



[12] 发明专利说明书

专利号 ZL 03809372.3

[45] 授权公告日 2007 年 10 月 31 日

[11] 授权公告号 CN 100346392C

[22] 申请日 2003.4.28 [21] 申请号 03809372.3

[30] 优先权

[32] 2002. 4. 26 [33] JP [31] 127541/2002

[32] 2002. 9. 12 [33] JP [31] 267436/2002

[86] 国际申请 PCT/JP2003/005419 2003. 4. 28

[87] 国际公布 WO2003/091989 日 2003. 11. 6

[85] 进入国家阶段日期 2004. 10. 26

[73] 专利权人 松下电器产业株式会社

地址 日本大阪府

[72] 发明人 押切正浩

[56] 参考文献

US - 6208957B1 2001. 3. 27

JP - 8 - 46517A 1996. 2. 16

EP - 1173028A2 2002. 1. 16

JP - 11 - 251917A 1999. 9. 17

CN - 1247415A 2000. 3. 15

审查员 臧自欣

[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

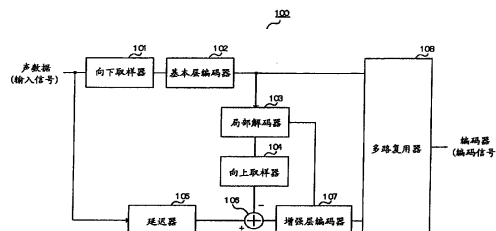
权利要求书 3 页 说明书 46 页 附图 39 页

[54] 发明名称

编码设备、解码设备、编码方法和解码方法

[57] 摘要

向下取样器(101)将输入信号的取样速率从取样速率 FH 降低到取样速率 FL。基本层编码器(102)编码取样速率 FL 的声信号。局部解码器(103)解码从基本层编码器(102)输出的编码信息。向上取样器(104)将解码信号的取样速率升高到 FH。减法器(106)从取样速率 FH 的声信号中减去解码信号。增强层编码器(107)利用从局部解码器(103)输出的解码所得参数编码从减法器(106)输出的信号。



1. 一种声音编码设备, 包括:

第一编码部分, 对声信号进行加权以使所述声信号的频谱包络掩蔽量化失真谱后进行编码来获取第一编码信息;

解码部分, 对所述第一编码信息进行解码来获取解码信号;

确定部分, 对从所述解码信号获取的解码谱计算听觉掩蔽阈值, 并通过对所述解码谱进行标度调整和平滑化, 来从所述解码谱生成与所述听觉掩蔽阈值进行比较的误差谱, 并且确定出所述误差谱中幅度显示为大于或等于该听觉掩蔽阈值的频域;

减法部分, 获取所述声信号和所述解码信号之间的残差信号; 以及

第二编码部分, 对所述残差信号中的由所述确定部分确定出的频域进行编码, 获取第二编码信息。

2. 如权利要求 1 所述的的声音编码设备, 其中,

所述第一编码部分对所述声信号的低频域进行编码;

所述第二编码部分对于所述残差信号中的低频域, 对由所述确定部分确定出的频域进行编码, 对于所述残差信号中的高频域, 则对预先决定的频域进行编码。

3. 如权利要求 1 所述的的声音编码设备, 其中, 所述第二编码部分求出各频率中的所述听觉掩蔽阈值和误差谱幅度之间的差, 并基于该差来决定编码位的分配。

4. 如权利要求 1 所述的的声音编码设备, 其中, 所述确定部分使所述听觉掩蔽阈值平滑化, 并且确定误差谱幅度显示为大于或等于平滑化的听觉掩蔽阈值的频域。

5. 如权利要求 1 所述的的声音编码设备, 其中,

所述第一编码部分根据 CELP 方式进行编码;

所述第二编码部分根据 MDCT 方式进行编码。

6. 一种声音解码设备, 包括:

第一解码部分, 对由权利要求 1 所述的的声音编码设备获取的第一编码信息进行解码来获取第一解码信号;

确定部分, 对从所述第一解码信号获取的解码谱计算听觉掩蔽阈值, 并

通过对所述解码谱进行标度调整和平滑化，来从所述解码谱生成与所述听觉掩蔽阈值进行比较的误差谱，并且确定所述误差谱中幅度显示为大于或等于该听觉掩蔽阈值的频域；

第二解码部分，对第二编码信息中的由所述确定部分确定出的频域进行解码，获取第二解码信号；以及

加法部分，将所述第一解码信号与所述第二解码信号相加来获取声信号。

7. 如权利要求 6 所述的声音解码设备，其中，

所述第一解码部分对所述第一编码信息进行解码来获取低频域的第一解码信号；

所述第二解码部分对于所述第二编码信息中的低频域，对由所述确定部分确定出的频域进行解码，对于所述第二编码信息中的高频域，则对预先决定的频域进行解码。

8. 如权利要求 6 所述的声音解码设备，其中，所述确定部分使所述听觉掩蔽阈值平滑化，并且确定误差谱幅度显示为大于或等于平滑化的听觉掩蔽阈值的频域。

9. 如权利要求 6 所述的声音解码设备，其中，

所述第一解码部分根据 CELP 方式进行解码；

所述第二解码部分根据 IMDCT 方式进行解码。

10. 一种通信终端设备，包括如权利要求 1 所述的声音编码设备或如权利要求 6 所述的声音解码设备。

11. 一种基站设备，包括如权利要求 1 所述的声音编码设备或如权利要求 6 所述的声音解码设备。

12. 一种声音编码方法，包括：

第一编码步骤，对声信号进行加权以使所述声信号的频谱包络掩蔽量化失真谱后进行编码来获取第一编码信息；

解码步骤，对所述第一编码信息进行解码来获取解码信号；

确定步骤，对从所述解码信号获取的解码谱计算听觉掩蔽阈值，并通过所述解码谱进行标度调整和平滑化，来从所述解码谱生成与所述听觉掩蔽阈值进行比较的误差谱，并且确定所述误差谱中幅度显示为大于或等于该听觉掩蔽阈值的频域；

减法步骤，获取所述声信号和所述解码信号之间的残差信号；以及

第二编码步骤，对所述残差信号中的由所述确定步骤确定出的频域进行编码，获取第二编码信息。

13. 一种声音解码方法，包括：

第一解码步骤，对由权利要求 12 所述的语音编码方法获取的第一编码信息进行解码来获取第一解码信号；

确定步骤，对从所述第一解码信号获取的解码谱计算听觉掩蔽阈值，并通过对所述解码谱进行标度调整和平滑化，来从所述解码谱生成与所述听觉掩蔽阈值进行比较的误差谱，并且确定所述误差谱中幅度显示为大于或等于该听觉掩蔽阈值的频域；

第二解码步骤，对第二编码信息中的由所述确定步骤确定出的频域进行解码，获取第二解码信号；以及

加法步骤，将所述第一解码信号与所述第二解码信号相加来获取声信号。

编码设备、解码设备、编码方法和解码方法

技术领域

本发明涉及对诸如音频信号或语音信号之类的声信号进行高效压缩编码的编码设备、解码设备、编码方法和解码方法，尤其涉及即使根据一部分编码信息也能够解码音频或语音的、适合于可伸缩（scalar）编码和解码的编码设备、解码设备、编码方法和解码方法。

背景技术

以低位速率压缩音频信号或语音信号的声音编码技术对于有效利用移动通信中的无线电和记录媒体是非常重要的。编码语音信号的语音编码方法包括由 ITU（国际电信联盟）标准化的 G726 和 G729。这些方法编码窄带信号（300 Hz - 3.4 kHz），并且能够以 8 kb/s（千位每秒）到 32 kb/s 的位速率进行高质编码。

宽带（50 Hz - 7 kHz）的标准编码包括 ITU 的 G722 和 G722.1 和 GPP（第三代伙伴项目）的 AMR-WB。这些方法能够以 6.6 kb/s 到 64 kb/s 的位速率高质编码宽带语音信号。

以低位速率对语音信号高效编码的有效方法是 CELP（码激励线性预测）。CELP 是根据通过工程技术模仿人类语音生成模型的模型进行编码的方法。具体地说，在 CELP 中，让由随机值组成的激励信号经过与周期性的强度相对应的音调滤波器和与声道特性相对应的合成滤波器，并且确定编码参数，以便在听觉特性加权下使输出信号和输入信号之间的平方误差达到最小。

在许多最新标准语音编码方法中，都是根据 CELP 进行编码。例如，G729 能够以 8 kb/s 进行窄带信号编码，和 AMR-WB 能够以 6.6 kb/s 到 23.85 kb/s 进行窄带信号编码。

同时，在编码音频信号的音频编码的情况下，共同使用将音频信号转换到频域和利用听觉心理声学模型进行编码的方法，譬如，由 MPEG（运动图像专家组）标准化的 Layer III 方法和 AAC 方法。众所周知，利用这些方法，对于 44.1 kHz 取样速率的信号，在 64 kb/s 到 96 kb/s 每信道上几乎不会变

差。

这种音频编码是对音乐进行高质编码的方法。音频编码也可以对如上所述，在背景中存在音乐或环境声音的语音信号进行高质编码，并且可以管理具有 CD 质量的、大约 22 kHz 的信号频带。

但是，当利用语音编码方法对语音信号占优势和在背景中叠加了音乐或环境声音的信号进行编码时，存在如下问题，由于背景音乐或环境声音，不仅背景信号变差了，而且语音信号也变差了，因此，总质量下降了。

出现这个问题是因为语音编码方法基于专用于 CELP 语音模型的方法。问题在于，语音编码方法只能管理直到 7 kHz 的信号频带，和对于复合信号，不能充分地管理作为更高频带中的成分的信号。

此外，对于音频编码方法，为了取得高质编码，必须使用高位速率。对于音频编码方法，如果应该利用下至 32 kb/s 的位速率进行编码，那么，存在解码信号质量大幅下降的问题。因此，存在问题不能在传输速率低的通信网络上使用的问题。

发明内容

本发明的目的是提供一种甚至在低位速率下也能够对语音信号占优势和在背景中叠加了音乐或环境声音的信号进行高质编码和解码的编码设备、解码设备、编码方法和解码方法。

这个目的是通过拥有两个层，即基本层和增强层，根据基本层中的 CELP，以低位速率对输入信号窄带或宽带频区进行高质编码，和在不能在基本层中得到表示的背景音乐或环境声音，以及存在比基本层覆盖的频区高的频率成分的信号的增强层中进行编码达到的。

根据本发明的一个方面，提供一种声音编码设备，包括：第一编码部分，对声信号进行加权以使所述声信号的频谱包络掩蔽量化失真谱后进行编码来获取第一编码信息；解码部分，对所述第一编码信息进行解码来获取解码信号；确定部分，计算对从所述解码信号获取的解码谱计算听觉掩蔽阈值，并通过对所述解码谱进行标度调整和平滑化，来从所述解码谱生成与所述听觉掩蔽阈值进行比较的误差谱，并且确定出确定所述解码信号误差谱中幅度显示为大于或等于该听觉掩蔽阈值的频域；减法部分，获取所述声信号和所述解码信号之间的残差信号；以及第二编码部分，对所述残差信号中的由所述

确定部分确定出的频域进行编码，获取第二编码信息。

根据本发明的一个方面，提供一种声音解码设备，包括：第一解码部分，对由上述声音编码设备获取的第一编码信息进行解码来获取第一解码信号；确定部分，对从所述第一解码信号获取的解码谱计算听觉掩蔽阈值，并通过对所述解码谱进行标度调整和平滑化，来从所述解码谱生成与所述听觉掩蔽阈值进行比较的误差谱，并且确定所述误差谱中幅度显示为大于或等于该听觉掩蔽阈值的频域；第二解码部分，对第二编码信息中的由所述确定部分确定出的频域进行解码，获取第二解码信号；以及加法部分，将所述第一解码信号与所述第二解码信号相加来获取声信号。

根据本发明的一个方面，提供一种通信终端设备，包括如上所述的声音编码设备或如上所述的声音解码设备。

根据本发明的一个方面，提供一种基站设备，包括如上所述的声音编码设备或如上所述的声音解码设备。

根据本发明的一个方面，提供一种声音编码方法，包括：第一编码步骤，对声信号进行加权以使所述声信号的频谱包络掩蔽量化失真谱后进行编码来获取第一编码信息；解码步骤，对所述第一编码信息进行解码来获取解码信号；确定步骤，对从所述解码信号获取的解码谱计算听觉掩蔽阈值，并通过对所述解码谱进行标度调整和平滑化，来从所述解码谱生成与所述听觉掩蔽阈值进行比较的误差谱，并且确定所述误差谱中幅度显示为大于或等于该听觉掩蔽阈值的频域；减法步骤，获取所述声信号和所述解码信号之间的残差信号；以及第二编码步骤，对所述残差信号中的由所述确定步骤确定出的频域进行编码，获取第二编码信息。

根据本发明的一个方面，提供一种声音解码方法，包括：第一解码步骤，对由上述声音编码方法获取的第一编码信息进行解码来获取第一解码信号；确定步骤，对从所述第一解码信号获取的解码谱计算听觉掩蔽阈值，并通过对所述解码谱进行标度调整和平滑化，来从所述解码谱生成与所述听觉掩蔽阈值进行比较的误差谱，并且确定所述误差谱中幅度显示为大于或等于该听觉掩蔽阈值的频域；第二解码步骤，对第二编码信息中的由所述确定步骤确定出的频域进行解码，获取第二解码信号；以及加法步骤，将所述第一解码信号与所述第二解码信号相加来获取声信号。

附图说明

图 1 是示出根据本发明第 1 实施例的信号处理设备的配置的方块图;

图 2 是示出输入信号成分的例子图形;

图 3 是示出根据上面实施例的信号处理设备的信号处理方法的例子图形;

图 4 是示出基本层编码器的配置的例子图形;

图 5 是示出增强层编码器的配置的例子图形;

图 6 是示出增强层编码器的配置的例子图形；

图 7 是示出增强层中的 LPC 系数计算的例子图形；

图 8 是示出根据本发明第 3 实施例的信号处理设备的增强层编码器的配置的方块图；

图 9 是示出根据本发明第 4 实施例的信号处理设备的增强层编码器的配置的方块图；

图 10 是示出根据本发明第 5 实施例的信号处理设备的配置的方块图；

图 11 是示出基本层解码器的例子方块图；

图 12 是示出增强层解码器的例子方块图；

图 13 是示出增强层解码器的例子图形；

图 14 是示出根据本发明第 7 实施例的信号处理设备的增强层解码器的配置的方块图；

图 15 是示出根据本发明第 8 实施例的信号处理设备的增强层解码器的配置的方块图；

图 16 是示出根据本发明第 9 实施例的声音编码设备的配置的方块图；

图 17 是示出声信号信息分布的例子图形；

图 18 是示出在基本层和增强层中经受编码的区域的例子图形；

图 19 是示出声（音乐）信号谱的例子图形；

图 20 是示出上面实施例的声音编码设备的频率确定部分的内部配置的例子方块图；

图 21 是示出上面实施例的声音编码设备的听觉掩蔽计算器的内部配置的例子图形；

图 22 是示出上面实施例的增强层编码器的内部配置的例子方块图；

图 23 是示出上面实施例的听觉掩蔽计算器的内部配置的例子方块图；

图 24 是示出根据本发明第 9 实施例的声音解码设备的配置的方块图；

图 25 是示出上面实施例的声音解码设备的增强层解码器的内部配置的例子方块图；

图 26 是示出根据本发明第 10 实施例的基本层编码器的内部配置的例子方块图；

图 27 是示出上面实施例的基本层解码器的内部配置的例子方块图；

图 28 是示出上面实施例的基本层解码器的内部配置的例子方块图；

图 29 是示出根据本发明第 11 实施例的声音编码设备的频率确定部分的内部配置的例子方块图；

图 30 是示出上面实施例的估计误差谱计算器计算的残留误差谱的例子图形；

图 31 是示出根据本发明第 12 实施例的声音编码设备的频率确定部分的内部配置的例子方块图；

图 32 是示出上面实施例的声音编码设备的频率确定部分的内部配置的例子方块图；

图 33 是示出根据本发明第 13 实施例的声音编码设备的增强层编码器的内部配置的例子方块图；

图 34 是示出上面实施例的定序部分排序估计失真值的例子图形；

图 35 是示出根据本发明第 13 实施例的声音解码设备的增强层解码器的内部配置的例子方块图；

图 36 是示出根据本发明第 14 实施例的声音编码设备的增强层编码器的内部配置的例子方块图；

图 37 是示出根据本发明第 14 实施例的声音解码设备的增强层解码器的内部配置的例子方块图；

图 38 是示出上面实施例的声音编码设备的频率确定部分的内部配置的例子方块图；

图 39 是示出根据本发明第 14 实施例的声音解码设备的增强层解码器的内部配置的例子方块图；

图 40 是示出根据本发明第 15 实施例的通信设备的配置的方块图；

图 41 是示出根据本发明第 16 实施例的通信设备的配置的方块图；

图 42 是示出根据本发明第 17 实施例的通信设备的配置的方块图；和

图 43 是示出根据本发明第 18 实施例的通信设备的配置的方块图。

具体实施方式

基本上，本发明拥有两个层，即基本层和增强层，根据基本层中的 CELP，以低位速率对输入信号窄带或宽带频区进行高质编码，然后，不能在基本层中得到表示的背景音乐或环境声音，以及存在比基本层覆盖的频区高的频率成分的信号的增强层中进行编码，增强层具有如同使用音频编码方法一样，

能够使所有信号得到管理的配置。

通过这种手段，可以对不能在基本层中得到表示的背景音乐或环境声音，以及存在比基本层覆盖的频区高的频率成分的信号进行高效编码。本发明的特性是，此时，利用通过基本层编码信息获得的信息进行增强层编码。通过这种手段，获得了能够减少增强层编码位的个数的效果。

现在参照附图详细描述本发明的实施例。

(第1实施例)

图1是示出根据本发明第1实施例的信号处理设备的配置的方块图。图1中的信号处理设备100主要包括向下取样器(down-sampler)101、基本层编码器102、局部解码器103、向上取样器(up-sampler)104、延迟器105、减法器106、增强层编码器107和多路复用器108。

向下取样器101从取样速率FH到取样速率FL向下取样(down-sample)输入信号取样速率，并且将取样速率FL的声信号输出到基本层编码器102。这里，取样速率FL是比取样速率FH低的频率。

基本层编码器102编码取样速率FL的声信号，并且将编码信息输出到局部解码器103和多路复用器108。

局部解码器103解码从基本层编码器102输出的编码信息，将解码信号输出到向上取样器104，并且将从解码结果中获得的参数输出到增强层编码器107。

向上取样器104将解码信号取样速率升高到FH，并且将结果输出到减法器106。

延迟器105将输入取样速率FH的声信号延迟预定时间，然后，将信号输出到减法器106。通过使这个延迟时间等于在向下取样器101、基本层编码器102、局部解码器103和向上取样器104中产生的时间延迟，可以防止在接着的相减处理中出现相移。

减法器106从取样速率FH的声信号中减去解码信号，并且将相减结果输出到增强层编码器107。

增强层编码器107利用从局部解码器103输出的解码结果参数解码从减法器106输出的信号，并且将所得结果输出到多路复用器108。多路复用器108多路复用和输出由基本层编码器102和增强层编码器107编码的信号。

现在说明基本层编码和增强层编码。图2是示出输入信号成分的例子

图形。在图 2 中，垂直轴表示信号成分信息量，而水平轴表示频率。图 2 示出了给出包含在输入信号中的语音信息和背景音乐/背景噪声信息的频带。

在语音信息的情况下，在低频区中存在大量信息，信息量随着频区增高而减少。相反，在背景音乐和背景噪声信息的情况下，与语音信息相比，在较低区域中存在相对少的信息，和大量信息处在较高区域中。

因此，本发明的信号处理设备使用数种编码方法，并且对各自编码方法适合的每个区域进行不同编码。

图 3 是示出根据本实施例的信号处理设备的信号处理方法的例子图形。在图 3 中，垂直轴表示信号成分信息量，而水平轴表示频率。

基本层编码器 102 被设计成有效表示从 0 到 FL 的频带中的语音信息，并且可以对该区域中的语音信息进行高质编码。但是，从 0 到 FL 的频带中背景音乐和背景噪声信息的编码质量不高。增强层编码器 107 编码基本层编码器 102 不能编码的部分和从 FL 到 FH 的频带中的信号。

因此，通过组合基本层编码器 102 和增强层编码器 107，可以在宽带中实现高质编码。此外，可以实现即使只利用至少基本层编码部分的编码信息也可以解码语音信息的可伸缩功能。

这样，局部解码器 103 中出自通过编码生成的那些参数当中的有用参数被供应给增强层编码器 107，和增强层编码器 107 利用这个参数进行编码。

由于这个参数是从编码信息中生成的，当解码本实施例的信号处理设备编码的信号时，在声音解码过程中可以获得相同参数，没有必要附加这个传输到解码方的参数。其结果是，增强层编码部分可以实现不会招致附加信息增加的有效编码处理。

例如，存在于局部解码器 103 解码的参数当中，指示输入信号是诸如元音之类具有明显周期性的信号还是诸如辅音之类具有明显噪声特性的信号的有声/无声标志用作增强层编码器 107 应用的参数。可以利用有声/无声标志进行调整，譬如，进行在有声部分中的增强层中强调较低区域多于较高区域的位分配，和进行在无声部分中强调较高区域多于较低区域的位分配。

因此，根据本实施例的信号处理设备，通过从输入信号中提取不超过预定频率的成分和进行适当于语音编码的编码，和利用解码所得编码信息的结果进行适合于音频编码的编码，可以以低位速率进行高质编码。

关于取样速率 FH 和 FL，只需要取样速率 FH 比取样速率 FL 高，并且对

这些值没有限制。例如，可以利用 $FH = 24 \text{ kHz}$ 和 $FL = 16 \text{ kHz}$ 的取样速率进行编码。

(第2实施例)

在本实施例中，描述在第1实施例的局部解码器103解码的参数当中，指示输入信号谱的LPC系数用作增强层编码器107利用的参数的例子。

本实施例的信号处理设备利用图1中的基本层编码器102中的CELP进行编码，并且在增强层编码器107中利用指示输入信号谱的LPC系数进行编码。

首先给出基本层编码器102的操作的详细描述，后面接着增强层编码器107的基本配置的描述。这里提到的“基本配置”旨在简化随后实施例的描述，和表示不使用局部解码器103编码参数的配置。此后，给出使用局部解码器103解码的LPC系数的增强层编码器107的描述，这是本实施例的特征。

图4是示出基本层编码器102的配置的例子的图形。基本层编码器102主要包括LPC分析器401、加权部分402、自适应码簿搜索单元403、自适应增益量化器404、目标矢量发生器405、噪声码簿搜索单元406、噪声增益量化器407和多路复用器408。

LPC分析器401从向下取样器101以取样速率 FL 取样的输入信号中获取LPC系数，并且将这些LPC系数输出到加权部分402。

加权部分402根据LPC分析器401获取的LPC系数，对输入信号进行加权，并且将加权输入信号输出到自适应码簿搜索单元403、自适应增益量化器404和目标矢量发生器405。

自适应码簿搜索单元403利用作为目标信号的加权输入信号进行自适应码簿搜索，并且将检索的自适应矢量输出到自适应增益量化器404和目标矢量发生器405。然后，自适应码簿搜索单元403将确定为存在最小量化失真的自适应矢量的代码输出到多路复用器408。

自适应增益量化器404量化乘以从自适应码簿搜索单元403输出的自适应矢量的自适应增益，并且将结果输出到目标矢量发生器405。然后，将这个代码输出到多路复用器408。

目标矢量发生器405对将自适应矢量乘以自适应增益的结果与从加权部分402输入的输入信号进行矢量相减，并且将相减结果作为目标矢量输出到噪声码簿搜索单元406和噪声增益量化器407。

噪声码簿搜索单元406从噪声码簿中检索与从目标矢量发生器405输出

的目标矢量相关的失真最小的噪声矢量。然后，噪声码簿搜索单元 406 将检索的噪声矢量输出到噪声增益量化器 407，并且还将那个代码输出到多路复用器 408。

噪声增益量化器 407 乘以噪声码簿搜索单元 406 检索的噪声矢量的噪声增益，并且将那个代码输出到多路复用器 408。

多路复用器 408 多路复用 LPC 系数、自适应矢量、自适应增益、噪声矢量和噪声增益编码信息，并且将所得信号输出到局部解码器 103 和多路复用器 108。

接着，描述图 4 中基本层编码器 102 的操作。首先，输入从向下取样器 101 输出的取样速率 FL，和 LPC 分析器 401 获取 LPC 系数。将 LPC 系数转换成诸如 LSP 系数之类适合量化的参数，并且量化它们。将通过这种量化获得的编码信息供应给多路复用器 408，并且，从编码信息中计算出量化 LSP 系数和将其转换成 LPC 系数。

通过这种量化，获得量化 LPC 系数。利用量化 LPC 系数、自适应码簿、自适应增益、噪声码簿和噪声增益进行编码。

然后，加权部分 402 根据 LPC 分析器 401 获得的 LPC 系数，对输入信号进行加权。这种加权的目的是进行谱成形，以便通过输入信号的频谱包络掩蔽量化失真谱。

然后，自适应码簿搜索单元 403 利用作为目标信号的加权输入信号搜索自适应码簿。以音调周期为基础重复旧激励序列的信号被称为自适应矢量，和自适应码簿由在预定范围的音调周期上生成的自适应矢量组成。

如果将加权输入信号指定为 $t(n)$ ，将包括 LPC 系数的加权合成滤波器的脉冲响应被卷积成音调周期为 i 的自适应矢量的信号指定为 $p_i(n)$ ，那么，将使如下方程 (1) 的估算函数 D 达到极小的自适应矢量的音调周期 i 作为参数发送到多路复用器 408。

$$D = \sum_{n=0}^{N-1} t^2(n) - \frac{\left(\sum_{n=0}^{N-1} t(n) p_i(n) \right)^2}{\sum_{n=0}^{N-1} p_i^2(n)} \quad \dots (1)$$

这里， N 表示矢量长度。

接着，自适应增益量化器 404 进行乘以自适应矢量的自适应增益的量化。自适应增益 β 用方程 (2) 表示。这个 β 值经受标量量化 (scalar quantization)，

并且所得代码被发送到多路复用器 408。

$$\beta = \frac{\sum_{n=0}^{N-1} t(n) p_i(n)}{\sum_{n=0}^{N-1} p_i^2(n)} \quad \dots (2)$$

然后，目标矢量发生器 405 从输入信号中减去自适应矢量产生的效果，生成噪声码簿搜索单元 406 和噪声增益量化器 407 使用的目标矢量。如果这里的 $p_i(n)$ 表示当方程 (1) 所表示的估算函数 D 达到极小时，合成滤波器的脉冲响应被卷积成自适应矢量的信号，和 βq 表示当方程 (2) 所表示的自适应矢量 β 经受标量量化时的量化值，那么，目标矢量 $t_2(n)$ 由如下方程 (2) 表示。

$$t_2(n) = t(n) - \beta q \cdot p_i(n) \quad \dots (3)$$

将前述目标矢量 $t_2(n)$ 和 LPC 系数供应给噪声码簿搜索单元 406，进行噪声码簿搜索。

这里，提供给噪声码簿搜索单元 406 的噪声码簿的典型成分是代数。在代数码簿中，幅度为 1 的脉冲由只具有预定极少数的矢量表示。此外，对于代数码簿，事先决定可以为每个相位保留的位置，以便不重叠。因此，代数码簿的特征是，通过小量计算就可以确定脉冲位置和脉冲代码（极性）的最佳组合。

如果将目标矢量指定为 $t_2(n)$ ，将加权合成滤波器的脉冲响应被卷积成与代码 j 相对应的噪声矢量的信号指定为 $c_j(n)$ ，那么，将使如下方程 (4) 的估算函数 D 达到极小的噪声矢量的指标 j 作为参数发送到多路复用器 408。

$$D = \sum_{n=0}^{N-1} t_2^2(n) - \frac{\left(\sum_{n=0}^{N-1} t_2(n) c_j(n) \right)^2}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (4)$$

接着，噪声增益量化器 407 进行乘以噪声矢量的噪声增益的量化。噪声增益 γ 用方程 (5) 表示。这个 γ 值经受标量量化，并且所得代码被发送到多路复用器 408。

$$\gamma = \frac{\sum_{n=0}^{N-1} t_2(n) c_j(n)}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (5)$$

多路复用器 408 多路复用发送的 LPC 系数、自适应码簿、自适应增益、

噪声码簿和噪声增益编码信息，并且将所得信号输出到局部解码器 103 和多路复用器 108。

当存在新输入信号时，重复上面的处理。当不存在新输入信号时，终止该处理。

现在描述增强层编码器 107。图 5 是示出增强层编码器 107 的配置的例子的图形。图 5 中的增强层编码器 107 主要包括 LPC 分析器 501、谱包络计算器 502、MDCT 部分 503、功率计算器 504、功率归一化器 505、谱归一化器 506、Bark 标度归一化器 508、Bark 标度形状计算器 507、矢量量化器 509 和多路复用器 510。

LPC 分析器 501 对输入信号进行 LPC 分析。并且，LPC 分析器 501 在 LSP 或其它适合于量化的参数的值域中有效地量化 LPC 系数，LPC 分析器将编码信息输出到多路复用器，和 LPC 分析器将量化 LPC 系数输出到谱包络计算器 502。谱包络计算器 502 从量化 LPC 系数中计算谱包络，并且将这个谱包络输出到矢量量化器 509。

MDCT 部分 503 对输入信号进行 MDCT (改进离散余弦变换) 处理，并且将获得的 MDCT 系数输出到功率计算 504 和功率归一化器 505。功率计算器 504 找出和量化 MDCT 系数的功率，并且将量化功率输出到功率归一化器 505 和将编码信息输出到多路复用器 510。

功率归一化器 505 利用量化功率归一化 MDCT 系数，并且将功率归一化 MDCT 系数输出到谱归一化器 506。谱归一化器 506 利用谱包络归一化根据功率归一化的 MDCT 系数，并且将归一化 MDCT 系数输出到 Bark 标度形状计算器 507 和 Bark 标度归一化器 508。

Bark 标度形状计算器 507 通过 Bark 标度计算以等间隔频带划分的频谱的形状，然后，量化这个谱形状，并且，将量化谱形状输出到 Bark 标度归一化器 508 和矢量量化器 509。并且，Bark 标度形状计算器 507 将编码信息输出到多路复用器 510。

Bark 标度归一化器 508 利用量化 Bark 标度形状归一化归一化 MDCT 系数，将结果输出到矢量量化器 509。

矢量量化器 509 对从 Bark 标度归一化器 508 输出的归一化 MDCT 系数进行矢量量化，找出失真最小的代码矢量，并且将代码矢量的指标作为编码信息输出到多路复用器 510。

多路复用器 510 多路复用所有编码信息，并且将所得信号输出到多路复用器 108。

现在描述图 5 中增强层编码器 107 的操作。图 1 中的减法器 106 获得的相减信号经受 LPC 分析器 501 的 LPC 分析。然后，通过 LPC 分析计算出 LPC 系数。将 LPC 系数转换成此后进行量化、诸如 LSP 系数之类适合于量化的参数。将与这里获得的 LPC 系数有关的编码信息供应给多路复用器 510。

谱包络计算器 502 根据解码的 LPC 系数，按照如下的方程 (6) 计算谱包络。

$$env(m) = \left| \frac{1}{1 - \sum_{i=1}^{NP} \alpha_q(i) e^{-j \frac{2\pi m i}{M}}} \right| \quad \dots (6)$$

这里， α_q 表示解码的 LPC 系数，NP 表示 LPC 系数的次序，和 M 表示谱分辨率。通过方程 (6) 获得的谱包络 $env(m)$ 供如后所述的谱归一化器 506 和矢量量化器 509 使用。

然后，输入信号在 MDCT 部分 503 中经受 MDCT 处理，获得 MDCT 系数。MDCT 处理的特征是，由于使用了每次一半地完全叠加连续帧的分析帧，和分析帧的前一半是奇函数，而分析帧的后一半是偶函数的正交基，不会出现帧边缘失真。当进行 MDCT 处理时，将输入信号与诸如正弦函数窗口那样的窗口函数相乘。当将 MDCT 系数指定为 $X(m)$ 时，MDCT 系数按照如下方程 (7) 计算。

$$X(m) = \sqrt{\frac{1}{N}} \sum_{n=0}^{2N-1} x(n) \cos \left\{ \frac{(2n+1+N) \cdot (2m+1)\pi}{4N} \right\} \quad \dots (7)$$

这里， $x(n)$ 表示将输入信号乘以窗口函数时的信号。

接着，功率计算器 504 求出和量化 MDCT 系数 $X(m)$ 的功率。然后，功率归一化器 505 利用方程 (8) 归一化具有那个量化之后的功率的 MDCT 系数。

$$pow = \sum_{m=0}^{M-1} X(m)^2 \quad \dots (8)$$

这里，M 表示 MDCT 系数的大小。在 MDCT 系数功率 pow 被量化之后，将编码信息发送到多路复用器 510。利用编码信息解码 MDCT 系数的功率，和利用所得值，按照如下方程 (9) 归一化 MDCT 系数。

$$X1(m) = \frac{X(m)}{\sqrt{powq}} \quad \dots (9)$$

这里， $X1(m)$ 代表功率归一化之后的 MDCT 系数，和 $powq$ 表示量化之后 MDCT 系数的功率。

然后，谱归一化器 506 利用谱包络归一化已经按照功率归一化的 MDCT 系数。谱归一化器 506 按照如下的方程 (10) 进行归一化。

$$X2(m) = \frac{X1(m)}{env(m)} \quad \dots (10)$$

接着，Bark 标度形状计算器 507 通过 Bark 标度计算以等间隔频带划分的频谱的形状，然后，量化这个谱形状。Bark 标度形状计算器 507 将这个编码信息发送到多路复用器 510，并且还利用解码值，对作为来自谱归一化器 506 的输出信号的 MDCT 系数 $X2(m)$ 进行归一化。Bark 标度和 Herz 标度之间的对应关系通过如下方程 (11) 所表示的转换表达式给出。

$$B = 13 \tan^{-1}(0.76f) + 3.5 \tan^{-1}\left(\frac{f}{7.5}\right) \quad \dots (11)$$

这里， B 表示 Bark 标度和 f 表示 Herz 标度。对于在 Bark 标度上以等间隔频带划分的子频带，Bark 标度形状计算器 507 按照如下的方程 (12) 计算形状。

$$B(k) = \sum_{m=f_l(k)}^{f_h(k)} X2(m)^2 \quad 0 \leq k < K \quad \dots (12)$$

这里， $f_l(k)$ 表示第 k 子频带的最低频率和 $f_h(k)$ 表示第 k 子频带的最高频率，和 K 表示子频带的个数。

然后，Bark 标度形状计算器 507 量化每个频带的 Bark 标度形状 $B(k)$ 并将编码信息发送到多路复用器 510，并且还解码 Bark 标度形状并将结果供应给 Bark 标度归一化器 508 和矢量量化器 509。利用归一化之后的 Bark 标度形状，Bark 标度归一化器 508 按照如下方程 (13) 生成归一化 MDCT 系数 $X3(m)$ 。

$$X3(m) = \frac{X2(m)}{\sqrt{B_q(k)}} \quad f_l(k) \leq m \leq f_h(k) \quad 0 \leq k < K \quad \dots (13)$$

这里， $B_q(k)$ 表示第 k 子频带量化之后的 Bark 标度形状。

接着，矢量量化器 509 将 $X3(m)$ 划分成数个矢量和利用与每个矢量相对应的码簿，找出失真最小的代码矢量，并且将这个指标作为编码信息发送到多路复用器 510。

当进行矢量量化时，矢量量化器 509 利用输入信号谱信息，确定两个重要参数。这些参数之一是量化位分配，和另一个是码簿搜索加权。量化位分配是利用谱包络计算器 502 获得的谱包络 $env(m)$ 确定的。

当利用谱包络确定量化位分配时，也可以作出这样的设置，使分配在与频率 0 到 FL 相对应的频谱中的位数很少。

实现这个过程的一个例子是设置可以分配在频率 0 到 FL 中的最大位数 MAX_LOWBAND_BIT，并且，施加一个限制，以便分配在这个频带中的最大位数不超过最大位数 MAX_LOWBAND_BIT 的方法。

在这种实现例子中，由于在频率为 0 到 FL 的基本层中已经进行了编码，没有必要分配大量位数，和通过进行有意使这个频带中的量化粗糙些并使位分配保持在低水平上的量化和将额外位分配给频率 FL 到 FH，可以提高总体质量。也可以使用通过组合谱包络 env(m) 和前述 Bark 标度形状 Bq(k) 确定这种位分配的配置。

利用应用谱包络计算器 502 获得的谱包络 env(m) 和从 Bark 标度形状计算器 507 获得的 Bark 标度形状 Bq(k) 中计算的权重的失真度量进行矢量量化。通过找出使如下方程 (14) 规定的失真 D 达到极小的代码矢量 C 的指标 j 实现矢量量化。

$$D = \sum_m w(m)^2 (C_j(m) - X3(m))^2 \quad \dots (14)$$

这里，w(m) 表示加权函数。

利用谱包络 env(m) 和 Bark 标度形状 Bq(k) 可以将加权函数 w(n) 表示成如下方程 (15) 所示那样。

$$w(m) = (\text{env}(m) \cdot Bq(\text{Herz_to_Bark}(m)))^p \quad \dots (15)$$

这里，p 表示 0 和 1 之间的常数，和 Herz-to-Bark() 表示从 Herz 标度转换到 Bark 标度的函数。

当加权函数 w(m) 确定下来时，也可以作出这样的设置，使将位分配给与频率 0 到 FL 相对应的频谱的加权函数很小。实现这个过程的一个例子是下面将与频率 0 到 FL 相对应的加权函数 w(m) 的可能最大值设置成 MAX_LOWBAND_WGT，并且，施加一个限制，以便这个频带的加权函数 w(m) 的值不超过 MAX_LOWBAND_WGT 的方法。在这种实现例子中，在频率为 0 到 FL 的基本层中已经进行了编码，通过有意降低这个频带的量化精度和相对提高频率 FL 到 FH 的量化精度，可以提高总体质量。

最后，多路复用器 510 多路复用编码信息，并且将所得信号输出到多路复用器 108。当存在新输入信号时，重复上面的处理。当不存在新输入信号时，终止该处理。

因此，根据本实施例的信号处理设备，通过从输入信号中提取不超过预定频率的成分和利用码激励线性预测进行编码，和利用解码所得编码信息的结果通过 MDCT 处理进行编码，可以以低位速率进行高质编码。

上面已经描述了从减法器 106 获得的相减信号中分析 LPC 系数的例子，但本发明的信号处理设备也可以利用局部解码器 103 解码的 LPC 系数进行解码。

图 6 是示出增强层编码器 107 的配置的例子的图形。将与图 5 中相同的标号指定给图 6 中与图 5 中的那些相同的部分，并且省略对它们的详细描述。

图 6 中的增强层编码器 107 与图 5 中的增强层编码器 107 的不同之处在于，配备了转换表 601、LPC 系数映射部分 602、谱包络计算器 603 和变换部分 604，并且利用局部解码器 103 解码的 LPC 系数进行编码。

转换表 601 存储基本层 LPC 系数和增强层 LPC 系数，以及指示它们之间的对应关系。

LPC 系数映射部分 602 参考转换表 601，将从局部解码器 103 输入的基本层 LPC 系数转换成增强层 LPC 系数，并且将增强层 LPC 系数输出到谱包络计算器 603。

谱包络计算器 603 根据增强层 LPC 系数获取谱包络，并且将这个谱包络输出到变换部分 604。变换部分 604 变换谱包络并将结果输出到谱归一化器 506 和矢量量化器 509。

现在描述图 6 中增强层编码器 107 的操作。基本层 LPC 系数是为信号带 0 到 FL 中的信号求的，并且与增强层信号（信号带 0 到 FH）所用的 LPC 系数不一致。但是，在两者之间存在强关联。因此，在 LPC 系数映射部分 602 中，利用这种关联事先独立设计示出信号带 0 到 FL 信号的 LPC 系数和信号带 0 到 FH 信号的 LPC 系数之间的对应关系的转换表 601。这个转换表 601 用于从基本层 LPC 系数中求出增强层 LPC 系数。

图 7 是示出增强层中的 LPC 系数计算的例子的图形。转换表 601 由表示增强层 LPC 系数（次序 M）的 J 个候选者 $\{Y_j(m)\}$ 和与 $\{Y_j(m)\}$ 指定了对应关系、与基本层 LPC 系数具有相同次数（ $=K$ ）的候选者 $\{y_j(k)\}$ 组成。 $\{Y_j(m)\}$ 和 $\{y_j(k)\}$ 是根据大规模音频和语音数据等事先设计和提供的。当输入基本层 LPC 系数 $x(k)$ 时，从 $\{y_j(k)\}$ 当中找出与 $x(k)$ 最相似的一系列 LPC 系数。通过输出与确定为最相似的 LPC 系数的指标 j 相对应的增强层 LPC 系数 $Y_j(m)$ ，

可以实现从基本层 LPC 系数到增强层 LPC 系数的映射。

接着，谱包络计算器 603 根据以这种方式找出的增强层 LPC 系数获得谱包络。然后，变换部分 604 变换这个谱包络。然后，将这个变换谱包络当作如上所述的实现例子的谱包络，由此加以处理。

实现变换谱包络的变换部分 604 的一个例子是使与经受基本层编码的信号带 0 到 FL 相对应的谱包络的作用很小的处理。如果将谱包络指定为 $env(m)$ ，变换 $env'(m)$ 由如下方程 (16) 表示。

$$env'(m) = \begin{cases} env(m)^p & \text{if } 0 \leq m \leq Fl \\ env(m) & \text{else} \end{cases} \quad \dots (16)$$

这里， p 表示 0 和 1 之间的常数。

在频率为 0 到 FL 的基本层中已经进行了编码，和经过增强层编码的相减信号的频率 0 到 FL 之间的频谱接近平坦。与此无关，在如在这个实现例子中所述的 LPC 系数映射中不考虑这样的动作。因此，通过利用利用方程 (16) 校正谱包络的技术可以提高质量。

因此，根据本实施例的信号处理设备，通过利用基本层量化器量化的 LPC 系数求出增强层 LPC 系数，和从增强层 LPC 系数分析中计算出谱包络，使 LPC 分析和量化变得多余了，并且可以减少量化位的个数。

(第 3 实施例)

图 8 是示出根据本发明第 3 实施例的信号处理设备的增强层编码器的配置的方块图。将与图 5 中相同的标号指定给图 8 中与图 5 中的那些相同的部分，并且省略对它们的详细描述。

图 8 中的增强层编码器 107 与图 5 中的增强层编码器 107 的不同之处在于，配备了谱精细结构计算器 801，并且，利用基本层编码器 102 编码和局部解码器 103 解码的音调周期计算谱精细结构，并将那个谱精细结构应用在谱归一化和矢量量化中。

谱精细结构计算器 801 从在基本层中编码的音调周期 T 和音调增益 β 中计算谱精细结构，并且将谱精细结构输出到谱归一化器 506。

前述音调周期 T 和音调增益 β 实际上是编码信息的组成部分，并且，通过局部解码器 (如图 1 所示) 可以获得相同信息。因此，即使利用音调周期 T 和音调增益 β 进行编码，位速率也不会增加。

利用音调周期 T 和音调增益 β 进行编码，谱精细结构计算器 801 按照如下

方程 (17) 计算谱精细结构 $har(m)$ 。

$$har(m) = \left| \frac{1}{1 - \beta \cdot e^{-j \frac{2\pi m T}{M}}} \right| \quad \dots (17)$$

这里, M 表示谱分辨率。由于方程 (17) 是 β 的绝对值大于等于 1 时的振荡滤波, 所以还存在设置一种限制, 使 β 绝对值的可能范围小于等于小于 1 的预定设置值 (例如, 0.8) 的方法。

谱归一化器 506 利用谱包络计算器 502 获得的谱包络 $env(m)$ 和谱精细结构计算器 801 获得的谱精细结构 $har(m)$ 两者, 按照如下方程 (18) 进行归一化。

$$X2(m) = \frac{X1(m)}{env(m) \cdot har(m)} \quad \dots (18)$$

利用谱包络计算器 502 获得的谱包络 $env(m)$ 和谱精细结构计算器 801 获得的谱精细结构 $har(m)$ 两者还可以确定矢量量化器 509 的量化位分配。谱精细结构还用在矢量量化中的加权函数 $w(m)$ 确定中。具体地说, 按照如下方程 (18) 定义加权函数 $w(m)$ 。

$$w(m) = (env(m) \cdot har(m) \cdot Bq(Herz_to_Bark(m)))^p \quad \dots (19)$$

这里, p 表示 0 和 1 之间的常数, 和 $Herz_to_Bark()$ 表示从 Herz 标度转换到 Bark 标度的函数。

因此, 根据本实施例的信号处理设备, 通过利用基本层编码器编码和局部解码器解码的音调周期计算谱精细结构, 和将那个谱精细结构应用在谱归一化和矢量量化中, 可以提高量化性能。

(第 4 实施例)

图 9 是示出根据本发明第 4 实施例的信号处理设备的增强层编码器的配置的方块图。将与图 5 中相同的标号指定给图 9 中与图 5 中的那些相同的部分, 并且省略对它们的详细描述。

图 9 中的增强层编码器 107 与图 5 中的增强层编码器的不同之处在于, 配备了功率估计单元 901 和功率涨落量量化器 902, 并且, 在局部解码器 103 中利用基本层编码器 102 获得的编码信息生成解码信号, 根据那个解码信号预测 MDCT 系数功率, 和根据那个预测值编码涨落量。

在图 1 中, 解码参数从局部解码器 103 输出到增强层编码器 107, 但是,

在本实施例中，将局部解码器 103 获得的解码信号输出到增强层编码器 107，而不是解码参数。

图 5 中局部解码器 103 解码的信号 $s1(n)$ 输入到功率估计单元 901。然后，功率估计单元 901 根据这个解码信号 $s1(n)$ 估计 MDCT 系数功率。如果将 MDCT 系数功率指定为 $powp$ ， $powp$ 由如下方程 (20) 表示。

$$powp = \alpha \cdot \sum_{n=0}^{N-1} sl(n)^2 \quad \dots (20)$$

这里， N 表示解码信号 $s1(n)$ 的长度，和 α 表示用于校正的预定常数。在使用从基本层 LPC 系数中求出的谱斜度的另一种方法中，MDCT 系数功率估计由如下方程 (21) 表示。

$$powp = \alpha \cdot \beta \cdot \sum_{n=0}^{N-1} sl(n)^2 \quad \dots (21)$$

这里， β 表示具有当谱斜度大时（当低频带中谱能量大时）接近 0，和当谱斜度小时（当相对高区域中存在功率时）接近 1 的特性、取决于从基本层 LPC 系数中求出的谱斜度的变量。

接着，功率涨落量量化器 902 通过功率估计单元 901 获得的功率估计 $powp$ ，归一化 MDCT 部分 503 获得的 MDCT 系数的功率，并且量化涨落量。涨落量 r 用如下方程 (22) 表示。

$$r = \frac{pow}{powp} \quad \dots (22)$$

这里， pow 表示 MDCT 系数功率，和通过方程 (23) 来计算。

$$pow = \sum_{m=0}^{M-1} X(m)^2 \quad \dots (23)$$

这里， $X(m)$ 表示 MDCT 系数，和 M 表示帧长度。功率涨落量量化器 902 量化涨落量 r ，将编码信息发送到多路复用器 510，并且还解码量化涨落量 r_q 。利用量化涨落量 r_q ，功率归一化器 505 利用如下方程 (24) 归一化 MDCT 系数。

$$X1(m) = \frac{X(m)}{\sqrt{r_q \cdot powp}} \quad \dots (24)$$

这里， $X1(m)$ 表示功率归一化之后的 MDCT 系数。

因此，根据本实施例的信号处理设备，通过利用基本层解码信号功率和增强层 MDCT 系数功率之间的关联，利用基本层解码信号预测 MDCT 系数功率，和根据那个预测值编码涨落量，可以减少 MDCT 系数功率量化所需的位数。

(第5实施例)

图10是示出根据本发明第5实施例的信号处理设备的配置的方块图。图10中的信号处理设备主要包括多路分用器1001、基本层解码器1002、向上取样器1003、增强层解码器1004和加法器1005。

多路分用器1001分离编码信息，生成基本层编码信息和增强层编码信息。然后，多路分用器1001将基本层编码信息输出到基本层解码器1002，并将增强层编码信息输出到增强层解码器1004。

基本层解码器1002利用多路分用器1001获得的基本层编码信息解码取样速率FL解码信号，并且将所得信号输出到向上取样器1003。同时，将基本层解码器1002解码的参数输入到增强层解码器1004。向上取样器1003将解码信号取样频率升高到FH，并且将它输出到加法器1005。

增强层解码器1004利用多路分用器1001获得的增强层编码信息和基本层解码器1002解码的参数，解码取样速率FH解码信号，并且将所得信号输出到加法器1005。

加法器1005对从向上取样器1003输出的解码信号和从增强层解码器1004输出的解码信号进行相加。

现在描述本实施例的信号处理设备的操作。首先，输入在第1到第4实施例任何一个的信号处理设备中编码的代码，并且，多路分用器1001分离那个代码，生成基本层编码信息和增强层编码信息。

接着，基本层解码器1002利用多路分用器1001获得的基本层编码信息解码取样速率FL解码信号。然后，向上取样器1003将那个解码信号的取样频率升高到FH。

在增强层解码器1004中，利用多路分用器1001获得的增强层编码信息和基本层解码器1002解码的参数解码取样速率FH解码信号。

加法器1005相加向上取样器1003向上取样的基本层解码信号和增强层解码信号。当存在新输入信号时，重复上面的处理。当不存在新输入信号时，终止该处理。

因此，根据本实施例的信号处理设备，通过利用基本层解码器1002解码的参数进行增强层解码器1004解码，可以从利用基本层编码中的解码参数进行增强层编码的声音编码单元的编码信息中生成解码信号。

现在描述基本层解码器1002。图11是示出基本层解码器1002的例子的

方块图。图 11 中的基本层解码器 1002 主要包括多路分用器 1101、激励发生器 1102 和合成滤波器 1103，和进行 CELP 解码处理。

多路分用器 1101 从多路分用器 1001 输入的基本层编码信息中分离出各种参数，并且将这些参数输出到和合成滤波器 1103。

激励发生器 1102 进行自适应矢量、自适应矢量增益、噪声矢量和噪声矢量增益解码，利用这些值生成激励信号，并且将这个激励信号输出到合成滤波器 1103。合成滤波器 1103 利用解码 LPC 系数生成合成信号。

现在描述图 11 中基本层解码器 1002 的操作。首先，多路分用器 1101 从基本层编码信息中分离出各种参数。

接着，激励发生器 1102 进行自适应矢量、自适应矢量增益、噪声矢量和噪声矢量增益解码。然后，激励发生器 1102 按照如下方程 (25) 生成激励矢量 $ex(n)$ 。

$$ex(n) = \beta_q \cdot q(n) + \gamma_q \cdot c(n) \quad \dots (25)$$

这里， $q(n)$ 表示自适应矢量， β_q 表示自适应矢量增益， $c(n)$ 表示噪声矢量，和 γ_q 表示噪声矢量增益。

然后，合成滤波器 1103 利用解码 LPC 系数，按照如下方程 (26) 生成合成信号 $syn(n)$ 。

$$syn(n) = ex(n) + \sum_{i=1}^{NP} \alpha_q(i) \cdot syn(n-i) \quad \dots (26)$$

这里， α_q 表示解码 LPC 系数，和 NP 表示 LPC 系数的次序。

将以这种方式解码的解码信号 $syn(n)$ 输出到向上取样器 1003，并且将作为解码结果获得的参数输出到增强层解码器 1004。当存在新输入信号时，重复上面的处理。当不存在新输入信号时，终止该处理。取决于 CELP 配置，在经过后置滤波器之后输出合成信号的模式也是可以的。所述的后置滤波器具有使编码失真更不易觉察的后处理功能。

现在描述增强层解码器 1004。图 12 是示出增强层解码器 1004 的例子的方块图。图 12 中的增强层解码器 1004 主要包括多路分用器 1201、LPC 系数解码器 1202、谱包络计算器 1203、矢量解码器 1204、Bark 标度形状解码器 1205、乘法器 1206、乘法器 1207、功率解码器 1208、乘法器 1209 和 IMDCT 部分 1210。

多路分用器 1201 从多路分用器 1001 输出的增强层编码信息中分离出各

种参数。LPC 系数解码器 1202 利用 LPC 系数相关编码信息解码 LPC 系数，并且将结果输出到谱包络计算器 1203。

谱包络计算器 1203 利用 LPC 系数，按照方程 (6) 计算谱包络，并且将谱包络 $env(m)$ 输出到矢量解码器 1204 和乘法器 1207。

矢量解码器 1204 根据谱包络计算器 1203 获得的谱包络 $env(m)$ 确定量化位分配，并且根据从多路分用器 1201 中获得的编码信息和前述量化位分配，解码归一化 MDCT 系数 $X3q(m)$ 。量化位分配方法与用在第 1 到第 4 实施例任何一个的编码方法中的增强层编码中的方法相同。

Bark 标度形状解码器 1205 根据从多路分用器 1201 中获得的编码信息，解码 Bark 标度形状 $Bq(k)$ ，并且将结果输出到乘法器 1206。

乘法器 1206 按照如下方程 (27)，将归一化 MDCT 系数 $X3q(m)$ 乘以 Bark 标度形状 $Bq(k)$ ，并且将结果输出到乘法器 1207。

$$X2_q(m) = X3_q(m) \sqrt{B_q(k)} \quad fl(k) \leq m \leq fh(k) \quad 0 \leq k < K \quad \dots (27)$$

这里， $fl(k)$ 表示第 k 子频带的最低频率和 $fh(k)$ 表示第 k 子频带的最高频率，和 K 表示子频带个数。

乘法器 1207 按照如下方程 (28)，将从乘法器 1206 中获得的归一化 MDCT 系数 $X2q(m)$ 乘以谱包络计算器 1203 获得的谱包络 $env(m)$ ，并且将相乘结果输出到乘法器 1209。

$$X1_q(m) = X2_q(m) env(m) \quad \dots (28)$$

功率解码器 1208 根据从多路分用器 1201 中获得的编码信息解码功率 $powq$ ，并且将解码结果输出到乘法器 1209。

乘法器 1209 按照如下方程 (29)，将归一化 MDCT 系数 $X1q(m)$ 乘以解码功率 $powq$ ，并且将相乘结果输出到 IMDCT 部分 1210。

$$X_q(m) = X1_q(m) \sqrt{powq} \quad \dots (29)$$

IMDCT 部分 1210 对以这种方式获得的解码 MDCT 系数进行 IMDCT (改进离散余弦逆变换)，重叠和相加一半在前一个帧中获得和一半在当前帧中获得的信号，并且，所得信号是输出信号。当存在新输入信号时，重复上面的处理。当不存在新输入信号时，终止该处理。

因此，根据本实施例的信号处理设备，通过利用基本层解码器解码的参数进行增强层解码器解码，可以从利用基本层编码中的解码参数进行增强层编码的编码单元的编码信息中生成解码信号。

(第6实施例)

图13是示出增强层解码器1004的例子的方块图。将与图12中相同的标号指定给图13中与图2中的那些相同的部分，并且省略对它们的详细描述。

图13中的增强层解码器1004与图12中的增强层编码器1004的不同之处在于，配备了转换表1301、LPC系数映射部分1302、谱包络计算器1303和变换部分1304，并且利用基本层解码器1002解码的LPC系数进行解码。

转换表1301存储基本层LPC系数和增强层LPC系数，以及指示它们之间的对应关系。

LPC系数映射部分1302参考转换表1301，将从局部解码器1002输入的基本层LPC系数转换成增强层LPC系数，并且将增强层LPC系数输出到谱包络计算器1303。

谱包络计算器1303根据增强层LPC系数获取谱包络，并且将这个谱包络输出到变换部分1304。变换部分1304变换谱包络并将结果输出到乘法器1207和矢量解码器1204。变换方法的一个例子是显示在第2实施例的方程(16)中的方法。

现在描述图13中增强层解码器1003的操作。基本层LPC系数是为信号带0到FL中的信号求的，并且与增强层信号(信号带0到FH)所用的LPC系数不一致。但是，在两者之间存在强关联。因此，在LPC系数映射部分1302中，利用这种关联事先独立设计示出信号带0到FL信号的LPC系数和信号带0到FH信号的LPC系数之间的对应关系的转换表1301。这个转换表1301用于从基本层LPC系数中求出增强层LPC系数。

转换表1301的细节与第2实施例中转换表601的细节相同。

因此，根据本实施例的信号处理设备，通过利用基本层解码器量化的LPC系数求出增强层LPC系数，和从增强层LPC系数中计算出谱包络，使LPC分析和量化变得多余了，并且可以减少量化位的个数。

(第7实施例)

图14是示出根据本发明第7实施例的信号处理设备的增强层解码器的配置的方块图。将与图12中相同的标号指定给图14中与图12中的那些相同的部分，并且省略对它们的详细描述。

图14中的增强层解码器1004与图12中的增强层解码器的不同之处在于，配备了谱精细结构计算器1401，并且，利用基本层解码器1002解码的

音调周期计算谱精细结构，将那个谱精细结构应用在解码中，并且进行与声音编码相对应的声音解码，从而提高量化性能。

谱精细结构计算器 1401 从基本层解码器 1002 解码的音调周期 T 和音调增益 β 中计算谱精细结构，并且将谱精细结构输出到矢量解码器 1204 和乘法器 1207。

利用音调周期 T_q 和音调增益 β_q ，谱精细结构计算器 1401 按照如下方程 (30) 计算谱精细结构 $har(m)$ 。

$$har(m) = \left| \frac{1}{1 - \beta_q \cdot e^{-j \frac{2\pi m T_q}{M}}} \right| \quad \dots (30)$$

这里， M 表示谱分辨率。由于方程 (30) 是 β_q 的绝对值大于等于 1 时的振荡滤波，所以还可以设置一种限制，使 β_q 绝对值的可能范围小于等于小于 1 的预定设置值（例如，0.8）。

利用谱包络计算器 1203 获得的谱包络 $env(m)$ 和谱精细结构计算器 1401 获得的谱精细结构 $har(m)$ 两者还可以确定矢量解码器 1204 的量化位分配。然后，根据那个量化位分配和从多路分用器 1201 中获得的编码信息解码归一化 MDCT 系数 $X3_q(m)$ 。此外，通过按照如下方程 (31) 将归一化 MDCT 系数 $X2_q(m)$ 乘以谱包络 $env(m)$ 和谱精细结构 $har(m)$ 求出归一化 MDCT 系数 $X1_q(m)$ 。

$$X1_q(m) = X2_q(m)env(m)har(m) \quad \dots (31)$$

因此，根据本实施例的信号处理设备，通过利用基本层编码器编码和局部解码器解码的音调周期计算谱精细结构，并将那个谱精细结构应用在谱归一化和矢量量化中，可以进行与声音编码相对应的声音解码，从而提高量化性能。

(第 8 实施例)

图 15 是示出根据本发明第 8 实施例的信号处理设备的增强层解码器的配置的方块图。将与图 12 中相同的标号指定给图 15 中与图 12 中的那些相同的部分，并且省略对它们的详细描述。

图 15 中的增强层解码器 1004 与图 12 中的增强层解码器的不同之处在于，配备了功率估计单元 1501、功率涨落量解码器 1502 和功率发生器 1503，并且，形成与利用基本层解码信号预测 MDCT 系数功率，并根据那个预测值编码涨落量的编码器相对应的解码器。

在图 10 中，解码参数从基本层解码器 1002 输出到增强层解码器 1004，但是，在本实施例中，将基本层解码器 1002 获得的解码信号输出到增强层解码器 1004，而不是解码参数。

功率估计单元 1501 利用方程 (20) 或方程 (21)，从基本层解码器 1002 解码的解码信号 $s_1(n)$ 中估计 MDCT 系数的功率。

功率涨落量量化器 1502 根据从多路分用器 1201 获得的编码信息解码功率涨落量，并且将这个功率涨落量输出到功率发生器 1503。功率发生器 1503 从功率涨落量中计算功率。

乘法器 1209 按照如下方程 (32) 求出 MDCT 系数。

$$X_q(m) = X_{1q}(m) \sqrt{rq \cdot powp} \quad \dots (32)$$

这里， rq 表示功率涨落量，和 $powp$ 表示功率估计。 $X_{1q}(m)$ 表示来自乘法器 1207 的输出信号。

因此，根据本实施例的信号处理设备，通过配置与利用基本层解码信号预测 MDCT 系数功率，和根据那个预测值编码涨落量的编码器相对应的解码器，可以减少 MDCT 系数功率量化所需的位数。

(第 9 实施例)

图 16 是示出根据本发明第 9 实施例的声音编码设备的配置的方块图。图 16 中的声音编码设备 1600 主要包括向下取样器 1601、基本层编码器 1602、局部解码器 1603、向上取样器 1604、延迟器 1605、减法器 1606、频率确定部分 1607、增强层编码器 1608 和多路复用器 1609。

基本层编码器 1602 以预定基本帧为单位编码取样速率 FL 输入数据，并且将第一编码信息输出到局部解码器 1603 和多路复用器 1609。基本层编码器 1602 可以利用，例如，CELP 方法编码输入数据。

局部解码器 1603 解码第一编码信息，并且将通过解码获得的解码信号输出到向上取样器 1604。向上取样器 1604 将解码信号取样速率升高到 FH ，并且将结果输出到减法器 1606 和频率确定部分 1607。

延迟器 1605 将输入信号延迟预定时间，然后，将信号输出到减法器 1606。通过使这个延迟时间等于在向下取样器 1601、基本层编码器 1602、局部解码器 1603 和向上取样器 1604 中产生的时间延迟，可以防止在接着的相减处理中出现相移。减法器 1606 进行输入信号和解码信号之间的相减，并且将相减结果作为误差信号输出到增强层编码器 1608。

频率确定部分 1607 根据取样速率已经升高到 FH 的解码信号确定进行误差信号编码的区域和不进行误差信号编码的区域，并且通知增强层编码器 1608。例如，频率确定部分 1607 根据取样速率已经升高到 FH 的解码信号确定听觉掩蔽的频率，并且将这个频率输出到增强层编码器 1608。

增强层编码器 1608 将误差信号转换到频域和生成误差谱，并且根据从频率确定部分 1607 中获得的频率信息进行误差谱编码。多路复用器 1609 多路复用通过基本层编码器 1602 编码获得编码信息和通过增强层编码器 1608 编码获得编码信息。

现在分别描述基本层编码器 1602 和增强层编码器 1608 编码的信号。图 17 是示出声信号信息分布的例子的图形。在图 17 中，垂直轴表示信息量，而水平轴表示频率。图 17 示出了在哪些频带中给出多少包含在输入信号中的语音信息和背景音乐和背景噪声信息。

如图 17 所示，在语音信息的情况下，在低频区中存在大量信息，信息量随着频区增高而减少。相反，在背景音乐和背景噪声信息的情况下，与语音信息相比，在较低区域中存在相对少的信息，和在较高区域中存在大量信息。

因此，在基本层中，利用 CELP 高质量地编码语音信号，和在增强层中，不能在基本层中得到表示的背景音乐或环境声音和存在比基本层覆盖的频区高的频率成分的信号得到有效编码。

图 18 是示出基本层和增强层中编码区的例子的图形。在图 18 中，垂直轴表示信息量，而水平轴表示频率。图 18 示出了作为分别由基本层编码器 1603 和增强层编码器 1608 编码的信息的对象的区域。

基本层编码器 1602 被设计成有效表示从 0 到 FL 的频带中的语音信息，并且可以对该区域中的语音信息进行高质编码。但是，对于基本层编码器 1602，从 0 到 FL 的频带中背景音乐和背景噪声信息的编码质量不高。

增强层编码器 1608 被设计成覆盖如上所述，基本层编码器 1602 的能力不足的部分和从 FL 到 FH 的频带中的信号。因此，通过组合基本层编码器 1502 和增强层编码器 1608，可以在宽带中实现高质编码。

如图 18 所示，通过基本层编码器 1602 中的编码获得的第一编码信息包含 0 和 FL 之间的频带中的语音信息，因此，可以实现即使只利用至少第一编码信息也可以获得解码信号的可伸缩功能。

此外，可以考虑利用增强层中的听觉掩蔽来升高编码频率。听觉掩蔽应

用了当供应某个信号时，频率在那个信号的频率附近的信号不能被听到（被掩蔽）的人听觉特性。

图 19 是示出声（音乐）信号谱的例子的例子。在图 19 中，实线表示听觉掩蔽，和虚线表示误差谱。这里的“误差谱”指的是输入信号和基本层解码信号的误差信号（增强层输入信号）的频谱。

在图 19 中阴影区所指的误差谱中，幅度值低于听觉掩蔽，因此，人的耳朵听不到声音，而在其它区域中，误差谱幅度值超过听觉掩蔽，因此，感觉得到量化失真。

在增强层中，只需编码包括在图 19 中的白区中的误差谱，使得那些区域的量化失真小于听觉掩蔽。属于阴影区的系数已经小于听觉掩蔽，因此，不需要量化。

在本实施例的声音编码设备 1600 中，根据听觉掩蔽等编码残留误差信号的频率不从编码方发送到解码方，编码方和解码方利用向上取样基本层解码信号分开确定进行增强层编码的误差谱频率。

在解码信号来源于对基本层编码信息的解码的情况下，编码方和解码方获得相同信号，因此，通过让编码方通过从这个解码信号中确定听觉掩蔽频率来编码信号和让解码方通过从这个解码信号中获取听觉掩蔽频率来解码信号，编码和发送作为附加信息的误差谱频率信息就变得多余了，从而能够实现位速率的降低。

接着，详细描述根据本实施例的声音编码设备的操作。首先，描述频率确定部分 1607 从向上取样基本层解码信号（下文称为“基本层解码信号”）中确定在增强层中编码的误差谱频率的操作。图 20 是示出本实施例的声音编码设备的频率确定部分的内部配置的例子的方块图。

在图 20 中，频率确定部分 1607 主要包括 FFT 部分 1901、估计听觉掩蔽计算器 1902 和确定部分 1903。

FFT 部分 1901 对从向上取样器 1604 输出的基本层解码信号 $x(n)$ 进行正交转换，计算幅度谱 $P(m)$ ，并且将幅度谱 $P(m)$ 输出到估计听觉掩蔽计算器 1902 和确定部分 1903。具体地说，FFT 部分 1901 利用如下方程 (33) 计算幅度谱 $P(m)$ 。

$$P(m) = \sqrt{\text{Re}^2(m) + \text{Im}^2(m)} \quad \cdots (33)$$

这里， $\text{Re}(m)$ 和 $\text{Im}(m)$ 表示基本层解码信号 $x(n)$ 的付里叶系数的实部和虚

部, 和 m 表示频率。

接着, 估计听觉掩蔽计算器 1902 利用基本层解码信号幅度谱 $P(m)$ 计算估计听觉掩蔽 $M'(m)$, 并且将估计听觉掩蔽 $M'(m)$ 输出到确定部分 1903。一般说来, 听觉掩蔽是根据输入信号的频谱计算的, 但在这个实现例子中, 利用基本层解码信号 $x(n)$ 而不是利用输入信号来估计听觉掩蔽。这基于这样的思想, 由于基本层解码信号 $x(n)$ 被确定成相对应于输入信号失真很小, 所以, 如果用基本层解码信号 $x(n)$ 取代输入信号, 将会取得足够好的近似, 并且主要问题也不会存在。

然后, 确定部分 1903 利用基本层解码信号幅度谱 $P(m)$ 和估计听觉掩蔽计算器 1902 获得的估计听觉掩蔽 $M'(m)$, 确定增强层编码器 1608 进行误差谱编码可应用的频率。确定部分 1903 把基本层解码信号幅度谱 $P(m)$ 当作误差谱的近似, 并且将使如下方程 (34) 成立的频率输出到增强层编码器 1608。

$$P(m) - M'(m) > 0 \quad \dots (34)$$

在方程 (34) 中, 项 $P(m)$ 估计误差谱的大小, 和项 $M'(m)$ 估计听觉掩蔽。然后, 确定部分 1903 比较估计误差谱和估计听觉掩蔽的值, 并且, 如果方程 (34) 得到满足 - 也就是说, 如果估计误差谱的值超过估计听觉掩蔽的值 - 假设那个频率的误差谱是可当作噪声感觉的, 并且让增强层编码器 1608 对它进行编码。

相反, 如果估计误差谱的值小于估计听觉掩蔽的大小, 确定部分 1903 认为由于掩蔽效应, 那个频率的误差谱将不会当作噪声感觉到, 并且确定不要对这个频率的误差谱进行量化。

现在描述估计听觉掩蔽计算器 1902 的操作。图 21 是示出本实施例的声音编码设备的听觉掩蔽计算器的内部配置的例子图形。在图 21 中, 估计听觉掩蔽计算器 1902 主要包括 Bark 谱计算器 2001、扩展函数卷积单元 2002、音调计算器 2003 和听觉掩蔽计算器 2004。

在图 21 中, Bark 谱计算器 2001 利用如下方程 (35) 计算 Bark 谱 $B(k)$ 。

$$B(k) = \sum_{m=f_l(k)}^{f_h(k)} P^2(m) \quad \dots (35)$$

这里, $P(m)$ 表示幅度谱, 并且从上面的方程 (33) 中求出, k 与 Bark 谱号相对应, 和 $f_l(k)$ 和 $f_h(k)$ 分别表示第 k Bark 谱的最低频率和最高频率。在频带分布在 Bark 标度上是等间隔的情况下, Bark 谱 $B(k)$ 表示谱强度。如果 Herz

标度用 h 表示和 Bark 标度用 B 表示, Herz 标度和 Bark 标度之间的关系用如下方程 (36) 表示。

$$B = 13 \tan^{-1}(0.76f) + 3.5 \tan^{-1}\left(\frac{f}{7.5}\right) \quad \dots (36)$$

扩展函数卷积单元 2002 利用如下方程 (37) 将扩展函数 $SF(k)$ 卷积成 Bark 谱 $B(k)$ 。

$$C(k) = B(k) * SF(k) \quad \dots (37)$$

音调计算器 2003 利用如下方程 (38) 求出每个 Bark 谱的谱平坦度 $SFM(k)$ 。

$$SFM(k) = \frac{\mu g(k)}{\mu a(k)} \quad \dots (38)$$

这里, $\mu g(k)$ 表示第 k Bark 谱中功率谱的几何平均, 和 $\mu a(k)$ 表示第 k Bark 谱中功率谱的算术平均。然后, 音调计算器 2003 利用如下方程 (39), 从谱平坦度 $SFM(k)$ 的分贝值 $SFMdB(k)$ 中计算音调系数 $\alpha(k)$ 。

$$\alpha(k) = \min\left(\frac{SFMdB(k)}{-60}, 1.0\right) \quad \dots (39)$$

利用如下方程 (40), 听觉掩蔽计算器 2004 从音调计算器 2003 计算的音调系数 $\alpha(k)$ 中求出每个 Bark 标度的偏移量 $O(k)$ 。

$$O(k) = \alpha(k) \cdot (14.5 - k) + (1.0 - \alpha(k)) \cdot 5.5 \quad \dots (40)$$

然后, 听觉掩蔽计算器 2004 利用如下方程 (41), 通过从扩展函数卷积单元 2002 求出的 $C(k)$ 中减去偏移量 $O(k)$ 计算听觉掩蔽 $T(k)$ 。

$$T(k) = \max\left(10^{\log_{10}(C(k)) - (O(k)/10)}, T_q(k)\right) \quad \dots (41)$$

这里, $T_q(k)$ 表示绝对阈值。绝对阈值代表作为人听觉特性观察的听觉掩蔽的最小值。听觉掩蔽计算器 2004 将在 Bark 标度上表达的听觉掩蔽 $T(k)$ 转换成 Herz 标度。并且求出输出到确定部分 1903 的估计听觉掩蔽 $M'(k)$ 。

增强层编码器 1608 利用以这种方式求出的经过量化的频率 m 进行 MDCT 系数编码。图 22 是示出本实施例的增强层编码器的内部配置的例子图形。图 22 中的增强层编码器 1608 主要包括 MDCT 部分 2101 和 MDCT 系数量化器 2102。

MDCT 部分 2101 将从减法器 1606 输出的输入信号乘以分析窗, 然后, 进行 MDCT (改进离散余弦变换) 处理以获得 MDCT 系数。在 MDCT 处理中, 供分析用的正交基用于相继的两个帧。并且, 分析帧一半重叠, 分析帧的前一半

是奇函数，而分析帧的后一半是偶函数。MDCT 处理的特征是，由于逆变换之后波形的叠加造成的相加，不会出现帧边缘失真。当进行 MDCT 时，输入信号被乘以诸如正弦函数窗口之类的窗口函数。如果将一系列 MDCT 系数指定为 $X(n)$ 时，MDCT 系数按照如下方程 (42) 计算。

$$X(m) = \sqrt{\frac{1}{N}} \sum_{n=0}^{2N-1} x(n) \cos \left\{ \frac{(2n+1+N) \cdot (2m+1)\pi}{4N} \right\} \quad \dots (42)$$

MDCT 系数量化器 2102 量化与来自频率确定部分 1607 的频率相对应的系数。然后，MDCT 系数量化器 2102 将量化 MDCT 系数编码信息输出到多路分用器 1609。

因此，根据本实施例的声音编码设备，由于利用基本层解码信号确定了增强层中用于量化的频率，没有必要将用于量化的频率信息从编码方发送到解码方，并且能够以低位速率进行高质编码。

在上面的实施例中，已经描述了使用 FFT 的听觉掩蔽计算方法，但是，也可以利用 MDCT 取代 FFT 来计算听觉掩蔽。图 23 是示出本实施例的听觉掩蔽计算器的内部配置的例子图形。将与图 20 中相同的标号指定给图 23 中与图 20 中的那些相同的部分，并且省略对它们的详细描述。

MDCT 部分 2201 利用 MDCT 系数近似计算幅度谱 $P(m)$ 。具体地说，MDCT 部分 2201 利用如下方程 (43) 近似计算幅度谱 $P(m)$ 。

$$P(m) = \sqrt{R^2(m)} \quad \dots (43)$$

这里， $R(m)$ 是通过向上取样器 1604 供应的信号进行 MDCT 处理求出的 MDCT 系数。

估计听觉掩蔽计算器 1902 从 $P(m)$ 中近似计算 Bark 谱 $B(k)$ 。此后，按照上述方法计算用于量化的频率信息。

因此，本实施例的声音编码设备可以利用 MDCT 计算听觉掩蔽。

现在描述解码方。图 24 是示出根据本发明第 9 实施例的声音解码设备的配置的方块图。图 24 中的声音解码设备 2300 主要包括多路分用器 2301、基本层解码器 2302、向上取样器 2303、频率确定部分 2304、增强层解码器 2305 和加法器 2306。

多路分用器 2301 将声音编码设备 1600 编码的代码分离成基本层第一编码信息和增强层第二编码信息，并且将第一编码信息输出到基本层解码器 2302，和将第二编码信息输出到增强层解码器 2305。

基本层解码器 2302 解码第一编码信息和获取取样速率 FL 解码信号。然后,基本层解码器 2302 将解码信号输出到向上取样器 2303。向上取样器 2303 将取样速率 FL 解码信号转换成取样速率 FH 解码信号,并且将这个信号输出到频率确定部分 2304 和加法器 2306。

利用向上取样基本层解码信号,频率确定部分 2304 确定要在增强层解码器 2305 中解码的误差谱频率。这个频率确定部分 2304 具有与图 16 中的频率确定部分 16 相同类型的配置。

增强层解码器 2305 解码第二编码信息和将取样速率 FH 解码信号输出到加法器 2306。

加法器 2306 相加向上取样器 2303 向上取样的基本层解码信号和增强层解码器 2305 解码的增强层解码信号,并且输出所得信号。

接着,详细描述根据本实施例的声音解码设备的每个方块的操作。图 25 是示出本实施例的声音解码设备的增强层解码器的内部配置的例子方块图。图 25 示出了图 24 中的增强层解码器 2305 的内部配置的例子。图 25 中的增强层解码器 2305 主要包括 MDCT 系数解码器 2401、IMDCT 部分 2402 和叠加加法器 2403。

MDCT 系数解码器 2401 根据从频率确定部分 2304 输出的频率,确定从多路分用器 2301 输出的第二编码信息中量化的 MDCT 系数。具体地说,定位与频率确定部分 2304 所指的频率相对应的解码 MDCT 系数,并且,对于其它频率填上零。

IMDCT 部分 2402 对从 MDCT 系数解码器 2401 输出的 MDCT 系数进行逆 MDCT 处理,生成时域信号,并且将这个信号输出到叠加加法器 2403。

叠加加法器 2403 在对来自 IMDCT 部分 2402 的时域信号开窗之后,进行叠加和相加操作,并且,它将解码信号输出到加法器 2306。具体地说,叠加加法器 2403 将解码信号乘以一个窗口和叠加在前一帧和当前帧中解码的时域信号,进行相加,并且生成输出信号。

因此,根据本实施例的声音解码设备,通过利用基本层解码信号确定用于增强层解码的频率,无需任何附加信息就可以确定用于增强层解码的频率,并且,能够以低位速率进行高质编码。

(第 10 实施例)

在本实施例中,描述 CELP 用在基本层编码中的例子。图 26 是示出根据

本发明第 10 实施例的基本层编码器的内部配置的例子方块图。图 26 示出了图 16 中的基本层编码器 1602 的内部配置的例子。图 26 中的基本层编码器 1602 主要包括 LPC 分析器 2501、加权部分 2502、自适应码簿搜索单元 2503、自适应增益量化器 2504、目标矢量发生器 2505、噪声码簿搜索单元 2506、噪声增益量化器 2507 和多路复用器 2508。

LPC 分析器 2501 计算取样速率 FL 输入信号的 LPC 系数，将 LPC 系数转换成诸如 LSP 系数之类适合于量化的参数，并且进行量化。然后，LPC 分析器 2501 将通过这种量化获得的编码信息输出到多路复用器 2508。

此外，LPC 分析器 2501 从编码信息中计算量化 LSP 系数，将这个量化 LSP 系数转换成 LSP 系数，并且将量化 LSP 系数输出到自适应码簿搜索单元 2503、自适应增益量化器 2504、噪声码簿搜索单元 2506 和噪声增益量化器 2507。LPC 分析器 2501 还将原 LPC 系数输出到加权部分 2502、自适应码簿搜索单元 2503、自适应增益量化器 2504、噪声码簿搜索单元 2506 和噪声增益量化器 2507。

加权部分 2502 根据 LPC 分析器 1501 获得的 LPC 系数，对从向下取样器 1601 输出的输入信号进行加权。这种操作的目的是进行谱成形，以便通过输入信号谱包络掩蔽量化失真谱。

然后，自适应码簿搜索单元 2503 利用作为目标信号的加权输入信号搜索自适应码簿。以音调周期为基础重复以前确定激励信号的信号被称为自适应矢量，和自适应码簿由在预定范围的音调周期上生成的自适应矢量组成。

如果将加权输入信号指定为 $t(n)$ ，和将包括原 LPC 系数和量化 LPC 系数的加权合成滤波器的脉冲响应被卷积成音调周期为 i 的自适应矢量的信号指定为 $p_i(n)$ ，那么，自适应码簿搜索单元 2503 将使如下方程 (44) 的估算函数 D 达到极小的自适应矢量的音调周期 i 作为编码信息输出到多路复用器 408。

$$D = \sum_{n=0}^{N-1} t^2(n) - \frac{\left(\sum_{n=0}^{N-1} t(n)p_i(n) \right)^2}{\sum_{n=0}^{N-1} p_i^2(n)} \quad \dots (44)$$

这里， N 表示矢量长度。由于方程 (44) 的第一项与音调周期 i 无关，自适应码簿搜索单元 2503 实际上只计算第二项。

自适应增益量化器 2504 进行乘以自适应矢量的自适应增益的量化。自适

应增益 β 用方程(45)表示。自适应增益量化器 2504 进行这个自适应增益 β 的标量量化, 并且将在量化过程中获得的编码信息输出到多路复用器 2508。

$$\beta = \frac{\sum_{n=0}^{N-1} t(n)p_i(n)}{\sum_{n=0}^{N-1} p_i^2(n)} \quad \dots (45)$$

目标矢量发生器 2505 从输入信号中减去自适应矢量产生的效果, 生成和输出噪声码簿搜索单元 2506 和噪声增益量化器 2507 使用的目标矢量。在目标矢量发生器 2505 中, 如果 $p_i(n)$ 表示当方程(44)所表示的估算函数 D 达到极小时, 加权合成滤波器脉冲响应被卷积成自适应矢量的信号, 和 βq 表示当方程(45)所表示的自适应矢量 β 经受标量量化时的量化自适应增益, 那么, 目标矢量 $t_2(n)$ 由如下方程(46)表示。

$$t_2(n) = t(n) - \beta q \cdot p_i(n) \quad \dots (46)$$

噪声码簿搜索单元 406 利用前述目标矢量 $t_2(n)$ 、原 LPC 系数和量化 LPC 系数, 进行噪声码簿搜索。噪声码簿搜索单元 406 可以使用, 例如, 随机噪声或利用大量语音信号学习的信号。此外, 可以使用代数码簿。代数码簿由一些脉冲组成。这样代数码簿的特征是, 通过小量计算就可以确定脉冲位置和脉冲代码(极性)的最佳组合。

如果将目标矢量指定为 $t_2(n)$, 和将加权合成滤波器的脉冲响应被卷积成与代码 j 相对应的噪声矢量的信号指定为 $c_j(n)$, 那么, 噪声码簿搜索单元 2506 将使如下方程(47)的估算函数 D 达到极小的噪声矢量的指标 j 输出到多路复用器 2508。

$$D = \sum_{n=0}^{N-1} t_2^2(n) - \frac{\left(\sum_{n=0}^{N-1} t_2(n)c_j(n) \right)^2}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (47)$$

噪声增益量化器 2507 量化乘以噪声矢量的噪声增益。噪声增益量化器 2507 利用如下方程(48)计算自适应增益 γ , 对这个噪声增益 γ 进行标量量化, 并且将编码信息输出到多路复用器 2508。

$$\gamma = \frac{\sum_{n=0}^{N-1} t_2(n) c_j(n)}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (48)$$

多路复用器 2508 多路复用 LPC 系数的编码信息、自适应矢量、自适应增益、噪声矢量和噪声增益编码信息，并且将所得信号输出到局部解码器 1603 和多路复用器 1609。

现在描述解码方。图 27 是示出本实施例的基本层解码器的内部配置的例子方块图。图 27 示出了基本层解码器 2302 的例子。图 27 中的基本层解码器 2302 主要包括多路分用器 2601、激励发生器 2602 和合成滤波器 2603。

多路分用器 2601 将来自多路分用器 2301 的第一编码信息分离成 LPC 系数、自适应矢量、自适应增益、噪声矢量和噪声增益编码信息，并且将自适应矢量、自适应增益、噪声矢量和噪声增益编码信息输出到激励发生器 2602。类似地，多路分用器 2601 将线性预测系数编码信息输出到合成滤波器 2603。

激励发生器 2602 解码自适应矢量、自适应矢量增益、噪声矢量和噪声矢量增益编码信息，和利用如下方程 (49) 生成激励矢量 $ex(n)$ 。

$$ex(n) = \beta_q \cdot q(n) - \gamma_q \cdot c(n) \quad \dots (49)$$

这里， $q(n)$ 表示自适应矢量， β_q 表示自适应矢量增益， $c(n)$ 表示噪声矢量，和 γ_q 表示噪声矢量增益。

合成滤波器 2603 对 LPC 系数编码信息进行 LPC 系数解码，和利用如下方程 (50)，从解码 LPC 系数中生成合成信号 $syn(n)$ 。

$$syn(n) = ex(n) + \sum_{i=1}^{NP} \alpha_q(i) \cdot syn(n-i) \quad \dots (50)$$

这里， α_q 表示解码 LPC 系数，和 NP 表示 LPC 系数的次序。然后，合成滤波器 2603 将以这种方式解码的解码信号 $syn(n)$ 输出到向上取样器 2303。

因此，根据本实施例的声音编码设备，通过在发送方在基本层中利用 CELP 编码输入信号，和在接收方利用 CELP 解码这个编码输入信号，可以以低位速率实现高质基本层。

为了抑制量化失真被人们感觉到，本实施例的编码设备也可以应用在合成滤波器 2603 之后附属连接后置滤波器的配置。图 28 是示出本实施例的基本层解码器的内部配置的例子方块图。将与图 27 中相同的标号指定给图

28 中与图 27 中的那些相同的部分，并且省略对它们的详细描述。

各种类型的配置可用于后置滤波器，以实现量化失真被人们感觉到的抑制，一种典型方法是利用包括通过多路分用器 2601 解码获得的 LPC 系数的共振峰强调滤波器的方法。共振峰强调滤波函数 $H_f(z)$ 用如下方程 (51) 表示。

$$H_f(z) = \frac{A(z/\gamma_n)}{A(z/\gamma_d)} (1 - \mu z^{-1}) \cdots \quad (51)$$

这里， $A(z)$ 表示包括解码 LPC 系数的分析滤波函数，和 γ_n 、 γ_d 和 μ 表示确定滤波器特性的常数。

(第 11 实施例)

图 29 是示出根据本发明第 11 实施例的声音编码设备的频率确定部分的内部配置的例子方块图。将与图 20 中相同的标号指定给图 29 中与图 20 中的那些相同的部分，并且省略对它们的详细描述。图 29 中的频率确定部分 1607 与图 20 中的频率确定部分的不同之处在于，配备了估计误差谱计算器 2801 和确定部分 2802，并且，从基本层解码信号幅度谱 $P(m)$ 中估计估计误差谱 $E'(m)$ ，和利用估计误差谱 $E'(m)$ 和估计听觉掩蔽 $M'(m)$ ，确定增强层编码器 1608 编码的误差谱的频率。

FFT 部分 1901 对从向上取样器 1604 输出的基本层解码信号 $x(n)$ 进行付里叶变换，计算幅度谱 $P(m)$ ，并且将幅度谱 $P(m)$ 输出到估计听觉掩蔽计算器 1902 和估计误差谱计算器 2801。具体地说，FFT 部分 1901 利用如下方程 (33) 计算幅度谱 $P(m)$ 。

估计误差谱计算器 2801 从 FFT 部分 1901 计算的基本层解码信号幅度谱 $P(m)$ 中计算估计误差谱 $E'(m)$ ，并且将估计误差谱 $E'(m)$ 输出到确定部分 2802。估计误差谱 $E'(m)$ 是通过执行使基本层解码信号幅度谱 $P(m)$ 接近平坦的处理计算的。具体地说，估计误差谱计算器 2801 利用如下方程 (52) 计算估计误差谱 $E'(m)$ 。

$$E'(m) = a \cdot P(m)^\gamma \quad \cdots (52)$$

这里， a 和 γ 是大于等于 0 和小于 1 的常数。

利用估计误差谱计算器 2801 获得的估计误差谱 $E'(m)$ 和估计听觉掩蔽计算器 1902 获得的估计听觉掩蔽 $M'(m)$ ，确定部分 1903 确定增强层编码器 1608 用于误差谱编码的频率。

接着，描述本实施例的估计误差谱计算器 2801 计算的估计误差谱。图

30 是示出本实施例的估计误差谱计算器计算的残留误差谱的例子的图形。

如图 30 所示，误差谱 $E(m)$ 的谱线形状比基本层解码信号幅度谱 $P(m)$ 的谱线形状光滑，并且，它的总频带功率较小。因此，通过使幅度谱 $P(m)$ 变平成 γ ($0 < \gamma < 1$) 的功率，和通过乘以 a ($0 < a < 1$) 降低总频带功率，可以提高误差谱估计的精度。

此外，在解码方，声音解码设备 2300 的频率确定部分 2304 的内部配置与图 29 中的编码方频率确定部分 1607 的内部配置相同。

因此，根据本实施例的声音编码设备，通过使从基本层解码信号谱中估计的残留误差谱变平滑，可以使估计误差谱接近残留误差谱，和可以在增强层中有效地编码误差谱。

在本实施例中，已经描述了使用 FFT 的情况，但是，像上述第 9 实施例中那样，MDCT 或其它变换用来取代 FFT 的配置也是可以的。

(第 12 实施例)

图 31 是示出根据本发明第 12 实施例的声音编码设备的频率确定部分的内部配置的例子方块图。将与图 20 中相同的标号指定给图 31 中与图 20 中的那些相同的部分，并且省略对它们的详细描述。图 31 中的频率确定部分 1607 与图 20 中的频率确定部分的不同之处在于，配备了估计听觉掩蔽校正部分 3001 和确定部分 3002，并且，在通过估计听觉掩蔽计算器 1902 从基本层解码信号幅度谱 $P(m)$ 中计算出估计听觉掩蔽 $M'(m)$ 之后，频率确定部分 1607 根据局部解码器 1603 解码参数信息，对这个估计听觉掩蔽 $M'(m)$ 加以校正。

FFT 部分 1901 对从向上取样器 1604 输出的基本层解码信号 $x(n)$ 进行付里叶变换，计算幅度谱 $P(m)$ ，并且将幅度谱 $P(m)$ 输出到估计听觉掩蔽计算器 1902 和确定部分 3002。估计听觉掩蔽计算器 1902 利用基本层解码信号幅度谱 $P(m)$ 计算估计听觉掩蔽 $M'(m)$ ，并且将估计听觉掩蔽 $M'(m)$ 输出到估计听觉掩蔽校正部分 3001。

因此，根据本实施例的声音编码设备，通过使从基本层解码信号谱中估计的残留误差谱变平滑，可以使估计误差谱接近残留误差谱，和可以在增强层中有效地编码误差谱。

利用从局部解码器 1603 输入的基本层解码参数信息，估计听觉掩蔽校正部分 3001 对估计听觉掩蔽计算器 1902 获得的估计听觉掩蔽 $M'(m)$ 加以校正。

这里假设供应从解码 LPC 系数中计算的一阶 PARCOR 系数作为基本层编码信息。一般说来, LPC 系数和 PARCOR 系数代表输入信号谱包络。由于 PARCOR 系数的特性, 随着 PARCOR 系数的阶降低, 谱包络的形状简化了, 并且, 当 PARCOR 系数的阶是 1 时, 指出了频谱的倾斜度。

另一方面, 在音频或语音输入信号的谱特性中, 存在与较高区相反, 功率朝着较低区方向偏置的情况 (例如, 对于元音), 和反过来的情况 (例如, 对于辅音)。基本层解码信号易受这样输入信号谱特性的影响, 并且, 存在过分强调谱功率偏置的倾向。

因此, 在本实施例的声音编码设备中, 通过在估计听觉掩蔽校正部分 3001 中利用前述一阶 PARCOR 系数校正过分强调谱偏置, 可以提高估计听觉掩蔽 $M'(m)$ 的精度。

估计听觉掩蔽校正部分 3001 利用如下方程 (53), 从基本层编码器 1602 输出的一阶 PARCOR 系数 $k(1)$ 中计算校正滤波函数 $H_k(z)$ 。

$$H_k(z) = 1 - \beta \cdot k(1) \cdot z^{-1} \quad \dots (53)$$

这里, β 表示小于 1 的正常数。接着, 估计听觉掩蔽校正部分 3001 利用如下方程 (54), 计算校正滤波函数 $H_k(z)$ 的幅度特性 $K(m)$ 。

$$K(m) = \left| 1 - \beta \cdot k(1) \cdot e^{-j \frac{2\pi m}{M}} \right| \quad \dots (54)$$

然后, 估计听觉掩蔽校正部分 3001 利用如下方程 (55), 从校正滤波函数幅度特性 $K(m)$ 中计算校正估计听觉掩蔽 $M''(m)$ 。

$$M''(m) = K(m) \cdot M'(m) \quad \dots (55)$$

然后, 取代估计听觉掩蔽 $M'(m)$, 估计听觉掩蔽校正部分 3001 将校正估计听觉掩蔽 $M''(m)$ 输出到确定部分 3002。

利用基本层解码信号幅度谱 $P(m)$ 和从估计听觉掩蔽校正部分 3001 输出的校正估计听觉掩蔽 $M''(m)$, 确定部分 3002 确定增强层编码器 1608 用于误差谱编码的频率。

因此, 根据本实施例的声音编码设备, 通过利用掩蔽效应特性, 从输入信号谱中计算听觉掩蔽, 和在增强层编码中进行使量化失真不超过掩蔽值的量化, 在不会使质量下降的情况下, 可以减少经受量化的 MDCT 系数的个数, 和以低位速率进行高质编码。

因此, 根据本实施例的声音编码设备, 通过根据基本层编码器解码参数

信息对估计听觉掩蔽加入校正，可以提高估计听觉掩蔽的精度，和在增强层中进行有效误差谱编码。

此外，在解码方，声音解码设备 2300 的频率确定部分 2304 的内部配置与图 31 中的编码方频率确定部分 1607 的内部配置相同。

对于本实施例的频率确定部分 1607，还可以应用将本实施例和第 11 实施例组合在一起的配置。图 32 是示出本实施例的声音编码设备的频率确定部分的内部配置的例子方块图。将与图 20 中相同的标号指定给图 32 中与图 20 中的那些相同的部分，并且省略对它们的详细描述。

FFT 部分 1901 对从向上取样器 1604 输出的基本层解码信号 $x(n)$ 进行付里叶变换，计算幅度谱 $P(m)$ ，并且将幅度谱 $P(m)$ 输出到估计听觉掩蔽计算器 1902 和估计误差谱计算器 2801。

估计听觉掩蔽计算器 1902 利用基本层解码信号幅度谱 $P(m)$ 计算估计听觉掩蔽 $M'(m)$ ，并且将估计听觉掩蔽 $M'(m)$ 输出到估计听觉掩蔽校正部分 3001。

在估计听觉掩蔽校正部分 3001 中，应用从局部解码器 1603 输入的基本层解码参数信息来校正估计听觉掩蔽计算器 1902 获得的估计听觉掩蔽 $M'(m)$ 。

估计误差谱计算器 2801 从 FFT 部分 1901 计算的基本层解码信号幅度谱 $P(m)$ 中计算估计误差谱 $E'(m)$ ，并且将估计误差谱 $E'(m)$ 输出到确定部分 3101。

利用估计误差谱计算器 2801 估计的估计误差谱 $E'(m)$ 和从估计听觉掩蔽校正部分 3001 输出的校正听觉掩蔽 $M''(m)$ ，确定部分 3101 确定增强层编码器 1608 进行误差谱编码的频率。

在本实施例中，已经描述了使用 FFT 的情况，但是，像上述第 9 实施例中那样，MDCT 或其它变换技术用来取代 FFT 的配置也是可以的。

(第 13 实施例)

图 33 是示出根据本发明第 13 实施例的声音编码设备的增强层编码器的内部配置的例子方块图。将与图 22 中相同的标号指定给图 33 中与图 22 中的那些相同的部分，并且省略对它们的详细描述。图 33 中的增强层编码器与图 22 中的增强层编码器的不同之处在于，配备了定序部分 3201 和 MDCT 系数量化器 3202，并且，按照估计失真值 $D(m)$ 的数量，通过频率对频率确定部分

1607 供应的频率进行加权。

在图 33 中，MDCT 部分 2101 将从减法器 1606 输出的输入信号乘以分析窗，然后，进行 MDCT（改进离散余弦变换）处理以获得 MDCT 系数，并且将 MDCT 系数输出到 MDCT 系数量化器 3202。

定序部分 3201 接收频率确定部分 1607 获得的频率信息，并且计算每个频率的估计误差谱 $E'(m)$ 超过估计听觉掩蔽 $M'(m)$ 的数量（下文称为“估计失真值”） $D(m)$ 。这个估计失真值 $D(m)$ 由如下方程（56）定义。

$$D(m) = E'(m) - M'(m) \quad \dots (56)$$

这里，定序部分 3201 只计算满足如下方程（57）的估计失真值 $D(m)$ 。

$$E'(m) - M'(m) > 0 \quad \dots (57)$$

然后，定序部分 3201 按从高到低估计失真值 $D(m)$ 次序进行定序，并且将相应频率信息输出到 MDCT 系数量化器 3202。MDCT 系数量化器 3202 进行量化，根据估计失真值 $D(m)$ ，将位成正比地分配给位于按从高到低估计失真值 $D(m)$ 次序排列的频率上的误差谱 $E(m)$ 。

作为一个例子，这里描述从频率确定部分发送的频率和估计失真值像图 34 所示那样的情况。图 34 是示出本实施例的定序部分排序估计失真值的例子的图形。

定序部分 3201 根据图 34 中的信息，按从高到低估计失真值 $D(m)$ 次序重新排列频率。在本例中，作为定序部分 3201 的处理结果获得的频率 m 次序是：7、8、4、9、1、11、3、12。定序部分 3201 将这个定序信息输出到 MDCT 系数量化器 3202。

在 MDCT 部分 2101 给出的误差谱 $E(m)$ 内，MDCT 系数量化器 3202 根据定序部分 3201 给出的定序信息，量化 $E(7)$ 、 $E(8)$ 、 $E(4)$ 、 $E(9)$ 、 $E(1)$ 、 $E(11)$ 、 $E(3)$ 、 $E(12)$ 。

同时，在该次序的开头分配许多用于误差谱量化的位，和朝着该次序的末端分配逐渐减少的位。也就是说，频率的估计失真值 $D(m)$ 越大，分配用于误差谱量化的位就越多，频率的估计失真值 $D(m)$ 越小，分配用于误差谱量化的位就越少。

例如，可以进行如下位分配：对于 $E(7)$ ，8 个位；对于 $E(8)$ 和 $E(4)$ ，7 个位；对于 $E(9)$ 和 $E(1)$ ，6 个位；对于 $E(11)$ 、 $E(3)$ 和 $E(12)$ ，5 个位。这样，根据估计失真值 $D(m)$ 进行自适应位分配提高了量化效率。

当应用矢量量化时，增强层编码器 1608 从位于该次序的开头上的误差谱开始依次配置矢量，并且对各自矢量进行矢量量化。同时，进行矢量配置和量化位分配，以便对于位于该次序的开头上的误差谱，分配的位较多，和对于位于该次序的末端上的误差谱，分配的位较少。在图 34 中的例子中，配置了三个矢量 - 二维、二维和四维，以及 $V1 = (E(7), E(8))$ 、 $V2 = (E(4), E(9))$ 和 $V3 = (E(1), E(11), E(3), E(12))$ ，并且，位分配是：对于 $V1$ ，10 个位；对于 $V2$ ，8 个位；和对于 $V3$ ，8 个位。

因此，根据本实施例的声音编码设备，通过在增强层编码中进行将大量信息分配给估计误差谱超过估计听觉掩蔽的数量大的频率的编码，可以实现量化效率的提高。

现在描述解码方。图 35 是示出根据本发明第 13 实施例的声音解码设备的增强层解码器的内部配置的例子方块图。将与图 25 中相同的标号指定给图 35 中与图 25 中的那些相同的部分，并且省略对它们的详细描述。图 35 中的增强层解码器 2305 与图 25 中的增强层解码器的不同之处在于，配备了定序部分 3401 和 MDCT 系数解码器 3402，并且，按照估计失真值 $D(m)$ 的数量定序频率确定部分 2304 供应的频率。

定序部分 3401 利用上面的方程 (56) 计算估计失真值 $D(m)$ 。定序部分 3401 具有与上述定序部分 3201 相同的配置。通过这种配置，可以解码能够进行自适应位分配和提高量化效率的上述声音编码方法的编码信息。

MDCT 系数解码器 3402 利用按照估计失真值 $D(m)$ 的数量定序的频率信息，解码从多路分用器 2301 输出的第二编码信息。具体地说，MDCT 系数解码器 3402 定位与频率确定部分 2304 供应的频率相对应的解码 MDCT 系数，并且，对于其它频率填上零。然后，IMDCT 部分 2402 对从 MDCT 系数解码器 2401 获得的 MDCT 系数进行逆 MDCT 处理，生成时域信号。

叠加加法器 2403 为了组合将前述信号乘以一个窗口函数，和叠加在前一帧和当前帧中解码的时域信号，进行相加，并且生成输出信号。叠加加法器 2403 将这个输出信号输出到加法器 2306。

因此，根据本实施例的声音解码设备，通过在增强层编码中进行按照估计误差谱超过估计听觉掩蔽的数量进行自适应位分配的矢量量化，可以实现量化效率的提高。

(第 14 实施例)

图 36 是示出根据本发明第 14 实施例的声音编码设备的增强层编码器的内部配置的例子方块图。将与图 22 中相同的标号指定给图 36 中与图 22 中的那些相同的部分，并且省略对它们的详细描述。图 36 中的增强层编码器与图 22 中的增强层编码器的不同之处在于，配备了固定频带指定部分 3501 和 MDCT 系数量化器 3502，并且，与从频率确定部分 1607 中获得的频率一起量化包括在事先指定的频带中的 MDCT 系数。

在图 36 中，在固定频带指定部分 3501 中事先设置就听觉感觉而言重要的频带。这里假设对于包括在所设频带中的频率，设置“ $m = 15, 16$ ”。

MDCT 系数量化器 3502 在来自 MDCT 部分 2101 的输入信号中，利用从频率确定部分 1607 输出的听觉掩蔽将输入信号分类成要量化的系数和不要量化的系数，并且，编码要量化的系数，以及固定频带指定部分 3501 设置的频带中的系数。

假设相关频率成为如图 34 所示那样，MDCT 系数量化器 3502 量化误差谱 $E(1)$ 、 $E(3)$ 、 $E(4)$ 、 $E(7)$ 、 $E(8)$ 、 $E(9)$ 、 $E(11)$ 、 $E(12)$ 和固定频带指定部分 3501 指定的频率的误差谱 $E(15)$ 、 $E(16)$ 。

因此，根据本实施例的声音编码设备，通过强迫量化不可能选作量化的对象、但从听觉的观点来说重要的频带，即使不选择应该真正选为编码的对象的频率，也必定可以量化位于包括在从听觉的观点来说重要的频带中的频率上的误差谱，从而使质量得到提高。

现在描述解码方。图 37 是示出根据本发明第 14 实施例的声音解码设备的增强层解码器的内部配置的例子方块图。将与图 25 中相同的标号指定给图 37 中与图 25 中的那些相同的部分，并且省略对它们的详细描述。图 37 中的增强层解码器与图 25 中的增强层解码器的不同之处在于，配备了固定频带指定部分 3601 和 MDCT 系数解码器 3602，并且，与从频率确定部分 2304 中获得的频率一起解码包括在事先指定的频带中的 MDCT 系数。

在图 37 中，在固定频带指定部分 3601 中事先设置就听觉感觉而言重要的频带。

MDCT 系数解码器 3602 根据从频率确定部分 1607 输出的经过解码的误差谱频率，解码从多路分用器 2301 输出的第二编码信息中量化的 MDCT 系数。具体地说，MDCT 系数解码器 3602 定位与频率确定部分 2304 和固定频带指定部分 3501 所指的频率相对应的解码 MDCT 系数，并且，对于其它频率填上零。

IMDCT 部分 2402 对从 MDCT 系数解码器 3601 输出的 MDCT 系数进行逆 MDCT 处理, 生成时域信号, 并且将这个信号输出到叠加加法器 2403。

因此, 根据本实施例的声音解码设备, 通过解码包括在事先指定的频带中的 MDCT 系数, 可以解码其中已经强迫量化了不可能选作量化的对象、但从听觉的观点来说重要的频带的信号, 并且, 即使不选择在编码方应该真正选为编码的对象的频率, 也必定可以量化位于包括在从听觉的观点来说重要的频带中的频率上的误差谱, 从而使质量得到提高。

对于本实施例的增强层编码器和增强层解码器, 还可以应用将本实施例和第 13 实施例组合在一起的配置。图 38 是示出本实施例的声音编码设备的频率确定部分的内部配置的例子方块图。将与图 22 中相同的标号指定给图 38 中与图 22 中的那些相同的部分, 并且省略对它们的详细描述。

在图 38 中, MDCT 部分 2101 将从减法器 1606 输出的输入信号乘以分析窗, 然后, 进行 MDCT (改进离散余弦变换) 处理以获得 MDCT 系数, 并且将 MDCT 系数输出到 MDCT 系数量化器 3701。

定序部分 3201 接收频率确定部分 1607 获得的频率信息, 并且, 计算每个频率的估计误差谱 $E'(m)$ 超过估计听觉掩蔽 $M'(m)$ 的数量 (下文称为“估计失真值”) $D(m)$ 。

在固定频带指定部分 3501 中事先设置就听觉感觉而言重要的频带。

MDCT 系数量化器 3701 进行量化, 根据按照估计失真值 $D(m)$ 定序的频率信息, 将位成正比地分配给位于按从高到低估计失真值 $D(m)$ 次序排列的频率上的误差谱 $E(m)$ 。MDCT 系数量化器 3701 还编码固定频带指定部分 3501 设置的频带中的系数。

现在描述解码方。图 39 是示出根据本发明第 14 实施例的声音解码设备的增强层解码器的内部配置的例子方块图。将与图 25 中相同的标号指定给图 39 中与图 25 中的那些相同的部分, 并且省略对它们的详细描述。

在图 39 中, 定序部分 3401 接收频率确定部分 2304 获得的频率信息, 并且, 计算每个频率的估计误差谱 $E'(m)$ 超过估计听觉掩蔽 $M'(m)$ 的数量 (下文称为“估计失真值”) $D(m)$ 。

然后, 定序部分 3401 按从高到低估计失真值 $D(m)$ 次序进行定序, 并且将相应频率信息输出到 MDCT 系数解码器 3801。在固定频带指定部分 3601 中事先设置就听觉感觉而言重要的频带。

MDCT 系数解码器 3801 根据从定序部分 3401 输出的经过解码的误差谱频率, 解码从多路分用器 2301 输出的第二编码信息中量化的 MDCT 系数。具体地说, MDCT 系数解码器 3801 定位与定序部分 3401 和固定频带指定部分 3601 所指的频率相对应的解码 MDCT 系数, 并且, 对于其它频率填上零。

IMDCT 部分 2402 对从 MDCT 系数解码器 3801 输出的 MDCT 系数进行逆 MDCT 处理, 生成时域信号, 并且将这个信号输出到叠加加法器 2403。

(第 15 实施例)

现在参照附图描述本发明的第 15 实施例。图 40 是示出根据本发明第 15 实施例的通信设备的配置的方块图。本实施例的特征是图 40 中的信号处理设备 3903 被配置成如上述第 1 到第 14 实施例所示的声音编码设备之一。

如图 40 所示, 根据本发明第 15 实施例的通信设备 3900 包括输入设备 3901、A/D 转换设备 3902 和与网络 3904 连接的信号处理设备 3903。

A/D 转换设备 3902 与输入设备 3901 的输出端相连接。信号处理设备 3903 的输入端与 A/D 转换设备 3902 的输出端相连接。信号处理设备 3903 的输出端与网络 3904 相连接。

输入设备 3901 将人耳朵可听见的声波转换成作为电信号的模拟信号, 并且将这个模拟信号供应给 A/D 转换设备 3902。A/D 转换设备 3902 将模拟信号转换成数字信号, 并且将这个数字信号供应给信号处理设备 3903。信号处理设备 3903 编码输入数字信号和生成代码, 并且将这个代码输出到网络 3904。

因此, 根据本发明这个实施例的通信设备, 可以在通信过程中获得像上述第 1 到第 14 实施例所示那样的效果, 并且, 可以提供用少量的位有效编码声信号的声音编码设备。

(第 16 实施例)

现在参照附图描述本发明的第 16 实施例。图 41 是示出根据本发明第 16 实施例的通信设备的配置的方块图。本实施例的特征是图 41 中的信号处理设备 4003 被配置成如上述第 1 到第 14 实施例所示的声音解码设备之一。

如图 41 所示, 根据本发明第 16 实施例的通信设备 4000 包括与网络 4001 连接的接收设备 4002、信号处理设备 4003、D/A 转换设备 4004 和输出设备 4005。

接收设备 4002 与网络 4001 相连接。信号处理设备 4003 的输入端与接收设备 4002 的输出端相连接。D/A 转换设备 4004 的输入端与信号处理设备 4003

的输出端相连接。输出设备 4005 的输入端与 D/A 转换设备 4004 的输出端相连接。

接收设备 4002 接收来自网络 4001 的数字编码声信号，生成数字接收声信号，并且将这个接收声信号供应给信号处理设备 4003。信号处理设备 4003 接收来自接收设备 4002 的接收声信号，对这个接收声信号进行解码处理和生成数字解码声信号，并且将这个数字解码声信号供应给 D/A 转换设备 4004。D/A 转换设备 4004 转换来自信号处理设备 4003 的数字解码声信号和生成模拟解码语音信号，并且将这个模拟解码语音信号供应给输出设备 4005。输出设备 4005 将作为电信号的模拟解码语音信号转换成空气振动，并且像声波那样输出这些空气振动，以便人的耳朵可听见。

因此，根据本实施例的通信设备，可以在通信过程中获得像上述第 1 到第 14 实施例所示那样的效果，并且，可以解码用少量的位有效编码的声信号，从而输出良好的声信号。

(第 17 实施例)

现在参照附图描述本发明的第 17 实施例。图 42 是示出根据本发明第 17 实施例的通信设备的配置的方块图。本实施例的特征是图 42 中的信号处理设备 4103 被配置成如上述第 1 到第 14 实施例所示的声音编码设备之一。

如图 42 所示，根据本发明第 17 实施例的通信设备 4100 包括输入设备 4101、A/D 转换设备 4102 和信号处理设备 4103、RF (射频) 调制设备 4104 和天线 4105。

输入设备 4101 将人耳朵可听见的声波转换成作为电信号的模拟信号，并且将这个模拟信号供应给 A/D 转换设备 4102。A/D 转换设备 4102 将模拟信号转换成数字信号，并且将这个数字信号供应给信号处理设备 4103。信号处理设备 4103 编码输入数字信号和生成编码声信号，并且将这个编码声信号输出到 RF 调制设备 4104。RF 调制设备 4104 调制编码声信号和生成调制编码声信号，并且将这个调制编码声信号供应给天线 4105。天线 4105 发送该调制编码声信号作为无线电波。

因此，根据本实施例的通信设备，可以在无线电通信过程中获得像上述第 1 到第 14 实施例所示那样的效果，并且，可以用少量的位有效编码声信号。

(第 18 实施例)

现在参照附图描述本发明的第 18 实施例。图 43 是示出根据本发明第 18

实施例的通信设备的配置的方块图。本实施例的特征是图 43 中的信号处理设备 4203 被配置成如上述第 1 到第 14 实施例所示的声音解码设备之一。

如图 43 所示, 根据本发明第 18 实施例的通信设备 4200 包括天线 4201、RF 解调设备 4202、信号处理设备 4203、D/A 转换设备 4204 和输出设备 4205。

天线 4201 接收作为无线电波的数字编码声信号, 生成作为电信号的数字接收编码声信号, 并且将这个数字接收编码声信号供应给 RF 解调设备 4202。RF 解调设备 4202 解调来自天线 4201 的接收编码声信号和生成解调编码声信号, 并且将这个解调编码声信号供应给信号处理设备 4203。

信号处理设备 4203 接收来自 RF 解调设备 4202 的数字解调编码声信号, 进行解码处理和生成数字解码声信号, 并且将这个数字解码声信号供应给 D/A 转换设备 4204。D/A 转换设备 4204 转换来自信号处理设备 4203 的数字解码声信号和生成模拟解码语音信号, 并且将这个模拟解码语音信号供应给输出设备 4205。输出设备 4205 将作为电信号的模拟解码语音信号转换成空气振动, 并且像声波那样输出这些空气振动, 以便人的耳朵可听见。

因此, 根据本实施例的通信设备, 可以在无线电通信过程中获得像上述第 1 到第 14 实施例所示那样的效果, 并且, 可以解码用少量的位有效编码的声信号, 从而输出良好的声信号。

本发明可应用于使用音频信号的接收设备、接收解码设备、或语音信号解码设备。本发明还可应用于移动台设备或基站设备。

本发明不局限于上述的实施例, 并且, 在不偏离本发明范围的情况下, 可以进行各种各样的改变和改进。例如, 在上面的实施例中, 已经描述了将本发明作为信号处理设备来实现的情况, 但是, 本发明不局限于此, 并且, 也可以将这种信号处理方法作为软件来实现。

例如, 事先将执行上述信号处理方法的程序存储在 ROM (只读存储器) 中, 和由 CPU (中央处理单元) 执行这个程序也是可以的。

将执行上述信号处理方法的程序存储在计算机可读存储媒体中, 将存储在存储媒体中的程序记录在计算机的 RAM (随机访问存储器) 中, 和按照那个程序操作计算机也是可以的。

在上面的描述中, 已经描述了 MDCT 用作从时域变换到频域的方法, 但是, 本发明不局限于此, 只要是正交的, 可以应用任何变换方法。例如, 也可以应用离散付里叶变换、离散余弦变换或小波变换方法。

本发明可应用于使用音频信号的接收设备、接收解码设备、或语音信号解码设备。本发明还可应用于移动台设备或基站设备。

从上面的描述中可清楚看出，根据本发明的编码设备，解码设备、编码方法和解码方法，通过利用从基本层编码信息中获得的信息进行增强层编码，即使在语音占优势和在背景中叠加了音乐或环境声音的信号的情况下，也可以以低位速率进行高质编码。

本申请基于2002年4月26日提出的日本专利申请第2002-127541号和2002年9月12日提出的日本专利申请第2002-267436号，特此全文引用，以供参考。

工业可应用性

本发明适用于编码和解码语音信号的设备 and 通信设备。

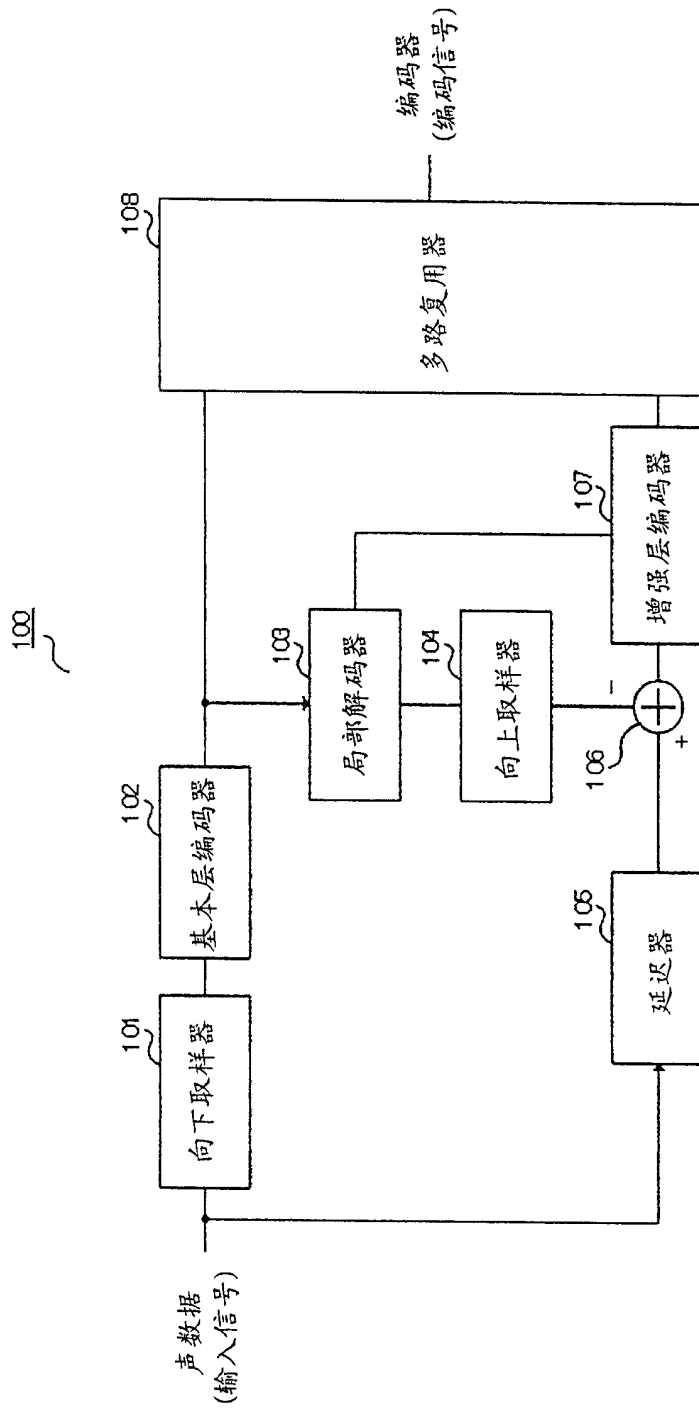


图 1

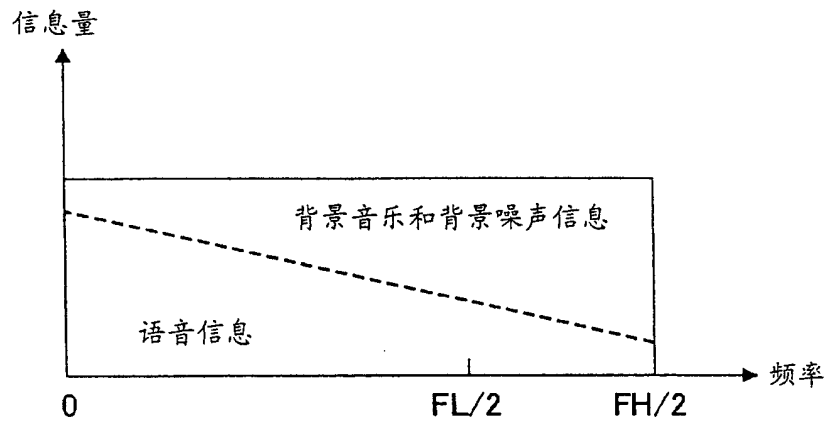


图 2

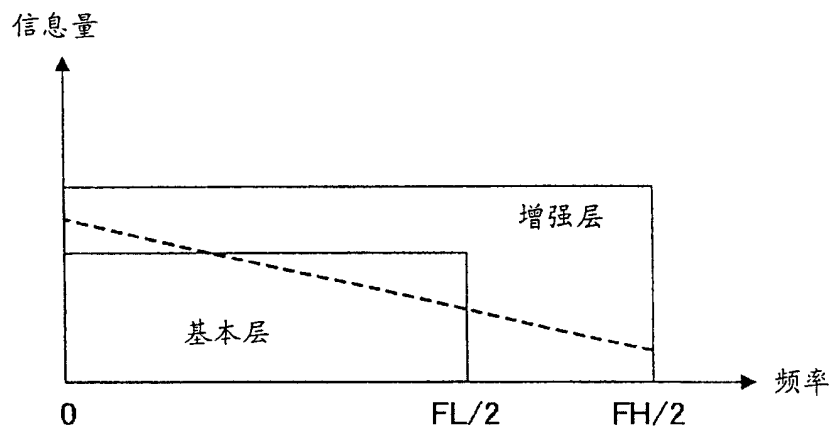


图 3

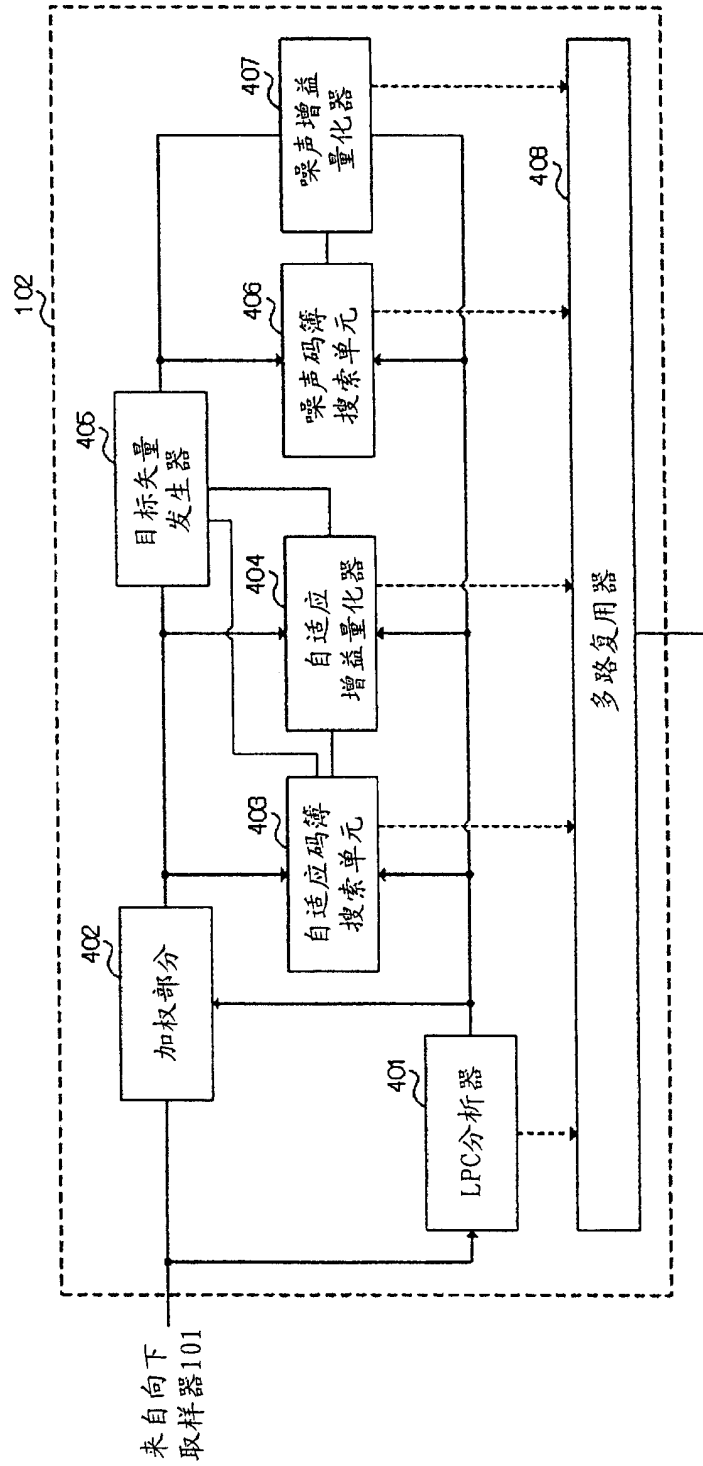


图 4

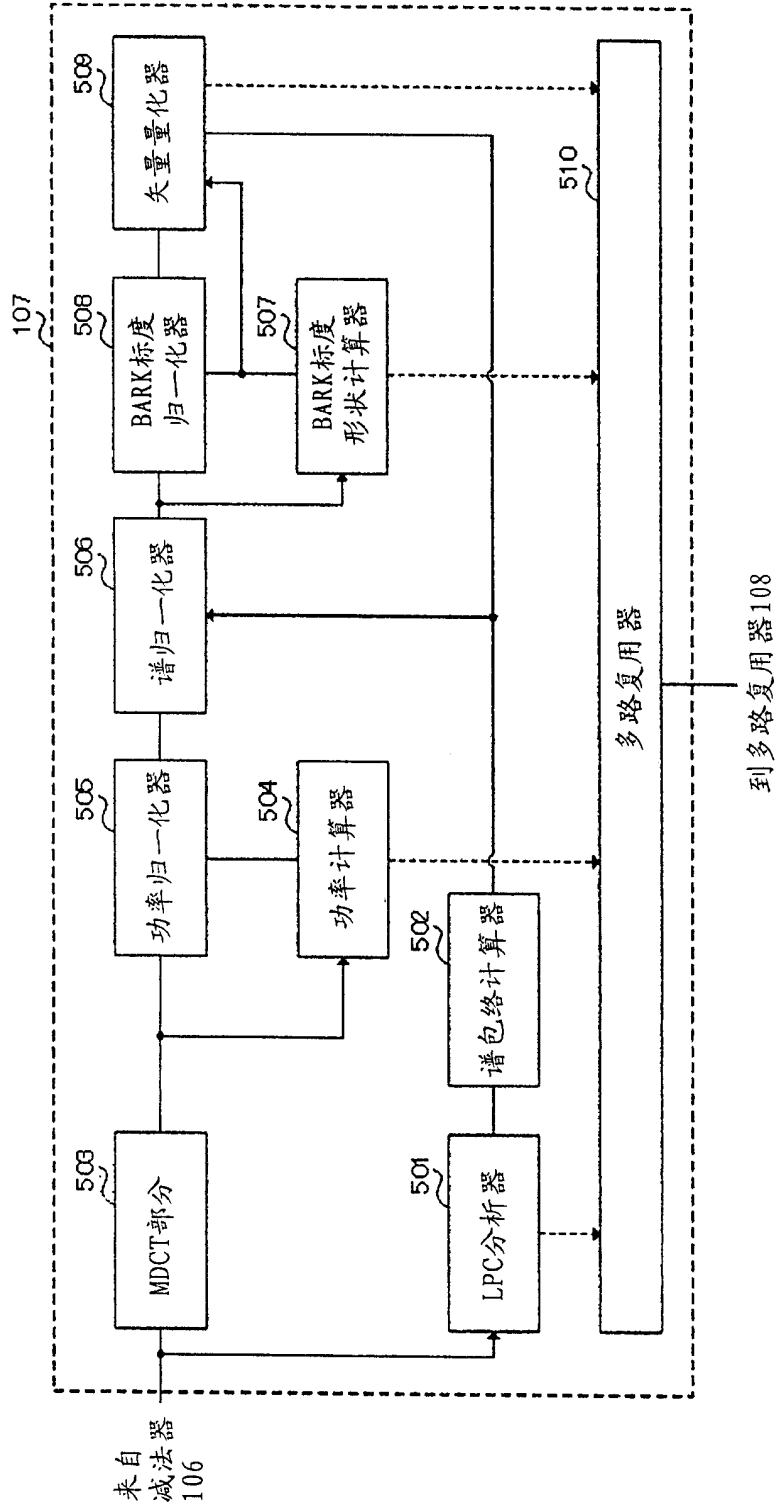


图 5

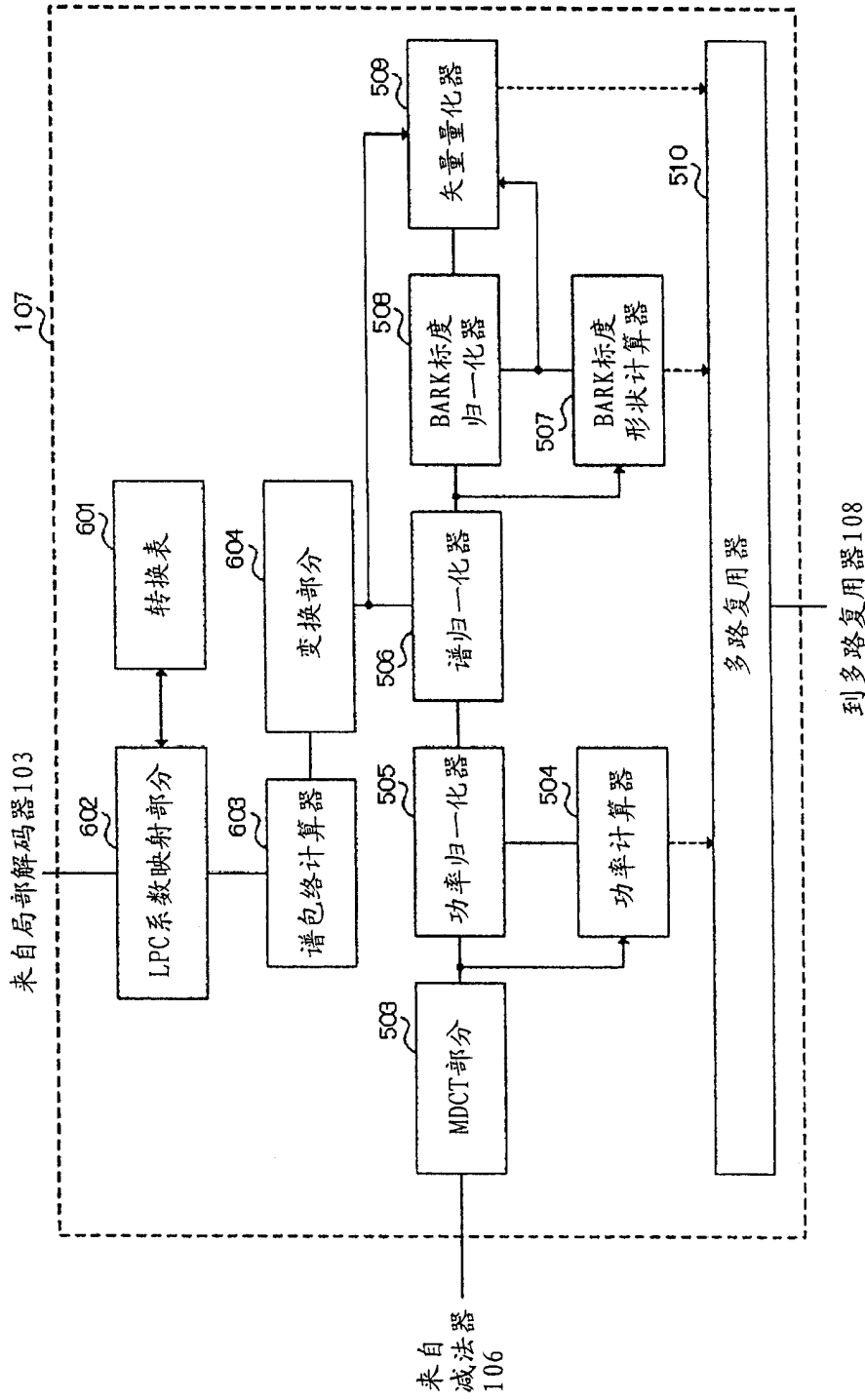


图 6

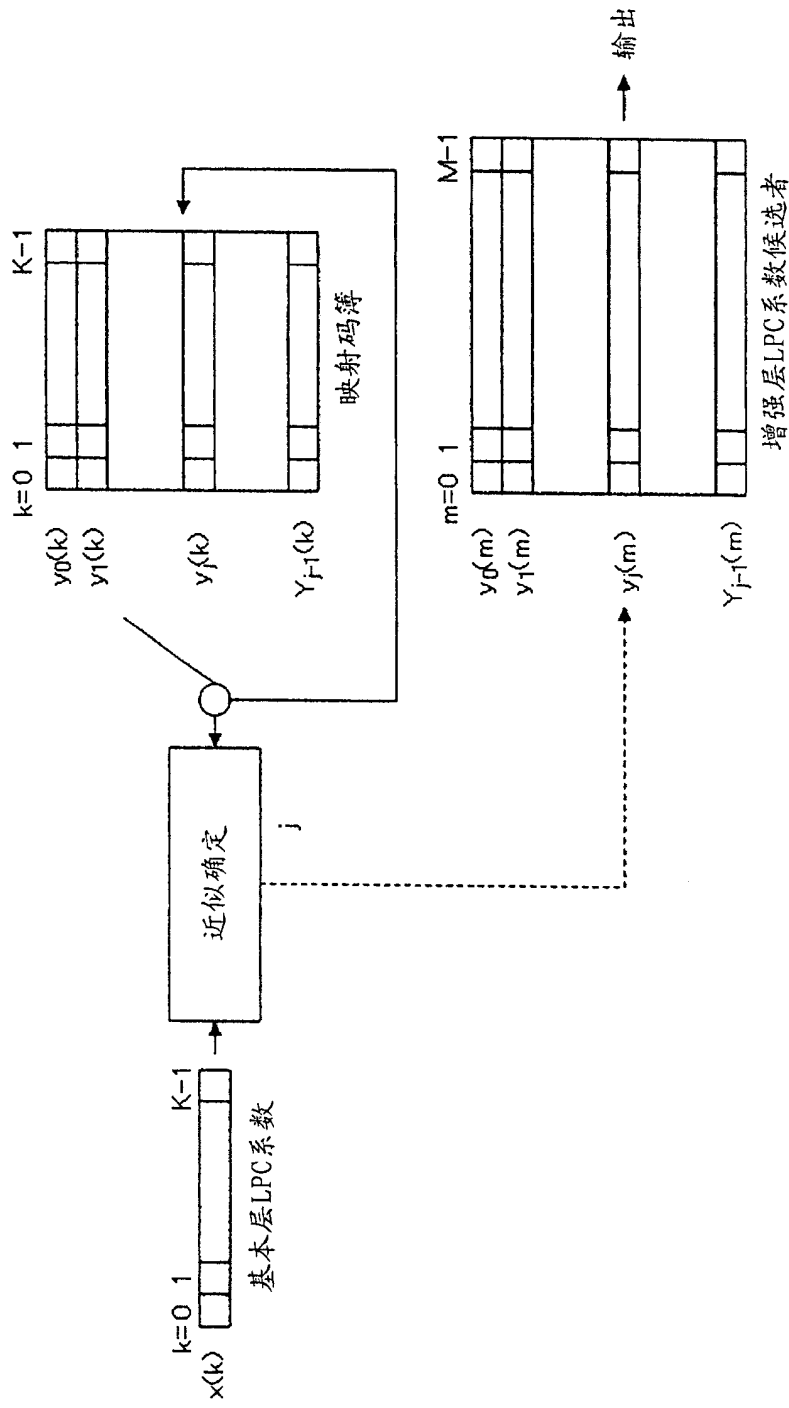


图 7

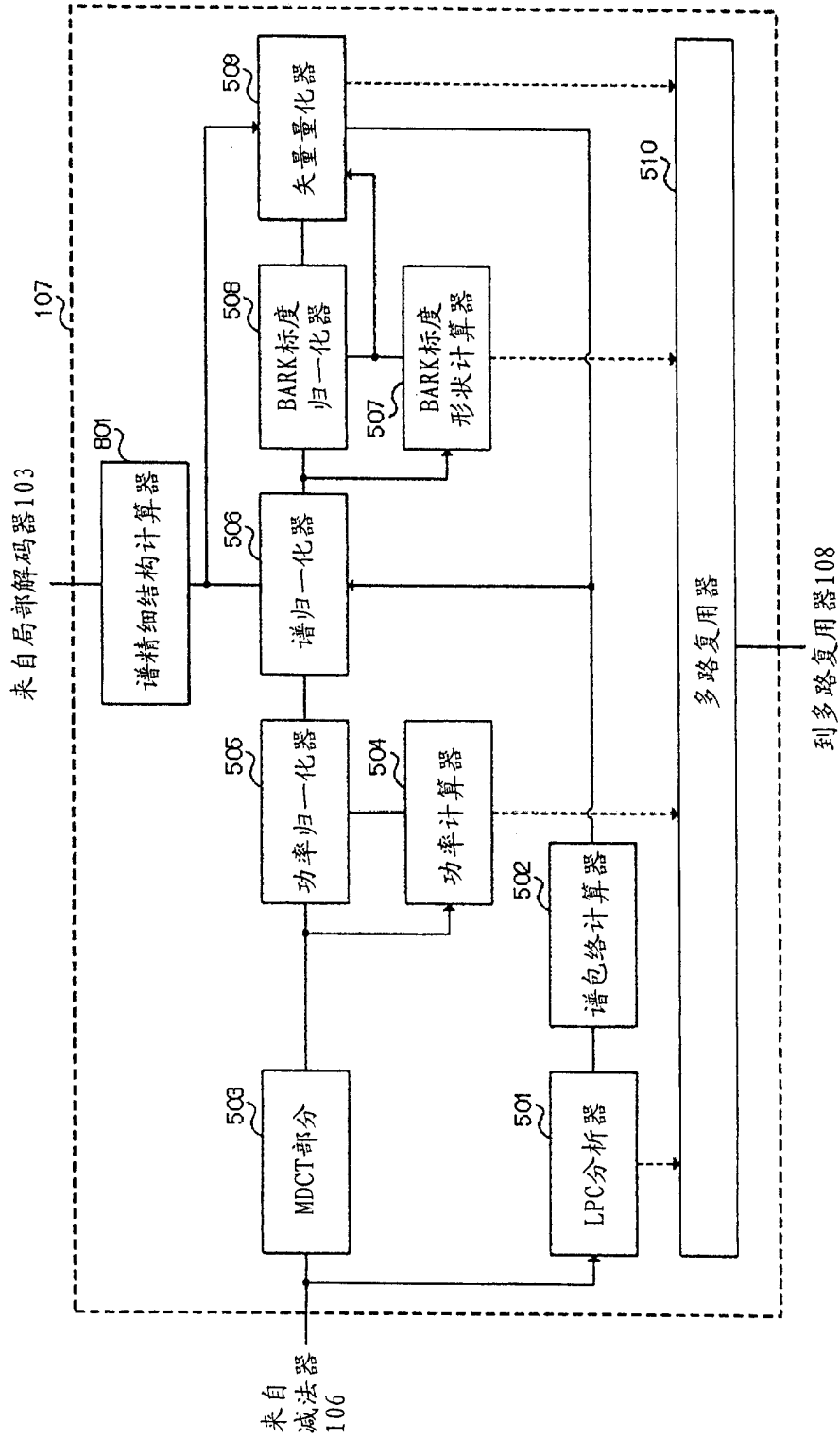


图 8

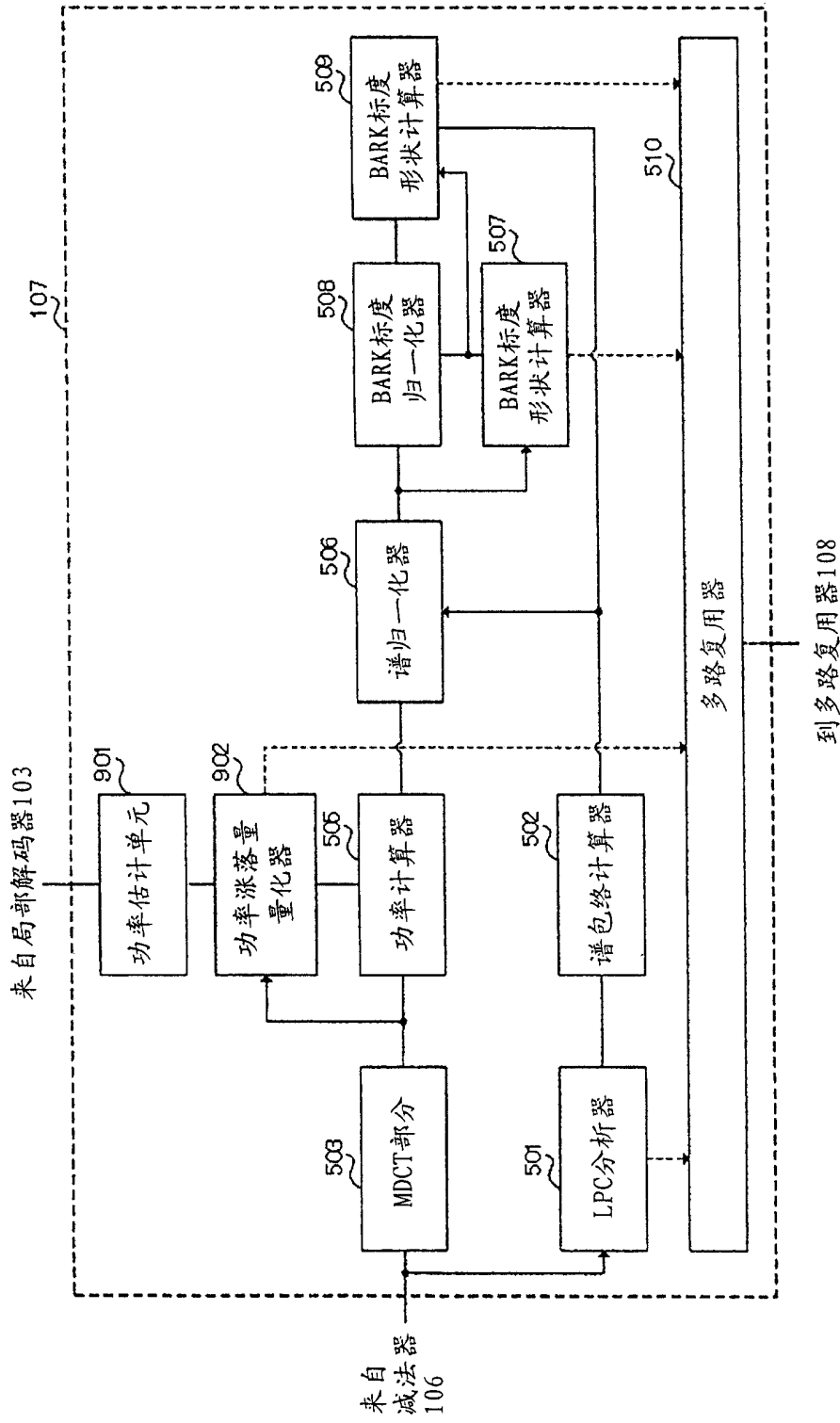


图 9

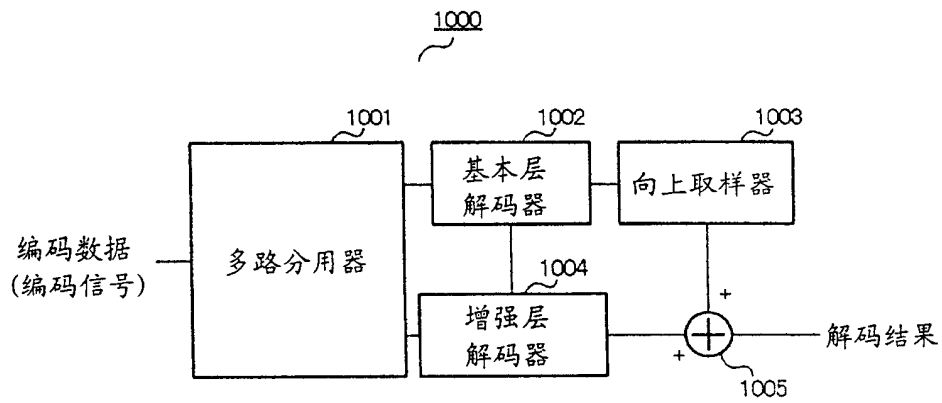


图 10

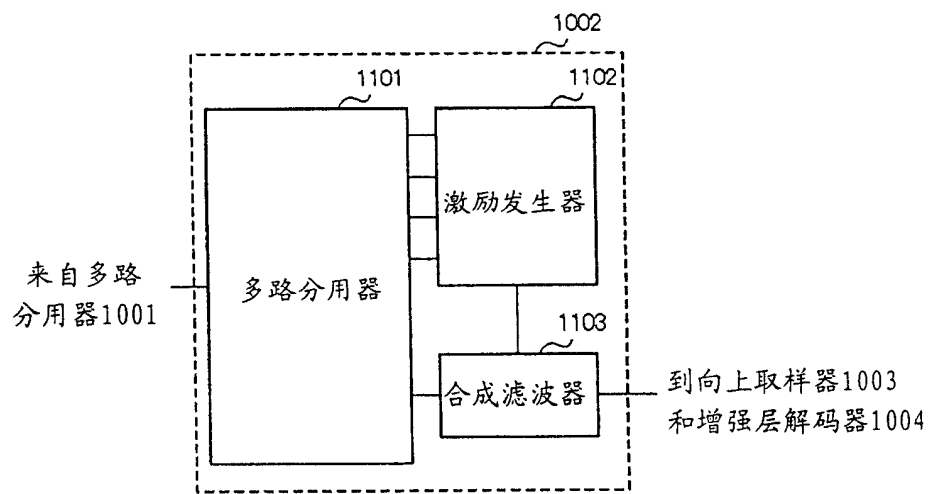


图 11

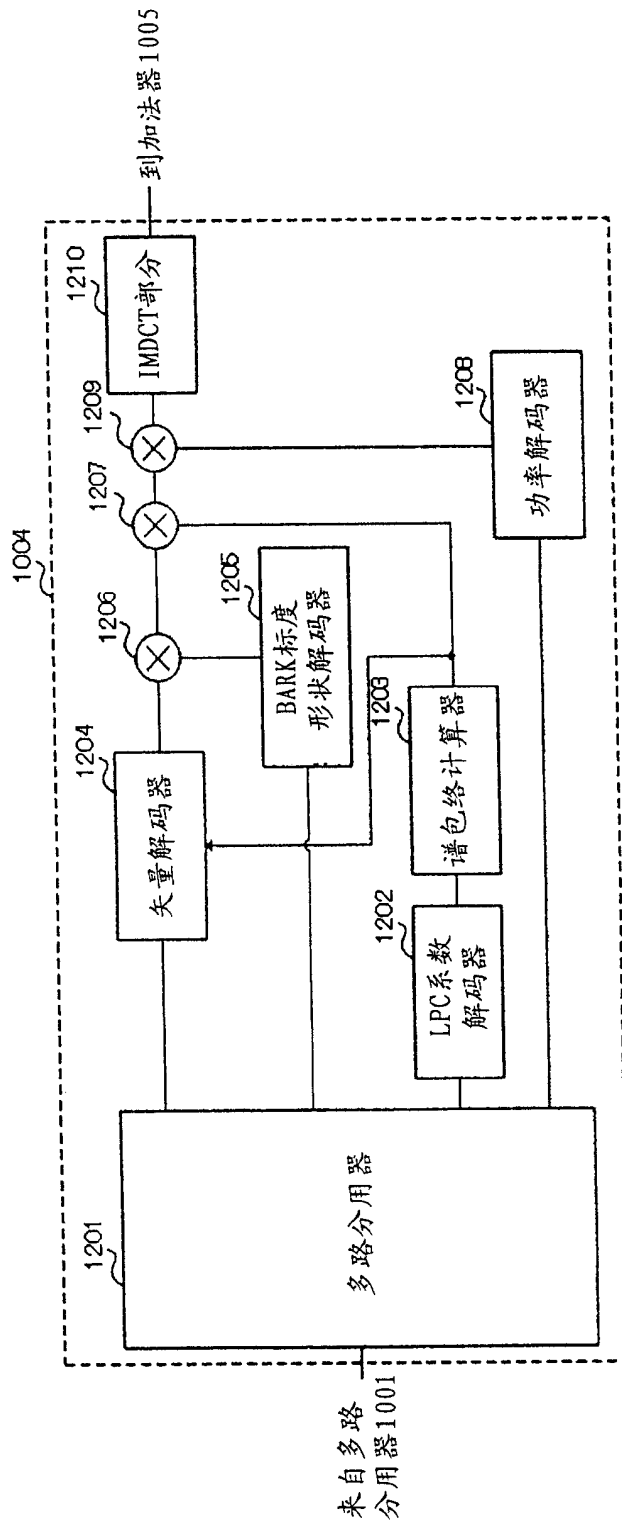


图 12

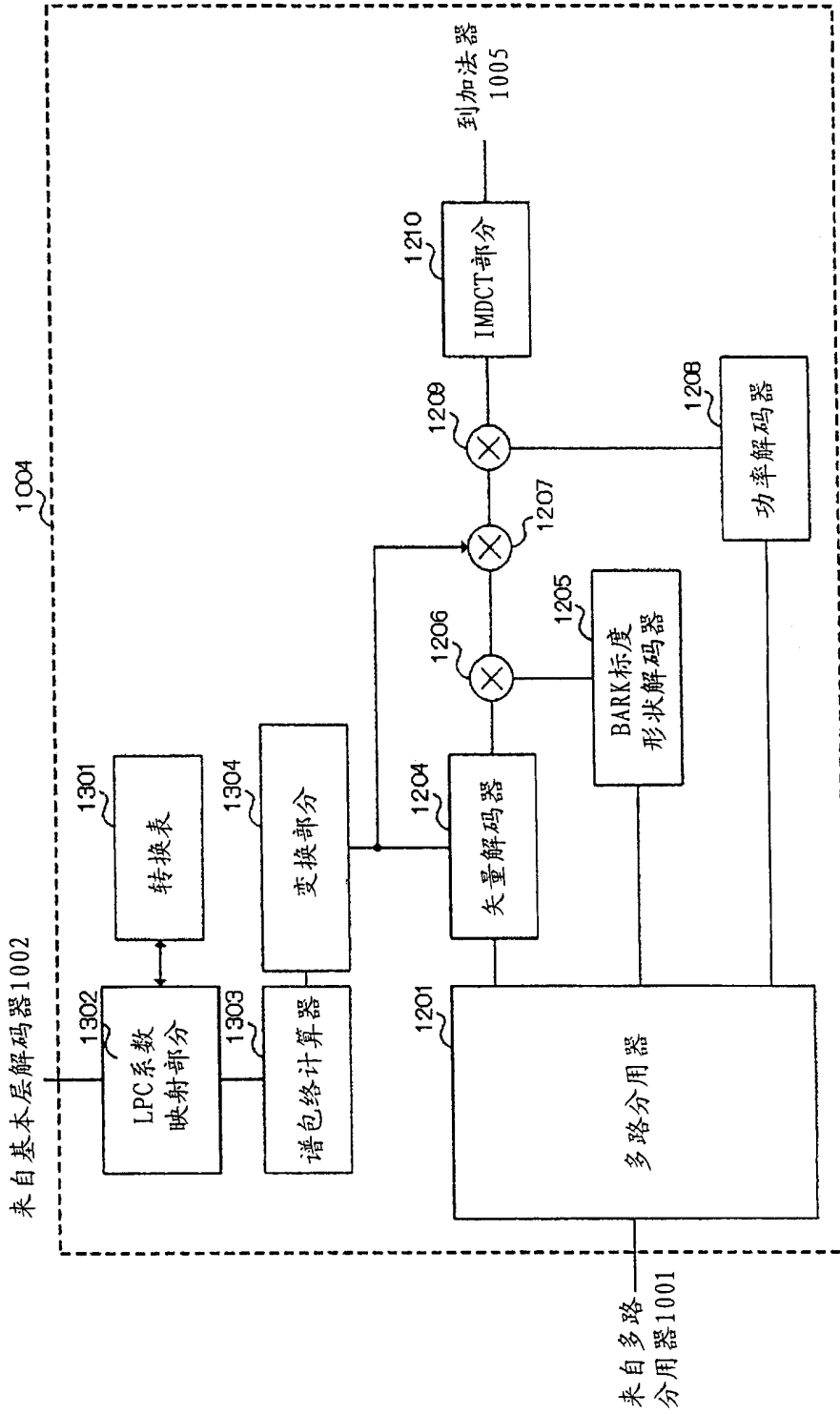


图 13

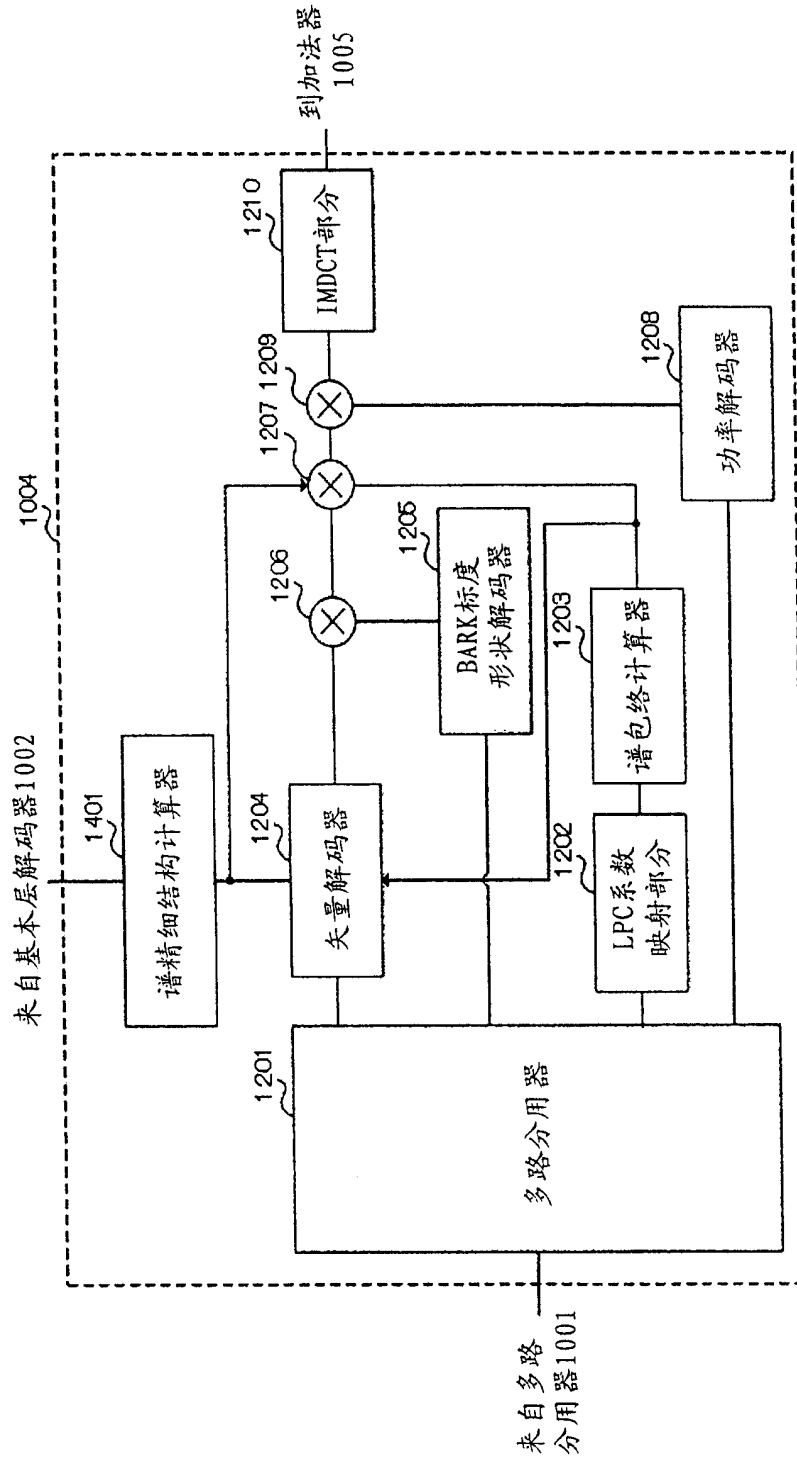


图 14

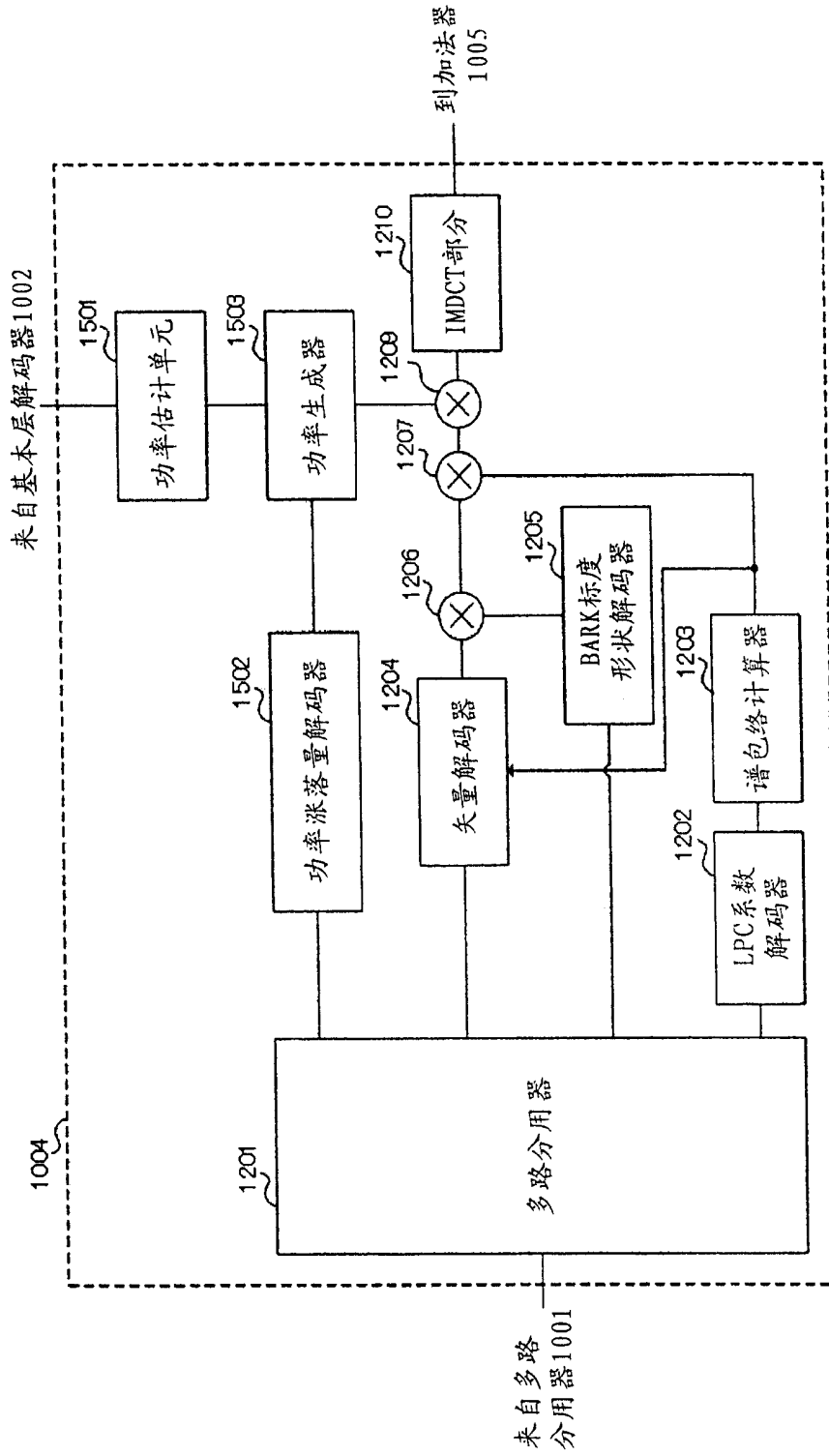


图 15

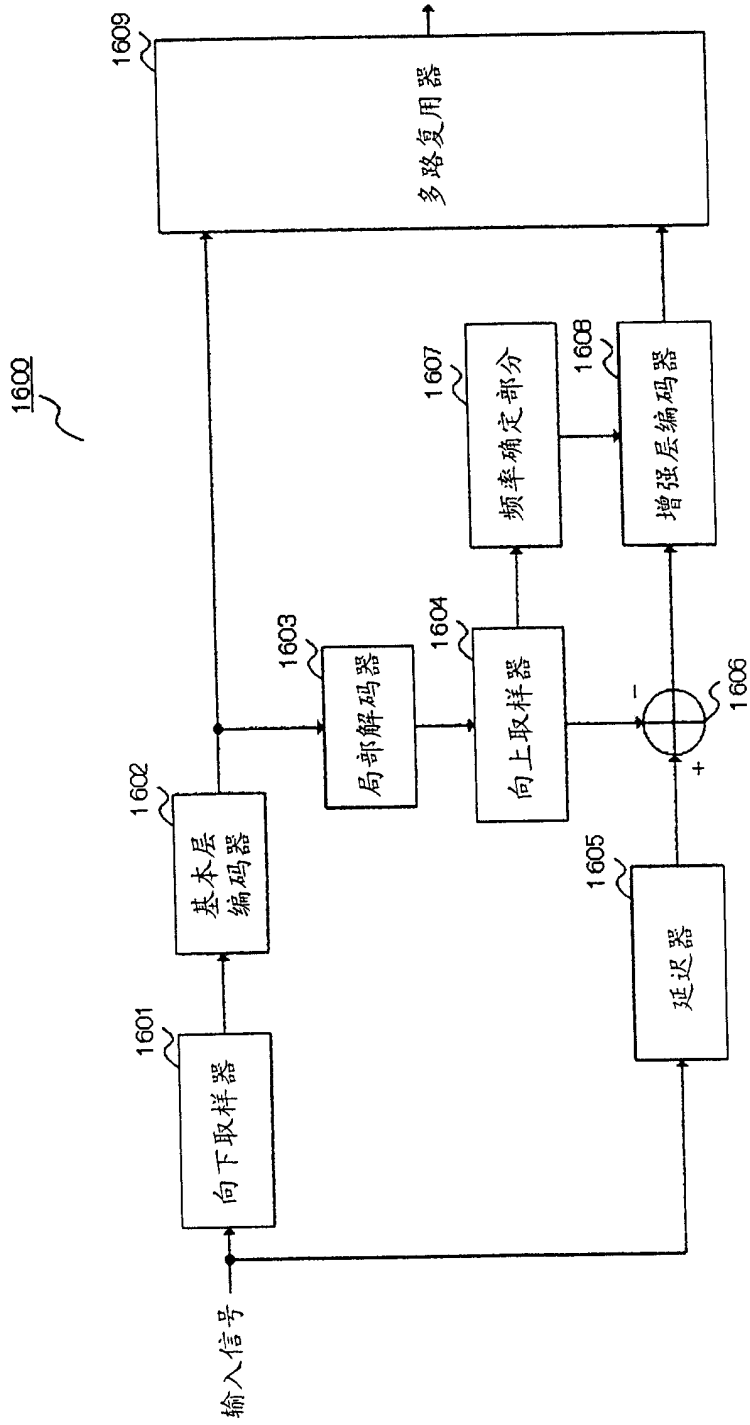


图 16

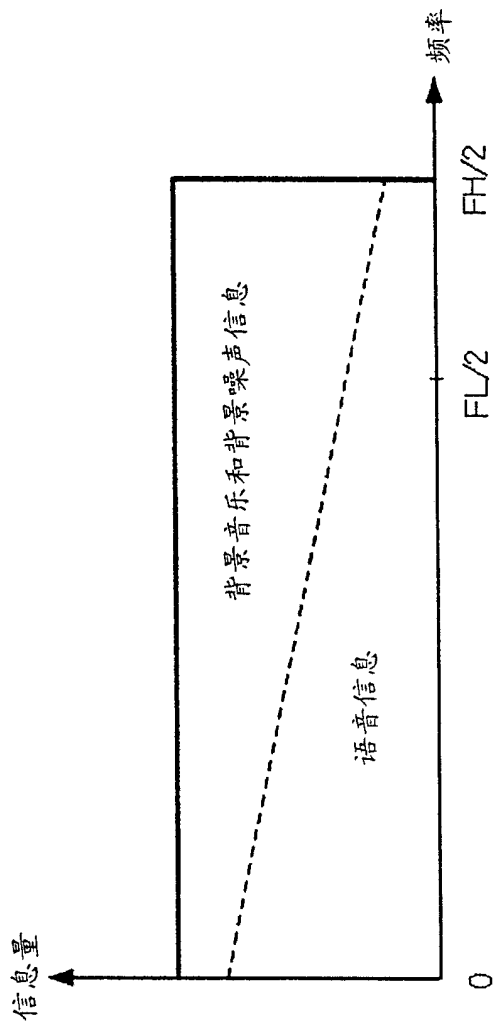


图 17

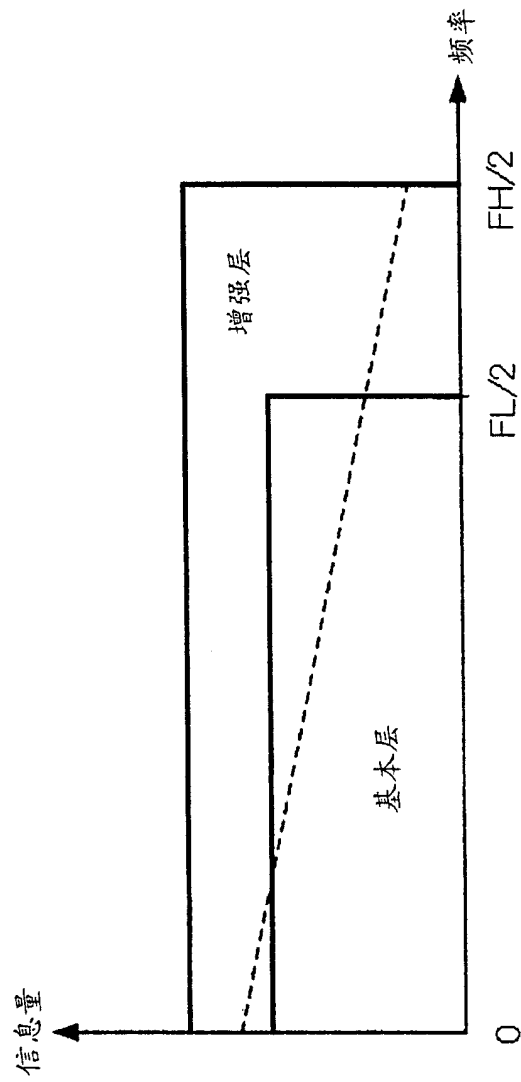


图 18

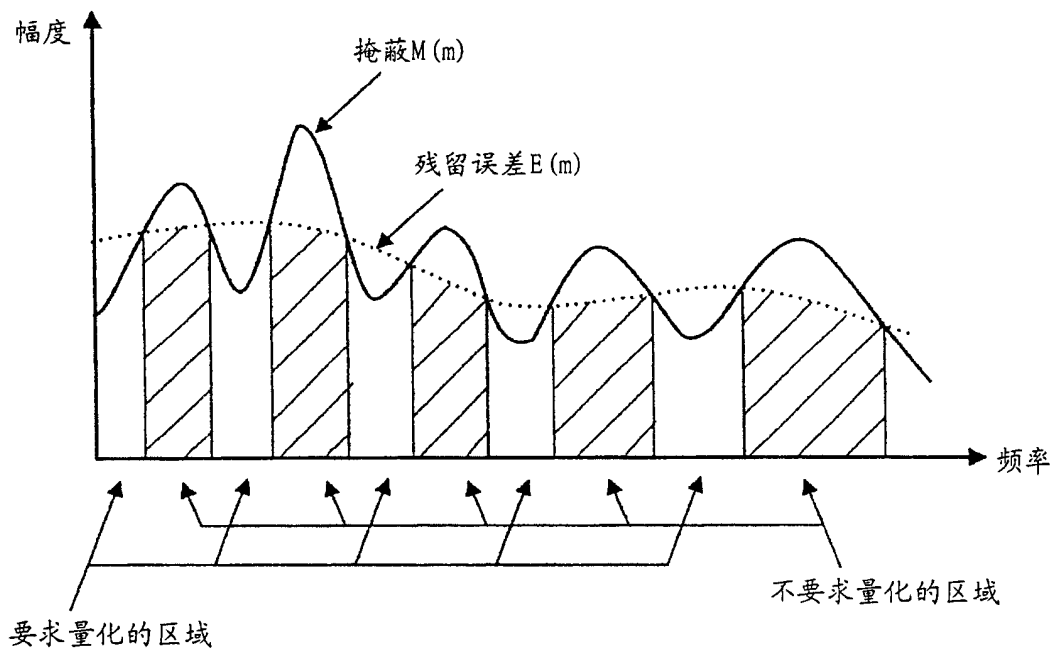


图 19

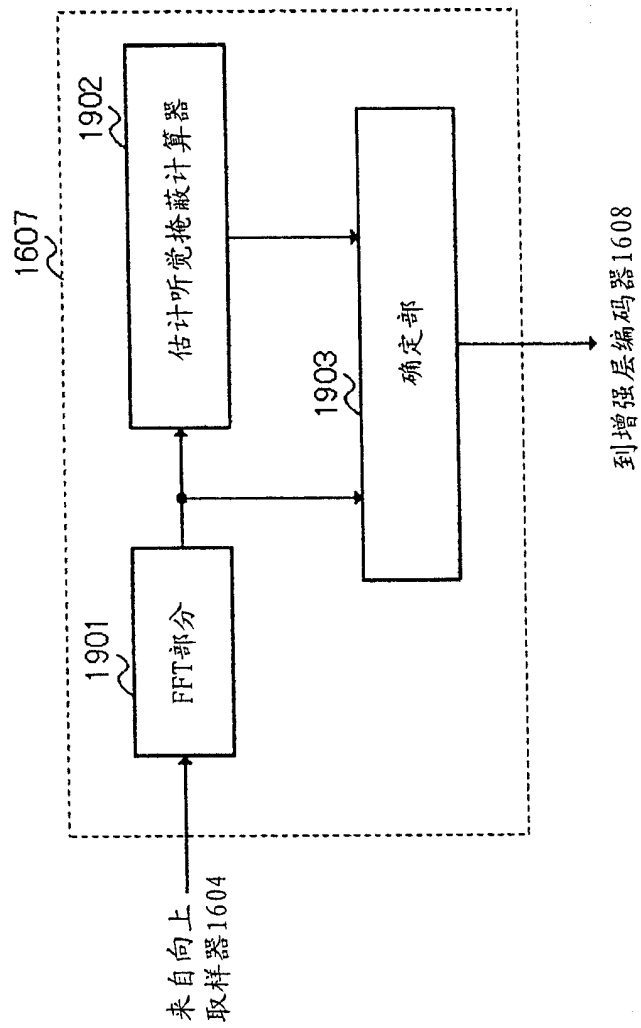


图 20

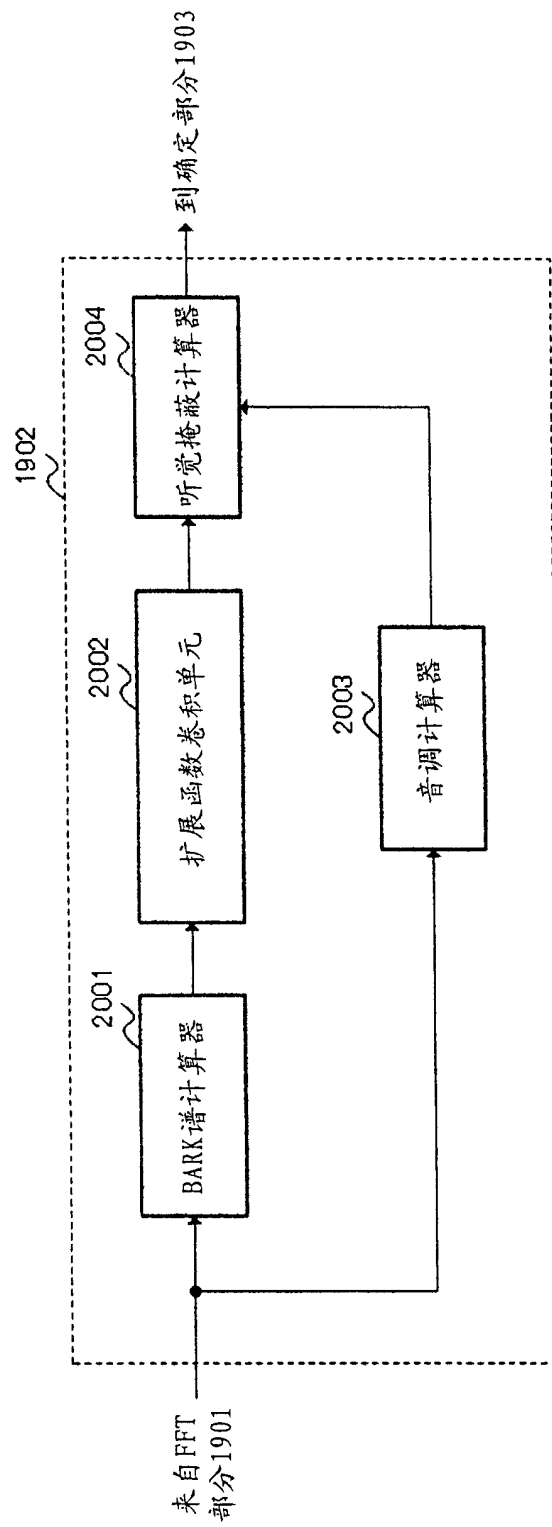


图 21

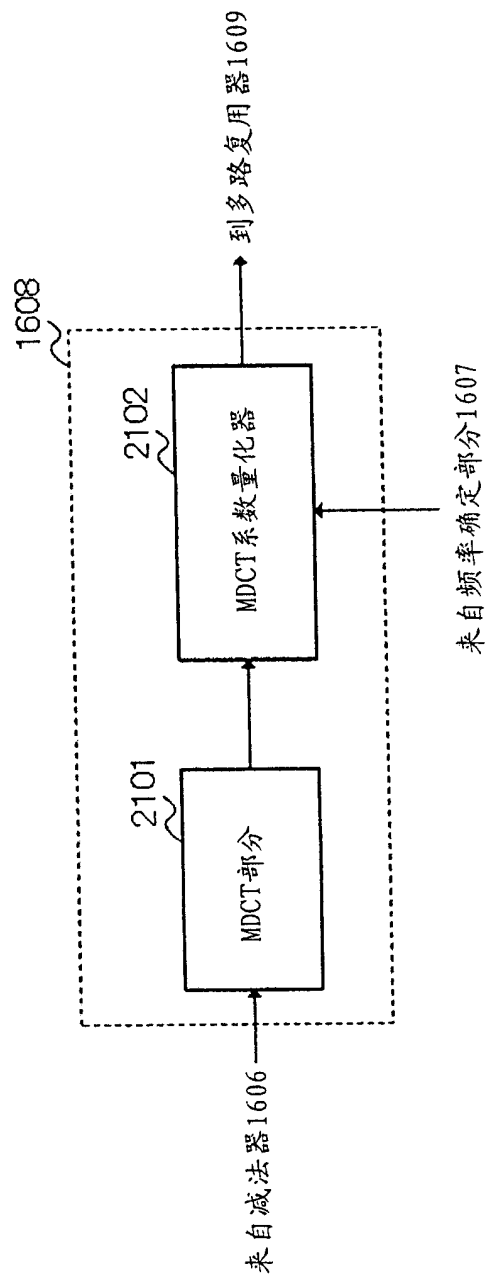


图 22

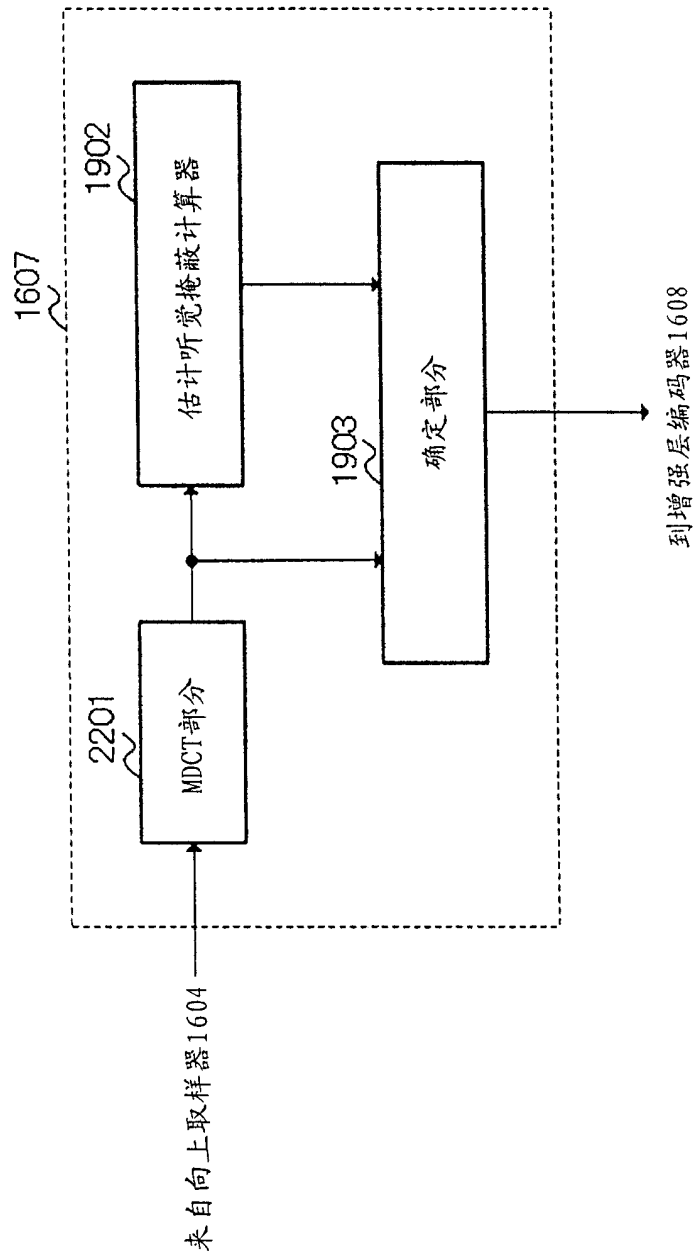


图 23

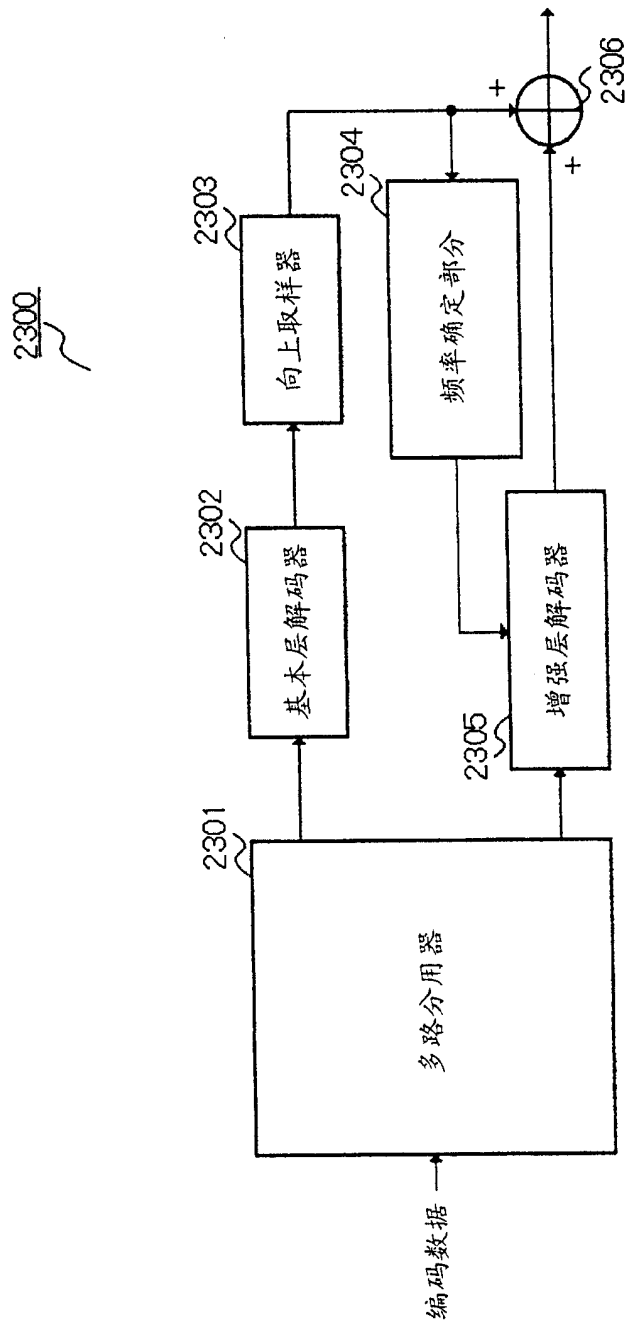


图 24

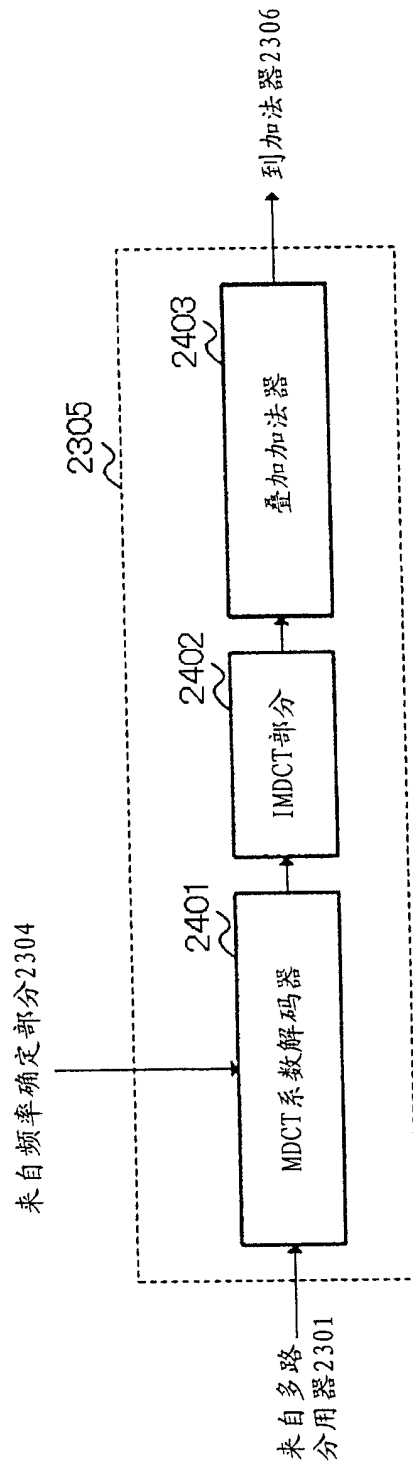
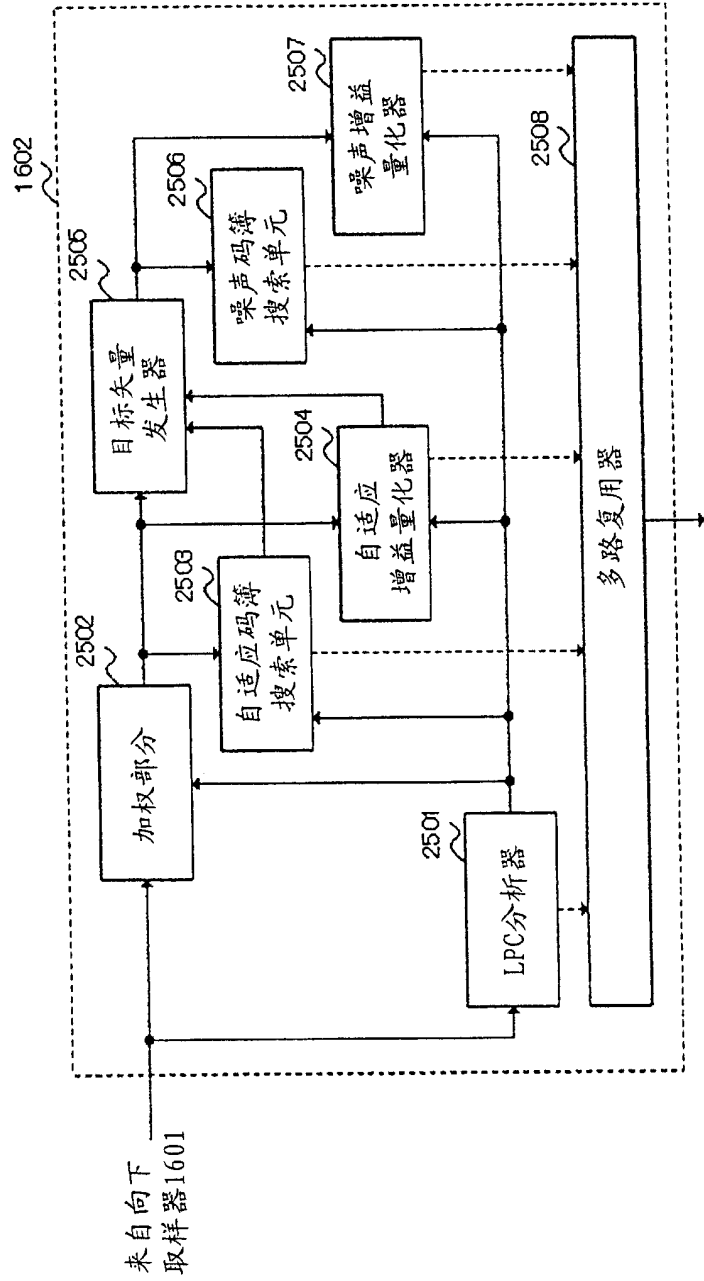


图 25



到局部解码器1603和多路复用器1609

图 26

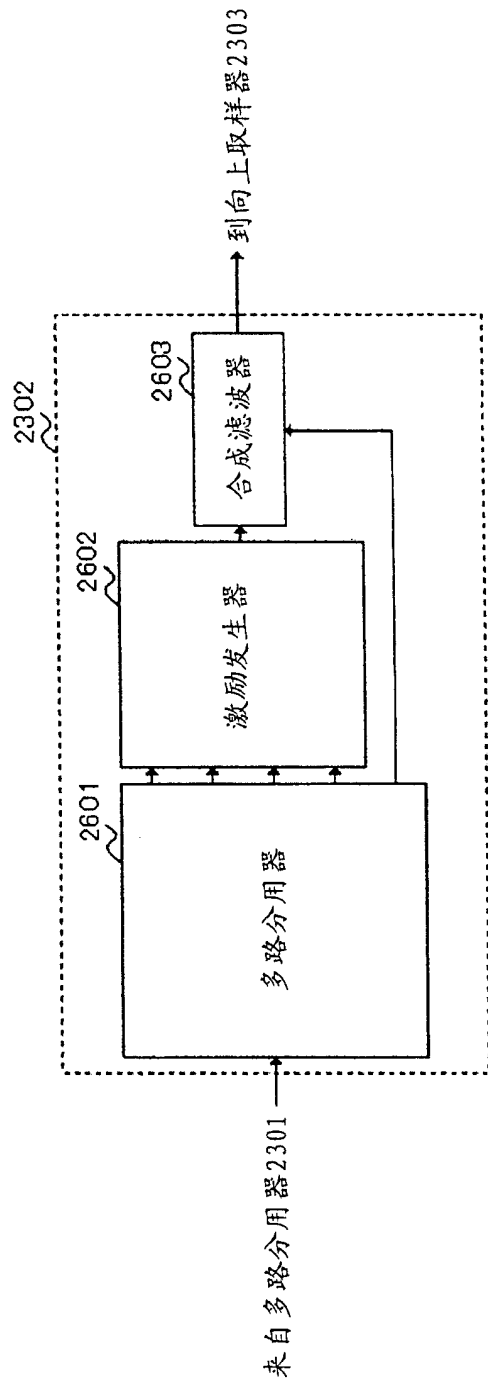


图 27

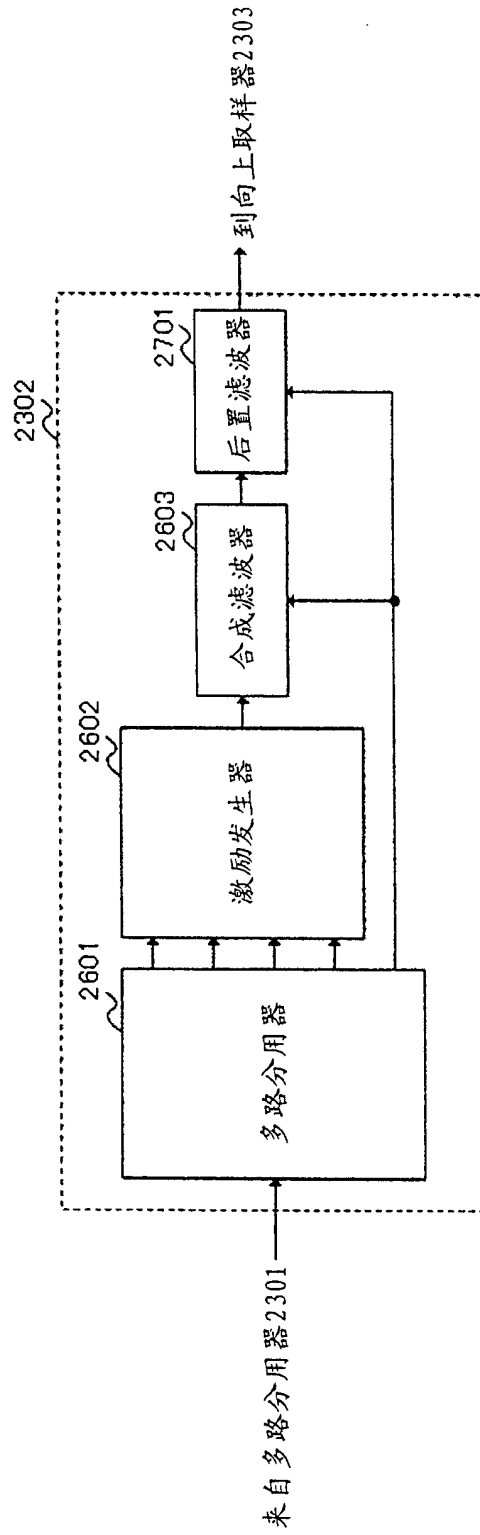


图 28

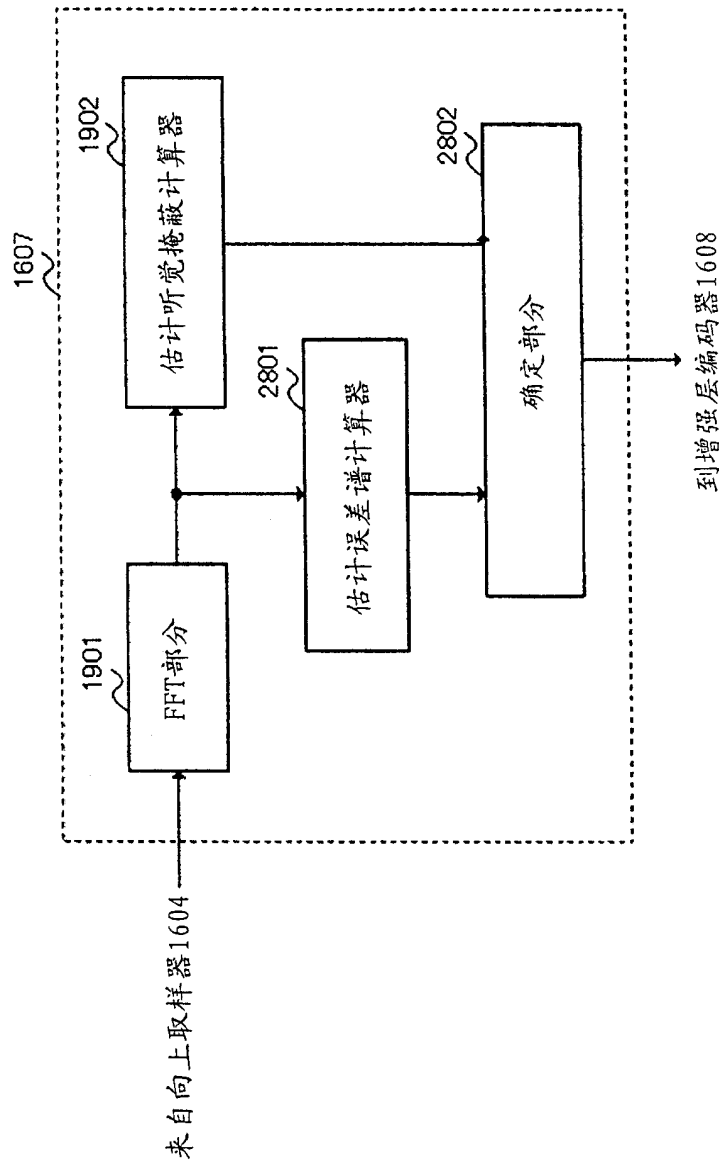


图 29

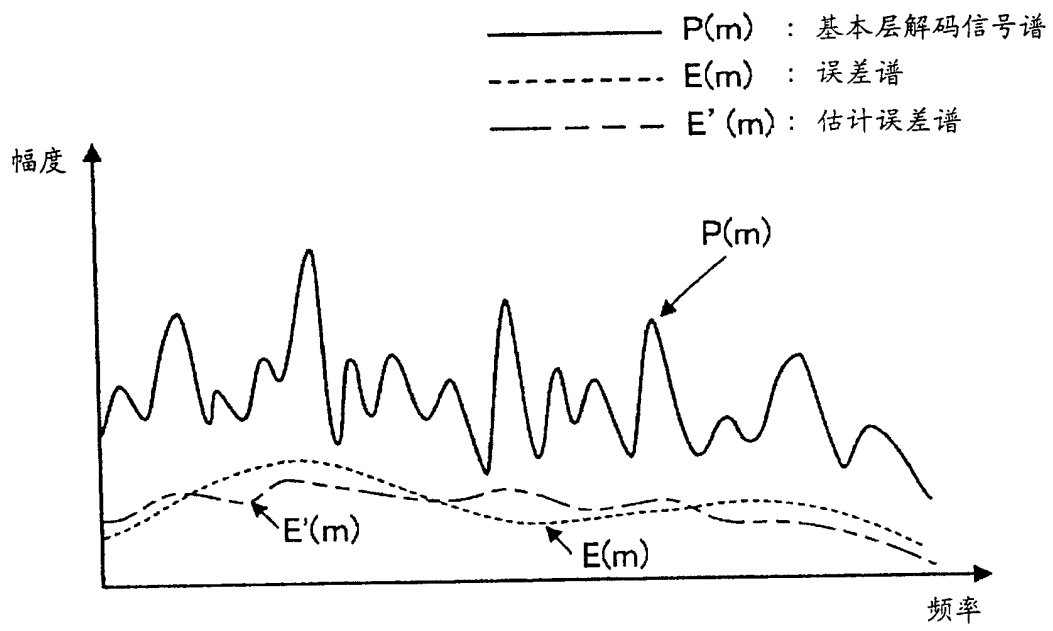


图 30

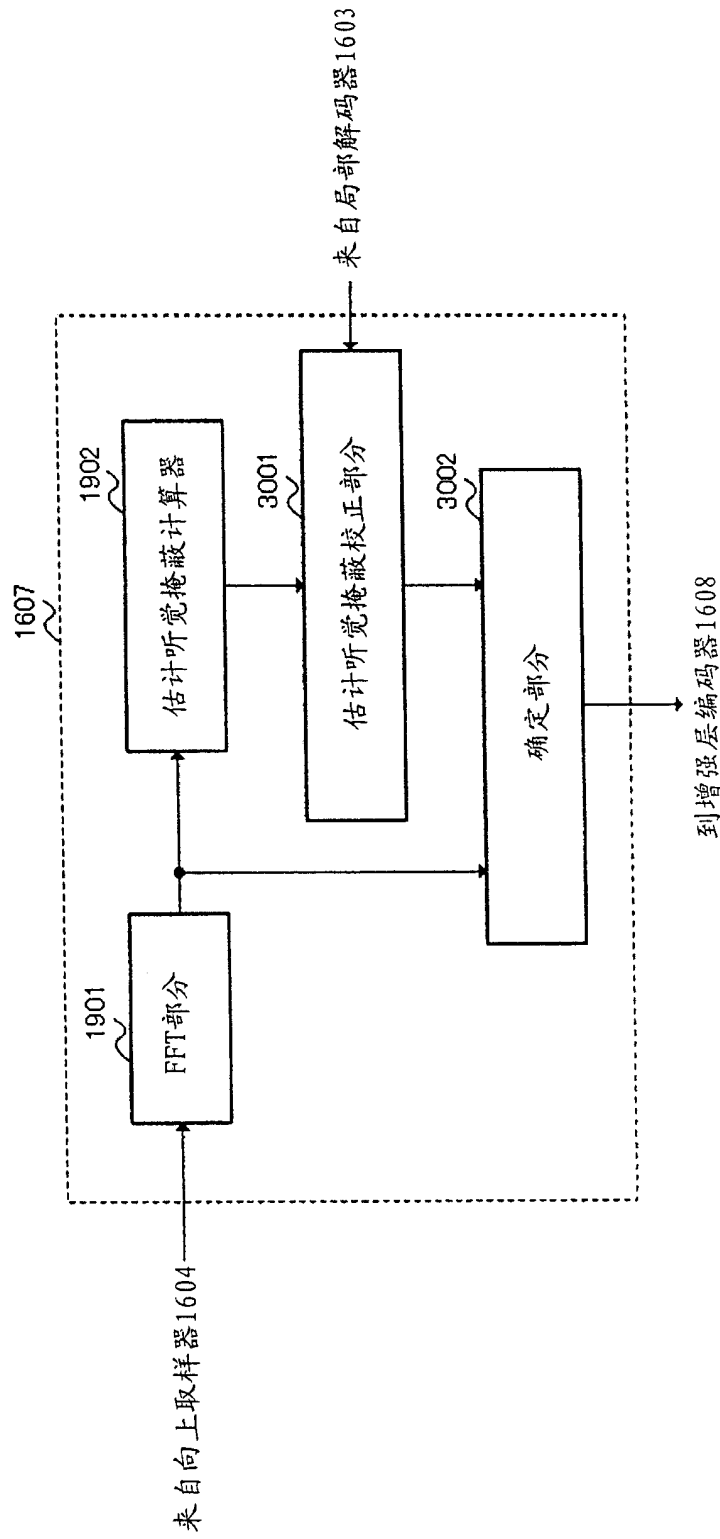


图 31

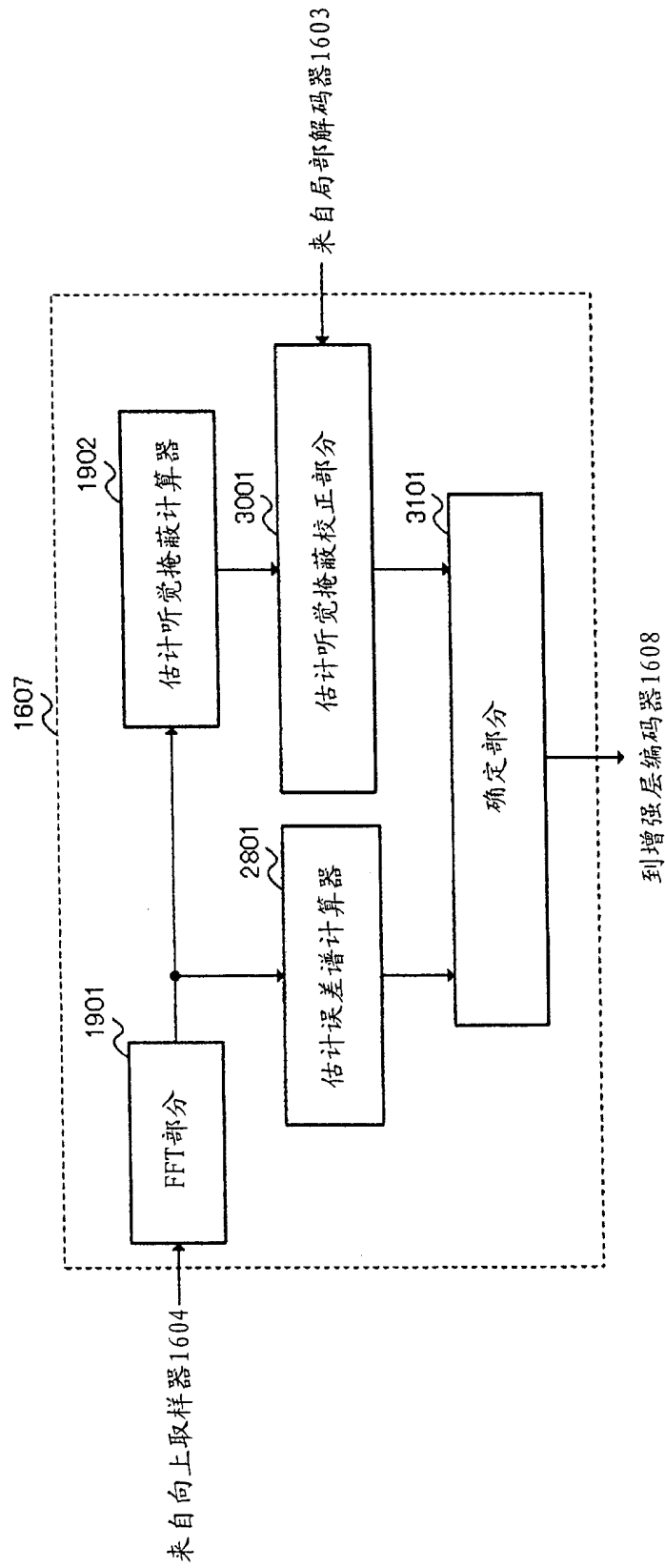


图 32

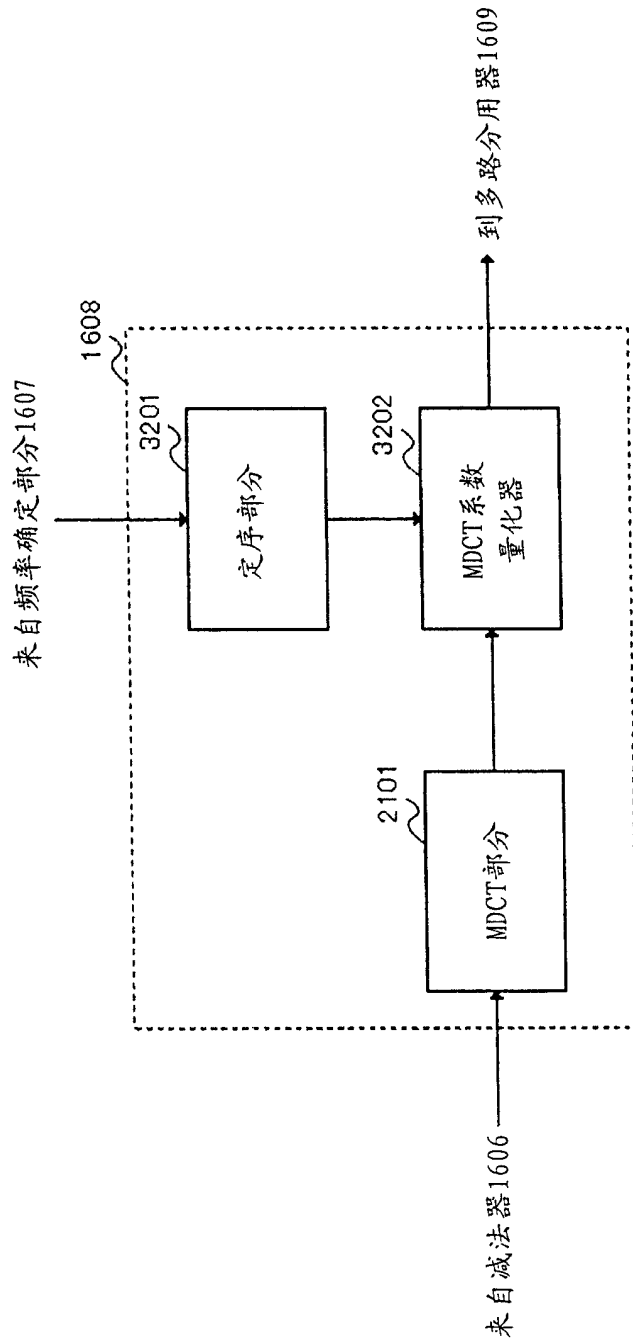


图 33

频率 (m)	估计失真值D (m)	次序
1	24.0	5
3	16.8	7
4	35.9	3
7	147.2	1
8	135.8	2
9	26.9	4
11	23.0	6
12	12.9	8

图 34

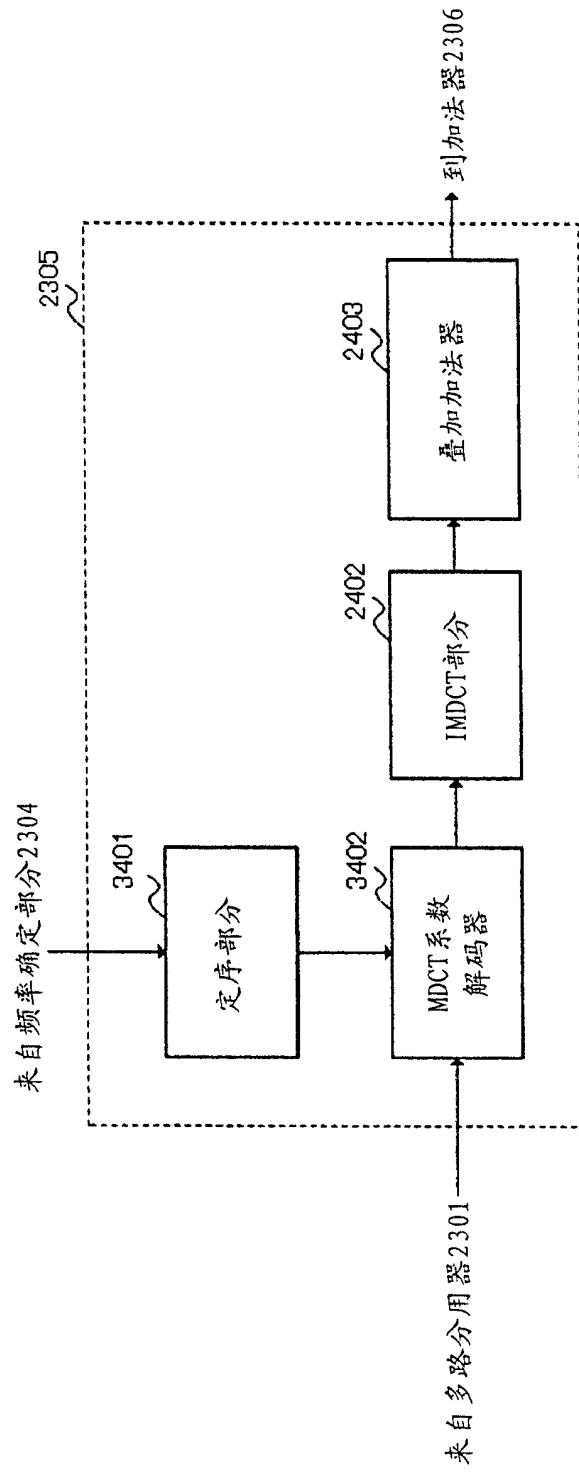


图 35

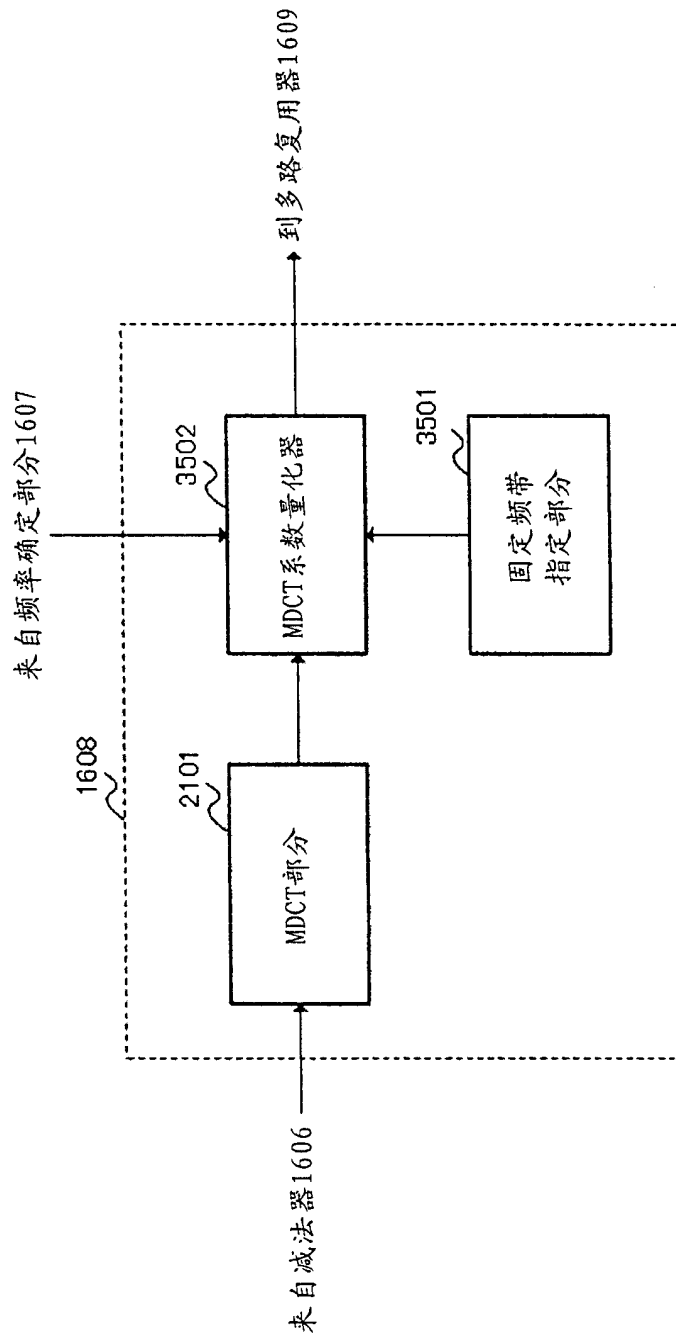


图 36

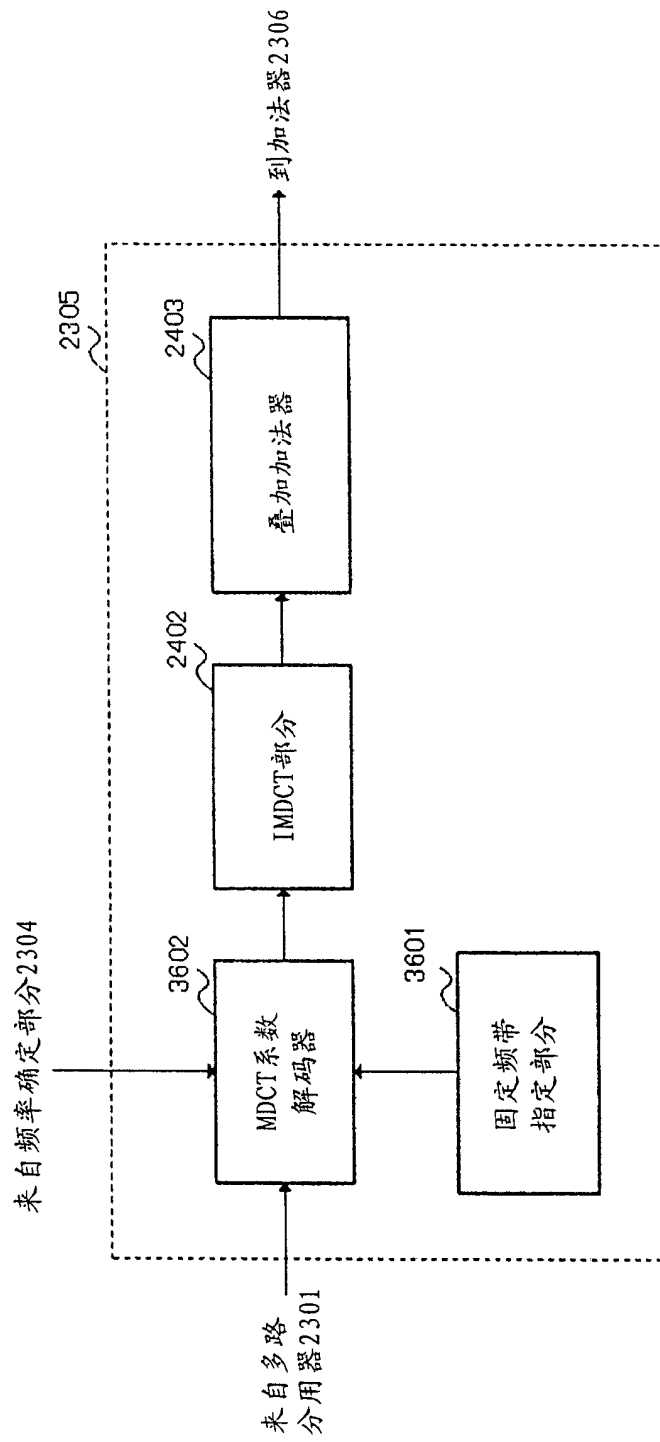


图 37

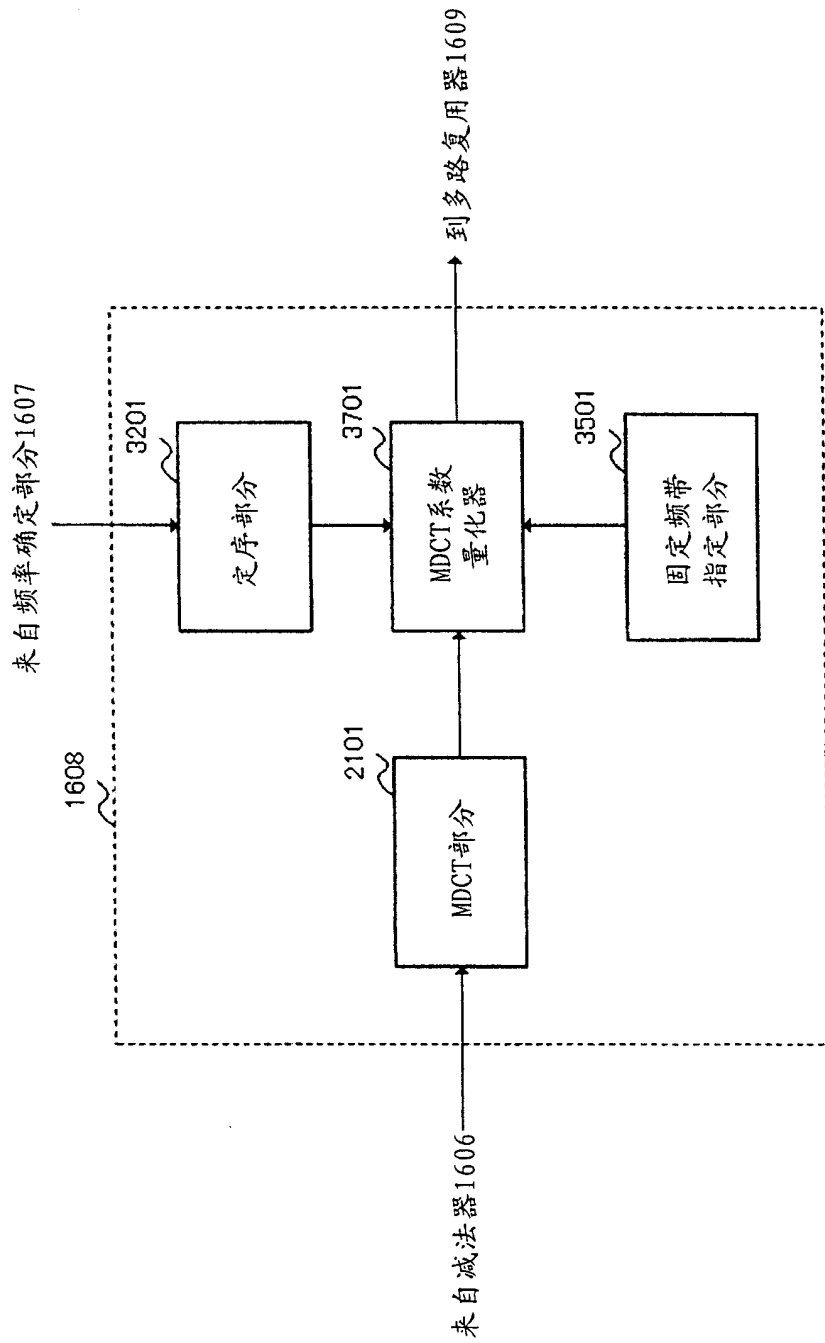


图 38

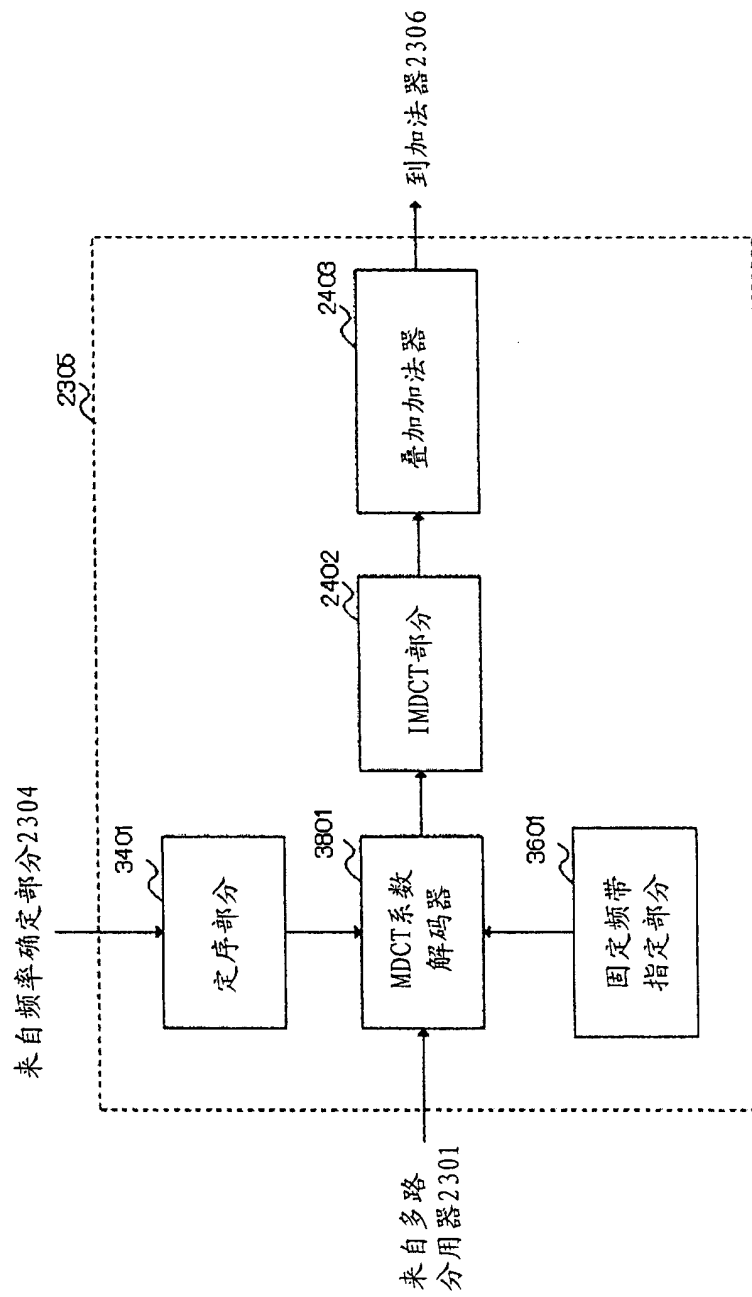


图 39

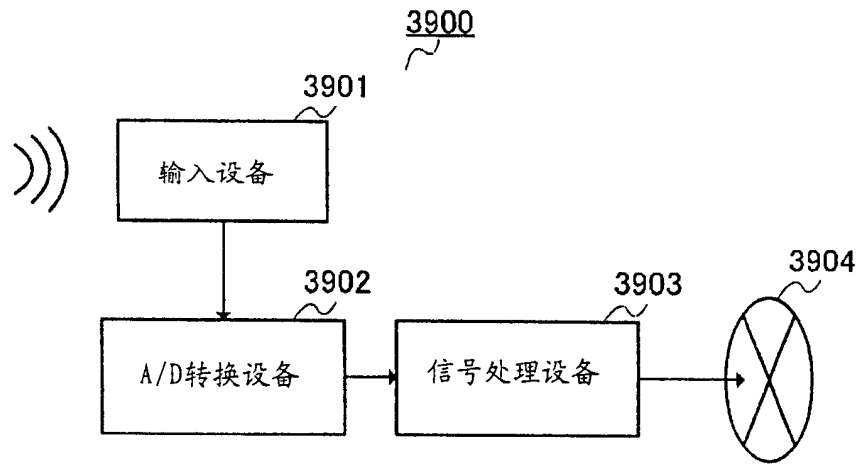


图 40

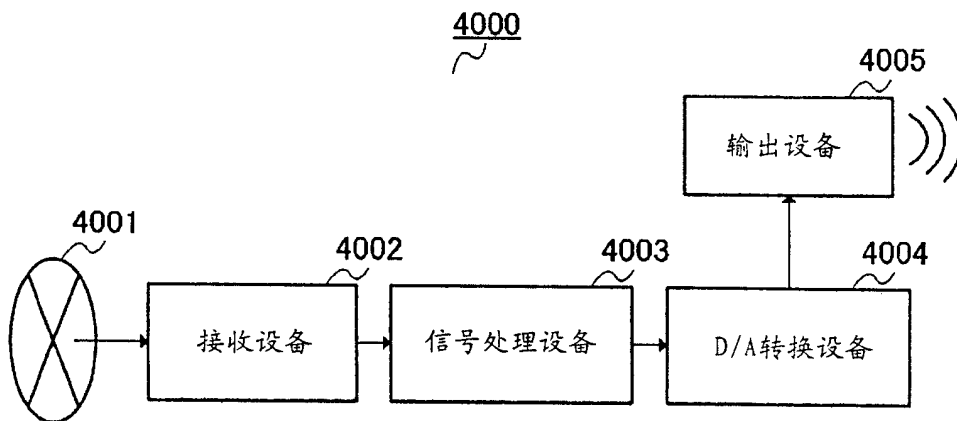


图 41

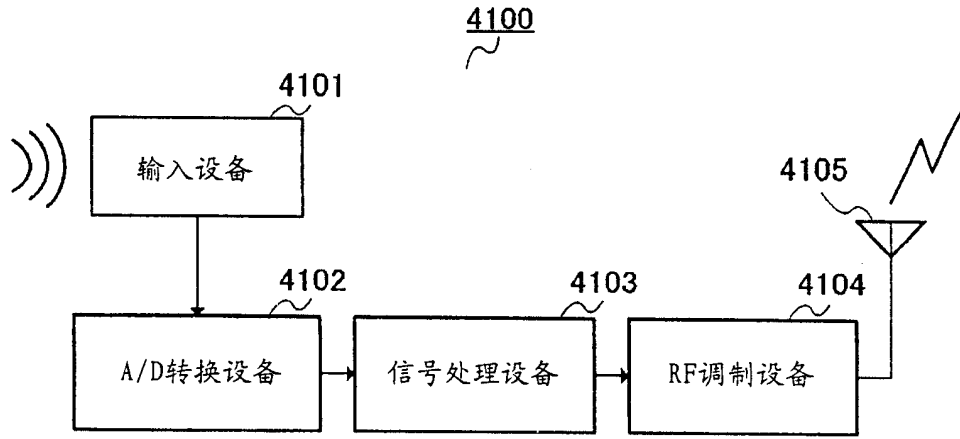


图 42

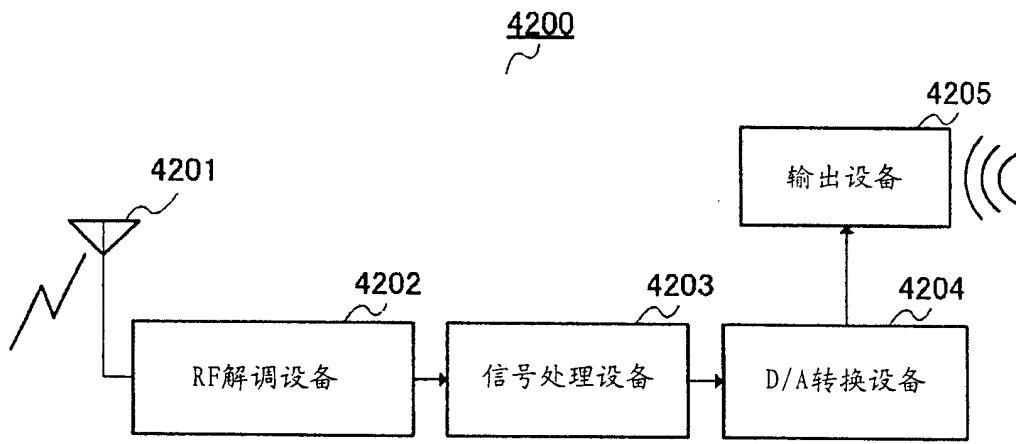


图 43