



(12) 发明专利申请

(10) 申请公布号 CN 102369573 A

(43) 申请公布日 2012. 03. 07

(21) 申请号 201080011617. 3

代理人 孙之刚 刘鹏

(22) 申请日 2010. 03. 05

(51) Int. Cl.

(30) 优先权数据

G11B 20/00 (2006. 01)

09155086. 3 2009. 03. 13 EP

G10L 19/00 (2006. 01)

(85) PCT申请进入国家阶段日

2011. 09. 13

(86) PCT申请的申请数据

PCT/IB2010/050955 2010. 03. 05

(87) PCT申请的公布数据

W02010/103442 EN 2010. 09. 16

(71) 申请人 皇家飞利浦电子股份有限公司

地址 荷兰艾恩德霍芬

(72) 发明人 F. M. J. 德 邦特 A. W. J. 乌门

E. G. P. 舒伊杰斯

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

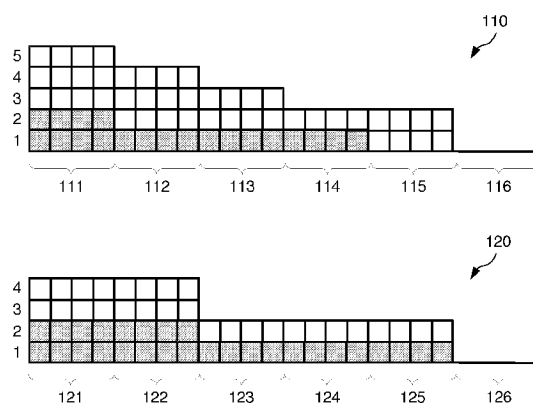
权利要求书 1 页 说明书 6 页 附图 5 页

(54) 发明名称

嵌入和提取辅助数据

(57) 摘要

本发明提出了一种用于将辅助数据嵌入到压缩音频信号中的方法。这是通过用辅助数据代替压缩音频信号的至少一个频率子带中的最低有效位(LSB)来实现的。当利用辅助数据代替压缩子带信号的LSB位时,有效地修改了子带信号,导致不同的解码的输出。代替的与辅助数据相应的LSB位作为位流的一部分而输送,并且可以在解码器处容易地提取。通过这种方式,解码器获得可以用于解码器处的更高级的音频复现的辅助数据。尽管代替了频率子带的LSB位,压缩音频本身仍然维持良好的音频质量,因为LSB位对可听假象不产生贡献。



1. 一种用于将辅助数据(202)嵌入到压缩音频信号(201)中的方法,特征在于,用辅助数据代替压缩音频信号的至少一个频率子带(111,112,113, ……)中的 LSB 位。

2. 依照权利要求 1 的方法,其中要由辅助数据(202)代替的 LSB 位基于心理声学准则而确定。

3. 依照权利要求 1 的方法,其中由辅助数据(202)代替的 LSB 位的分配由嵌入到 LSB 位中的指示信息指示。

4. 依照权利要求 1 的方法,其中使用子带编码的编码获得压缩音频信号(201)。

5. 依照权利要求 1 的方法,其中辅助数据(202)包括要用于处理解码的压缩音频信号的数据。

6. 依照权利要求 1 的方法,其中辅助数据包括 MPEG 环绕数据。

7. 一种用于将辅助数据(202)嵌入到压缩音频信号(201)中的嵌入设备(200),特征在于,该嵌入设备包括代替电路(220),该代替电路用于产生其中压缩音频信号的至少一个频率子带中的 LSB 位被辅助数据代替的输出压缩音频信号。

8. 一种用于从输入压缩音频信号(304)中提取辅助数据(302)的方法,特征在于,从输入压缩音频信号的至少一个频率子带的 LSB 位中提取辅助数据。

9. 依照权利要求 8 的方法,其中 LSB 位中辅助数据(302)的分配由嵌入到 LSB 位中的指示信息指示。

10. 依照权利要求 8 的方法,其中辅助数据(302)包括要用于处理解码的压缩音频信号的数据。

11. 依照权利要求 10 的方法,其中辅助数据(302)包括 MPEG 环绕数据。

12. 一种用于从输入压缩音频信号(304)中提取辅助数据(302)的提取设备(300),特征在于,该提取设备包括提取电路(320),该提取电路用于从输入压缩音频信号的至少一个频率子带的 LSB 位中提取辅助数据。

13. 一种用于解码输入压缩音频信号(304)的解码器(700),该解码器(700)包括:

- 依照权利要求 12 的用于提取辅助数据的提取设备(300);
- 用于解码输入压缩音频信号的第一解码器(400);以及
- 用于组合第一解码器的输出信号和辅助数据的处理电路(500)。

14. 依照权利要求 13 的解码器(700),其中处理电路(500)包括第二解码器,该第二解码器用于将第一解码器的输出信号和辅助数据解码成多通道音频信号与双耳音频信号之一。

嵌入和提取辅助数据

技术领域

[0001] 本发明涉及嵌入辅助数据。本发明还涉及提取辅助数据。

背景技术

[0002] 如 ISO/IEC 23003-1:2007, MPEG 环绕中所规定的 MPEG 环绕是一种利用空间图像的参数表示的多通道音频编码方案。由于其高的编码效率, MPEG 环绕可以用来以后向兼容的方式将单声道 / 立体声编码器朝多通道扩展, 这只需低的附加位率。MPEG 环绕数据可以作为单独的流存储或发送或者嵌入到缩混数据的辅助数据部分中。为了将 MPEG 环绕数据作为核心编码器位流的部分而传输, 核心编码器需要支持辅助数据嵌入。然而, 存在通过蓝牙 A2DP 的高质量音频流送所必须的诸如例如子带编码 (SBC) 之类的许多缩混编码器, 其并没有将辅助数据存储到位流中的能力。第 7.3 节中的 MPEG 环绕规范指出了称为“隐藏数据”的技术如何可以用来在位流中传输 MPEG 环绕数据。然而, 该技术只能应用于编码为 PCM 的缩混。该技术基于以下假设: 位流中的位在 PCM 数据与 MPEG 环绕数据之间共享。对于 MPEG 环绕数据的较高的位分配导致较低的音频质量, 因为较少的位可用于编码音频信号。“隐藏数据”技术具有以下缺点: 它不能用于压缩音频信号。

发明内容

[0003] 本发明的目的是提供将辅助数据嵌入到压缩音频信号中, 以及从压缩音频信号中提取辅助数据。本发明由独立权利要求限定。从属权利要求限定了有利的实施例。

[0004] 本发明的一个方面提出了一种用于将辅助数据嵌入到压缩音频信号中的方法。这是通过用辅助数据代替压缩音频信号的至少一个频率子带中的最低有效位 (LSB) 来实现的。

[0005] 当利用辅助数据代替压缩子带信号的 LSB 位时, 有效地修改了子带信号, 导致不同的解码的输出。代替的与辅助数据相应的 LSB 位作为位流的一部分而输送, 并且可以在解码器处容易地提取。通过这种方式, 解码器获得可以用于解码器处的更高级的音频复现的辅助数据。尽管代替了频率子带的 LSB 位, 压缩音频本身仍然维持良好的音频质量, 因为 LSB 位对潜在可听假象贡献最小。

[0006] 在一个实施例中, 要由辅助数据代替的 LSB 位基于心理声学准则而确定。作为 LSB 修改的结果的输出差异造成的主观影响通过应用心理声学准则控制可以修改的 LSB 位的位置和数量而最小化。于是, 尽管代替了频率子带的 LSB 位, 压缩音频本身仍然维持良好的音频质量, 因为那些选择的 LSB 位并不对可听假象产生贡献。LSB 位的分配通过采用编码器中使用的相同准则而在解码器中隐式地确定。解码器侧的 LSB 位分配的相似性可以事先在编码器处进行估计。因此, 无需用于 LSB 位分配的附加指示信息, 或者在编码器处使用的分配和解码器处的期望分配之间存在差异的情况下只需有限的附加指示信息以便指示这些差异。

[0007] 在另一实施例中, 要由辅助数据代替的 LSB 位的分配由嵌入到 LSB 位中的指示信

息指示。在解码器侧,需要指示信息以识别构成辅助数据的 LSB 位的位置和数量。缺省分配给特定子带的固定数量的 LSB 位用来输送该指示信息。这些位被分配用于每帧。

[0008] 在另一实施例中,使用 SBC 编码获得压缩音频信号。SBC 编码对于辅助数据没有内在的支持。可以修改 SBC 编码以接受要在一个或多个子带信号的 LSB 位中输送的辅助数据。换言之,利用辅助数据代替 LSB 位变成音频压缩的一部分。通过这种方式,SBC 编码器可以创建拥有辅助数据的位流。LSB 位分配可以随时间变化以便高效地使用频率子带,使得分配的 LSB 位不对潜在的可听假象产生贡献。可替换地,利用辅助数据代替 LSB 位可以在编码之后作为后处理步骤而执行。应当清楚的是,得到的 SBC 位流与现有的 SBC 解码器兼容。

[0009] 在另一优选的实施例中,辅助数据包括要用于处理解码的压缩音频信号的数据。这允许改变音频信号的特性的附加处理,诸如解码的压缩音频信号的后处理,比如参数控制虚拟化处理。

[0010] 在另一实施例中,辅助数据包括 MPEG 环绕数据。

[0011] MPEG 环绕缩混使用例如 SBC 编码器来编码。MPEG 环绕数据也输入到 SBC 编码器并且在 SBC 编码的缩混信号的一个或多个子带信号的 LSB 位中输送。在传输和 / 或存储得到的位流之后,SBC 解码器解码立体缩混并且提取 MPEG 环绕数据。MPEG 环绕解码器将立体缩混和 MPEG 环绕数据组合成多通道音频信号。

[0012] 本发明的另一个方面提供了一种用于从输入压缩音频信号中提取辅助数据的方法。应当理解的是,上面描述的特征、优点、评论等等同样适用于本发明的这个方面。

[0013] 本发明进一步提供了一种嵌入设备和提取设备以及包括依照本发明的提取设备的解码器。

[0014] 本发明的这些和其他方面、特征和优点根据下面描述的实施例将是清楚明白的并且将参照所述实施例进行阐述。

附图说明

[0015] 图 1 示出了依照本发明的用于将辅助数据嵌入到压缩音频信号中的方法的实施例的流程图;

图 2 示出了用辅助数据代替压缩音频的至少一个频率子带中的 LSB 位的实例;

图 3 示出了用于将辅助数据嵌入到压缩音频信号中的方法的实施例的流程图,该压缩音频信号被修改以便通过嵌入到 LSB 位中的指示信息指示要由辅助数据代替的 LSB 位的分配;

图 4 示意性地示出了依照本发明的用于将辅助数据嵌入到压缩音频信号中的嵌入设备的实例;

图 5 示意性地示出了用于从输入压缩音频信号中提取辅助数据的提取设备的实例;

图 6 示出了用于解码输入压缩音频信号的解码器的实例,其包括依照本发明的提取设备。

具体实施方式

[0016] 图 1 示出了依照本发明的用于将辅助数据嵌入到压缩音频信号中的方法的实施

例的流程图。该方法包括用辅助数据代替压缩音频的至少一个频率子带中的 LSB 位的步骤 101。压缩音频信号可以由 SBC、AAC、MP3 或 HE-AAC 编码器获得。压缩音频信号包括至少一个频率子带。在这里,频率子带指的是例如 SBC 提供的滤波器组子带表示以及例如 AAC 提供的变换表示。通常,来自子带滤波器的子带称为子信号,而来自变换的子带称为频率系数。应当指出的是,两种情况下的 LSB 位都指量化频谱系数的位。辅助数据可以是任何类型。然而,优选地它应当包括与可以用来提高压缩音频的空间音频质量的空间音频信息有关的数据。这种辅助数据的一个实例是例如格式化成与 ISO/IEC 23003-1:2007, MPEG 环绕的第 7.3.2 节中规定的类似的数据结构的 MPEG 环绕数据。可替换地,辅助数据可以包括例如谱带复制数据、参数立体声数据、诸如定时信息或响度级之类的元数据或者允许解码器侧的交互混合的空间音频对象编码数据。

[0017] 图 2 示出了用辅助数据代替压缩音频的至少一个频率子带中的 LSB 位的实例。在图 2 中,绘出了压缩音频信号的一个实例。这样的压缩音频信号可以由具有以下配置参数的 SBC 编码器获得:采样频率 48kHz,立体声通道模式,8 个子带以及块长 4。图 110 与左通道音频相应,而图 120 与右通道音频相应。对于每个通道,分别绘出了用于左通道和右通道的六个子带 111-116 和 121-126。出于表示清楚的原因,仅仅绘出了六个子带(而不是规定的八个子带),因为在当前实例中没有位分配给其余的子带。用于左通道音频 110 的第一子带 111 的压缩音频信号需要规定的 4 位的块长和 5 位的块宽,导致 20 个位。应当指出的是,块长与子带中的子带样本数相应。子带 112 需要规定的 4 位的块长和 4 位的块宽,导致 16 个位。同时,对于子带 113、114 和 115,分别需要 12 位、8 位和 8 位。类似地,对于右音频通道 120,子带 121、122、123、124、125 分别需要 16 位、16 位、8 位、8 位和 8 位。如本发明所规定的,一些子带的 LSB 位可以用于嵌入辅助数据。这些位在图 2 中标记为灰色。因此,子带 111 中的 8 个 LSB 位、子带 112 中的 4 个 LSB 位、子带 113 中的 4 个 LSB 位和子带 114 中的 4 个 LSB 位用于嵌入辅助数据。嵌入辅助数据在这里表示利用辅助数据代替指明的 LSB 位。尽管要利用辅助数据代替的 LSB 位的分配在子带中是变化的,但是同样可能的是使用固定的 LSB 位分配。变化的 LSB 位分配的优点在于,该位分配可以适应压缩音频中的实际音频内容以便不损害音频质量。通过在频率子带上变化 LSB 位分配,可以控制子带内的代替的 LSB 位产生的失真。LSB 位分配的控制允许在频谱域中对失真定形,使得失真保持被掩盖。

[0018] 在一个实施例中,要由辅助数据代替的 LSB 位基于心理声学准则来确定。该心理声学准则具有以下目标:选择用于利用辅助数据代替的子带和 LSB 位,对于该辅助数据,预期对于感知的影响最小。该心理声学准则可以例如通过在子带表示的网格上确定原始音频信号的掩盖曲线来实现。这样的掩盖曲线指示可以在每个频带中添加多少噪声。其中可以添加大多数噪声的频带例如被选择用于嵌入辅助数据。可替换地,该准则可以进一步通过将使用例如 SBC 编码而编码的压缩音频信号的失真与确定的掩盖曲线进行比较而改进。因此,要由辅助数据代替的 LSB 位可以被选择成使得与掩盖曲线相比,总体失真(包括 SBC 编码的量化以及将辅助数据嵌入到子带的 LSB 位中)在所有子带上近似相等。将 SBC 编码与辅助数据嵌入相组合是有利的,因为它允许最小化辅助数据嵌入对感知音频质量的影响。如果压缩音频信号是预编码的信号,例如 SBC 位流,那么较高的频率已经被粗量化,从而给嵌入辅助数据留下很少的空间。然而,如果将辅助数据的嵌入与使用例如 SBC 编码的音频信号的压缩相组合,那么存在用于嵌入辅助数据的空间,其优选地由编码和嵌入参数控制。

[0019] 图 3 示出了用于将辅助数据嵌入到压缩音频信号中的方法的实施例的流程图,该压缩音频信号被修改以便通过嵌入到 LSB 位中的指示信息指示要由辅助数据代替的 LSB 位的分配。该方法包括用辅助数据代替压缩音频的至少一个频率子带中的 LSB 位的步骤 101。步骤 102 包括嵌入指示信息以指示压缩音频信号中要由辅助数据代替的 LSB 位的分配。该指示信息类似于嵌入到压缩音频信号的 LSB 位中的辅助数据。尽管步骤 102 在步骤 101 之后,但是这两个步骤的顺序可以互换。

[0020] 指示信息可以例如以预定数量(例如 16 位)包含在帧中第一子带的 LSB 位的预定固定位置处。可替换地,可以采用 ISO/IEC 23003-1:2007, MPEG 环绕的第 7.3.2 节中描述的方法以指示包括具有嵌入的辅助数据的压缩音频信号的位流中的指示信息。

[0021] 在另一实施例中,使用 SBC 编码获得压缩音频。SBC 编码提供了相对较高的位率的可能性,从而允许更多空间用于嵌入辅助数据。此外,对于 SBC 编码而言,需要较少的注意以确保不出现可听假象(例如可以使用简化的心理声学模型)。SBC 作为各种通信设备(例如电话或者汽车无线电)之间的通信编解码器也变得越来越流行。

[0022] 然而,紧跟在 SBC 编码之后,可以使用任何其他的变换或子带编码。尤其是不支持辅助数据的属于该类别的编码技术可以受益于依照本发明的辅助数据的嵌入。

[0023] 在另一实施例中,辅助数据包括要用于处理解码的压缩音频信号的数据。如前面所指出的,辅助数据优选地应当包括与可以用来提高压缩音频的空间音频质量的空间音频信息有关的数据。这种辅助数据的一个实例是例如格式化成与 ISO/IEC 23003-1:2007, MPEG 环绕的第 7.3.2 节中规定的类似的数据结构的 MPEG 环绕数据。相同规范的第 6 节描述了如何采用 MPEG 环绕数据以便从单声道或立体声缩混信号和 MPEG 环绕数据创建多通道或双耳音频信号。

[0024] 在将包括 MPEG 环绕数据的辅助数据嵌入到包括 SBC 编码音频 PCM 样本的压缩音频信号中的情况下,需要多个 SBC 帧以用于嵌入包含在一个 MPEG 环绕帧中的 MPEG 环绕数据。假设除了块长现在为 16 之外,如针对图 2 所描述的使用了 SBC 配置。这导致 SBC 帧长 $8 \times 16 (=128)$ 子带样本,其中 8 为子带数,并且 16 为块长。MPEG 环绕数据的帧长为 1024 个 PCM 样本,其与 SBC 帧的 1024 个子带样本相应。假设依照 MPEG 环绕标准编码的 1024 个 PCM 帧导致 888 位。此外,假设需要 72 位以用于编码指示信息。因此,需要 8 个 SBC 帧以容纳辅助数据的 888 位和指示信息的 72 位。为了高效地使用可用的位,将 8 个 SBC 帧分组成 4 个具有 2 个 SBC 帧的组。对于每个具有 2 个帧的组,使用指示信息。因此,对于二通道和用于每个通道的 4 组,使用总共 8 个单位的指示信息。对于具有比指示信息中指示的数量更少的可用于子带样本的位的子带而言,这两个值中的最小值用于辅助数据在子带中的实际嵌入。假设如图 2 中绘出的子带样本用于每个通道的 8 个 SBC 帧。此外,假设使用分配 2、1、0 和 1 位用于左通道,并且使用分配 1、0、1 和 0 用于右通道。分配 2 位用于左通道意味着对于具有 2 个 SBC 帧的第一组而言,将每子带 2 位分配给辅助数据。这导致 $2 \text{ (对于 2 个 SBC 帧)} \times 5 \text{ (对于 5 个子带)} \times 16 \text{ (对于块长)} \times 2 \text{ (对于每个子带中分配的位)} = 320$ 位可用于辅助数据。随后,每通道分配 1 位导致 160 位可用于辅助数据。

[0025] 这继而对于左通道的 2、1、0、1 位分配以及右通道的 1、0、1、0 位分配而言,导致总共 960 位,这些位足以容纳实际所需的 888 位的辅助数据。

[0026] 图 4 示意性地示出了依照本发明的用于将辅助数据 202 嵌入到压缩音频信号 201

中的嵌入设备 200 的实例。嵌入设备 200 包括分配电路 210, 该分配电路用于基于提供给电路 210 的心理声学准则 203 确定利用辅助数据代替的 LSB 位分配。这种准则 203 的一个实例是所有子带上关于掩盖阈值的嵌入数据的能量最小化。嵌入设备 200 进一步包括代替电路 220, 该代替电路利用辅助数据 202 代替压缩音频信号 201 中的由分配电路 210 分配的 LSB 位, 导致输出压缩音频信号 204。

[0027] 应当清楚的是, 当 LSB 位分配固定时, 分配电路 210 是多余的, 并且无需包含在嵌入设备 200 中。然而, 在这种情况下, 该固定 LSB 位分配应当传送到解码器侧以便允许在解码器侧从压缩音频信号 204 中正确地提取辅助数据 202。

[0028] 本发明的另一方面是一种用于从输入压缩音频信号中提取辅助数据的方法, 该方法的特征在于, 从输入压缩音频的至少一个频率子带的 LSB 位中提取辅助数据。基本上, 该提取方法是嵌入方法的逆方法。基于固定的或适应性的向辅助数据的 LSB 位分配, 从已经依照本发明嵌入了辅助数据的输入压缩音频中检测并且提取辅助数据。

[0029] 用于将辅助数据嵌入到压缩音频信号中的方法的优选实施例同样适用于从输入压缩音频信号中提取辅助数据的方法。

[0030] 图 5 示意性地示出了用于从输入压缩音频信号 304 中提取辅助数据 302 的提取设备 300 的实例。输入压缩音频信号 304 与压缩音频信号 204 相应, 该压缩音频信号 204 被修改成具有嵌入到压缩音频信号 201 的至少一个频率子带中的 LSB 位中的辅助数据 202。提取设备 300 包括分配-提取电路 310, 该分配-提取电路用于提取 LSB 位到辅助数据 302 的分配。分配-提取电路 310 确定的分配馈送到提取电路 320 中, 该提取电路基于该分配从输入压缩音频信号 304 中提取辅助数据 302。

[0031] 应当清楚的是, 当 LSB 位分配固定时, 分配-提取电路 310 是多余的并且无需包含在提取设备 300 中。然而, 在这种情况下, 该固定 LSB 位分配应当传送到提取设备侧以便允许从输入压缩音频信号 304 中正确地提取辅助数据 302。

[0032] 图 6 示出了用于解码输入压缩音频信号 304 的解码器 700 的实例, 其包括依照本发明的提取设备。解码器 700 包括用于提取辅助数据的提取设备 300。此外, 解码器 700 包括用于解码输入压缩音频信号的第一解码器 400 以及用于组合第一解码器 400 的输出信号 301 和辅助数据 302 的处理电路 500。特别地, 处理电路 500 可以包括第二解码器, 其将第一解码器 400 的输出信号 301 和辅助数据 302 解码成多通道音频信号、双耳音频信号或者任何其他适当的音频信号。第一解码器 400 的一个实例是 SBC 解码器。第二解码器 500 的一个实例是 MPEG 环绕解码器。第二解码器接收单声道或立体声信号 301 以及 MPEG 环绕数据 302。然后, 它将单声道或立体声信号 301 再现为多通道信号 620 或者如 MPEG 环绕数据所规定的双耳音频信号 610。优选地, 在作为辅助数据嵌入到压缩音频信号中之前, MPEG 环绕数据被随机化。ISO/IEC 23003-1:2007, MPEG 环绕的第 7.3.4.2 节中规定了 MPEG 环绕数据的随机化。

[0033] 本发明也可以应用于译码, 例如从其中使用所谓的辅助数据通道将 MPEG 环绕数据嵌入到位流中的 HE-AAC/MPEG 环绕译码成其中使用本发明嵌入 MPEG 环绕数据的 SBC/MPEG 环绕。

[0034] 尽管已经结合一些实施例描述了本发明, 但是本发明并不预限于本文阐述的特定形式。相反地, 本发明的范围仅由所附权利要求书限制。此外, 虽然特征可能看起来结合

特定实施例而被描述,但是本领域技术人员将认识到,依照本发明可以组合所描述的实施例的各种特征。在权利要求书中,措词“包括 / 包含”并没有排除其他元件或步骤的存在。

[0035] 此外,尽管单独地被列出,但是多个电路、元件或方法步骤可以由例如单个单元或处理器实现。此外,尽管单独的特征可以包含于不同的权利要求中,但是这些特征可能地可以有利的加以组合,并且包含于不同的权利要求中并不意味着特征的组合不可行和 / 或不是有利的。此外,特征包含于一种权利要求类别中并不意味着限于该类别,而是表示该特征同样可适当地应用于其他权利要求类别。此外,单数引用并没有排除复数。因此,对于“一”、“一个”、“第一”、“第二”等等的引用并没有排除复数。权利要求中的附图标记仅仅作为澄清的实例而被提供,不应当以任何方式被解释为限制了权利要求的范围。本发明可以通过包含若干不同元件的硬件电路以及通过经过适当编程的计算机或其他可编程设备的电路实现。

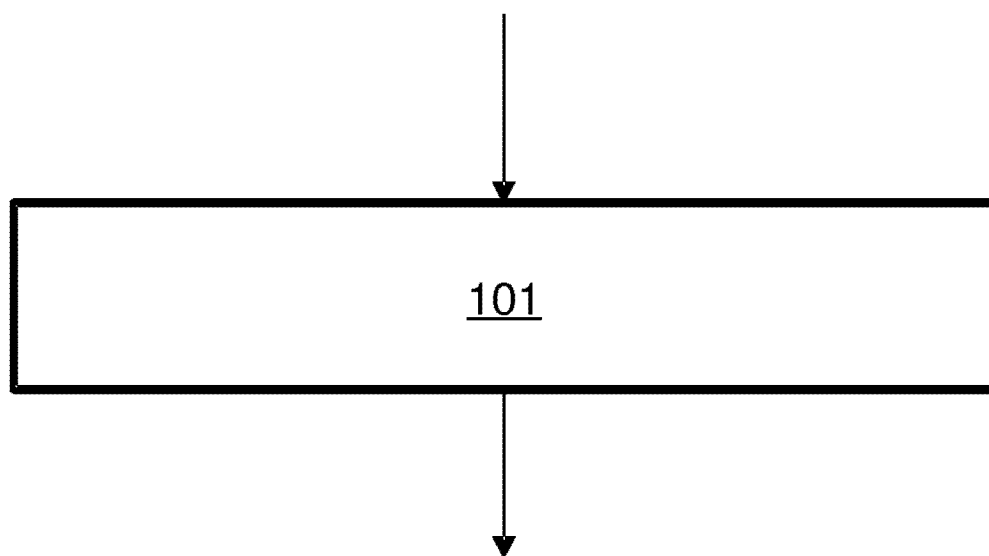


图 1

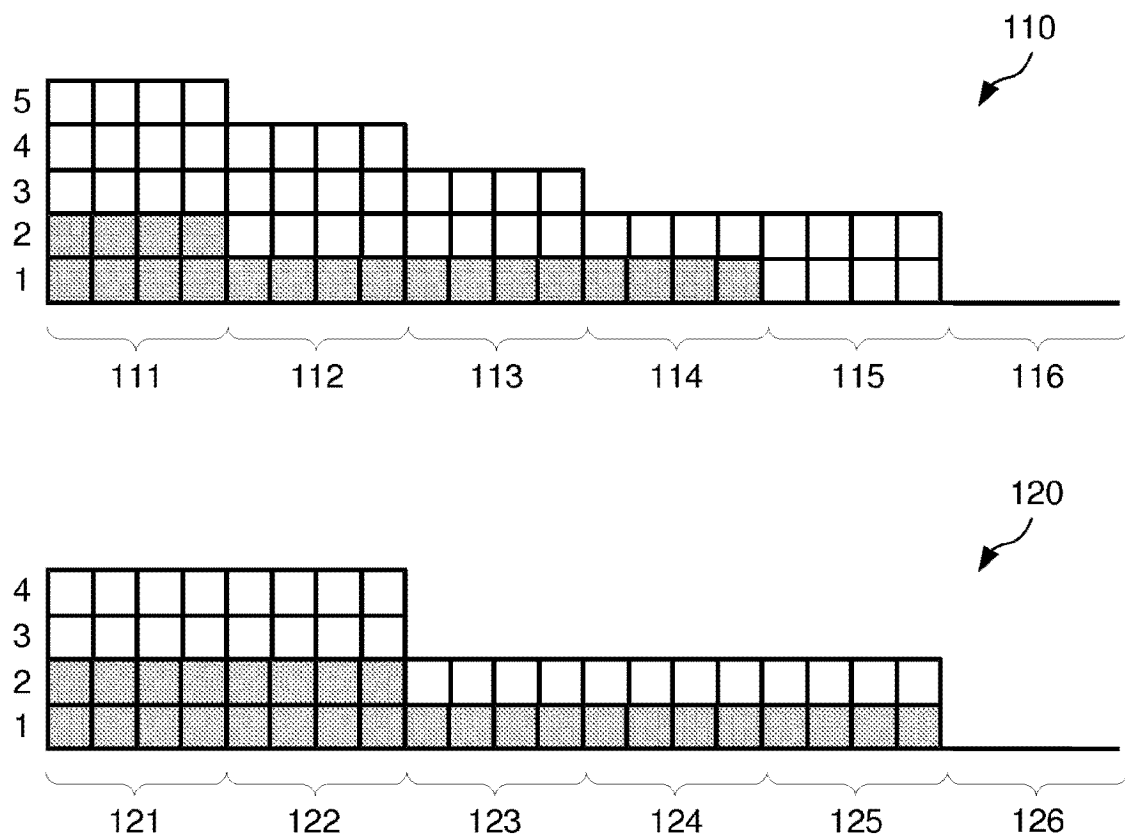


图 2

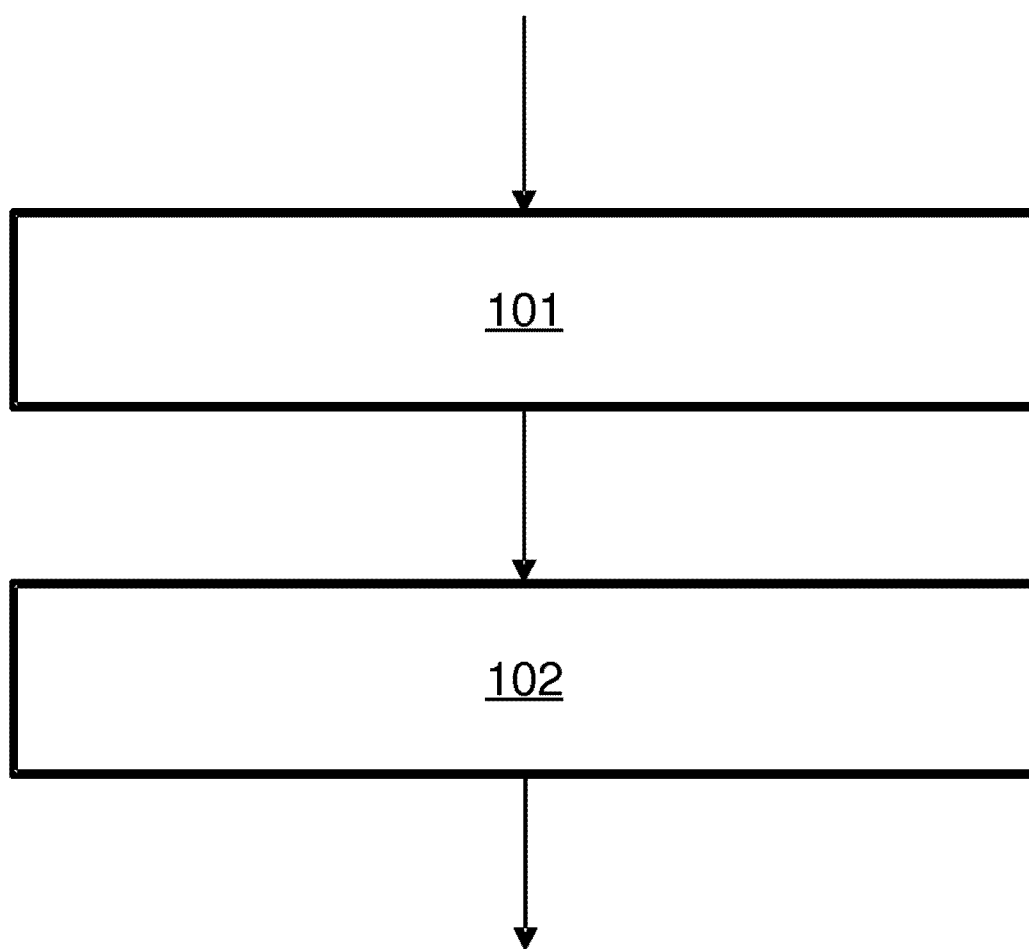


图 3

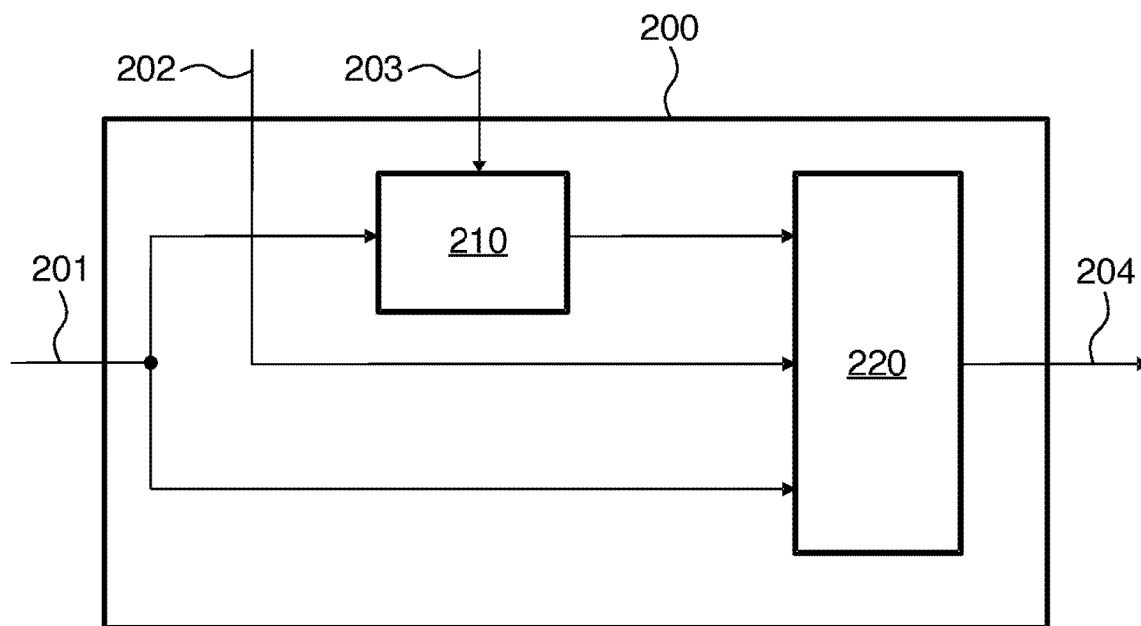


图 4

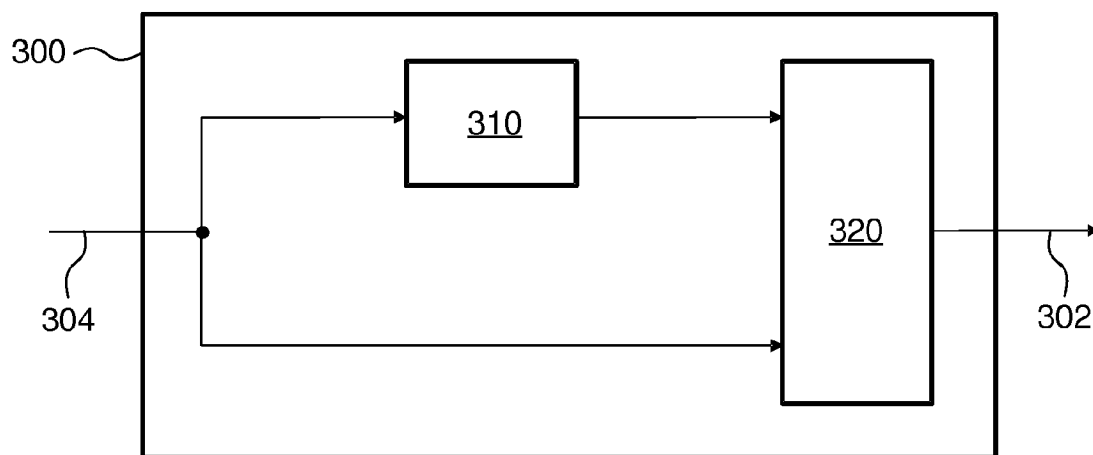


图 5

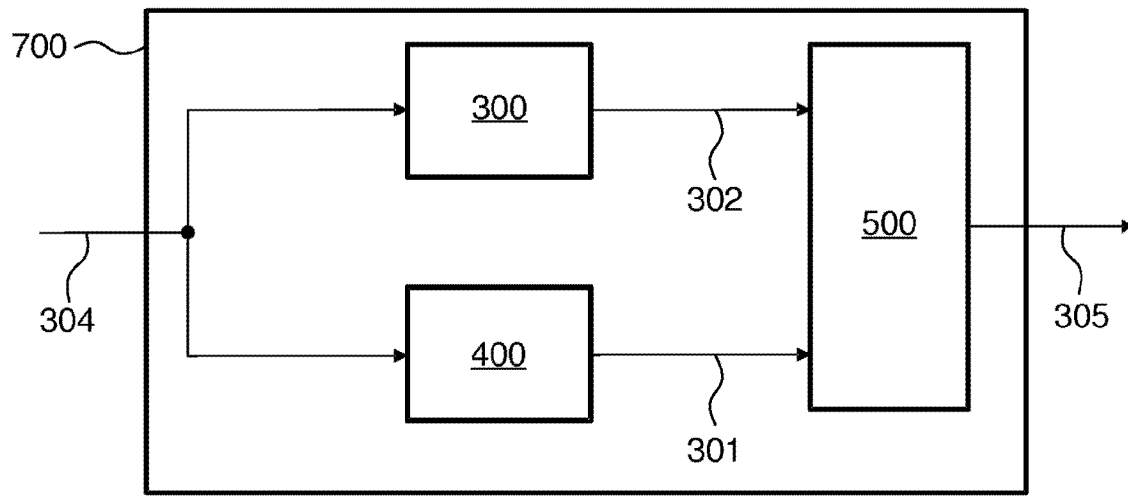


图 6