

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

G10L 19/00

G10L 19/02



[12] 发明专利申请公开说明书

[21] 申请号 200310101234.7

[43] 公开日 2004年6月16日

[11] 公开号 CN 1504993A

[22] 申请日 2003.10.13

[21] 申请号 200310101234.7

[30] 优先权

[32] 2002.11.29 [33] KR [31] 0075529/2002

[71] 申请人 三星电子株式会社

地址 韩国京畿道

[72] 发明人 吴润学 马修·马努

[74] 专利代理机构 北京市柳沈律师事务所

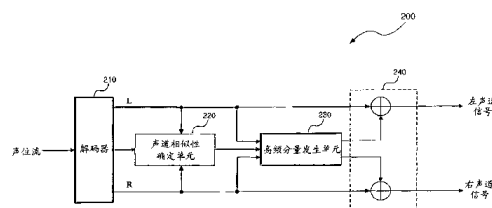
代理人 黄小临 蒲迈文

权利要求书 2 页 说明书 8 页 附图 8 页

[54] 发明名称 用较少的计算量重构高频分量的声频解码方法和装置

[57] 摘要

提供一种以较小的计算量重构高频分量的声频信号解码方法和装置。所述解码装置包括：解码器；声道相似性确定单元；高频分量发生单元；和声频合成单元。所述方法对每个声道信号每隔一帧跳跃地产生各帧的高频分量；在左右声道信号彼此相似时，使用另一个声道信号的相对应的帧产生的高频分量，产生任何一个声道信号跳跃帧的高频分量；在左右声道信号彼此不相似时，使用相关声道信号的前帧，产生每个声道信号的跳跃帧的高频分量。



- 1.一种在解码声频数据时产生高频分量的方法，包括步骤：
利用第一和第二声道信号之间的相似性，产生高频分量。
- 5 2.如权利要求1所述的方法，其中基于第一和第二声道信号之间的和及差的信息获得的信噪比（SNR），确定声道信号之间的相似性。
- 3.如权利要求1所述的方法，其中声频数据包括模式信息。
- 4.如权利要求3所述的方法，还包括步骤：确定模式信息是否是表示在第一和第二声道信号之间大的相似性的共同立体声值，或是表示第一和第二声道信号之间没有相似性的立体声模式值。
- 10 5.如权利要求1所述的方法，还包括步骤：在第一和第二声道信号彼此相似时，
 产生每个声道信号的仅一些帧的高频分量；和
 使用另一声道信号的某些帧的产生的高频分量，产生每个声道信号的其他一些帧的高频分量。
- 15 6.如权利要求5所述的方法，其中通过适当修改一些帧的高频分量产生其它帧的高频分量。
- 7.如权利要求1所述的方法，还包括步骤：在第一和第二声道信号彼此不相似时，
- 20 产生每个声道信号的仅一些帧的高频分量；和
 使用相关声道信号的一些帧的产生的高频分量，产生每个声道信号的其他帧的高频分量。
- 8.如权利要求7所述的方法，其中通过适当修改一些帧的高频分量产生其它帧的高频分量。
- 25 9.一种重构高频分量的声频解码方法，包括步骤：
 (a)接收编码的声频数据，解码接收的数据，和输出第一和第二声道的解码的声频信号；
 (b)产生每个第一和第二声道信号的仅一些帧的高频分量；
 (c)确定在第一声道和第二声道信号之间的相似性；
- 30 (d)在第一和第二声道信号彼此相似时，使用另一声道信号的一些帧的产生的高频分量，产生每个声道信号的另一一些帧的高频分量；和

(e)将产生的高频分量与解码的声频信号结合，并输出结合的声频信号。

10.如权利要求9所述的方法，其中步骤(c)包括步骤：基于从第一和第二声道信号的和及差信息获得的信噪比(SNR)，确定声道信号之间的相似性。

5 11.如权利要求9所述的方法，其中声频数据包括模式信息。

12.如权利要求9所述的方法，其中步骤(c)包括步骤：确定模式信息是否是表示在第一和第二声道信号之间大的相似性的共同立体声值，或是表示第一和第二声道信号之间无相似性的立体声模式值。

10 13.如权利要求9所述的方法，还包括步骤：在确定第一和第二声道信号彼此不相似时，使用相关声道信号的一些帧的产生的高频分量，产生每个声道信号的其它帧的高频分量。

14.一种重构高频分量的声频解码装置，包括：

声频解码器，接收编码的声频数据，解码接收的数据，和输出第一和第二声道的解码的声频数据；

15 声道相似性确定单元，确定第一和第二声道信号之间的相似性；

高频分量发生单元，基于第一和第二声道信号之间的相似性，产生每个声道的声频信号的高频分量；和

声频合成单元，将解码的声频信号与产生的高频分量结合，并输出结合的声频数据。

20 15.如权利要求14所述的装置，其中高频分量发生单元被配置成在第一和第二声道信号彼此相似时，产生每个第一和第二声道信号的仅一些帧的高频分量，然后使用另一声道信号的一些帧的产生的高频分量，产生每个声道信号的其它帧的高频分量。

25 16.如权利要求14所述的装置，其中高频分量发生单元被配置成在第一和第二声道信号彼此不相似时，产生每个声道的仅一些帧的高频分量，然后使用相关声道信号的一些帧的产生的高频分量，产生每个声道信号其它帧的高频分量。

17.一种计算机可读记录介质，其中记录了在计算机中执行权利要求1-13任何一项权利要求所述方法的程序。

30

用较少的计算量重构高频分量的
声频解码方法和装置

5

技术领域

本发明涉及声频编码方法和装置，特别涉及一通过用较少的计算量重构声频信号的高频分量而能够获得和输出高质量的声频信号的声频解码方法和装置。

10

背景技术

一般，在声频编码中使用心理学模型高效地压缩声频数据，使得少量位分配到人耳不能收听的高频分量中。此时，增加了压缩比，但是损失了高频声信号。由于损失高频声信号，所以在再现声频数据时，声音的音调改变，可读性降低，产生模糊的声音。因此，需要后处理方法，重构损失的高频分量来加强声音质量，以便完全地再现原始声音的音调和提高声频信号的可读性。

结合图 1 说明加强声频信号的声音质量的后处理方法。参见图 1，如果输入编码的信号，它们被分成左右声道信号，并通过解码器 110 分别解码被分开的信号。然后，分别用第一和第二高频分量发生器单元 120 和 130 重构解码的左右声道信号的高频分量。

但是，因为大多数的声频信号的左右声道的声频信号彼此基本相似和高度冗余，所以不对它们单独编码。因此，存在的问题是，分开地重构左右声道信号的现有技术的后处理方法不能够有效地利用声道信号之间地相似性，因而不必要地增加了计算时间。

发明内容

本发明的一个目的是提供一种即使使用较少的计算也能加强声频信号的声音质量的声频解码方法和装置。

为实现上述目的，根据本发明的一个方面，提供了一种声频解码方法，该方法包括步骤：对每个声道信号每隔一帧跳跃地产生帧的高频分量；在左

右声道信号彼此相似时, 使用所产生的另一个声道信号的相应帧的高频分量, 产生任何一个声道信号的跳跃的帧的高频分量; 和在左右声道信号彼此不相似时, 用相关的声道信号的前帧, 产生每个声道信号的跳跃帧的高频分量。

- 5 根据本发明的另一方面, 提供一种重构高频分量的声频解码装置, 它包括: 声频解码器, 它接收编码的声频数据, 解码接收的数据, 和输出第一和第二声道的解码的声频数据; 声道相似性确定单元, 它确定在第一和第二声道信号之间的相似性; 高频分量发生单元, 它基于第一和声道信号之间的相似性, 产生每个声道的声频信号的高频分量; 和声频合成单元, 它将解码的
- 10 声频信号与产生的高频分量结合并输出结合的声频信号。

附图说明

通过结合以下优选实施例的说明将明了本发明的上述目的和其它特征。

- 图 1 是应用现有技术后处理算法的声频解码装置的方框图;
- 15 图 2 是根据本发明的声音解码装置的配置示意图;
- 图 3 是 MPEG (活动图像专家组) -1 层 3 的格式示意图;
- 图 4 是根据本发明的声频解码方法的整个处理过程流程图;
- 图 5 是根据本发明对每个声道信号每隔一个帧跳跃地, 产生高频分量的处理图;
- 20 图 6 是在声道信号彼此不相似时, 产生左右声道信号的高频分量的方法示意图;
- 图 7 是在声道信号彼此相似时, 产生左右声道信号的高频分量的方法示意图; 和
- 图 8 的图表, 示出了与现有技术比较, 本发明的声频解码方法声音质量
- 25 的加强。

具体实施方式

下面参照附图详细说明本发明的声频解码装置的配置和工作。

- 图 2 示出声频解码装置 200 的配置示意图。参见图 2, 声频解码装置 200
- 30 包括: 解码器 210; 声道相似性确定单元 220; 高频分量发生单元 230; 和声频合成单元 240。装置 200 被配置成解码声频位流, 然后重构来自解码的声

频信号的各声道信号的高频分量。

解码器 210 解码输入的声频位流并产生声频信号。即，从输入的声频位流解码声频数据，然后将解码的数据去量化 (dequantize)，以恢复在声频数据的编码处理中以前进行的量化操作，使得输出原始声频信号。

- 5 这里，在解码器 210 中采用的解码方法可以根据压缩声频信号使用的编码类型，例如比例因子编码、AC-3、MEPG 和霍夫曼 (Huffman) 编码等变化。然而，因为在声频信号处理中使用的编码器的配置和操作基本彼此相同，所以，在此省略其详细说明。

- 10 同时，已知 SBR (频谱带复制 Spectral Band Replication)，即，从声频信号的低频范围重构高频范围的算法是迄今提出的声音质量强化的几种后处理算法当中的最有效的技术。但是，不能够将 SBR2 应用到各种声频编解码器，因为它是依靠 MPEG-1 层 3 的后处理算法。另外，与 SBR2 比较，SBR1 能够用于各种声频编解码器，但是它将对每帧的左右声道信号进行后处理操作。因此，操作中不能够有效地利用两个声道之间的相似性，结果，增加了
- 15 计算时间。因此，存在着几乎不能够将这种算法用于相关的产品的限制。

因此，为了减少与能够用于各种声频编解码器的相应于 SBR1 (下面简称为 SBR) 的问题的大计算量，和具有优良的重构性能，将本发明配置成这样，通过声道相似性确定单元 220 和 230，有效地利用声道的相似性，即使使用小的计算量也能够重构高频分量。

- 20 在输入解码的声频信号时，声道相似性确定单元 220 分析输入的声频信号是否包括模式信息。然后，声道相似性确定单元 220 根据该模式信息确定左右声道信号之间的相似性。否则，基于从声道信号之间的和及差信息获得的 SNR (信噪比)，确定声道信号之间的相似性。

- 25 这里，在声频信号不包括模式信息时使用 SNR 确定声道信号之间相似性的原因是，基于从声道信号之间的和及差获得的 SNR 值，能够容易地确定左右声道之间的相似性，因为在压缩比高时，在通常的声频编解码器中频繁编码该和及差信息。

下面，为了更好理解本发明，通过 MPEG-1 层 3 声频信号的例子说明确定左右声道信号之间相似性的方法。

- 30 图 3 示出了 MPEG-1 层 3 声频流的格式。

MEPG-1 层 3 声频流由多个 AAU (声音访问单元) 300 构成。每个 AAU300

是能够单个解码并含有预定的固定数目的压缩数据的取样的最小的单元。

每个 AAU300 包括: 标题 310, 循环冗余校验 (CRC) 320, 声频数据 330, 和辅助数据 340。

标题 310 含有关于同步字、ID、层、保护位的存在、位速率指数、取样
5 频率、填充位的存在、专用使用位、模式、模式扩展、版权、原始/复制和强调特征的信息。

CRC320 是可任选的, 16 位长, 并且标题 310 定义在每个 AAU 300 中是否包括 CRC 320。

声频数据 330 是含有压缩声音数据的部分。

10 辅助数据 340 是在每个声频数据 330 的尾端未达到相关 AAU 的尾端时的剩余的部分。在辅助数据 340 中能够包括除了 MPEG 声频数据外的任何数据

。如图 3 所示, MP3 声频位流的标题 310 含有模式信息, 该模式信息表示所述流是否使用声道信号之间的相似性被压缩。因此, 通过分析输入的 MP3
15 声频位流的模式信息, 能够容易地确定在声道信号之间的相似性。

因此, 在包括上述模式信息的 MPEG-1 层 3 声频信号输入时, 声道相似性确定单元 220 分析在输入信号中的模式信息, 并根据模式信息是否是具有左右声道信号之间很大相似性的共同立体声 (joint stereo) 模式值, 或是具有在声道信号之间很小相似性的立体声模式值, 来确定在声道信号之间的
20 相似性。

在另一方面, 在解码的声频信号中不包括模式信息的情况, 声道相似性确定单元 220 基于从声频信号获得的声道信号之间的和及差信息, 计算与代表声道信号之间相似性的参数相对应的 SNR。然后, 如果计算的 SNR 值小于声道信号之间相似性的阈值, 则确定这两个声道信号彼此相似。否则确定这
25 两个声道信号不彼此相似。

即, 将从关于声道的信号之间的和及差的信息获得的 SNR 值用作代表声道之间相似性的参数。现在详细说明基于关于两个声道信号之间的和及差的信息计算 SNR 值的方法。

首先, 计算在这两个声道信号之间的和及差的能量值。然后, 对通过用
30 这两个声道信号之间的和及差的相加值除以声道信号之间的差的能量值获得的值取对数。然后, 用 10 乘这个对数值。此时, 为了减小计算能量值需

要的计算,最好是使用这个信号之间的和及差的数值。

这里,可对声道信号之间相似性的阈值赋予试验值。在本发明中,20分贝的值确定为声道信号之间相似性的阈值。

因此,声道相似性确定单元 220 分析声频信号是否包括模式信息。如果是,则确定单元基于该模式信息确定左右声道信号之间的相似性。否则,确定单元基于从这两个声道信号之间的和及差的信息获得的 SNR,确定相似性。

本领域的技术人员能够作出确定左右声道信号之间相似性的方法的各种修改和等同的变型。例如,如果不是 MPEG-1 层 3,而是 AC-3 声频信号包括在左右声道信号之间的差的信息中,也能够确定左右声道信号之间的相似性。另外,如果在声频位流中存在线性预测系数,则通过解码这个线性预测系数和建模谱包络信号也能够确定左右声道信号之间的相似性。

另外,高频分量发生单元 230 使用 SBR 算法,对每个声道每隔一个帧跳跃地重构左右声道信号的高频分量。然后,在左右声道信号彼此相似时,在一个声道中产生的高频分量重构另外一个声道信号的跳跃的帧的高频分量。在左右声道信号彼此不相似时,用每个声道信号的前帧的高频分量,重构相关声道信号的跳跃帧的高频分量。稍后参照图 5 和 7 说明其详细情况。

在高频发生单元 230 重构每个声道信号的高频分量时,声频合成单元 240 产生通过将发生的高频分量加到解码的视频信号获得的输出。因此,依靠在声道信号之间的相似性,合适地重构高频分量,从而能够减少不必要的计算,并且也能够加强声音质量。

下面参照附图详细说明本发明的声频信号解码方法。

图 4 是根据本发明的声频信号解码方法的整个处理过程流程图。

首先,解码器 210 解码输入的声频位流并输出声频信号(S10)。这里,这个解码方法能够根据压缩声频信号使用的如 AC-3、MEPG 和霍夫曼 编码等编码类型变化。

然后,高频分量发生单元 230 使用 SBR 算法对每个声道每隔一个帧跳跃地重构左右声道信号的高频分量(S20)。下面参照图 5 对上述更具体说明。

图 5 是根据本发明对每个声道信号每隔一个帧跳跃地产生高频分量的处理过程示意图。参见图 5,高频发生单元 230 分别对左右声道信号每隔一个帧跳跃地重构高频分量。

即，从时间 t_1 的帧产生左声道的高频分量 (L_{t1})，而在时间 t_2 的帧产生右声道 (R_{t2}) 的高频分量。相似地，这个处理在时间 t_3 、 t_4 、 t_5 等被重复地进行。

然后，声道相似性确定单元 220 确定左右声道信号之间的相似性 (S30 5)。下面简要说明确定声道信号之间相似性的方法。

首先，声道相似性确定单元 220 分析解码的声频信号是否包括模式信息。如果是，确定单元 220 基于模式信息确定声道信号之间的相似性，即，根据模式信息是具有左右声道信号之间大的相似性的共同立体声模式值，或是具有声道信号之间小的相似性的立体声模式值，确定声道信号之间的相似性 10。

另一方面，在解码的声频信号中不包括模式信息时，声道相似性确定单元 220 基于从声频信号获得的声道信号之间的和及差的信息，计算相应于表示声道信号之间相似性的参数的 SNR。然后，如果计算的 SNR 值小于声道信号之间的相似性的阈值，则确定这两个声道信号是彼此相似的。否则，确定 15 这两个声道信号彼此不相似。即，如果在解码声频信号中不含有模式信息，则将从声道信号的和及差信息获得的 SNR 作为表示声道信号之间相似性的参数，然后与 20 分贝的阈值比较，确定声道信号之间的相似性。

已就图 2 和 3 说明了依靠模式信息确定声道信号之间相似性的方法，在此省略其详细说明。

另外，在声道相似性确定单元 220 确定左右声道信号彼此不相似时，高频分量发生单元 230，使用每个声道信号的前帧的高频分量，重构跳跃帧的高频分量，从而产生各声道信号的高频分量 (S40)。下面参照图 6 更详细说明这个处理。

图 6 是在两个声道信号彼此不相似时的产生左右声道信号的方法示意图。参见图 6，在两个声道信号彼此不相似时，高频分量发生单元 230 用对 25 每个声道信号的前帧产生的高频分量 (在每隔一个帧跳跃时产生的高频分量)，重构跳跃的帧的高频分量。

也就是说，用在时间 t_1 的左声道信号信号的高频分量 L_{t1} 代替跳跃的帧的高频分量，即在时间 t_2 的左声道的高频分量 L_{t2} 。相似地，在时间 t_2 30 的右声道信号的高频分量 R_{t2} 代替在时间 t_3 的高频分量 R_{t3} 。

在另一方面，在声道相似性确定单元 220 确定左右声道信号彼此相似

时，高频分量发生单元 230 利用从一个声道信号产生的高频分量，重构另一声道信号的高频分量 (S50)。下面参照图 7 详细说明这个处理。

图 7 是在左右声道信号彼此相似时的重构每个声道信号的高频分量的方法示意图。参见图 7，在确定左右声道彼此相似时，高频分量发生单元 230 使得左声道信号的高频分量分别代替左右声道信号的跳跃帧的高频分量。此时，从每个声道信号产生的高频分量能够乘以预定的修改值（例如，特定的常数）并用于产生另一声道信号的高频分量。

即，左声道信号高频分量 (Lt1) 代替在时间 t1 的右声道信号相对应的高频分量 Rt1，并且右声道信号的高频分量 (Rt2) 代替时间 t2 的左声道的高频分量 (Lt2)。

此时，因为左右声道信号通常彼此很相似，所以声音质量的降低变得最小。而且，对每个声道信号每隔一帧跳跃地产生高频分量，并且将它们有效地用于另一声道信号的高频分量。因此，与传统的 SBR 算法相比较，计算量约能够减少 30%。

最后，产生的高频分量与解码的声频信号结合，然后输出结合的信号 (S60)。

总的来说，因为多数声频信号的左右声道信号彼此相似，所以根据本发明的解码方法解码声频信号位流，使得重构高频分量需要的计算，与现有技术相比，约减少 30%。

图 8 示出了与现有技术的 SBR 和 MP3 方法相比较，本发明加强声音质量的例子。为了评价在 64kbps 的比率压缩的包括 3 首爵士音乐、9 首流行音乐、7 首摇滚音乐和 6 首古典音乐的各种歌曲的声频信号的声音质量，进行了 14 次试验。使用用于测量压缩数字声响/声频信号的周知的系统的歌剧工具 (opera tool)，作为声音质量评价程序。按照由歌剧工具测量的值接近零，确定改进重构的声音质量。

如图 8 所示，能够理解，根据本发明重构高频分量的方法再现的声频信号的声音质量，几乎与现有技术 SBR 和 MP3 方法再现的声频信号的声音质量相同或它们的差别可忽略。

因此，与传统的 SBR 算法相比较，本发明使得即使在计算量减少约 30% 的情况下也能够输出高质量的声频信号，而现有技术的 SBR 算法，尽管也具有有良好的加强声音质量的作用，但是由于过量的计算时间，实际应用到相

关产品是困难的。

而且，本发明的优选实施例能够以计算机可执行的程序的形式实现。另外，通过计算机可读记录介质，所述程序能够在数字计算机上运行。

所述计算机可读记录介质包括磁性记录介质（如 ROM、软盘、硬盘等），
5 光学可读介质（如 CD ROM、DVD 等）和载波（如通过因特网传输）。

虽然现有技术能够使得声音质量改善，但是由于过多的计算时间，现有技术应用到相关产品是困难的，但是根据上述的本发明，能够解决这个关键问题。因此，优点是重构高频分量需要的计算时间能够约减少了 30%。

虽然结合图示的优选实施例说明了本发明，但是本领域的技术人员应当
10 理解，在不偏离本发明的范围和精神的情况下，能够做出各种改变。因此，本发明的优选实施例是示范说明，而不是限定本发明。本发明由权利要求限定，各种修改方案应包括在本发明内。

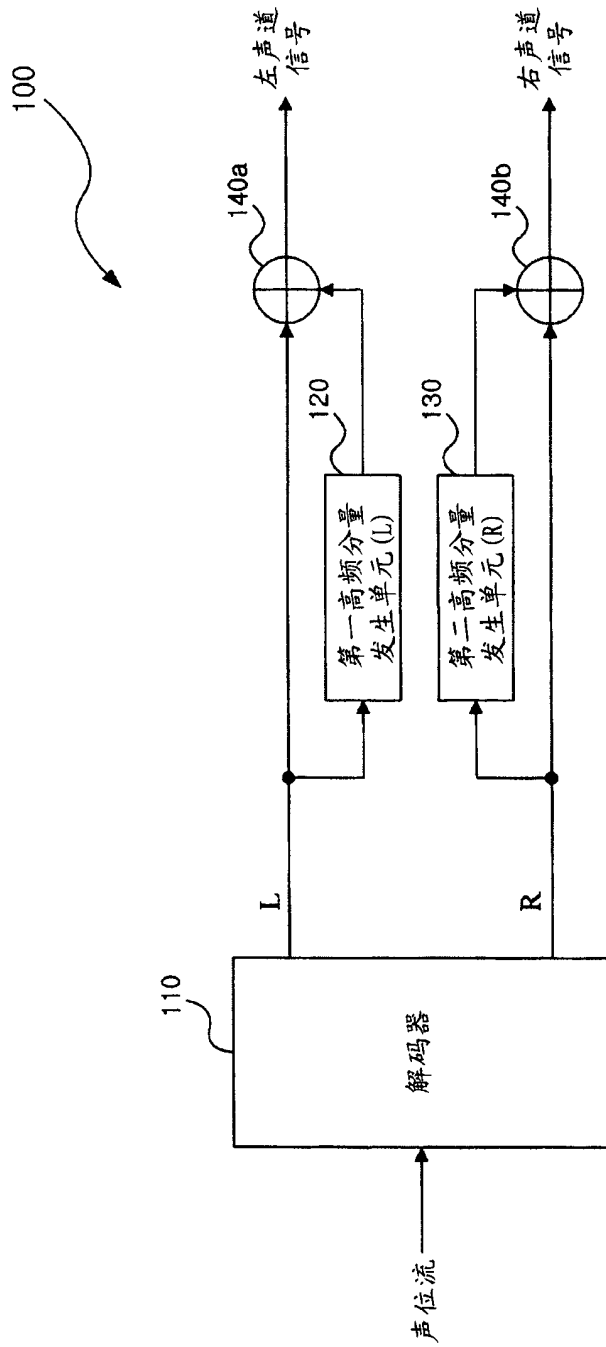


图 1

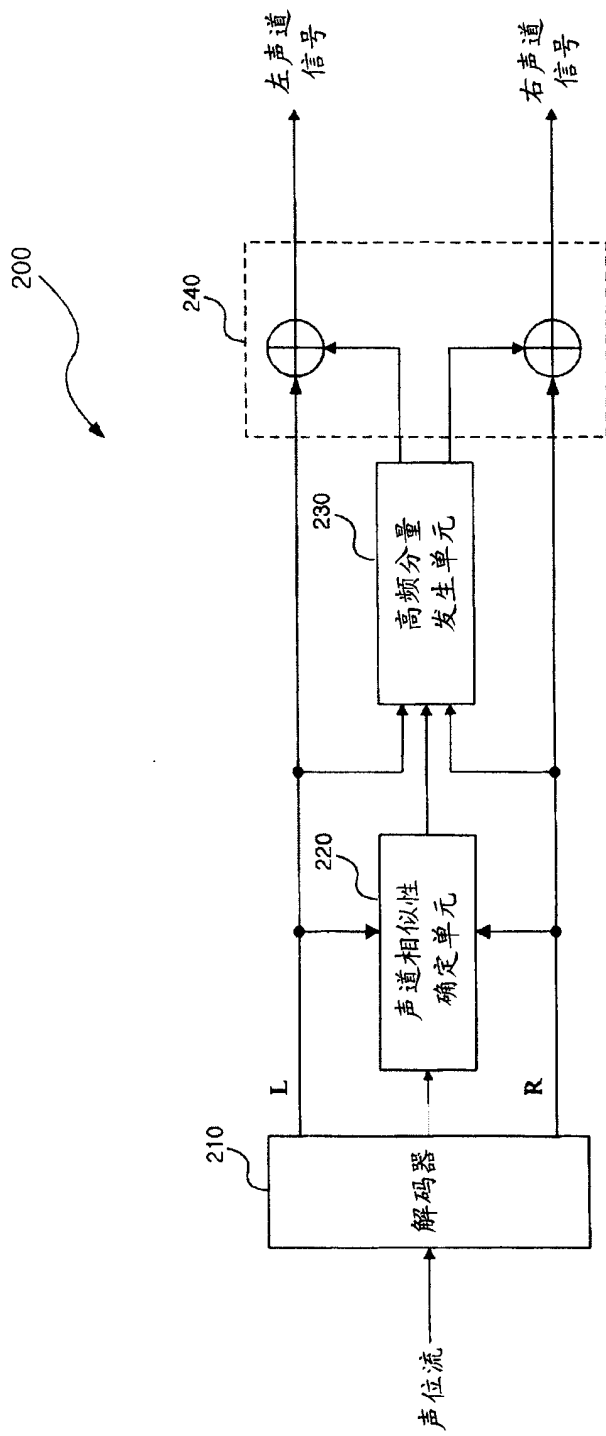


图 2

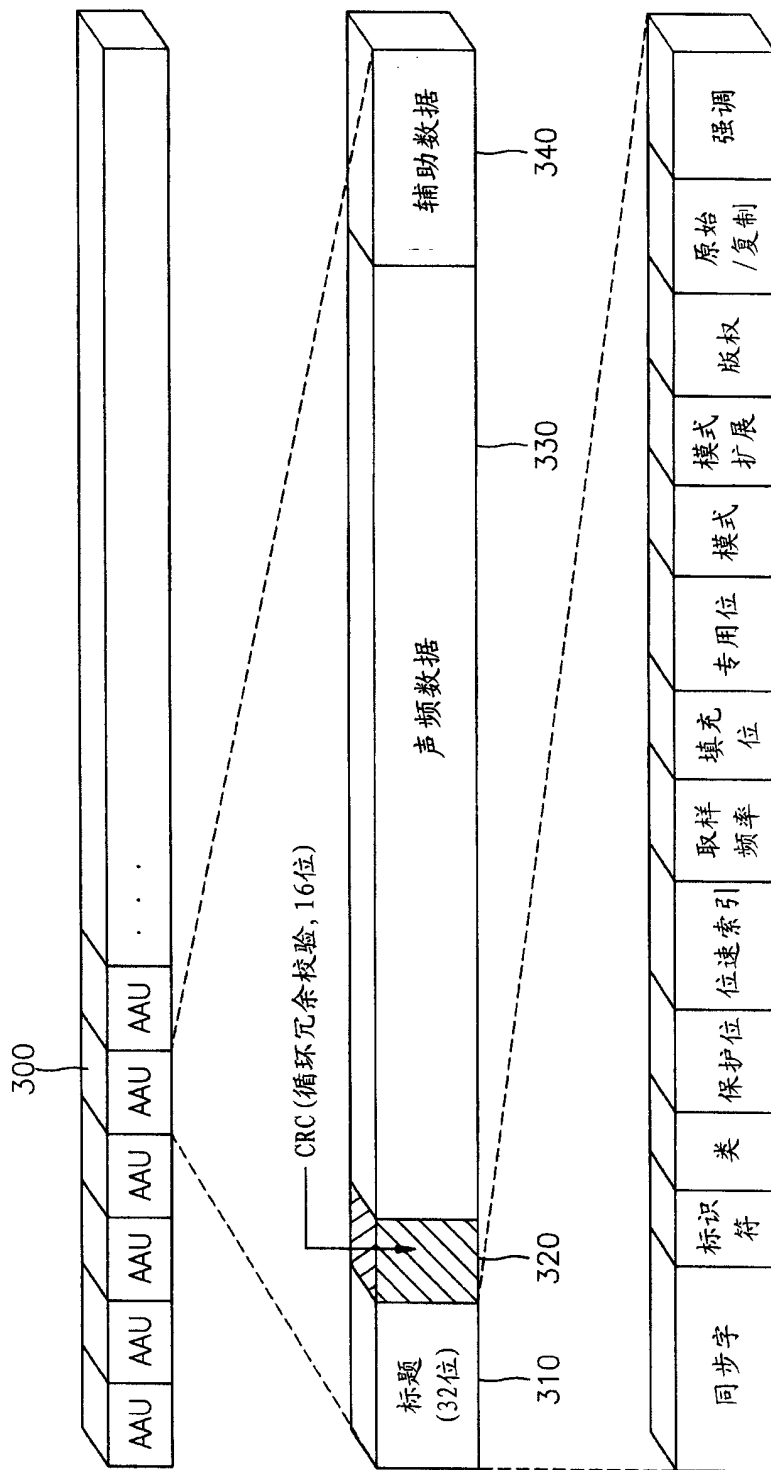


图 3

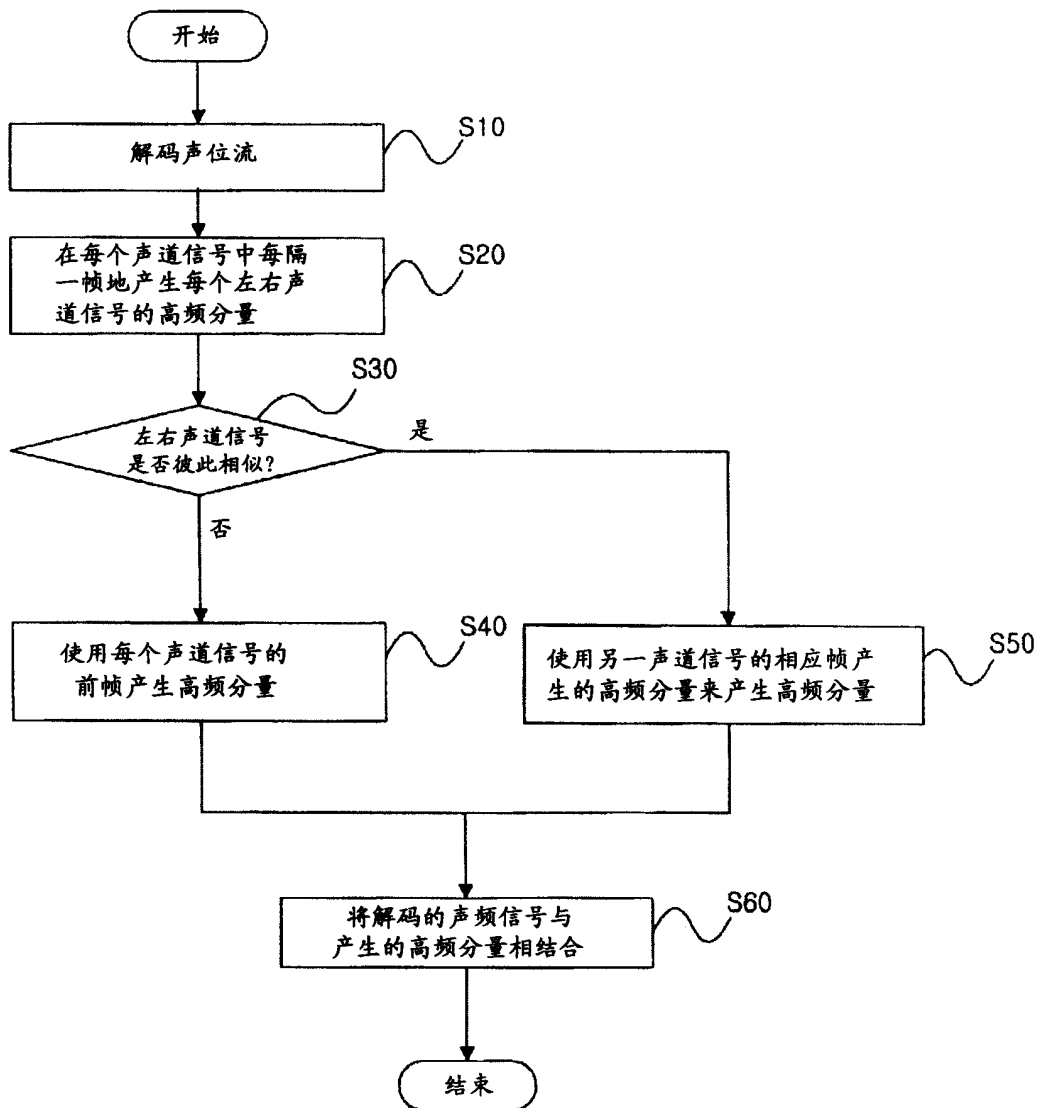


图 4

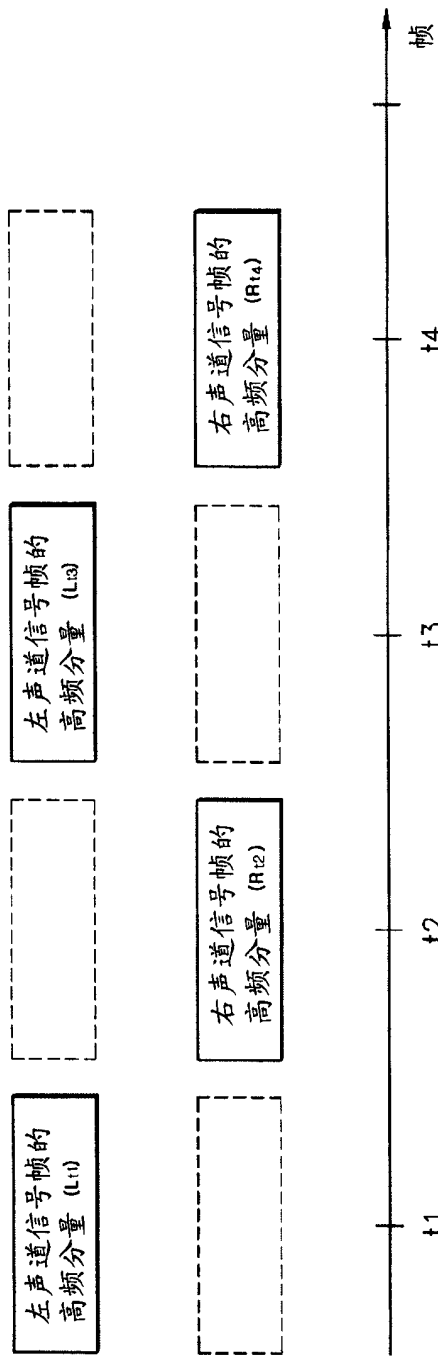


图 5

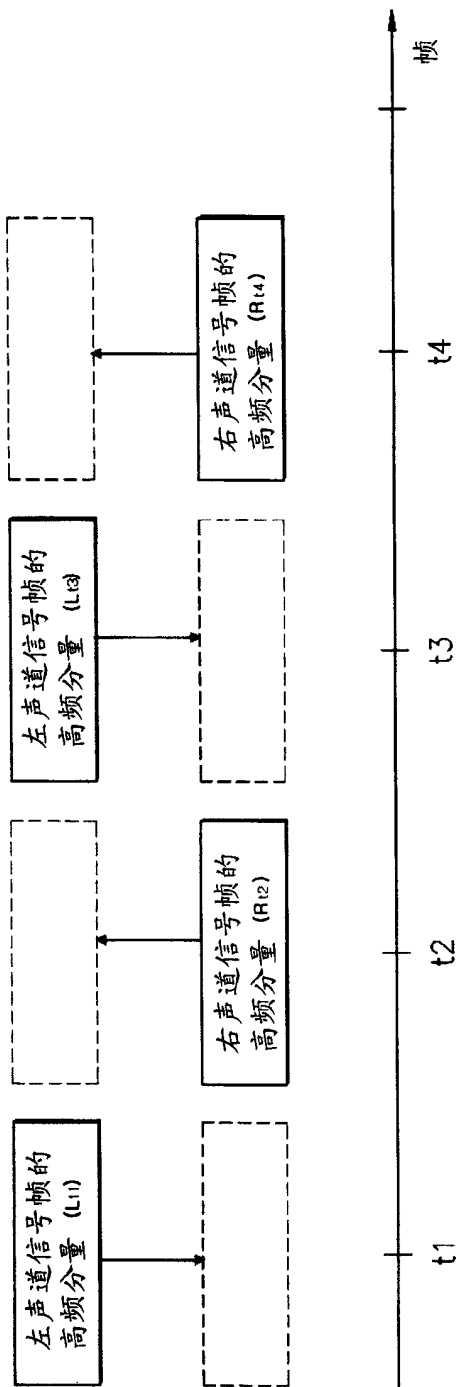


图 6

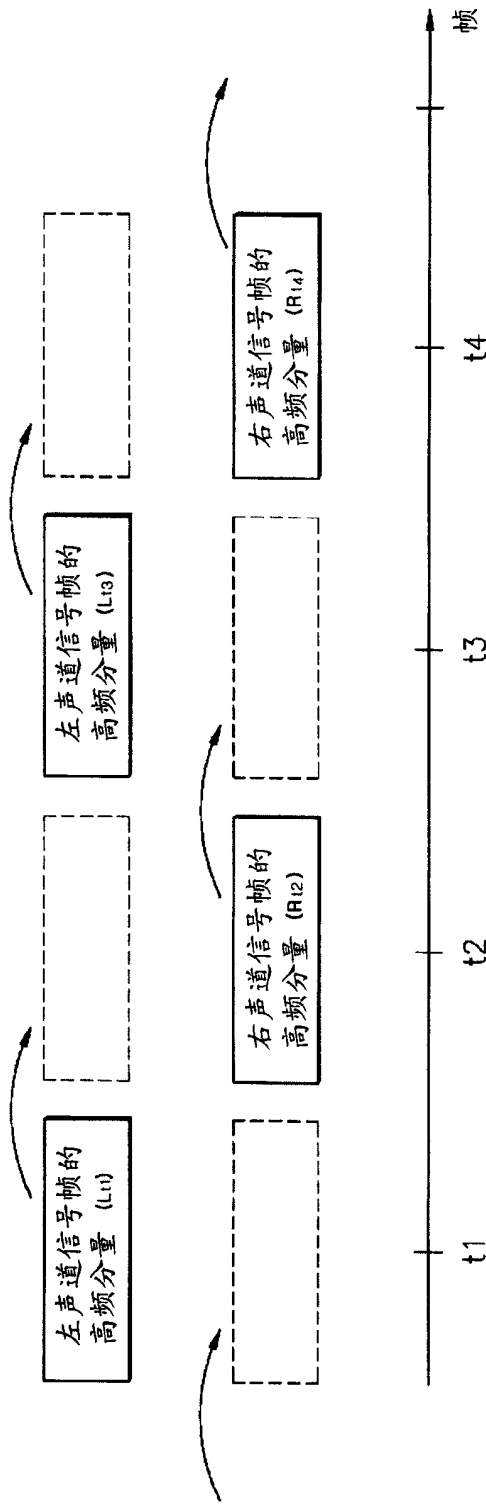


图 7

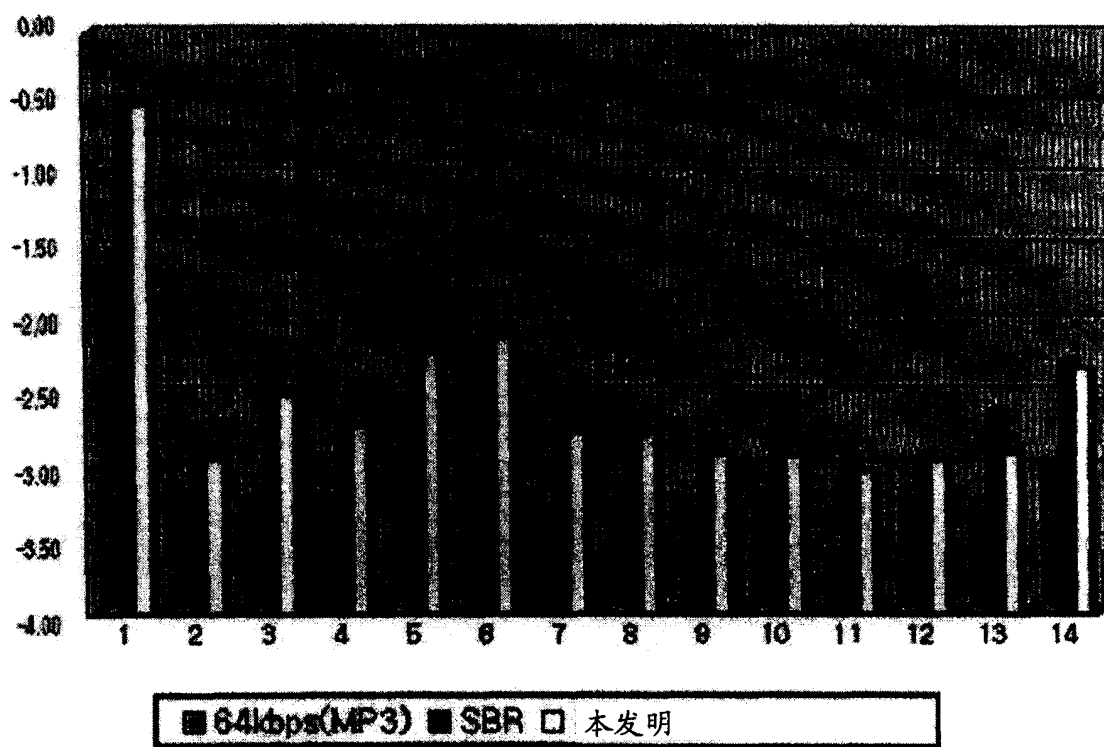


图 8