



(12) 发明专利

(10) 授权公告号 CN 102099855 B

(45) 授权公告日 2012. 09. 26

(21) 申请号 200980128382. 3

(51) Int. Cl.

(22) 申请日 2009. 08. 07

G10L 19/02(2006. 01)

G10L 11/00(2006. 01)

(30) 优先权数据

205645/08 2008. 08. 08 JP

096222/09 2009. 04. 10 JP

(56) 对比文件

US 7277550 B1, 2007. 10. 02, 全文.

CN 1476180 A, 2004. 02. 18, 全文.

CN 1684143 A, 2005. 10. 19, 全文.

JP 特开 2006-11456 A, 2006. 01. 12, 全文.

JP 特开 2002-244695 A, 2002. 08. 30, 全文.

(85) PCT申请进入国家阶段日

2011. 01. 20

(86) PCT申请的申请数据

PCT/JP2009/003799 2009. 08. 07

审查员 王馨宁

(87) PCT申请的公布数据

W02010/016271 JA 2010. 02. 11

(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 山梨智史 押切正浩 森井利幸

江原宏幸

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 邸万奎

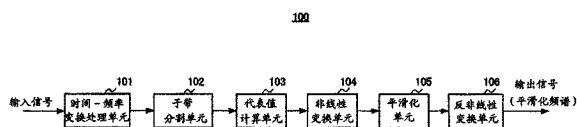
权利要求书 3 页 说明书 16 页 附图 10 页

(54) 发明名称

频谱平滑化装置、编码装置、解码装置、通信终端装置、基站装置以及频谱平滑化方法

(57) 摘要

公开了在对从语音信号算出的频谱进行了非线性变换之后,进行平滑的结构中,维持良好的语音质量,并且大幅度削减处理运算量的频谱平滑化装置。在该频谱平滑化装置中,子带分割单元(102)将输入频谱分割为多个子带,代表值计算单元(103)对每个子带使用算术平均以及几何平均计算代表值,非线性变换单元(104)对于各个代表值,进行其值越大越增强特性的非线性变换,平滑化单元(105)将该每个子带的非线性变换后的代表值在频域上进行平滑。



1. 频谱平滑化装置,包括:

时间-频率变换单元,将输入语音信号进行时间-频率变换而生成频率分量;

子带分割单元,将所述频率分量分割为多个子带;

代表值计算单元,对于所述分割出的各个子带,使用算术平均的计算及利用了该计算结果的乘法运算计算子带的代表值;

非线性变换单元,对每个所述子带的代表值进行非线性变换;以及

平滑化单元,将进行了所述非线性变换的代表值在频域上进行平滑,

所述代表值计算单元将各个子带进一步分割为多个子群,对所述多个子群的每个子群计算算术平均值,计算将所述每个子群的算术平均值相乘所得的值作为每个所述子带的代表值。

2. 频谱平滑化装置,包括:

时间-频率变换单元,将输入语音信号进行时间-频率变换而生成频率分量;

子带分割单元,将所述频率分量分割为多个子带;

代表值计算单元,对于所述分割出的各个子带,使用算术平均的计算及利用了该计算结果的乘法运算计算子带的代表值;

非线性变换单元,对每个所述子带的代表值进行非线性变换;以及

平滑化单元,将进行了所述非线性变换的代表值在频域上进行平滑,

所述代表值计算单元通过将各个子带进一步分割为多个子群,对所述多个子群的每个子群计算算术平均值,使用进行了所述乘法运算所得的结果,计算几何平均值,从而计算每个所述子带的代表值,所述乘法运算利用了所述每个子群的算术平均值。

3. 如权利要求1或2所述的频谱平滑化装置,还包括:

反非线性变换单元,将平滑后的代表值进行与所述非线性变换相反特性的反非线性变换。

4. 如权利要求1或2所述的频谱平滑化装置,

所述非线性变换单元对于所述各个代表值,进行其值越大越增强特性的非线性变换。

5. 如权利要求1或2所述的频谱平滑化装置,

所述非线性变换单元进行对数变换作为所述非线性变换。

6. 如权利要求1或2所述的频谱平滑化装置,

所述非线性变换单元通过对每个所述子带的代表值进行所述非线性变换,计算每个子带的中间值,并且计算对于每个所述子带的中间值乘以各个子带内的子群数的倒数所得的值作为所述非线性变换后的代表值。

7. 编码装置,其进行频带扩展编码,该编码装置包括:

第1编码单元,对输入语音信号的规定频率以下的低频部分进行编码而生成第1编码信息;

解码单元,对所述第1编码信息进行解码而生成解码信号;以及

第2编码单元,通过将所述输入语音信号的比所述规定频率高的高频部分分割为多个子带,从所述输入语音信号或所述解码信号分别估计所述多个子带,从而生成第2编码信息,

所述第2编码单元

具备输入所述解码信号并进行平滑的权利要求 1 至权利要求 6 中的任一项权利要求所述的频谱平滑化装置，

从所述输入语音信号或平滑后的所述解码信号分别估计所述多个子带。

8. 解码装置，其进行频带扩展解码，该解码装置包括：

接收单元，接收在编码装置中生成的第 1 编码信息和在编码装置中生成的第 2 编码信息，所述第 1 编码信息为对编码侧输入语音信号的规定频率以下的低频部分进行编码所得的编码信息，所述第 2 编码信息为通过将所述编码侧输入语音信号的比所述规定频率高的高频部分分割为多个子带，从所述编码侧输入语音信号或对所述第 1 编码信息进行解码所得的第 1 解码信号，分别估计所述多个子带，从而生成的编码信息；

第 1 解码单元，对所述第 1 编码信息进行解码而生成第 2 解码信号；以及

第 2 解码单元，通过使用所述第 2 编码信息，从所述第 2 解码信号估计所述编码侧输入语音信号的高频部分，从而生成第 3 解码信号，

所述第 2 解码单元

具备输入所述第 2 解码信号并进行平滑的权利要求 1 至权利要求 6 中的任一项权利要求所述的频谱平滑化装置，

从平滑后的所述第 2 解码信号估计所述编码侧输入语音信号的高频部分。

9. 通信终端装置，包括权利要求 1 至权利要求 6 中的任一项权利要求所述的频谱平滑化装置。

10. 基站装置，包括权利要求 1 至权利要求 6 中的任一项权利要求所述的频谱平滑化装置。

11. 频谱平滑化方法，包括：

时间 - 频率变换步骤，将输入语音信号进行时间 - 频率变换而生成频率分量；

子带分割步骤，将所述频率分量分割为多个子带；

代表值计算步骤，对于所述分割出的各个子带，使用算术平均的计算及利用了该计算结果的乘法运算计算子带的代表值；

非线性变换步骤，对于每个所述子带的代表值进行非线性变换；以及

平滑化步骤，将进行了所述非线性变换的代表值在频域上进行平滑，

在所述代表值计算步骤中，将各个子带进一步分割为多个子群，对所述多个子群的每个子群计算算术平均值，计算将所述每个子群的算术平均值相乘所得的值作为每个所述子带的代表值。

12. 频谱平滑化方法，包括：

时间 - 频率变换步骤，将输入语音信号进行时间 - 频率变换而生成频率分量；

子带分割步骤，将所述频率分量分割为多个子带；

代表值计算步骤，对于所述分割出的各个子带，使用算术平均的计算及利用了该计算结果的乘法运算计算子带的代表值；

非线性变换步骤，对于每个所述子带的代表值进行非线性变换；以及

平滑化步骤，将进行了所述非线性变换的代表值在频域上进行平滑，

在所述代表值计算步骤中，通过将各个子带进一步分割为多个子群，对所述多个子群的每个子群计算算术平均值，使用进行了所述乘法运算所得的结果，计算几何平均值，从而

计算每个所述子带的代表值,所述乘法运算利用了所述每个子群的算术平均值。

频谱平滑化装置、编码装置、解码装置、通信终端装置、基站装置以及频谱平滑化方法

技术领域

[0001] 本发明涉及将语音信号的频谱平滑化的频谱平滑化装置、编码装置、解码装置、通信终端装置、基站装置以及频谱平滑化方法。

背景技术

[0002] 通过以因特网通信为代表的分组通信系统、或移动通信系统等传输语音 / 音频信号时,为了提高语音 / 音频信号的传输效率,常使用压缩 / 编码技术。另外,近年来,仅以低比特率对语音 / 音频信号进行编码,另一方面,对更高画质的语音 / 音频信号进行编码的技术的需求日益高涨。

[0003] 对于这样的需求,开发了用于下述目的的各种技术,即,对语音信号进行正交变换(时间-频率变换),计算语音信号的频率分量(频谱),对计算出的频谱进行线性变换及非线性变换等处理而提高解码信号的质量(例如,参照专利文献1)。在专利文献1公开的方法中,首先从一定时间长度的语音信号分析该语音信号所包含的频谱,并且对于分析得出的频谱,进行其频谱强度的值越大越增强非线性变换处理。接着,对于进行了非线性变换处理的频谱,在频域上进行线性的平滑化处理。其后,进行用于消除非线性变换特性的反非线性变换处理,还进行用于消除平滑化特性的反平滑化处理,由此抑制语音信号所包含的全频带的噪音分量。这样,在专利文献1公开的方法中,通过在对于从语音信号获得的频谱的全部样本进行了非线性变换处理之后,进行频谱的平滑化,由此获得良好音质的语音信号。另外,专利文献1中列举了乘幂、对数变换等变换方法作为非线性处理的例子。

[0004] 现有技术文献

[0005] 专利文献

[0006] 专利文献1:特开2002-244695号公报

[0007] 专利文献2:国际公开第2007/037361号小册子

[0008] 非专利文献

[0009] 非专利文献1:Yuichiro TAKAMIZAWA, Toshiyuki NOMURA and Masao IKEKAWA, "High-Quality and Processor-Efficient Implementation of and MPEG-2 AAC Encoder", IEICE TRANS. INF. & SYST., VOL. E86-D, No. 3 MARCH 2003

发明内容

[0010] 发明需要解决的问题

[0011] 但是,专利文献1公开的方法中存在下述问题,即,对从语音信号获得的频谱的全部样本进行非线性变换处理,所以处理运算量庞大。为了削减处理运算量,单纯地从频谱的样本中提取一部分的样本,并仅对于提取出的样本进行非线性变换处理,即使在非线性变换后进行频谱的平滑化,也不一定能够获得足够好的语音质量。

[0012] 本发明的目的在于,提供在对于从语音信号计算出的频谱进行了非线性变换之

后,进行平滑化的结构中,能够维持良好的语音质量,而且大幅度削减处理运算量的频谱平滑化装置、编码装置、解码装置、通信终端装置、基站装置以及频谱平滑化方法。

[0013] 解决问题的方案

[0014] 本发明的频谱平滑化装置采用的结构包括:时间-频率变换单元,将输入语音信号进行时间-频率变换而生成频率分量;子带分割单元,将所述频率分量分割为多个子带;代表值计算单元,对于所述分割出的各个子带,使用算术平均的计算及利用了该计算结果的乘法运算计算子带的代表值;非线性变换单元,对每个所述子带的代表值进行非线性变换;以及平滑化单元,将进行了所述非线性变换的代表值在频域上进行平滑,所述代表值计算单元将各个子带进一步分割为多个子群,对所述多个子群的每个子群计算算术平均值,计算将所述每个子群的算术平均值相乘所得的值作为每个所述子带的代表值。

[0015] 本发明的频谱平滑化装置包括:时间-频率变换单元,将输入语音信号进行时间-频率变换而生成频率分量;子带分割单元,将所述频率分量分割为多个子带;代表值计算单元,对于所述分割出的各个子带,使用算术平均的计算及利用了该计算结果的乘法运算计算子带的代表值;非线性变换单元,对每个所述子带的代表值进行非线性变换;以及平滑化单元,将进行了所述非线性变换的代表值在频域上进行平滑,所述代表值计算单元通过将各个子带进一步分割为多个子群,对所述多个子群的每个子群计算算术平均值,使用进行了所述乘法运算所得的结果,计算几何平均值,从而计算每个所述子带的代表值,所述乘法运算利用了所述每个子群的算术平均值。

[0016] 本发明的频谱平滑化方法包括:时间-频率变换步骤,将输入语音信号进行时间-频率变换而生成频率分量;子带分割步骤,将所述频率分量分割为多个子带;代表值计算步骤,对于所述分割出的各个子带,使用算术平均的计算及利用了该计算结果的乘法运算计算子带的代表值;非线性变换步骤,对于每个所述子带的代表值进行非线性变换;以及平滑化步骤,将进行了所述非线性变换的代表值在频域上进行平滑,在所述代表值计算步骤中,将各个子带进一步分割为多个子群,对所述多个子群的每个子群计算算术平均值,计算将所述每个子群的算术平均值相乘所得的值作为每个所述子带的代表值。

[0017] 本发明的频谱平滑化方法,包括:时间-频率变换步骤,将输入语音信号进行时间-频率变换而生成频率分量;子带分割步骤,将所述频率分量分割为多个子带;代表值计算步骤,对于所述分割出的各个子带,使用算术平均的计算及利用了该计算结果的乘法运算计算子带的代表值;非线性变换步骤,对于每个所述子带的代表值进行非线性变换;以及平滑化步骤,将进行了所述非线性变换的代表值在频域上进行平滑,在所述代表值计算步骤中,通过将各个子带进一步分割为多个子群,对所述多个子群的每个子群计算算术平均值,使用进行了所述乘法运算所得的结果,计算几何平均值,从而计算每个所述子带的代表值,所述乘法运算利用了所述每个子群的算术平均值。

[0018] 发明的效果

[0019] 根据本发明,能够维持良好的语音质量,而且大幅度削减处理运算量。

附图说明

[0020] 图 1A ~ 图 1D 是表示本发明实施方式 1 的处理的概要的频谱概略图。

[0021] 图 2 是表示实施方式 1 的频谱平滑化装置的主要结构的方框图。

- [0022] 图 3 是表示实施方式 1 的代表值计算单元的主要结构的方框图。
- [0023] 图 4 是表示实施方式 1 的输入信号的子带及子群的结构概略图。
- [0024] 图 5 是表示本发明实施方式 2 的具有编码装置和解码装置的通信系统的结构的方框图。
- [0025] 图 6 是表示实施方式 2 的图 5 所示的编码装置的内部主要结构的方框图。
- [0026] 图 7 是表示实施方式 2 的图 6 所示的第 2 层编码单元的内部主要结构的方框图。
- [0027] 图 8 是表示实施方式 2 的图 7 所示的频谱平滑单元的主要结构的方框图。
- [0028] 图 9 是用于说明实施方式 2 的图 7 所示的滤波单元中的滤波处理的细节的图。
- [0029] 图 10 是表示在实施方式 2 的图 7 所示的搜索单元中对子带 SB_p 搜索最合适音调系数 T_p' 的处理的步骤的流程图。
- [0030] 图 11 是表示实施方式 2 的图 5 所示的解码装置的内部主要结构的方框图。
- [0031] 图 12 是表示实施方式 2 的图 11 所示的第 2 层解码单元的内部主要结构的方框图。
- [0032] 标号说明
- [0033] 100 频谱平滑化装置
- [0034] 101、315、334、357 时间 - 频率变换处理单元
- [0035] 102 子带分割单元
- [0036] 103 代表值计算单元
- [0037] 104 非线性变换单元
- [0038] 105 平滑化单元
- [0039] 106 反非线性变换单元
- [0040] 201 相加平均计算单元
- [0041] 202 相乘平均计算单元
- [0042] 301 编码装置
- [0043] 302 传输路径
- [0044] 303 解码装置
- [0045] 311 下采样处理单元
- [0046] 312 第 1 层编码单元
- [0047] 313、332 第一层解码单元
- [0048] 314、333 上采样处理单元
- [0049] 316 第二层编码单元
- [0050] 317 编码信息汇集单元
- [0051] 318 延迟单元
- [0052] 331 编码信息分离单元
- [0053] 335 第 2 层解码单元
- [0054] 351 分离单元
- [0055] 352、361 频谱平滑化单元
- [0056] 353、362 滤波器状态设定单元
- [0057] 354、363 滤波单元
- [0058] 355 增益解码单元

- [0059] 356 频谱调整单元
- [0060] 360 频带分割单元
- [0061] 364 搜索单元
- [0062] 365 音调系数设定单元
- [0063] 366 增益编码单元
- [0064] 367 复用单元

具体实施方式

[0065] 以下,参照附图详细说明本发明的实施方式。

[0066] (实施方式 1)

[0067] 首先使用图 1 说明本发明实施方式的频谱平滑化方法的概要。图 1 是用于说明本实施方式的频谱平滑化方法的概要的频谱图。

[0068] 图 1A 表示输入信号的频谱。在本实施方式中,首先将输入信号的频谱分割为多个子带。图 1B 表示已分割为多个子带的输入信号的频谱的情形。另外,图 1 的频谱图是用于说明本发明的概要的图,例如,本发明并不限定图中的子带数。

[0069] 接着,对各个子带计算代表值。具体而言,将子带内的样本进一步分割为多个子群。另外,对每个子群计算频谱的绝对值的算术平均(相加平均)。

[0070] 接着,对每个子带计算各个子群的算术平均值的几何平均(相乘平均)。另外,上述的几何平均值此时还不是正确的几何平均值,计算将各个子群的算术平均值单纯地相乘所得的值,在后述的非线性变换之后求得正确的几何平均值。上述处理是为了进一步削减运算量,当然也可以在此时求取正确的几何平均值。

[0071] 将上述的几何平均值设为各个子带的代表值。图 1C 中,与以虚线表示的输入信号的频谱重叠而表示各个子带的代表值。另外,为了使说明易于理解,图 1C 将正确的几何平均值表示为代表值,替代将各个子群的算术平均值单纯地相乘所得的值。

[0072] 接着,对于各个子带的代表值,在相对于输入信号的频谱,进行频谱强度的值越大越增强非线性变换(例如,对数变换)之后,在频域进行平滑化处理。此后,进行反非线性变换(例如,对数逆变换),对各个子带计算平滑化频谱。图 1D 中,与以虚线表示的输入信号的频谱重叠而表示各个子带的平滑化频谱。

[0073] 通过这样的处理,对数区域内的频谱的平滑化能够抑制语音质量的劣化,而且大幅度削减处理运算量。以下,说明获得该效果的本发明的实施方式的频谱平滑化装置的结构。

[0074] 本实施方式的频谱平滑化装置将输入频谱平滑化,并输出平滑化后的频谱(以下称为“平滑化频谱”)作为输出信号。更具体而言,频谱平滑化装置将输入信号每 N 样本(N 为自然数)为单位进行划分,将 N 样本作为 1 帧而对每帧进行平滑化处理。这里,将作为平滑化处理的对象的输入信号表示为 $x_n(n = 0, \dots, N-1)$ 。 x_n 表示每 N 样本为单位划分的输入信号中第 n+1 样本。

[0075] 图 2 表示本实施方式的频谱平滑化装置 100 的主要结构。

[0076] 图 2 所示的频谱平滑化装置 100 主要包括:时间-频率变换处理单元 101、子带分割单元 102、代表值计算单元 103、非线性变换单元 104、平滑化单元 105 以及反非线性变换

单元 106。

[0077] 时间-频率变换处理单元 101 对输入信号 x_n 进行快速傅立叶变换 (FFT:Fast Fourier Transform), 计算频率分量的频谱 $S1(k)$ (以下称为输入频谱)。

[0078] 另外, 时间-频率变换处理单元 101 将输入频谱 $S1(k)$ 输出到子带分割单元 102。

[0079] 子带分割单元 102 将从时间-频率变换处理单元 101 输入的输入频谱 $S1(k)$ 分割为 P (P 为 2 以上的整数) 个子带。以下, 以下述情况为例进行说明, 即, 子带分割单元 102 将输入频谱 $S1(k)$ 分割, 使各个子带的样本数相等。另外, 各个子带的样本数也可以是在各个子带不同。子带分割单元 102 将已分割为子带的频谱 (以下也称为“子带频谱”) 输出到代表值计算单元 103。

[0080] 代表值计算单元 103 对于从子带分割单元 102 输入的、已分割为子带的输入频谱的各个子带, 计算代表值, 将并计算出每个子带的代表值输出到非线性变换单元 104。在后面叙述代表值计算单元 103 的详细处理。

[0081] 图 3 表示代表值计算单元 103 的内部结构。图 3 表示的代表值计算单元 103 包括: 相加平均计算单元 201 以及相乘平均计算单元 202。

[0082] 首先, 从子带分割单元 102, 子带频谱被输入到相加平均计算单元 201。

[0083] 相加平均计算单元 201 将输入的子带频谱的各个子带进一步分割为 Q 个 (Q 为 2 以上的整数) 的子群 (第 0 子群~第 $Q-1$ 子群)。另外, 以下, 以 Q 个的各个子群分别由 R 个 (R 为 2 以上的整数) 的样本构成情况为例进行说明。另外, 这里说明 Q 个的各个子群都由 R 个样本构成的情况, 但各个子群内的样本当然也可以为不同的数。

[0084] 图 4 表示子带和子群的结构例。图 4 表示作为一例, 构成 1 子带的样本数为 8, 构成子带的子群数 Q 为 2, 子群内的样本数 R 为 4 的情况。

[0085] 接着, 相加平均计算单元 201 对 Q 个子群的各个子群, 使用式 (1), 计算各个子群所包含的频谱 (FFT 系数) 的绝对值的算术平均 (相加平均)。

$$[0086] \quad AVE1_q = \frac{1}{R} \sum_{i=0}^{R-1} |S1_{BS_q+i}| \quad (q=0, \dots, Q-1) \quad \dots (1)$$

[0087] 另外, 在式 (1) 中, $AVE1_q$ 是第 q 子群所包含的频谱 (FFT 系数) 的绝对值的算术平均 (相加平均), BS_q 表示第 q 子群的开头样本的索引。

[0088] 接着, 相加平均计算单元 201 将计算出的每个子带的算术平均 (相加平均) 值频谱 $AVE1_q$ ($q=0 \sim Q-1$) (子带算术平均值频谱) 输出到相乘平均计算单元 202。

[0089] 相乘平均计算单元 202 将从相加平均计算单元 201 输入的每个子带的算术平均值 (相加平均) 频谱 $AVE1_q$ ($q=0 \sim Q-1$), 如式 (2) 所示那样全部相乘, 对每个子带计算代表值频谱 (子带代表值频谱) $AVE2_p$ ($p=0 \sim P-1$)。

$$[0090] \quad AVE2_p = \prod_{i=0}^{Q-1} AVE1_i \quad (p=0, \dots, P-1) \quad \dots (2)$$

[0091] 式 (2) 中, P 是子带数。

[0092] 接着, 相乘平均计算单元 202 将计算出的子带代表值频谱 $AVE2_p$ ($p=0 \sim P-1$) 输出到非线性变换单元 104。

[0093] 非线性变换单元 104 对于从相乘平均计算单元 202 输入的子带代表值频谱 $AVE2_p$ ($p=0 \sim P-1$), 使用式 (3), 对于各个代表值, 进行其值越大越增强特性的非线性变

换,计算第 1 子带对数代表值频谱 $AVE3_p$ ($p = 0 \sim P-1$)。这里,说明进行对数变换作为非线性变换处理的情况。

$$[0094] \quad AVE3_p = \log_{10}(AVE2_p) \quad (p = 0, \dots, P-1) \quad \dots (3)$$

[0095] 接着,非线性变换单元 104 使用式 (4),通过对于计算出的第 1 子带对数代表值频谱 $AVE3_p$ ($p = 0 \sim P-1$) 乘以子群数 Q 的倒数,计算第 2 子带对数代表值频谱 $AVE4_p$ ($p = 0 \sim P-1$)。

$$[0096] \quad AVE4_p = \frac{AVE3_p}{Q} \quad (p = 0, \dots, P-1) \quad \dots (4)$$

[0097] 相乘平均计算单元 202 中的式 (2) 的处理中,仅单纯地使各个子带的子带算术平均值频谱 $AVE1_p$ 相乘,但通过非线性变换单元 104 中的式 (4) 的处理而计算几何平均(相乘平均)。这样,在本实施方式中,使用式 (3) 变换为对数区域之后,使用式 (4) 乘以子群数 Q 的倒数。由此,能够将运算量大的方根的计算置换为单纯的除法运算。进而,在子群数 Q 是常数时,预先计算 Q 的倒数,由此能够将方根的计算置换为单纯的乘法运算,所以能够更削减运算量。

[0098] 接着,非线性变换单元 104 将使用式 (4) 计算出的第 2 子带对数代表值频谱 $AVE4_p$ ($p = 0 \sim P-1$) 输出到平滑化单元 105。

[0099] 再次返回到图 2,平滑化单元 105 对于从非线性变换单元 104 输入的第 2 子带对数代表值频谱 $AVE4_p$ ($p = 0 \sim P-1$),使用式 (5),在频域上进行平滑,并计算对数平滑化频谱 $AVE5_p$ ($p = 0 \sim P-1$)。

$$[0100] \quad AVE5_p = \frac{1}{MA_LEN} \cdot \sum_{i=p-\frac{MA_LEN-1}{2}}^{p+\frac{MA_LEN-1}{2}} AVE4_i \cdot W_i \left(\frac{MA_LEN-1}{2} \leq p \leq P-1 - \frac{MA_LEN-1}{2} \right) \quad \dots (5)$$

[0101] 另外,式 (5) 表示平滑化滤波处理,在式 (5) 中 MA_LEN 表示平滑化滤波的阶数, W_i 表示平滑化滤波器的权重。

[0102] 另外,式 (5) 是子带索引 p 为 $p \geq (MA_LEN-1)/2$, 并且 $p \leq P-1-(MA_LEN-1)/2$ 的情况下的对数平滑化频谱的计算方法。子带索引 p 在开头或最末尾附近的情况下,考虑边界条件,使用式 (6) 和式 (7) 分别将频谱平滑化。

[0103]

$$AVE5_p = \frac{1}{p + \frac{MA_LEN-1}{2} + 1} \cdot \sum_{i=0}^{p+\frac{MA_LEN-1}{2}} AVE4_i \cdot W_i \left(0 \leq p < \frac{MA_LEN-1}{2} \right) \quad \dots (6)$$

[0104]

$$AVE5_p = \frac{1}{P-1-p + \frac{MA_LEN-1}{2} + 1} \cdot \sum_{i=p-\frac{MA_LEN-1}{2}}^{P-1} AVE4_i \cdot W_i \left(P-1 - \frac{MA_LEN-1}{2} < p \leq P-1 \right) \quad \dots (7)$$

[0105] 另外,平滑化单元 105 也可以进行基于单纯的移动平均的平滑化作为如上所述进行基于平滑化滤波处理的平滑化处理 (W_i 对所有 i 为 1 时,为基于移动平均的平滑化)。另

外,窗口函数(权重)也可以利用汉宁窗(Hanning Window)或其他的窗口函数。

[0106] 接着,平滑化单元 105 将计算出的对数平滑化频谱 $AVE5_p (p = 0 \sim P-1)$ 输出到反非线性变换单元 106。

[0107] 反非线性变换单元 106 对从平滑化单元 105 输入的对数平滑化频谱 $AVE5_p (p = 0 \sim P-1)$ 进行对数逆变换并将对数平滑化频谱从对数区域的值变换为线性区域的值,作为反非线性变换。反非线性变换单元 106 使用式 (8),将对数平滑化频谱 $AVE5_p (p = 0 \sim P-1)$ 进行对数逆变换,计算平滑化频谱 $AVE6_p (p = 0 \sim P-1)$ 。

$$[0108] \quad AVE6_p = 10^{AVE5_p} \quad (p = 0, \dots, P-1) \quad \dots (8)$$

[0109] 进而,反非线性变换单元 106 将各个子带内的样本的值作为计算出的线性区域的平滑化频谱 $AVE6_p (p = 0 \sim P-1)$ 的值,计算全部样本的平滑化频谱。

[0110] 反非线性变换单元 106 输出全部样本的平滑化频谱值作为频谱平滑化装置 100 的处理结果。

[0111] 以上,说明了本发明的频谱平滑化装置及频谱平滑化方法。

[0112] 如上所述,在本实施方式中,子带分割单元 102 将输入频谱分割为多个子带,代表值计算单元 103 对每个子带,使用算术平均、以及乘法运算或几何平均来计算代表值,非线性变换单元 104 对于各个代表值,进行其值越大越增强特性的非线性变换,平滑化单元 105 将该每个子带的非线性变换后的代表值在频域上进行平滑。

[0113] 这样,将频谱的全部样本分割为多个子带,对于各个子带,通过组合算术平均(相加平均)和乘法运算或几何平均(相乘平均)而获得代表值,并在将该代表值进行非线性变换之后进行平滑,由此能够维持良好的语音质量,并且大幅度地减少处理运算量。

[0114] 如上所述,本发明中采用将子带内的样本的算术平均和乘法运算或几何平均结合而计算子带的代表值的结构,由此在将子带内的样本值的算术平均值(相加平均值)、即线性区域的平均值单纯地作为各个子带的代表值时,能够避免因子带内的样本值的大小的偏差而可能产生的语音质量的劣化。

[0115] 另外,在本实施方式中,举例说明了快速傅立叶变换(FFT)作为时间-频率变换处理,但本发明并不限于此,也同样适用于利用快速傅立叶变换(FFT)以外的时间-频率变换方法的情况。例如,在非专利文献 1 中,在计算听觉屏蔽(masking)值时(参照图 2),不是使用快速傅立叶变换(FFT),而是使用改进离散余弦变换(MDCT:Modified Discrete Cosine Transform)计算频率分量(频谱)。这样,在时间-频率变换处理单元中,即使对使用改进离散余弦变换(MDCT)或其他的时间-频率变换方法的结构,也能够同样地适用本发明。

[0116] 另外,在上述的结构中,相乘平均计算单元 202 仅将算术平均(相加平均)值频谱 $AVE1_q (q = 0 \sim Q-1)$ 相乘,而不进行方根的计算。因此,相乘平均计算单元 202 并不是准确地计算相乘平均值。这是因为,如上所述,在非非线性变换单元 104 中,作为非线性变换处理使用式 (3) 变换为对数区域之后,使用式 (4) 乘以子群数 Q 的倒数,由此能够将方根的计算置换为单纯的除法运算(乘法运算),从而能够更加削减运算量。

[0117] 因此,本发明并不限于上述的结构。例如,在下述结构中,也能够同样地适用本发明,即:在相乘平均计算单元 202 中,对于算术平均(相加平均)值频谱 $AVE1_q (q = 0 \sim Q-1)$,对每个子带将其全部子群的算术平均值频谱的值相乘之后,计算子群数的方根,并且将计算出的方根作为子带代表值频谱 $AVE2_p (p = 0 \sim P-1)$ 而输出到非线性变换单元 104 的

结构。也就是说,在任何情况下,平滑化单元 105 能够获得非线性变换后的每个子带的代表值。另外,在这样的情况下,在非线性变换单元 104 中,省略式 (4) 的运算即可。

[0118] 另外,在本实施方式中说明了下述情况,首先求子群的算术平均值,接着将子带内的全部子群的算术平均值的几何平均值作为每个子带的代表值的情况。但是,本发明并不限于此,也能够同样适用于在构成子群的样本数为 1 的情况,即,不计算各个子群的算术平均值,而将子带内的全部样本的几何平均值作为子带的代表值的情况。另外,在该结构中,如上所述,并未准确地计算几何平均值,而可以通过在进行非线性变换后乘以子群数的倒数,由此在对数区域计算几何平均值。

[0119] 另外,在上述说明中,在反非线性变换单元 106 中将同一子带内的样本的频谱值全都设为相同的值。但是,本发明并不限于此,在反非线性变换单元 106 的后级设置反平滑化处理单元,反平滑化处理单元也可以在各个子带内对每个样本附加权重而进行反平滑化处理。另外,该反平滑化处理也可以不进行与平滑化单元 105 完全相反的变换。

[0120] 另外,在以上的说明中以下述情况为例进行了说明,即:非线性变换单元 104 进行对数变换作为非线性变换处理,反非线性变换单元 106 进行对数逆变换作为反非线性变换处理的情况,但非线性变换处理并不限于此,也可以使用乘幂等,反非线性变换处理中进行该非线性变换处理的逆处理即可。但是,通过使用式 (4) 并乘以子群数 Q 的倒数,能够将方根的计算单纯地替换为除法运算(乘法运算),从而能够更加削减运算量,这是因为非线性变换单元 104 进行对数变换作为非线性变换。因此,在进行对数变换以外的处理作为非线性变换处理的情况下,通过对每个子群的算术平均值计算几何平均值,计算每个子带的代表值,对该代表值进行非线性处理即可。

[0121] 另外,作为子带数、子群数,例如举出下述情况为一例,即:输入信号的采样频率为 32kHz,1 帧长度为 20msec 时,也就是说输入信号为 640 样本时,将子带数设定为 80,将子群数设定为 2,将各个子群的样本数设定为 4,并将平滑化滤波的阶数设定为 7。但是,本发明并不限于该设定,也能够同样地适用于将这些值设定为其他数值的情况。

[0122] 另外,本发明的频谱平滑化装置及频谱平滑化方法能够适用于所有的语音编码装置及语音编码方法、语音解码装置及语音解码方法、语音识别装置及语音识别方法等在频谱区域中进行平滑化的频谱平滑化部分。例如,在专利文献 2 公开的频带扩展技术中,作为为了计算用于生成高频频谱的参数而进行的对低频频谱的预处理,进行下述处理,即:根据 LPC(Linear Predictive Coefficient:线性预测系数)计算频谱包络,使用计算出的频谱包络,从低频频谱中去除频谱包络,但是也能够使用将本发明的频谱平滑化方法适用于低频频谱而计算出的平滑化频谱,替代在专利文献 2 的频谱包络去除处理中利用的频谱包络。

[0123] 另外,在本实施方式中,说明了将输入的输入频谱 $S1(k)$ 分割为各个子带的样本数相等的 P 个(P 为 2 以上的整数)的子带的结构,但本发明并不限于此,也能够同样地适用于各个子带的样本数不同的结构。例如,举出了下述结构作为例子,即,将子带分割,以使越低频侧的子带,样本数越少,而越高频侧的子带,样本数越多。一般地,可以说人的听感,越高频侧,频率分辨能力越低,所以通过采用上述那样的结构,从而能够将频谱更高效率地进行平滑。另外,对于构成各个子带的子群也是一样的。也就是说,在本实施方式中说明了 Q 个的各个子群全部由 R 个的样本构成的情况,但本发明并不限于此,也能够同样地适用于

下述结构,即:将子群进行分割,以使越低频侧的子群,样本数越少,而越高频侧的子群,样本数越多。

[0124] 另外,在本实施方式中,作为平滑化处理以加权移动平均为例进行了说明,但本发明并不限于此,也能够同样地适用于各种平滑化处理。例如,如上所述,在各个子带的样本数不同(越高频,样本数越多)的结构中,移动平均的滤波器的抽头数不是左右对称,也可以越高频,抽头数越小。在越高频的子带,样本数越多时,通过使用高频侧的抽头数小的移动平均滤波器,能够进行听感上更合适的平滑化处理。当然,本发明也能够同样适用于利用越是高频、抽头越大的左右非对称的移动平均滤波器的情况。

[0125] (实施方式 2)

[0126] 在本实施方式中,说明将在实施方式 1 中说明了的频谱平滑化处理利用于专利文献 2 等公开了的频带扩展编码时的预处理的情况下的结构。

[0127] 图 5 是表示本发明实施方式 2 的具有编码装置和解码装置的通信系统的结构的方框图。在图 5 中,通信系统包括编码装置和解码装置,并分别处于经由传输路径能够进行通信的状态。另外,编码装置和解码装置通常都可以搭载于基站装置或通信终端装置等上使用。

[0128] 编码装置 301 将输入信号以 N 样本 (N 为自然数)为单位进行划分,将 N 样本作为 1 帧而对每帧进行编码。这里,将作为编码的对象输入信号表示为 x_n ($n = 0, \dots, N-1$)。 n 表示以 N 样本为单元划分的输入信号中第 $n+1$ 的信号元素。编码后的输入信息(编码信息)经由传输路径 302 发送给解码装置 303。

[0129] 解码装置 303 经由传输路径 302 接收从编码装置 301 发送的编码信息,并将其解码而获得输出信号。

[0130] 图 6 是表示图 5 所示的编码装置 301 的内部的主要结构的方框图。将输入信号的采样频率设为 SR_{input} ,下采样处理单元 311 对输入信号的采样频率从 SR_{input} 至 SR_{base} 为止进行下采样 ($SR_{base} < SR_{input}$),将下采样后的输入信号输入到第 1 层编码单元 312 作为下采样后输入信号。

[0131] 第 1 层编码单元 312 对于从下采样处理单元 311 输入的下采样后输入信号,使用例如 CELP(Code Excited Linear Prediction:码激励线性预测)方式的语音编码方法进行编码而生成第 1 层编码信息,并将生成的第 1 层编码信息输出到第 1 层解码单元 313 和编码信息汇集单元 317。

[0132] 第 1 层解码单元 313 对于从第 1 层编码单元 312 输入的第 1 层编码信息,使用例如 CELP 方式的语音解码方法进行解码而生成第 1 层解码信号,并将生成了的第 1 层解码信号输出到上采样处理单元 314。

[0133] 上采样处理单元 314 将从第一层解码单元 313 输入的第 1 层解码信号的采样频率进行从 SR_{base} 至 SR_{input} 为止的上采样,将上采样后的第 1 层解码信号输出到时间-频率变换处理单元 315 作为上采样后第 1 层解码信号。

[0134] 延迟单元 318 将规定的长度的延迟赋予输入信号。该延迟是用于校正下采样处理单元 311、第 1 层编码单元 312、第 1 层解码单元 313 以及上采样处理单元 314 中产生的时间迟延。

[0135] 时间-频率变换处理单元 315 在其内部具有缓冲器 $buf1_n$ 和 $buf2_n$ ($n = 0, \dots,$

N-1), 将输入信号 x_n 和从上采样处理单元 314 输入的上采样后第 1 层解码信号 y_n 进行改进离散余弦变换 (MDCT: Modified Discrete Cosine Transform)。

[0136] 接着, 对于时间-频率变换处理单元 315 中的正交变换处理, 说明其计算步骤和输出到内部缓冲器的数据输出。

[0137] 首先, 时间-频率变换处理单元 315 通过下述的式 (9) 和式 (10), 将缓冲器 $buf1_n$ 和缓冲器 $buf2_n$ 分别以“0”作为初始值进行初始化。

$$[0138] \quad buf1_n = 0 \quad (n = 0, \dots, N-1) \quad \dots (9)$$

$$[0139] \quad buf2_n = 0 \quad (n = 0, \dots, N-1) \quad \dots (10)$$

[0140] 接着, 时间-频率变换处理单元 315 对于输入信号 x_n 、上采样后第 1 层解码信号, 按照下述的式 (11) 和式 (12) 进行 MDCT, 求输入信号的 MDCT 系数 (以下称为“输入频谱”) $S2(k)$ 以及上采样后第 1 层解码信号 y_n 的 MDCT 系数 (以下称为“第 1 层解码频谱”) $S1(k)$ 。

$$[0141] \quad S2(k) = \frac{2}{N} \sum_{n=0}^{2N-1} x_n' \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (k = 0, \dots, N-1) \quad \dots (11)$$

$$[0142] \quad S1(k) = \frac{2}{N} \sum_{n=0}^{2N-1} y_n' \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (k = 0, \dots, N-1) \quad \dots (12)$$

[0143] 其中, k 表示 1 帧中的各个样本的索引。时间-频率变换处理单元 315 通过下述的式 (13) 求将输入信号 x_n 和缓冲器 $buf1_n$ 结合所得的矢量即 x_n' 。另外, 时间-频率变换处理单元 315 通过下述的式 (14) 求将上采样后第 1 层解码信号 y_n 和缓冲器 $buf2_n$ 结合所得的矢量即 y_n' 。

$$[0144] \quad x_n' = \begin{cases} buf1_n & (n = 0, \dots, N-1) \\ x_{n-N} & (n = N, \dots, 2N-1) \end{cases} \quad \dots (13)$$

$$[0145] \quad y_n' = \begin{cases} buf2_n & (n = 0, \dots, N-1) \\ y_{n-N} & (n = N, \dots, 2N-1) \end{cases} \quad \dots (14)$$

[0146] 接着, 时间-频率变换处理单元 315 通过式 (15) 和式 (16), 将缓冲器 $buf1_n$ 和 $buf2_n$ 更新。

$$[0147] \quad buf1_n = x_n \quad (n = 0, \dots, N-1) \quad \dots (15)$$

$$[0148] \quad buf2_n = y_n \quad (n = 0, \dots, N-1) \quad \dots (16)$$

[0149] 另外, 时间-频率变换处理单元 315 将输入频谱 $S2(k)$ 和第 1 层解码频谱 $S1(k)$ 输出到第 2 层编码单元 316。

[0150] 第 2 层编码单元 316 使用从时间-频率变换处理单元 315 输入输入频谱 $S2(k)$ 和第 1 层解码频谱 $S1(k)$ 生成第 2 层编码信息, 并将生成的第 2 层编码信息输出到编码信息汇集单元 317。另外, 在后面叙述第 2 层编码单元 316 的细节。

[0151] 编码信息汇集单元 317 将从第 1 层编码单元 312 输入的第 1 层编码信息和从第 2 层编码单元 316 输入的第 2 层编码信息进行汇集, 并且对于汇集后的信息源码, 如果有必要则附加了传输差错码等之后将其输出到传输路径 302 作为编码信息。

[0152] 接着, 使用图 7 说明图 6 所示的第 2 层编码单元 316 的内部的主要结构。

[0153] 第 2 层编码单元 316 包括: 频带分割单元 360、频谱平滑化单元 361、滤波器状态设定单元 362、滤波单元 363、搜索单元 364、音调系数设定单元 365、增益编码单元 366 以及复

用单元 367, 各个单元进行以下的动作。

[0154] 频带分割单元 360 将从时间 - 频率变换处理单元 315 输入的输入频谱 $S_2(k)$ 的高频部 ($FL \leq k < FH$) 分割为 P 个子带 SB_p ($p = 0, 1, \dots, P-1$)。另外, 频带分割单元 360 将分割后的各个子带的带宽 BW_p ($p = 0, 1, \dots, P-1$) 和开头索引 BS_p ($p = 0, 1, \dots, P-1$) ($FL \leq BS_p < FH$) 输出到滤波单元 363、搜索单元 364 以及复用单元 367 作为频带分割信息。以下, 将输入频谱 $S_2(k)$ 中的、与子带 SB_p 对应的部分记为子带频谱 $S_{2p}(k)$ ($BS_p \leq k < BS_p + BW_p$)。

[0155] 频谱平滑化单元 361 对于从时间 - 频率变换处理单元 315 输入的第 1 层解码频谱 $S_1(k)$ ($0 \leq k < FL$) 进行平滑化处理, 并将平滑化处理后的平滑化第 1 层解码频谱 $S_1'(k)$ ($0 \leq k < FL$) 输出到滤波器状态设定单元 362。

[0156] 图 8 表示频谱平滑化单元 361 的内部结构。频谱平滑化单元 361 主要由子带分割单元 102、代表值计算单元 103、非线性变换单元 104、平滑化单元 105 以及反非线性变换单元 106 构成。这里, 各个处理单元与实施方式 1 中说明了的处理单元相同, 所以附加相同的标号并省略其说明。

[0157] 滤波器状态设定单元 362 将从频谱平滑化单元 361 输入的平滑化第 1 层解码频谱 $S_1'(k)$ ($0 \leq k < FL$) 设定为在后级的滤波单元 363 中使用的滤波器的内部状态。在滤波单元 363 中的全频带的频谱 $S(k)$ 的 $0 \leq k < FL$ 的频带中, 存储平滑化第 1 层解码频谱 $S_1'(k)$ 作为滤波器的内部状态 (滤波器状态)。

[0158] 滤波单元 363 包括多抽头的音调滤波器, 基于由滤波器状态设定单元 362 设定的滤波器状态、从音调系数设定单元 365 输入的音调系数以及从频带分割单元 360 输入的频带分割信息, 对第 1 层解码频谱进行滤波, 并计算各个子带 SB_p ($p = 0, 1, \dots, P-1$) 的估计值频谱 $S_{2p}'(k)$ ($BS_p \leq k < BS_p + BW_p$) ($p = 0, 1, \dots, P-1$) (以下称为“子带 SB_p 的估计频谱”)。滤波单元 363 将子带 SB_p 的估计频谱 $S_{2p}'(k)$ 输出到搜索单元 364。另外, 在后面叙述滤波单元 363 中的滤波处理的细节。另外, 假设多抽头的抽头数能够取 1 以上的任意值 (整数)。

[0159] 搜索单元 364 基于从频带分割单元 360 输入的频带分割信息, 计算从滤波单元 363 输入的子带 SB_p 的估计频谱 $S_{2p}'(k)$ 和从时间 - 频率变换处理单元 315 输入的输入频谱 $S_2(k)$ 的高频部 ($FL \leq k < FH$) 中的各个子带频谱 $S_{2p}(k)$ 之间的相似度。通过例如相关运算等进行该相似度的计算。另外, 滤波单元 363、搜索单元 364 以及音调系数设定单元 365 的处理, 对每个子带构成闭环的搜索处理, 在各个闭环中, 搜索单元 364 通过使从音调系数设定单元 365 输入到滤波单元 363 的音调系数 T 产生各种变化, 计算与各个音调系数对应的相似度。搜索单元 364 在每个子带的闭环中, 例如, 在与子带 SB_p 对应的闭环中求相似度为最大的最合适音调系数 T_p' (其中范围为 $T_{min} \sim T_{max}$), 并将 P 个最合适音调系数输出到复用单元 367。搜索单元 364 使用各个最合适音调系数 T_p' , 计算与各个子带 SB_p 相似的、第 1 层解码频谱的一部分频带。另外, 搜索单元 364 将与各个最合适音调系数 T_p' ($p = 0, 1, \dots, P-1$) 对应的估计频谱 $S_{2p}'(k)$ 输出到增益编码单元 366。另外, 在后面叙述搜索单元 364 中的最合适音调系数 T_p' ($p = 0, 1, \dots, P-1$) 的搜索处理的细节。

[0160] 音调系数设定单元 365 在搜索单元 364 的控制之下, 与滤波单元 363 和搜索单元 364 一起进行与第 1 子带 SB_0 对应的闭环的搜索处理时, 使音调系数 T 在预先决定的搜索范围 $T_{min} \sim T_{max}$ 中每次稍稍变化, 并将其依次输出到滤波单元 363。

[0161] 增益编码单元 366 计算从时间-频率变换处理单元 315 输入的输入频谱 $S2(k)$ 的高频部 ($FL \leq k < FH$) 的增益信息。具体而言,增益编码单元 366 将频带 $FL \leq k < FH$ 分割为 J 个子带,并求输入频谱 $S2(k)$ 的每个子带的频谱功率。此时,用下述的式 (17) 表示第 $j+1$ 子带的频谱功率 B_j 。

$$[0162] \quad B_j = \sum_{k=BL_j}^{BH_j} S2(k)^2 \quad (j=0, \dots, J-1) \quad \dots (17)$$

[0163] 在式 (17) 中, BL_j 表示第 $j+1$ 子带的最小频率, BH_j 表示第 $j+1$ 子带的最大频率。另外,增益编码单元 366 使从搜索单元 364 输入的各个子带的估计频谱 $S2'_p(k)$ ($p=0, 1, \dots, P-1$) 在频域连续而构成输入频谱的高频部的估计频谱 $S2'(k)$ 。另外,增益编码单元 366 与对于输入频谱 $S2(k)$ 计算出频谱功率时同样,按照下述的式 (18) 计算估计频谱 $S2'(k)$ 的每个子带的频谱功率 B'_j 。接着,增益编码单元 366 按照式 (19) 计算对输入频谱 $S2(k)$ 的估计频谱 $S2'(k)$ 的每个子带的频谱功率的变动量 V_j 。

$$[0164] \quad B'_j = \sum_{k=BL_j}^{BH_j} S2'(k)^2 \quad (j=0, \dots, J-1) \quad \dots (18)$$

$$[0165] \quad V_j = \sqrt{\frac{B_j}{B'_j}} \quad (j=0, \dots, J-1) \quad \dots (19)$$

[0166] 另外,增益编码单元 366 将变动量 V_j 编码,将与编码后的变动量 VQ_j 对应的索引输出到复用单元 367。

[0167] 复用单元 367 将从频带分割单元 360 输入的频带分割信息、从搜索单元 364 输入的各个子带 SB_p ($p=0, 1, \dots, P-1$) 的最适合音调系数 T_p' 以及从增益编码单元 366 输入的变动量 VQ_j 的索引进行复用,作为第 2 层编码信息,并将其输出到编码信息汇集单元 317。另外,也可以将 T_p' 和 VQ_j 的索引直接输入到编码信息汇集单元 317,并通过编码信息汇集单元 317 将其与第 1 层编码信息复用。

[0168] 接着,使用图 9 说明图 7 所示的滤波单元 363 中的滤波处理的细节。

[0169] 滤波单元 363 使用从滤波器状态设定单元 362 输入的滤波器状态、从音调系数设定单元 365 输入的音调系数 T 以及从频带分割单元 360 输入的频带分割信息,对于子带 SB_p ($p=0, 1, \dots, P-1$),生成频带 $BS_p \leq k < BS_p+BW_p$ ($p=0, 1, \dots, P-1$) 中的估计频谱。用下述的式 (20) 表示滤波单元 363 中使用的滤波器的传递函数 $F(z)$ 。

[0170] 以下,以子带 SB_p 为例子,说明生成子带频谱 $S2_p(k)$ 的估计频谱 $S2'_p(k)$ 的处理。

$$[0171] \quad F(z) = \frac{1}{1 - \sum_{i=-M}^M \beta_i z^{-T+i}} \quad \dots (20)$$

[0172] 在式 (20) 中, T 表示由音调系数设定单元 365 给出的音调系数, β_i 表示预先存储在内部的滤波系数。例如,抽头数为 3 时,滤波系数的候补例举为 $(\beta_{-1}, \beta_0, \beta_1) = (0.1, 0.8, 0.1)$ 。其他, $(\beta_{-1}, \beta_0, \beta_1) = (0.2, 0.6, 0.2)$ 、 $(0.3, 0.4, 0.3)$ 等值也合适。另外,也可以是 $(\beta_{-1}, \beta_0, \beta_1) = (0.0, 1.0, 0.0)$ 的值,此时意味着:对于频带 $0 \leq k < FL$ 的第 1 层解码频谱的一部分频带,不使其形状变化而直接将其复制到 $BS_p \leq k < BS_p+BW_p$ 的频带。另外,在式 (20) 中设为 $M=1$ 。 M 是与抽头数有关的指标。

[0173] 在滤波单元 363 中的全频带的频谱 $S(k)$ 的 $0 \leq k < FL$ 的频带中, 存储平滑化第 1 层解码频谱 $S1'(k)$ 作为滤波器的内部状态 (滤波器状态)。

[0174] 在 $S(k)$ 的 $BS_p \leq k < BS_p + BW_p$ 的频带中, 通过下述步骤的滤波处理, 存储子带 SB_p 的估计频谱 $S2_p'(k)$ 。也就是说, 一般将比该 k 低 T 的频率的频谱 $S(k-T)$ 代入 $S2_p'(k)$ 。但是, 为了增加频谱的平滑性, 实际上, 对所有的 i 将规定的滤波系数 β_i 乘以与频谱 $S(k-T)$ 相距 i 的附近的频谱 $S(k-T+i)$ 所得的频谱 $\beta_i \cdot S(k-T+i)$ 相加, 将相加所得的频谱代入 $S2_p'(k)$ 。用下述的式 (21) 表示该处理。

$$[0175] \quad S2_p'(k) = \sum_{i=-1}^1 \beta_i \cdot S2(k-T+i)^2 \quad \dots (21)$$

[0176] 从频率低的 $k = BS_p$ 开始, 依序使 k 在 $BS_p \leq k < BS_p + BW_p$ 的范围内变化而进行上述运算, 从而计算 $BS_p \leq k < BS_p + BW_p$ 中的估计频谱 $S2_p'(k)$ 。

[0177] 每次由音调系数设定单元 365 给出音调系数 T 时, 在 $BS_p \leq k < BS_p + BW_p$ 的范围内, 每次将 $S(k)$ 清零而进行上述的滤波处理。也就是说, 每次音调系数 T 变化时计算 $S(k)$, 并将其输出到搜索单元 364。

[0178] 图 10 是表示图 7 所示的搜索单元 364 中对于子带 SB_p 搜索最合适音调系数 T_p' 的处理的步骤的流程图。另外, 搜索单元 364 通过反复进行图 10 所示的步骤, 搜索与各个子带 SB_p ($p = 0, 1, \dots, P-1$) 对应的最合适音调系数 T_p' ($p = 0, 1, \dots, P-1$)。

[0179] 首先, 搜索单元 364 将用于保存相似度的最小值的变量即最小相似度 D_{\min} 初始化为“+ ∞ ”(ST110)。接着, 搜索单元 364 按照下述的式 (22), 计算某一音调系数中的输入频谱 $S2(k)$ 的高频部 ($FL \leq k < FH$) 和估计频谱 $S2_p'(k)$ 之间的相似度 D (ST120)。

[0180]

$$D = \sum_{k=0}^{M'} S2(BS_p + k) \cdot S2(BS_p + k) - \frac{\left(\sum_{k=0}^{M'} S2(BS_p + k) \cdot S2'(BS_p + k) \right)^2}{\sum_{k=0}^{M'} S2'(BS_p + k) \cdot S2'(BS_p + k)} \quad (0 < M' \leq BW_p) \quad \dots (22)$$

[0181] 在式 (22) 中, M' 表示计算相似度 D 时的样本数, 可以是各个子带的带宽以下的任意值。另外, 在式 (22) 中, $S2_p'(k)$ 不存在, 但这是因为使用 BS_p 和 $S2'(k)$ 表示 $S2_p'(k)$ 。

[0182] 接着, 搜索单元 364 判定计算出的相似度 D 是否小于最小相似度 D_{\min} (ST130)。在 ST120 中计算出的相似度 D 小于最小相似度 D_{\min} 时 (ST130: “是”), 搜索单元 364 将相似度 D 代入最小相似度 D_{\min} (ST140)。另一方面, 在 ST120 中计算出的相似度 D 为最小相似度 D_{\min} 以上时 (ST130: “否”), 搜索单元 364 判定整个搜索范围的处理是否结束。也就是说, 搜索单元 364 对于搜索范围内的所有音调系数的各个音调系数, 判定在 ST120 中是否按照上述的式 (22) 计算出相似度 (ST150)。在整个搜索范围的处理未结束时 (ST150: “否”), 搜索单元 364 将处理再次返回到 ST120。另外, 搜索单元 364 在上一次的 ST120 的步骤中根据式 (22) 计算出相似度的情况是指: 对于不同的音调系数, 根据式 (22) 计算相似度。另一方面, 在整个搜索范围的处理结束时 (ST150: “是”), 搜索单元 364 将与最小相似度 D_{\min} 对应的音调系数 T 输出到复用单元 367 作为最合适音调系数 T_p' (ST160)。

[0183] 接着, 说明图 5 所示的解码装置 303。

[0184] 图 11 是表示解码装置 303 的内部的主要结构的方框图。

[0185] 在图 11 中,编码信息分离单元 331 从输入的编码信息中分离第 1 层编码信息和第 2 层编码信息,将第 1 层编码信息输出到第 1 层解码单元 332,并将第 2 层编码信息输出到第 2 层解码单元 335。

[0186] 第一层解码单元 332 对于从编码信息分离单元 331 输入的第 1 层编码信息进行解码,将生成的第 1 层解码信号输出到上采样处理单元 333。这里,由于第 1 层解码单元 332 的动作与图 6 所示的第 1 层解码单元 313 相同,所以省略详细的说明。

[0187] 上采样处理单元 333 对于从第一层解码单元 332 输入的第 1 层解码信号,进行采样频率从 SR_{base} 至 SR_{input} 为止的上采样的处理,并将获得的上采样后第 1 层解码信号输出到时间-频率变换处理单元 334。

[0188] 时间-频率变换处理单元 334 对于从上采样处理单元 333 输入的上采样后第 1 层解码信号进行正交变换处理 (MDCT),并将获得的上采样后第 1 层解码信号的 MDCT 系数 (以下,称为第 1 层解码频谱) $S1(k)$ 输出到第 2 层解码单元 335。这里,由于时间-频率变换处理单元 334 的动作与图 6 所示的时间-频率变换处理单元 315 的对上采样后第 1 层解码信号的处理相同,所以省略详细的说明。

[0189] 第 2 层解码单元 335 使用从时间-频率变换处理单元 334 输入的第 1 层解码频谱 $S1(k)$ 以及从编码信息分离单元 331 输入的第 2 层编码信息,生成含有高频分量的第 2 层解码信号而将其作为输出信号输出。

[0190] 图 12 是表示图 11 所示的第二层解码单元 335 的内部的主要结构的方框图。

[0191] 分离单元 351 将从编码信息分离单元 331 输入的第 2 层编码信息分离为含有各个子带的带宽 BW_p ($p = 0, 1, \dots, P-1$) 和开头索引 BS_p ($p = 0, 1, \dots, P-1$) ($FL \leq BS_p < FH$) 的频带分割信息、与滤波有关的信息即最合适音调系数 T_p' ($p = 0, 1, \dots, P-1$) 以及与增益有关的信息即编码后变动量 VQ_j ($j = 0, 1, \dots, J-1$) 的索引。另外,分离单元 351 将频带分割信息和最合适音调系数 T_p' ($p = 0, 1, \dots, P-1$) 输出到滤波单元 354,并将编码后变动量 VQ_j ($j = 0, 1, \dots, J-1$) 的索引输出到增益解码单元 355。另外,在编码信息分离单元 331 中,在已经分离为频带分割信息、 T_p' ($p = 0, 1, \dots, P-1$) 以及 VQ_j ($j = 0, 1, \dots, J-1$) 的索引时,也可以不配置分离单元 351。

[0192] 频谱平滑化单元 352 对于从时间-频率变换处理单元 334 输入的第 1 层解码频谱 $S1(k)$ ($0 \leq k < FL$) 进行平滑化处理,并将平滑化后的平滑化第 1 层解码频谱 $S1'(k)$ ($0 \leq k < FL$) 输出到滤波器状态设定单元 353。由于频谱平滑化单元 352 的处理与第 2 层编码单元 316 内的频谱平滑化单元 361 相同,所以这里省略其说明。

[0193] 滤波器状态设定单元 353 将从频谱平滑化单元 352 输入的平滑化第 1 层解码频谱 $S1'(k)$ ($0 \leq k < FL$) 设定为在滤波单元 354 中使用的滤波器状态。这里,为了方便而将滤波单元 354 中的全频带 $0 \leq k < FH$ 的频谱称为 $S(k)$ 时,在 $S(k)$ 的 $0 \leq k < FL$ 的频带中存储平滑化第 1 层解码频谱 $S1'(k)$ 作为滤波器的内部状态 (滤波器状态)。这里,由于滤波器状态设定单元 353 的结构和动作与图 7 所示的滤波器状态设定单元 362 相同,所以省略详细的说明。

[0194] 滤波单元 354 包括多抽头 (抽头数大于 1) 的音调滤波器。滤波单元 354 基于从分离单元 351 输入的频带分割信息、由滤波器状态设定单元 353 设定的滤波器状态、从分离

单元 351 输入的音调系数 T_p' ($p = 0, 1, \dots, P-1$) 以及预先存储在内部的滤波系数, 对平滑化第 1 层解码频谱 $S1'(k)$ 进行滤波, 计算如上述的式 (21) 所示的、各个子带 SB_p ($p = 0, 1, \dots, P-1$) 的估计值频谱 $S2_p'(k)$ ($BS_p \leq k < BS_p + BW_p$) ($p = 0, 1, \dots, P-1$)。滤波单元 354 也使用上述的式 (20) 所示的滤波函数。但是, 假设此时的滤波处理和滤波函数是将式 (20)、式 (21) 中的 T 替换为 T_p' 。

[0195] 增益解码单元 355 将从分离单元 351 输入的、编码后变动量 VQ_j 的索引进行解码, 求变动量 V_j 的量化值即变动量 VQ_j 。

[0196] 频谱调整单元 356 使从滤波单元 354 输入的各个子带 SB_p ($p = 0, 1, \dots, P-1$) 的估计值频谱 $S2_p'(k)$ ($BS_p \leq k < BS_p + BW_p$) ($p = 0, 1, \dots, P-1$) 在频域连续而求输入频谱的估计频谱 $S2'(k)$ 。另外, 频谱调整单元 356 根据下述的式 (23), 将从增益解码单元 355 输入的每个子带的变动量 VQ_j 乘以估计频谱 $S2'(k)$ 。由此, 频谱调整单元 356 调整估计频谱 $S2'(k)$ 的频带 $FL \leq k < FH$ 中的频谱形状, 生成解码频谱 $S3(k)$ 并将其输出到时间-频率变换处理单元 357。

$$[0197] \quad S3(k) = S2'(k) \cdot VQ_j \quad (BL_j \leq k \leq BH_j, \text{ for all } j) \quad \dots (23)$$

[0198] 接着, 如式 (24) 所示, 频谱调整单元 356 将从时间-频率变换处理单元 334 输入的第 1 层解码频谱 $S1(k)$ ($0 \leq k < FL$) 代入解码频谱 $S3(k)$ 的低频部 ($0 \leq k < FL$)。这里, 解码频谱 $S3(k)$ 的低频部 ($0 \leq k < FL$) 由第 1 层解码频谱 $S1(k)$ 构成, 解码频谱 $S3(k)$ 的高频部 ($FL \leq k < FH$) 由频谱形状调整后的估计频谱 $S2'(k)$ 构成。

$$[0199] \quad S3(k) = S1(k) \quad (0 \leq k \leq FL) \quad \dots (24)$$

[0200] 时间-频率变换处理单元 357 将从频谱调整单元 356 输入的了解码频谱 $S3(k)$ 正交变换为时域的信号, 并将获得的第 2 层解码信号作为输出信号输出。这里, 根据需要进行适当的开窗及叠加加法等处理, 避免在帧间产生的中断。

[0201] 以下, 说明时间-频率变换处理单元 357 中的具体处理。

[0202] 时间-频率变换处理单元 357 在内部具有缓冲器 $buf'(k)$, 如下述的式 (25) 所示那样, 将缓冲器 $buf'(k)$ 初始化。

$$[0203] \quad buf'(k) = 0 \quad (k = 0, \dots, N-1) \quad \dots (25)$$

[0204] 另外, 时间-频率变换处理单元 357 使用从频谱调整单元 356 输入的第 2 层解码频谱 $S3(k)$ 并根据下述的式 (26), 求第 2 层解码信号 y_n'' 并将其输出。

$$[0205] \quad y_n'' = \frac{2}{N} \sum_{k=0}^{2N-1} Z4(k) \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (n = 0, \dots, N-1) \quad \dots (26)$$

[0206] 在式 (26) 中, 如下述的式 (27) 所示, $Z4(k)$ 是将解码频谱 $S3(k)$ 和缓冲器 $buf'(k)$ 结合所得的矢量。

$$[0207] \quad Z4(k) = \begin{cases} buf'(k) & (k = 0, \dots, N-1) \\ S3(k) & (k = N, \dots, 2N-1) \end{cases} \quad \dots (27)$$

[0208] 接着, 时间-频率变换处理单元 357 根据下述的式 (28), 对缓存器 $buf'(k)$ 进行更新。

$$[0209] \quad buf'(k) = S3(k) \quad (k = 0, \dots, N-1) \quad \dots (28)$$

[0210] 接着, 时间-频率变换处理单元 357 将解码信号 y_n'' 作为输出信号输出。

[0211] 这样, 根据本实施方式, 在使用低频部的频谱进行频带扩展并估计高频部的频谱

的编码 / 解码中,对于低频部的频谱进行组合了相加平均和相乘平均的平滑化处理作为预处理。由此,即使对于频带扩展编码方式,也不使解码信号中产生大的质量劣化,并能够大幅度地削减处理运算量。

[0212] 另外,在本实施方式中说明了下述结构,即,在频带扩展编码时,对于解码所得的低频解码频谱进行平滑化处理,使用平滑化后的低频解码频谱估计高频频谱并进行编码的结构,但本发明并不限于此,本发明也同样能够适用于下述结构,即,对于输入信号的低频频谱进行平滑化处理,根据平滑化后的输入频谱估计高频频谱并进行编码的结构。

[0213] 另外,本发明的频谱平滑化装置和频谱平滑化方法并不限于上述实施方式,也能够进行各种变更后实施。例如,也可以适当地组合各个实施方式而实施。

[0214] 另外,将信号处理程序记录和写入到存储器、磁盘、磁带、CD、DVD 等可进行机械性读取的记录媒体上并进行动作时,也能够适用本发明,并能够获得与本实施方式相同的作用和效果。

[0215] 另外,在上述实施方式中,举例说明了以硬件构成本发明的情况,但本发明也可以通过软件来实现。

[0216] 另外,用于上述实施方式的说明中使用的各功能块通常被作为集成电路的 LSI 来实现。这些功能块既可以被单独地集成为一个芯片,也可以包含一部分或全部地被集成为一个芯片。虽然这里称为 LSI,但根据集成程度,可以被称为 IC、系统 LSI、超大 LSI (Super LSI)、或特大 LSI (Ultra LSI)。

[0217] 另外,实现集成电路化的方法不仅限于 LSI,也可使用专用电路或通用处理器来实现。也可以使用可在 LSI 制造后编程的 FPGA (Field Programmable Gate Array :现场可编程门阵列),或者可重构 LSI 内部的电路单元的连接和设定的可重构处理器。

[0218] 再者,随着半导体的技术进步或随之派生的其它技术的出现,如果出现能够替代 LSI 的集成电路化的新技术,当然可利用该新技术进行功能块的集成化。还存在着适用生物技术等的可能性。

[0219] 2008 年 8 月 8 日提出的日本专利申请第 2008-205645 以及 2009 年 4 月 10 日提出的日本专利申请第 2009-096222 号所包含的说明书、附图以及说明书摘要的公开内容全部被引用于本申请。

[0220] 工业实用性

[0221] 本发明的频谱平滑化装置、编码装置、解码装置、通信终端装置、基站装置以及频谱平滑化方法能够以小运算量实现在频谱区域的平滑化,例如能够适用于分组通信系统、移动通信系统等。

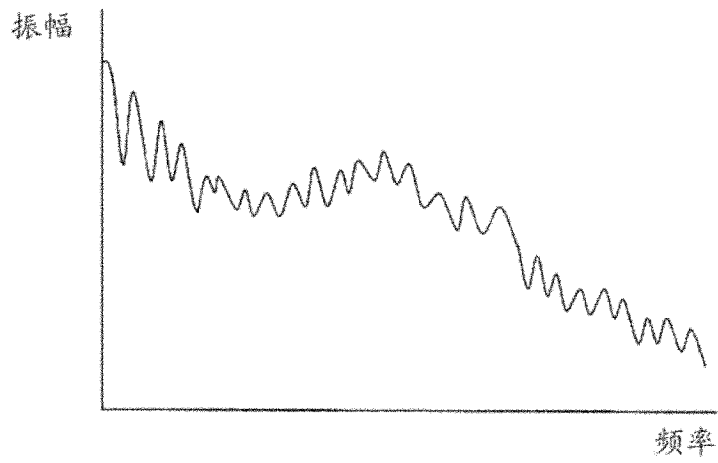


图 1A

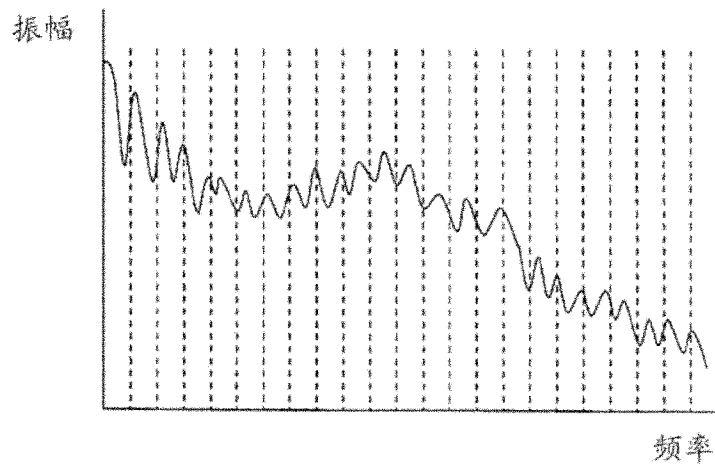


图 1B

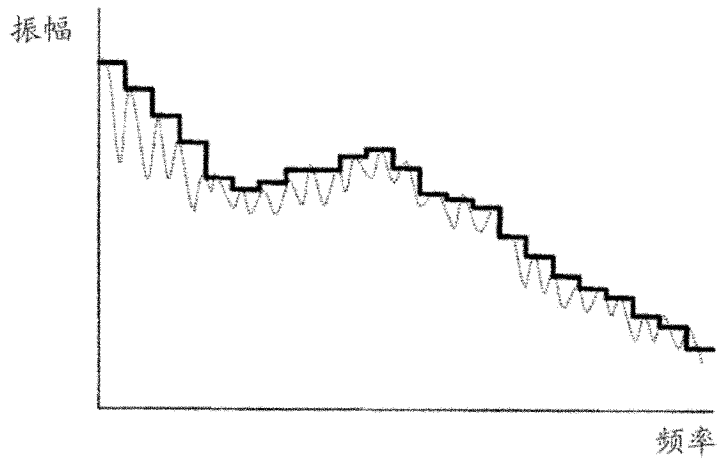


图 1C

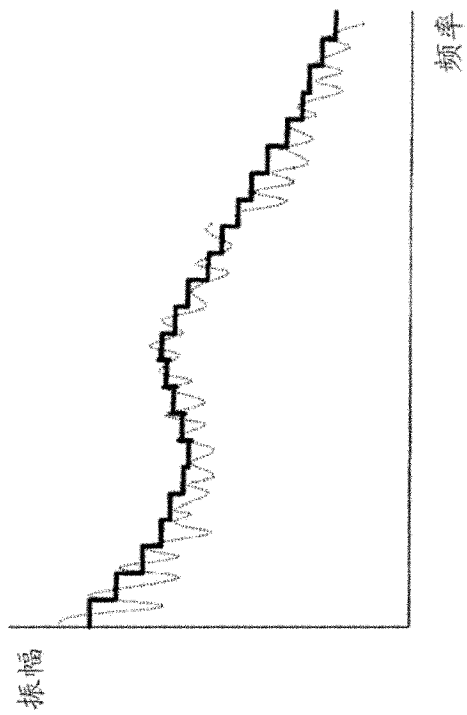


图 1D

100

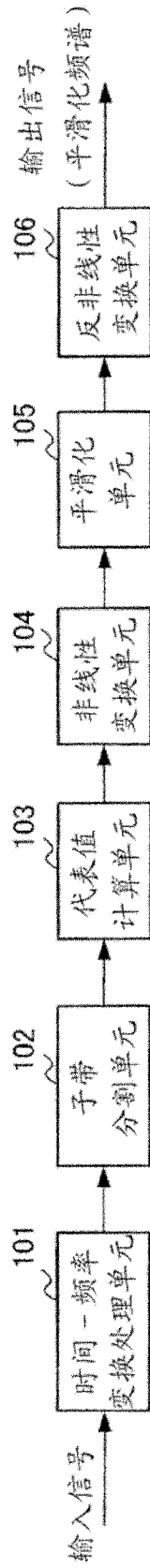


图 2

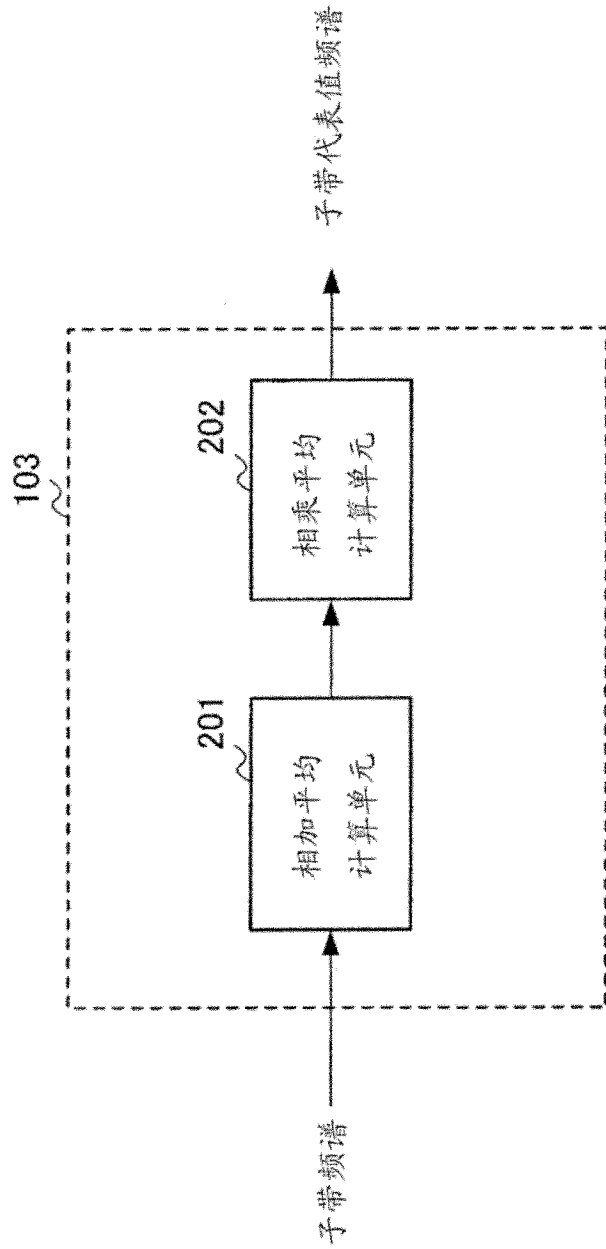


图 3

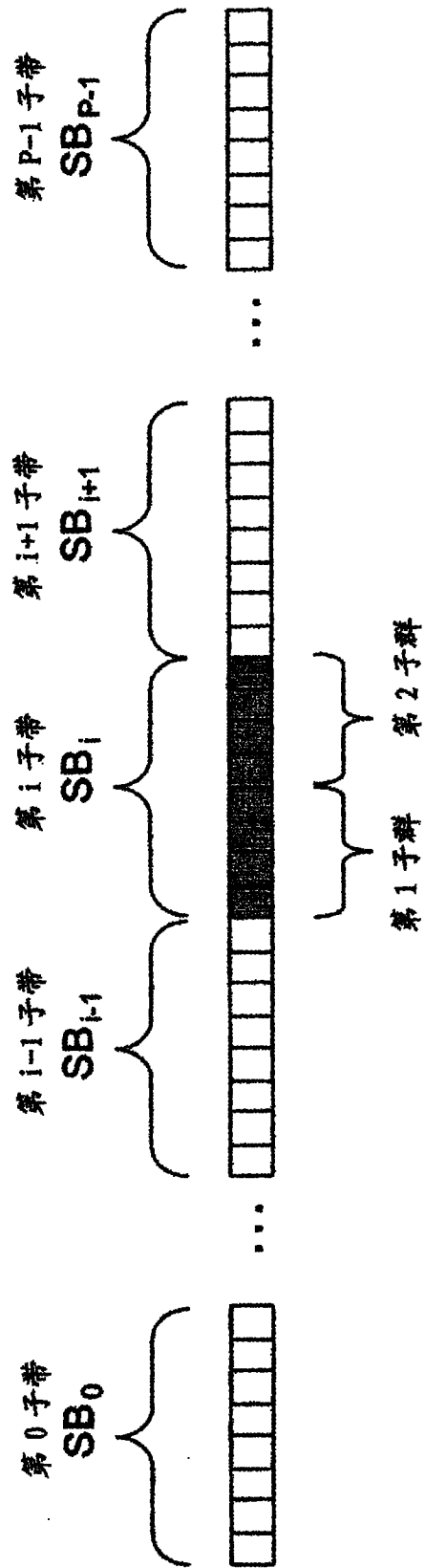


图 4

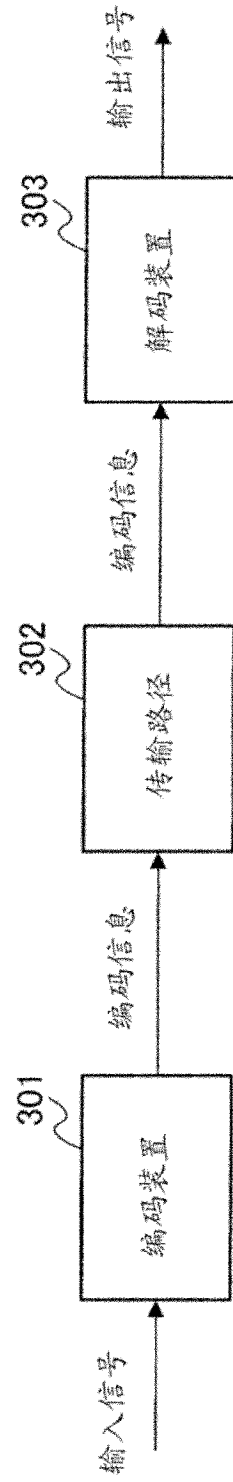


图 5

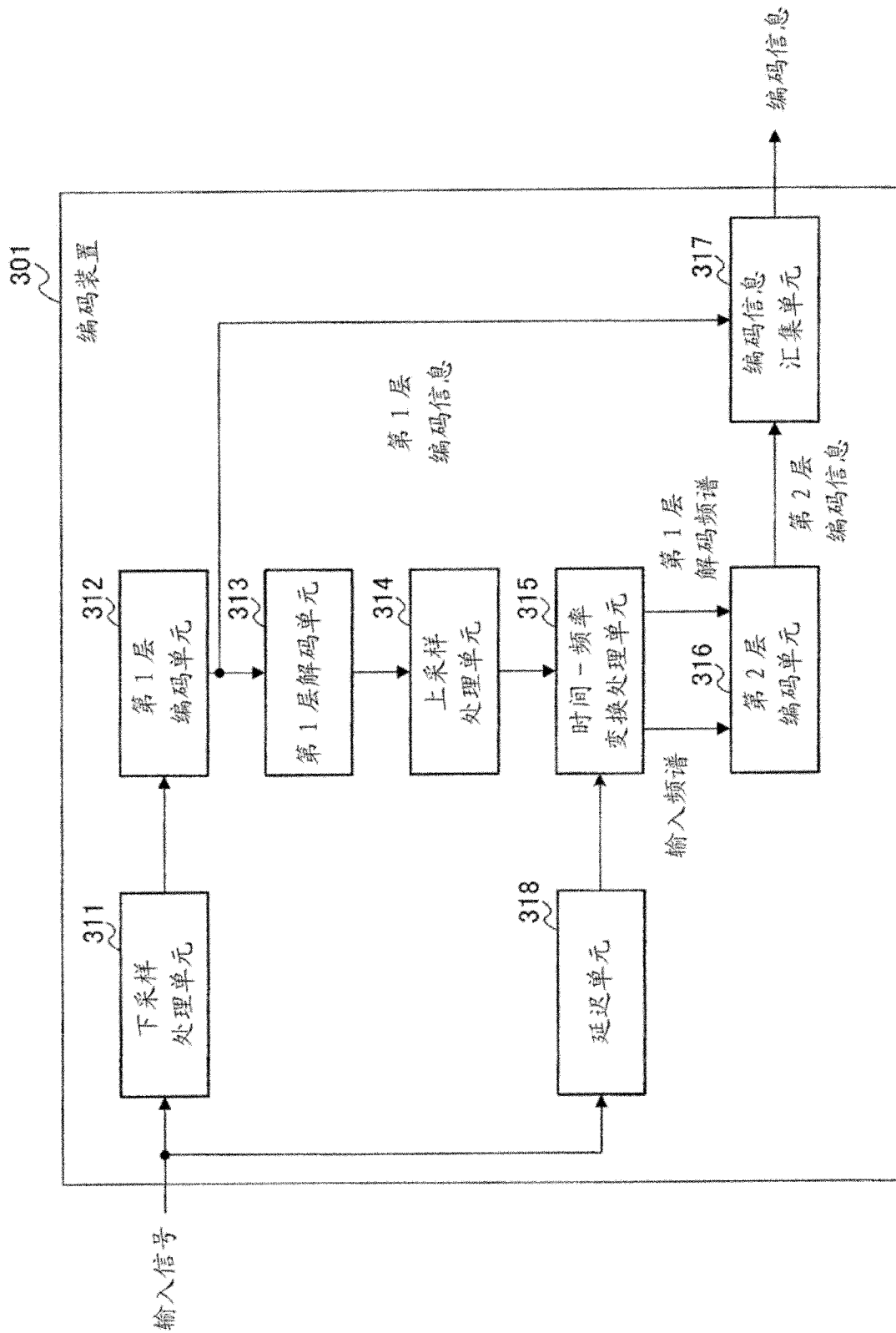


图 6

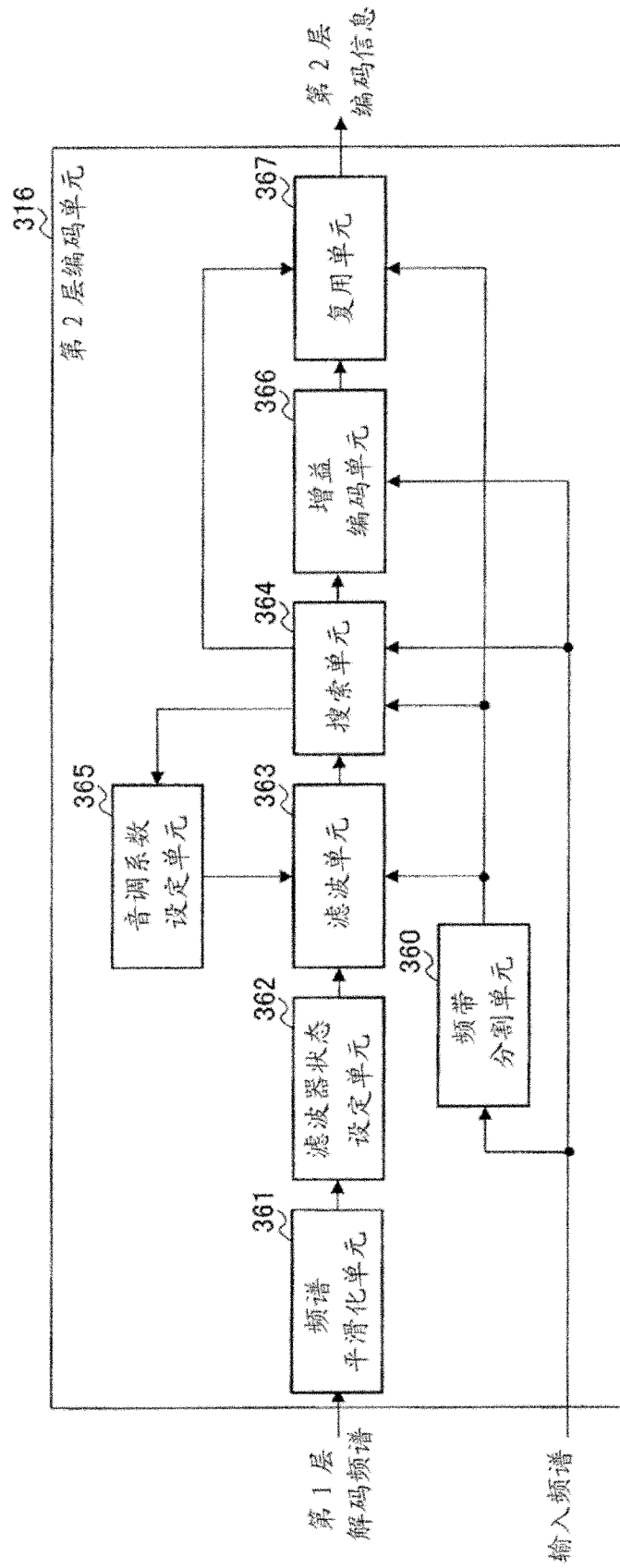


图 7

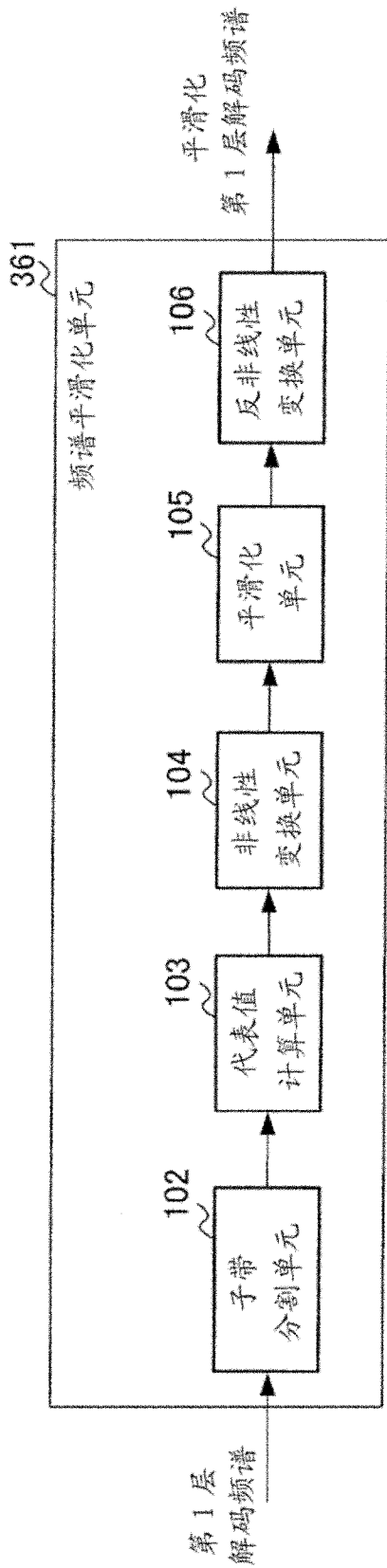


图 8

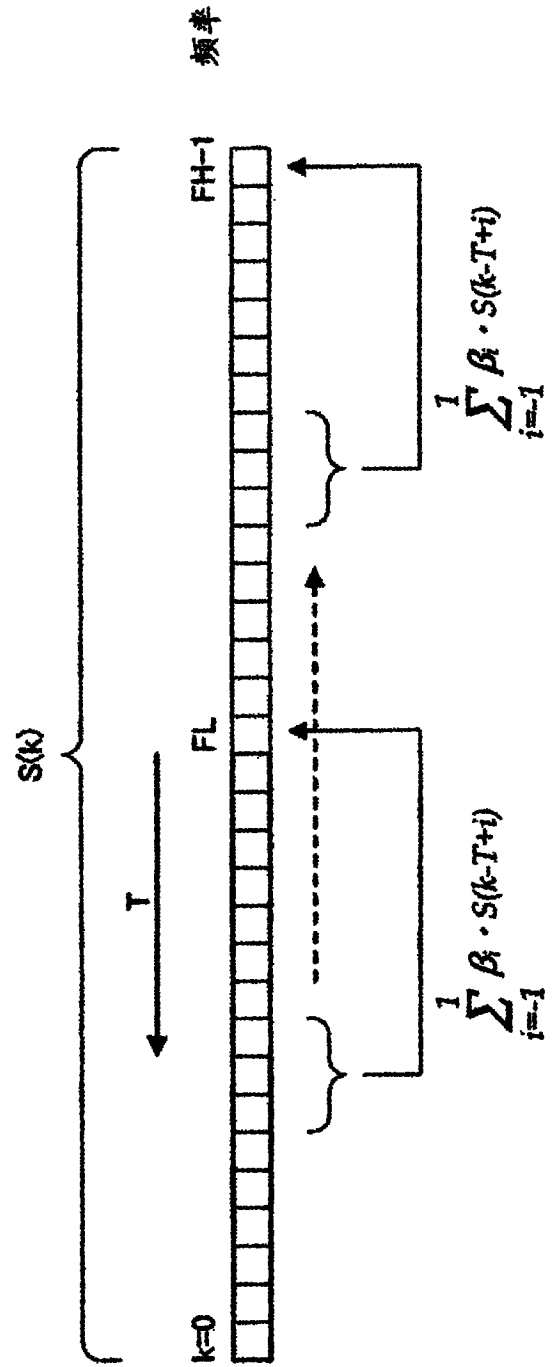


图 9

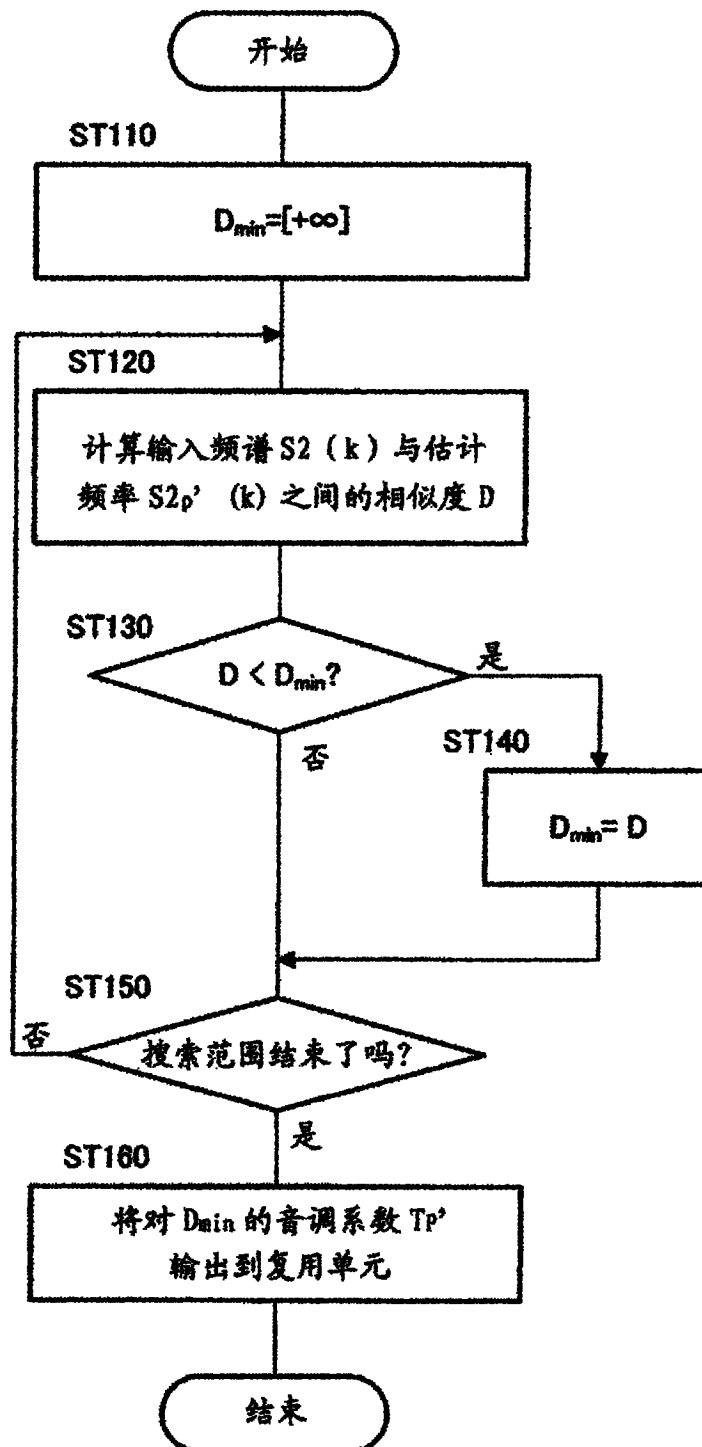


图 10

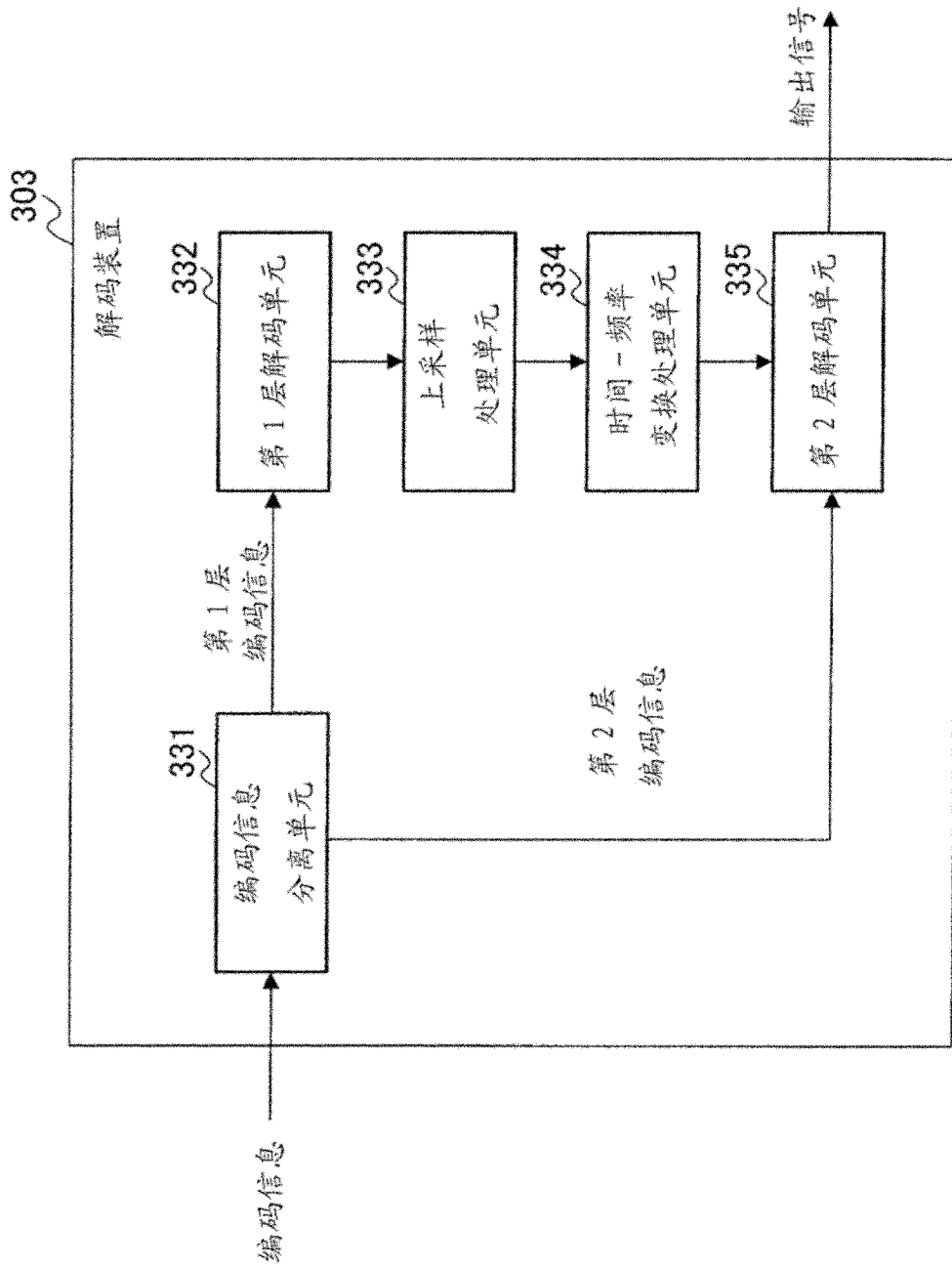


图 11

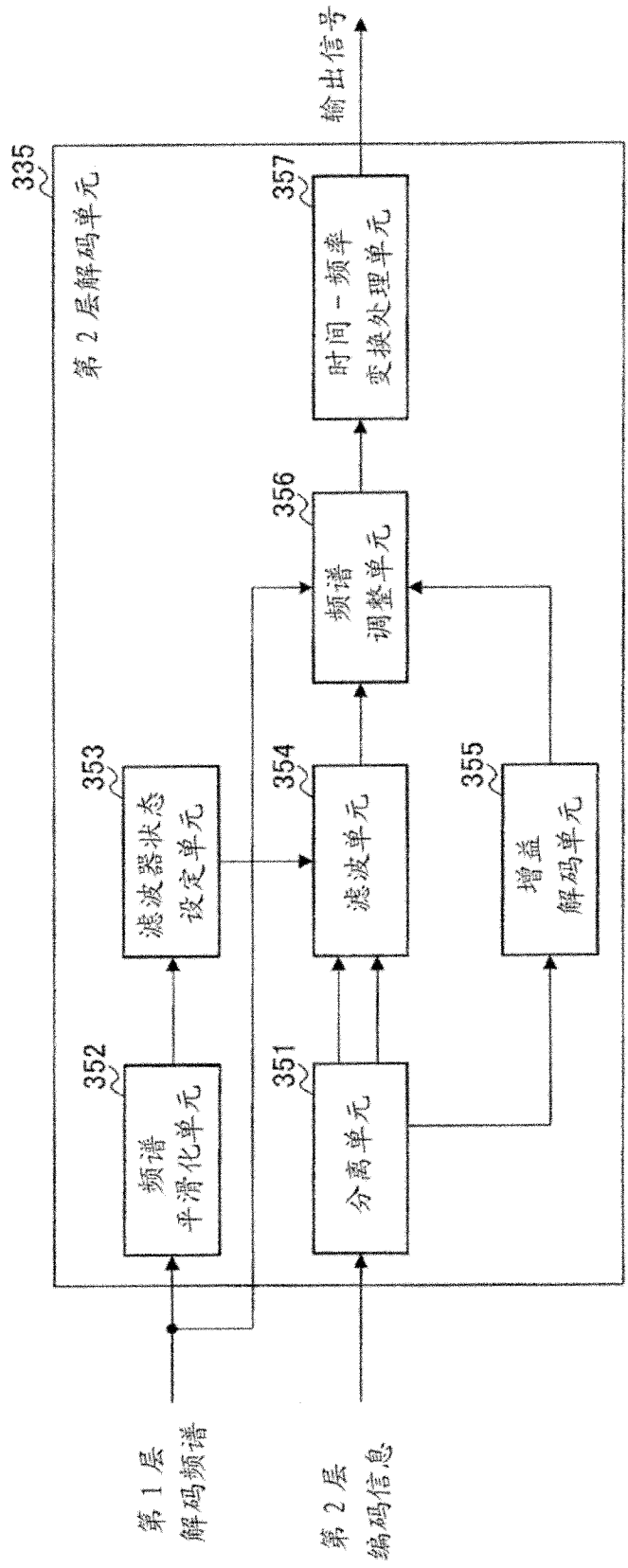


图 12