

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200680031861.X

[43] 公开日 2008 年 8 月 27 日

[51] Int. Cl.

G10L 19/00 (2006.01)

G10L 19/02 (2006.01)

[11] 公开号 CN 101253556A

[22] 申请日 2006.8.31

[21] 申请号 200680031861.X

[30] 优先权

[32] 2005.9.2 [33] JP [31] 254357/2005

[32] 2006.7.11 [33] JP [31] 190127/2006

[86] 国际申请 PCT/JP2006/317218 2006.8.31

[87] 国际公布 WO2007/026821 日 2007.3.8

[85] 进入国家阶段日期 2008.2.29

[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 高木良明 张国成 则松武志

宫阪修二 川村明久 小野耕司郎

石川智一

[74] 专利代理机构 永新专利商标代理有限公司

代理人 徐殿军

权利要求书 6 页 说明书 23 页 附图 9 页

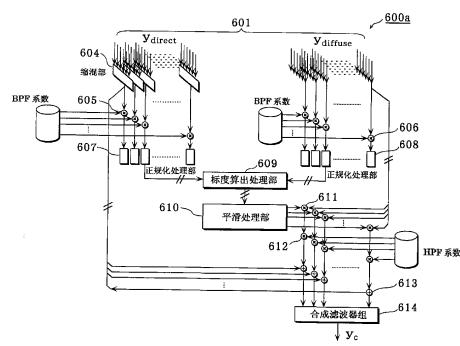
[54] 发明名称

能量整形装置以及能量整形方法

[57] 摘要

时间上的处理装置(能量整形装置)(600a)包括：分离器(601)，将通过时频混合转换而得到的子频带区域的音频信号划分为，表示混响成分的扩散信号和表示非混响成分的直接信号；缩混部(604)，通过缩混所述直接信号来生成缩混信号；BPF(605、606)，通过以每个子频带为单位，对所述缩混信号和按照每个所述子频带划分的扩散信号施行带通滤波处理，从而分别生成带通缩混信号和带通扩散信号；正规化处理部(607、608)，通过对所述带通缩混信号和所述带通扩散信号各自的能量进行正规化，从而分别生成正规化缩混信号和正规化扩散信号；标度算出处理部(609)，按每个预先规定的时间间隙算出标度因子，该标度因子示出对于所述正规化扩散信号的能量的、所述正规化缩混信号的能量的大小；运算部(611)，通过使所述标

度因子与所述扩散信号相乘，从而生成标度扩散信号；HPF(612)，通过对所述标度扩散信号施行高通滤波处理，从而生成高通扩散信号；加法部(613)，通过使所述高通扩散信号和所述直接信号相加，从而生成相加信号；以及合成滤波器组(614)，通过对所述相加信号施行合成滤波处理，从而使所述相加信号转换为时域信号。



1. 一种能量整形装置，在多通道音频信号的解码中进行能量整形，其特征在于，包括：

分离单元，将通过时频混合转换而得到的子频带区域的音频信号划分为，表示混响成分的扩散信号和表示非混响成分的直接信号；

缩混单元，通过缩混所述直接信号来生成缩混信号；

滤波处理单元，通过以每个子频带为单位，对所述缩混信号和按照每个所述子频带划分的扩散信号施行带通滤波处理，从而分别生成带通缩混信号和带通扩散信号；

正规化处理单元，通过对所述带通缩混信号和所述带通扩散信号各自的能量进行正规化，从而分别生成正规化缩混信号和正规化扩散信号；

标度因子算出单元，按每个预先规定的时间间隙算出标度因子，该标度因子示出对于所述正规化扩散信号的能量的、所述正规化缩混信号的能量的大小；

乘法单元，通过使所述标度因子与所述扩散信号相乘，从而生成标度扩散信号；

高通滤波处理单元，通过对所述标度扩散信号施行高通滤波处理，从而生成高通扩散信号；

加法单元，通过使所述高通扩散信号和所述直接信号相加，从而生成相加信号；以及

合成滤波处理单元，通过对所述相加信号施行合成滤波处理，从而使所述相加信号转换为时域信号。

2. 如权利要求1所述的能量整形装置，其特征在于，

所述能量整形装置进一步包括平滑单元，通过对所述标度因子施

行平滑处理，从而生成平滑的标度因子，所述平滑处理是指抑制每个时间间隙的所述标度因子的变动。

3. 如权利要求 2 所述的能量整形装置，其特征在于，
所述平滑单元通过使以下两个值相加来施行所述平滑处理，所述两个值是指：当前的时间间隙中的标度因子与 α 相乘得到的值，和当前的时间间隙的前一个时间间隙中的标度因子与 $(1 - \alpha)$ 相乘得到的值。

4. 如权利要求 1 所述的能量整形装置，其特征在于，
所述能量整形装置进一步包括限幅处理单元，在所述标度因子超过预先规定的上限值的情况下，将所述标度因子限制为上限值，并且，在所述标度因子低于预先规定的下限值的情况下，将所述标度因子限制为下限值，通过进行以上所述限制来对所述标度因子施行限幅处理。

5. 如权利要求 4 所述的能量整形装置，其特征在于，
所述限幅处理单元在设上限值为 β 的情况下，设下限值为 $1 / \beta$ ，来施行所述限幅处理。

6. 如权利要求 1 所述的能量整形装置，其特征在于，
所述直接信号包含：所述音频信号的低频中的混响成分和非混响成分，以及所述音频信号的高频中的非混响成分。

7. 如权利要求 1 所述的能量整形装置，其特征在于，
所述扩散信号包含所述音频信号的高频中的混响成分，不包含所述音频信号的低频成分。

8. 如权利要求 1 所述的能量整形装置，其特征在于，

所述能量整形装置进一步包括控制单元，对所述音频信号是否施行能量整形进行切换。

9. 如权利要求8所述的能量整形装置，其特征在于，
所述控制单元按照表示在每个音频帧是否施行能量整形处理的控制标志进行选择，在不施行能量整形处理的情况下选择所述扩散信号，在施行能量整形处理的情况下选择所述高通扩散信号；
所述加法单元使在所述控制单元选择的信号和所述直接信号相加。

10. 一种能量整形方法，在多通道音频信号的解码中进行能量整形，其特征在于，包括：

分离步骤，将通过时频混合转换而得到的子频带区域的音频信号划分为，表示混响成分的扩散信号和表示非混响成分的直接信号；
缩混步骤，通过缩混所述直接信号来生成缩混信号；
滤波处理步骤，通过以每个子频带为单位，对所述缩混信号和按照每个所述子频带划分的扩散信号施行带通滤波处理，从而分别生成带通缩混信号和带通扩散信号；
正规化处理步骤，通过对所述带通缩混信号和所述带通扩散信号各自的能量进行正规化，从而分别生成正规化缩混信号和正规化扩散信号；

标度因子算出步骤，按每个预先规定的时间间隙算出标度因子，该标度因子示出对于所述正规化扩散信号的能量的、所述正规化缩混信号的能量的大小；

乘法步骤，通过使所述标度因子与所述扩散信号相乘，从而生成标度扩散信号；

高通滤波处理步骤，通过对所述标度扩散信号施行高通滤波处理，从而生成高通扩散信号；

加法步骤，通过使所述高通扩散信号和所述直接信号相加，从而生成相加信号；以及

合成滤波处理步骤，通过对所述相加信号施行合成滤波处理，从而使所述相加信号转换为时域信号。

1 1 . 如权利要求 1 0 所述的能量整形方法，其特征在于，

所述能量整形方法进一步包括平滑步骤，通过对所述标度因子施行平滑处理，从而生成平滑的标度因子，所述平滑处理是指抑制每个时间间隙的所述标度因子的变动。

1 2 . 如权利要求 1 1 所述的能量整形方法，其特征在于，

在所述平滑步骤通过使以下两个值相加来施行所述平滑处理，所述两个值是指：当前的时间间隙中的标度因子与 α 相乘得到的值，和当前的时间间隙的前一个时间间隙中的标度因子与 $(1 - \alpha)$ 相乘得到的值。

1 3 . 如权利要求 1 0 所述的能量整形方法，其特征在于，

所述能量整形方法进一步包括限幅处理步骤，在所述标度因子超过预先规定的上限值的情况下，将所述标度因子限制为上限值，并且，在所述标度因子低于预先规定的下限值的情况下，将所述标度因子限制为下限值，通过进行以上所述限制来对所述标度因子施行限幅处理。

1 4 . 如权利要求 1 3 所述的能量整形方法，其特征在于，

在所述限幅处理步骤，在设上限值为 β 的情况下，设下限值为 $1 / \beta$ ，来施行所述限幅处理。

1 5 . 如权利要求 1 0 所述的能量整形方法，其特征在于，

所述直接信号包含：所述音频信号的低频中的混响成分和非混响

成分，以及所述音频信号的高频中的非混响成分。

1 6 . 如权利要求 1 0 所述的能量整形方法，其特征在于，
所述扩散信号包含所述音频信号的高频中的混响成分，不包含所
述音频信号的低频成分。

1 7 . 如权利要求 1 0 所述的能量整形方法，其特征在于，
所述能量整形方法进一步包括控制步骤，对所述音频信号是否施
行能量整形进行切换。

1 8 . 如权利要求 1 7 所述的能量整形方法，其特征在于，
在所述控制步骤按照表示在每个音频帧是否施行能量整形处理的
控制标志进行选择，在不施行能量整形处理的情况下选择所述扩散信
号，在施行能量整形处理的情况下选择所述高通扩散信号；
在所述加法步骤使在所述控制步骤选择的信号和所述直接信号相
加。

1 9 . 一种程序，其特征在于，该程序是用于在多通道音频信号
的解码中进行能量整形的程序，并使计算机执行权利要求 1 0 所述的
能量整形方法中所包含的步骤。

2 0 . 一种集成电路，在多通道音频信号的解码中进行能量整形，
其特征在于，该集成电路集成了能量整形装置，该能量整形装置包括：
分离器，将通过时频混合转换而得到的子频带区域的音频信号划
分为，表示混响成分的扩散信号和表示非混响成分的直接信号；
缩混电路，通过缩混所述直接信号来生成缩混信号；
滤波器，通过以每个子频带为单位，对所述缩混信号和按照每个
所述子频带划分的扩散信号施行带通滤波处理，从而分别生成带通缩

混信号和带通扩散信号；

正规化处理电路，通过对所述带通缩混信号和所述带通扩散信号各自的能量进行正规化，从而分别生成正规化缩混信号和正规化扩散信号；

标度因子算出电路，按每个预先规定的时间间隙算出标度因子，该标度因子示出对于所述正规化扩散信号的能量的、所述正规化缩混信号的能量的大小；

乘法器，通过使所述标度因子与所述扩散信号相乘，从而生成标度扩散信号；

高通滤波处理电路，通过对所述标度扩散信号施行高通滤波处理，从而生成高通扩散信号；

加法器，通过使所述高通扩散信号和所述直接信号相加，从而生成相加信号；以及

合成滤波器，通过对所述相加信号施行合成滤波处理，从而使所述相加信号转换为时域信号。

能量整形装置以及能量整形方法

技术领域

本发明涉及能量整形装置以及能量整形方法，尤其涉及在多通道音频信号的解码中进行能量整形的技术。

背景技术

近年，在 MPEG 音频标准中，被称为 Spatial Audio Codec（空间音频编解码）的技术正在被标准化。其目的在于要以非常少的信息量来对表现出临场感的多通道信号进行压缩及编码。例如，在作为数字电视的声音方式已被广泛使用的多声道编解码方式的 AAC (Advanced Audio Coding : 高级音频编码) 方式，5.1 声道要有 512kbps 或 384kbps 的比特率，然而，在 Spatial Audio Codec 则以用 128kbps 或 64kbps 甚至于 48kbps 这样非常少的比特率来对多通道音频信号进行压缩及编码为目标（例如参照非专利文献 1）。

图 1 是利用空间编解码的基本原理的音频装置的整体结构方框图。

音频装置 1 包括：音频编码器 10 和音频解码器 20，音频编码器 10 输出对音频信号的组进行空间音频编码后而得到的编码信号，音频解码器 20 对从音频编码器 10 输出的编码信号进行解码。

音频编码器 10 以由 1024 个采样或 2048 个采样等所示出的帧为单位，对多个声道的音频信号（例如双声道的音频信号 L、R）进行处理，且该音频编码器 10 包括：缩混部 11、双声列（Binaural Cue）检测部 12、编码器 13、以及多路复用部 14。

缩混部 11 通过对以谱表示的例如左右双声道的音频信号 L、R 取

平均，即通过 $M = (L+R)/2$ ，而生成缩混音频信号 L、R 后而得到的缩混信号 M。

双声列检测部 12 通过按照各个谱带对音频信号 L、R 以及缩混信号 M 进行比较，从而生成用于将缩混信号 M 复原到音频信号 L、R 的 BC 信息（双声列）。

BC 信息中包含：示出声道间强度 / 强度差（inter-channel level/intensity difference）的强度信息 IID、示出声道间相干 / 相关（inter-channel coherence/correlation）的相关信息 ICC、以及示出声道间相位延迟差（inter-channel phase/delay difference）的相位信息 IPD。

在此，相关信息 ICC 示出两个音频信号 L、R 的类似性，强度信息 IID 示出音频信号 L、R 的相对强度。一般而言，强度信息 IID 是用于控制声音的平衡和定位的信息，相关信息 ICC 是用于控制声音的幅度和扩散性的信息。这些信息均为帮助听者在头脑中构成听觉情景的空间参数。

在最新的空间编码中，以谱表示的音频信号 L、R 以及缩混信号 M 被划分为由“参数频带（parameter band）”构成的通常的多个组。因此，BC 信息是按照各个参数频带被算出的。并且，“BC 信息（双声列）”和“空间参数”会经常被作为具有互换性的同义词语来使用。

编码器 13 通过例如 MP3 (MPEG Audio Layer-3) 或 AAC(Advanced Audio Coding : 先进音频编码) 等对缩混信号 M 进行压缩编码。即，编码器 13 编码缩混信号 M，并生成被压缩的编码串。

多路复用部 14 在量化 BC 信息的同时，还通过对缩混信号 M 和被量化了的 BC 信息进行多路复用而生成比特流，并将该比特流作为所述的编码信号来输出。

音频解码器 20 包括：逆多路复用部 21、解码器 22、以及多声道合成部 23。

逆多路复用部 21 获得所述的比特流，并从该比特流中将被量化的

BC 信息和被编码的缩混信号 M 分离出来后输出。并且，逆多路复用部 21 对被量化的 BC 信息进行逆量化后输出。

解码器 22 将被编码的缩混信号 M 解码后输出到多声道合成部 23。

多声道合成部 23 获得从解码器 22 输出的缩混信号 M 和从逆多路复用部 21 输出的 BC 信息。并且，多声道合成部 23 利用所述 BC 信息，将缩混信号 M 复原为两个音频信号 L、R。这些将缩混信号复原为原来的两个音频信号的处理是与后述的“声道分离技术”结合在一起进行的。

并且，在上述的例子中只是说明了(1)是怎样在编码器中以一个缩混信号和空间参数的组来表示两个音频信号的，以及(2)是怎样通过处理空间参数和缩混信号，在解码器中将缩混信号分离为两个音频信号的。在所述技术中，可以在进行编码处理时，将多于两个声道的声音（例如 5.1 声道声源的六个声道）压缩为一个或两个缩混声道，并可以在进行解码处理时复原。

即，在以上所述中，以对两个声道的音频信号进行编码及解码为例对音频装置 1 进行了说明，不过，音频装置 1 也可以对两个声道以上的声道的音频信号（例如构成 5.1 声道声源的六个声道的音频信号）进行编码及解码。

图 2 是六个声道中多声道合成部 23 的功能结构框图。

多声道合成部 23 例如在将缩混信号 M 分离为六个声道的音频信号的情况下，包括：第一声道分离部 241、第二声道分离部 242、第三声道分离部 243、第四声道分离部 244、以及第五声道分离部 245。并且，缩混信号 M 是对以下的音频信号进行缩混后而得到的，这些音频信号是指：与设置在视听者正面的扬声器相对应的中置音频信号 C、与设置在视听者左前方的扬声器相对应的前左音频信号 Lf、与设置在视听者右前方的扬声器相对应的前右音频信号 Rf、与设置在视听者左侧的扬声器相对应的左环绕音频信号 Ls、与设置在视听者右侧的扬声器相对应的右环绕音频信号 Rs、以及与用于输出低音的重低音扬声器相对

应的低音音频信号 LFE。

第一声道分离部 241 从缩混信号 M 中将中途的第一缩混信号 M1 和中途的第四缩混信号 M4 分离出来后输出。第一缩混信号 M1 由中置音频信号 C、前左音频信号 Lf、前右音频信号 Rf、以及低音音频信号 LFE 缩混而成。第四缩混信号 M4 由左环绕音频信号 Ls 和右环绕音频信号 Rs 缩混而成。

第二声道分离部 242 从第一缩混信号 M1 中将中途的第二缩混信号 M2 和中途的第三缩混信号 M3 分离出来后输出。第二缩混信号 M2 由前左音频信号 Lf 和前右音频信号 Rf 缩混而成。第三缩混信号 M3 由中置音频信号 C 和低音音频信号 LFE 缩混而成。

第三声道分离部 243 从第二缩混信号 M2 中将前左音频信号 Lf 和前右音频信号 Rf 分离出来后输出。

第四声道分离部 244 从第三缩混信号 M3 中将中置音频信号 C 和低音音频信号 LFE 分离出来后输出。

第五声道分离部 245 从第四缩混信号 M4 中将左环绕音频信号 Ls 和右环绕音频信号 Rs 分离出来后输出。

这样，多声道合成部 23 通过多阶段的方法在各个声道分离部以相同的分离处理，将一个缩混信号分离为两个缩混信号，直至分离到单声道的音频信号为止重复进行递归的信号分离。

图 3 是用于说明多声道合成部 23 的原理的功能框图。

多声道合成部 23 包括：全通滤波器 261、BCC 处理部 262、以及运算部 263。

全通滤波器 261 获得缩混信号 M，并对该缩混信号 M 生成没有相关性的无相关信号 Mrev 并输出。在听觉上对缩混信号 M 和无相关信号 Mrev 进行比较可知它们互不相干。并且，无相关信号 Mrev 具有与缩混信号 M 相等的能量，含有能够制作出好像声音被传播得很远这种幻觉的有限时间的混响成分。

BCC 处理部 262 获得 BC 信息，并根据该 BC 信息中所包含的强度

信息 IID 或相关信息 ICC 等，生成混合系数 H_{ij} 并输出，该混合系数 H_{ij} 用于维持 L、R 之间的相关程度或 L、R 的方向性。

运算部 263 获得并利用缩混信号 M、无相关信号 M_{rev} 、以及混合系数 H_{ij} ，进行（公式 1）所示的运算，并输出音频信号 L、R。这样，通过利用混合系数 H_{ij} ，从而使音频信号 L、R 间的相关程度或这些信号的方向性成为希望的状态。

（公式 1）

$$L = H_{11} * M + H_{12} * M_{rev}$$

$$R = H_{21} * M + H_{22} * M_{rev}$$

图 4 是多声道合成部 23 的详细构成的方框图。并且，解码器 22 也在此图中示出。

解码器 22 将编码缩混信号解码为时域的缩混信号 M，并将解码后的缩混信号 M 输出到多声道合成部 23。

多声道合成部 23 包括：解析滤波器组 231、声道扩大部 232、以及时空上的处理装置（能量整形装置）900。声道扩大部 232 包括：前矩阵处理部 2321、后矩阵处理部 2322、第一运算部 2323、无相关处理部 2324、以及第二运算部 2325。

解析滤波器组 231 获得从解码器 22 输出的缩混信号 M，并将该缩混信号 M 的表示形式转换为以时间和频率表示的混合表示形式，并作为以简略方式的矢量 x 表示的第一频带信号 x 来输出。并且，此解析滤波器组 231 包括第一阶段和第二阶段。例如，第一阶段是 QMF 滤波器组，第二阶段是奈奎斯特滤波器组。在这些阶段中，首先以 QMF 滤波器（第一阶段）划分为多个频带，进而以奈奎斯特滤波器（第二阶段）将低频侧的子频带分为更窄的子频带，从而可以提高位于低频的子频带的频谱分辨率。

声道扩大部 232 的前矩阵处理部 2321 利用 BC 信息生成作为标度因子的矩阵 R1，所述标度因子示出向各声道的信号强度的分配（定标）。

例如，前矩阵处理部 2321 利用强度信息 IID 来生成矩阵 R1，所述强度信息 IID 示出以下的信号强度的比率，即缩混信号 M 的信号强度分别和第一缩混信号 M1、第二缩混信号 M2、第三缩混信号 M3 以及第四缩混信号 M4 的信号强度的比率。

即，前矩阵处理部 2321 的目的在于生成中间信号，该中间信号可以用在图 2 所示的第一到第五声道分离部 241 到 245 生成无相关信号中，且前矩阵处理部 2321 根据定标输入缩混信号 M 的能量强度的 ILD 空间参数来算出标度因子的矢量 R1，该标度因子的矢量 R1 包括合成信号 M1~M4 的 ILD 空间参数的矢量元素（vector elements）R1[0] 到 R1[4]。

第一运算部 2323 获得从解析滤波器组 231 输出的时频混合表示的第一频带信号 x，例如（公式 2）和（公式 3）所示，算出所述第一频带信号 x 和矩阵 R1 的乘积。并且，第一运算部 2323 输出示出矩阵运算结果的中间信号 v。即，第一运算部 2323 从由解析滤波器组 231 输出的时频混合表示的第一频带信号 x 中分离四个缩混信号 M1 到 M4。

（公式 2）

$$v = \begin{bmatrix} M \\ M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix} = R_1 x$$

在此, M1 到 M4, 以下列(公式3)来表示。

(公式3)

$$M_1 = L_f + R_f + C + LFE$$

$$M_2 = L_f + R_f$$

$$M_3 = C + LFE$$

$$M_4 = L_s + R_s$$

无相关处理部 2324 具有图 3 所示的全通滤波器 261 所具有的功能, 通过对中间信号 v 施行全通滤波处理, 从而如(公式4所示), 生成并输出无相关信号 w。并且, 无相关信号 w 的构成要素 Mrev 以及 Mi, rev 是对缩混信号 M 以及 Mi 施行无相关处理的信号。

(公式4)

$$w = \begin{bmatrix} M \\ decorr(v) \end{bmatrix} = \begin{bmatrix} M \\ M_{rev} \\ M_{1,rev} \\ M_{2,rev} \\ M_{3,rev} \\ M_{4,rev} \end{bmatrix} = \begin{bmatrix} M \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ M_{rev} \\ M_{1,rev} \\ M_{2,rev} \\ M_{3,rev} \\ M_{4,rev} \end{bmatrix} = w_{Dry} + w_{Wet}$$

并且, 上述(公式4)的 wDry 由原来的缩混信号构成(以后称为“Dry”信号), wWet 由无相关信号的集合构成(以后称为“Wet”信号)。

后矩阵处理部 2322 利用 BC 信息生成矩阵 R2, 该矩阵 R2 示出混

响在各个声道中的分配。即，后矩阵处理部 2322 为了导出各个信号，算出混合 M 和 Mi, rev 的混合系数的矩阵 R2。例如，后矩阵处理部 2322 通过示出声音的幅度或扩散性的相关信息 ICC 导出混合系数 Hi j，并生成由该混合系数 Hi j 构成的矩阵 R2。

第二运算部 2325 算出无相关信号 w 和矩阵 R2 的乘积，并输出示出矩阵运算结果的输出信号 y。即，第二运算部 2325 从无相关信号 w 分离六个音频信号，即 Lf、Rf、Ls、Rs、C、以及 LFE。

例如，如图 2 所示，要想从第二缩混信号 M2 分离前左音频信号 Lf，就要在该前左音频信号 Lf 的分离中利用第二缩混信号 M2 和与其相对应的无相关信号 w 的构成要素 M2, rev。同样，要想从第一缩混信号 M1 分离第二缩混信号 M2，就要在该第二缩混信号 M2 的算出中利用第一缩混信号 M1 和与其相对应的无相关信号 w 的构成要素 M1, rev。

因此，前左音频信号 Lf 以以下的（公式 5）来表示。

（公式 5）

$$L_f = H_{11,A} * M_2 + H_{12,A} * M_{2,rev}$$

$$M_2 = H_{11,D} * M_1 + H_{12,D} * M_{1,rev}$$

$$M_1 = H_{11,E} * M + H_{12,E} * M_{rev}$$

在此，（公式 5）中的 Hi j, A 是第三声道分离部 243 中的混合系数，Hi j, D 是第二声道分离部 242 中的混合系数，Hi j, E 是第一声道分离部 241 中的混合系数。（公式 5）中所示出的三个公式可以归纳为以下（公式 6）所示出的一个矢量乘法算式。

（公式 6）

$$\begin{aligned} L_f &= [H_{11,A} H_{11,D} H_{11,E} \quad H_{11,A} H_{11,D} H_{12,E} \quad H_{11,A} H_{12,D} \quad H_{12,A} \quad 0 \quad 0] w \\ &= R_{2,Lf} w \end{aligned}$$

除前左音频信号 Lf 以外，其它的音频信号 Rf、C、LFE、Ls、以及 Rs 也可以通过上述的矩阵和无相关信号 w 的矩阵的运算来算出。

即，输出信号 y 可以用以下的（公式 7）来表示。

（公式 7）

$$y = \begin{bmatrix} Lf \\ Rf \\ Ls \\ Rs \\ C \\ LFE \end{bmatrix} = \begin{bmatrix} R_{2,Lf} \\ R_{2,Rf} \\ R_{2,Ls} \\ R_{2,Rs} \\ R_{2,C} \\ R_{2,LFE} \end{bmatrix} w = R_2 w = R_2 w_{Dry} + R_2 w_{Wet} = y_{Dry} + y_{Wet}$$

由第一到第五声道分离部 241 到 245 的混合系数的倍数集合构成的矩阵 R2 为了生成多通道信号，而线性组合 M、Mrev、M2, rev、… M4, rev。为了以后的能量整形处理，yDry 和 yWet 被分别存储。

时间上的处理装置 900 将被复原的各个音频信号的表示形式从时频混合表示转换为时域表示形式，并将以时域表示的多个音频信号作为多通道信号来输出。并且，时间上的处理装置 900 为了与解析滤波器组 231 相匹配，例如可以由两个阶段构成。并且，矩阵 R1、R2 是按各个上述的参数频带 b 作为矩阵 R1(b)、R2(b) 而被生成的。

在此，Wet 信号和 Dry 信号在被合并之前，Wet 信号是按照 Dry 信号的时间上的包络被整形的。时间上的处理装置 900 这一模块对于像攻击音等这种具有快速时间变化特性的信号而言是不可缺少的。

即，时间上的处理装置 900 在像攻击音或声音信号这样的时间变化急剧的信号的情况下，为了改善声音的失真，为了适合于直接信号的时间包络，而通过使整形扩散信号的时间包络的信号和直接信号相

加并输出，来保持原有声音的音质。

图 5 是图 4 所示的时间上的处理装置 900 的详细构成方框图。

如图 5 所示，时间上的处理装置 900 包括：分离器 901、合成滤波器组 902 和 903、缩混部 904、带通滤波器(BPF) 905 和 906、正规化处理部 907 和 908、标度算出处理部 909、平滑处理部 910、运算部 911、高通滤波器(HPF) 912、以及加法部 913。

分离器 901，将被复原了的信号 y 像以下的（公式 8）和（公式 9）所示那样，划分为直接信号 y_{direct} 和扩散信号 $y_{diffuse}$ 。

（公式 8）

$$y_{direct} = \begin{bmatrix} y_{1,direct} \\ y_{2,direct} \\ y_{3,direct} \\ y_{4,direct} \\ y_{5,direct} \\ y_{6,direct} \end{bmatrix} = \begin{cases} y_{Dry} + y_{Wet} & \text{低频用} \\ y_{Dry} & \text{高频用} \end{cases}$$

（公式 9）

$$y_{diffuse} = \begin{bmatrix} y_{1,diffuse} \\ y_{2,diffuse} \\ y_{3,diffuse} \\ y_{4,diffuse} \\ y_{5,diffuse} \\ y_{6,diffuse} \end{bmatrix} = \begin{cases} 0 & \text{低频用} \\ y_{Wet} & \text{高频用} \end{cases}$$

合成滤波器组 902 将六个直接信号向时域转换。合成滤波器组 903 与合成滤波器组 902 一样，将六个扩散信号向时域转换。

缩混部 904 根据以下(公式 1 0)，使时域中的六个直接信号相加，从而得到一个直接缩混信号 M_{direct} 。

(公式 1 0)

$$M_{direct} = \sum_{i=1}^6 y_{i,direct}$$

BPF905 对一个直接缩混信号施行带通处理。BPF906 与 BPF905 一样，对六个扩散信号均施行带通处理。被施行了带通处理的直接缩混信号以及扩散信号以以下的（公式 1 1 ）来表示。

(公式 1 1)

$$M_{direct,BP} = Bandpass(M_{direct})$$

$$y_{i,diffuse,BP} = Bandpass(y_{i,diffuse})$$

正规化处理部 907 根据以下所示的(公式 1 2)进行正规化处理，以使直接缩混信号在一个处理帧中具有一个能量。

(公式 1 2)

$$M_{direct,norm}(t) = \frac{M_{direct,BP}(t)}{\sqrt{\sum_t M_{direct,BP}(t) \cdot M_{direct,BP}(t)}}$$

正规化处理部 908 与正规化处理部 907 一样，根据以下所示的(公式 1 3)，对六个扩散信号进行正规化。

(公式 1 3)

$$y_{1,diffuse,norm}(t) = \frac{y_{i,diffuse,BP}(t)}{\sqrt{\sum_t y_{i,diffuse,BP}(t) * y_{i,diffuse,BP}(t)}}$$

被正规化的信号在标度算出处理部 909 被划分为时间块。并且，标度算出处理部 909 根据以下（公式 1 4）分别对各个时间块算出标度因子。

（公式 1 4）

$$scale_i(b) = \sqrt{\frac{\sum_{t \in b} M_{direct,norm}(t) \cdot M_{direct,norm}(t)}{\sum_{t \in b} y_{i,diffuse,norm}(t) \cdot y_{1,diffuse,norm}(t)}}$$

另外，图 6 示出了所述（公式 1 4）的时间块 b 示出“块索引”时的所述划分处理。

最后，所述扩散信号在运算部 911 被定标，并且在于加法部 913 被组合到所述直接信号之前，在 HPF912 根据以下（公式 1 5）被施行高通滤波处理。

（公式 1 5）

$$y_{i,diffuse,scaled,HP} = Highpass(y_{i,diffuse} \cdot scale_i)$$

$$y_i = y_{i,direct} + y_{i,diffuse,scaled,HP}$$

并且，平滑处理部 910 是一种提高连续的时间块中标的平滑性的附加技术。例如，连续的时间块可以是如图 6 中所示，具有重叠的 a 部分，在此重叠的部分，“被加权”的标度因子利用窗功能被运

算。

在定标处理部 911 也可以利用本领域技术人员所熟知的重叠加法技术。

这样，在以往的时间上的处理装置 900，对于原来的各个信号，通过整形时域的各个无相关信号，来提示上述的能量整形方法。

非专利文献 1 J. Herre, et al, "The Reference Model Architecture for MPEG Spatial Audio Coding", 118th AES Convention, Barcelona

然而，在以往的能量整形装置中，一半是直接信号，另一半是扩散信号，由于需要对于这十二个信号进行合成滤波处理，因此运算负荷非常重。并且，由于频带的不同以及高通滤波器的使用，导致了滤波处理的延迟。

即，在以往的能量整形装置中，通过合成滤波器组 902 和 903，分别将分离器 901 所划分的直接信号和扩散信号转换为时域信号。因此，出现的问题是：例如输入音频信号为六个声道的情况下，在每个时间帧中需要进行 $6 \times 2 = 12$ 个合成滤波处理，这样，处理量就会很大。

并且，还会出现的问题是：由于对由合成滤波器组 902 和 903 转换的时域的直接信号以及扩散信号施行带通处理或高通滤波处理，因此，导致在这些通过处理中发生延迟。

发明内容

因此，本发明为了解决上述问题，目的在于提供一种能量整形装置以及能量整形方法，可以降低合成滤波处理的处理量，并可以防止在通过处理中所发生的延迟。

为了达成上述的目的，本发明所涉及的能量整形装置在多通道音频信号的解码中进行能量整形，其中，包括：分离单元，将通过时频混合转换而得到的子频带区域的音频信号划分为，表示混响成分的扩散信号和表示非混响成分的直接信号；缩混单元，通过缩混所述直接

信号来生成缩混信号；滤波处理单元，通过以每个子频带为单位，对所述缩混信号和按照每个所述子频带划分的扩散信号施行带通滤波处理，从而分别生成带通缩混信号和带通扩散信号；正规化处理单元，通过对所述带通缩混信号和所述带通扩散信号各自的能量进行正规化，从而分别生成正规化缩混信号和正规化扩散信号；标度因子算出单元，按每个预先规定的时间间隙算出标度因子，该标度因子示出对于所述正规化扩散信号的能量的、所述正规化缩混信号的能量的大小；乘法单元，通过使所述标度因子与所述扩散信号相乘，从而生成标度扩散信号；高通滤波处理单元，通过对所述标度扩散信号施行高通滤波处理，从而生成高通扩散信号；加法单元，通过使所述高通扩散信号和所述直接信号相加，从而生成相加信号；以及合成滤波处理单元，通过对所述相加信号施行合成滤波处理，从而使所述相加信号转换为时域信号。

这样，在进行合成滤波处理之前，可以对各声道的直接信号以及扩散信号，按每个子频带进行带通滤波处理。因此，可以以单纯的乘法运算来实现带通滤波处理，并可以防止带通滤波处理所产生的延迟。并且，对于各声道的直接信号以及扩散信号，在处理结束后对相加信号施行合成滤波处理，由此可以进行转换为时域信号的合成滤波处理。因此，例如在六个声道的情况下，可以将合成滤波处理的个数减少为六个，从而使合成滤波处理的处理量比以往减少一半。

并且，可以是，在本发明所涉及的能量整形装置中，所述能量整形装置进一步包括平滑单元，通过对所述标度因子施行平滑处理，从而生成平滑的标度因子，所述平滑处理是指抑制每个时间间隙的所述标度因子的变动。

据此，可以防止在频域求出的标度因子的值急剧变化或上溢，或因音质降低而引起的问题。

并且，可以是，在本发明所涉及的能量整形装置中，所述平滑单元通过使以下两个值相加来施行所述平滑处理，所述两个值是指：当

前的时间间隙中的标度因子与 α 相乘得到的值，和当前的时间间隙的前一个时间间隙中的标度因子与 $(1 - \alpha)$ 相乘得到的值。

据此，可以以简单的处理来防止在频域求出的标度因子的值的急剧变化或上溢。

并且，可以是，在本发明所涉及的能量整形装置中，所述能量整形装置进一步包括限幅处理单元，在所述标度因子超过预先规定的上限值的情况下，将所述标度因子限制为上限值，并且，在所述标度因子低于预先规定的下限值的情况下，将所述标度因子限制为下限值，通过进行以上所述限制来对所述标度因子施行限幅处理。

据此，也可以防止在频域求出的标度因子的值急剧变化或上溢，或因音质降低而引起的问题。

并且，可以是，在本发明所涉及的能量整形装置中，所述限幅处理单元在设上限值为 β 的情况下，设下限值为 $1 / \beta$ ，来施行所述限幅处理。

据此，也可以以简单的处理来防止在频域求出的标度因子的值的急剧变化或上溢。

并且，可以是，在本发明所涉及的能量整形装置中，所述直接信号包含：所述音频信号的低频中的混响成分和非混响成分，以及所述音频信号的高频中的非混响成分。

并且，可以是，在本发明所涉及的能量整形装置中，所述扩散信号包含所述音频信号的高频中的混响成分，不包含所述音频信号的低频成分。

并且，可以是，在本发明所涉及的能量整形装置中，所述能量整形装置进一步包括控制单元，对所述音频信号是否施行能量整形进行切换。据此，通过对是否施行能量整形进行切换，从而可以实现声音的时间上变动的尖锐性以及确实地声像定位。

并且，可以是，在本发明所涉及的能量整形装置中，所述控制单元按照控制是否施行能量整形处理的控制标志，对所述扩散信号以及

所述高通扩散信号的某一方进行选择，所述加法单元使在所述控制单元选择的信号和所述直接信号相加。

据此，可以随时简单地对是否施行能量整形进行切换。

并且，本发明不仅可以作为这样能量整形装置来实现，而且可以作为将具有这样的能量整形装置的特征性单元作为步骤的能量整形方法来实现，也可以作为使计算机执行这些步骤的程序来实现，还可以对能量整形装置所具备的特征性单元进行集成电路化。并且，这样的程序可以通过 CD-ROM 等记录介质或互联网等传输介质来分发。

通过以上的说明，可以明确知道，根据本发明所涉及的能量整形装置，可以在不使比特流的句法变形，且维持高音质的基础上，降低合成滤波处理的处理量，并可以防止在通过处理中所发生的延迟。

因此，本发明在向移动电话或移动信息终端分发音乐内容或在视听已经普及的今天，具有非常高的实用价值。

附图说明

图 1 是利用空间编解码的基本原理的音频装置的全体构成的方框图。

图 2 是在六个声道时的多声道合成部 23 的功能构成方框图。

图 3 是用于说明多声道合成部 23 的原理的其它的功能构成方框图。

图 4 是多声道合成部 23 的详细构成的方框图。

图 5 是图 4 所示的时间上的处理装置 900 的详细构成的方框图。

图 6 是在以往的整形方法中根据加窗重叠处理的平滑技术示意图。

图 7 是本实施例 1 中的时间上的处理装置（能量整形装置）的构成示意图。

图 8 是考虑在子频带区域中频带滤波处理以及运算节约的示意图。

图 9 是本实施例 1 中的时间上的处理装置（能量整形装置）的构成示意图。

符号说明

600 a , 600b 时间上的处理装置

601 分离器

604 缩混部

605, 606 BPF

607, 608 正规化处理部

609 标度算出处理部

610 平滑处理部

611 运算部

612 HPF

613 加法部

614 合成滤波器组

615 控制部

具体实施方式

以下，利用附图对本发明的实施例进行说明。并且，以下所示的实施例只不过用于说明各种创造性的原理。在此所述的详细变形可以理解为是本领域技术人员可以知道的。因此，本发明仅限定于权利要求的范围，并非受以下具体的、详细说明所限。

（实施例 1）

图 7 是本发明的实施例 1 中时间上的处理装置（能量整形装置）的构成图。

该时间上的处理装置 600a 是代替图 5 的时间上的处理装置 900，构成多声道合成部 23 的装置，如图 7 所示，该时间上的处理装置 600a 包括：分离器 601、缩混部 604、BPF605、BPF606、正规化处理部 607、正规化处理部 608、标度算出处理部 609、平滑处理部 610、运算部 611、

HPF612、加法部 613、以及合成滤波器组 614。

在该时间上的处理装置 600a 中，将来自声道扩大部 232 的、以时频混合表示的子频带区域的输出信号作为直接输入，最后通过以合成滤波处理返回到时间信号，从而可以节省以往所需要的合成滤波处理负荷的 50%，并且，各部的处理也可以变得简单。

分离部 601 的工作因与图 5 的分离器 901 相同，故省略说明。即，分离部 601 将通过时频混合转换而得到的子频带区域的音频信号划分为，表示混响成分的扩散信号和表示非混响成分的直接信号。

在此，直接信号中含有：音频信号的低频频带中的混响成分和非混响成分，以及所述音频信号的高频频带中的非混响成分。并且，扩散信号中含有音频信号的高频频带中的混响成分，不含有音频信号的低频成分。据此，可以对攻击音等时间变化剧烈的声音施行恰当地音质失真防止处理。

在非专利文献 1 中所述的缩混部 904 和本发明中的缩混部 604 的区别在于：处理的信号是时域信号还是子频带信号。然而，不论是哪一个都采用的是一般的多声道缩混处理方法。即，缩混部 604 通过缩混直接信号而生成缩混信号。

BPF605 以及 BPF606 对于缩混信号以及按照每个所述子频带划分的扩散信号，通过在各个子频带中施行带通处理，从而生成各个带通缩混信号以及带通扩散信号。

如图 8 所示，在 BPF605 和 BPF606 中的频带滤波处理是通过在各个子频带对频带滤波器所对应的频率响应进行单纯的相乘而被简单的。从广义上来讲，频带滤波器可以被视为乘法器。在此，800 表示频带滤波器的频率响应。而且，在此乘法运算只要对具有重要的频带响应的区域 801 进行即可，因此，可以消减运算量。例如，在外部阻带区域 802 和 803，若假定乘法运算结果为 0，通带的振幅为 1，则乘法运算可以视为是单纯的复制处理。

即，在 BPF605 和 BPF606 的频带滤波处理可以根据以下（公式 1

6) 进行。

(公式 1 6)

$$M_{direct,BP}(ts, sb) = M_{direct}(ts, sb) \cdot Bandpass(sb)$$

$$y_{i,diffuse,BP}(ts, sb) = y_{i,diffuse}(ts, sb) \cdot Bandpass(sb)$$

在此, ts 是时间间隙索引, sb 是子频带索引。Bandpass(sp)可以是如以上所说明的单纯的乘法运算器。

正规化处理部 607 和 608 对带通缩混信号以及带通扩散信号各自的能量进行正规化, 并分别生成正规化缩混信号以及正规化扩散信号。

正规化处理部 607 以及正规化处理部 608 与非专利文献 1 中公开的正规化处理部 907 以及正规化处理部 908 的区别在于: (1) 处理的信号的区域不同, 正规化处理部 607 以及正规化处理部 608 处理的是子频带区域的信号, 而正规化处理部 907 以及正规化处理部 908 处理的是时域的信号; (2) 除利用以下所示的复共轭以外, 一般的正规化处理方法不同, 即按照以下 (公式 1 7) 的处理方法不同。

在这种情况下, 需要在每个子频带进行正规化处理, 根据正规化处理部 607 以及正规化处理部 608 的优点, 在具有“0”数据的空间区域中运算可以被省略。因此, 与必需要对成为正规化对象的所有样本进行处理的、以往的文献中所公开的正规化模块相比, 从整体上来看几乎没有增加运算负荷。

(公式 1 7)

$$M_{direct,norm}(ts, sb) = \frac{M_{direct,BP}(ts, sb)}{\sqrt{\sum_{ts \in T} \sum_{sb \in BP} M_{direct,BP}(ts, sb) \cdot M_{direct,BP}^*(ts, sb)}}$$

$$y_{i,diffuse,norm}(ts, sb) = \frac{y_{i,diffuse,BP}(ts, sb)}{\sqrt{\sum_{ts \in T} \sum_{sb \in BP} y_{i,diffuse,BP}(ts, sb) \cdot y_{i,diffuse,BP}^*(ts, sb)}}$$

标度算出处理部 609 按照预先规定的各个时间间隙算出标度因子，该标度因子是示出对于正规化扩散信号的能量的正规化缩混信号的能量的大小的标度因子。具体而言，如以下所示，与其说是以时间块为单位来执行，不如说是以时间间隙为单位来执行，除此之外，标度算出处理部 609 的运算也如以下（公式 1 8）所示，原则上与标度算出处理部 909 相同。

（公式 1 8）

$$scale_i(ts) = \sqrt{\frac{\sum_{sb \in BP} M_{direct,norm}(ts, sb) \cdot M_{direct,norm}^*(ts, sb)}{\sum_{sb \in BP} y_{i,diffuse,norm}(ts, sb) \cdot y_{i,diffuse,norm}^*(ts, sb)}}$$

在成为处理对象的时域数据相当少的情况下，基于平滑处理部 910 的加窗重叠处理的平滑技术也必需被平滑处理部 610 所运行。

然而，本实施例中所涉及的平滑处理部 610 的情况是，若为了以非常小的单位进行平滑处理，而采用的标度因子与以往文献中所述的标度因子相同的话（公式 1 4），标度因子会很散乱，因此，标度因子本身也需要进行平滑处理。

为此，例如（公式 1 9）所示的单纯的低通滤波器可以用于按照时间间隙抑制 $scale_i(ts)$ 的大幅度变动。

（公式 1 9）

$$scale_i(ts) = \alpha \cdot scale_i(ts) + (1 - \alpha) \cdot scale_i(ts - 1)$$

即，平滑处理部 610 通过对标度因子施行抑制各个时间间隙的变动的平滑处理，从而生成平滑的标度因子。具体而言，平滑处理部 610 通过使以下两个值相加来施行平滑处理，所述两个值是指：对当前的

时间间隙中的标度因子乘以 α 而得到的值，和对当前的时间间隙的前一个时间间隙中的标度因子乘以 $(1 - \alpha)$ 而得到的值。

在此，例如设 α 为 0.45。并且，通过改变 α 的大小，从而可以控制效果 $(0 \leq \alpha \leq 1)$ 。

上述 α 的值可以由编码装置一方即音频编码器 10 发送，也可以在发送方控制平滑处理，可以在众多方面发挥效用。当然，以上所述的预先规定的 α 值也可以保持在平滑处理装置中。

但是，在平滑处理中处理的信号能量大的情况下等，能量集中于特定的频带，平滑处理的输出会出现上溢。为此，进行以下的（公式 20）所示的 $scale_i(ts)$ 的限幅处理。

（公式 20）

$$scale_i(ts) = \min(\max(scale_i(ts), 1/\beta), \beta)$$

在此， β 是限幅系数， $\min()$ 和 $\max()$ 分别表示最小值和最大值。

即，限幅处理单元（未图示）对标度因子施行的限幅处理是：在标度因子超过预先规定的上限值的情况下则限制为上限值，在低于预先规定的下限值的情况下则限制为下限值。

在（公式 20）中，按每个通道计算的 $scale_i(ts)$ 表示的意思是：例如 $\beta = 2.82$ 的情况下，上限值被设定为 2.82，下限值被设定为 $1/2.82$ ，则 $scale_i(ts)$ 的值被限制在这个范围。并且，所述阈值 2.82 和 $1/2.82$ 在此仅为一个例子，并非受这些值所限。

运算部 611 通过使标度因子与扩散信号相乘，从而生成标度扩散信号。HPF612 通过对标度扩散信号施行高通滤波处理，从而生成高频扩散信号。加法部 613 使高频扩散信号和直接信号相加，从而生成相加后的信号。

具体而言，运算部 611、HPF612、以及与直接信号进行加法运算的加法部 613，分别像合成滤波器组 902、HPF912、以及加法部 913

那样工作。

然而，上述处理可以像以下（公式 2 1）所示进行组合。

（公式 2 1）

$$\begin{aligned}y_{i,diffuse,scaled,HP}(ts, sb) &= y_{i,diffuse}(ts, sb) \cdot scale_i(ts) \cdot Highpass(sb) \\y_i &= y_{i,direct} + y_{i,diffuse,scaled,HP}\end{aligned}$$

上述的 BPF605 和 BPF606 中的为了节约运算的想法（例如在阻带适用“0”，在通带适用复制处理）也可以适用于高通滤波器 612。

合成滤波器组 614 通过对相加后的信号施行合成滤波处理，从而使相加后的信号转换为时域信号。即，最后通过合成滤波器组 614 而使新的直接信号 y_1 转换为时域信号。

另外，本发明所包括的各个构成构件可以由 LSI (Large Scale Integration) 等集成电路来构成。

而且在本发明，这些装置以及各构成构件的工作可以作为使计算机执行的程序来实现。

（实施例 2）

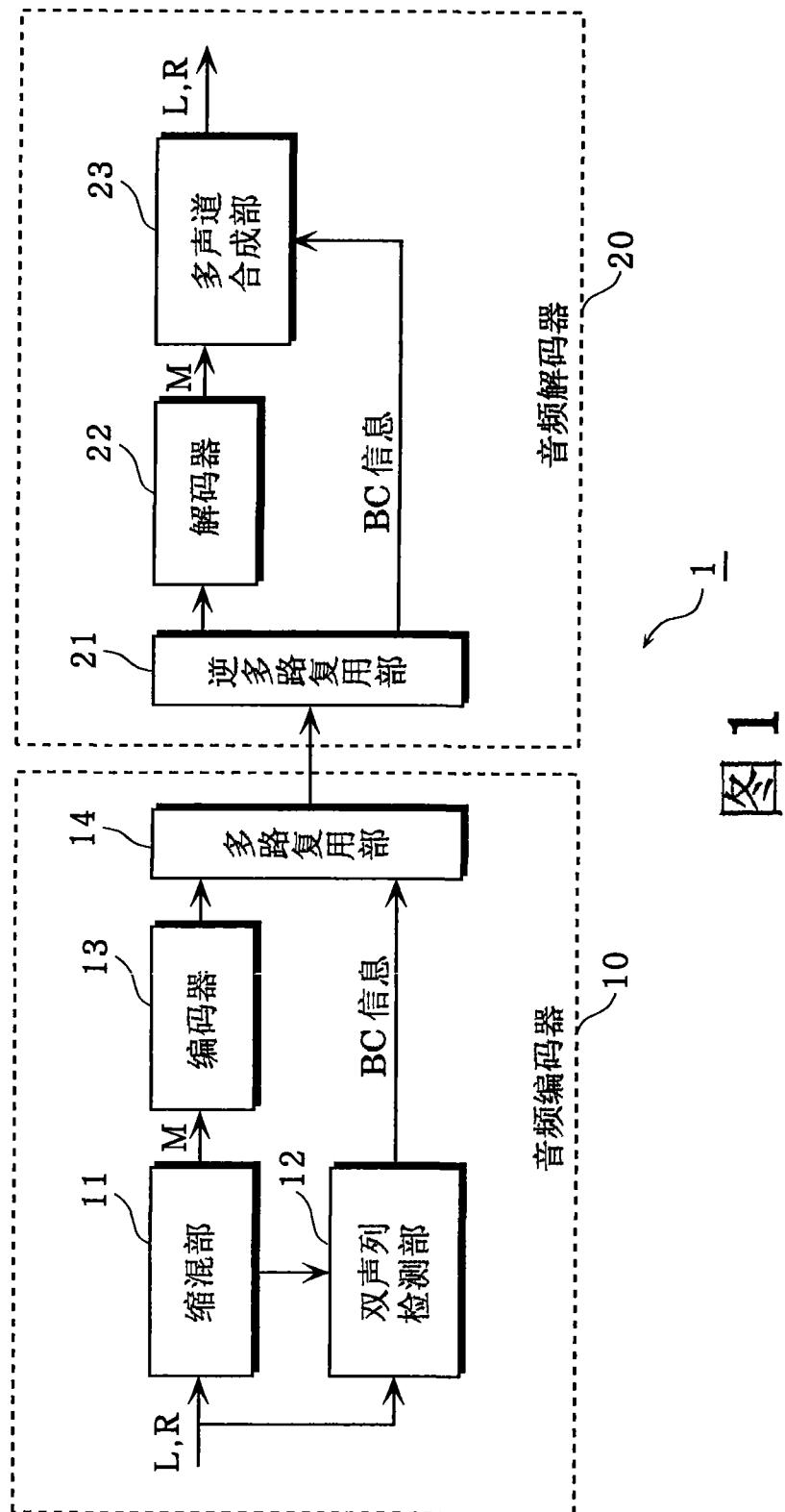
并且，对于决定本发明是否适用，可以设定比特流中的若干个控制标志，并在图 9 所示的时间上的处理装置 600b 的控制部 615，可以通过这些标志，对一部分再建信号的每个帧的工作或不工作进行控制。即，控制部 615 可以按照时间帧或通道来切换是对音频信号施行能量整形还是不施行能量整形。据此，通过对施行或不施行能量整形进行切换，可以实现声音的时间上变动的尖锐性以及确实地声像定位。

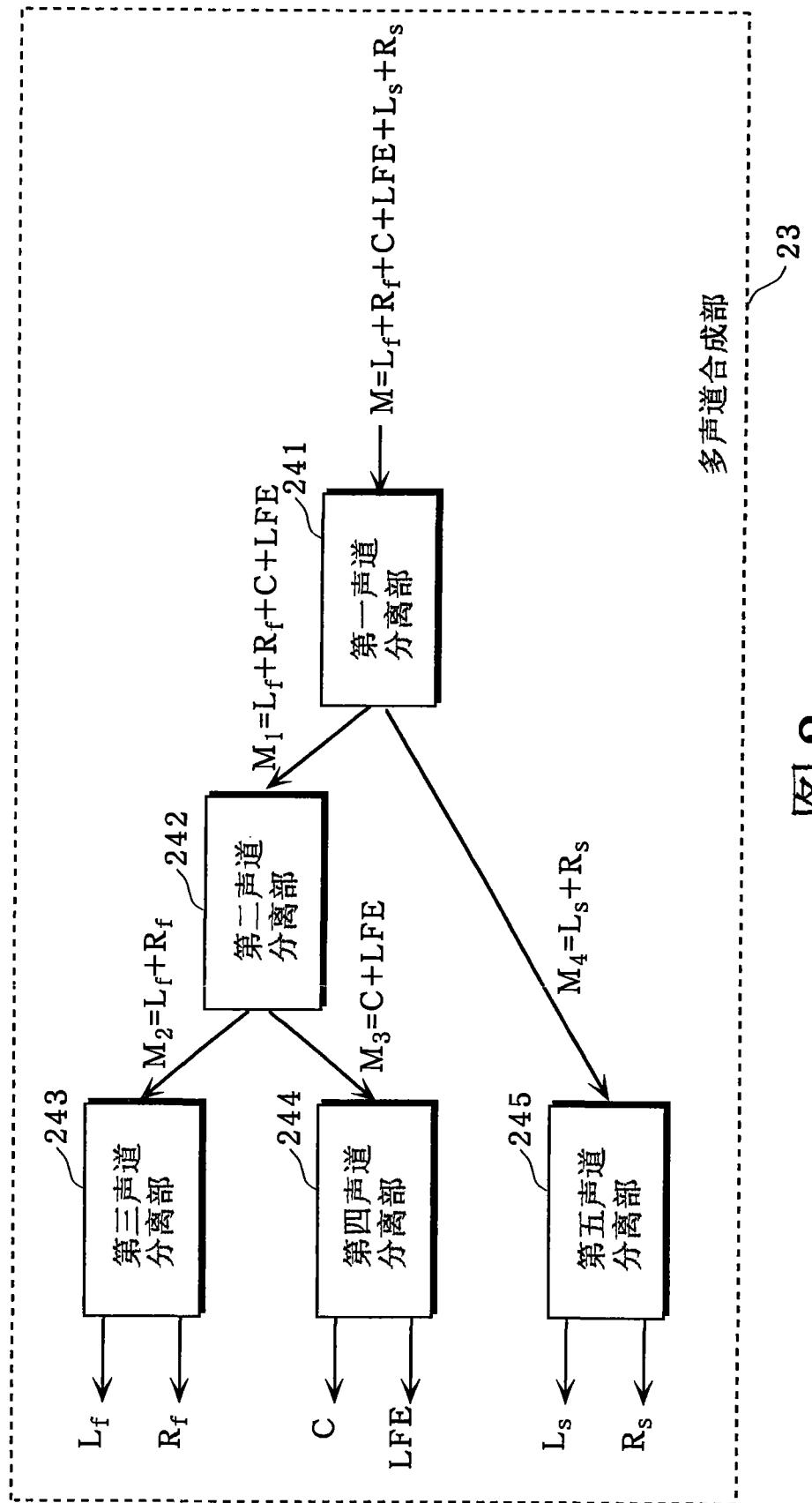
为此，也可以是，例如在编码处理过程中，分析音频通道，进行是否具有伴随急剧变化的能量包络的判断，在存在有相对应的音频通道的情况下，由于需要能量整形，因此将所述控制标志设定为“ON”，使整形处理按照解码时的控制标志来适用。

即也可以是，控制单元 615 按照所述控制标志，选择扩散信号以

及高频扩散信号的任一个，加法部 613 使控制部 615 所选择的信号和直接信号相加。这样，可以随时简单地对施行或不施行能量整形进行切换。

本发明所涉及的能量整形装置是一种可以比芯片尺寸还要小，并且可以减少存储器的必要容量的技术，可以适用于家庭影院系统、车载音响系统、电子游戏系统或移动电话等，并可以适用于希望多声道再生的装置。





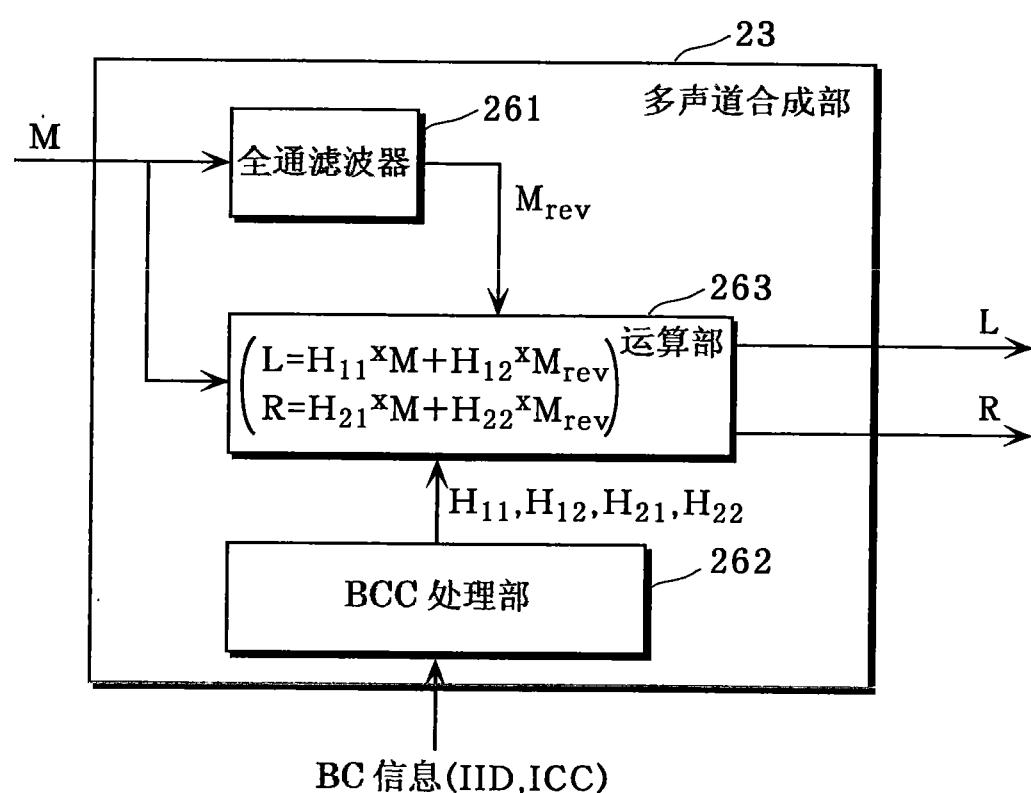


图 3

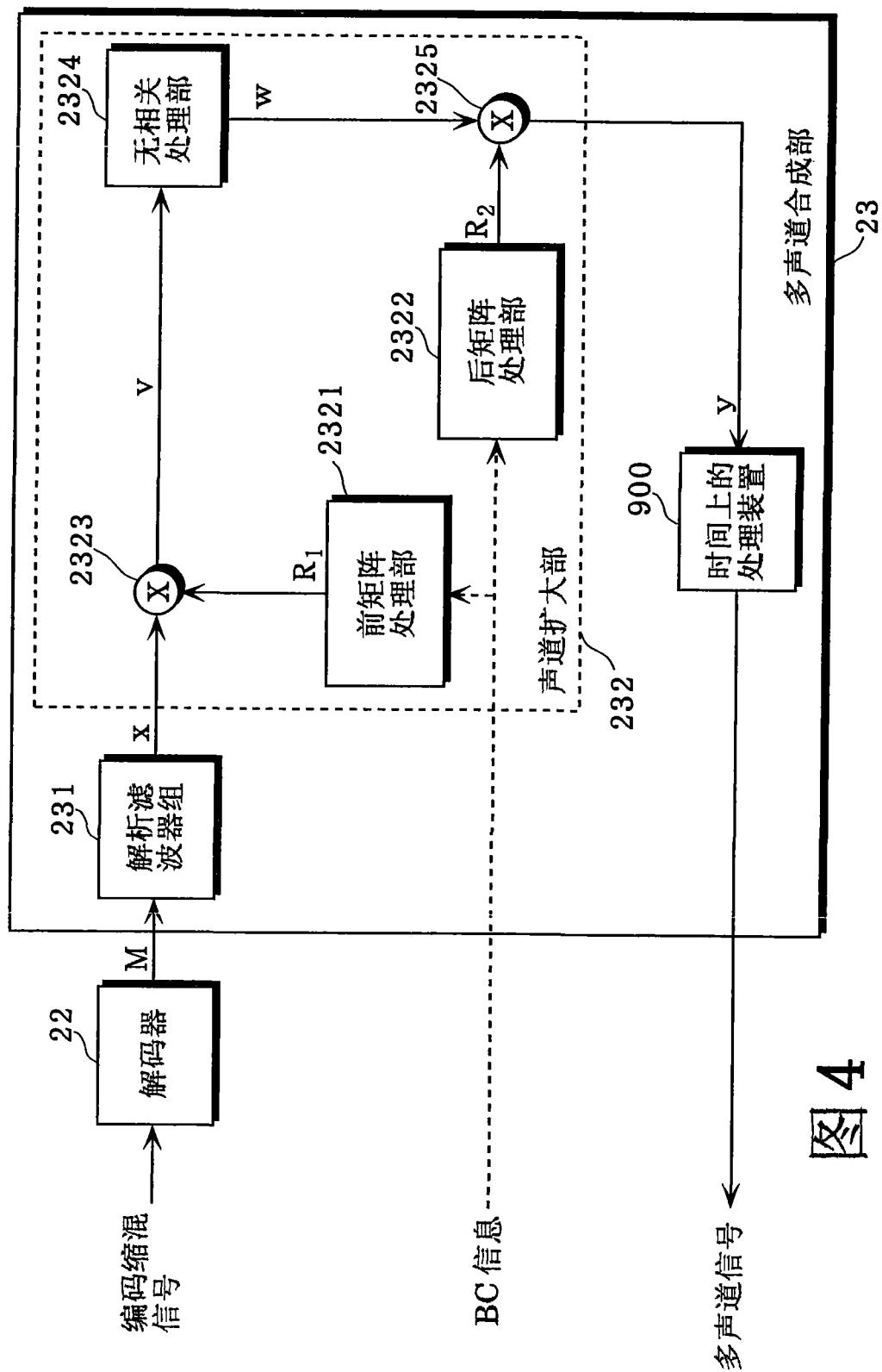


图 4

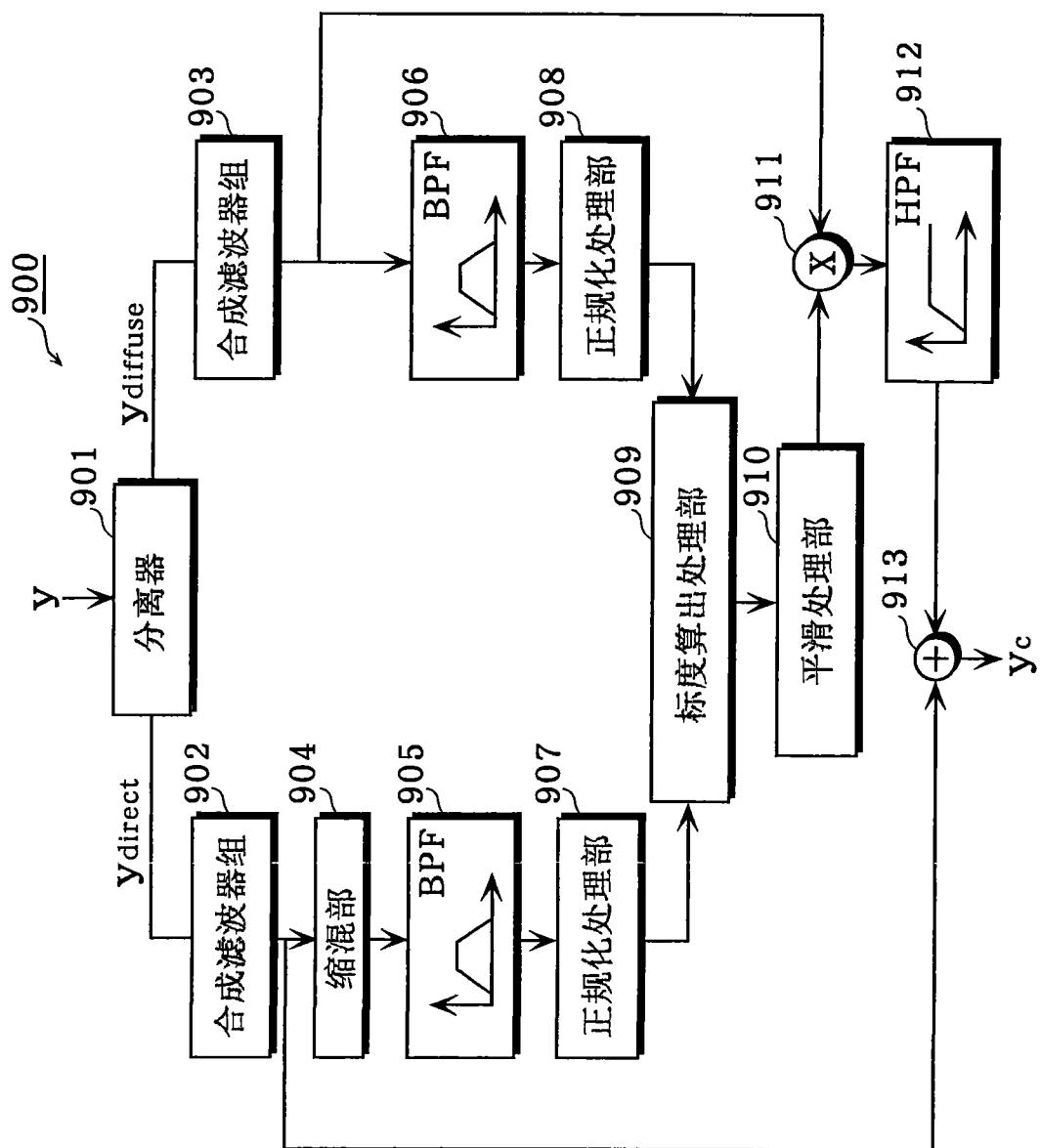


图 5

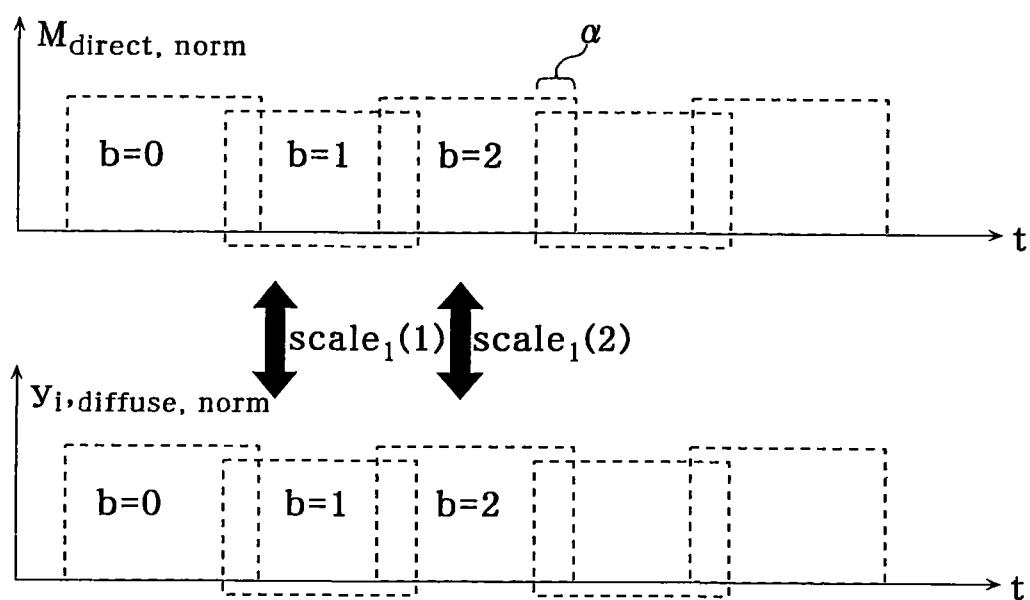
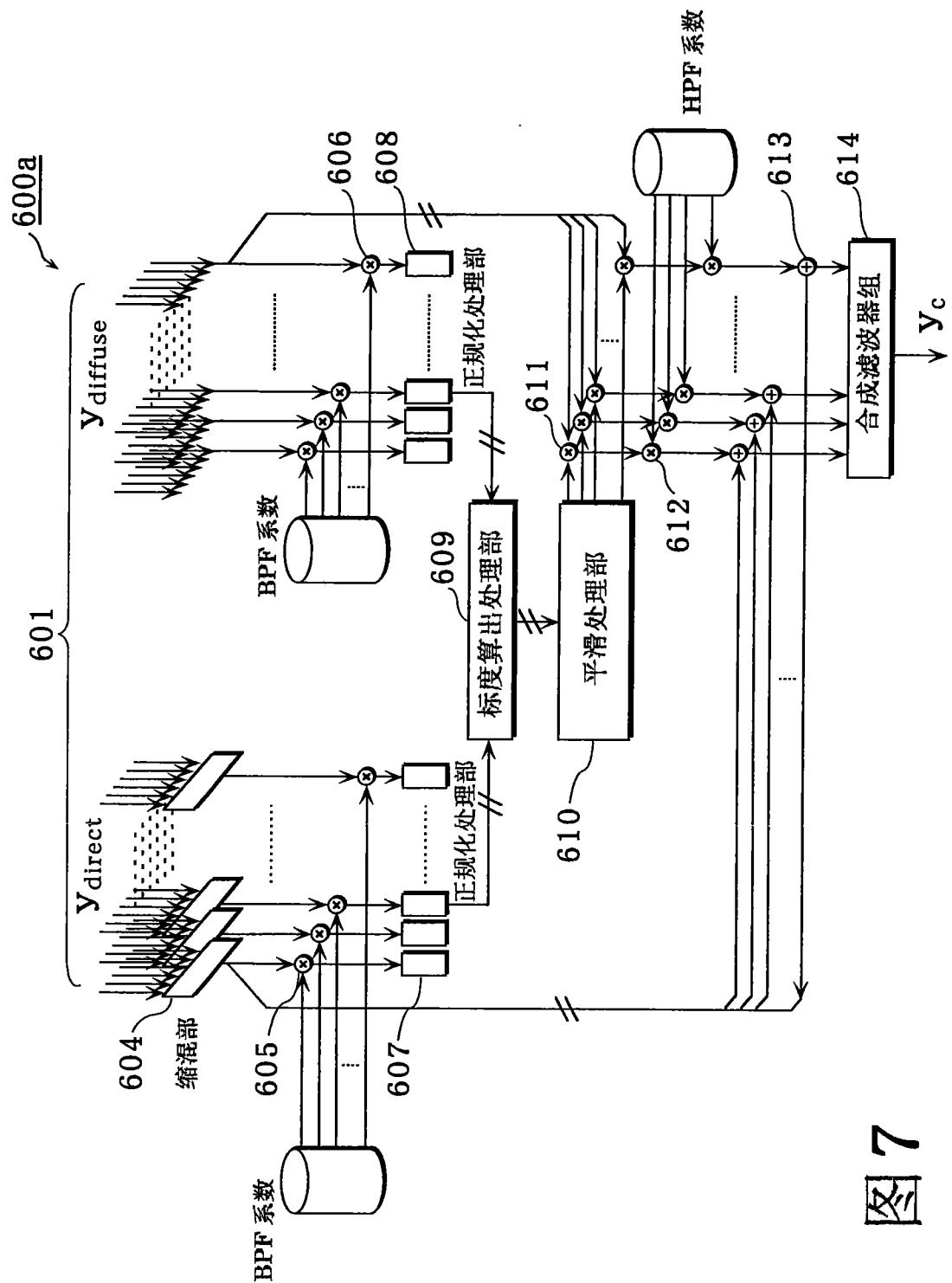
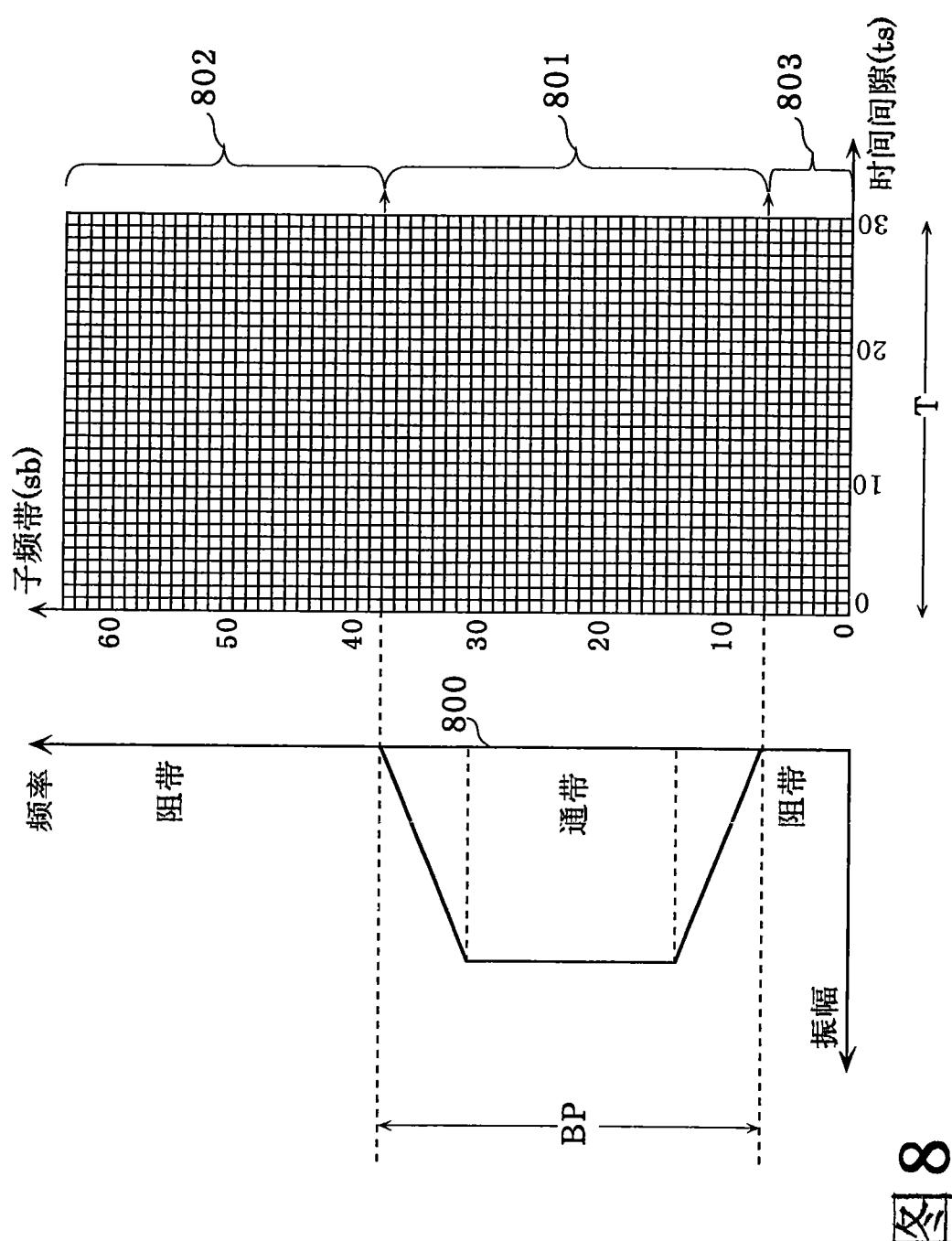


图 6





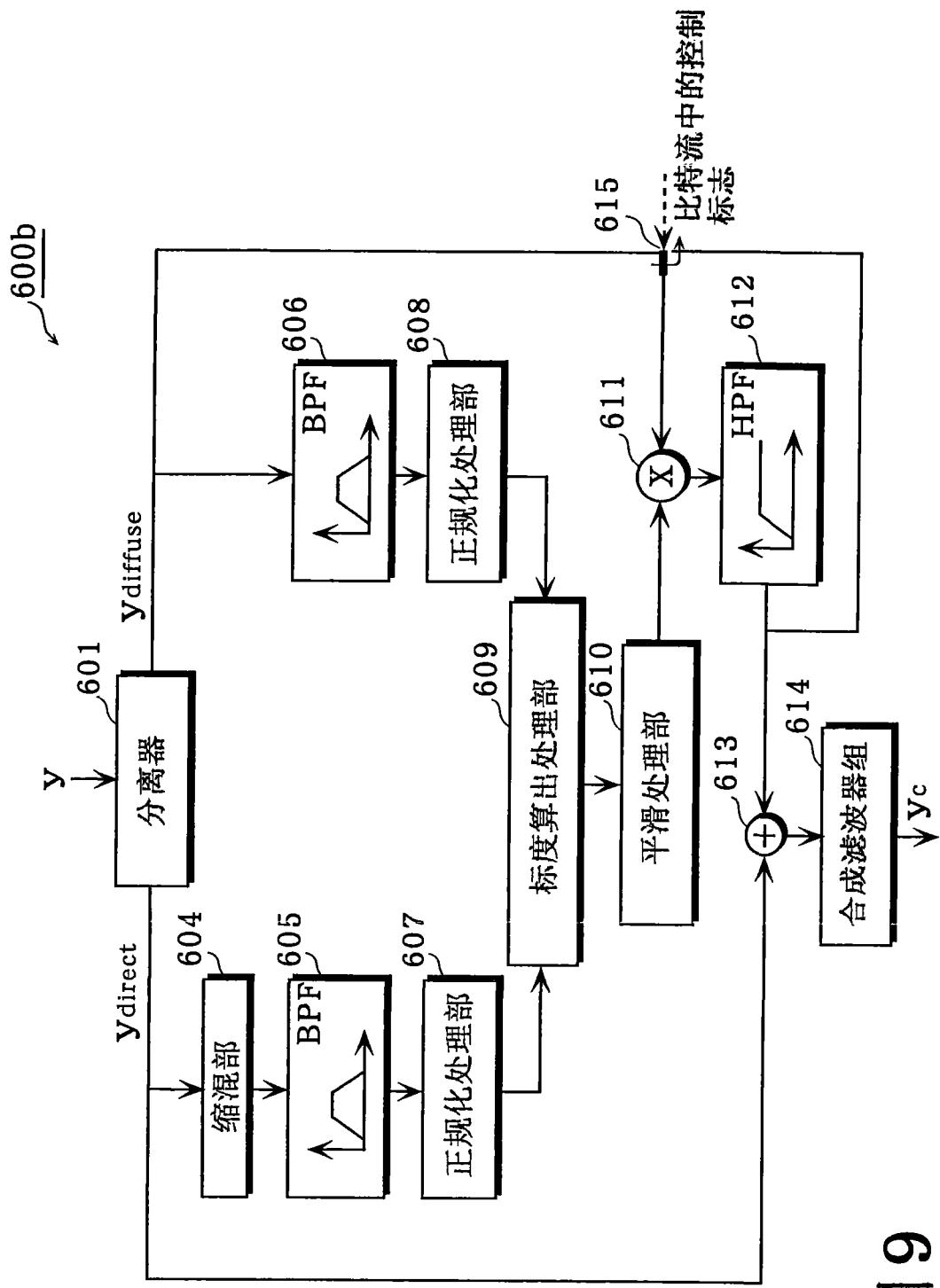


图 9