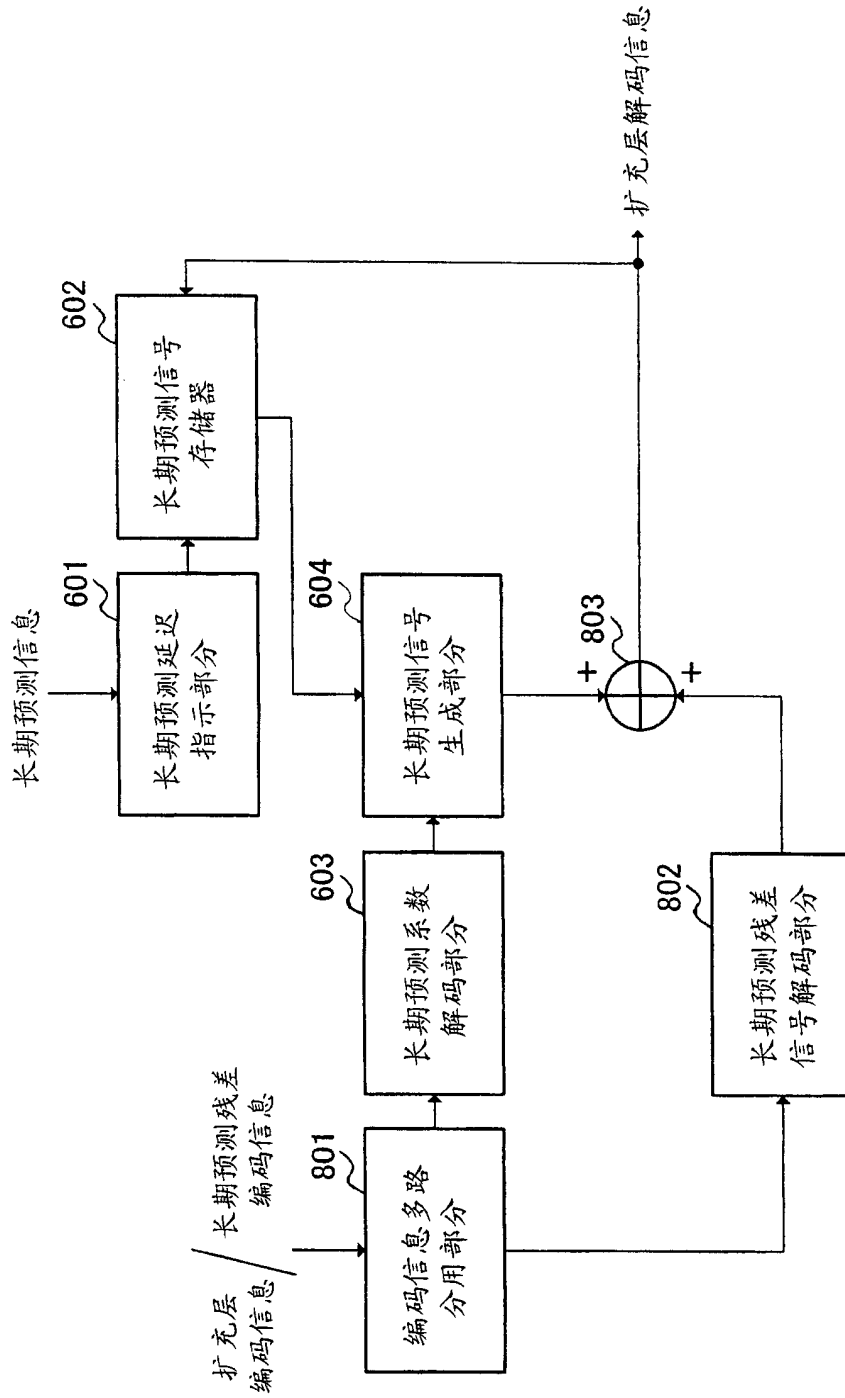


图 8

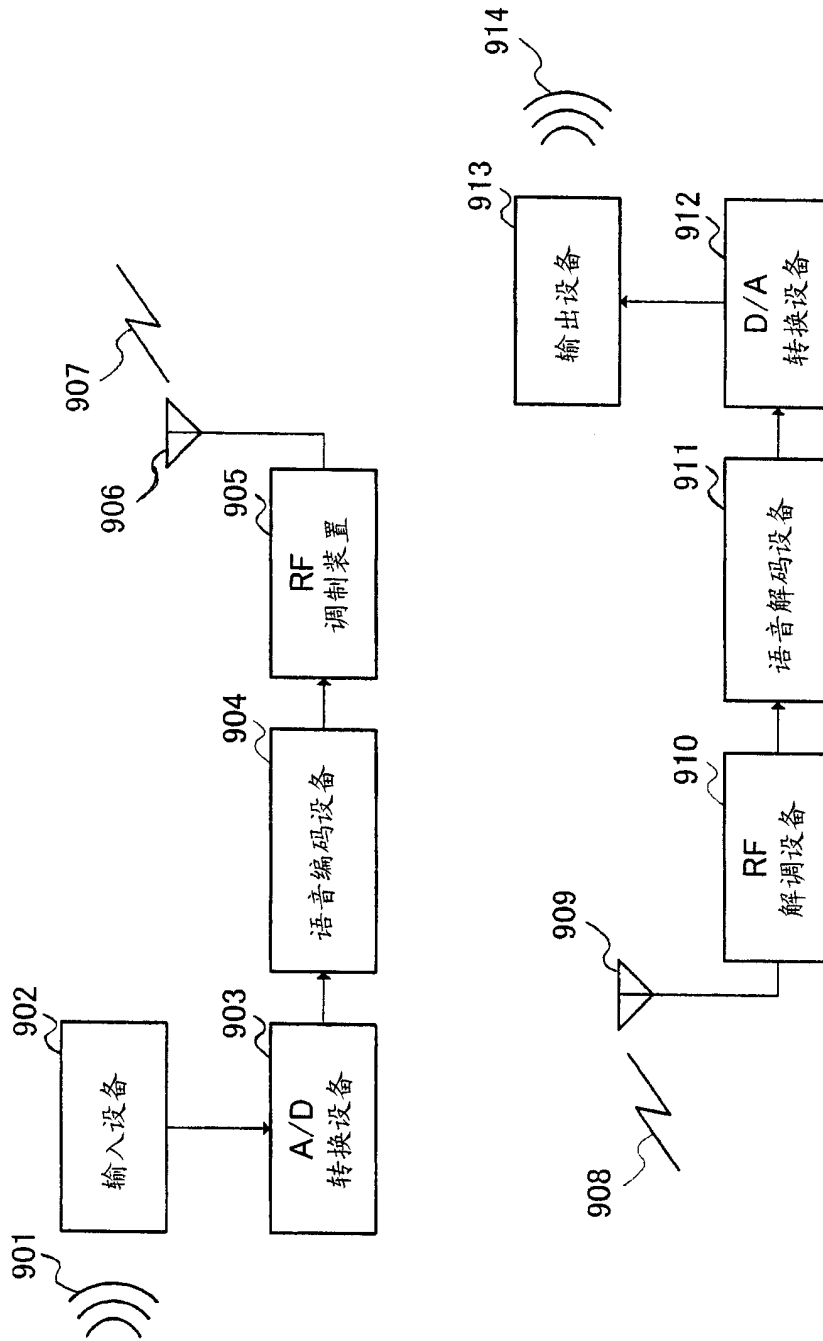


图 9