



(12)发明专利

(10)授权公告号 CN 105684085 B

(45)授权公告日 2019.11.05

(21)申请号 201480056387.0

(74)专利代理机构 北京集佳知识产权代理有限公司 11227

(22)申请日 2014.10.10

代理人 唐京桥 陈炜

(65)同一申请的已公布的文献号

申请公布号 CN 105684085 A

(51)Int.Cl.

G11B 27/10(2006.01)

(43)申请公布日 2016.06.15

H04N 5/04(2006.01)

(30)优先权数据

H04N 5/60(2006.01)

2013-218267 2013.10.21 JP

H04N 21/236(2011.01)

2013-257508 2013.12.12 JP

H04N 21/41(2011.01)

2014-000831 2014.01.07 JP

H04N 21/434(2011.01)

2014-115406 2014.06.04 JP

H04N 21/439(2011.01)

(85)PCT国际申请进入国家阶段日

H04N 21/8547(2011.01)

2016.04.13

(56)对比文件

(86)PCT国际申请的申请数据

US 2013135526 A1, 2013.05.30,

PCT/JP2014/005186 2014.10.10

CN 102177726 A, 2011.09.07,

(87)PCT国际申请的公布数据

US 7907212 B2, 2011.03.15,

W02015/059891 EN 2015.04.30

CN 101189661 A, 2008.05.28,

JP 2012244413 A, 2012.12.10,

(73)专利权人 索尼公司

审查员 郭晓冰

地址 日本东京都

(72)发明人 井上晃 高桥秀介 西口正之

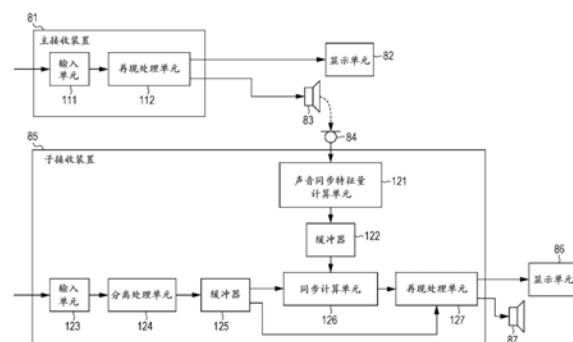
权利要求书3页 说明书35页 附图29页

(54)发明名称

信息处理设备、方法及程序

(57)摘要

一种信息处理设备,用于以与不同于该信息处理设备的第二信息处理设备进行的第一内容的再现同步的方式来再现第二内容,第一内容包括音频内容。该信息处理设备包括电路,该电路被配置成:从音频内容中提取第一特征;获得音频内容的第二特征,第二特征与第二内容一起;比较第一特征与第二特征;以及基于比较的结果生成用于以与第一内容同步的方式再现第二内容的同步信息。



1. 一种信息处理设备, 用于以与由不同于所述信息处理设备的第二信息处理设备进行的第一内容的再现同步的方式来再现第二内容, 所述第一内容包括音频内容, 所述信息处理设备包括:

电路, 被配置成:

从所述音频内容中提取第一特征;

获得所述第二内容和从所述第二内容中提取的第二特征;

比较所述第一特征与所述第二特征; 以及

基于所述比较的结果生成用于以与所述第一内容同步的方式来再现所述第二内容的同步信息,

其中, 所述电路被配置成在对所述第一特征和所述第二特征中的至少一个执行帧率转换使得所述第一特征的帧率与所述第二特征的帧率一致之后比较所述第一特征和所述第二特征。

2. 根据权利要求1所述的信息处理设备, 其中, 所述电路被配置成:

从经由通过所述第二信息处理设备对所述音频内容的再现生成的声音中获取音频信号; 以及

从所述音频信号中提取所述第一特征。

3. 根据权利要求1所述的信息处理设备, 其中, 所述第二内容和所述第二特征在基于所述第一内容的到达时间与所述第二内容和所述第二特征的到达时间之间的差而确定的时间处被传输至所述信息处理设备。

4. 根据权利要求1所述的信息处理设备, 其中, 所述电路被配置成:

请求所述第二内容和所述第二特征的传输; 以及

接收响应于所述请求而被传输的所述第二内容和所述第二特征。

5. 根据权利要求4所述的信息处理设备, 其中, 所述电路被配置成:

请求所述第一内容的传输; 以及

接收响应于针对所述第一内容的传输的请求而被传输的所述第一内容。

6. 根据权利要求1所述的信息处理设备, 其中, 所述电路被配置成:

从所述音频内容中提取第一多个特征, 所述第一多个特征包括所述第一特征;

从所述第二内容中提取第二多个特征, 所述第二多个特征包括所述第二特征;

比较所述第一多个特征与所述第二多个特征; 以及

基于比较所述第一多个特征与所述第二多个特征的结果来生成用于以与所述第一内容同步的方式来再现所述第二内容的所述同步信息。

7. 根据权利要求1所述的信息处理设备, 其中, 所述电路被配置成控制所述第二内容的再现。

8. 根据权利要求7所述的信息处理设备, 其中, 所述电路被配置成基于所述同步信息来校正所述第二内容的再现位置。

9. 根据权利要求1所述的信息处理设备, 其中, 所述电路被配置成:

获取与所述第一内容关联的呈现时间信息; 以及

基于比较与所述第一内容关联的呈现时间信息和与所述第二内容关联的呈现时间信息的结果来获得时间范围,

其中,所述电路被配置成至少部分地通过在所述时间范围中比较所述第二特征与从所述第一内容中出现的所述第一内容的一个或多个部分中提取的一个或多个特征来生成所述同步信息。

10.一种信息处理方法,用于以与第一内容的再现同步的方式来再现第二内容,所述第一内容包括音频内容,所述信息处理方法包括:

从所述音频内容中提取第一特征;

获得所述第二内容和从所述第二内容中提取的第二特征;

比较所述第一特征与所述第二特征;以及

基于所述比较的结果生成用于以与所述第一内容同步的方式再现所述第二内容的同步信息,

其中,在对所述第一特征和所述第二特征中的至少一个执行帧率转换使得所述第一特征的帧率与所述第二特征的帧率一致之后比较所述第一特征和所述第二特征。

11.至少一种非暂态计算机可读存储介质,其存储有处理器可执行指令,当所述处理器可执行指令由至少一个计算机执行时,使所述至少一个计算机执行用于以与第一内容的再现同步的方式来再现第二内容的信息处理方法,所述第一内容包括音频内容,所述信息处理方法包括:

从所述音频内容中提取第一特征;

获得所述第二内容和从所述第二内容中提取的第二特征;

比较所述第一特征与所述第二特征;以及

基于所述比较的结果生成用于以与所述第一内容同步的方式再现所述第二内容的同步信息,

其中,在对所述第一特征和所述第二特征中的至少一个执行帧率转换使得所述第一特征的帧率与所述第二特征的帧率一致之后比较所述第一特征和所述第二特征。

12.一种信息处理设备,包括:

电路,被配置成:

从通过第一内容的再现生成的声音中提取特征;

以与所述第一内容同步的方式将所述特征与和所述第一内容具有时间同步关系的第二内容进行关联;以及

输出所述第二内容和所述特征,

其中,所述电路被配置成:

从通过再现所述第一内容的多个部分生成的声音中提取多个特征;以及

将所提取的特征与所述第二内容相关联并且输出所提取的特征和所述第二内容,

其中,所述电路被配置成:对所述特征进行下采样以获得下采样的特征;以及输出所述第二内容和所述下采样的特征。

13.根据权利要求12所述的信息处理设备,其中,所述电路被配置成输出所述第一内容。

14.根据权利要求13所述的信息处理设备,其中,所述电路被配置成在基于所述第一内容的到达时间与所述第二内容的到达时间之间的差而确定的时间处输出所述第二内容和所述特征。

15. 根据权利要求13所述的信息处理设备,其中,所述电路被配置成响应于针对所述第二内容和所述特征的传输的请求来输出所述第二内容和所述特征。

16. 根据权利要求15所述的信息处理设备,其中,所述电路被配置成响应于针对所述第一内容的传输的请求来输出所述第一内容。

17. 一种信息处理方法,包括:

从通过第一内容的再现生成的声音中提取特征;

以与所述第一内容同步的方式将所述特征与和所述第一内容具有时间同步关系的第二内容进行关联;以及

输出所述第二内容和所述特征,

其中,所述信息处理方法包括:

从通过再现所述第一内容的多个部分生成的声音中提取多个特征;以及

将所提取的特征与所述第二内容相关联并且输出所提取的特征和所述第二内容,

其中,所述信息处理方法还包括:对所述特征进行下采样以获得下采样的特征;以及输出所述第二内容和所述下采样的特征。

18. 至少一种非暂态计算机可读存储介质,当所述至少一种非暂态计算机可读存储介质由至少一个计算机执行时,使所述至少一个计算机执行:

从通过第一内容的再现生成的声音中提取特征;

以与所述第一内容同步的方式将所述特征与和所述第一内容具有时间同步关系的第二内容进行关联;以及

输出所述第二内容和所述特征,

其中,当所述至少一种非暂态计算机可读存储介质由至少一个计算机执行时,使所述至少一个计算机执行:

从通过再现所述第一内容的多个部分生成的声音中提取多个特征;以及

将所提取的特征与所述第二内容相关联并且输出所提取的特征和所述第二内容,

其中,当所述至少一种非暂态计算机可读存储介质由至少一个计算机执行时,使所述至少一个计算机执行:对所述特征进行下采样以获得下采样的特征;以及输出所述第二内容和所述下采样的特征。

信息处理设备、方法及程序

技术领域

[0001] 本技术涉及信息处理设备、方法及程序,具体地涉及能够对通过不同路径获取的多条内容进行同步的信息处理设备、方法及程序。

[0002] <相关申请的交叉引用>

[0003] 本申请要求于2013年10月21日提交的日本优先专利申请JP 2013-218267、于2013年12月12日提交的日本优先专利申请JP 2013-257508、于2014年1月7日提交的日本优先专利申请JP 2014-000831以及于2014年6月4日提交的日本优先专利申请JP 2014-115406的权益,以上专利申请中的每个专利申请的全部内容通过引用并入本文中。

背景技术

[0004] 近年来,可以在网络连接的前提下再现各种媒体内容的装置例如多功能移动电话以及平板终端装置已经有所增多。此外,基于使用网络功能的多个装置的组合的利用成为必需,使用网络功能的多个装置包括已经在相关技术中存在的电视接收器等。

[0005] 作为多个装置的组合,可以呈现下述应用程序:用于通过多个设备经由广播、因特网等接收具有时间同步关系的多条媒体内容例如以下示出的(A1)至(A4),以及以同步的方式再现各条内容的应用程序。

[0006] (A1) 外语声音内容、评论声音内容、隐藏字幕以及针对主视频内容和声音内容的字符信息;

[0007] (A2) 通过由每个乐器播放乐曲以及捕获每个乐器在播放的场景的图像来获取的多条视频内容和声音内容;

[0008] (A3) 通过以多个角度对一个场景进行成像来获得的视频内容和声音内容;

[0009] (A4) 主视频内容和声音内容以及其高分辨率版本的视频内容和声音内容。

[0010] 需要在再现期间维持同步的状态下来再现这样的多条内容。例如,已经公开了通过从由多个不同的成像设备同时成像的多条内容中提取特征量并且计算特征量的相似度来对相应的各条内容进行同步的技术作为同步多条内容的技术(例如,参见PTL 1)。

[0011] 引用列表

[0012] 专利文献

[0013] PTL 1:日本未审查专利申请公开第2013-174765号

发明内容

[0014] 技术问题

[0015] 顺便提及地,当试图通过多个装置经由相应不同的路径来接收如上所述的媒体内容时,在实践中由于传输延迟、传输处理与接收处理中的延迟、接收装置的操作时钟的不同等而难以在维持同步的同时再现内容。根据PTL 1中公开的技术,在待同步并且再现的各条内容没有任何相似的特征的情况下难以同步各条内容。

[0016] 期望能够对通过不同路径获取的多条内容进行同步。

[0017] 问题的解决方案

[0018] 根据本技术的第一实施方式,提供了一种信息处理设备,包括:特征量计算单元,从第一内容的声音信号中提取特征量;以及同步计算单元,通过比较在和第一内容具有时间同步关系的第二内容的同步中获取的特征量与由特征量计算单元提取的特征量,基于用于以与第一内容同步的方式再现第二内容的声音特征量来生成同步校正信息。

[0019] 特征量计算单元可以从通过收集被再现的第一内容中的声音获得的声音信号中提取特征量。

[0020] 该信息处理设备还可以包括:第一输入单元,获取第二内容和以与第二内容同步的方式与第二内容关联的特征量。

[0021] 第二内容和该特征量可以在考虑与第一内容的到达时间的到达时间的差确定的定时处被传输至该信息处理设备。

[0022] 第一输入单元可以请求第二内容和该特征量的传输,以及接收响应于该请求被传输的第二内容和特征量。

[0023] 该信息处理设备还可以包括:第二输入单元,请求第一内容的传输,以及接收根据该请求被传输的第一内容。

[0024] 特征量计算单元可以提取根据声音信号再现的一条第一内容的特征量,以及同步计算单元可以通过比较与第二内容关联的多条第一内容的相应的特征量和由特征量计算单元提取的特征量来指定被再现的第一内容,以及基于用于以彼此同步的方式再现所指定的第一内容和第二内容的声音特征量来生成同步校正信息。

[0025] 该信息处理设备还可以包括:再现处理单元,控制第二内容的再现。

[0026] 该再现处理单元可以取决于基于声音特征量的同步校正信息来校正第二内容中的再现位置。

[0027] 该信息处理设备还可以包括:获取单元,获取第一内容的呈现时间信息;以及比较单元,比较第一内容的呈现时间信息与第二内容的呈现时间信息,以及基于呈现时间信息生成同步校正信息,以及同步计算单元可以通过比较从一系列获取的特征量中由基于呈现时间信息的同步校正信息指示的范围内的特征量与由特征量计算单元提取的特征量来生成基于声音特征量的同步校正信息。

[0028] 同步计算单元可以在对所获取的特征量和由特征量计算单元提取的特征量中的至少一个执行帧率转换使得所获取的特征量的帧率与由特征量计算单元提取的特征量的帧率一致之后来比较特征量。

[0029] 根据本技术的第一实施方式,提供了一种信息处理方法及程序,包括:从第一内容的声音信号中提取特征量;以及通过比较在和第一内容具有时间同步关系的第二内容的同步中获取的特征量与在提取中提取的特征量来基于用于与第一内容同步的方式再现第二内容的声音特征量生成同步校正信息。

[0030] 在本技术的第一实施方式中,从第一内容的声音信号中提取特征量,以及通过比较在和第一内容具有时间同步关系的第二内容的同步中获取的特征量与在提取中提取的特征量来基于用于以与第一内容同步的方式再现第二内容的声音特征量生成同步校正信息。

[0031] 根据本技术的第二实施方式,提供了一种信息处理设备,包括:特征量计算单元,

从第一内容的声音信号中提取特征量；以及第一输出单元，输出与第一内容具有时间同步关系的第二内容和以与第二内容同步的方式与第二内容关联的特征量。

[0032] 该信息处理设备还包括：第二输出单元，输出第一内容。

[0033] 第一输出单元可以在考虑与第一内容的到达时间的到达时间的差确定的定时输出第二内容和特征量。

[0034] 当第二内容和特征量的传输被请求时，第一输出单元可以响应于该请求输出第二内容和特征量。

[0035] 当第一内容的传输被请求时，第二输出单元可以响应于该请求输出第一内容。

[0036] 特征量计算单元可以从声音信号中提取多条第一内容的特征量，以及第一输出单元可以关联多条第一内容的特征量与第二内容的特征量以及输出多条第一内容的特征量和第二内容。

[0037] 特征量计算单元可以对特征量进行下采样，以及第一输出单元可以输出第二内容和所下采样的特征量。

[0038] 根据本技术的第二实施方式，提供了一种信息处理方法及程序，包括：从第一内容的声音信号中提取特征量；以及输出与第一内容具有时间同步关系的第二内容和以与第二内容同步的方式与第二内容关联的特征量。

[0039] 在本技术的第二实施方式中，从第一内容的声音信号中提取特征量，以及输出与第一内容具有时间同步关系的第二内容和以与第二内容同步的方式与第二内容关联的特征量。

[0040] 发明的有益效果

[0041] 根据本技术的第一实施方式和第二实施方式，可以对通过不同路径获取的多条内容进行同步。

[0042] 此外，本技术不必限制为上述效果，并且可以获得在本技术中描述的任何效果。

附图说明

[0043] [图1]图1是示出了提供设备的配置示例的图。

[0044] [图2]图2是示出了声音同步特征量计算单元的配置示例的图。

[0045] [图3]图3是示出了声音同步特征量的下采样的图。

[0046] [图4]图4是示出了内容再现系统的配置示例的图。

[0047] [图5]图5是示出了声音同步特征量计算单元的配置示例的图。

[0048] [图6]图6是示出了同步计算单元的配置示例的图。

[0049] [图7]图7是示出了声音同步特征量的同步计算的图。

[0050] [图8]图8是示出了声音同步特征量的同步计算的图。

[0051] [图9]图9是示出了声音同步特征量的同步计算的图。

[0052] [图10]图10是示出了作为相似度计算的目标的块的图。

[0053] [图11]图11是示出了相似度计算的图。

[0054] [图12]图12是示出了传输处理的流程图。

[0055] [图13]图13是示出了子通道信号和声音同步特征量的多路复用的图。

[0056] [图14]图14是示出了声音同步特征量计算处理的流程图。

- [0057] [图15]图15是示出了主内容再现处理的流程图。
- [0058] [图16]图16是示出了子内容再现处理的流程图。
- [0059] [图17]图17是示出了声音同步特征量计算处理的流程图。
- [0060] [图18]图18是示出了同步校正信息生成处理的流程图。
- [0061] [图19]图19是示出了本技术的应用示例的图。
- [0062] [图20]图20是示出了本技术的应用示例的图。
- [0063] [图21]图21是示出了本技术的应用示例的图。
- [0064] [图22]图22是示出了提供设备的配置示例的图。
- [0065] [图23]图23是示出了内容再现系统的配置示例的图。
- [0066] [图24]图24是示出了传输处理的流程图。
- [0067] [图25]图25是示出了主内容再现处理的流程图。
- [0068] [图26]图26是示出了子内容再现处理的流程图。
- [0069] [图27]图27是示出了同步校正信息生成处理的流程图。
- [0070] [图28]图28是示出了作为相似度计算的目标的块的图。
- [0071] [图29]图29是示出了计算机的配置示例的图。

具体实施方式

- [0072] 在下文中,将参照附图给出应用本技术的实施方式的描述。
- [0073] <第一实施方式>
- [0074] <本技术的特征>
- [0075] 首先,将给出本技术的特征的描述。
- [0076] 特别地,本技术包括下面的特征B1至特征B6。
- [0077] (特征B1)
 - [0078] 根据本技术,可以实现具有下述配置的方法及装置:用于在包含不同内容的多条媒体内容经由不同传输路径传输并且由多个不同装置接收时通过使用声音来执行自动同步。
 - [0079] (1) 媒体内容是通过对视频、声音、图像、字符信息等进行多路复用而获取的数据流。
 - [0080] (2) 如前述示例(A1)至(A4)中的一样,作为传输目标的多条媒体内容具有时间同步关系。
 - [0081] (3) 作为传输目标的多条媒体内容中的至少一条媒体内容被确定为主通道信号,根据声音信号来计算声音同步特征量,以及以由系统限定的传输格式根据主通道信号来生成主传输信号。
 - [0082] (4) 为了满足每条剩余的媒体内容(子通道信号)与主通道信号之间的时间同步关系,以由系统限定的传输格式来执行主通道信号的声音同步特征量和子通道信号的多路复用处理,以及生成子传输信号。
 - [0083] (5) 接收主传输信号的主接收装置通过扬声器等以再现主通道信号的方式输出主通道信号的声音信号。
 - [0084] (6) 接收包括主通道信号的声音同步特征量的子传输信号的子接收装置收集(由

主接收装置通过扬声器、通过麦克风等输出的)主通道信号的声音,计算声音同步特征量,用所接收的主通道信号的声音同步特征量来执行自动同步计算,以及基于声音特征量计算同步校正信息(时差信息)。

[0085] (7)子接收装置用所接收的子通道信号上的主通道信号执行同步校正处理,并且参照基于声音特征量的同步校正信息执行再现。

[0086] 此外,如前述(1)的数据流的传输,呈现出在网络例如广播或因特网中的媒体内容的传输,并且由多路复用数据流占用的逻辑传输路径将被称为传输路径。

[0087] 此外,例如,由在日本未审查的专利申请公开第2013-174765号中公开的技术实现了上述“声音同步特征量的计算”以及“自动同步计算”。还可以在传输之前对声音同步特征量进行下采样或者必要时在自动同步计算期间利用声音同步特征量来执行声音同步特征量的帧率转换。

[0088] 通过使用这样的技术,在子接收装置收集主通道信号的声音时,即使在具有噪声或不想要的声音的不利环境中,也可以以健壮的方式执行自动同步计算。此外,不需要使用该技术。

[0089] 在这种情况下,需要在主传输信号的传输之前传输子传输信号。

[0090] (特征B2)

[0091] 在前述特征中(特征B1),在传输侧的系统分别以单侧的方式向主接收装置和子接收装置传输主传输信号和子传输信号。

[0092] 在这种情况下,需要在主传输信号之前传输子传输信号。

[0093] (特征B3)

[0094] 在前述特征中(特征B1),在传输侧的系统以单侧的方式向主接收装置传输主传输信号,并且子接收装置以子接收装置自身的定时经由网络来获取子传输信号,执行自动同步计算,以及执行子通道信号的同步再现。

[0095] 该配置的优点是子接收装置可以考虑通过网络等的传输中的延迟取决于子接收装置自身的便利性来控制子传输信号的获取。

[0096] (特征B4)

[0097] 在前述特征中(特征B1),主接收装置以主接收装置自身的定时经由网络来获取主传输信号以及再现主通道信号;并且子接收装置也以子接收装置自身的定时经由网络来获取子传输信号,执行自动同步计算,以及执行子通道信号的同步再现。

[0098] 该配置的优点是子接收装置可以考虑经由网络等的传输中的延迟取决于子传输信号自身的便利性来控制子传输信号的获取。

[0099] (特征B5)

[0100] 在前述特征中(特征B1),存在多系列的主通道信号的声音信号。

[0101] 例如,多系列的主通道信号与针对双语广播的主声音和补充声音相对应。针对所有系列的声音信号来计算声音同步特征量,并且多路复用声音同步特征量和利用子通道信号传输声音同步特征量。在所收集的声音与所有接收的声音同步特征量之间执行同步计算时,子接收装置确定主通道信号的声音中的哪些声音被重现。还在同步计算中检测由主接收装置输出的声音信号的切换。

[0102] (特征B6)

[0103] 在前述特征中(特征B1),通过子接收装置在自动同步计算中检测“同步中的偏差”,并且在子接收装置侧执行实时校正处理。

[0104] 由于主接收装置和子接收装置独立地操作,音频时钟不同,因而在同步中发生偏差。因此,可以在维持同步的同时通过对同步中的偏差进行检测和校正来再现多条内容。

[0105] <提供设备的配置示例>

[0106] 接下来,将给出应用本技术的特定实施方式的描述。

[0107] 首先,将给出提供如前述(A1)至(A4)中所示的示例一样具有时间同步关系的内容的提供设备的配置示例的描述。

[0108] 图1是示出了提供设备的配置示例的图。向提供设备11供给了主通道信号和子通道信号,主通道信号是用于再现主要内容(在下文中,被称为主内容)的信号,并且子通道信号是用于再现具有主内容中的相关内容的内容(在下文中,被称为子内容)的信号。

[0109] 在此,主内容和子内容配置有视频和声音中的至少任一种以及彼此具有时间同步关系。也就是说,期望在再现期间在同步状态下再现主内容和子内容。

[0110] 此外,将在主内容和子内容分别配置有用于再现伴随有图像信号的视频信号和声音信号的图像信号的前提下继续下面的描述。因此,在该示例中,主通道信号和子通道信号分别地配置有图像信号和声音信号。

[0111] 提供设备11包括:转换单元21、输出单元22、声音同步特征量计算单元23、多路复用处理单元24以及输出单元25。

[0112] 转换单元21将供给的主通道信号转换成由预定广播规则等限定的格式的信号,并且将获取的主传输信号作为结果供给至输出单元22。输出单元22经由广播波来广播从转换单元21供给的主传输信号,或者经由通信网络例如因特网来传输从转换单元21供给的主传输信号。

[0113] 声音同步特征量计算单元23从构成所供给的主通道信号的声音信号中提取声音同步特征量,并且将声音同步特征量供给至多路复用处理单元24。在此,声音同步特征量是在再现主内容和子内容时用于随着主内容一起同步和再现子内容的特征量。

[0114] 多路复用处理单元24通过使用所供给的主通道信号对来自声音同步特征量计算单元23的声音同步特征量与所供给的子通道信号之间的时间同步关系进行调整。也就是说,由于在提供设备11中主通道信号和子通道信号预先处于同步状态,因此多路复用处理单元24在声音同步特征量和子通道信号以时间同步关系的方式同步的状态下通过使用主通道信号将声音同步特征量与子通道信号关联起来。例如,在MPEG-4系统中,音频信号、视频信号等中的每种信号被处理为单媒体目标(基本流(ES))并且被多路复用。由于时间属性被限定在通过划分ES获取的被称为访问单元(AU)的最小单元中,因此可以通过将声音同步特征量处理为包括时间属性信息的一个媒体目标来容易地将具有作为子通道信号的媒体目标的声音同步特征量进行多路复用。

[0115] 此外,多路复用处理单元24在时间同步状态下对声音同步特征量和子通道信号进行多路复用,然后,必要时执行格式转换,并且将获取的子传输信号作为结果供给至输出单元25。

[0116] 例如,输出单元25通过广播波或通过包括因特网的通信网络传输从多路复用处理单元24供给的子传输信号。在此,主传输信号和子传输信号经由彼此不同的传输路径被传

输至内容再现侧的系统。

[0117] 尽管在图1中所示的示例中的提供设备11被配置为单个设备,但是提供设备11可以被配置为多个设备,或者可以由云计算执行相应的处理。

[0118] <声音同步特征量计算单元的配置示例>

[0119] 更具体地,例如,在图1中所示的声音同步特征量计算单元23被配置为如图2中所示。

[0120] 声音同步特征量计算单元23包括:频带划分单元51、周期性检测单元52-1至周期性检测单元52-4、周期性强度检测单元53-1至周期性强度检测单元53-4、周期性信息合并单元54、峰值检测单元55以及下采样单元56。

[0121] 频带划分单元51通过使用窗函数将构成所供给的主通道信号的声音信号划分成从大约几十毫秒至大约100毫秒的时间段。

[0122] 在此,针对一个时间段执行从频带划分单元51至峰值检测单元55执行的处理。可以通过对施加窗函数的时间位置进行转移来获取在时间方向上连续的多个时间段(时间帧),施加窗函数的时间位置被转移至延迟大约几毫秒至大约100毫秒。反之,下采样单元56将多个连续的时间段的结果合并成一个时间段,并且在合并之后计算新时间段的特征量。

[0123] 频带划分单元51通过使用多个带通滤波器将针对每个时间段的声音信号划分成四个频带,并且将相应频带上的声音信号供给至周期性检测单元52-1至周期性检测单元52-4。

[0124] 作为带通滤波器,使用随着频率变得更高进一步扩大通过频率的带宽的滤波器例如倍频带滤波器是有效的。

[0125] 周期性检测单元52-1至周期性检测单元52-4通过计算从频带划分单元51供给的预定频带中的每个时间段的声音信号的自相关函数来提取指示每个时间段的周期性的周期性信息。

[0126] 在本文中,虽然使用在由指标b指示的频带中具有由指标tau指示的时间延迟的声音信号的自相关函数x(b,tau)自身作为周期性信息,但是还可以使用通过用自相关函数x(b,tau)除以x(b,0)获取的值。此外,可以使用下述方法作为计算自相关函数x(b,tau)的方法:使用通过对预定频带中的声音信号执行离散傅里叶变换而获取的频谱的峰值。

[0127] 周期性检测单元52-1至周期性检测单元52-4向周期性强度检测单元53-1至周期性强度检测单元53-4和周期性信息合并单元54供给针对每个时间段所提取的周期性信息。在下文中,当不需要特别地区分周期性检测单元52-1至周期性检测单元52-4时,周期性检测单元52-1至周期性检测单元52-4将被简单地称为周期性检测单元52。

[0128] 周期性强度检测单元53-1至周期性强度检测单元53-4基于从周期性检测单元52-1至周期性检测单元52-4供给的针对每个时间段的周期性信息来计算每个时间段的周期性的强度。特别地,作为除了tau=0附近的频率信息之外的针对tau的频率信息的自相关函数x(b,tau)的最大值被计算为周期性的强度。随着周期性的强度增加,作为处理目标的声音信号的周期性增加。随着周期性的强度减小,作为处理目标的声音信号的周期性变得更可能是噪声的周期性。

[0129] 周期性强度检测单元53-1至周期性强度检测单元53-4取决于是否超过阈值来对作为针对每个时间段的周期性强度信息的每个时间段上的周期性的强度进行二值化。也就

是说,当每个时间段上的周期性的强度超过预定阈值时,周期性强度信息被设定成一。当周期性的强度等于或小于预定阈值时,周期性强度信息被设定成零。周期性强度检测单元53-1至周期性强度检测单元53-4向周期性信息合并单元54供给针对每个时间段的周期性强度信息。

[0130] 在下文中,当不需要特别地区分周期性强度检测单元53-1至周期性强度检测单元53-4时,周期性强度检测单元53-1至周期性强度检测单元53-4将被简单地称为周期性强度检测单元53。

[0131] 周期性信息合并单元54基于从周期性检测单元52供给的针对每个时间段的周期性信息以及从周期性强度检测单元53供给的针对每个时间段的周期性强度信息来执行合并针对每个时间段的周期性信息的周期性合并处理,特别地,周期性信息合并单元54通过使用下面的等式(1)来获取自相关函数x(b, tau)的和作为针对时间段的频率信息。

[0132] [数学式1]

$$[0133] s(\tau) = \frac{1}{N_p} \sum_{b=1}^{N_b} x(b, \tau) \cdot p(b) \quad \dots (1)$$

[0134] 在等式(1)中,N_b表示频带的总数,并且p(b)表示周期性强度信息。此外,N_p表示在p(b)是一时的频带的数量。

[0135] 周期性信息合并单元54将获取的针对每个时间段的周期性信息的和S(tau)作为周期性合并处理的结果供给至峰值检测单元55。

[0136] 峰值检测单元55针对每个时间段对从周期性信息合并单元54供给的周期性信息的和S(tau)执行峰值检测,并且生成峰值信息P(tau),其中,在峰值位置tau_p处的值是一并且除了峰值位置tau_p之外的位置处的值是零。作为峰值检测方法,例如,存在通过以下假定来检测峰值的方法:当周期性信息的和S(tau)的微分值从正值转变成负值时的指标tau处于峰值位置tau_p。

[0137] 此外,当在峰值位置tau_p处的周期性信息的和S(tau_p)小于预定阈值时,峰值检测单元55可以将峰值位置tau_p处的峰值信息P(tau_p)当作零。在这样配置的情况下,可以减少峰值信息P(tau_p)中的噪声。此外,峰值信息可以是周期性信息的和S(tau_p)自身。

[0138] 峰值检测单元55向下采样单元56供给针对每个时间段的峰值信息P(tau)作为在每个时间段中的声音同步特征量的时间序列数据。

[0139] 下采样单元56合并从峰值检测单元55供给的多个时间段的声音同步特征量即多个时间段上的峰值信息P(tau)作为针对新时间段的信息,并且生成峰值信息P'_i(tau)作为最终声音同步特征量。换言之,下采样单元56通过对峰值信息P(tau)进行下采样来生成峰值信息P'_i(tau)。

[0140] 在P'_i(tau)中,tau是指示时间延迟的指标,并且i是指示时间段的指标。下采样单元56向多路复用处理单元24供给这样获得的时间段上的峰值信息P'_i(tau)作为每个时间段上的声音同步特征量的时间序列数据。

[0141] 在此,将参照图3给出生成峰值信息P'_i(tau)的描述。在图3中,纵轴表示指示时间延迟的指标tau,并且横轴表示指示时间即时间段的指标i。

[0142] 在该示例中,在附图的上侧示出了峰值信息P(tau)的序列,并且在附图的下侧示

出了峰值信息 $P'_i(\tau)$ 的序列。具体地,在图3中,由 τ 表示时间延迟,并且由 $P_i(\tau)$ 表示由指标 i 指定的时间段上的峰值信息 $P(\tau)$ 。此外,每个方形表示时间段上的峰值信息。特别地,白色方形表示由该方形表示的峰值信息是零,并且黑色方形表示由该方形表示的峰值信息是一。

[0143] 在附图中,在上侧示出的针对峰值信息 $P_i(\tau)$ 的时间段的长度是8毫秒。也就是说,以8毫秒的时间间隔来计算峰值信息 $P_i(\tau)$ 。此外,在时间方向上(时间段方向)相邻的具有相同时间延迟 τ 的四个峰值信息项 $P_i(\tau)$ 被合并成一个,并获得峰值信息项 $P'_i(\tau)$ 。因此,针对每个峰值信息项 $P'_i(\tau)$ 的时间段是32毫秒。

[0144] 例如,下采样单元56通过计算下面的等式(2)来合并(下采样)峰值信息 $P_i(\tau)$ 以及获取峰值信息 $P'_i(\tau)$ 。

[0145] [数学式2]

$$P'_i(\tau) = P_{4i}(\tau) \cup P_{4i+1}(\tau) \cup P_{4i+2}(\tau) \cup P_{4i+3}(\tau) \dots (2)$$

[0147] 在等式(2)的计算中,如果来自待合并的四个连续的峰值信息项 $P_i(\tau)$ 中的一个或更多个峰值信息项 $P_i(\tau)$ 具有值“一”,则在合并之后获得的峰值信息 $P'_i(\tau)$ 的值被认为是“一”。反之,如果所有待合并的四个连续峰值信息项 $P_i(\tau)$ 的值都是“零”,则在合并之后获得的峰值信息 $P'_i(\tau)$ 的值被认为是“零”。

[0148] 通过获得在时间段方向上连续的峰值信息项 $P_i(\tau)$ 的逻辑和以及对如上所述的逻辑和进行下采样,可以消除由于下采样而使与峰值位置有关的信息被移除的情况,峰值位置包括在时间方向上排列的峰值信息项的序列中。在这样的配置下,即使在下采样之后也可以维持与峰值位置有关的位置如何在时间延迟方向上转移。

[0149] 例如,当来自多个时间段上的峰值信息项 $P_i(\tau)$ 的一个峰值信息项 $P_i(\tau)$ 的值在用于对峰值信息 $P_i(\tau)$ 进行下采样的下采样之后仅仅被用作峰值信息 $P'_i(\tau)$ 的值时,信息量是不足够的,并且同步计算的精确性恶化。也就是说,性能恶化。

[0150] 特别地,例如,当峰值位置在四个时间段上转移时,如果峰值信息 $P_i(\tau)$ 仅仅细化为峰值信息 $P'_i(\tau)$,则仅仅采用在转移的过程中的时间段上的峰值信息 $P_i(\tau)$ 作为最终的特征量,并且峰值位置已经转移的信息丢失。

[0151] 反之,根据前述基于用于下采样的多个时间段上的峰值信息 $P_i(\tau)$ 输出适当的值作为峰值信息 $P'_i(\tau)$ 的方法,可以在下采样之后的时间段内维持时间段上已经发生转移的信息。因此,即使执行下采样处理也可以维持检测性能。

[0152] 此外,通过执行这样的下采样处理可以减少声音同步特征量的传输量。此外,可以减少在存储器或存储装置中用于维持所计算的声音同步特征量的容量。

[0153] 此外,可以减少用于在两个声音同步特征量之间执行同步处理的计算量。由于同步处理的计算量随着输入特征量的长度增加 n 倍而增加 n^2 倍,因此下采样处理的效果是重要的。反之,由于如果仅仅执行细化处理则同步检测性能会恶化,因此需要由下采样单元56在维持如下采样方法中需要的信息的同时执行下采样的处理。

[0154] 尽管在图3中描述了作为声音同步特征量的峰值信息被下采样成1/4的示例,但是还可以以任何其他比率例如1/2或1/8执行转换(下采样)。

[0155] 此外,除了前述等式(2)的计算方法之外的方法可以用于对峰值信息进行下采样。

[0156] 例如,当来自四个时间段的两个或更多个时间段上的峰值信息 $P_i(\tau)$ 的值是

“一”时,在下采样之后的峰值信息 $P'_i(\tau)$ 的值可以被认为是“一”。可替选地,当三个或更多个时间段上的峰值信息 $P_i(\tau)$ 的值是“一”时或者当所有四个段上的峰值信息 $P_i(\tau)$ 的值都是“一”时,在下采样之后的峰值信息 $P'_i(\tau)$ 的值可以被认为是“一”。

[0157] 此外,当在下采样之前来自四个时间段的两个或更多个连续时间段上的连续的峰值信息 $P_i(\tau)$ 的值是“一”时,在下采样之后的峰值信息 $P'_i(\tau)$ 的值可以被认为是“一”,或者当三个或更多个连续时间段上的峰值信息 $P_i(\tau)$ 的值是“一”时,在下采样之后的峰值信息 $P'_i(\tau)$ 的值可以被认为是“一”。

[0158] 尽管上面的描述给出了在时间轴方向上(时间段方向上)对峰值信息 $P_i(\tau)$ 进行下采样的方法,但是,可以在时间延迟 τ 方向上对峰值信息 $P_i(\tau)$ 进行下采样。

[0159] 在这样的情况下,下采样单元56通过计算下面的等式(3)来对峰值信息 $P_i(\tau)$ 进行下采样并且获得峰值信息 $P'_i(\tau)$ 。

[0160] [数学式3]

$$[0161] P'_i(\tau) = P_i(4\tau) \cup P_i(4\tau+1) \cup P_i(4\tau+2) \cup P_i(4\tau+3)$$

$$[0162] \dots (3)$$

[0163] 在等式(3)的计算中,在时间延迟 τ 方向上连续排列的、相同时间段上的四个峰值信息项 $P_i(\tau)$ 被合并,并且获得一个峰值信息项 $P'_i(\tau)$ 。

[0164] 此时,当来自待合并的四个连续的峰值信息项 $P_i(\tau)$ 中的一个或更多个峰值信息 $P_i(\tau)$ 的值是“一”时,通过合并获得的峰值信息 $P'_i(\tau)$ 的值被认为是“一”。反之,当所有待合并的四个连续峰值信息项 $P_i(\tau)$ 的值都是“零”时,通过合并获得的峰值信息 $P'_i(\tau)$ 的值被认为是“零”。

[0165] 此外,可以在时间段*i*方向和时间延迟 τ 方向两者上对峰值信息 $P_i(\tau)$ 进行下采样。

[0166] 在这样的情况下,下采样单元56通过计算下面的等式(4)来对峰值信息 $P_i(\tau)$ 进行下采样并且获得峰值信息 $P'_i(\tau)$ 。

[0167] [数学式4]

$$[0168] P'_i(\tau) = P_{2i}(2\tau) \cup P_{2i+1}(2\tau) \cup P_{2i}(2\tau+1) \cup P_{2i+1}(2\tau+1)$$

$$[0169] \dots (4)$$

[0170] 在等式(4)的计算中,包括在时间段*i*方向上连续排列的具有相同时间差 τ 的两个峰值信息项 $P_i(\tau)$ 以及在时间延迟 τ 方向上被排列成与上面两个峰值信息项 $P_i(\tau)$ 相邻的两个峰值信息项 $P_i(\tau)$ 的总共四个峰值信息项 $P_i(\tau)$ 被合并,并且获得一个峰值信息项 $P'_i(\tau)$ 。

[0171] 此时,当来自待合并的四个峰值信息项 $P_i(\tau)$ 中的一个或更多个峰值信息 $P_i(\tau)$ 的值是“一”时,通过合并获得的峰值信息 $P'_i(\tau)$ 的值被认为是“一”。反之,当所有待合并的四个峰值信息项 $P_i(\tau)$ 的值都是“零”时,通过合并获得的峰值信息 $P'_i(\tau)$ 的值被认为是“零”。

[0172] 下采样单元56通过对峰值信息 $P(\tau)$ 进行下采样来获得峰值信息 $P'_i(\tau)$ 并且向多路复用处理单元24供给针对每个新获得的时间段的峰值信息 $P'_i(\tau)$ 作为针对每个时间段的声音同步特征量的时间序列数据。

[0173] <内容再现系统的配置示例>

[0174] 接下来,将给出下述内容再现系统的配置的描述:接收从提供设备11传输的主传输信号和子传输信号分别作为主接收信号和子接收信号,并且再现主内容和子内容。例如,这样的内容再现系统被配置为如图4所示。

[0175] 图4中所示的内容再现系统包括:主接收装置81、显示单元82、扬声器83、麦克风84、子接收装置85、显示单元86以及扬声器87。虽然在此示出的是被配置为多个设备的内容再现系统的情况,但是内容再现系统可以被配置为单个设备。

[0176] 主接收装置81接收从提供设备11传输的主接收信号,并且控制从主接收信号获得的主内容的再现。

[0177] 主接收装置81包括输入单元111和再现处理单元112。

[0178] 输入单元111接收从提供设备11传输的主传输信号作为主接收信号,并且将主接收信号供给至再现处理单元112。再现处理单元112提取包括在从输入单元111供给的主接收信号中的主内容的图像信号和声音信号,将图像信号供给至显示单元82,使显示单元82再现图像信号,将声音信号供给至扬声器83,并且使扬声器83再现声音信号。也就是说,再现处理单元112控制主内容的再现。

[0179] 例如,显示单元82配置有液晶显示装置,并且基于从再现处理单元112供给的图像信号来显示主内容的图像(视频)。扬声器83是声音再现装置,并且基于从再现处理单元112供给的声音信号来输出主内容的声音。

[0180] 麦克风84收集从扬声器83输出的主内容的声音,并且将获取的声音信号作为结果供给至子接收装置85。

[0181] 子接收装置85接收从提供设备11传输的子传输信号作为子接收信号,并且控制从子接收信号中获得的子内容的再现。

[0182] 子接收装置85包括:声音同步特征量计算单元121、缓冲器122、输入单元123、分离处理单元124、缓冲器125、同步计算单元126以及再现处理单元127。

[0183] 声音同步特征量计算单元121计算从麦克风84供给的声音信号的声音同步特征量,并且将声音同步特征量供给至缓冲器122。缓冲器122暂时地记录从声音同步特征量计算单元121供给的声音同步特征量。

[0184] 输入单元123接收从提供设备11传输的子接收信号,并且将子接收信号供给至分离处理单元124。分离处理单元124将从输入单元123供给的子接收信号分离成声音同步特征量和子通道信号,并且将声音同步特征量和子通道信号供给至缓冲器125。缓冲器125暂时地记录从分离处理单元124供给的声音同步特征量和子通道信号。

[0185] 同步计算单元126基于在缓冲器122中记录的声音同步特征量和在缓冲器125中记录的声音同步特征量取决于用于同步主内容和子内容的声音特征量来生成同步校正信息,并且将同步校正信息供给至再现处理单元127。也就是说,同步计算单元126通过对从通过声音收集获得的声音信号中提取的声音同步特征量与包括在子接收信号中的声音同步特征量之间的处理进行匹配来检测主内容与子内容之间在再现时间上的偏差,并且基于指示偏差的声音特征量来生成同步校正信息。

[0186] 再现处理单元127基于从同步计算单元126供给的同步校正信息来对记录在缓冲器125中的子通道信号的再现定时(时间)进行校正,并且将图像信号和声音信号作为子通道信号分别供给至显示单元86和扬声器87。也就是说,再现处理单元127控制子内容的再

现。例如,当声音同步特征量被处理为媒体目标并且通过使用MPEG-4系统来与子通道信号的媒体目标进行同步和多路复用时,针对作为每个媒体目标的最小单元的每个访问单元(AU)限定时间属性,因此,可以根据前述同步校正信息来计算子通道信号的媒体目标的适当的再现定时(时间)。

[0187] 例如,显示单元86配置有液晶显示装置,并且基于从再现处理单元127供给的图像信号来显示子内容的图像(视频)。扬声器87是声音再现装置,并且基于从再现处理单元127供给的声音信号来输出子内容的声音。

[0188] <声音同步特征量计算单元的配置示例>

[0189] 更具体地,例如,在图4中所示的计算单元121被配置为如图5中所示。

[0190] 声音同步特征量计算单元121包括:频带划分单元151、周期性检测单元152-1至周期性检测单元152-4、周期性强度检测单元153-1至周期性强度检测单元153-4、周期性信息合并单元154以及峰值检测单元155。

[0191] 此外,由于频带划分单元151至峰值检测单元155与在图2中所示的频带划分单元51至峰值检测单元55相同,因此将省略其描述。然而,还可以针对频带划分单元151和频带划分单元51将窗函数的转移时间设定成不同值。例如,当子接收装置85具有足够的计算资源时,可以由频带划分单元151通过使用较短转移时间来提取具有较细间隔的声音同步特征量。

[0192] 在下文中,当不需要特别地区分周期性检测单元152-1至周期性检测单元152-4时,周期性检测单元152-1至周期性检测单元152-4将简单地被称为周期性检测单元152,并且当不需要特别地区分周期性强度检测单元153-1至周期性强度检测单元153-4时,周期性强度检测单元153-1至周期性强度检测单元153-4将被称为周期性强度检测单元153。

[0193] <同步计算单元的配置示例>

[0194] 更具体地,例如,在图4中所示的同步计算单元126被配置为如图6中所示。

[0195] 在图6中所示的同步计算单元126包括:帧率转换单元181、帧率转换单元182、块合并单元183、块合并单元184、相似度计算单元185以及最优路径搜索单元186。

[0196] 帧率转换单元181从缓冲器122读取针对主内容的每个时间段的声音同步特征量的时间序列数据,转换声音同步特征量的帧率,并且将具有经转换的帧率的声音同步特征量供给至块合并单元183。本文中描述的帧率表示在声音同步特征量的时间序列数据上每单位时间的时间段的数量,即时间段长度。

[0197] 帧率转换单元182从缓冲器125读取针对主内容的每个时间段的声音同步特征量的时间序列数据,转换声音同步特征量的帧率,并且将具有经转换的帧率的声音同步特征量供给至块合并单元184。

[0198] 在缓冲器122和缓冲器125中保持的声音同步特征量在一些情况下分别地具有不同帧率,即不同的时间段长度。

[0199] 例如,可以考虑下述情况:包括在子传输信号中的声音同步特征量被设定成具有低比率以降低从提供设备11提供的子内容(子传输信号)的传输比特率,而由于不需要传输声音同步特征量,所以根据由麦克风84收集的声音计算的声音同步特征量被设定成具有高比率。

[0200] 在这样的情况下,例如,可以考虑通过与如图7中所示的下采样单元56的方法相同

的方法来对根据由麦克风84收集的声音计算的声音同步特征量进行下采样。此外,作为由图7中的箭头Q11至箭头Q14中的每个箭头表示的声音同步特征量的峰值信息的纵轴表示时间延迟tau,并且横轴表示时间段i。此外,方形表示时间段上的峰值信息。

[0201] 在该示例中,如由箭头Q11所表示,获取作为声音同步特征量的峰值信息,然后,在提供设备11侧对该峰值信息进行下采样,并且由箭头Q12所示,该峰值信息作为较长时间段上的峰值信息被传输至子接收装置85。在此,在与8毫秒对应的时间段上的峰值信息经受帧率转换(下采样)以获得与32毫秒对应的时间段上的峰值信息。

[0202] 反之,子接收装置85中的声音同步特征量计算单元121根据通过收集由主接收装置81再现的主内容的声音而获取的声音信号来计算声音同步特征量,并且获得作为声音同步特征量的由箭头Q13表示的峰值信息作为结果。在此,针对与8毫秒对应的每个时间段来计算由箭头Q13表示的峰值信息。

[0203] 如上所述由声音同步特征量计算单元121获取的声音同步特征量和从提供设备11接收的声音同步特征量具有不同的时间段长度,即不同的帧率。因此,帧率转换单元181按照帧率转换对由声音同步特征量计算单元121获取的声音同步特征量进行下采样,使得声音同步特征量的帧率彼此一致,并且获取峰值信息作为由箭头Q14表示的声音同步特征量。由箭头Q14表示的声音同步特征量是32毫秒的时间段上的峰值信息。

[0204] 在使帧率(时间段长度)彼此一致之后,声音同步特征量用于执行同步计算。通过如上所述在子接收装置85侧对声音同步特征量进行下采样,可以处理任意帧率(比特率)。

[0205] 此外,还存在下述情况:传输至子接收装置85的声音同步特征量具有高比率,而根据由麦克风84收集的声音计算的声音同步特征量具有低比率。例如,存在下述情况:子接收装置85的计算资源是不够的,从而增加帧转移量以减少用于计算声音同步特征量所需要的计算量。

[0206] 在这样的情况下,例如,通过使用与下采样单元56的方法相同的方法来由帧率转换单元182对由图8中箭头Q21表示的包括在子传输信号中的声音同步特征量的帧率进行下采样,并且获取由箭头Q22表示的声音同步特征量。在图8中,由箭头Q21至箭头Q23中的每个箭头表示的作为声音同步特征量的峰值信息的纵轴表示时间延迟tau,并且其横轴表示时间段i。此外,方形表示时间段上的峰值信息。

[0207] 在该示例中,8毫秒的时间段上的峰值信息经受帧率转换(下采样)以获取32毫秒的时间段上的峰值信息。

[0208] 此外,子接收装置85中的声音同步特征量计算单元121计算来自通过收集由主接收装置81再现的主内容的声音而获取的声音信号的声音同步特征量,并且获取作为声音同步特征量的由箭头Q23表示的峰值信息作为结果。在此,针对32毫秒的每个时间段来计算由箭头Q23表示的峰值信息。

[0209] 可以通过如上所述对包括在子传输信号中的声音同步特征量进行下采样来使包括在子传输信号中的声音同步特征量的帧率和由子接收装置85计算的声音同步特征量的帧率彼此一致。

[0210] 此外,虽然在上面的描述中通过对具有较高帧率的声音同步特征量进行下采样来使帧率彼此一致,但是可以通过对具有较低帧率的声音同步特征量进行上采样来使帧率彼此一致。

[0211] 在这样的情况下,例如,如图9中所示,由帧率转换单元182对包括在子传输信号中的声音同步特征量的帧率进行上采样。在图9中,由箭头Q31至箭头Q34中的每个箭头表示的作为声音同步特征量的峰值信息的纵轴表示时间延迟 τ ,并且其横轴表示时间段i。此外,方形表示时间段上的峰值信息。

[0212] 在该示例中,如由箭头Q31所表示,获取作为声音同步特征量的峰值信息,然后,在提供设备11侧对该峰值信息进行下采样,以及如由箭头Q32所示,该峰值信息作为较长时间段上的峰值信息被传输至子接收装置85。在此,8毫秒的时间段上的峰值信息经受帧率转换(下采样)以获取32毫秒的时间段上的峰值信息。

[0213] 反之,子接收装置85中的声音同步特征量计算单元121计算通过收集由主接收装置81再现的主内容的声音而获取的声音信号的声音同步特征量,并且获取作为声音同步特征量的由箭头Q33表示的峰值信息作为结果。在此,针对8毫秒的每个时间段来计算由箭头Q33表示的峰值信息。

[0214] 在该示例中,由声音同步特征量计算单元121计算的声音同步特征量的帧率和从提供设备11接收的声音同步特征量彼此不一致。

[0215] 因此,帧率转换单元182通过对从提供设备11接收的作为声音同步特征量的峰值信息进行上采样并且计算由箭头Q34表示的8毫秒的时间段上的峰值信息来使用于同步计算的声音同步特征量的时间同步间隔均匀。例如,帧率转换单元182对通过计算下述等式(5)获得的峰值信息进行上采样。

[0216] [数学式5]

[0217] $P_{4i}(\tau) = P_i'(\tau)$

[0218] $P_{4i+1}(\tau) = P_i'(\tau)$

[0219] $P_{4i+2}(\tau) = P_i'(\tau)$

[0220] $P_{4i+3}(\tau) = P_i'(\tau) \dots (5)$

[0221] 在等式(5)的计算中,在没有任何改变的情况下,与在时间方向上(时间段方向)相邻的在上采样之后的四个峰值信息项 $P_i(\tau)$ 具有相同时间延迟 τ 并且具有相同位置关系的在上采样之前的峰值信息项 $P_i'(\tau)$ 的值被认为是在上采样之后的四个峰值信息项 $P_i(\tau)$ 中的每个峰值信息项的值。

[0222] 如上所述,通过适当地根据较高帧率对用于同步计算的声音同步特征量进行上采样,可以实现具有按照推测的高分辨率的同步精度。

[0223] 此外,还可以对包括在子传输信号中的声音同步特征量和由声音同步特征量计算单元121计算的声音同步特征量两者进行下采样以减少子接收装置85中的计算资源。

[0224] 通过提供如上所述的帧率转换单元181和帧率转换单元182,可以利用不同帧率对声音同步特征量进行同步。此外,可以根据计算资源、传输频带等指定各种帧率,以及增强系统的灵活性。

[0225] 返回图6的说明,块合并单元183接收来自帧率转换单元181的主内容的每个时间段上的声音同步特征量的时间序列数据的供给,并且在将多个(例如,六十四个)连续的时间段当做块的同时执行块单元的合并。块合并单元183将块单元上的时间同步特征量的时间序列数据供给至相似度计算单元185。

[0226] 块合并单元184接收来自帧率转换单元182的主内容的每个时间段上的声音同步

特征量的时间序列数据的供给,并且将多个(例如,六十四个)连续的时间段当做块的同时执行块单元的合并。块合并单元184将块单元上的声音同步特征量的时间序列数据供给至相似度计算单元185。

[0227] 此外,构成块的多个时间段不需要是连续的时间段。例如,多个偶数的时间段可以被认为是一个块,并且多个奇数的时间段可以被认为是一个块。在这样的情况下,可以针对每个时间段对声音同步特征量的时间序列数据执行细化处理,从而减少计算量。

[0228] 相似度计算单元185计算从块合并单元183和块合并单元184供给的块单元上的声音同步特征量的时间序列数据项之间的相似度,并且生成示出了块之间的相似度的相似度矩阵。相似度计算单元185将相似度矩阵供给至最优路径搜索单元186。

[0229] 最优路径搜索单元186根据从相似度计算单元185供给的相似度矩阵搜索具有最优相似度的路径,并且基于声音特征量生成指示与路径上的相似度对应的两个块之间的时间差的信息作为同步校正信息。然后,最优路径搜索单元186将基于声音特征量的同步校正信息供给至再现处理单元127。

[0230] 如上所述,同步计算单元126基于声音信号的音高信息取决于声音特征量生成同步校正信息。因此,即使在每个声音信号包括不同噪声的情况下,也可以以健壮的方式来基于声音特征量生成同步校正信息。

[0231] 也就是说,当人听到具有频率特性的多个声音时,即使声音中包含噪声,人也可以通过识别具有相同基频的声音即通过识别具有相同音高的声音来容易地识别共同分量。考虑这样的事实来获得本技术,并且通过基于音高信息生成同步校正信息来相对于噪声以健壮的方式生成同步校正信息。

[0232] <关于相似度计算和搜索具有最优相似度的路径>

[0233] 在此,将给出关于相似度计算和搜索具有最优相似度的路径的描述。

[0234] 图10是示出了作为相似度计算的目标的块的图。

[0235] 在图10中,i是通过声音同步特征量计算单元121获取的声音同步特征量的块的指标,并且j是包括在子接收信号中的声音同步特征量的块的指标。更具体地,虽然这些声音同步特征量通过帧率转换单元181或帧率转换单元182适当地经受帧率转换,但是为了简化说明,在此将在不执行帧率转换的前提下继续相似性计算的描述。

[0236] 此外,X(i)表示来自通过声音同步特征量计算单元121获取的声音同步特征量中的具有指标i的块上的声音同步特征量的时间序列数据,并且Y(j)表示来自包括在子接收信号中的声音同步特征量中的具有指标j的块上的声音同步特征量的时间序列数据。

[0237] 如图10中所示,相似度计算的目标是n个X(i)中的每个X(i)与m个Y(j)中的每个Y(j)之间的n×m个组合。

[0238] 图11是示出了计算相似度的方法的图。

[0239] 在图11的矩阵中,横轴表示块内时间段数量,块内时间段数量表示块中的最大时间段的数量,并且纵轴表示指标tau。此外,白色方形表示具有在相应的块内时间段数量的时间段上的指标tau的声音同步特征量的时间序列数据P(tau)是零,并且黑色方形表示时间序列数据P(tau)是一。此外,在图11的示例中,构成块的时间段的数量是四,并且tau是从零至三。

[0240] 当如图11所示计算X(i)与Y(j)之间的相似度时,首先,计算X(i)与Y(j)之间的逻

辑积 $X(i) \cap Y(j)$ ，然后计算 $X(i)$ 与 $Y(j)$ 之间的逻辑和 $X(i) \cup Y(j)$ 。如图11所示，当 $X(i)$ 中的每个 $X(i)$ 和 $Y(j)$ 中的每个 $Y(j)$ 由九个0和七个1构成时计算 $X(i)$ 与 $Y(j)$ 之间的相似度，首先，计算 $X(i) \cap Y(j)$ 的逻辑积，该逻辑积由十二个0和四个1构成，然后计算 $X(i) \cup Y(j)$ 的逻辑和，该逻辑和由六个0与十个1构成。

[0241] 然后，通过下面的等式(6)，基于为逻辑积 $X(i) \cap Y(j)$ 中的1的数量的Number($X(i) \cap Y(j)$)和为逻辑和中的1的数量的Number($X(i) \cup Y(j)$)来计算 $X(i)$ 与 $Y(j)$ 之间的相似度 $A(i, j)$ 。

[0242] [数学式6]

$$A(i, j) = \text{Number}(X(i) \cap Y(j)) / \text{Number}(X(i) \cup Y(j)) \quad \dots (6)$$

[0245] 在图11的示例中，Number($X(i) \cap Y(j)$)是四，Number($X(i) \cup Y(j)$)是十，因此，相似度 $A(i, j)$ 是0.4。

[0246] 当周期性信息的和 $S(\tau)$ 用作声音同步特征量的时间序列数据时，可以将通过使用余弦距离计算相似度的方法用作计算相似度的方法。

[0247] 此外，相似度矩阵是表示在与指标 i 和指标 j 对应的每个点处的相似度 $A(i, j)$ 的信息，其中，横轴表示指标 j ，并且纵轴表示指标 i 。

[0248] 最优路径搜索单元186通过使用动态规划法搜索具有使相似度矩阵上的相似度的积分值最大化的相似度的路径作为具有最优相似度的路径。最优路径搜索单元186基于声音特征量生成与具有最优相似度的路径上的相似度对应的不同的 $i-j$ 指标作为同步校正信息。

[0249] <传输处理的说明>

[0250] 接下来，将给出提供设备11的操作的描述。

[0251] 当供给彼此时间同步的主通道信号和子通道信号时，提供设备11执行传输处理并且传输主传输信号和子传输信号。在下文中，将参照图12中的流程图来给出由提供设备11进行传输处理的描述。

[0252] 在步骤S11中，声音同步特征量计算单元23执行声音同步特征量计算处理以计算来自构成供给的主通道信号的声音信号的声音同步特征量，并且将声音同步特征量供给至多路复用处理单元24。

[0253] 稍后将给出声音同步特征量计算量的详细描述。

[0254] 在步骤S12中，转换单元21通过将所供给的主通道信号转换成以由系统限定的预定传输格式的信号来生成主传输信号，并且将所获得的主传输信号供给至输出单元22。

[0255] 在步骤S13中，输出单元22传输从转换单元21供给的主传输信号。

[0256] 在步骤S14中，多路复用处理单元24执行声音同步特征量与子通道信号之间的多路复用处理，并且将获得的子传输信号作为结果供给至输出单元25。

[0257] 例如，多路复用处理单元24通过使用所供给的主通道信号对声音同步特征量和以由系统限定的传输的格式的子通道信号进行多路复用，使得来自声音同步特征量计算单元23的声音同步特征量与所供给的子通道信号之间的时间同步关系被满足。

[0258] 在这样配置的情况下，例如，获取如图13中所示的子传输信号。

[0259] 在图13的示例中，作为子传输信号以比特流形式的段T11和段T12中的每个在一个

帧中包括图像信号、声音信号以及声音同步特征量。

[0260] 例如,包括在段T11中的图像信号和声音信号与一个帧中的子通道信号对应,并且包括在段T11中的声音同步特征量是从帧中的主通道信号中提取的声音同步特征量,该声音同步特征量暂时地与子通道信号对应。如上所述,在相同帧中的子通道信号和声音同步特征量被关联并且在子传输信号中被多路复用,并且可以在子传输信号的接收侧指定与每个帧中的子通道信号对应的声音同步特征量。

[0261] 返回图12中的流程图的说明,在步骤S15中,输出单元25传输从多路复用单元24供给的子传输信号,并且完成传输处理。

[0262] 如上所述,提供设备11通过关联和多路复用从主通道信号中获取的声音同步特征量与子通道信号来生成子传输信号,并且传输子传输信号和主传输信号。

[0263] 通过如上所述关联和传输声音同步特征量与子通道信号,即使在接收侧由多个不同装置经由不同传输路径来接收主通道信号和子通道信号的情况下,通过使用声音同步特征量,也可以在维持同步的同时再现主内容和子内容。

[0264] <声音同步特征量计算处理的说明>

[0265] 接下来,将参照图14中的流程图给出与图12中的步骤S11中的处理对应的声音同步特征量计算处理的描述。

[0266] 在步骤S41中,频带划分单元51通过使用窗函数将供给的声音信号划分成从大约几十毫秒至大约一百毫秒的时间段。

[0267] 在步骤S42中,频带划分单元51通过使用多个带通滤波器将每个时间段上的声音信号划分成四个频带。频带划分单元51将相应频带上的声音信号供给至相应的周期性检测单元52-1至周期性检测单元52-4。

[0268] 在步骤S43中,每个周期性检测单元52通过计算从频带划分单元51供给的预定频带上的每个时间段上的声音信号的自相关函数 $x(b, \tau)$ 来提取针对每个时间段的周期性信息,并且将该周期性信息供给至周期性强度检测单元53和周期性信息合并单元54。此外,通过周期性检测单元52中的每个周期性检测单元来执行步骤S43中的处理。

[0269] 在步骤S44中,周期性强度检测单元53基于从周期性检测单元52供给的针对每个时间段的周期性信息来计算针对每个时间段的周期性的强度。然后,周期性强度检测单元53通过取决于强度是否超过阈值对针对每个时间段的周期性强度进行二值化来生成针对每个时间段的周期性强度信息,并且将周期性强度信息供给至周期性信息合并单元54。此外,通过周期性强度检测单元53中的每个周期性强度检测单元来执行步骤S44中的处理。

[0270] 在步骤S45中,周期性信息合并单元54基于从周期性检测单元52供给的针对每个时段的周期性信息以及从周期性强度检测单元53供给的针对每个时段的周期性强度信息通过使用前述等式(1)来执行周期性合并处理。周期性信息合并单元54将获取的针对每个时段的周期性信息的和 $S(\tau)$ 作为周期性合并处理的结果供给至峰值检测单元55。

[0271] 在步骤S46中,峰值检测单元55针对每个时间段对从周期性信息合并单元54供给的周期性信息的和 $S(\tau)$ 执行峰值检测,生成峰值信息 $P(\tau)$,并且将峰值信息 $P(\tau)$ 供给至下采样单元56。

[0272] 在步骤S47中,下采样单元56通过将从峰值检测单元55供给的多个时间段上的峰值信息 $P(\tau)$ 合并成一个时间段上的峰值信息来对峰值信息执行下采样处理。

[0273] 下采样单元56将如上所述获取的针对每个时间段的峰值信息作为针对每个时间段的声音同步特征量的时间序列数据供给至多路复用处理单元24，并且完成声音同步特征量计算处理。如果完成了声音同步特征量计算处理，则然后处理进行至图12中的步骤S12。

[0274] 声音同步特征量计算单元23基于如上所述的周期性信息计算声音同步特征量，因此，可以以健壮的方式生成声音同步特征量。

[0275] <主内容再现处理的说明>

[0276] 如果从提供设备11传输主传输信号，则内容再现系统获取主传输信号作为主接收信号，并且再现主内容。在下文中，将参照图15中的流程图来给出由内容再现系统进行的主内容再现处理的描述。

[0277] 在步骤S71中，输入单元111获取主接收信号，并且将主接收信号供给至再现处理单元112。例如，输入单元111通过接收从提供设备11传输的主接收信号来获取主接收信号。

[0278] 在步骤S72中，再现处理单元112基于从输入单元111供给的主接收信号来再现主内容，并且完成主内容再现处理。

[0279] 例如，再现处理单元112从主接收信号中提取主内容的图像信号和声音信号，将图像信号供给至显示单元82，使显示单元82再现图像信号，将声音信号供给至扬声器83，以及使扬声器83再现声音信号。在这样配置的情况下，再现了主内容。

[0280] 如上所述，内容再现系统获取主接收信号并且再现主内容。

[0281] <子内容再现处理的说明>

[0282] 此外，内容再现系统以与主内容的再现同步的方式获取子接收信号并且再现子内容。在下文中，将参照图16中的流程图来给出由内容再现系统进行的子内容再现处理的描述。

[0283] 在步骤S101中，输入单元123获取子接收信号，并且将子接收信号供给至分离处理单元124。例如，输入单元123通过接收从提供设备11传输的子传输信号来获取子接收信号作为子接收信号。

[0284] 在步骤S102中，分离处理单元124将从输入单元123供给的子接收信号分离成子通道信号和声音同步特征量，将所分离的子通道信号和声音同步特征量供给至缓冲器125，以及使缓冲器125记录子通道信号和声音同步特征量。

[0285] 在步骤S103中，麦克风84收集从扬声器83输出的主内容的声音，并且将获取的声音信号作为结果供给至声音同步特征量计算单元121。在步骤S103中，例如，收集由图15中的步骤S72中的处理所再现的主内容的声音。

[0286] 在步骤S104中，声音同步特征量计算单元121执行声音同步特征量计算处理以计算来自麦克风84供给的声音信号的声音同步特征量，将声音同步特征量供给至缓冲器122，以及使缓冲器122记录声音同步特征量。

[0287] 虽然执行图17中的流程图所示的步骤S131至步骤S136中的处理作为声音同步特征量计算处理，但是该处理与图14中的步骤S41至步骤S46中的处理相同，因此，将省略其中的描述。然而，在图17所示的声音同步特征量计算处理中，根据从麦克风84供给的声音信号计算声音同步特征量，并且在缓冲器122中积累声音同步特征量。此外，声音同步特征量计算单元121将由峰值检测单元155获取的峰值信息作为声音同步特征量。

[0288] 返回图16中的流程图的说明，在步骤S105中，同步计算单元126基于声音特征量执

行同步校正信息生成处理以生成同步校正信息,并且将同步校正信息供给至再现处理单元127。虽然稍后描述同步校正信息生成处理的详细描述,但是在该处理中,通过比较在缓冲器122中记录的声音同步特征量与在缓冲器125中记录的声音同步特征量来生成基于用于同步主内容与子内容的声音特征量的同步校正信息。

[0289] 在步骤S106中,再现处理单元127根据从同步计算单元126供给的基于声音特征量的同步校正信息对在缓冲器125中记录的子通道的再现定时进行校正,并且基于校正之后的子通道信号来再现子内容。

[0290] 也就是说,再现处理单元127较早地或在具有由基于声音特征量的同步校正信息指示的时间的延迟的情况下将构成子通道信号的图像信号和声音信号供给至显示单元86和扬声器87,并且使显示单元86和扬声器87再现图像信号和声音信号。换言之,与在当前时刻被再现的主内容的一部分对应的在再现时刻的子内容的一部分被再现,在当前时刻被再现的主内容的该部分由基于声音特征量的同步校正信息来指定。

[0291] 例如,在子内容的无声音段和主内容的无声音段中执行用于同步子内容与主内容的再现位置的调整(校正)。

[0292] 显示单元86基于从再现处理单元127供给的图像信号来显示子内容的图像,并且扬声器87基于从再现处理单元127供给的声音信号来输出子内容的声音。

[0293] 如果如上所述以与主内容同步的方式再现子内容,则完成了子内容再现处理。

[0294] 如上所述,内容再现系统计算来自通过收集被再现的主内容的声音而获取的声音信号的声音同步特征量,并且通过使用所获取的声音同步特征量和包含在子接收信号中的声音同步特征量来计算基于声音特征量的同步校正信息。此外,内容再现系统通过使用所获得的同步校正信息来以与主内容同步的方式再现子内容。

[0295] 如上所述,通过使用从经由声音收集获取的声音信号中提取的声音同步特征量和包含在子接收信号中的声音同步特征量来计算基于声音特征量的同步校正信息,即使在主内容的传输路径和子内容的传输路径不同的情况下也可以以同步的方式来再现主内容和子内容。

[0296] 虽然在该示例中,针对每个帧执行声音同步特征量的同步计算即匹配处理,但是不需要以时间连续的方式执行声音同步特征量的同步计算,而是可以间断地执行声音同步特征量的同步计算。然而,当连续地执行同步计算时,可以在没有不舒适的感觉的情况下校正子内容的再现时间(再现位置)。

[0297] <同步校正信息生成处理的说明>

[0298] 此外,将参照图18中的流程图给出与图16中的步骤S105中的处理对应的同步校正信息生成处理的描述。

[0299] 在步骤S161中,帧率转换单元181和帧率转换单元182根据需要执行帧率转换处理。

[0300] 也就是说,帧率转换单元181从缓冲器122读取针对主内容的每个时间段的声音同步特征量的时间序列数据,执行帧率转换,即根据需要对声音同步特征量进行下采样,以及将声音同步特征量供给至块合并单元183。此外,帧率转换单元182从缓冲器125读取针对主内容的每个时间段的声音同步特征量的时间序列数据,执行帧率转换,即根据需要对声音同步特征量进行下采样或上采样,以及将声音同步特征量供给至块合并单元184。

[0301] 在步骤S162中,块合并单元183和块合并单元184合并声音同步特征量的时间序列数据。

[0302] 特别地,块合并单元183接收来自帧率转换单元181的针对主内容的每个时间段的声音同步特征量的时间序列数据的供给。然后,块合并单元183将多个(例如,六十四个)连续的时间段作为一个块,合并所供给的针对块单元的每个时间段的声音同步特征量的时间序列数据,以及将时间序列数据供给至相似度计算单元185。

[0303] 此外,块合并单元184接收来自帧率转换单元182的针对主内容的每个时间段的声音同步特征量的时间序列数据的供给。然后,块合并单元184将多个(例如,六十四个)连续的时间段作为一个块,合并所供给的针对块单元的每个时间段的声音同步特征量的时间序列数据,以及将时间序列数据供给至相似度计算单元185。

[0304] 在步骤S163中,相似度计算单元185计算从块合并单元183和块合并单元184供给的块单元上的声音同步特征量的时间序列数据项之间的相似度,并且生成表示相应块之间的相似度的相似度矩阵。相似度计算单元185将相似度矩阵供给至最优路径搜索单元186。

[0305] 在步骤S164中,最优路径搜索单元186根据从相似度计算单元185供给的相似度矩阵来搜索具有最优相似度的路径,并且基于声音特征量来生成同步校正信息。然后,最优路径搜索单元186将基于声音特征量的同步校正信息供给至再现处理单元127,并且完成同步校正信息生成处理。

[0306] 如上所述,内容再现系统基于根据周期性信息的声音特征量生成同步校正信息,因此,内容再现系统可以以健壮的方式生成同步校正信息。

[0307] 尽管上面的描述给出了主内容的条数是一的情况,但是主内容的条数可以是两条或多条。

[0308] 在这样的情况下,提供设备11中的声音同步特征量计算单元23计算针对多条主内容中的每条主内容的声音同步特征量,并且多路复用处理单元24对一条子内容的子通道信号与多条主内容的声音同步特征量进行多路复用,并且获取子传输信号。此外,输出单元22传输从多条主内容的主通道信号中获取的主传输信号。

[0309] 此外,在这种情况下,图4中所示的内容再现系统中的再现处理单元112选择并且再现多条主内容中的一条主内容。此外,输入单元123接收子接收信号,其中,多条主内容的声音同步特征量与单个子通道信号相关联。

[0310] 此外,同步计算单元126通过比较由输入单元123获取的每条主内容的声音同步特征量与通过声音同步特征量计算单元121获取的声音同步特征量来计算相似度,并且指定由扬声器83再现的主内容。例如,作为与声音同步特征量相匹配的结果的含有具有最高相似度的声音同步特征量的主内容被认为是正在再现的主内容。

[0311] 如果指定被正在再现的主内容,则基于根据所指定的主内容的声音同步特征量获取的同步校正信息来对子内容的再现位置进行校正。也就是说,同步计算单元126基于用于同步所指定的主内容与子内容的声音特征量来生成同步校正信息。

[0312] <现有技术的应用示例1>

[0313] 到目前为止描述的本技术可以应用到各种系统。

[0314] 例如,本技术可以应用到图19中所示的系统。

[0315] 例如,在图19所示的系统中,提供设备211例如广播站与图1中的提供设备11对应。

提供设备211提供主内容和子内容。

[0316] 在该示例中,例如,提供设备211通过由广播波来广播主内容的主传输信号来将主传输信号传输至主接收装置212。然后,主接收装置212接收由广播波传输的主传输信号作为主接收信号,并且再现主内容。此时,主接收装置212从在主接收装置212中设置的扬声器213输出主内容的声音。

[0317] 因此,在该示例中,主接收装置212配置有如图4中所示的主接收装置81、显示单元82以及扬声器83。在这样的情况下,输入单元111接收由广播波广播的主接收信号。此外,扬声器213与图4中的扬声器83对应。

[0318] 例如,主接收装置212是电视接收器等,并且用户观看并且收听由主接收装置212再现的主内容。

[0319] 反之,提供设备211还传输子传输信号。在该示例中,提供设备211经由包括因特网的通信网络214通过流分布等将子传输信号传输至子接收装置215。在此,通过所谓的推送型通信来传输子传输信号。

[0320] 此外,例如,子接收装置215配置有平板终端装置,子接收装置215接收通过通信网络214传输的子传输信号作为子接收信号,并且再现子内容。也就是说,子接收装置215使内置式显示单元显示子内容的图像,并且使内置式扬声器输出子内容的声音。

[0321] 此时,子接收装置215通过收集从扬声器213输出的主内容的声音来计算声音同步特征量,并且通过使用所获取的声音同步特征量和包含在子接收信号中的声音同步特征量来生成基于声音特征量的同步校正信息。然后,子接收装置215通过使用基于声音特征量的同步校正信息来以与主内容同步的方式再现子内容。

[0322] 在这样配置的情况下,由主接收装置212再现的主内容和由子接收装置215再现的子内容在同步状态下被再现,并且用户可以观看和收听主内容,同时适当地观看和收听子内容。也就是说,例如,用户可以欣赏主内容,同时利用子内容作为主内容的补充信息。

[0323] 在该示例中,子内容的子通道信号是与主内容的视频的图像信号的角度不同的视频的图像信号、关于主内容的评论声音的声音信号、与主内容有关的字符信息等。

[0324] 在这样的情况下,例如,子接收装置215配置有如图4中所示的麦克风84、子接收装置85、显示单元86以及扬声器87。因此,输入单元123接收经由通信网络214传输的子传输信号作为子接收信号。

[0325] 在图19的示例中,可以通过如上所述的在接收侧的推送型通信来以同步的方式简单并且准确地再现经由彼此不同的传输路径传输的主内容和子内容。在该示例中,需要在主传输信号之前传输子传输信号。也就是说,需要在考虑主传输信号至主接收装置212的到达时间与子传输信号至子接收装置215的到达时间之间的差所确定的时间差的情况下传输主传输信号和子传输信号。

[0326] <本技术的应用示例2>

[0327] 此外,例如,本技术可以应用到图20中所示的系统。在图20中,对与图19中的部件对应的部件给出相同的附图标记,并且将省略该部件的描述。

[0328] 在图20的示例中,主传输信号以与图19的示例相同的方式通过来自提供设备211的广播波即经由推送型通信传输至主接收装置212。

[0329] 反之,子传输信号经由通信网络214从服务器241传输至子接收装置215。此外,服

务器241通过确定的方法预先从提供设备211等获取子传输信号,并且在服务器241上记录子传输信号。

[0330] 在该示例中,子传输信号经由所谓的推送型通信传输。因此,服务器241在子接收装置215请求子传输信号的传输时将子传输信号经由通信网络214传输至子接收装置215。

[0331] 也就是说,图4中与子接收装置215对应的子接收装置85中的输入单元123将针对子传输信号的传输的请求传输至服务器241,并且响应于该传输请求接收从服务器241传输的子传输信号作为子接收信号。

[0332] 在这样的情况下,子接收装置215可以在广播主内容之前预先接收并且记录子传输信号。因此,如果预先接收并且记录子传输信号,则可以防止在主内容被广播时由于通信网络214等的条件不能以与主内容同步的方式再现子内容的情况。

[0333] 如果由主接收装置212开始主内容的再现,则子接收装置215收集从扬声器213输出的主内容的声音,并且计算声音同步特征量。然后,子接收装置215通过使用所获取的声音同步特征量和包含在子接收信号中的声音同步特征量来生成基于声音特征量的同步校正信息,并且通过使用同步校正信息以与主内容同步的方式再现子内容。

[0334] 在图20的示例中,子接收装置215可以在对于子接收装置215自身便利的定时处获取子接收信号。

[0335] <本技术的应用示例3>

[0336] 此外,例如,本技术可以应用到图21中所示的系统。在图21中,对与图20中的部件对应的部件给出相同的附图标记,并且将省略该部件的描述。

[0337] 在图21的示例中,由与服务器241不同的服务器271提供主内容即主传输信号。也就是说,服务器271响应于来自主接收装置212的请求将所记录的主传输信号经由通信网络272传输至主接收装置212。也就是说,在该示例中,通过推送型通信来传输主传输信号。

[0338] 特别地,图4中与主接收装置212对应的主接收装置81中的输入单元111将主传输信号的传输请求传输至服务器271,并且响应于该传输请求接收从服务器271传输的主传输信号作为主接收信号。

[0339] 在这样的情况下,主接收装置212可以预先接收并且记录主传输信号。因此,如果预先接收并且记录主传输信号,则可以防止在再现主内容时由于通信网络272等的条件使主内容的再现暂停或停止的情况。

[0340] 此外,由服务器241以与图20的示例中的方式相同的方式经由推送型通信传输子传输信号。

[0341] 如果由主接收装置212开始主内容的再现,则子接收装置215收集从扬声器213输出的主内容的声音,并且计算声音同步特征量。然后,子接收装置215通过使用所获取的声音同步特征量和包含在子接收信号中的声音同步特征量来生成基于声音特征量的同步校正信息,并且通过使用同步校正信息以与主内容同步的方式再现子内容。

[0342] 在图21的示例中,主接收装置212和子接收装置215可以在对于主接收装置212和子接收装置215自身便利的定时处获取主接收信号和子接收信号。

[0343] 即使当通信网络272与通信网络214相同时,如果主传输信号的传输定时、接收装置等和子传输信号的传输定时、接收装置等彼此不同,则主传输信号的传输路径与子传输信号的传输路径通常彼此不同。

[0344] <第二实施方式>

[0345] <本技术的特征>

[0346] 顺便提及地,可以采用下述应用程序,根据所述应用程序,由多个装置通过广播、因特网协议(IP)网络等接收具有时间同步关系的多条媒体内容,并且按照前述示例(A1)至(A4)的同步方式再现所接收的媒体内容。

[0347] 为了实现这样的功能,已经研究和开发了用于广播通信协作服务的系统,在该系统中,通过广播发布内容同时经由IP网络例如经由混合广播来单独地发布另外的内容,并且内容接收器以时间同步的方式同时地输出由广播发布的内容和经由IP网络发布的另外的内容。

[0348] 例如,在“Kinji Matsumura, Yoshiaki Shishikui and Michael J. Evans, “Personalization System of Broadcast Content using Synchronized Internet Streaming” in Proceedings of the Institute of Image Information and Television Engineers annual convention, 2009年8月26日,第3页至第8页”(在下文中,也被称为NPL 1)中描述了混合广播。

[0349] 此外,还在下述文献中描述了混合广播:“Japan Broadcasting Corporation, “Summary and Technology of HybridcastTM”, Science and Technology Research Laboratories of Japan Broadcasting Corporation, R&D, No.124, 第10页至第17页, 2010年11月, NHK Publishing, Inc., <http://www.nhk.or.jp/strl/publica/rd/rd124/PDF/P10-17.pdf>”(在下文中,也被称为NPL 2), “Japan Broadcasting Corporation, “Technologies Supporting HybridcastTM”, Science and Technology Research Laboratories of Japan Broadcasting Corporation, R&D, No.133, 第20页至第27页, 2012年5月, NHK Publishing, Inc., <http://www.nhk.or.jp/strl/publica/rd/rd133/PDF/P20-27.pdf>”(在下文中,也被称为NPL 3)等。

[0350] 根据混合广播的基本原理,增加基于广播流的参考时钟的呈现时间信息(呈现时间戳(PTS))的另外的内容是与广播内容的传输同时或者稍微在广播内容的传输之前流分发的,接收器设置有具有用于吸收通信内容的延迟和变化的足够容量的缓冲器,并且通过使广播内容延迟和比较广播内容的时间戳与另外的内容的时间戳来建立同步。

[0351] 例如,根据NPL 2,确认在接收器处于相同装置的实验环境下在大约一个视频帧(33ms)内的精度的情况下建立同步。

[0352] 接收另外的内容的装置可以是独立于广播内容接收装置的装置,例如以无线方式与IP网络连接的智能电话或平板个人计算机。在这样的情况下,对于广播内容接收装置需要将呈现时间信息(时间戳)提供至接收另外的内容的装置。该操作通常经由IP网络来共同起作用。

[0353] 此外,可以容易地设想实现下述系统,该系统仅经由网络例如IP网络代替广播发布多条内容,通过使用协调世界时间(UTC)作为参考时钟来增加时间戳,在接收器侧执行同步,以及输出多条内容。

[0354] 当实践中通过独立的接收器使用这样的广播通信协作服务时,由于下面两个原因而难以通过比较时间戳的方法来精确地建立同步。

[0355] 首先,只要广播内容接收器和另外的内容接收器是独立的电子装置,那么系统时

钟上就存在差异,以及随着时间而发生同步偏差。

[0356] 其次,可能呈现下述使用状态:用户位于距广播内容接收器例如电视接收器一定程度上的距离,通过用户的手持有另外的内容接收器例如智能电话或平板个人计算机,并且欣赏经由IP网络分发的另外的内容。如果在该使用状态下广播内容和另外的内容中包含声音信号,则难以在用户观看和收听内容的位置处建立精确的同步。

[0357] 例如,当用户距广播内容接收器10m时,对于从广播内容接收器输出的声音信号到达用户位置需要 $10\text{ (m)} / 340\text{ (m/s)} = \text{大约} 30\text{ (ms)}$ 。在此,声速是大约340 (m/s)。

[0358] 第一实施方式采用下述方法:通过另外的内容接收器收集从广播内容接收器输出的声音,计算声音同步特征量,并且用经由IP网络分发的广播内容的声音同步特征量执行同步计算。然而,需要在宽的范围上搜索同步位置,并且当IP网络的传输延迟、波动等严重时,处理量增加。

[0359] 因此,可以通过针对前述提供设备和内容再现系统提供下面的特征B11至特征B20来对经由具有更少处理量的不同路径获取的多条内容进行同步。

[0360] (特征B11)

[0361] 媒体内容是通过对视频、声音、图像、字符信息等进行多路复用而获取的数据流。

[0362] 对于(特征B11)的数据流的传输,采用通过广播波或经由包括因特网的网络的媒体内容的传输,以及由多路复用数据流占用的逻辑传输路径将被称为传输路径。

[0363] (特征B12)

[0364] 作为传输目标的多条媒体内容具有时间同步关系。

[0365] (特征B13)

[0366] 作为传输目标的多条媒体内容中的至少一条被确定为主通道信号,并且其余的各条媒体内容相应地被当做子通道信号。

[0367] (特征B14)

[0368] 根据参考时间信号针对主通道信号中的每个主通道信号和子通道信号中的每个子通道信号来生成呈现时间信息(PTC)。

[0369] 在此,广播流的参考时钟(PCR)、协调世界时间(UTC)等被用作参考时间信号。

[0370] (特征B15)

[0371] 主通道信号的呈现时间信息与主通道信号进行多路复用以生成主传输信号,并且主传输信号被传输。反之,还根据主通道信号的声音信号来计算声音同步特征量。

[0372] (特征B16)

[0373] 以由系统限定的传输格式来执行子通道信号的呈现时间信息、主通道信号的声音同步特征量以及子通道信号的多路复用处理,使得满足主通道信号与子通道信号之间的时间同步关系,并且生成子传输信号。

[0374] (特征B17)

[0375] 主接收装置获取并且分离主接收信号,并且当再现主通道信号时通过扬声器等输出基于声音信号的声音。同时,主接收装置对所接收的主通道信号的呈现时间信息进行呈现使得可以从外部参考或者获取呈现时间信息。

[0376] 例如,可以由其中提供获取功能的软件的应用程序接口(API)通过无线通信经由IP网络连接从外部参考主通道信号的呈现时间信息。

[0377] (特征B18)

[0378] 子接收装置获取并且分离子接收信号,比较所接收的子通道信号的呈现时间信息与从主接收装置获取的主通道信号的呈现时间信息,并且基于呈现时间信息来生成同步校正信息。

[0379] (特征B19)

[0380] 子接收装置通过麦克风等收集由主接收装置从扬声器输出的主通道信号的声音,计算声音同步特征量,考虑在(特征B18)中生成的基于呈现时间信息的同步校正信息的情况下相对于所接收的主通道信号的声音同步特征量执行自动同步计算,并且基于声音特征量计算同步校正信息(时间差信息)。

[0381] 由于可以根据通过比较呈现时间信息获取的基于呈现时间信息的同步校正信息来识别同步位置,则在稍后阶段的针对基于声音同步特征量的自动同步计算处理需要的处理量会减少。

[0382] (特征B20)

[0383] 子接收装置根据基于声音特征量的同步校正信息对所接收的子通道信号相对于主通道信号执行同步校正处理,并且执行再现。

[0384] <提供设备的配置示例>

[0385] 接下来,将给出具有前述特征B11至B20的提供设备和内容再现系统的特定实施方式的描述。

[0386] 图22是示出了按照前述示例(A1)至(A4)中的时间同步关系提供内容的提供设备的配置示例的图。在图22中,对与图1中的部件对应的部件给出相同的附图标记,并且将适当地省略该部件的描述。

[0387] 提供设备301包括:参考时间信号生成单元311、多路复用处理单元312、输出单元22、声音同步特征量计算单元23、多路复用处理单元24以及输出单元25。

[0388] 提供设备301的配置与提供设备11的配置的不同之处在于用多路复用处理单元312替换提供设备11中的转换单元21,并且新提供了参考时间信号生成单元311。

[0389] 参考时间信号生成单元311基于PCR和UTC生成指示主通道信号的内容呈现定时和子通道信号的内容呈现定时的呈现时间信息,并且将呈现时间信息供给至多路复用处理单元312和多路复用处理单元24。例如,呈现时间信息是PTS,并且呈现时间信息用于在重现侧建立主通道信号与子通道信号之间的同步。

[0390] 多路复用处理单元312将所供给的主通道信号转换成以由预定广播标准等限定的格式的信号。此外,多路复用处理单元312通过对以经转换的格式的主通道信号与从参考时间信号生成单元311供给的呈现时间信息进行多路复用生成主传输信号,并且将该主传输信号供给至输出单元22。包含在主传输信号中的呈现时间信息是主通道信号的呈现时间信息。

[0391] 此外,多路复用处理单元24在时间同步的状态下对从声音同步特征量计算单元23供给的声音同步特征量、所供给的子通道信号以及从参考时间信号生成单元311供给的呈现时间信息进行多路复用,然后按照需要执行格式转换,并且生成子传输信号。多路复用处理单元24将所获取的子传输信号供给至输出单元25。包含在子传输信号中的呈现时间信息是子通道信号的呈现时间信息。

[0392] 以与提供设备11相同的方式,多路复用处理单元24可以通过使用主通道信号来对声音同步特征量、子通道信号以及呈现时间信息的时间同步关系进行调整。

[0393] <内容再现系统的配置示例>

[0394] 此外,例如接收从提供设备301传输的主传输信号和子传输信号作为主接收信号和子接收信号并且再现主内容和子内容的内容再现系统被配置为如图23中所示。在图23中,对与图4中的部件对应的部件给出相同的附图标记,并且将适当地省略该部件的描述。

[0395] 图23中所示的内容再现系统包括:主接收装置341、显示单元82、扬声器83、麦克风84、子接收装置342、显示单元86以及扬声器87。

[0396] 主接收装置341接收从提供设备301传输的主接收信号,并且控制从主接收信号获取的主内容的再现。

[0397] 主接收装置341包括:输入单元111、分离处理单元351、呈现单元352以及再现处理单元112。主接收装置341的配置与主接收装置81的配置的不同之处在于新提供了分离处理单元351和呈现单元352。

[0398] 分离处理单元351将从输入单元111供给的主接收信号分离成主通道信号和主通道信号的呈现时间信息,将主通道信号供给至再现处理单元112,并且将呈现时间信息供给至呈现单元352。

[0399] 呈现单元352将从分离处理单元351供给的呈现时间信息经由包括互联网的有线通信网络或无线通信网络呈现至子接收装置342。也就是说,响应于来自通信对方的请求来传输呈现时间信息。

[0400] 此外,子接收装置342接收从提供设备301传输的子传输信号作为子接收信号,并且控制从子接收信号获取的子内容的再现。

[0401] 子接收装置342包括:获取单元361、呈现时间信息比较单元362、声音同步特征量计算单元121、缓冲器122、输入单元123、分离处理单元124、缓冲器125、同步计算单元126以及再现处理单元127。

[0402] 子接收装置342的配置与子接收装置85的配置的不同之处在于新提供了获取单元361和呈现时间信息比较单元362。

[0403] 获取单元361通过使用API等经由有线或无线通信网络获取由呈现单元352呈现的呈现时间信息,并且将呈现时间信息供给至呈现时间信息比较单元362。也就是说,获取单元361接收从呈现单元352传输的呈现时间信息。

[0404] 分离处理单元124将从输入单元123供给的子接收信号分离成声音同步特征量、子通道信号以及呈现时间信息,将呈现时间信息供给至呈现时间信息比较单元362,并且将声音同步特征量和子通道信号供给至缓冲器125。

[0405] 呈现时间信息比较单元362比较从分离处理单元124供给的呈现时间信息与从获取单元361供给的呈现时间信息,基于用于同步主通道信号与子通道信号的呈现时间信息来生成同步校正信息,并且将同步校正信息供给至同步计算单元126。

[0406] 基于呈现时间信息的同步校正信息可以对主通道信号与子通道信号之间的偏差进行校正,并且通过基于呈现时间信息的同步校正信息本身建立主通道信号与子通道信号之间的同步。然而,在该示例中,由同步计算单元126使用基于呈现时间信息的同步校正信息以确定待从缓冲器125中读取的声音同步特征量的范围以更精确地同步信号。换言之,利

用基于呈现时间信息的同步校正信息来大致地对记录在缓冲器125中的声音同步特征量与记录在缓冲器122中的声音同步特征量进行同步。通过使用如上所述的基于呈现时间信息的同步校正信息,可以用较少的处理量来执行声音同步特征量的匹配处理。

[0407] <传输处理的说明>

[0408] 接下来,将给出上述的提供设备301的特定操作和内容再现系统的特定操作的描述。

[0409] 首先,将参照图24中的流程图来给出由提供设备301执行的传输处理的描述。

[0410] 在步骤S191中,参考时间信号生成单元311生成主通道信号的呈现时间信息和子通道信号的呈现时间信息,并且将呈现时间信息供给至多路复用处理单元312和多路复用处理单元24。

[0411] 在步骤S192中,声音同步特征量计算单元23执行声音同步特征量计算处理以计算来自构成所供给的主通道信号的声音信号的声音同步特征量,并且将声音同步特征量供给至多路复用处理单元24。此外,由于在步骤S192中执行的声音同步特征量计算处理与参照图14描述的声音同步特征量计算处理相同,所以将省略其中的描述。

[0412] 在步骤S193中,多路复用处理单元312通过对所供给的主通道信号与从参考时间信号生成单元311供给的呈现时间信息进行多路复用来生成主传输信号,并且将该主传输信号供给至输出单元22。此时,多路复用处理单元312根据需要对主通道信号执行格式转换。

[0413] 在步骤S194中,输出单元22传输从多路复用处理单元312供给的主传输信号。

[0414] 在步骤S195中,多路复用处理单元24通过对声音同步特征量、子通道信号以及呈现时间信息进行多路复用来生成子传输信号,并且将该子传输信号供给至输出单元25。

[0415] 也就是说,多路复用处理单元24对来自声音同步特征量计算单元23的声音同步特征量、所供给的子通道信号以及从参考时间信号生成单元311供给的呈现时间信息进行多路复用以获取子传输信号。

[0416] 在步骤S196中,输出单元25传输从多路复用处理单元24供给的子传输信号,并且完成传输处理。

[0417] 如上所述,提供设备301生成共同地用于主通道信号和子通道信号的呈现时间信息,并且生成包含呈现时间信息的主传输信号和子传输信号。

[0418] 在这样配置的情况下,可以通过在内容再现侧使用呈现时间信息来用较少的处理量对主内容和子内容进行同步。

[0419] <主内容再现处理的说明>

[0420] 如果从提供设备301传输主传输信号,则内容再现系统获取主传输信号作为主接收信号,并且再现主内容。在下文中,将参照图25中的流程图来给出由内容再现系统进行的主要内容再现处理的描述。

[0421] 在步骤S221中,输入单元111获取主接收信号,并且将主接收信号供给至分离处理单元351。例如,输入单元111通过接收从提供设备301传输的主接收信号来获取主接收信号。

[0422] 在步骤S222中,分离处理单元351将从输入单元111供给的主接收信号分离成主通道信号和呈现时间信息。分离处理单元351将分离之后的主通道信号供给至再现处理单元

112,并且将呈现时间信息供给至呈现单元352。

[0423] 在步骤S223中,再现处理单元112基于从分离处理单元351供给的主通道信号来再现主内容。在步骤S223中,执行与图15中的步骤S72的处理相同的处理。

[0424] 在步骤S224中,呈现单元352呈现从分离处理单元351供给的呈现时间信息,并且完成主内容再现处理。例如,在与主内容的再现同步的状态下,以无线的方式将呈现时间信息传输至子接收装置342。

[0425] 如上所述,内容再现系统获取主接收信号,再现主内容,并且呈现主内容即主通道信号的再现时间信息。

[0426] 通过与如上所述的主内容的再现同时地呈现主内容的呈现时间信息,获取呈现时间信息的子接收装置342可以在较少处理量的情况下通过使用声音同步特征量来执行同步计算。

[0427] <子内容再现处理的说明>

[0428] 在与主内容的再现的同步中,内容再现系统获取子接收信号,并且再现子内容。在下文中,将参照图26中的流程图来给出由内容再现系统进行的子内容再现处理的描述。

[0429] 由于步骤S251中的处理与图16中的步骤S101中的处理相同,所以将省略其中的描述。

[0430] 在步骤S252中,分离处理单元124将从输入单元123供给的子接收信号分成子通道信号、声音同步特征量以及呈现时间信息。然后,分离处理单元124将子通道信号和声音同步特征量供给至缓冲器125,使缓冲器125在缓冲器125中记录子通道信号和声音同步特征量,并且将子通道信号的呈现时间信息供给至呈现时间信息比较单元362。

[0431] 在步骤S253中,获取单元361通过接收从呈现单元352传输的呈现时间信息来获取主通道信号的呈现时间信息,并且将呈现时间信息供给至呈现时间信息比较单元362。

[0432] 在步骤S254中,呈现时间信息比较单元362比较从分离处理单元124供给的呈现时间信息与从获取单元361供给的呈现时间信息,基于呈现时间信息来生成同步校正信息,并且将同步校正信息供给至同步计算单元126。

[0433] 例如,基于呈现时间信息的同步校正信息是下述信息:指示在缓冲器125中以时间序列方式记录的相应时间处的声音同步特征量的序列中的作为由同步计算单元126进行的同步计算的目标的范围(在下文中,还被称为搜索范围)。

[0434] 该搜索范围与具有预定长度的声音同步特征量的序列对应,该声音同步特征量的序列包括与指示和当前被再现的主内容的时间相同的时间的子通道信号的呈现时间信息即由获取单元361获取的最新的呈现时间信息关联的声音同步特征量。

[0435] 在与相同呈现时间对应的主通道信号中的位置和子通道信号中的位置是彼此同步的信号位置,即待被同时再现的再现位置(帧位置)。因此,可以通过比较呈现时间信息和检测在子通道信号中与主通道信号的呈现时间信息相同的呈现时间信息对应的位置来指定在子内容中大致上与正在被再现的主内容同步的再现位置。

[0436] 如果通过比较呈现时间信息来生成基于呈现时间信息的同步校正信息,则然后执行在步骤S255和步骤S256中的处理。然而,由于该处理与图16中的步骤S103和步骤S104中的处理相同,因此将省略其中的描述。在该处理中,收集主内容的声音,并且根据该声音计算声音同步特征量。

[0437] 在步骤S257中,同步计算单元126基于声音特征量执行同步校正信息生成处理以生成同步校正信息,并且将同步校正信息供给至再现处理单元127。稍后将给出同步校正信息生成处理的详细描述。在该处理中,通过使用基于呈现时间信息的同步校正信息来比较在缓冲器122中记录的声音同步特征量与在缓冲器125中记录的声音同步特征量,并且生成基于声音特征量的同步校正信息。

[0438] 在步骤S258中,再现处理单元127根据从同步计算单元126供给的基于声音特征量的同步校正信息对在缓冲器125中记录的子通道的再现定时进行校正,并且基于校正之后的子通道信号来再现子内容。在步骤S258中,执行与图16中的步骤S106中的处理相同的处理。

[0439] 如果如上所述以与主内容同步的方式再现子内容,则完成了子内容再现处理。

[0440] 内容再现系统通过获取主内容的呈现时间信息以及比较主内容的呈现时间信息与包含在如上所述的子接收信号中的子内容的呈现时间信息来生成基于呈现时间信息的同步校正信息。然后,内容再现系统对包含在作为目标的由基于呈现时间信息的同步校正信息指示的搜索范围中的声音同步特征量执行匹配处理,并且基于声音特征量计算同步校正信息。

[0441] 在这样配置的情况下,即使在主内容和子内容经由不同传输路径传输的情况下也可以用较少的处理量来计算同步校正信息,并且以同步的方式再现内容。

[0442] 也就是说,根据内容再现系统,子接收装置342首先基于呈现时间信息大致上对主通道信号与子通道信号进行同步,并且计算来自通过收集主内容的声音获取的声音信号的声音同步特征量。

[0443] 然后,子接收装置342可以通过相对于所接收的主通道信号的声音同步特征量执行自动同步计算来在正在观看和收听由子接收装置342再现的子内容的用户的位置处执行精确的内容同步。由于实践中呈现时间信息用于限制同步位置的近似范围,因此针对基于声音同步特征量的自动同步计算处理需要的处理量减少。

[0444] 例如,假定在内容再现系统中的显示单元82和扬声器83布置在远离显示单元86和扬声器87的位置处,并且用户在显示单元86和扬声器87的附近观看和收听内容。在这样的情况下,对于从扬声器83输出的声音到达用户正在观看和收听内容的位置会花费一定程度上的时间。

[0445] 因此,在这样的情况下,难以仅仅通过比较呈现时间信息来精确地对用户正在观看和收听内容的位置处的主内容的再现和子内容的再现进行同步。也就是说,例如,由于扬声器83和扬声器87同时分别再现主内容的声音和子内容的声音,因此如果主内容的声音到达用户花费时间,则用户在有偏差的情况下听到主内容的声音和子内容的声音。

[0446] 反之,根据应用本技术的内容再现系统,由连接至子接收装置342并且布置在子接收装置342附近的麦克风84来收集主内容的声音,并且执行同步计算。因为这个原因,内容再现系统可以在用户正在观看和收听内容的位置处在同步的状态下再现主内容和子内容。此外,内容再现系统可以通过比较呈现时间信息、生成基于呈现时间信息的同步校正信息以及限制用于匹配处理的搜索范围来用较小的处理量对内容进行同步。

[0447] <同步校正信息生成处理的说明>

[0448] 此外,将参照图27中的流程图给出与图26中的步骤S257中的处理对应的同步校正

信息生成处理的描述。

[0449] 在步骤S281中,帧率转换单元181和帧率转换单元182根据需要执行帧率转换处理。

[0450] 也就是说,帧率转换单元181从缓冲器122读取针对主内容的每个时间段的声音同步特征量的时间序列数据,根据需要执行帧率转换,即对声音同步特征量进行下采样,以及将声音同步特征量供给至块合并单元183。

[0451] 此外,帧率转换单元182仅在缓冲器125中记录的声音同步特征量的时间序列数据中读取包含在由基于呈现时间信息的同步校正信息指示的搜索范围中的时间序列数据,该搜索范围由呈现时间信息比较单元362供给。

[0452] 然后,帧率转换单元182根据需要对读取的声音同步特征量执行帧率转换即下采样或上采样,并且将声音同步特征量供给至块合并单元184。

[0453] 在步骤S282中,块合并单元183和块合并单元184合并声音同步特征量的时间序列数据。

[0454] 特别地,块合并单元183从帧率转换单元181接收针对主内容的每个时间段的声音同步特征量的时间序列数据的供给。然后,块合并单元183将多个(例如,六十四个)连续的时间段作为一个块,将所供给的块单元中的每个时间段上的声音同步特征量的时间序列数据进行合并,并且将所合并的时间序列数据供给至相似度计算单元185。

[0455] 此外,块合并单元184从帧率转换单元182接收主内容的每个时间段的声音同步特征量的时间序列数据的供给。然后,块合并单元184将多个(例如,六十四个)连续的时间段作为一个块,将所供给的块单元中的每个时间段的声音同步特征量的时间序列数据进行合并,并且将所合并的时间序列数据供给至相似度计算单元185。

[0456] 例如,由图28中的箭头A11所示,记录在缓冲器122中的n个块X(i)和记录在缓冲器125中的m个块Y(j)是图18中的步骤S162中的处理的处理目标。也就是说,作为搜索目标的声音同步特征量的块的组合是n×m个组合。更具体地,适当地对声音同步特征量执行帧率转换。然而,为了简化图28的说明,将在不执行帧率转换的前提下继续描述。

[0457] 在此,作为匹配处理的目标的块Y(j)是在缓冲器125中记录的所有块或在足够宽的范围上的块。

[0458] 在图28中,i是通过声音同步特征量计算单元121获取的声音同步特征量的块的指标,并且j是包括在子接收信号中的声音同步特征量的块的指标。

[0459] 此外,X(i)表示来自通过声音同步特征量计算单元121获取的声音同步特征量中的具有指标i的块上的声音同步特征量的时间序列数据,并且Y(j)表示来自包括在子接收信号中的声音同步特征量中的具有指标j的块上的声音同步特征量的时间序列数据。

[0460] 反之,在步骤S282中,仅来自记录在缓冲器125中的m个块中的包括在由基于呈现时间信息的同步校正信息指示的搜索范围中的m'个块是由箭头A12表示的匹配处理的目标。也就是说,仅m'个块是由相似度计算单元185进行的相似度计算的目标。

[0461] 在该示例中,PTS_i表示呈现时间信息,并且由呈现时间信息表示的位置与当前被再现的主内容中的位置对应。此外,具有包括与主内容的呈现时间信息的时间相同的时间处的子内容的呈现时间信息对应的位置的预定长度的范围即配置有m'个块的范围是搜索范围。因此,作为搜索目标的声音同步特征量的块的组合是n×m'个组合。

[0462] 可以通过使用通过比较呈现时间信息获取的基于呈现时间信息的同步校正信息来使作为匹配处理的目标的声音同步特征量的范围最小化,从而显著地减少用于相似度计算的搜索所需要的处理时间。

[0463] 返回图27中的流程图的说明,如果声音同步特征量的时间序列数据被合并,则然后处理进行至步骤S283。虽然执行步骤S283和步骤S284中的处理,并且完成同步校正信息生成处理,但是该处理与图18中的步骤S163和S164中的处理相同,因此将省略其中的描述。如果完成了同步校正信息生成处理,则然后处理进行至图26中的步骤S258。

[0464] 如上所述,内容再现系统通过使用由基于呈现时间信息的同步校正信息指示的搜索范围内的声音同步特征量来生成基于声音特征量的同步校正信息。在这样配置的情况下,可以用较少的处理量以健壮的方式生成同步校正信息。

[0465] 此外,在图22中所示的提供设备301和在图23中所示的内容再现系统还可以应用到图19至图21所示的相应的系统。

[0466] 顺便提及地,可以由硬件或软件来执行前述系列的处理。在通过软件执行该系列处理的情况下,配置软件的程序安装在计算机中。在此,计算机的示例包括包含在专用硬件中的计算机和能够通过安装各种程序来执行各种功能的通用目的的计算机。

[0467] 图29是示出了在通过软件执行前述系列处理的计算机中的硬件的配置示例的框图。

[0468] 在该计算机中,中央处理单元(CPU)501、只读存储器(ROM)502以及随机存取存储器(RAM)503经由总线504互相连接。

[0469] 此外,输入与输出接口505连接至总线504。此外,输入单元506、输出单元507、记录单元508、通信单元509以及驱动器510连接至输入与输出接口505。

[0470] 输入单元506配置由键盘、鼠标、麦克风、成像装置等。输出单元507配置有显示器、扬声器等。记录单元508配置有硬盘、非易失性储存器等。通信单元509配置有网络接口等。驱动器510驱动可移除介质511例如磁盘、光盘、磁光盘或半导体存储器。

[0471] 例如,如上所述配置的计算机通过CPU 501将记录在记录单元508中的程序经由输入与输出接口505以及总线504加载到RAM 503上并且执行该程序来执行前述序列处理。

[0472] 例如,由计算机(CPU 501)执行的程序可以记录在作为程序包介质的可移除介质511中,以及可以提供由计算机(CPU 501)执行的程序。此外,可以经由有线或无线传输介质例如局域网、因特网或数字卫星广播来提供程序。

[0473] 计算机可以通过将可移除介质511安装在驱动器510上来经由输入与输出接口505将程序安装在记录单元508中。此外,可以通过在通信单元509处经由有线或无线传输介质接收程序来将程序安装在记录单元508中。此外,程序可以预先安装在ROM 502或记录单元508中。

[0474] 此外,由计算机执行的程序可以是下述程序,根据所述程序以在本文中描述的顺序的时间序列的方式来执行处理,或者可以是下述程序,根据所述程序并行地或在需要的定时例如在访问程序的定时执行处理。

[0475] 此外,本技术的实施方式不限于前述实施方式,并且可以在不偏离本技术主旨的范围内进行各种修改。

[0476] 例如,本技术可以应用由多个设备共享功能并且协作地处理的云计算的配置。

[0477] 此外,可以由一个设备执行在前述流程图中描述的相应的步骤,或者可以由多个设备共享并且执行在前述流程图中描述的相应的步骤。

[0478] 此外,当在一个步骤中包括多个过程时,可以由一个设备执行包括在该步骤中的多个过程,或者可以由多个设备共享并且执行包括在该步骤中的多个过程。

[0479] 此外,本文中描述的效果仅是为了说明的目的而举例说明,而不意为呈现该效果的限制,并且可以获得其他效果。

[0480] 此外,本技术还可以被如下地配置。

[0481] (1)一种信息处理设备,用于以与由不同于所述信息处理设备的第二信息处理设备进行的第一内容的再现同步的方式来再现第二内容,所述第一内容包括音频内容,所述信息处理设备包括:

[0482] 电路,被配置成:

[0483] 从所述音频内容中提取第一特征;

[0484] 获得所述音频内容的第二特征,所述第二特征与所述第二内容一起;

[0485] 比较所述第一特征与所述第二特征;以及

[0486] 基于所述比较的结果生成用于以与所述第一内容同步的方式来再现所述第二内容的同步信息。

[0487] (2)根据(1)所述的信息处理设备,其中,所述电路被配置成:

[0488] 从经由通过所述第二信息处理设备对所述音频内容的再现生成的声音中获取音频信号;以及从所述音频信号中提取所述第一特征。

[0489] (3)根据(1)或(2)所述的信息处理设备,其中,所述第二内容和所述第二特征在基于所述第一内容的到达时间与所述第二内容和所述第二特征的到达时间之间的差而确定的时间处被传输至所述信息处理设备。

[0490] (4)根据(1)所述的信息处理设备,其中,所述电路被配置成:

[0491] 请求所述第二内容和所述第二特征的传输;以及

[0492] 接收响应于所述请求而被传输的所述第二内容和所述第二特征。

[0493] (5)根据(4)所述的信息处理设备,其中,所述电路被配置成:

[0494] 请求所述第一内容的传输;以及

[0495] 接收响应于针对所述第一内容的传输的请求而被传输的所述第一内容。

[0496] (6)根据(1)所述的信息处理设备,其中,所述电路被配置成:

[0497] 从所述音频内容中提取第一多个特征,所述第一多个特征包括所述第一特征;

[0498] 获得与所述第二内容一起的第二多个特征,所述第二多个特征包括所述第二特征;

[0499] 比较所述第一多个特征与所述第二多个特征;以及

[0500] 基于比较所述第一多个特征与所述第二多个特征的结果来生成用于以与所述第一内容同步的方式来再现所述第二内容的所述同步信息。

[0501] (7)根据(1)所述的信息处理设备,其中,所述电路被配置成控制所述第二内容的再现。

[0502] (8)根据(7)所述的信息处理设备,其中,所述电路被配置成基于所述同步信息来校正所述第二内容的再现位置。

- [0503] (9) 根据(1)所述的信息处理设备,其中,所述电路被配置成:
- [0504] 获取与所述第一内容关联的呈现时间信息;以及
- [0505] 基于比较与所述第一内容关联的呈现时间信息和与所述第二内容关联的呈现时间信息的结果来获得时间范围,
- [0506] 其中,所述电路被配置成至少部分地通过在所述时间范围中比较所述第二特征与从所述第一内容中出现的所述第一内容的一个或多个部分中提取的一个或多个特征来生成所述同步信息。
- [0507] (10) 根据(1)所述的信息处理设备,其中,所述电路被配置成在对所述第一特征和所述第二特征中的至少一个执行帧率转换使得所述第一特征的帧率与所述第二特征的帧率一致之后比较所述第一特征和所述第二特征。
- [0508] (11) 一种信息处理方法,用于以与第一内容的再现同步的方式来再现第二内容,所述第一内容包括音频内容,所述信息处理方法包括:
- [0509] 从所述音频内容中提取第一特征;
- [0510] 获得所述音频内容的第二特征,所述第二特征与所述第二内容一起;
- [0511] 比较所述第一特征与所述第二特征;以及
- [0512] 基于所述比较的结果生成用于以与所述第一内容同步的方式来再现所述第二内容的同步信息。
- [0513] (12) 至少一种非暂态计算机可读存储介质,其存储有处理器可执行指令,当所述处理器可执行指令由至少一个计算机执行时,使所述至少一个计算机执行用于以与第一内容的再现同步的方式来再现第二内容的信息处理方法,所述第一内容包括音频内容,所述信息处理方法包括:
- [0514] 从所述音频内容中提取第一特征;
- [0515] 获得所述音频内容的第二特征,所述第二特征与所述第二内容一起;
- [0516] 比较所述第一特征与所述第二特征;以及
- [0517] 基于所述比较的结果生成用于以与所述第一内容同步的方式来再现所述第二内容的同步信息。
- [0518] (13) 一种信息处理设备,包括:
- [0519] 电路,被配置成:
- [0520] 从通过第一内容的再现生成的声音中提取特征;
- [0521] 以与所述第一内容同步的方式将所述特征与和所述第一内容具有时间同步关系的第二内容进行关联;以及
- [0522] 输出所述第二内容和所述特征。
- [0523] (14) 根据(13)所述的信息处理设备,其中,所述电路被配置成输出所述第一内容。
- [0524] (15) 根据(14)所述的信息处理设备,其中,所述电路被配置成在基于所述第一内容的到达时间与所述第二内容的到达时间之间的差而确定的时间处输出所述第二内容和所述特征。
- [0525] (16) 根据(14)所述的信息处理设备,其中,所述电路被配置成响应于针对所述第二内容和所述特征的传输的请求来输出所述第二内容和所述特征。
- [0526] (17) 根据(16)所述的信息处理设备,其中,所述电路被配置成响应于针对所述第

一内容的传输的请求来输出所述第一内容。

[0527] (18) 根据 (13) 所述的信息处理设备,其中,所述电路被配置成:

[0528] 从通过再现所述第一内容的多个部分生成的声音中提取多个特征;以及

[0529] 将所提取的特征与所述第二内容相关联并且输出所提取的特征和所述第二内容。

[0530] (19) 根据权利要求 (13) 所述的信息处理设备,其中,所述电路被配置成:

[0531] 对所述特征进行下采样以获得下采样的特征;以及输出所述第二内容和所述下采样的特征。

[0532] (20) 一种信息处理方法,包括:

[0533] 从通过第一内容的再现生成的声音中提取特征;

[0534] 以与所述第一内容同步的方式将所述特征与和所述第一内容具有时间同步关系的第二内容进行关联;以及

[0535] 输出所述第二内容和所述特征。

[0536] (21) 至少一种非暂态计算机可读存储介质,当所述至少一种非暂态计算机可读存储介质由至少一个计算机执行时,使所述至少一个计算机执行:

[0537] 从通过第一内容的再现生成的声音中提取特征;

[0538] 以与所述第一内容同步的方式将所述特征与和所述第一内容具有时间同步关系的第二内容进行关联;以及

[0539] 输出所述第二内容和所述特征。

[0540] 还要认识到的是,本文中描述的本技术的各方面不限于允许在两个装置之间的内容的同步再现。因此,在一些实施方式中,本文中描述的技术可以应用到在多于两个装置(例如,三个、四个、五个、至少五个、至少十个、两个与十个之间等)之间的内容的同步再现。

[0541] 本领域技术人员要理解的是,取决于设计要求和其他因素可能发生各种修改、组合、子组合以及改变,在设计要求和其他因素的范围内,各种修改、组合、子组合以及改变在所附权利要求及其等同性的范围内。

[0542] 附图标记列表

[0543] 11 提供设备

[0544] 22 输出单元

[0545] 23 声音同步特征量计算单元

[0546] 24 多路复用处理单元

[0547] 25 输出单元

[0548] 81 主接收装置

[0549] 85 子接收装置

[0550] 111 输入单元

[0551] 112 再现处理单元

[0552] 121 声音同步特征量计算单元

[0553] 123 输入单元

[0554] 126 同步计算单元

[0555] 127 再现处理单元

[0556] 311 参考时间信号生成单元

- [0557] 352 呈现单元
- [0558] 361 获取单元
- [0559] 362 呈现时间信息比较单元

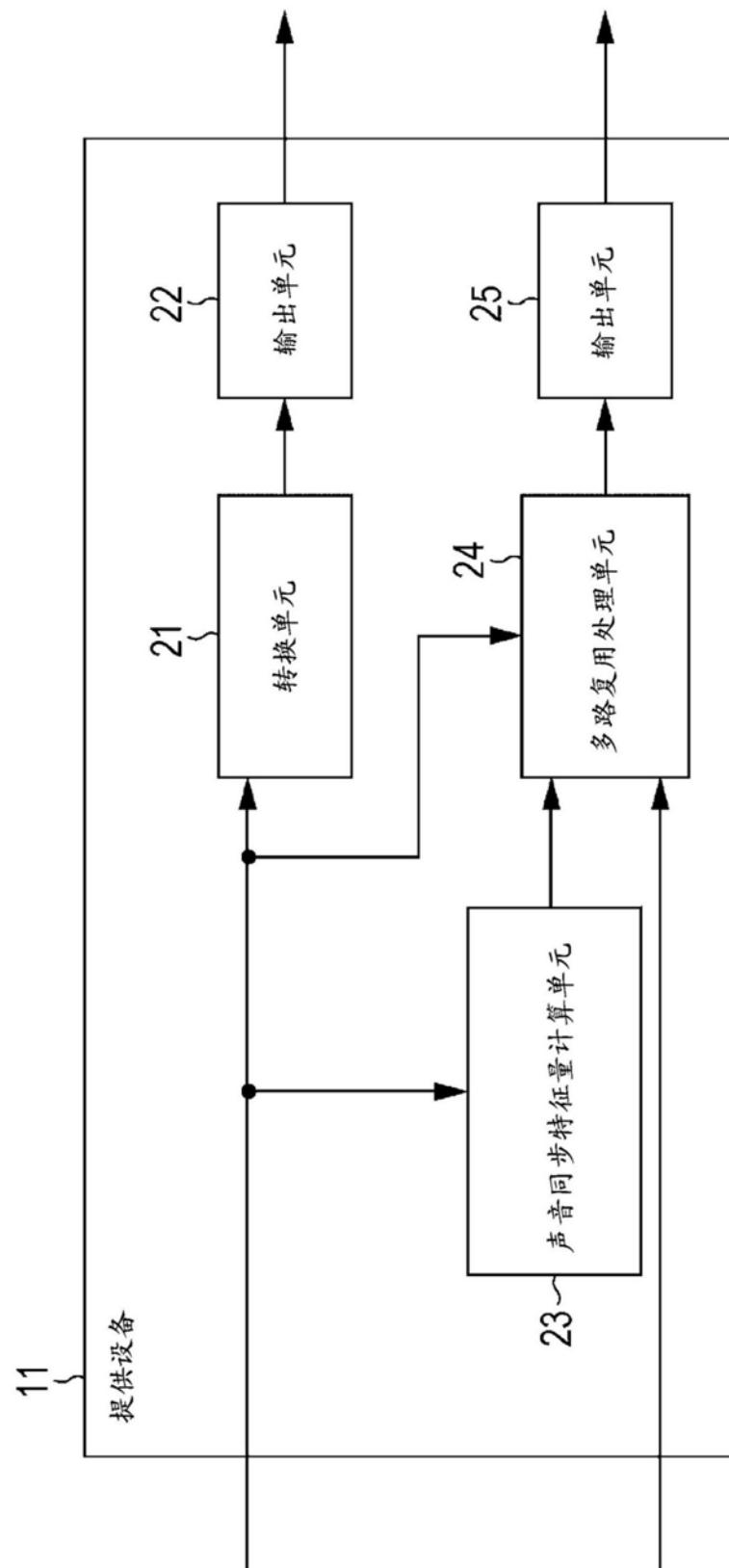


图1

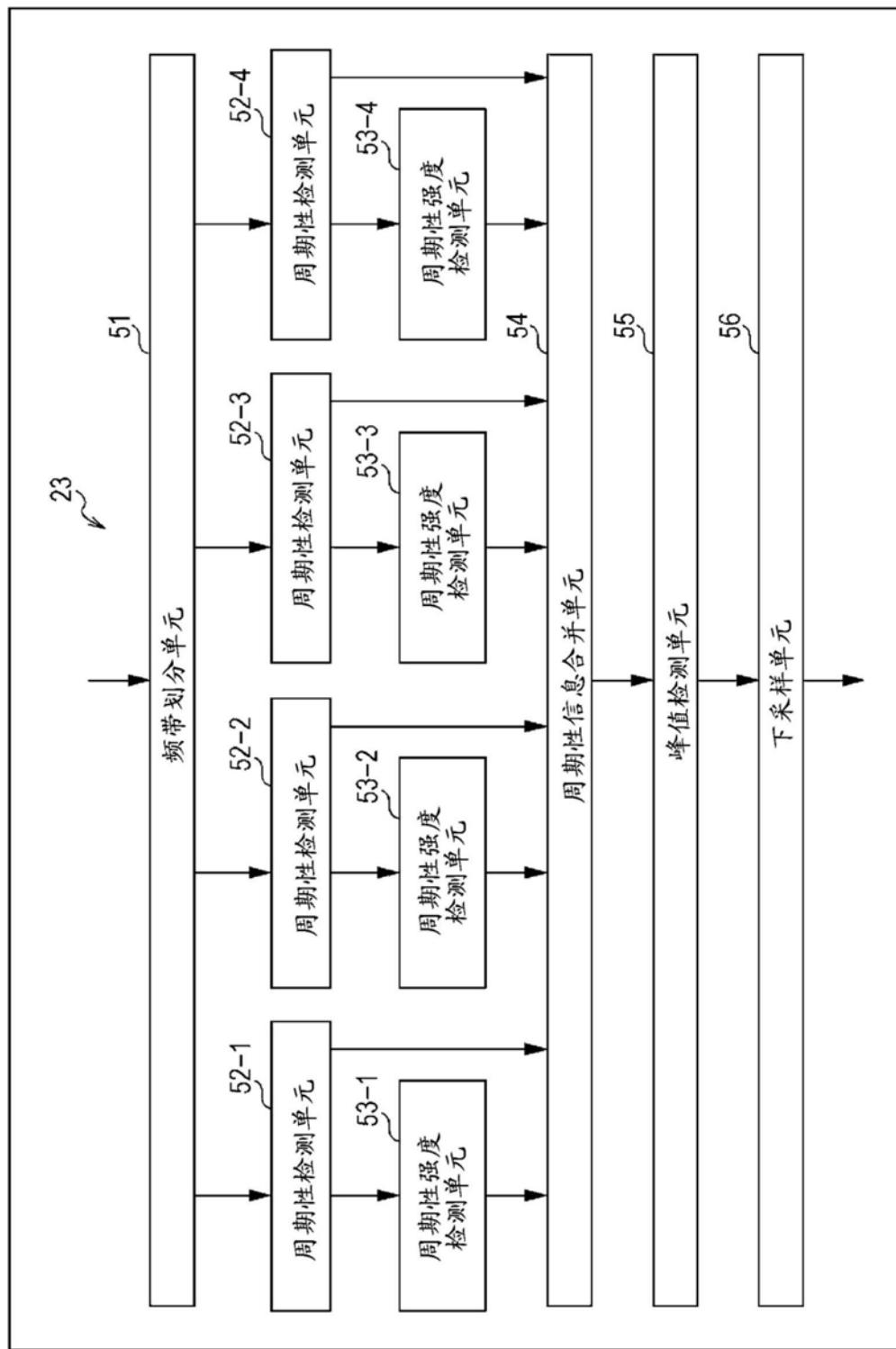


图2

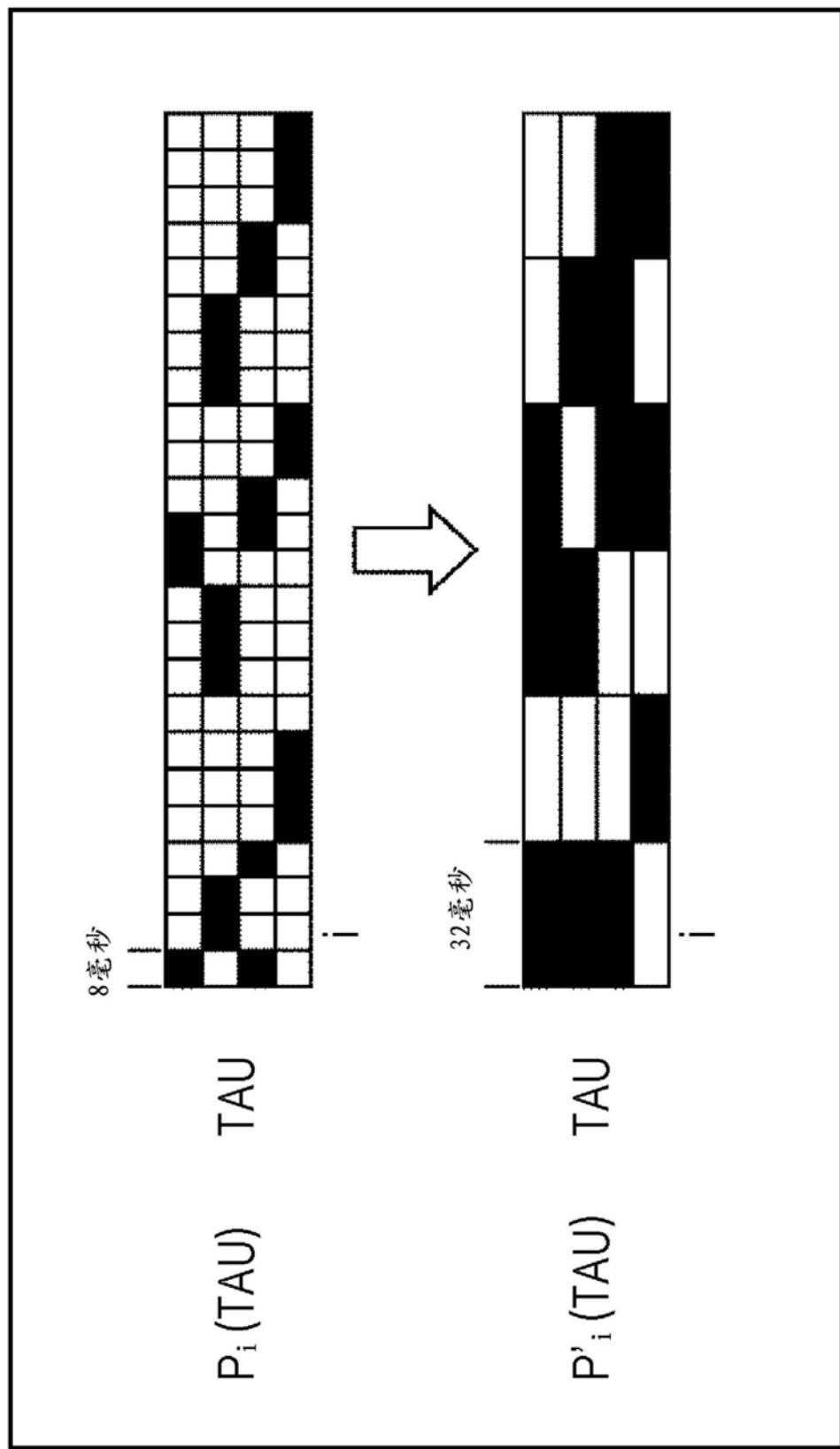


图3

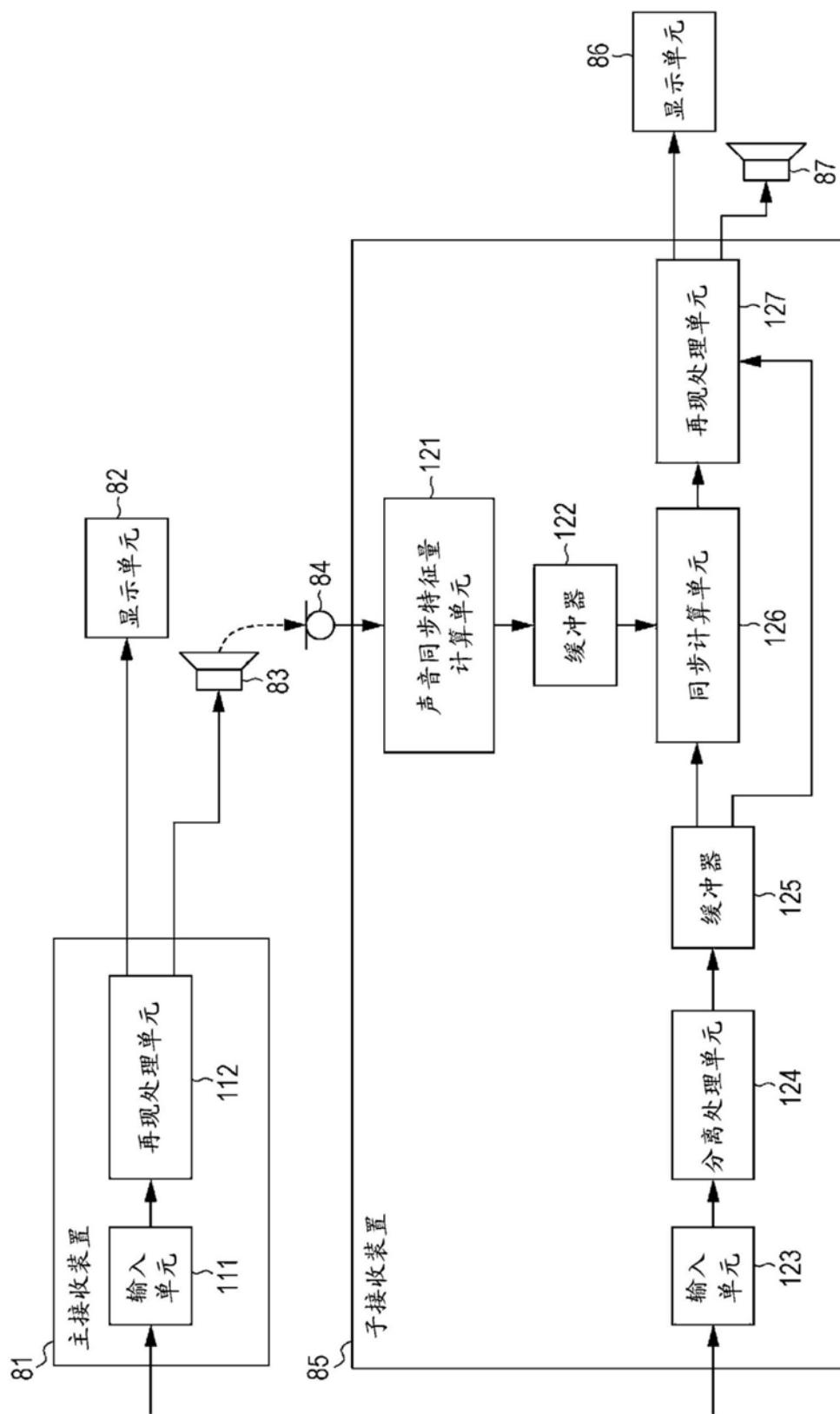


图4

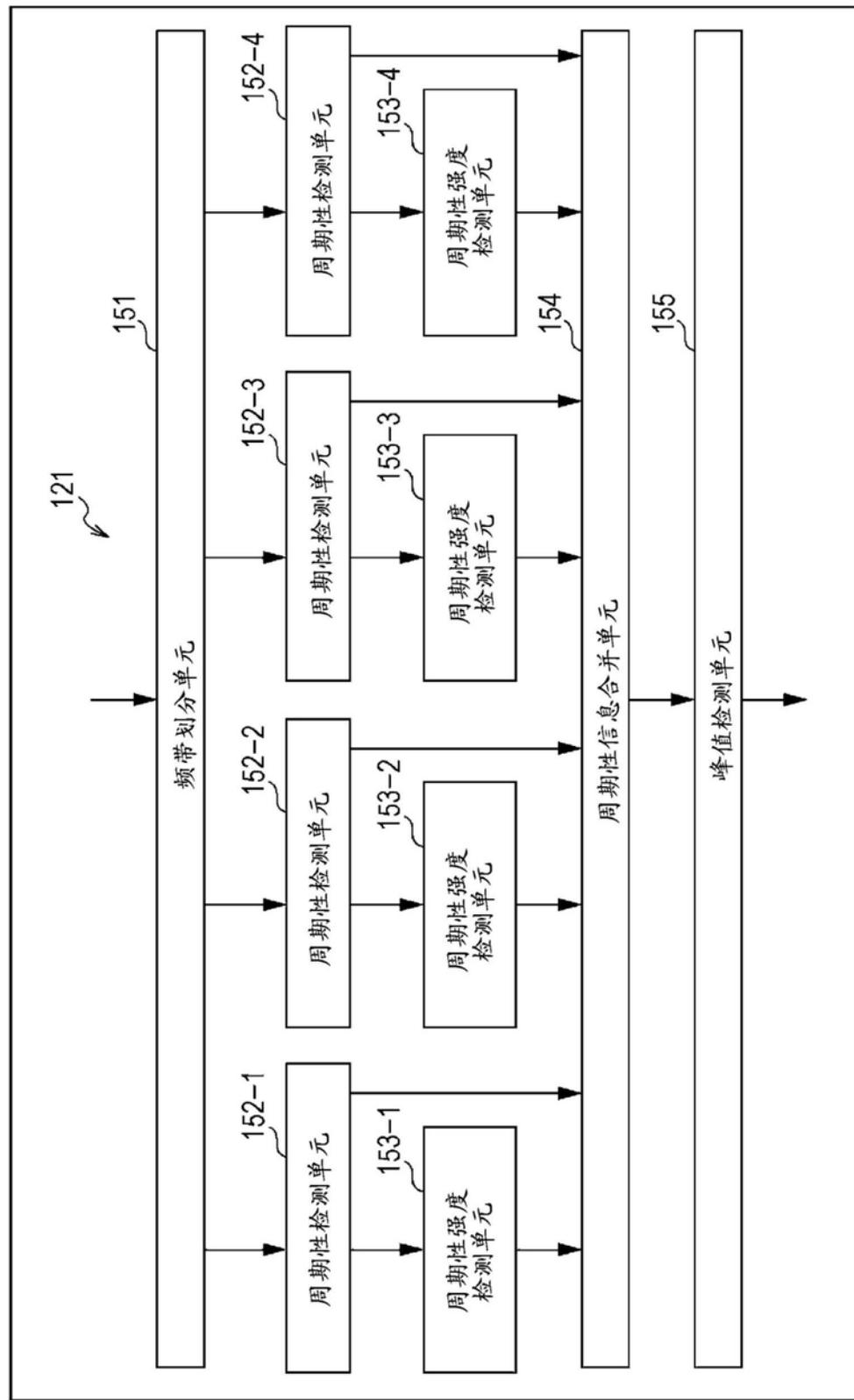


图5

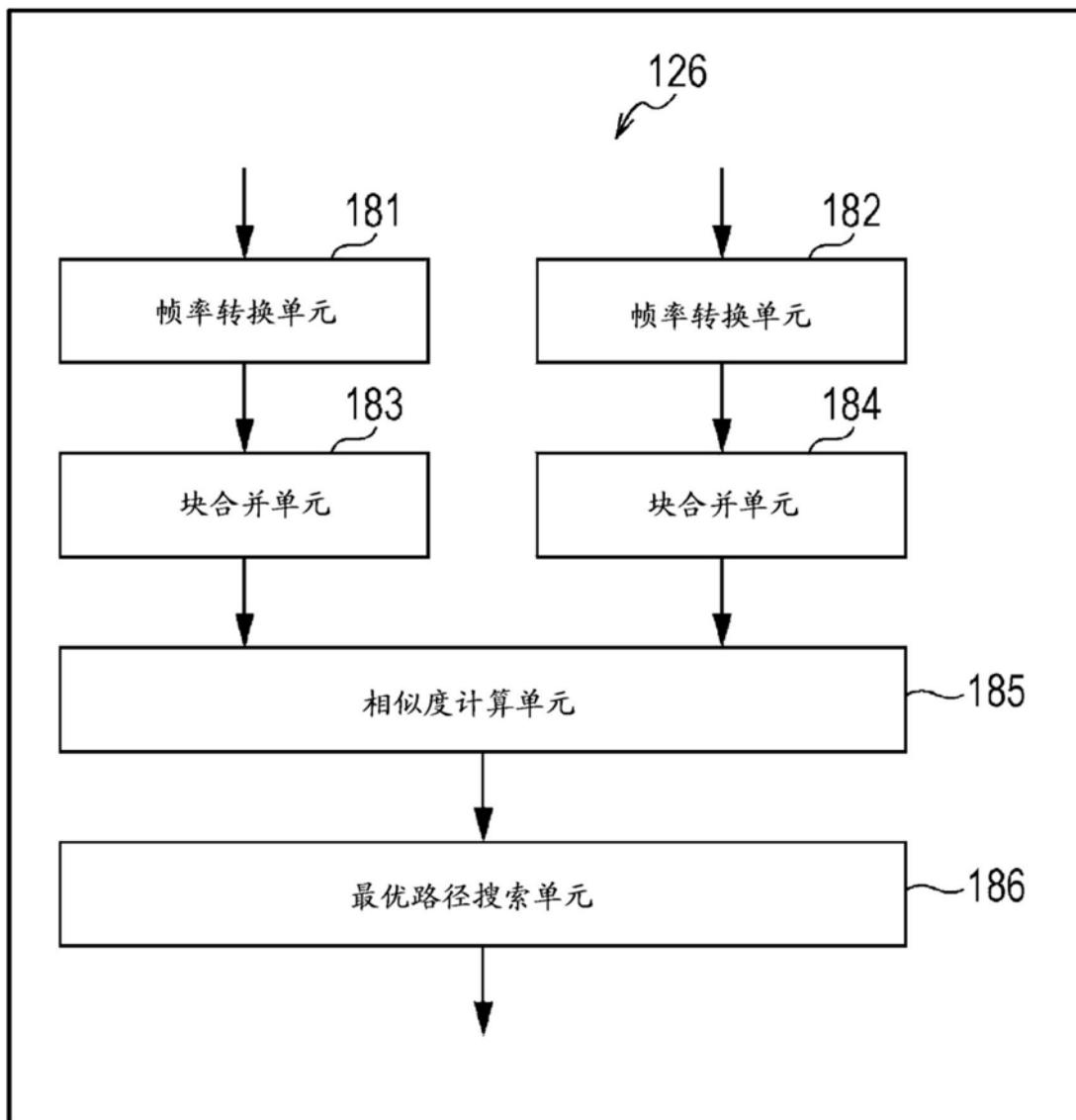


图6

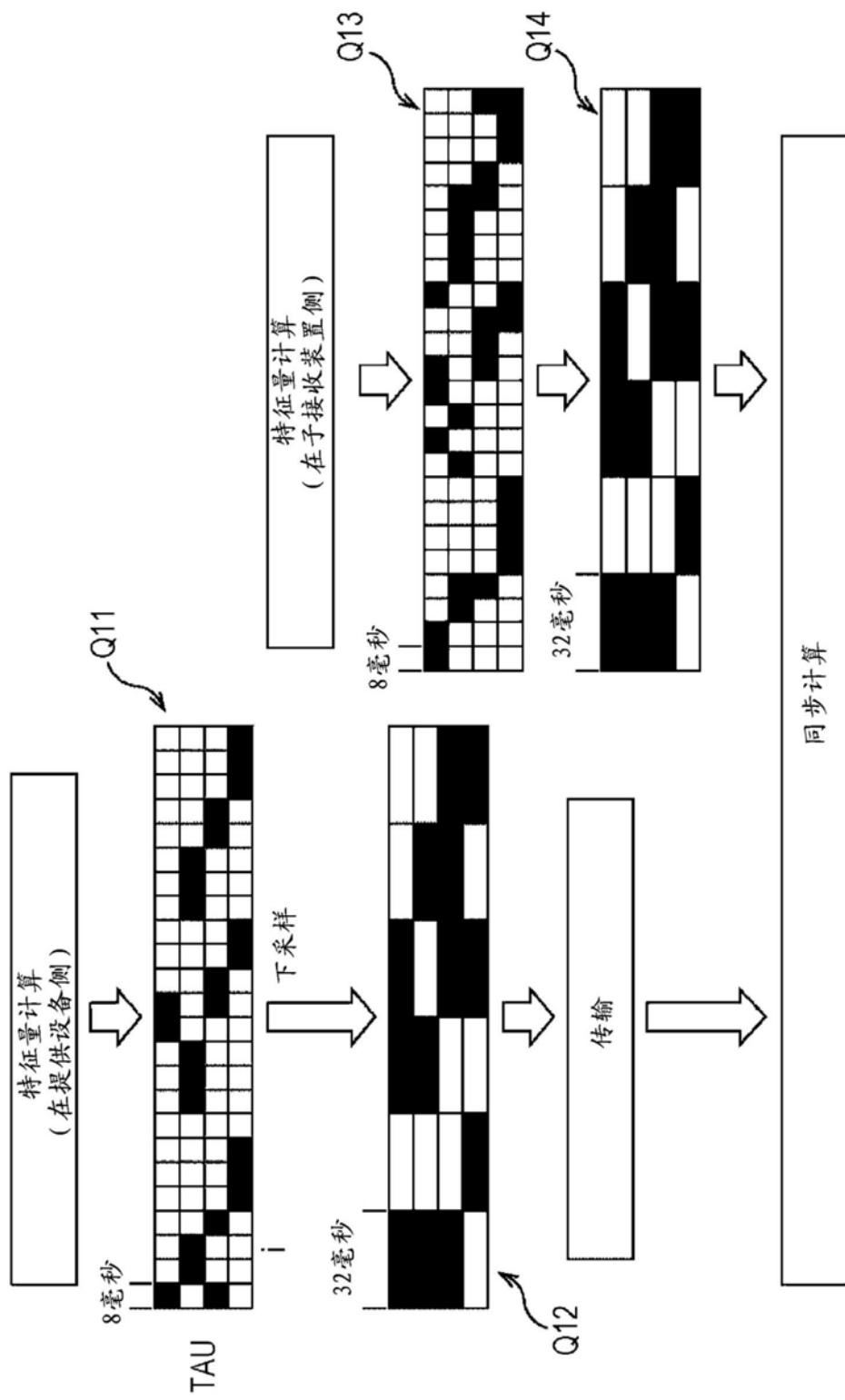


图7

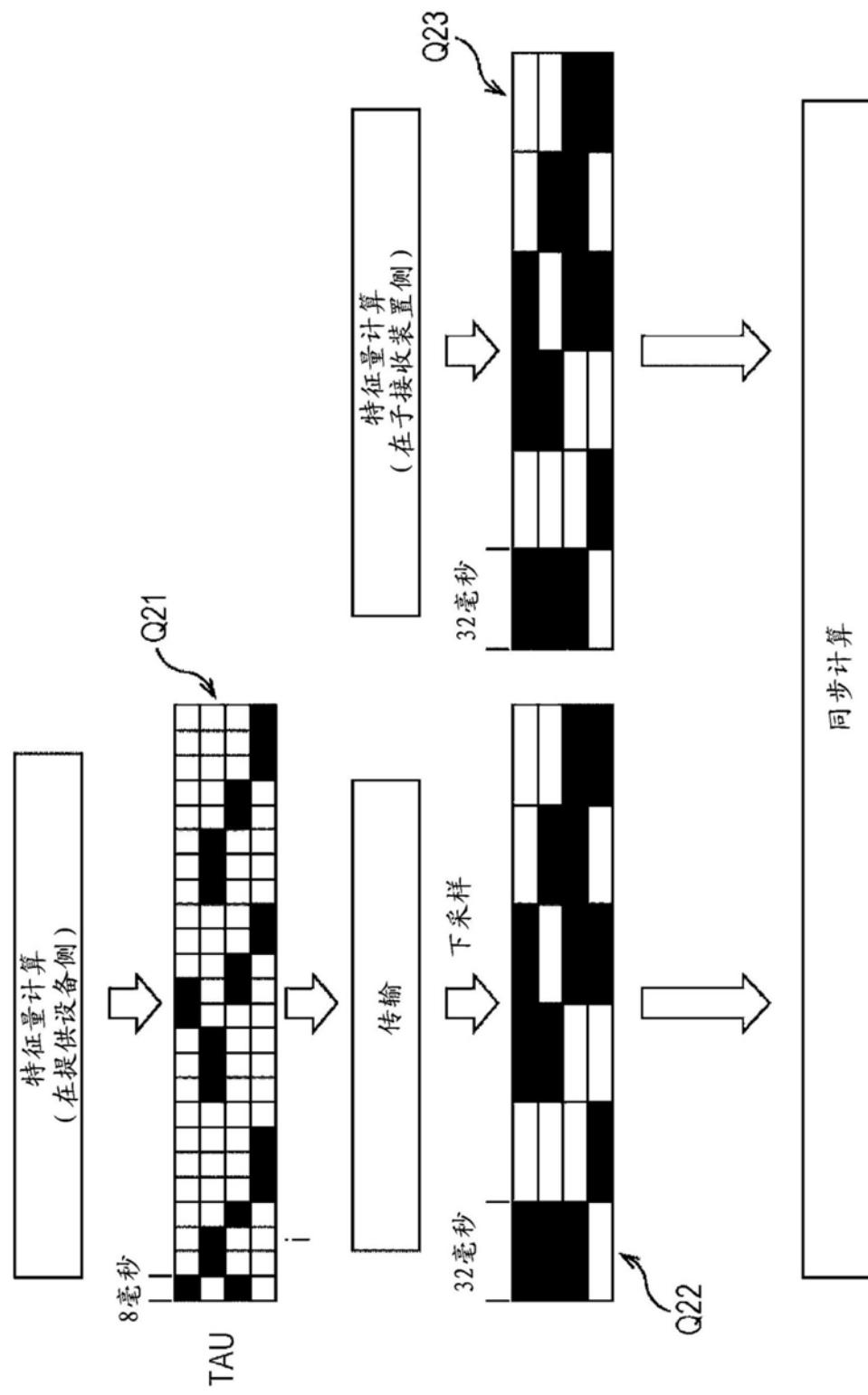


图8

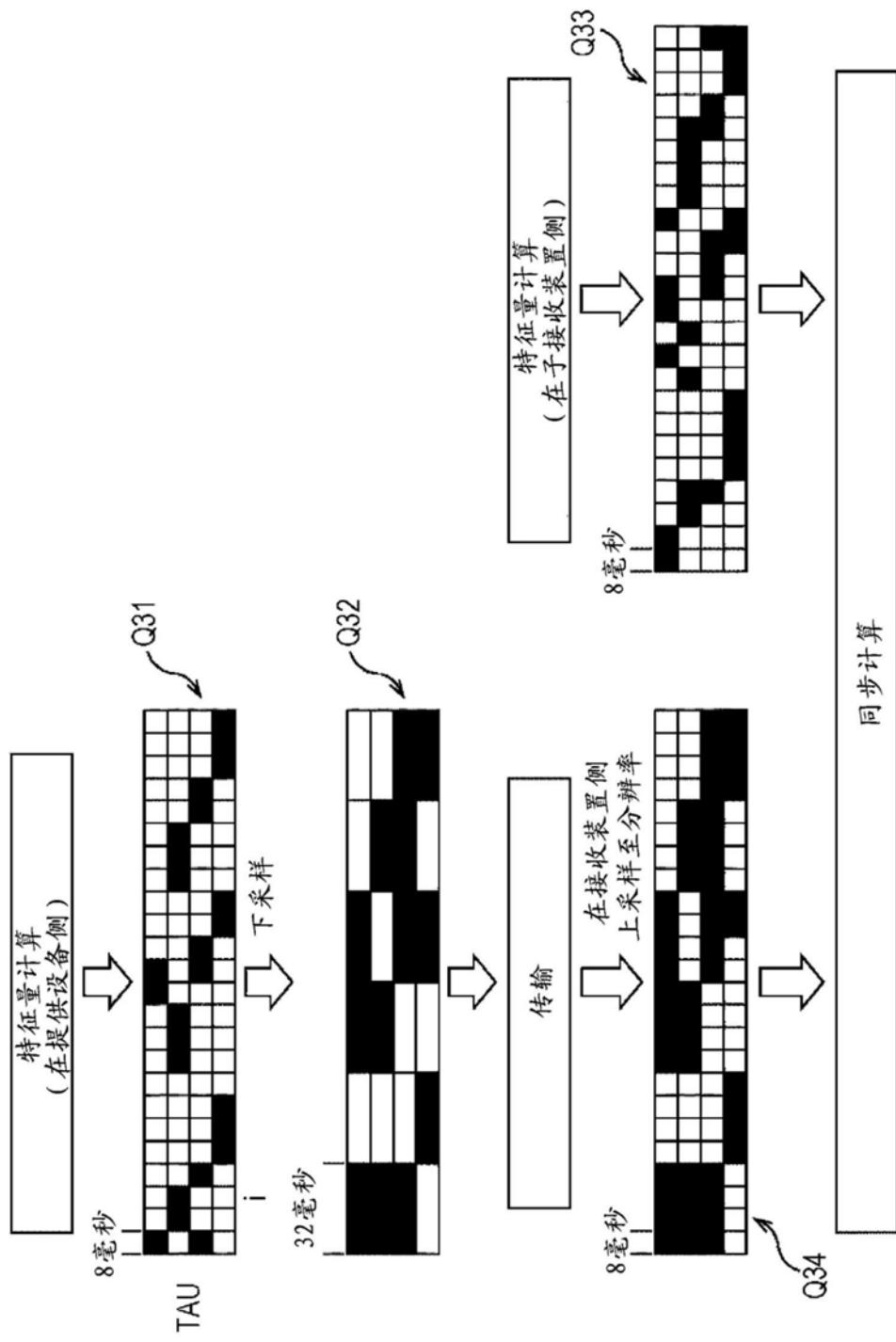


图9

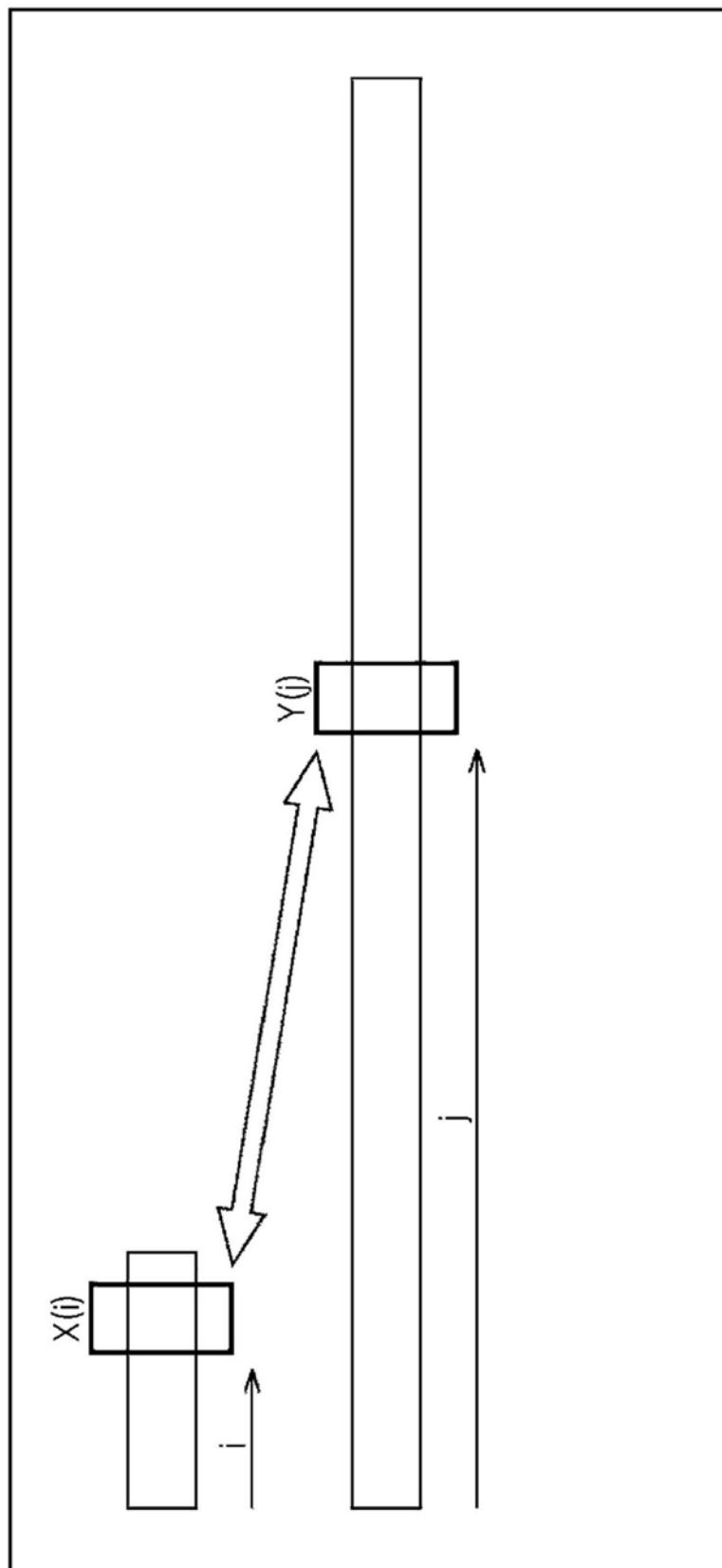


图10

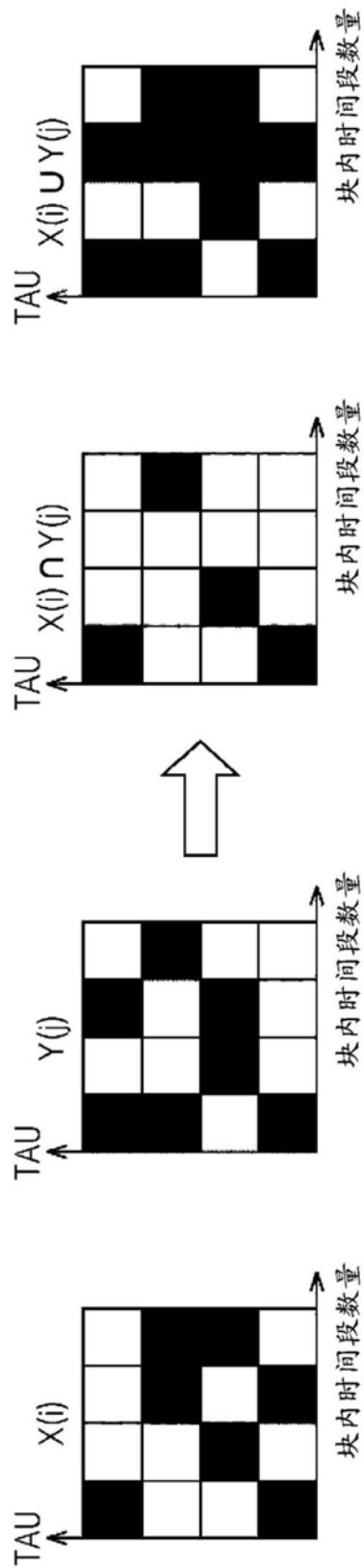


图11

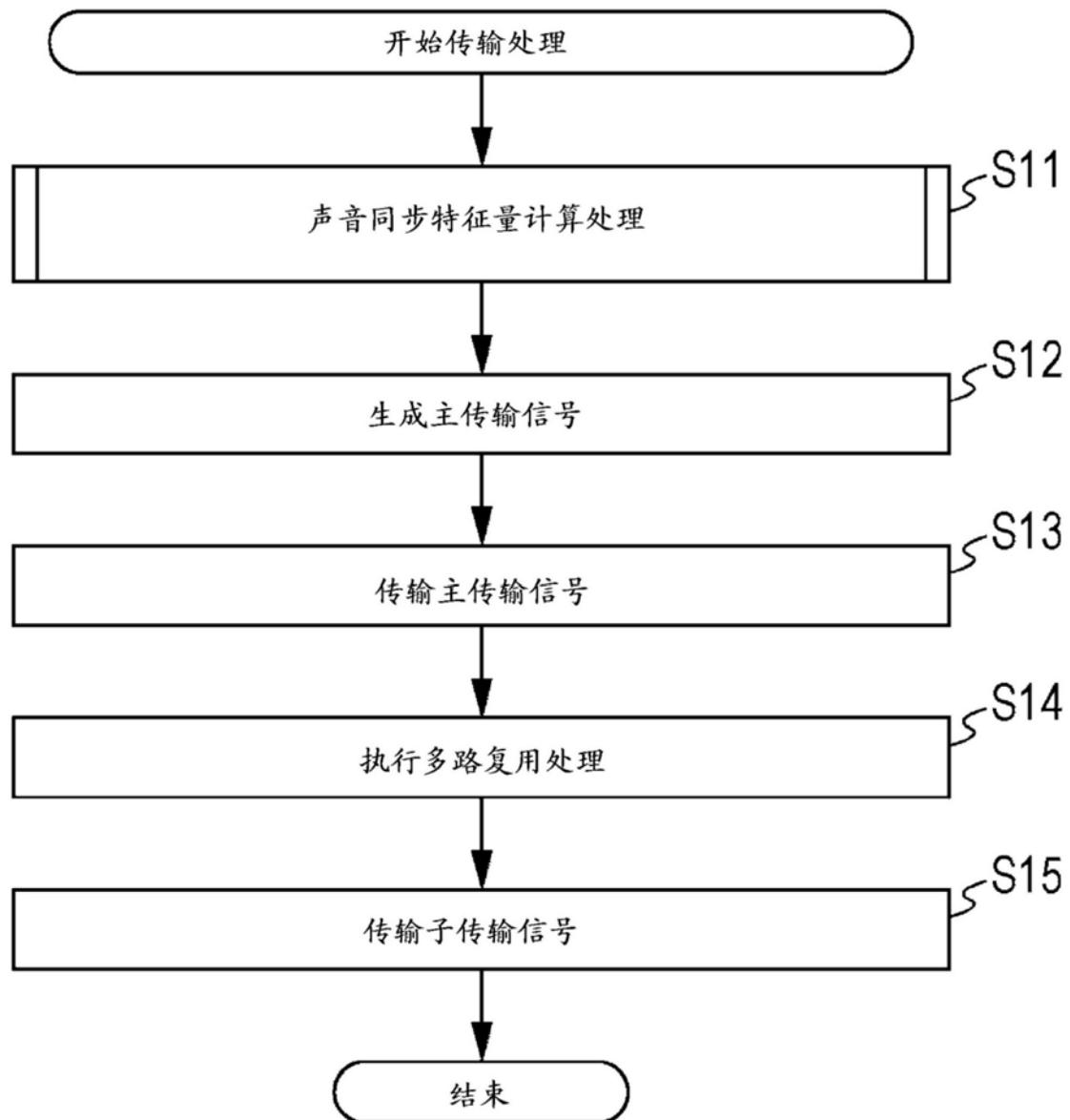


图12

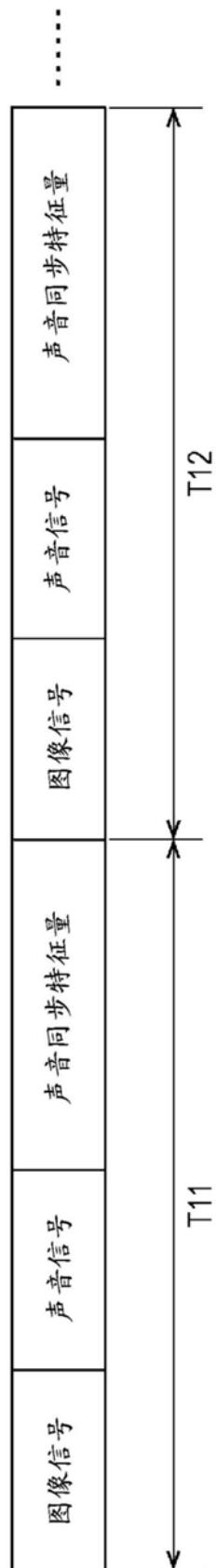


图13

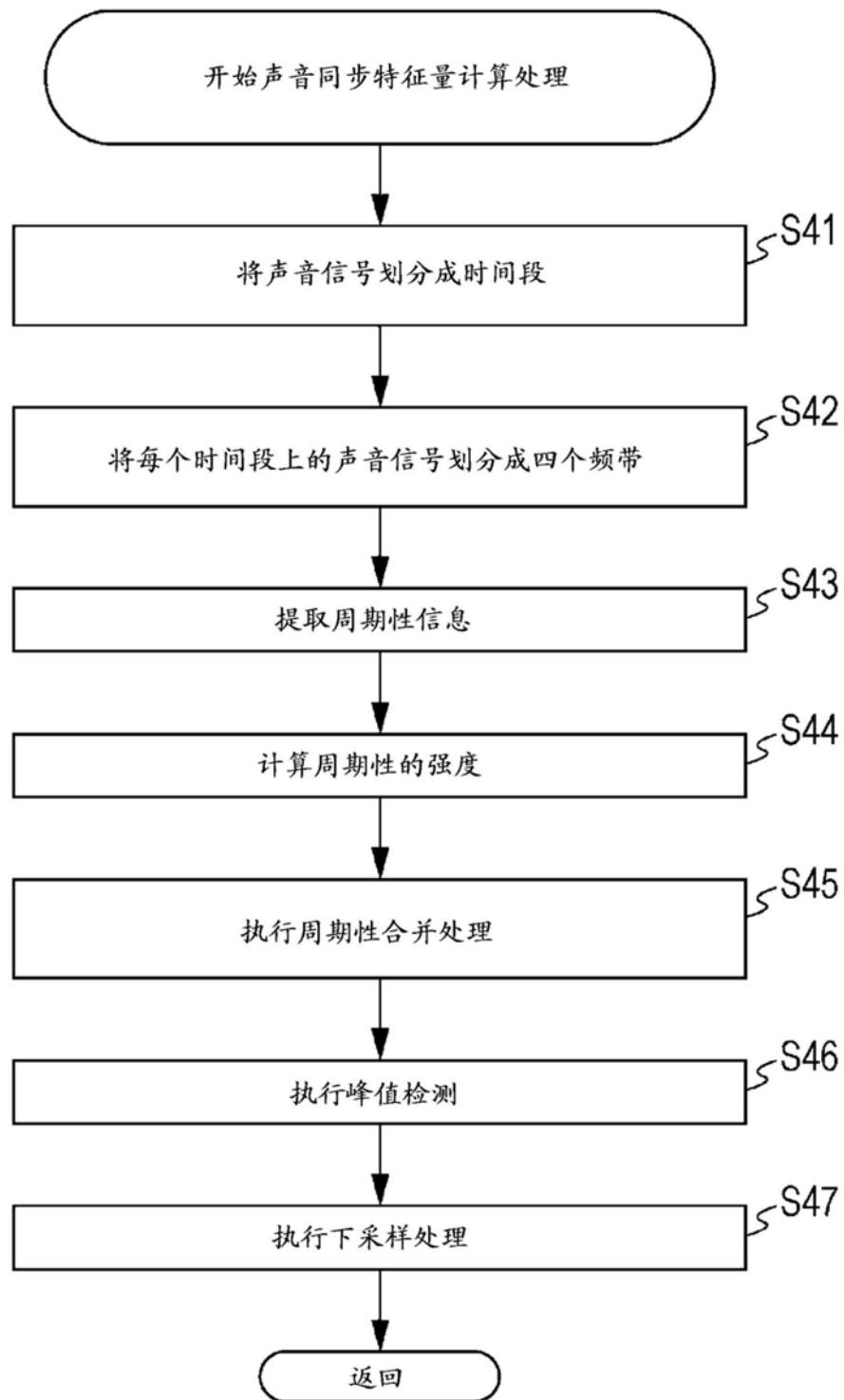


图14

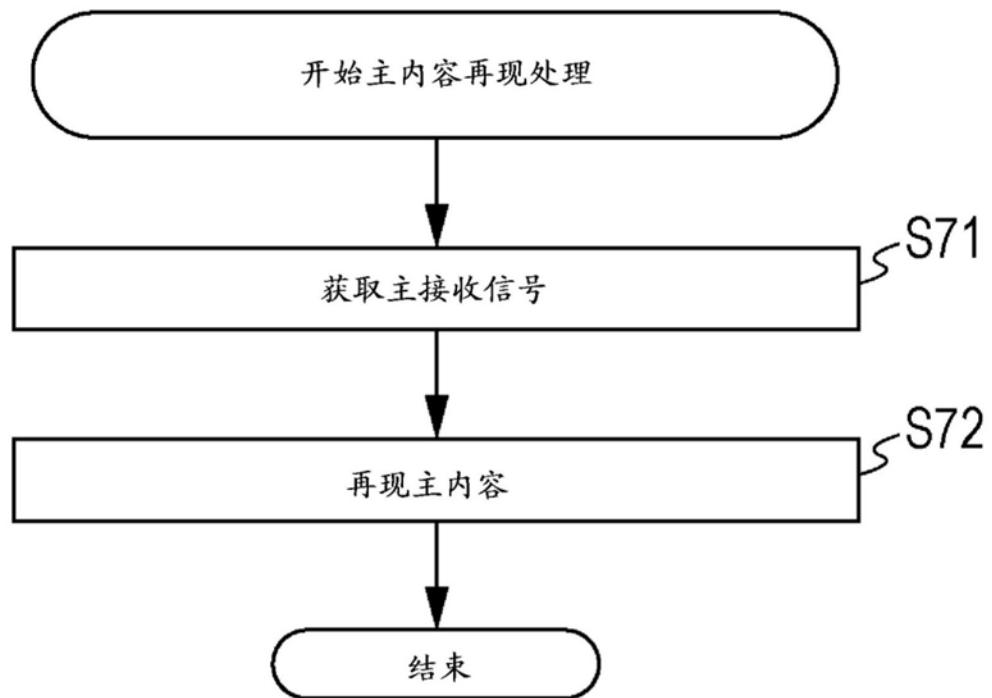


图15

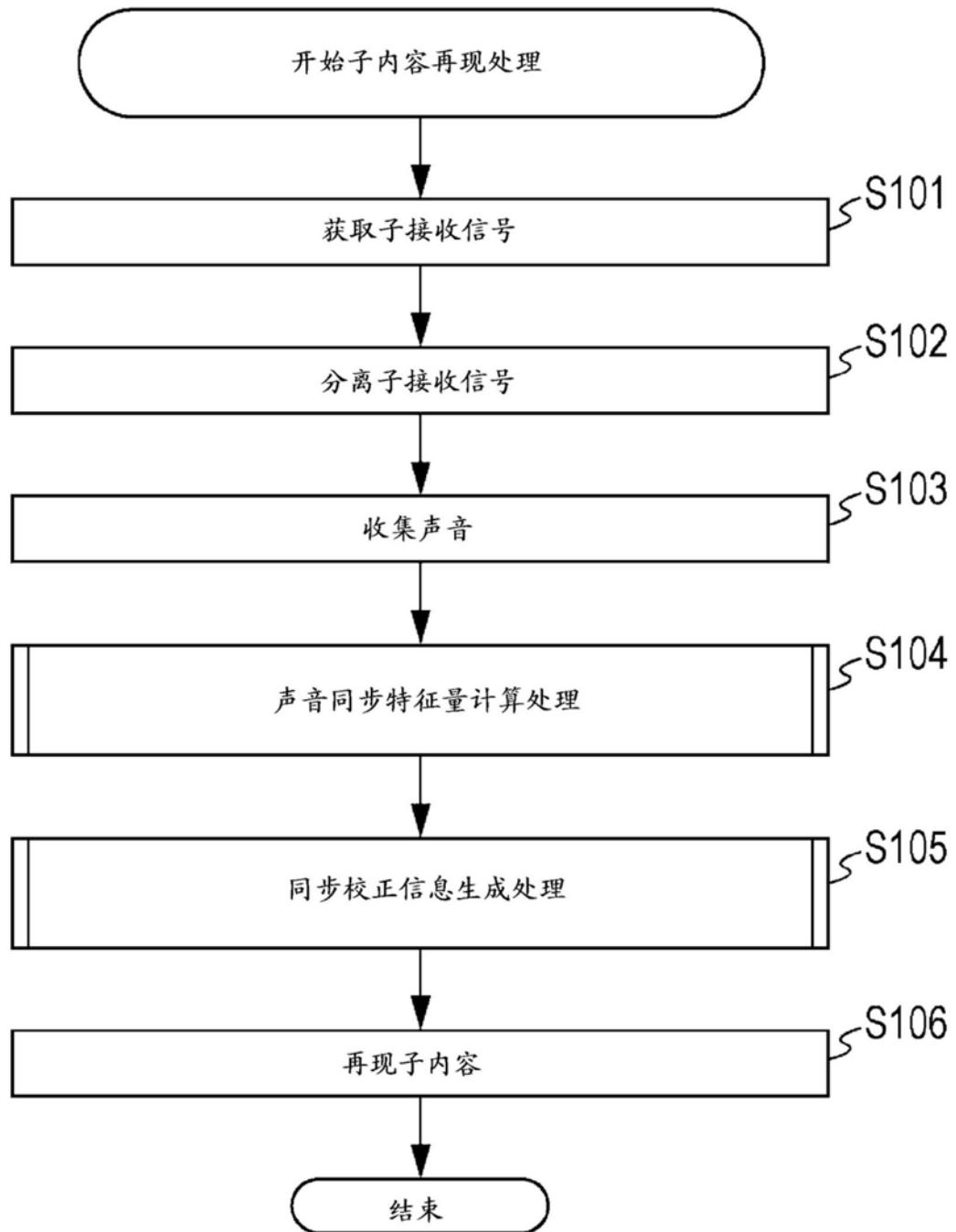


图16

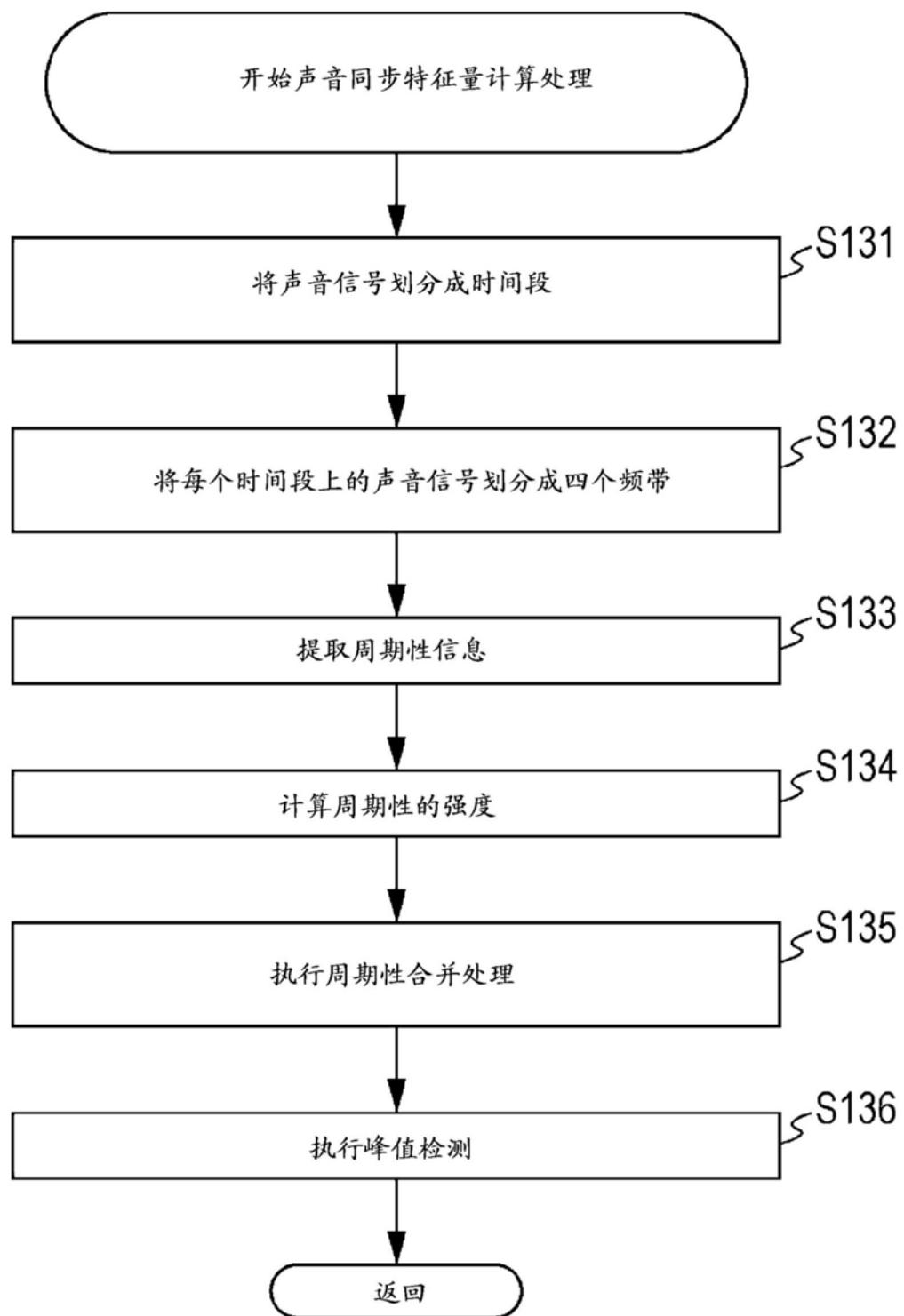


图17

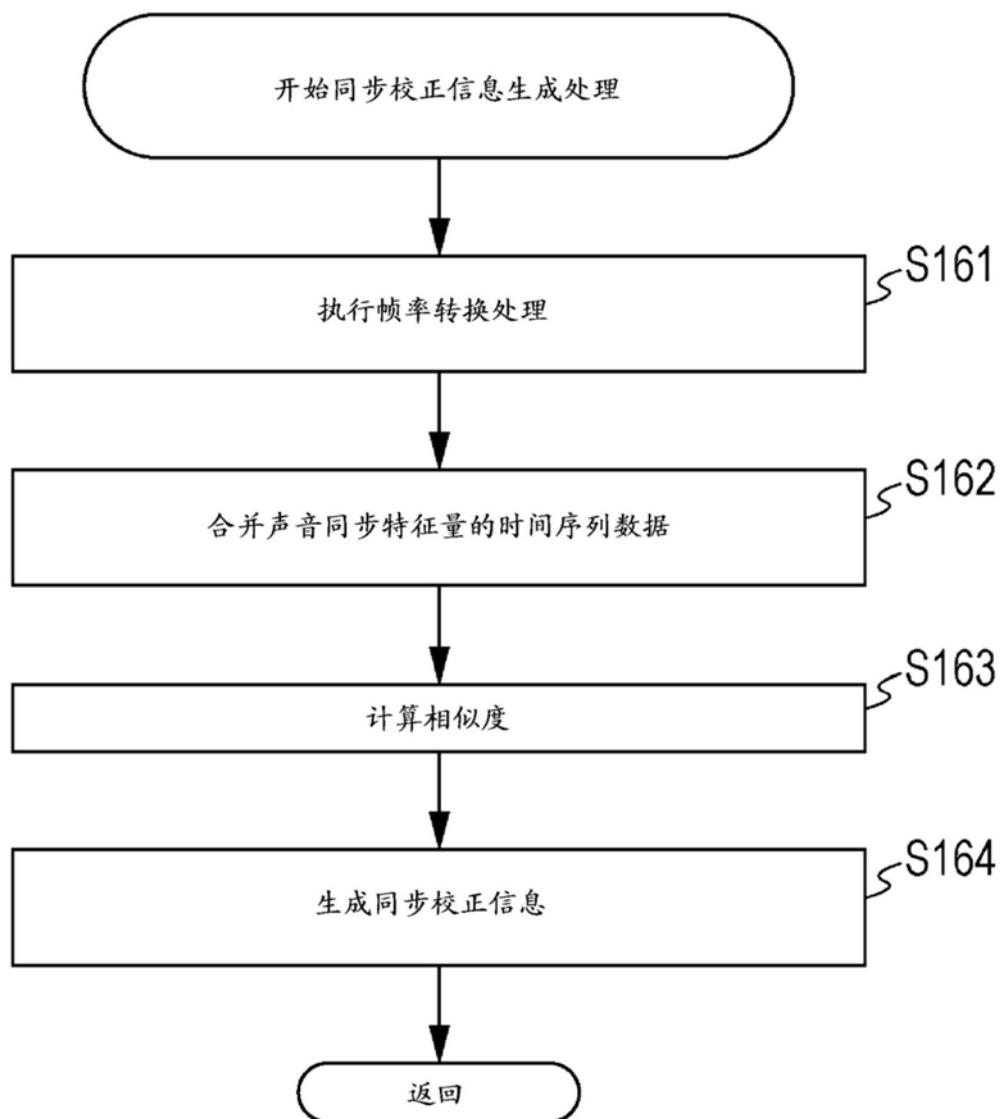


图18

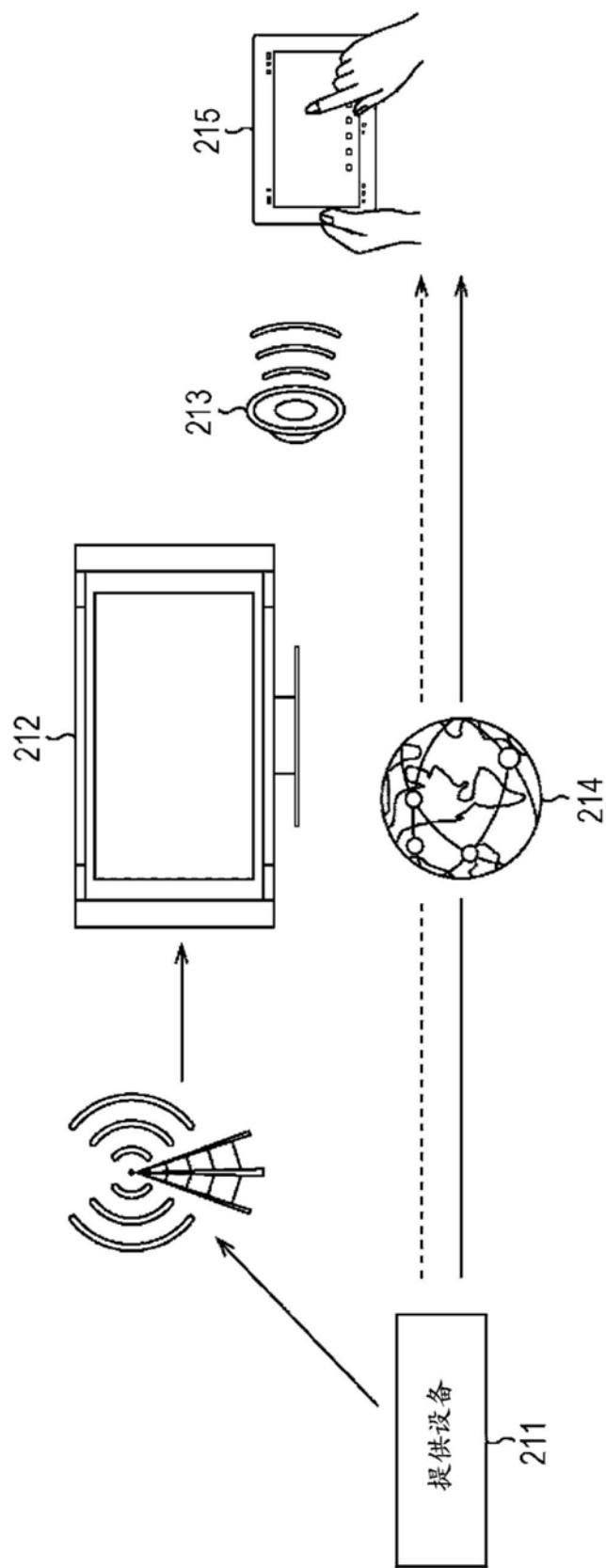


图19

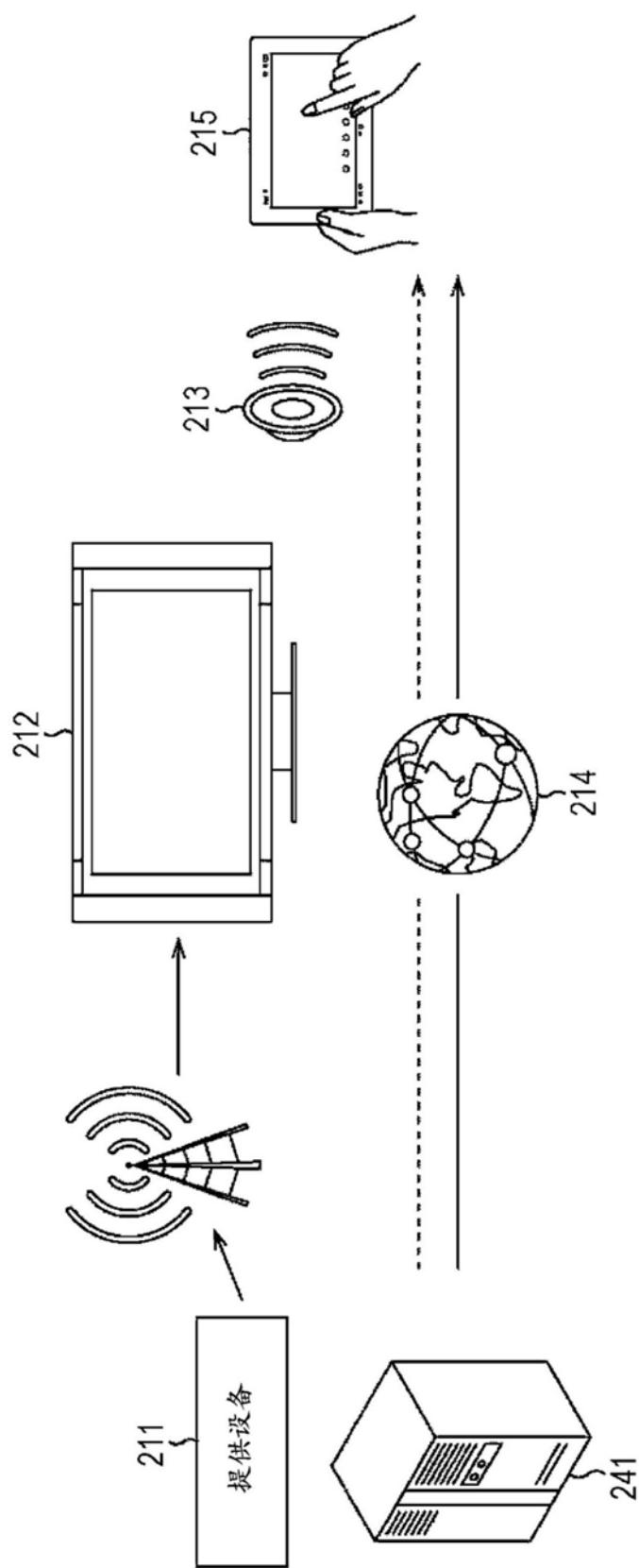


图20

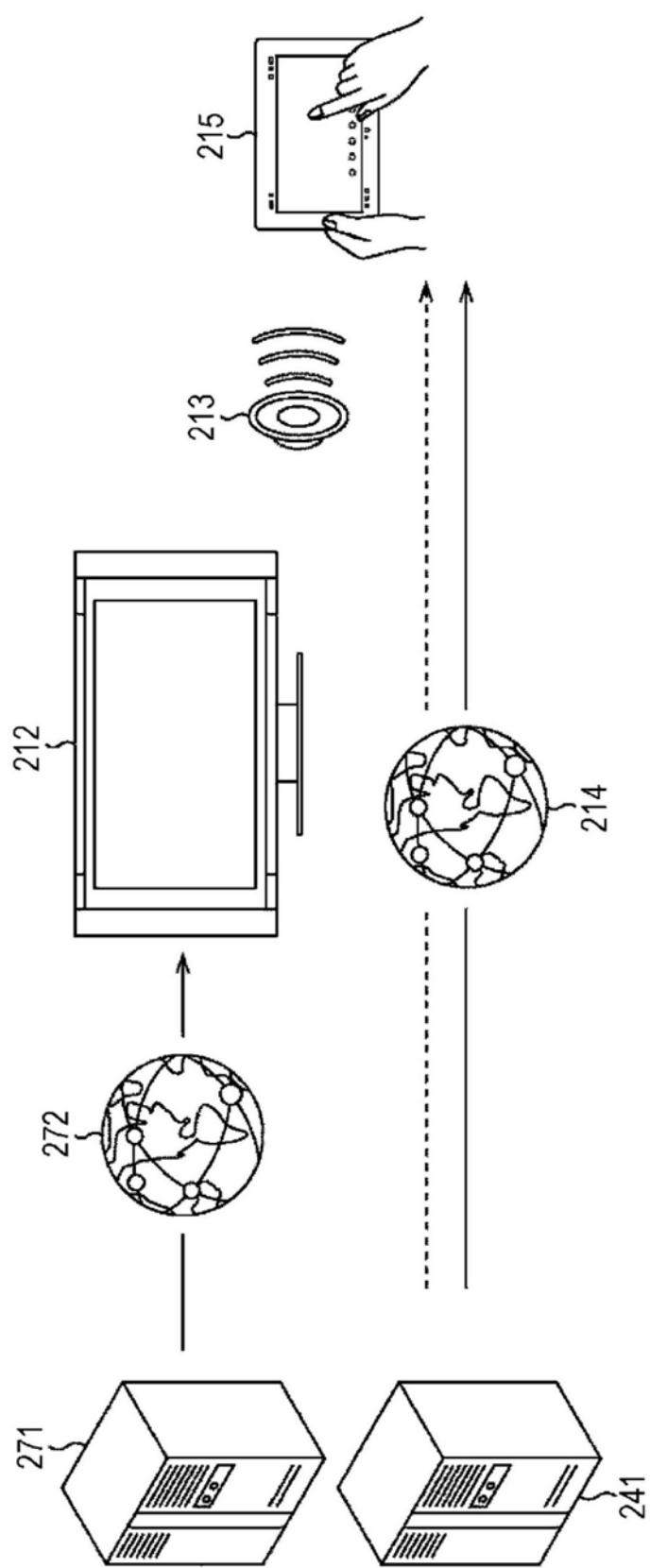


图21

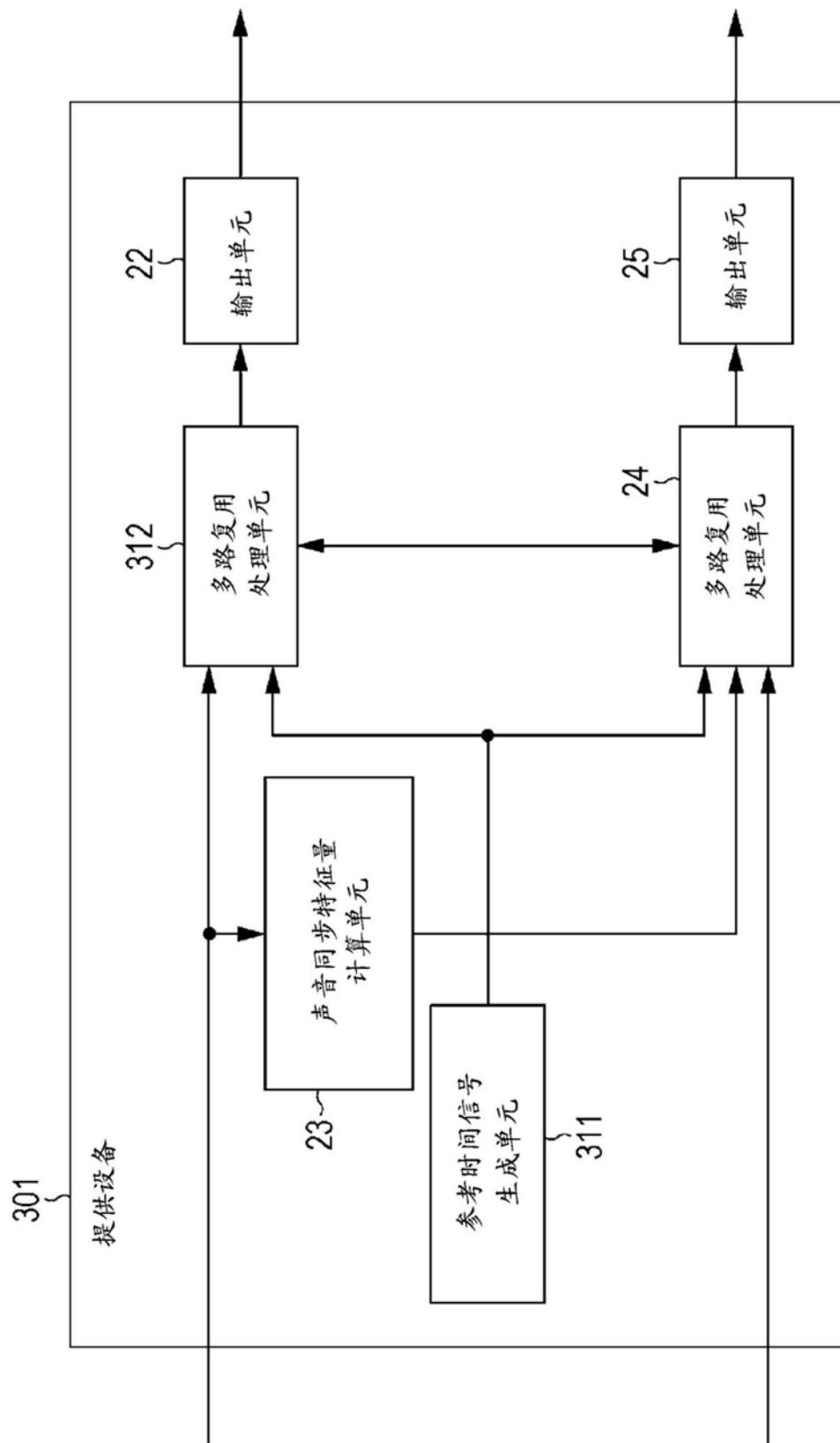


图22

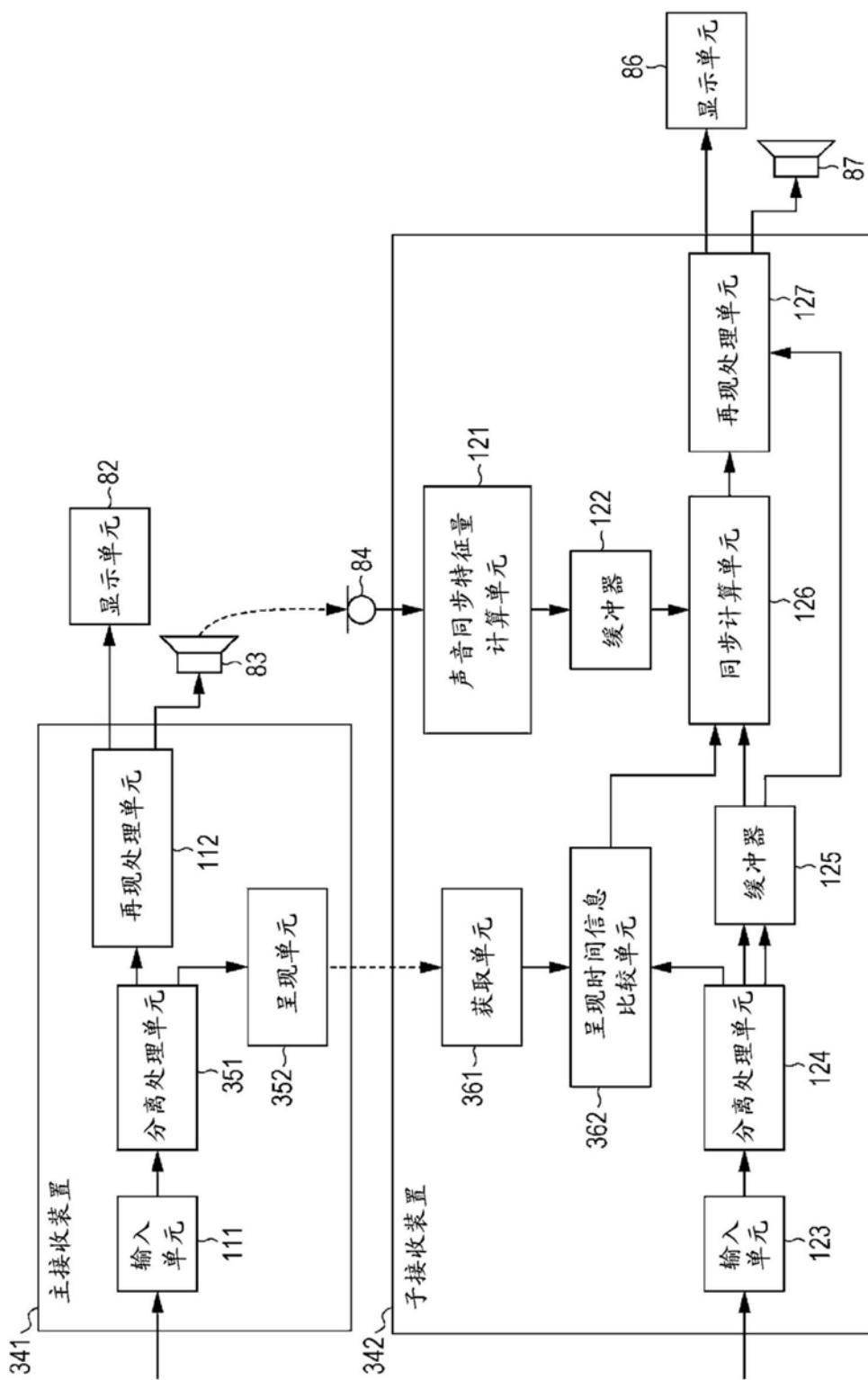


图23

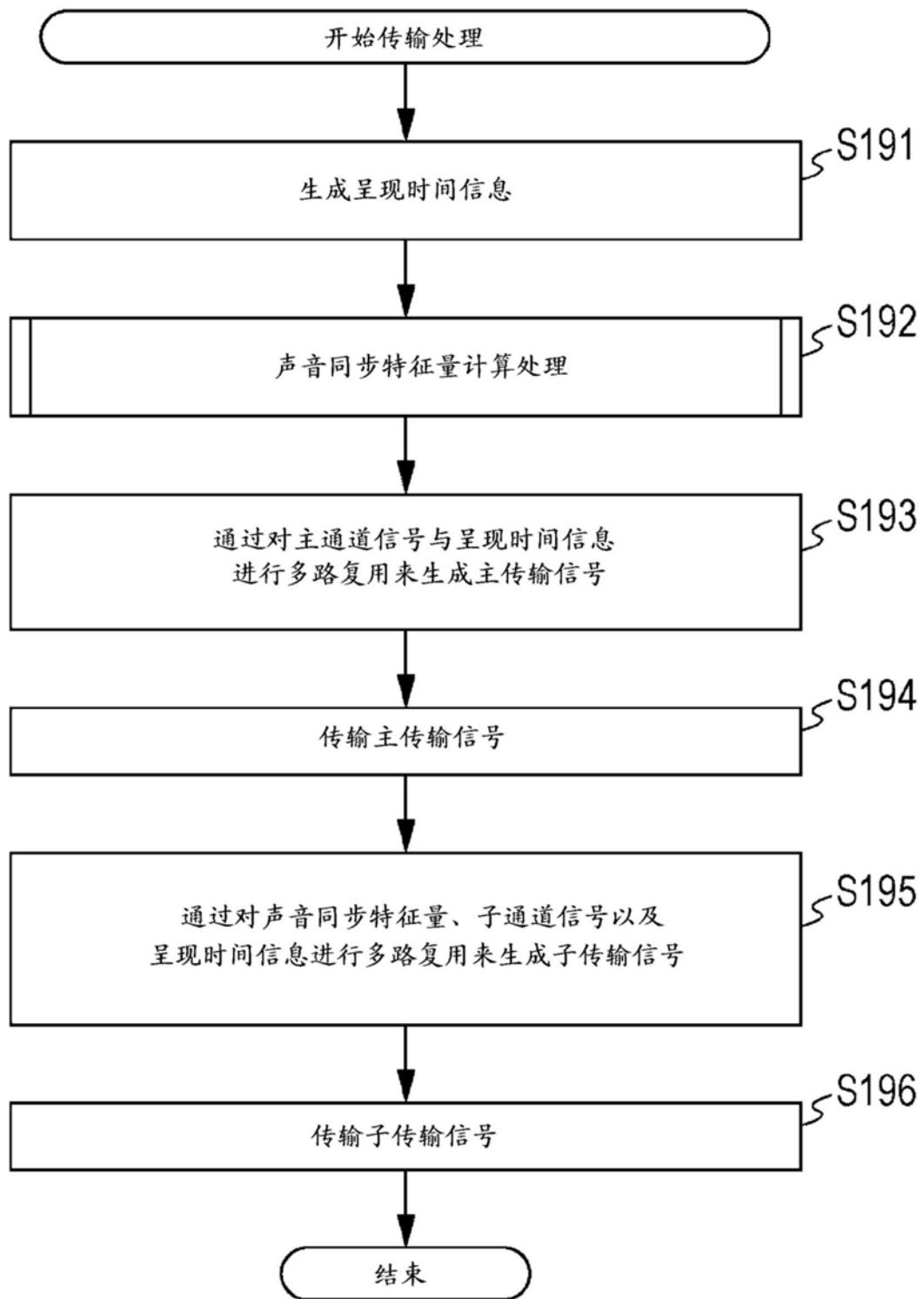


图24

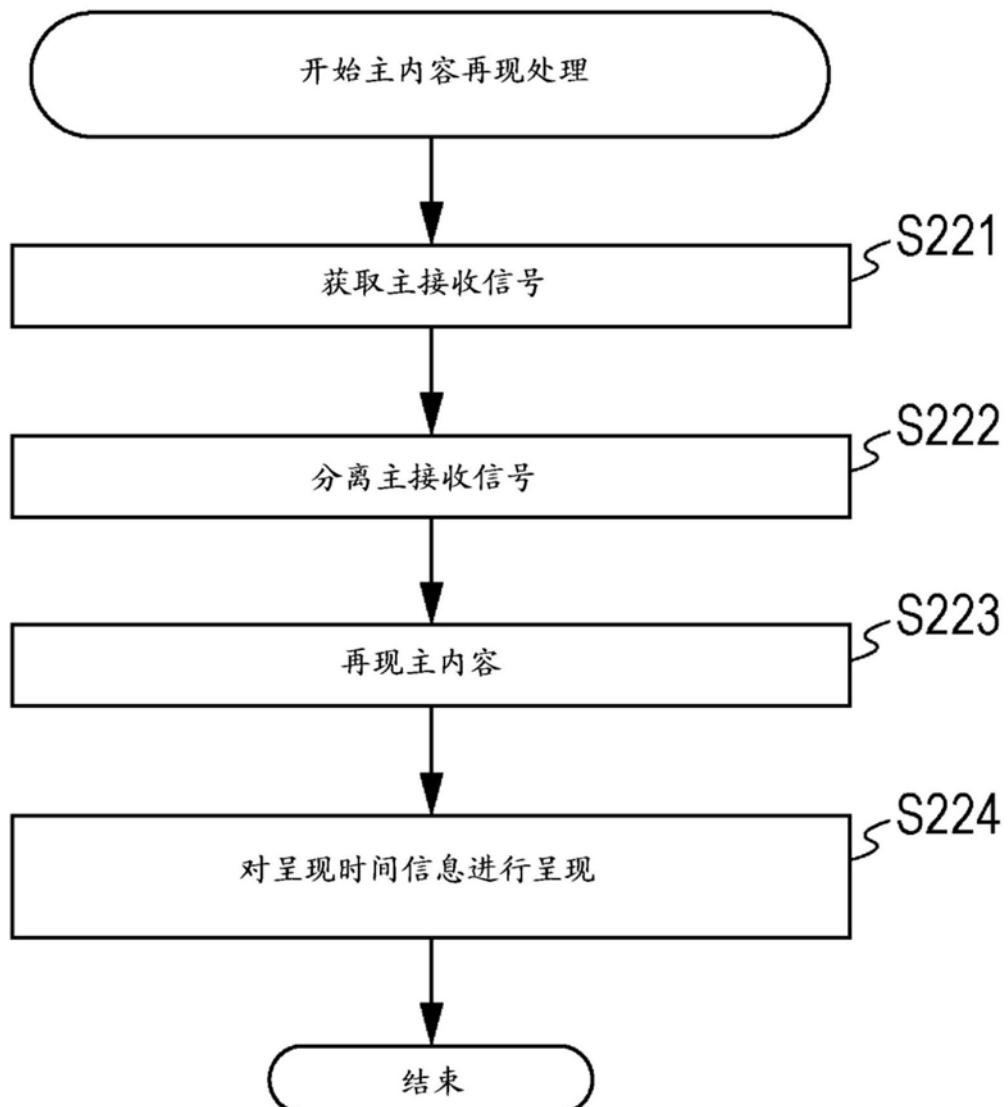


图25

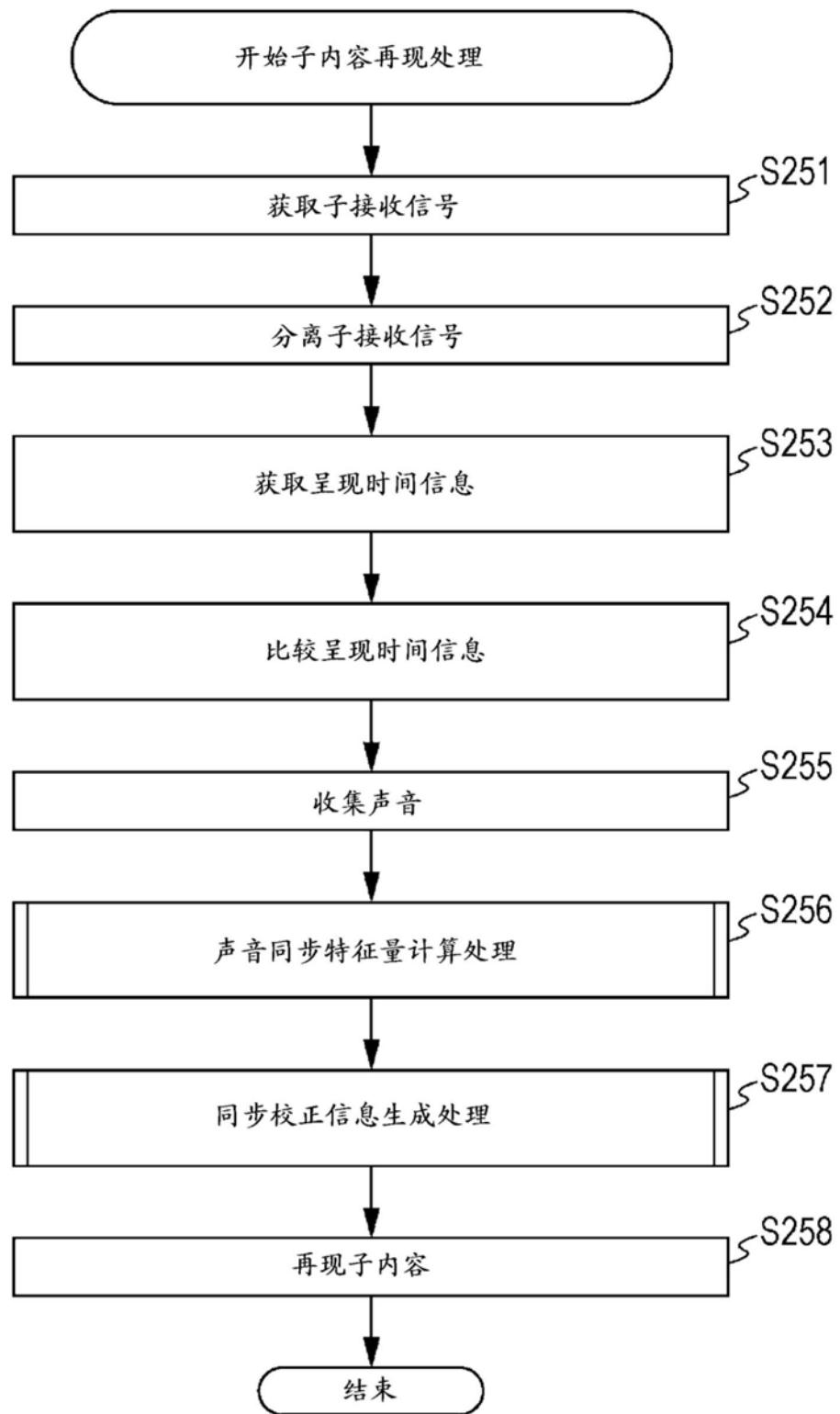


图26

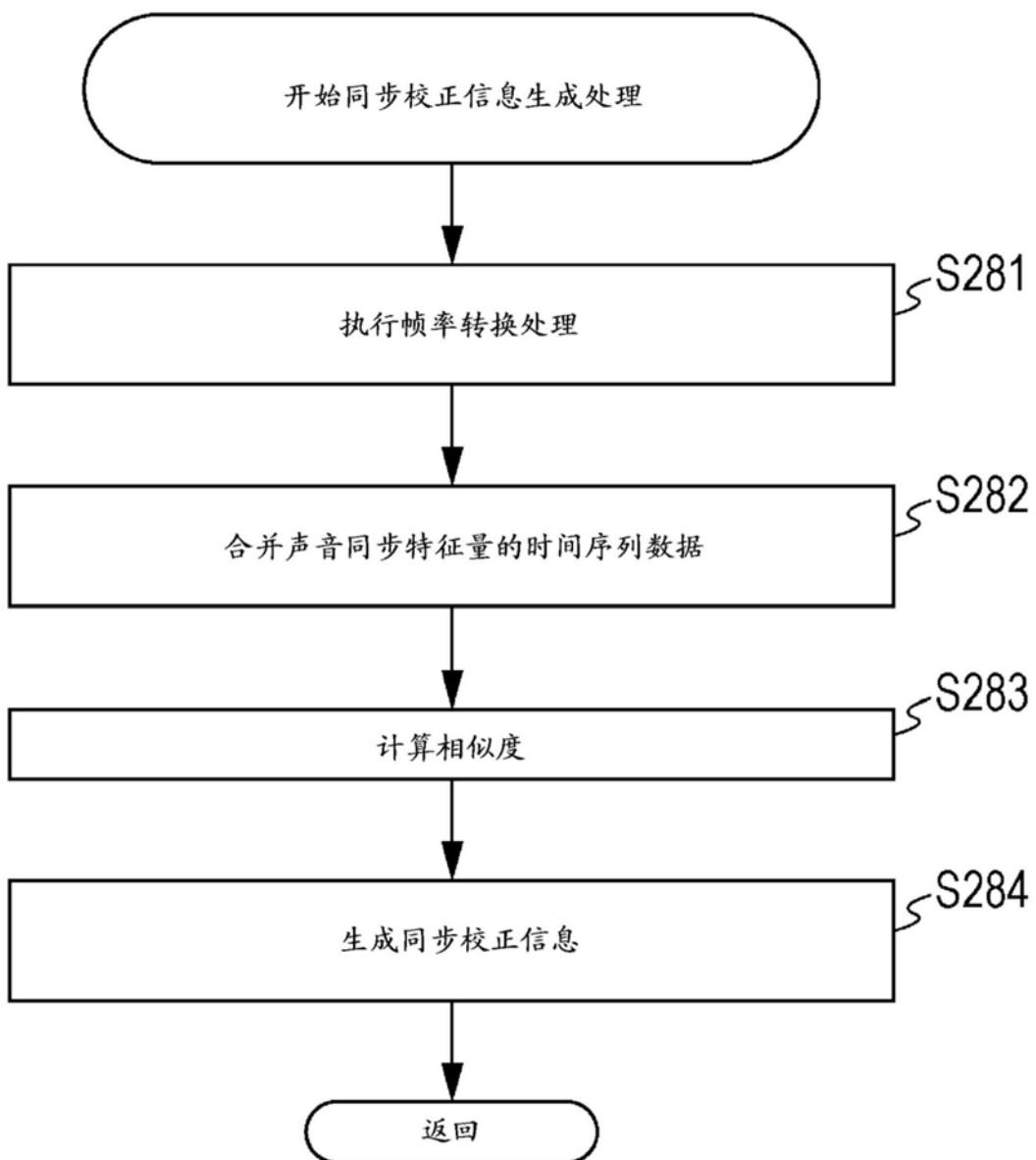


图27

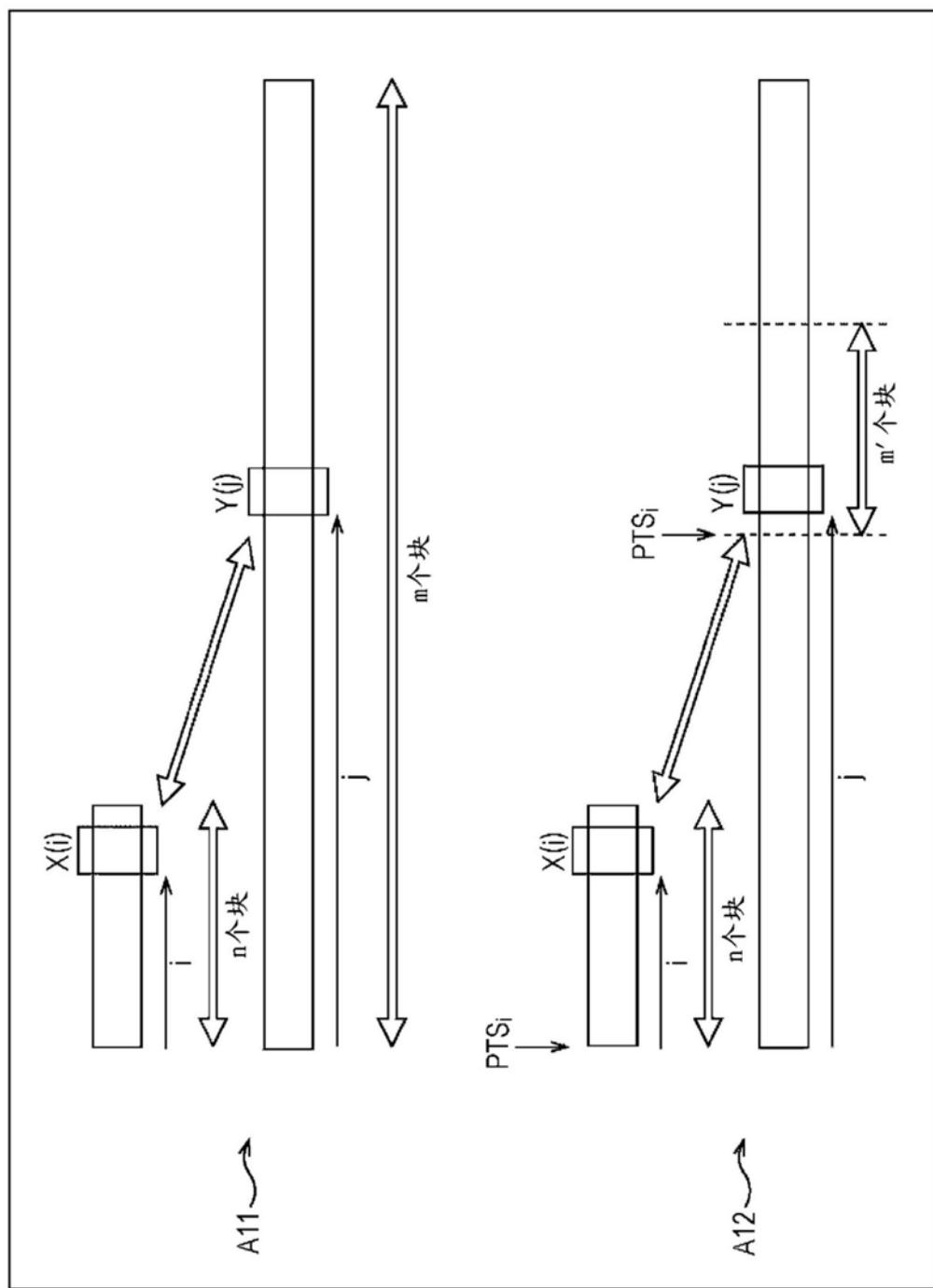


图28

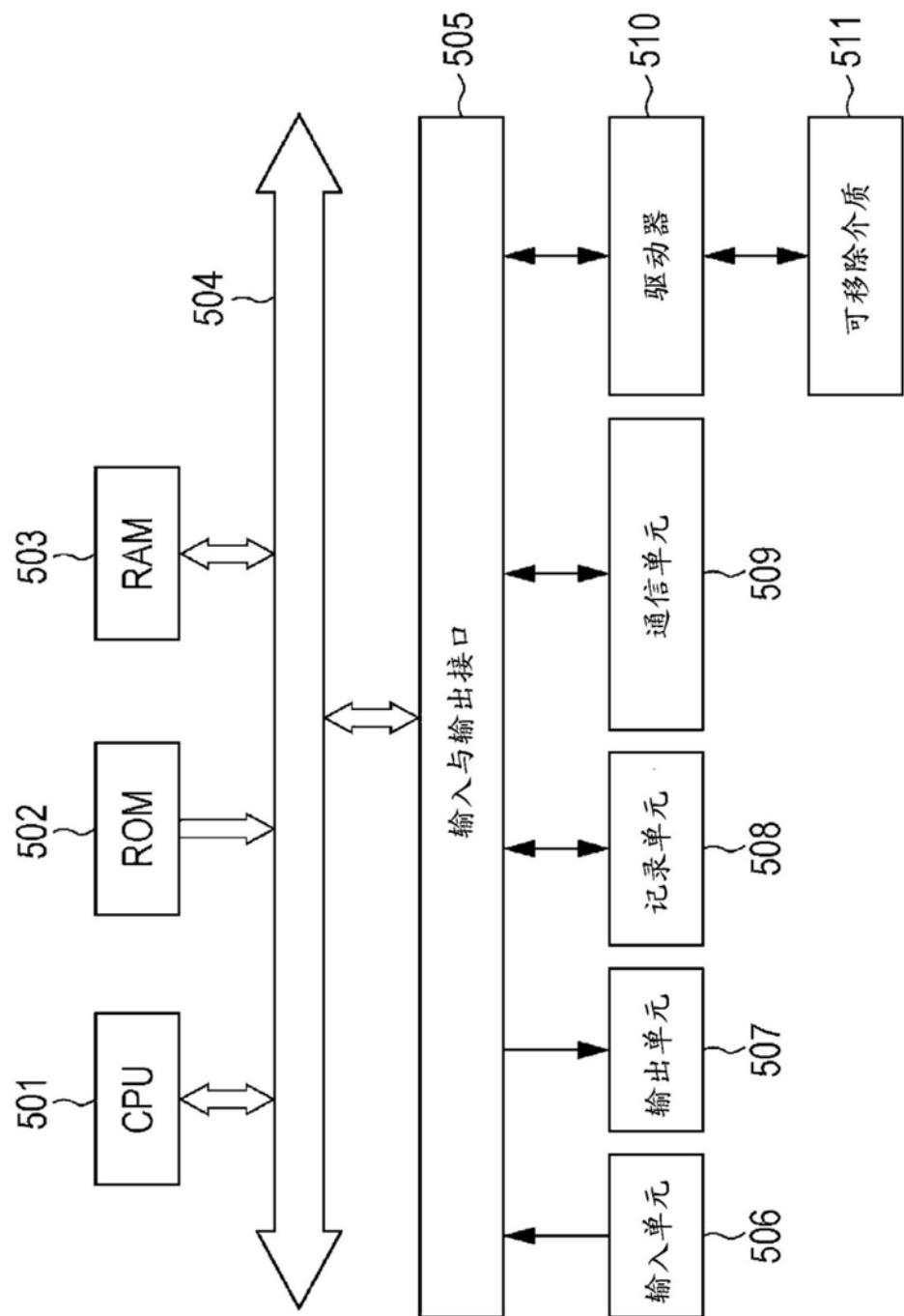


图29