



## (12)发明专利申请

(10)申请公布号 CN 108701455 A

(43)申请公布日 2018.10.23

(21)申请号 201680081621.4

(22)申请日 2016.12.13

(30)优先权数据

2016-028899 2016.02.18 JP

(85)PCT国际申请进入国家阶段日

2018.08.10

(86)PCT国际申请的申请数据

PCT/JP2016/087096 2016.12.13

(87)PCT国际申请的公布数据

W02017/141530 JA 2017.08.24

(71)申请人 索尼公司

地址 日本东京都

(72)发明人 桐原丽子

(74)专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 王萍 王鹏

(51)Int.Cl.

G10L 15/22(2006.01)

G10L 15/28(2013.01)

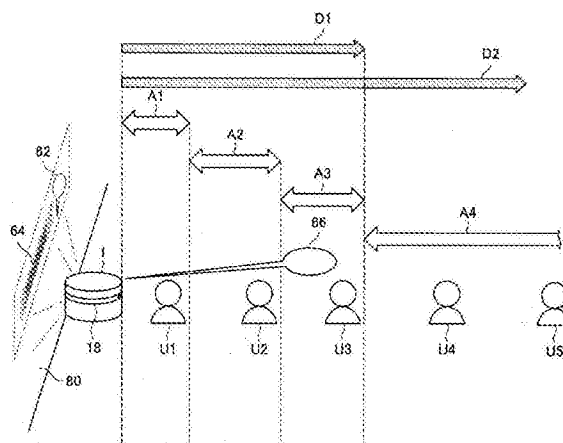
权利要求书2页 说明书16页 附图17页

(54)发明名称

信息处理装置、信息处理方法和程序

(57)摘要

为了提供能够输出更符合用户的意图的动作的信息处理装置、信息处理方法和程序。提供了一种设置有以下的信息处理装置：用于基于距用户的距离来确定对用户的动作的动作决定单元，以及用于输出动作的输出控制单元。



1. 一种信息处理装置,包括:  
动作决定单元,其被配置成基于距用户的距离来决定对所述用户的动作;以及  
输出控制单元,其被配置成输出所述动作。
2. 根据权利要求1所述的信息处理装置,其中,距用户的距离是所述信息处理装置和所述用户之间的距离。
3. 根据权利要求1所述的信息处理装置,还包括:  
用户管理单元,其被配置成指定用作所述动作的目标的目标用户。
4. 根据权利要求3所述的信息处理装置,其中,所述用户管理单元基于用户的状态来指定所述目标用户。
5. 根据权利要求4所述的信息处理装置,其中,所述用户管理单元还基于距用户的距离或用户的状态的持续时间来指定所述目标用户。
6. 根据权利要求3所述的信息处理装置,其中,所述用户管理单元基于用户的状态将用户分类为多个类型,以及  
所述动作决定单元还基于用户的分类类型来决定对所述用户的动作。
7. 根据权利要求6所述的信息处理装置,其中,所述动作决定单元根据距如下用户的距离来决定动作:该用户距被分类为预定类型的用户中的用户距离最小。
8. 根据权利要求7所述的信息处理装置,其中,所述用户管理单元基于被分类为所述预定类型的用户来指定所述目标用户。
9. 根据权利要求3所述的信息处理装置,其中,在用户的讲话被预测出的情况下,所述用户管理单元将讲话被预测出的所述用户指定为所述目标用户。
10. 根据权利要求9所述的信息处理装置,其中,在用户的讲话被预测出的情况下,所述动作决定单元将请求所述用户的响应的动作决定为对所述用户的动作。
11. 根据权利要求1所述的信息处理装置,其中,所述动作决定单元还基于用户的属性信息来决定所述动作。
12. 根据权利要求11所述的信息处理装置,其中,在存在具有预定属性的用户的情况下,所述动作决定单元将不需要所述用户的响应的动作决定为对所述用户的动作。
13. 根据权利要求1所述的信息处理装置,还包括:  
语音识别单元,其被配置成根据由所述动作决定单元决定的动作来开始语音识别。
14. 根据权利要求1所述的信息处理装置,其中,在距用户的距离为第一距离或更小距离的情况下,所述动作决定单元将请求所述用户的响应的动作决定为对所述用户的动作。
15. 根据权利要求14所述的信息处理装置,其中,在距用户的距离大于所述第一距离的情况下,所述动作决定单元将不需要所述用户的响应的动作决定为对所述用户的动作。
16. 根据权利要求15所述的信息处理装置,其中,在距用户的距离为第二距离或更小距离的情况下,所述动作决定单元将不请求所述用户的响应的显示决定为对所述用户的动作。
17. 根据权利要求16所述的信息处理装置,其中,在距用户的距离大于所述第二距离且为第三距离或更小距离的情况下,所述动作决定单元将不请求所述用户的响应的讲话决定为对所述用户的动作。
18. 一种信息处理方法,包括:

由处理器基于距用户的距离来决定对所述用户的动作;以及  
输出所述动作。

19. 一种使计算机实现以下功能的程序:

基于距用户的距离来决定对所述用户的动作的功能;以及  
输出所述动作的功能。

## 信息处理装置、信息处理方法和程序

### 技术领域

[0001] 本发明涉及信息处理装置、信息处理方法和程序。

### 背景技术

[0002] 已知向用户输出动作诸如讲话或信息提供的机器人。例如,在专利文献1中公开了根据周围环境来控制对用户的动作的机器人。

[0003] 引用列表

[0004] 专利文献

[0005] 专利文献1:JP 2008-254122A

### 发明内容

[0006] 技术问题

[0007] 在上述机器人(信息处理装置的示例)中,期望根据用户的意图输出动作。

[0008] 问题的解决方案

[0009] 根据本公开内容,提供了一种信息处理装置,包括:动作决定单元,其被配置成基于距用户的距离来决定对用户的动作;以及被配置成输出动作的输出控制单元。

[0010] 另外,根据本公开内容,提供了一种信息处理方法,包括:由处理器基于距用户的距离来决定对用户的动作;并且输出动作。

[0011] 另外,根据本公开内容,提供了一种使计算机实现以下功能的程序:基于距用户的距离来决定对用户的动作的功能;以及输出动作的功能。发明的有益效果

[0012] 如上所述,根据本公开内容,可以根据用户的意图输出动作。

[0013] 注意,上述效果不一定是限制性的。利用或替代以上效果,可以实现本说明书中描述的效果中的任何一种或者可以从本说明书中掌握的其他效果。

### 附图说明

[0014] [图1]图1是用于描述根据本公开内容的一个实施方式的信息处理装置的概况的说明图。

[0015] [图2]图2是示出根据实施方式的信息处理装置1的配置的示例的框图。

[0016] [图3]图3是示出用户管理单元103的关注分数的具体示例的表格。

[0017] [图4]图4是示出根据实施方式的信息处理装置1的操作示例的流程图。

[0018] [图5]图5是示出根据实施方式的信息处理装置1的操作示例的流程图。

[0019] [图6]图6是示出在根据实施方式的信息处理装置1附近仅存在一个用户的情况下信息处理装置1的操作示例的流程图。

[0020] [图7]图7是示出在根据实施方式的信息处理装置1附近仅存在一个用户的情况下信息处理装置1的操作示例的流程图。

[0021] [图8]图8是示出在根据实施方式的信息处理装置1附近存在多个家庭用户的场景

的说明图。

[0022] [图9]图9是示出在根据实施方式的信息处理装置1附近存在多个家庭用户的场景的说明图。

[0023] [图10]图10是示出在根据实施方式的信息处理装置1附近存在多个家庭用户的场景的说明图。

[0024] [图11]图11是根据实施方式的信息处理装置1在图8至图10中示出的场景中的操作示例的流程图。

[0025] [图12]图12是根据实施方式的信息处理装置1在图8至图10中示出的场景中的操作示例的流程图。

[0026] [图13]图13是根据实施方式的信息处理装置1在图8至图10中示出的场景中的操作示例的流程图。

[0027] [图14]图14是根据实施方式的信息处理装置1在图8至图10中示出的场景中的操作示例的流程图。

[0028] [图15]图15是示出在根据实施方式的信息处理装置1附近存在家庭用户和访客用户的场景的说明图。

[0029] [图16]图16是示出根据实施方式的信息处理装置1在图15中示出的场景中的操作示例的流程图。

[0030] [图17]图17是示出硬件配置示例的说明图。

## 具体实施方式

[0031] 在下文中,将参照附图详细描述本公开内容的(a)优选实施方式。注意,在本说明书和附图中,具有基本相同的功能和结构的结构元件用相同的附图标记表示,并且省略这些结构元件的重复说明。

[0032] 此外,将按照以下顺序进行描述。

[0033] <<1.概述>>

[0034] <<2.配置示例>>

[0035] <<3.操作>>

[0036] <3-1.第一操作示例>

[0037] <3-2.第二操作示例>

[0038] <<4.修改示例>>

[0039] <4-1.第一修改示例>

[0040] <4-2.第二修改示例>

[0041] <4-3.第三修改示例>

[0042] <<5.硬件配置示例>>

[0043] <<6.结论>>

[0044] <<1.概述>>

[0045] 已知输出对用户的动作的机器人(信息处理装置的示例)。机器人对用户执行的动作的示例包括与用户讲话以及向用户显示。此外,动作的示例可以包括讲话或显示询问用户想要什么(“我能为你做什么?”)的内容或者可以包括不请求用户响应的讲话或显示,例

如问候或嗫语。

[0046] 期望上述机器人根据用户的意图输出动作。例如,如果在用户没有任何用户想要机器人做的事情时执行询问用户想要什么的讲话,则用户可能会感到不舒服。

[0047] 就此而言,本实施方式是根据上述内容进行的。根据本实施方式,可以基于距用户的距离通过决定动作来根据用户的意图输出动作。下面将参照图1描述具有这种效果的根据本实施方式的信息处理装置的概述。

[0048] 图1是用于描述根据本公开内容的一个实施方式的信息处理装置的概况的说明图。图1中示出的信息处理装置1检测其周围的用户,并且通过语音或对用户显示来向用户输出动作诸如讲话。此外,信息处理装置1具有能够对用户的讲话(用户响应的示例)执行语音识别/语义分析并且通过语音或显示执行响应的语音UI代理功能。在下文中也将基于用户的讲话和信息处理装置1的响应的一系列交互称为“任务”。

[0049] 信息处理装置1可以识别在从信息处理装置1到图1中示出的距离D1的范围内的用户的语音并且识别用户(人)。距离D1可以是例如300cm。此外,信息处理装置1可以检测在从信息处理装置1到图1中示出的距离D2的范围内的用户(人)。距离D2可以是例如500cm。

[0050] 信息处理装置1的外观没有特别限制,并且可以是例如如图1中示出的圆柱形状,并且例如,信息处理装置1被安装在房间的地板或桌子上。此外,在信息处理装置1中,由发光元件诸如发光二极管(LED)构成的发光单元18被安装成在水平方向上以带的形式围绕侧表面的中心区域。信息处理装置1可以通过使发光单元18的全部或部分发光来向用户通知信息处理装置1的状态。例如,当信息处理装置1与用户交互时,信息处理装置1可以通过使发光单元18的部分在用户的方向或扬声器的方向上发光来使得看起来像信息处理装置1的视线面向正在进行交互的用户。此外,当信息处理装置1正在生成响应或搜索数据时,信息处理装置1可以通过执行控制使得光通过发光单元18在侧表面上方转动来向用户通知它正在处理中。此外,信息处理装置1具有例如在墙80上投影和显示图像的功能,如图1所示,并且除了通过语音执行动作的输出(讲话)以外,还可以通过显示执行动作的输出。

[0051] 例如,图1中示出的信息处理装置1基于信息处理装置1与用户之间的距离来决定并输出对用户的动作。例如,信息处理装置1可以向位于到信息处理装置1的距离是第一距离或更小距离的区域A1中的用户U1显示询问用户U1想要什么(例如,“我能为你做什么?”)的消息62。此外,信息处理装置1向位于到信息处理装置1的距离大于第一距离和第二距离或更小距离的区域A2中的用户U2显示用于呈现信息诸如推荐信息的信息64。此外,信息处理装置1还可以向位于到信息处理装置1的距离大于第二距离和第三距离或更小距离的区域A3中的用户U3输出不请求响应的讲话诸如问候。

[0052] 在此,第三距离可以是识别用户的语音的距离D1。在向位于不能够识别语音的区域A4中的用户输出动作的情况下,即使当用户响应时,信息处理装置1也可能不能够识别语音。就此而言,信息处理装置1可能不向位于不能够识别语音的区域A4中的用户U4和用户U5输出动作。

[0053] 此外,例如,第一距离可以是80cm,第二距离可以是200cm,并且第三距离可以是300cm。

[0054] 如上所述,根据本实施方式的信息处理装置1可以输出基于信息处理装置1与用户之间的距离而决定的动作。信息处理装置1和用户之间的距离被认为是指示用户的意图,并

且例如靠近信息处理装置1的用户可能具有用户想要信息处理装置1去做的事情。因此,根据这样的配置,可以根据用户的意图输出动作。

[0055] 上面已经描述了根据本公开内容的实施方式的信息处理装置1的概况。此外,信息处理装置1的形状不限于图1中示出的圆柱形状,并且可以是例如立方体、球体、多面体等。接下来,将依次描述根据本公开内容的第一实施方式的信息处理装置1的配置示例。

[0056] <<2.配置示例>>

[0057] 图2是示出根据本实施方式的信息处理装置1的配置的示例的框图。如图2所示,信息处理装置1包括控制单元10、通信单元11、声音收集单元12、扬声器13、摄像装置14、距离传感器15、投影单元16、存储单元17以及发光单元18。

[0058] 控制单元10控制信息处理装置1的各个组件。此外,如图2所示,根据本实施方式的控制单元10还作用户识别单元101、讲话预测单元102、用户管理单元103、动作决定单元104、语音识别单元105、语义分析单元106和输出控制单元107。

[0059] 用户识别单元101检测并且识别信息处理装置1附近的用户。例如,用户识别单元101基于由摄像装置14获取的图像和由距离传感器15获取的距离使用公知的面部检测技术、人检测技术等检测用户。此外,用户识别单元101使用公知的面部识别技术等基于由摄像装置14获取的图像来识别用户。例如,用户识别单元101可以通过在存储单元17中存储的已知用户的标识信息与从在图像中检测到的用户中提取的信息之间进行匹配来识别用户。

[0060] 在检测到用户的情况下,用户识别单元101向讲话预测单元102提供信息诸如检测到的用户的数目和用户的位置。此外,在检测到用户的情况下,用户识别单元101向用户管理单元103提供根据检测到的用户的数目和用户标识获得的标识信息。

[0061] 讲话预测单元102对由用户识别单元101检测到的用户执行用户讲话预测(例如,预测用户具有与信息处理装置1交互的意图)。可以用各种方法来执行通过讲话预测单元102对用户的讲话预测。

[0062] 例如,讲话预测单元102可以基于用于传达用户的显式交互意图的行为来预测用户的讲话。例如,在从由将在后面描述的声音收集单元12收集的用户的语音中检测到预定激活词(例如,“你好代理”等)的情况下,讲话预测单元102可以预测用户的讲话。此外,在通过触摸传感器(未示出)或物理按钮检测到触摸信息处理装置1的用户的情况下,讲话预测单元102可以预测用户的讲话。此外,在检测到触摸信息处理装置1的用户而用户识别单元101没有检测到用户的情况下,讲话预测单元102可以预测用户的讲话。此外,在检测到挥动他/她的手的用户的情况下,讲话预测单元102可以基于由将在后面描述的摄像装置14或距离传感器15获得的数据预测用户的讲话。此外,在检测到拍他/她的手的用户的情况下,讲话预测单元102基于由将在后面描述的摄像装置14或距离传感器15获得的数据或者由声音收集单元12收集到的声音来预测用户的讲话。此外,在用户直接来到信息处理装置1的前面的情况下,讲话预测单元102可以预测用户的讲话(例如,在信息处理装置1与用户之间的距离等于或小于预定值的情况下)。

[0063] 同时,用于传达上述显式交互意图的行为可能给用户带来负担。例如,在通过检测预定激活词来预测讲话的情况下,在开始他/她原先必须做什么之前,用户必须说出激活词,并且因此存在讲话负担。此外,在通过检测触摸信息处理装置1的用户来预测讲话的情况下,用户应当位于可以触摸信息处理装置1的范围内,并且触摸或操纵信息处理装置1也

变成用户的负担。

[0064] 就此而言,即使在用户没有明确执行如上所述向信息处理装置1传达交互意图的行为的情况下,根据本实施方式的讲话预测单元102也可以预测用户的讲话。根据以下描述的讲话预测示例,与如上所述用户明确地执行向信息处理装置1传达交互意图的行为的情况相比,减少了用户的负担。

[0065] 例如,讲话预测单元102可以基于由将在后面描述的摄像装置14或距离传感器15获得的数据来估计用户的视线的方向或面部的方向,或者可以在用户的视线或面部面向(用户注视)信息处理装置1的情况下估计用户的讲话。此外,在用户的视线或面部面向(用户注视)信息处理装置1达预定时间段或更长时间的情况下,讲话预测单元102可以预测用户的讲话。

[0066] 此外,在由距离传感器15获得的距用户的距离为预定值或更小的情况下,讲话预测单元102可以预测用户的讲话。

[0067] 此外,在用户以预定值或更高的速度靠近信息处理装置1的情况下,讲话预测单元102可以基于由距离传感器15获得的数据来指定用户的移动方向或速度并且预测用户的讲话。

[0068] 此外,讲话预测单元102可以确定作为与用户的一系列交互的任务(例如日程登记、信息检索等)是否结束并且在任务未结束(正在执行)的情况下预测用户的讲话。

[0069] 用户管理单元103基于由用户识别单元101检测到的用户的标识信息来执行与用户有关的管理。

[0070] 例如,用户管理单元103可以基于用户的标识信息和存储单元17中存储的用户的标识信息来管理用户的属性信息。例如,用户的属性可以是以下中的任何一个:属于具有信息处理装置1的家庭的家庭用户、除了其用户信息被存储在存储单元17中的家庭用户以外的已知用户或者其用户信息未被存储在存储单元17中的访客用户。

[0071] 此外,在由用户识别单元101检测到多个用户的情况下,用户管理单元103可以将目标用户指定为动作的目标。用户管理单元103可以从多个检测到的用户之中指定一个用户作为目标用户或者可以指定多个用户作为目标用户。

[0072] 用户管理单元103可以基于用户的状态针对每个检测到的用户指定指示与信息处理装置1的交互意图的关注分数,并且基于所指定的关注分数来指定目标用户。

[0073] 图3是示出用户管理单元103的关注分数的具体示例的表格。如图3所示,用户管理单元103可以基于用户的状态、距用户的距离以及状态的持续时间(状态持续时间)来指定关注分数。此外,在图3中,is\_watching指示用户的视线或面部面向(用户注视)信息处理装置1的状态。此外,在图3中,not\_watching指示用户的视线或面部不面向(用户不注视)信息处理装置1的状态。

[0074] 根据图3中示出的表格,用户管理单元103可以针对每个检测到的用户指定关注分数并且将具有最高关注分数的用户指定为目标用户。此外,在所有检测到的用户的关注分数为0或更小的情况下,用户管理单元103可以指定不存在目标用户(无动作目标)。

[0075] 此外,用户管理单元103对目标用户的指定不限于以上示例。例如,在由讲话预测单元102预测出用户的讲话的情况下,可以将用户指定为目标用户。此外,用户管理单元103可以基于用户的状态将用户分类为多个类型,并且基于用户的分类类型来指定目标用户。

将在后面将这样的示例作为第二操作示例描述。

[0076] 此外,用户管理单元103管理对每个用户的动作的输出信息。例如,在将在后面描述的输出控制单元107的控制下输出针对用户的动作的情况下,用户管理单元103将包括向用户输出动作的时间的动作的输出信息存储在存储单元17中。此外,用户管理单元103可以基于输出信息来识别目标用户。例如,在预定时间段内向同一用户输出由动作决定单元104决定的动作的情况下,用户不被指定为目标用户,并且因此防止向同一用户频繁地输出相同的动作。

[0077] 动作决定单元104基于距用户的距离来决定对用户的动作。例如,动作决定单元104可以基于从距离传感器15获得的信息处理装置1与用户之间的距离来决定动作。

[0078] 例如,由动作决定单元104决定的动作可以是请求来自用户的响应的动作诸如询问用户想要什么的显示或讲话,或者可以是不请求来自用户的响应的动作诸如信息呈现、问候或啜语。

[0079] 此外,信息呈现的内容可以是例如与每个用户对应的信息,诸如未读电子邮件(未读消息)、社交服务的新到信息以及用户经常访问的网站的更新信息。此外,在检测到仅一个用户作为用户的情况下,可以显示私人信息诸如未读电子邮件。此外,在检测到多个用户的情况下,可以显示允许在位于同一地点的人之间共享的信息例如现在可以观看的电视节目。

[0080] 在距用户的距离是第一距离或更小距离的情况下(例如,在用户位于以上参照图1描述的区域A1中的情况下),动作决定单元104可以将请求用户响应的动作决定为对用户的动作。根据这样的配置,可以更顺利地开始与更靠近信息处理装置1的用户的交互。

[0081] 此外,在距用户的距离大于第一距离的情况下,动作决定单元104可以将不请求用户响应的动作决定为对用户的动作。

[0082] 此外,在距用户的距离大于第一距离并且等于或小于第二距离的情况下(例如,在用户位于以上参照图1描述的区域A2中的情况下),动作决定单元104将不请求用户响应的显示决定为对用户的动作。此外,在距用户的距离大于第二距离并且等于或小于第三距离的情况下(例如,在用户位于以上参照图1描述的区域A3中的情况下),动作决定单元104将不请求用户响应的讲话决定为对用户的动作。例如,在信息处理装置1的输出功能中,在通过讲话的动作比通过显示的动作更可能被远离的用户识别的情况下,根据这样的配置,对用户的动作更可能被用户识别。

[0083] 此外,在讲话预测单元102预测出用户的讲话并且用户管理单元103将用户指定为目标用户的情况下,动作决定单元104可以将请求用户响应的动作决定为对用户的动作。根据这样的配置,可以更顺利地开始与其讲话已经被预测出的用户的交互。

[0084] 此外,在由用户识别单元101检测到多个用户的情况下,动作决定单元104可以基于到目标用户的距离来决定动作。此外,在由用户识别单元101检测到多个用户的情况下,动作决定单元104可以基于由用户管理单元103分类的用户类型来决定对用户的动作。例如,动作决定单元104可以根据到距被分类为预定类型的用户中的用户距离最小的用户的距离来决定对用户的动作。

[0085] 此外,在用户管理单元103基于以上参照图3描述的关注分数来指定目标用户的情况下,动作决定单元104可以基于关注分数来决定动作。例如,在其关注分数为预定值或更

大的用户为目标用户的情况下,动作决定单元104可以将请求用户响应的动作决定为对用户的动作。

[0086] 此外,动作决定单元104可以基于由用户管理单元103管理的用户的属性信息来决定动作。例如,在检测到的用户中存在具有预定属性的用户的情况下,动作决定单元104可以将不请求用户响应的动作决定作为对用户的动作。具有预定属性的用户可以是例如不期望给予信息处理装置1操纵权限的用户或者可以是例如已知用户或访客用户。根据这种配置,可以限制信息处理装置1使其不被例如除了家庭用户以外的用户操纵。

[0087] 动作决定单元104向输出控制单元107提供与所决定动作有关的动作信息。此外,在所决定动作是请求用户响应的动作的情况下,动作决定单元104可以向语音识别单元105给出指示所决定动作是请求用户响应的动作的通知。

[0088] 语音识别单元105识别由将在后面描述的声音收集单元12收集的用户的语音,将语音转换成字符串并且获取讲话文本。此外,还可以基于语音识别单元105的语音特征来识别正在讲话的人,或者估计语音的来源即扬声器的方向。

[0089] 此外,在所获取的讲话文本中包括预定激活词的情况下,语音识别单元105可以向讲话预测单元102给出指示已经检测到激活词的通知。此外,语音识别单元105可以将激活词与其他讲话文本进行比较并且相对于噪声更可靠地检测激活词。

[0090] 此外,在动作决定单元104决定将请求用户响应的动作作为对用户的动作的情况下,语音识别单元105可以开始除了上述激活词以外的语音识别。根据这样的配置,语音识别单元105可以在输出动作之后顺利地识别由用户执行的讲话响应。

[0091] 语义分析单元106使用自然语言处理等对由语音识别单元105获取的讲话文本执行语义分析。向输出控制单元107提供语义分析单元106的语义分析的结果。

[0092] 输出控制单元107控制将在后面描述的扬声器13、投影单元16和发光单元18,使得输出由动作决定单元104决定的对用户的动作。例如,输出控制单元107可以控制扬声器13使得语音的讲话被输出作为动作,或者可以控制投影单元16使得消息或信息的显示被输出作为动作。

[0093] 输出控制单元107可以向由用户管理单元103指定的目标用户输出动作。此外,在输出对用户的动作的情况下,输出控制单元107可以向用户管理单元103提供与输出动作有关的信息和与向其输出动作的用户有关的信息。

[0094] 输出控制单元107根据从语义分析单元106提供的语义分析结果来控制对用户的讲话的响应或与任务有关的输出诸如用户所需的操作。例如,在用户的讲话是要获得“明天的天气”的情况下,输出控制单元107从网络上的天气预报服务器获取与“明天的天气”有关的信息,并且控制扬声器13、投影单元16或发光单元18以使得所获取的信息被输出。在这样的示例中,可以认为任务已经在输出与“明天的天气”有关的信息的时间点处结束。

[0095] 通信单元11与外部装置交换数据。例如,通信单元11经由通信网络(未示出)建立与预定服务器(未示出)的连接并且接收由输出控制单元107进行的输出控制所需的信息。

[0096] 声音收集单元12具有收集环境声音并且向控制单元10输出所收集的声音作为音频信号的功能。此外,例如,声音收集单元12可以由一个或更多个麦克风来实现。

[0097] 扬声器13具有在输出控制单元107的控制下将音频信号转换成语音并且输出语音的功能。

[0098] 摄像装置14具有利用安装在信息处理装置1中的成像透镜对周围区域进行成像并且向控制单元10输出所捕获的图像的功能。此外,摄像装置14可以例如由360度摄像装置、广角摄像装置等实现。

[0099] 距离传感器15具有测量信息处理装置1与用户或用户周围的人之间的距离的功能。距离传感器15例如由光学传感器(基于发光/接收定时处的相位差信息来测量到对象的距离的传感器)来实现。

[0100] 投影单元16是显示装置的示例,并且具有通过将图像投影(放大)到墙壁或屏幕上来进行显示的功能。

[0101] 存储单元17存储使信息处理装置1的各个组件起作用的程序和参数。例如,存储单元17可以存储与用户有关的信息诸如用户的标识信息和用户的属性信息、要输出的讲话内容、与动作有关的信息诸如显示内容等。

[0102] 发光单元18由发光元件诸如LED实现,并且可以执行全部照亮、部分照亮、闪烁、照亮位置控制等。例如,发光单元18可以根据控制单元10的控制通过在由语音识别单元105识别的扬声器的方向上照亮其部分来使得看起来像视线面向扬声器的方向。

[0103] 以上具体描述了根据本实施方式的信息处理装置1的配置。此外,图2中示出的信息处理装置1的配置是示例,并且本实施方式不限于此。例如,信息处理装置1还可以包括红外(IR)摄像装置、深度摄像装置、立体摄像装置、人类传感器等,以便获得与周围环境有关的信息。此外,信息处理装置1还可以包括触摸面板显示器、物理按钮等作为用户界面。此外,信息处理装置1中安装的声音收集单元12、扬声器13、摄像装置14、发光单元18等的安装位置不受特别限制。此外,根据本实施方式的控制单元10的功能可以通过经由通信单元11连接的另一信息处理装置来执行。

[0104] <<3.操作>>

[0105] 接下来,将参照图4至图16描述根据本实施方式的信息处理装置1的几个操作示例。

[0106] <3-1.第一操作示例>

[0107] 首先,将参照图4和图5来将执行请求用户响应的动作并且开始语音识别的操作示例作为第一操作示例描述。

[0108] 图4和图5是示出根据本实施方式的信息处理装置1的操作示例的流程图。如图4所示,首先,控制单元10重复用户检测处理(S110),直到在信息处理装置1附近检测到用户。在检测到的用户的数目是1的情况下(S110中为是并且S120中为是),如果讲话预测单元102预测出用户的讲话(S130中为是),则指定并输出请求用户响应的动作,并且开始语音识别(S140)。此外,对用户的讲话的预测不限于基于用户明确地传达交互意图的行为的预测,并且可以是基于如上所述的用户的注视、距用户的距离或用户的移动速度的预测。在本操作示例中,在没有预测出用户的讲话的情况下(S130中为否),处理可以返回到步骤S110。

[0109] 在输出动作之后,如果用户的讲话和作为通过信息处理装置1的响应进行的一系列交互的任务结束(S150中为是),则处理结束。此外,如果任务和用户的讲话均未结束(S150中为否并且S160中为否),则继续进行语音识别直到任务结束。另一方面,在任务未结束但是用户的讲话结束的情况下(S150中为否并且S160中为是),处理结束。此外,例如,控制单元10可以确定用户的讲话是否结束,并且在非讲话时间段持续预定时间段(例如15秒)

的情况下,控制单元10可以确定用户的讲话结束。

[0110] 此外,在用户识别单元101检测到多个用户的情况下(S120中为否),用户识别单元101和用户管理单元103执行用户识别(图5中的S210)。接下来,用户管理单元103对每个用户指定以上参照图3描述的关注分数(S220),并且将具有最高关注分数的用户识别为目标用户(S230)。此外,在讲话预测单元102基于用户的明确传达交互意图的行为(激活词的讲话等)预测用户的讲话的情况下,在步骤S230中,用户可以被指定为目标用户。

[0111] 接下来,由动作决定单元104指定请求用户响应的动作,根据输出控制单元107的控制输出动作,并且开始语音识别(S240)。

[0112] 在输出动作之后,如果用户的讲话和作为通过信息处理装置1的响应进行的一系列交互的任务结束(S250中为是),则处理结束。此外,在任务和用户的讲话均未结束的情况下(S250中为否并且S260中为否),处理返回到步骤S210,同时继续进行语音识别。另一方面,在任务未结束但是用户的讲话结束的情况下(S250中为否并且S260中为是),处理结束。此外,控制单元10可以如上所述确定用户的讲话是否结束。

[0113] <3-2.第二操作示例>

[0114] 在上述第一操作示例中,已经描述了基于讲话预测执行请求用户响应的动作并且开始语音识别的操作示例。接下来,在将在下面描述的第二操作示例中,将具体描述根据距用户的距离的动作的具体示例。

[0115] 首先,将参照图6和图7描述由信息处理装置1检测到一个用户的情况下的操作示例,并且将参照图8至图16描述由信息处理装置1检测到多个用户的情况下的操作示例。

[0116] 图6和图7是示出在根据本实施方式的信息处理装置1附近仅存在一个用户的情况下信息处理装置1的操作示例的流程图。如图6所示,首先,控制单元10检测并且识别用户(S302)。在用户管理单元1031确定用户注视信息处理装置1的状态(is\_watching)持续预定时间段或更长时间的情况下(S304中为是),处理进行到步骤S306。另一方面,例如在未确定用户注视信息处理装置1的状态(is\_watching)持续预定时间段或更长时间的情况下(S304中为否),处理结束。

[0117] 在步骤S306中,在用户管理单元103确定用户是家庭用户的情况下(S306中为是),处理进行到步骤S308。另一方面,在用户管理单元103确定用户不是家庭用户的情况下,处理进行到图7中的步骤S350。

[0118] 在动作决定单元104确定用户位于区域A1中的情况下(S308中为是),在步骤S310中,用户管理单元103确定用户在预定时间段内是否已经被询问用户想要什么。在用户在预定时间段内没有被询问用户想要什么的情况下(S310中为否),输出控制单元107显示询问要做什么的消息并且开始由语音识别单元105进行的语音识别(S312)。此外,此处,在下文中,开始语音识别之后的处理可以与以上参照图4描述的步骤S150和步骤S160中的处理类似,并且省略其描述。在用户在预定时间段内已经被询问用户想要什么的情况下(S310中为是),处理结束。

[0119] 在动作决定单元104确定用户位于区域A2中的情况下(S308中为否并且S314中为是),处理进行到步骤S316。在步骤S316中,控制单元10确定是否存在给用户的推荐信息(能够呈现信息的信息)。在不存在推荐信息的情况下(S316中为否),处理结束,并且在存在推荐信息的情况下(S316中为是),输出控制单元107显示推荐信息(S318)。

[0120] 在动作决定单元104确定用户位于区域A3中的情况下(S308中为否、S314中为否并且S320中为是),处理进行到步骤S322。在步骤S322中,用户管理单元103确定用户在预定时间段内是否已经被问候。在用户在预定时间段内没有被询问用户想要什么的情况下(S322中为否),输出控制单元107使得通过语音讲话向用户发送问候(S324)。在步骤S322中在预定时间段内用户已经被问候的情况下(S322中为是),处理结束。

[0121] 此外,在区域A1至区域A3中的任何一个中不存在用户的情况下(S308中为否、S314中为否并且S320中为否),处理结束。

[0122] 在步骤S306中在用户管理单元103确定用户不是家庭用户的情况下(S306中为否),处理进行到图7中示出的步骤S350。

[0123] 在用户管理单元103确定用户是已知用户并且动作决定单元104确定用户位于区域A1至区域A3中的任何一个中的情况下(S350中为是并且S352中为是),处理进行到S354。由于步骤S354和步骤S356中的处理与步骤S322和步骤S324中的处理类似,所以省略其描述。另一方面,在确定在区域A1至区域A3中的任何一个中不存在用户的情况下(S352中为否),处理结束。

[0124] 在用户管理单元103确定用户是已知用户并且动作决定单元104确定用户位于区域A1至区域A3中的任何一个中的情况下(S350中为是并且S352中为是),处理进行到S354。由于步骤S354和步骤S356中的处理与步骤S322和步骤S324中的处理类似,所以省略其描述。

[0125] 在用户管理单元103确定用户既不是家庭用户也不是已知用户(访客用户)的情况下(步骤S350中为否),处理进行到步骤S358。在步骤S358中在动作决定单元104确定用户位于区域A1或区域A2中的情况下(S358中为是),由输出控制单元107执行向用户的单边语音讲话输出(S360),并且处理结束。由信息处理装置1嘀咕的内容的示例如下。例如,可以逐一讲出以下示例。

[0126] “那?有一个陌生人……”

[0127] “他是客户吗?”

[0128] “我莫名感到紧张……”

[0129] “好吧,我觉得自己像个人……”。

[0130] “这是很好的衣服。”

[0131] 另一方面,在步骤S358中在动作决定单元104确定用户不位于区域A1或区域A2中的情况下(S358中为否),处理结束。

[0132] 以上已经描述了由信息处理装置1检测到一个用户的情况下的操作示例。接下来,将参照图8至图16描述由信息处理装置1检测到多个用户的情况下的操作示例。此外,在下面将参照图8至图14描述在信息处理装置1附近仅存在家庭用户的情况下的操作示例,并且然后将参照图15和图16描述在信息处理装置1附近存在访客用户的情况下的操作示例。

[0133] 图8至图10是示出在信息处理装置1附近存在多个家庭用户的场景的说明图。在图8至图10中,在信息处理装置1附近存在家庭用户X、家庭用户Y和家庭用户Z。此外,例如,家庭用户X和家庭用户Y注视着信息处理装置1,并且家庭用户Z注视着电视接收机8而未注视信息处理装置1。在以下描述中,图8至图10被认为是适当的。

[0134] 图11至图14是示出根据本实施方式的信息处理装置1在图8至图10中示出的场景

中的任何一个中的操作示例的流程图。如图11所示,首先,用户识别单元101和用户管理单元103识别家庭用户X、家庭用户Y和家庭用户Z (S402)。然后,用户管理单元103将家庭用户X、家庭用户Y和家庭用户Z的类型分类为存在用户 (S402)。

[0135] 此处,假设家庭用户X、家庭用户Y和家庭用户Z中的家庭用户X的状态变成is\_watching (S406),家庭用户Y的状态变成is\_watching (S408),并且家庭用户Z的状态是未检测到面部的状态(not\_watching) (S410)。然后,在确定家庭用户X的状态作为is\_watching持续了预定时间段(例如2秒)的时间点(S412)处,将处于is\_watching状态下的家庭用户X和家庭用户Y的类型分类作为注视用户(S414)。此外,在家庭用户Z的状态是未检测到面部的状态(not\_watching)时经过了预定时间段(例如一分钟)的情况下,可以从存在用户中删除家庭用户Z。

[0136] 然后,确定注视用户是否包括访客用户,但是在本操作示例中,由于仅家庭用户X和家庭用户Y被分类作为注视用户,因此确定注视用户不包括访客用户(S416)。此外,将参照图15和图16来描述注视用户包括访客用户的示例。

[0137] 如图8所示,在注视用户中的一个位于区域A1中的情况下(S418中为是),处理进行到图12中的步骤S420。在家庭用户X和家庭用户Y两者在预定时间段内未被询问他们想要什么的情况下(S420中为否并且S422中为否),输出控制单元107使用于询问家庭用户X和家庭用户Y想要什么的消息被显示,并且由语音识别单元105开始语音识别(S424)。另外,在家庭用户X未被询问但是家庭用户Y在预定时间段内被询问家庭用户Y想要什么的情况下(S420中为否并且S422中为是),输出控制单元107使用于询问家庭用户X想要什么的消息被显示,并且由语音识别单元105开始语音识别(S426)。此外,在家庭用户X被询问但是家庭用户Y在预定时间段内未被询问家庭用户Y想要什么的情况下(S420中为是并且S428中为否),输出控制单元107使用于询问家庭用户Y想要什么的消息被显示,并且由语音识别单元105开始语音识别(S430)。此外,在家庭用户X和家庭用户Y两者在预定时间段内均被询问他们想要什么的情况下(S420中为是并且S428中为是),处理结束。

[0138] 此外,如图9所示,在注视用户不位于区域A1中的情况下(S418中为否),处理进行到图13中的步骤S440。在步骤S440中在注视用户中的任何一个不位于区域A2中的情况下(S440中为是),处理进行到步骤S442。由于步骤S422和步骤S444的处理与图6中的步骤S316和步骤S318的处理类似,所以省略其描述。

[0139] 此外,如图10所示,在注视用户不位于区域A1和区域A2中的情况下(S440中为否),处理进行到图14中的步骤S460。在步骤S460中在注视用户不位于区域A3中的情况下(S460中为否),处理结束,并且在注视用户中的任何一个位于区域A3中的情况下(S460中为是),处理进行到步骤S462。

[0140] 在家庭用户X和家庭用户Y两者在预定时间段内都未被问候的情况下(S462中为否并且S464中为否),输出控制单元107使得通过语音讲话来输出对家庭用户X和家庭用户Y的问候(S466)。另外,在家庭用户X在预定时间段内未被问候但是家庭用户Y被问候的情况下(S462中为否并且S464中为是),输出控制单元107使得通过语音讲话输出对家庭用户X的问候(S468)。另外,在家庭用户X在预定时间段内被问候但是家庭用户Y未被问候(S462中为是并且S470中为否)的情况下,输出控制单元107使得通过语音讲话输出对家庭用户Y的问候(S472)。此外,在家庭用户X和家庭用户Y在预定时间段内被询问他们想要什么的情况下

(S462中为是并且S470中为是),处理结束。

[0141] 上面已经描述了在信息处理装置1附近仅存在多个家庭用户的情况下的操作示例。接下来,将参照图15和图16描述在信息处理装置1附近存在访客用户的情况下的操作示例。

[0142] 图15是示出在信息处理装置1附近存在家庭用户和访客用户的场景的说明图。在图15中,在信息处理装置1附近存在家庭用户X和访客用户P和访客用户Q。此外,家庭用户X和访客用户P和访客用户Q正在注视信息处理装置1。在以下描述中,图15将被认为是适当的。

[0143] 图16是示出根据本实施方式的信息处理装置1在图15中示出的场景中的操作示例的流程图。如图16所示,首先识别访客用户P和访客用户Q(S502)。例如,用户识别单元101检测用户P和用户Q,并且用户管理单元103指定用户P和用户Q的属性是访客用户。此外,用户识别单元101和用户管理单元103中的每一个识别家庭用户X(S504)。然后,用户管理单元103将访客用户P和访客用户Q以及家庭用户X的类型分类作为存在用户(S506)。

[0144] 在此,状态变成按照访客用户P和访客用户Q以及家庭用户X的顺序的is\_watching(S508至S512)。此外,在确定访客用户P的状态在is\_watching状态下持续了预定时间段(例如2秒)的时间点处(S514),处于is\_watching状态下的访客用户P和访客用户Q以及家庭用户X的类型被分类作为注视用户(S516)。

[0145] 然后,确定访客用户是否被包括在注视用户中,但在本操作示例中,由于访客用户P和访客用户Q被分类作为注视用户,因此确定访客用户被包括在注视用户中(S518)。

[0146] 例如,如图15所示,在注视用户中的任何一个位于区域A1或区域A2中的情况下(S520中为是),处理进行到步骤S522。此外,由于步骤S522中的处理与步骤S360中的处理类似,所以省略其描述。另一方面,在注视用户不位于区域A1和区域A2中的任何一个中的情况下(S520中为否),处理结束。

[0147] 以上已经描述了在信息处理装置1附近存在家庭用户和访客用户的情况下的操作示例。如上所述,在信息处理装置1附近存在访客用户的情况下,不请求用户响应的动作(嘀咕讲话)被决定作为对用户的动作,并且因此可以限制信息处理装置1使其不被除了家庭用户以外的用户操纵。

[0148] <<4. 修改示例>>

[0149] 以上已经描述了本公开内容的实施方式。下面将描述本公开内容的实施方式的几个修改示例。此外,将在下面描述的修改示例中的每一个可以单独应用于本公开内容的实施方式或者可以组合应用于本公开内容的实施方式。此外,可以替代本公开内容的实施方式中描述的配置来应用每个修改示例,或者可以将每个修改示例另外应用于本公开内容的实施方式中描述的配置。

[0150] <4-1. 第一修改示例>

[0151] 在以上示例中,动作决定单元104基于距用户的距离来决定对用户的动作,但是本公开内容不限于该示例。以下将描述另一动作决定示例。

[0152] 例如,可以基于用户的注视时间来决定对用户的动作。例如,如果注视时间在第一时间段(例如1秒)内,则可以在注视时间大于第一时间段并且小于第二时间段(例如3秒)的情况下决定用于询问要做什么的动作,并且可以在注视时间大于第二时间段并且小于第三

时间段(例如5秒)的情况下决定用于说明如何使用它的动作。

[0153] 此外,可以基于用户靠近主体的速度来决定对用户的动作。例如,如果用户靠近主体的速度低于第一速度(例如2km/h),则可以在速度高于第一速度并且低于第二速度(例如3km/h)的情况下决定用于询问要做什么的动作。此外,如果用户靠近主体的速度高于第二速度并且低于第三速度,则可以立即开始语音识别而不输出动作。

[0154] <4-2. 第二修改示例>

[0155] 此外,在以上实施方式中,已经将询问用户想要什么的显示或讲话作为请求用户响应的动作的示例进行描述,但是本公开内容不限于这样的示例。例如,信息处理装置1可以基于用户的信息来估计用户想要什么,并且输出与检查估计要完成的内容有关的显示或讲话。例如,对于每天在预定时间检查次日的日程的用户,信息处理装置1可以估计用户想要的是次日日程检查,并且在显示次日日程的同时显示用于检查要做的事是否正确的消息。根据这样的配置,实现了与用户的更顺利的交互。

[0156] <4-3. 第三修改示例>

[0157] 此外,在以上实施方式中,已经描述了仅向目标用户输出动作的示例,但是本公开内容不限于这样的示例。例如,信息处理装置1可以向不是目标用户的用户输出讲话诸如“请等待一会直到与前一用户的交互完成”。此外,信息处理装置1可以不开始针对不是目标用户的用户的交互,而是执行指示其注意到用户具有交互意图的输出(例如,发光单元18在用户的方向上发光等)。

[0158] <<4. 硬件配置示例>>

[0159] 以上已经描述了本公开内容的实施方式。信息处理诸如处理诸如用户识别处理、讲话预测处理、用户管理处理、动作决定处理、语音识别处理、语义分析处理和输出控制处理与软件和信息处理装置1至信息处理装置3协作实现。信息处理装置1000的硬件配置示例将被描述为作为根据本实施方式的信息处理装置的信息处理装置1至信息处理装置3的硬件配置的示例。

[0160] 图17是示出信息处理装置1000的硬件配置的示例的说明图。如图11所示,信息处理装置1000包括中央处理单元(CPU) 1001、只读存储器(ROM) 1002、随机存取存储器(RAM) 1003、输入装置1004、输出装置1005、存储装置1006、成像装置1007和通信装置1008。

[0161] CPU 1001用作操作处理装置和控制装置,并且根据各种程序来控制信息处理装置1000的整体操作。此外,CPU 1001可以是微处理器。ROM 1002存储由CPU 1001使用的程序、操作参数等。RAM 1003临时存储要在CPU 1001的执行中使用的程序、在其执行中适当改变的参数等。这些组件经由包括CPU总线等的主机总线相互连接。控制单元10的功能主要通过CPU 1001、ROM 1002和RAM 1003以及软件的协作来实现。

[0162] 输入装置1004包括:用于输入信息的输入装置诸如鼠标、键盘、触摸面板、按钮、麦克风、开关、杠杆等,用于基于用户的输入生成输入信号并且向CPU 1001输出输入信号的输入控制电路。通过操作输入装置1004,信息处理装置1000的用户可以向信息处理装置1000输入各种数据或给出执行处理操作的指令。

[0163] 输出装置1005包括显示装置诸如例如液晶显示(LCD)装置、OLED装置、透视显示器或灯。此外,输出装置1005包括音频输出装置诸如扬声器和头戴式耳机。例如,显示装置显示所捕获的图像、生成图像等。另一方面,音频输出装置将语音数据等转换成语音并且输出

语音。例如,输出装置1005与以上参照图2描述的扬声器13、投影单元16和发光单元18对应。

[0164] 存储装置1006是用于存储数据的装置。存储装置1006可以包括存储介质、用于将数据记录在存储介质中的记录装置,用于从存储介质中读取数据的读取装置、用于删除存储介质中记录的数据的删除装置等。存储装置1006存储由CPU 1001执行的程序和各种数据。存储装置1006与以上参照图2描述的存储单元17对应。

[0165] 成像装置1007包括成像光学系统,诸如用于收集光的拍摄透镜和变焦透镜以及信号转换元件诸如电荷耦合装置(CCD)或互补金属氧化物半导体(CMOS)。成像光学系统收集从对象发出的光并且在信号转换单元中形成对象图像,并且信号转换元件将所形成的对象图像转换成电图像信号。成像装置1007与以上参照图2描述的摄像装置14对应。

[0166] 通信装置1008例如是由通信装置等构成的用于建立与通信网络的连接的通信接口。此外,通信装置1008可以包括与无线局域网(LAN)兼容的通信装置、与长期演进(LTE)兼容的通信装置、执行有线通信的有线通信装置或蓝牙(注册商标)通信装置。通信装置1008例如与以上参照图2描述的通信单元11对应。

[0167] <<6. 结论>>

[0168] 如上所述,根据本公开内容,可以基于距用户的距离通过决定动作来根据用户的意图输出动作。

[0169] 以上已经参照附图描述了本公开内容的优选实施方式,然而本公开内容不限于以上示例。本领域技术人员可以在所附权利要求的范围内发现各种改变和修改,并且应该理解,各种改变和修改将自然落入本公开内容的技术范围内。

[0170] 例如,以上实施方式中的各个步骤不一定必须按照如流程图所描述的顺序按时间顺序来处理。例如,以上实施方式的处理中的各个步骤可以按照与流程图所描述的顺序不同的顺序处理或者可以并行处理。

[0171] 此外,在以上实施方式中,已经描述了距用户的距离是信息处理装置与用户之间的距离的示例,但是本公开内容不限于这样的示例。例如,在信息处理装置控制具有扬声器或显示单元的另一终端的输出的情况下,距用户的距离可以是另一终端与用户之间的距离。

[0172] 此外,根据以上实施方式,还可以提供使硬件诸如CPU 1001、ROM 1002和RAM 1003执行与上述信息处理装置1的功能类似的功能的计算机程序。此外,还提供了其中记录有计算机程序的记录介质。

[0173] 此外,在本说明书中描述的效果仅仅是说明性或示例性的效果,而不是限制性的。即利用或替代以上效果,根据本公开内容的技术可以实现本领域技术人员从本说明书的描述中清楚的其他效果。

[0174] 另外,本技术还可以被配置如下。

[0175] (1) 一种信息处理装置,包括:

[0176] 动作决定单元,其被配置成基于距用户的距离来决定对所述用户的动作;以及

[0177] 输出控制单元,其被配置成输出所述动作。

[0178] (2) 根据(1)所述的信息处理装置,其中,距所述用户的距离是所述信息处理装置和所述用户之间的距离。

[0179] (3) 根据(1)或(2)所述的信息处理装置,还包括:

- [0180] 用户管理单元,其被配置成指定用作所述动作的目标的目标用户。
- [0181] (4) 根据(3)所述的信息处理装置,其中,所述用户管理单元基于所述用户的状态来指定所述目标用户。
- [0182] (5) 根据(4)所述的信息处理装置,其中,所述用户管理单元还基于距所述用户的距离或所述用户的状态的持续时间来指定所述目标用户。
- [0183] (6) 根据(3)至(5)中任一项所述的信息处理装置,其中,所述用户管理单元基于所述用户的状态将所述用户分类为多个类型,并且
- [0184] 所述动作决定单元还基于所述用户的分类类型来决定对所述用户的动作。
- [0185] (7) 根据(6)所述的信息处理装置,其中,所述动作决定单元根据到距被分类为预定类型的用户中的用户距离最小的用户的距离来决定动作。
- [0186] (8) 根据(7)所述的信息处理装置,其中,所述用户管理单元基于被分类为所述预定类型的用户来指定所述目标用户。
- [0187] (9) 根据(3)至(8)中任一项所述的信息处理装置,其中,在所述用户的讲话被预测出的情况下,所述用户管理单元将其讲话被预测出的用户指定为所述目标用户。
- [0188] (10) 根据(9)所述的信息处理装置,其中,在所述用户的讲话被预测出的情况下,所述动作决定单元将请求所述用户的响应的动作决定为对所述用户的动作。
- [0189] (11) 根据(1)至(10)中任一项所述的信息处理装置,其中,所述动作决定单元还基于所述用户的属性信息来决定所述动作。
- [0190] (12) 根据(11)所述的信息处理装置,其中,在存在具有预定属性的用户的情况下,所述动作决定单元将不需要所述用户的响应的动作决定为对所述用户的动作。
- [0191] (13) 根据(1)至(12)中任一项所述的信息处理装置,还包括:
- [0192] 语音识别单元,其被配置成根据由所述动作决定单元决定的动作来开始语音识别。
- [0193] (14) 根据(1)至(13)中任一项所述的信息处理装置,其中,在距所述用户的距离为第一距离或更小距离的情况下,所述动作决定单元将请求所述用户的响应的动作决定为对所述用户的动作。
- [0194] (15) 根据(14)所述的信息处理装置,其中,在距所述用户的距离大于所述第一距离的情况下,所述动作决定单元将不需要所述用户的响应的动作决定为对所述用户的动作。
- [0195] (16) 根据(15)所述的信息处理装置,其中,在距所述用户的距离为第二距离或更小距离的情况下,所述动作决定单元将不请求所述用户的响应的显示决定为对所述用户的动作。
- [0196] (17) 根据(16)所述的信息处理装置,其中,在距所述用户的距离大于所述第二距离和第三距离或更小距离的情况下,所述动作决定单元将不请求所述用户的响应的讲话决定为对所述用户的动作。
- [0197] (18) 一种信息处理方法,包括:
- [0198] 由处理器基于距用户的距离来决定对所述用户的动作;以及
- [0199] 输出所述动作。
- [0200] (19) 一种使计算机实现以下功能的程序:

- [0201] 基于距用户的距离来决定对所述用户的动作的功能;以及
- [0202] 输出所述动作的功能。
- [0203] 附图标记列表
- [0204] 1 信息处理装置
- [0205] 10 控制单元
- [0206] 11 通信单元
- [0207] 12 声音收集单元
- [0208] 13 扬声器
- [0209] 14 摄像装置
- [0210] 15 距离传感器
- [0211] 16 投影单元
- [0212] 17 存储单元
- [0213] 18 发光单元
- [0214] 101 用户识别单元
- [0215] 102 讲话预测单元
- [0216] 103 用户管理单元
- [0217] 104 动作决定单元
- [0218] 105 语音识别单元
- [0219] 106 语义分析单元
- [0220] 107 输出控制单元

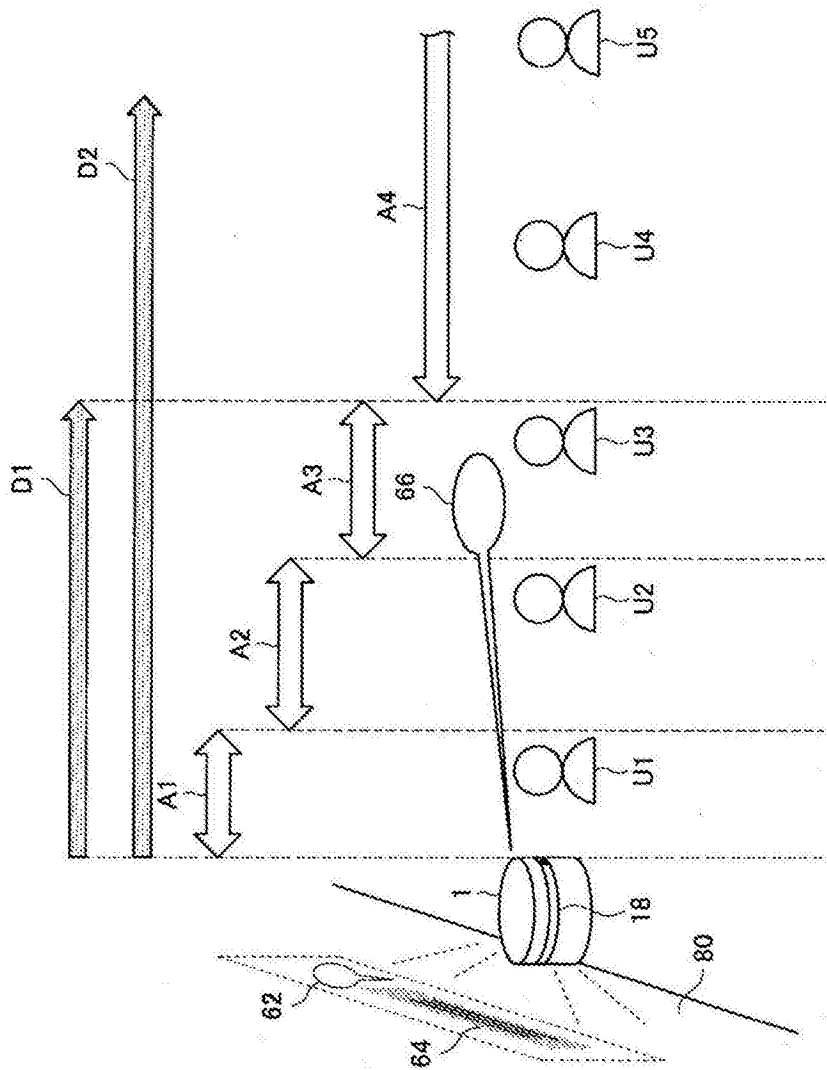


图1

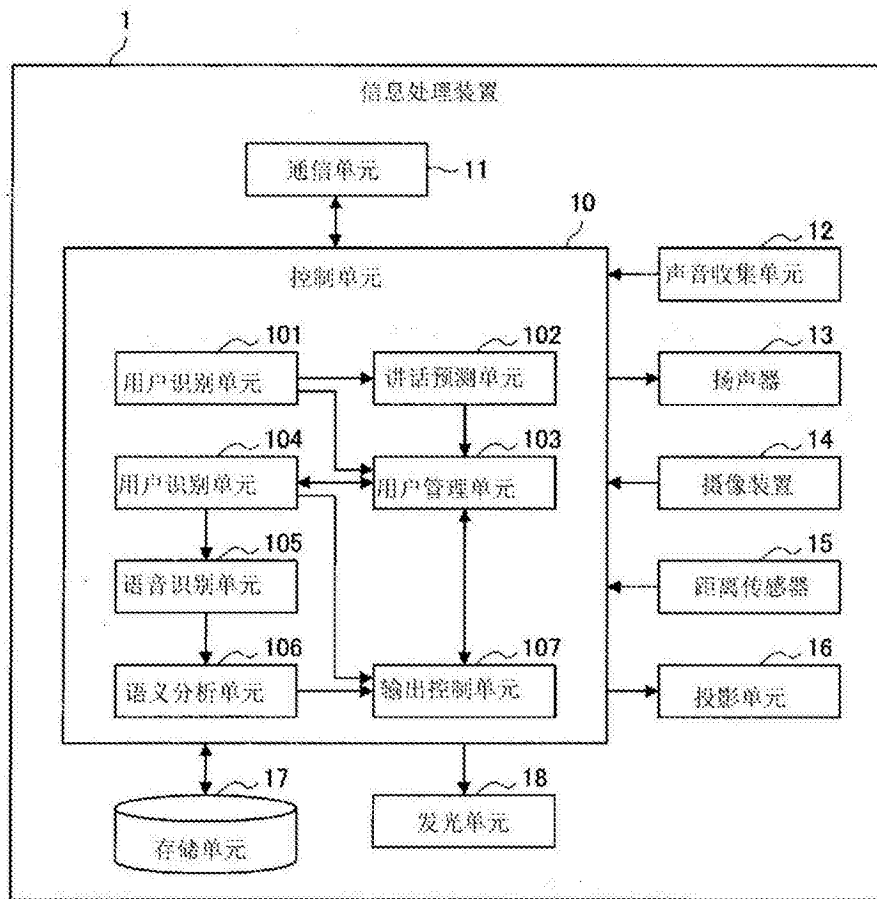


图2

| 关注分数 | 用户的状态              | 距用户的距离 | 状态持续时间 |
|------|--------------------|--------|--------|
| 5    | 直到刚刚才有交互           | -      | -      |
| 4    | 靠近主体               | 1m内    | -      |
| 3    | 靠近主体               | 2m内    | -      |
| 2    | 刚开始之后的is_watching  | -      | -      |
| 1    | is_watching持续      | -      | 少于5秒   |
| 0    | is_watching持续      | -      | 少于10秒  |
| -1   | is_watching持续      | -      | 少于15秒  |
| -2   | is_watching持续      | -      | 15秒或更多 |
| -3   | 刚开始之后的not_watching | -      | -      |
| -4   | 远离主体               | -      | -      |
| -5   | not_watching持续     | -      | 少于5秒   |
| -6   | not_watching持续     | -      | 少于10秒  |
| -7   | not_watching持续     | -      | 少于15秒  |
| -8   | not_watching持续     | -      | 15秒或更多 |

图3

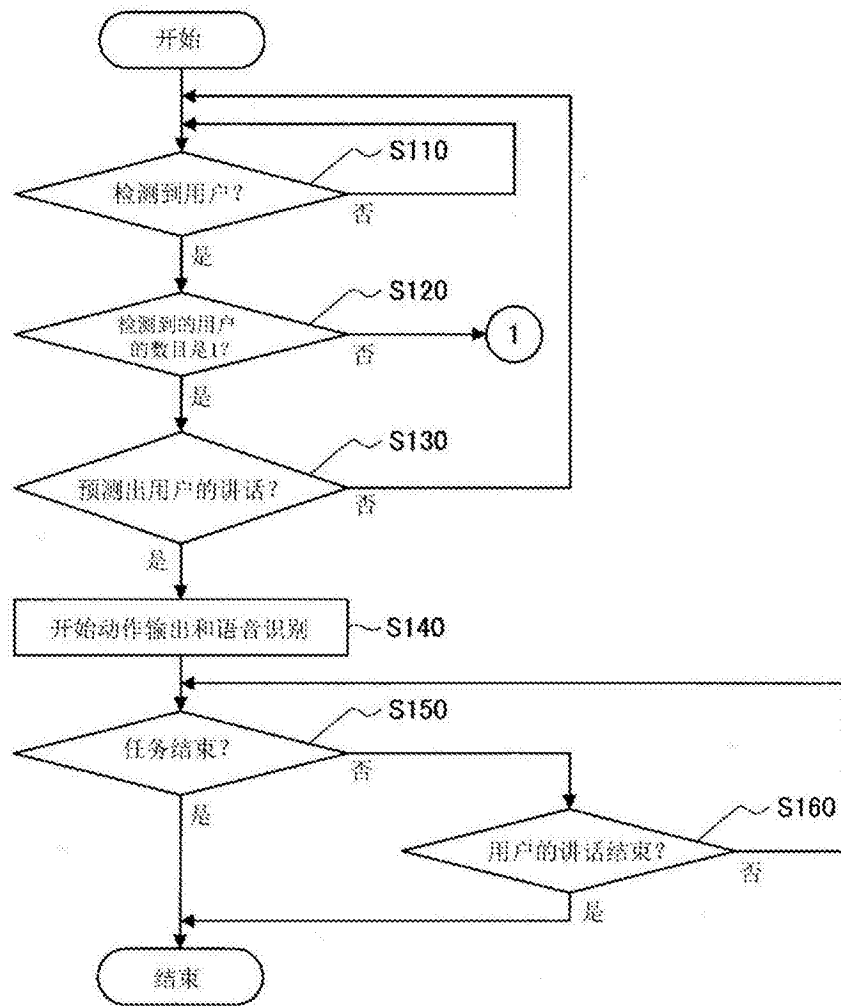


图4

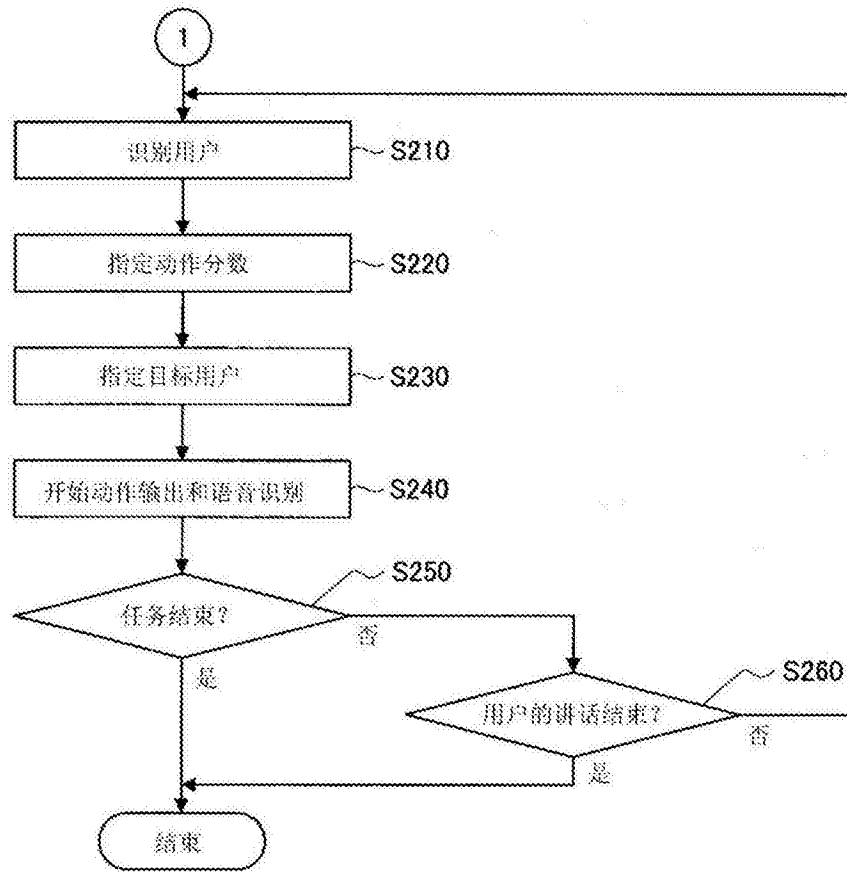


图5

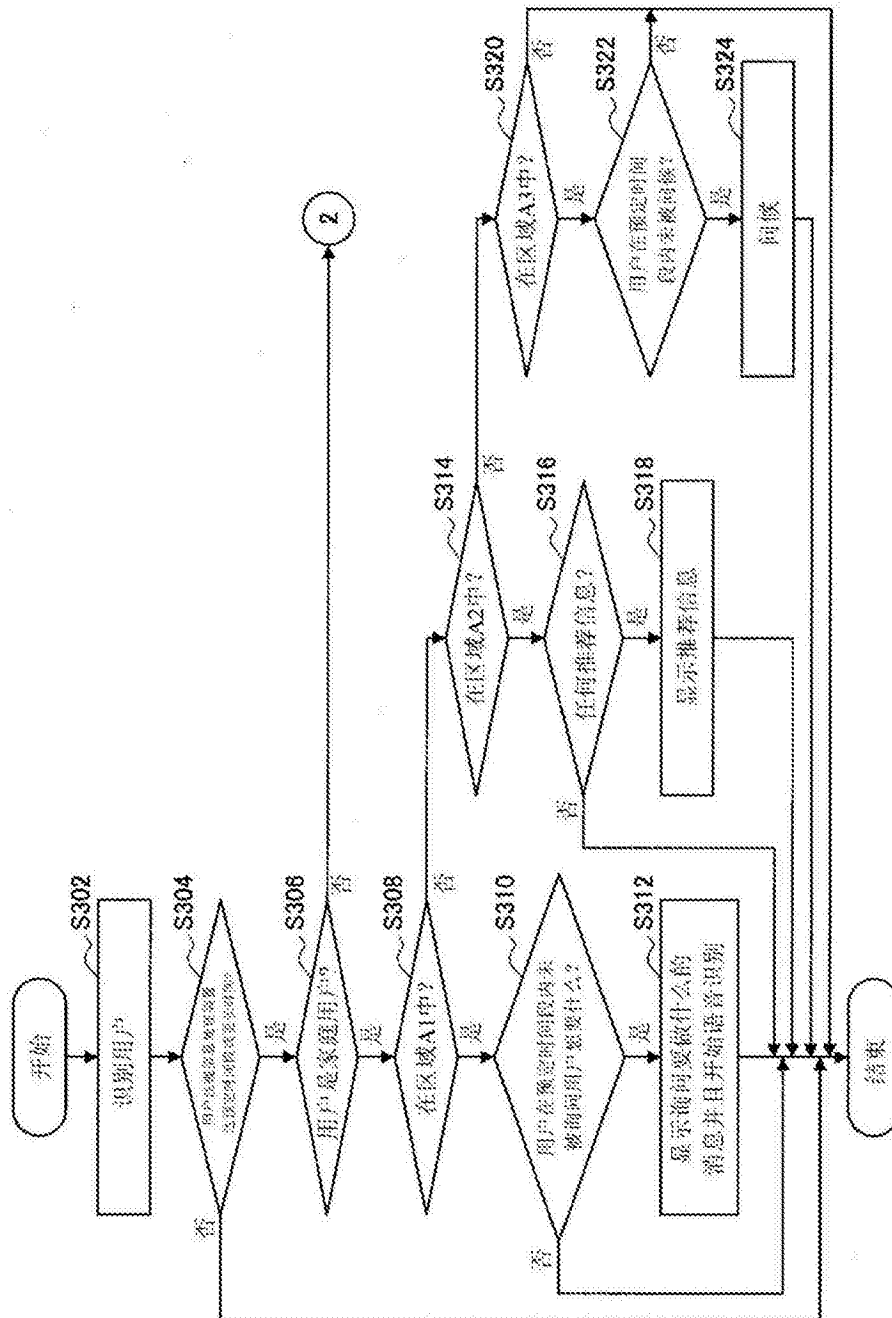


图6

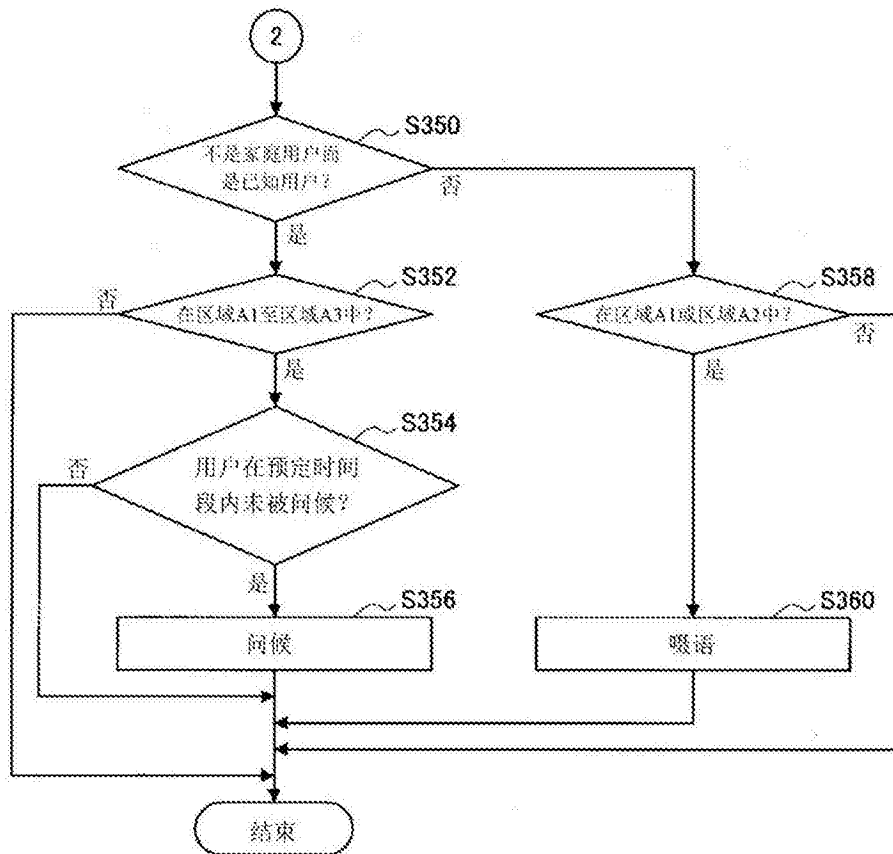


图7

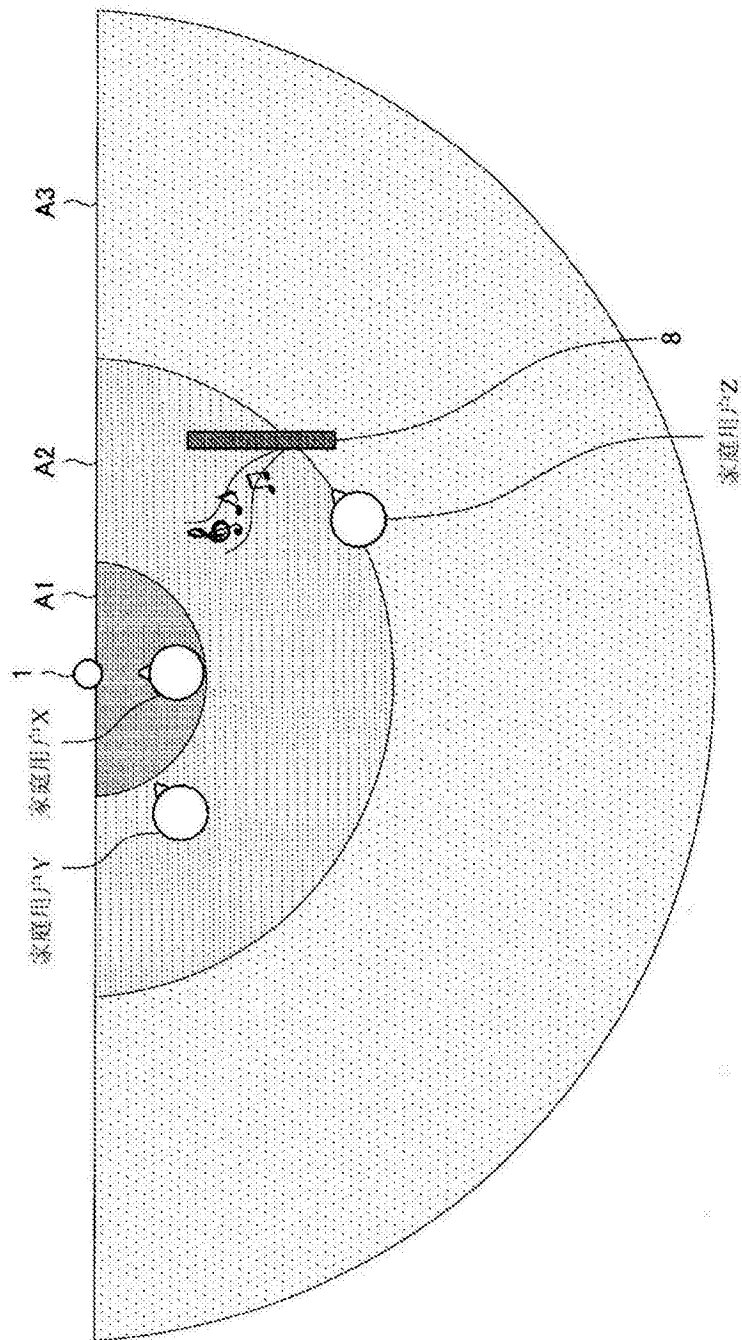


图8

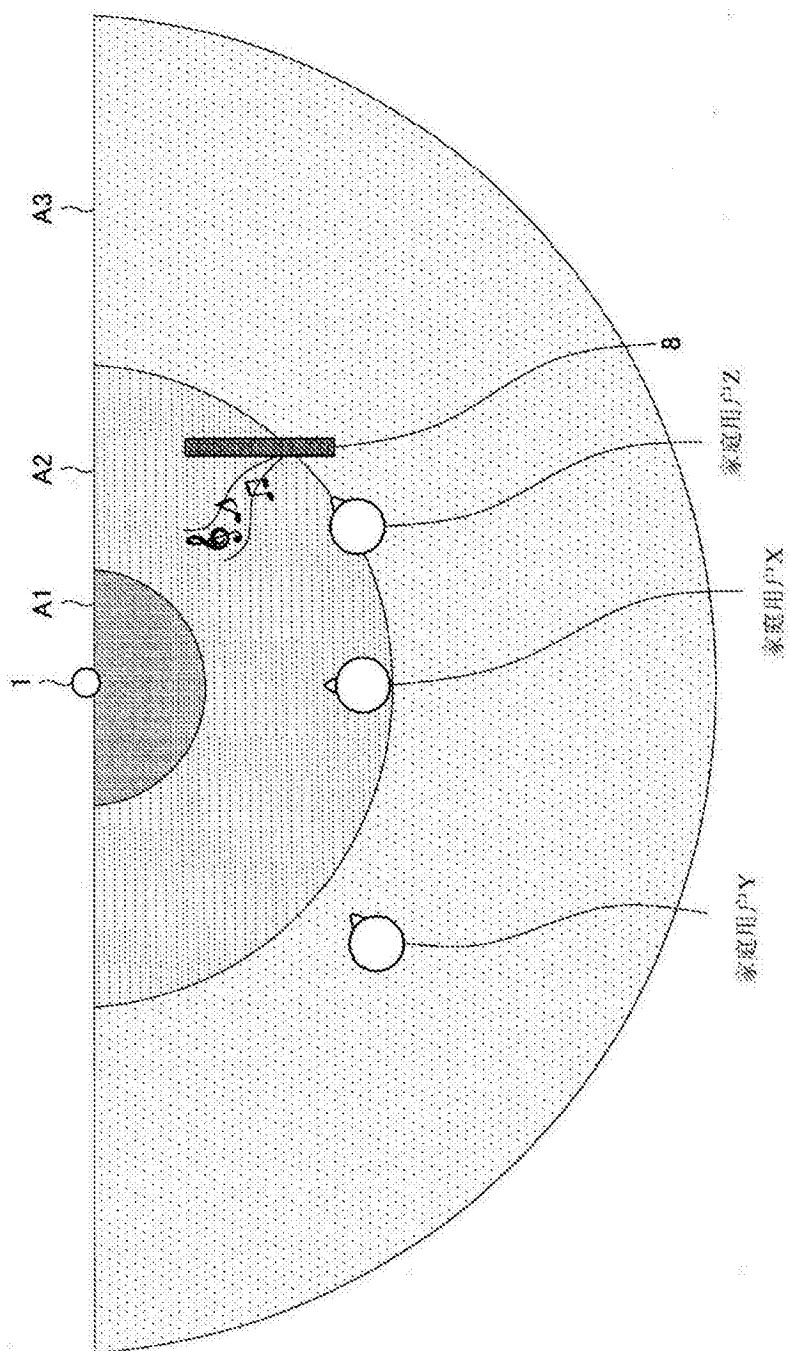


图9

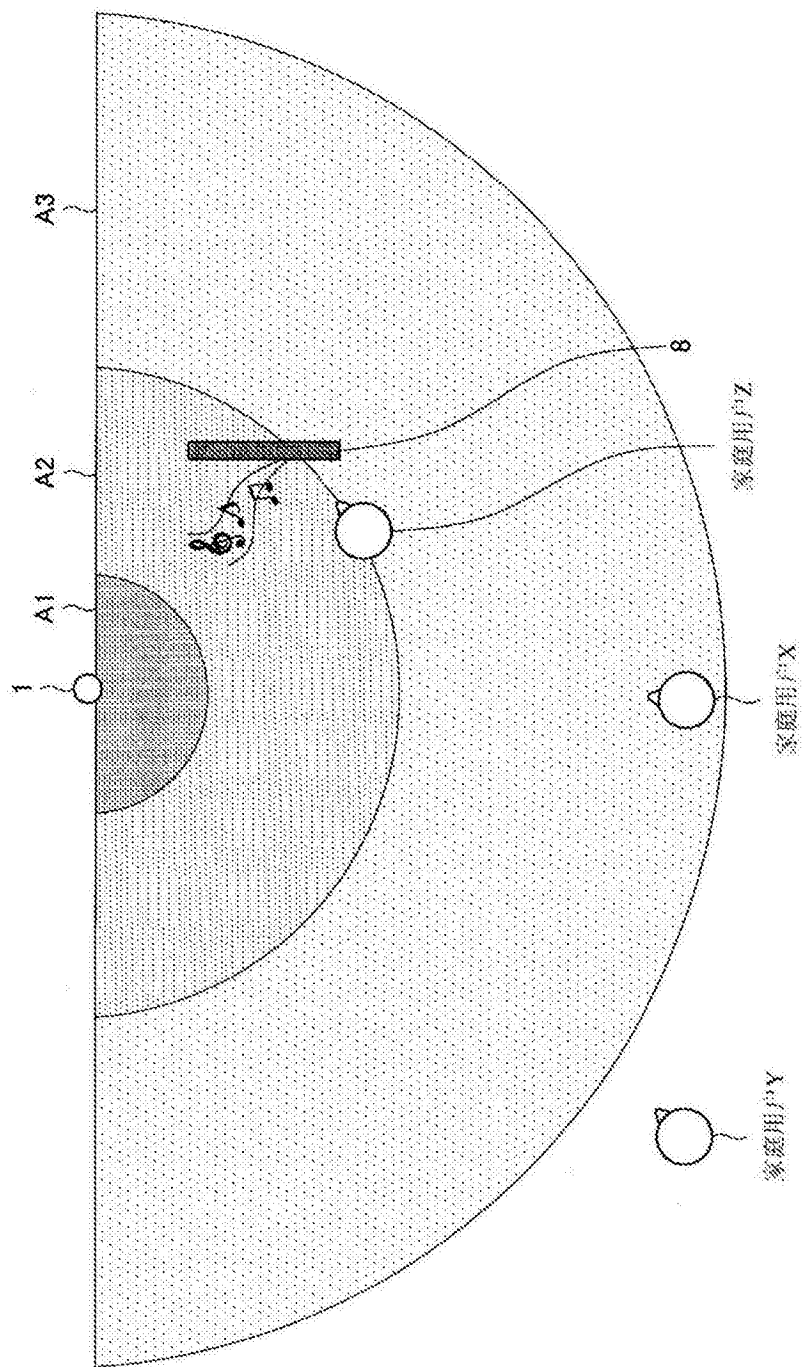


图10

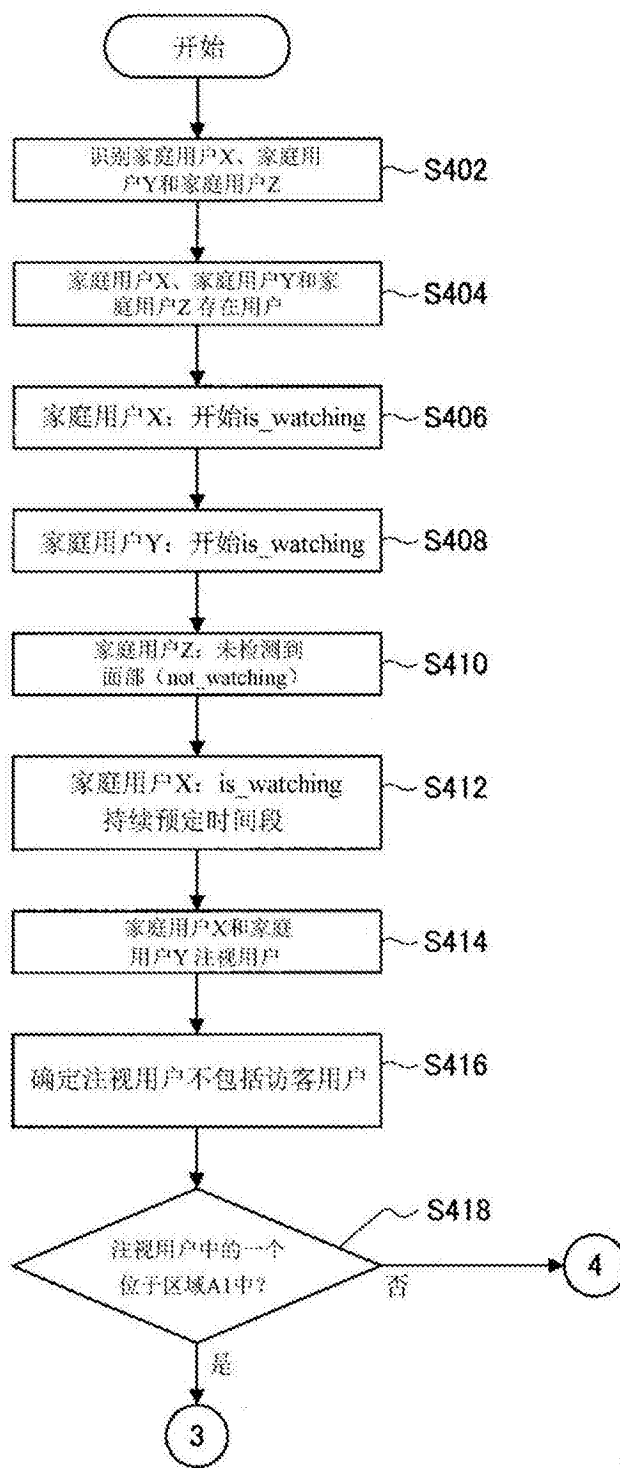


图11

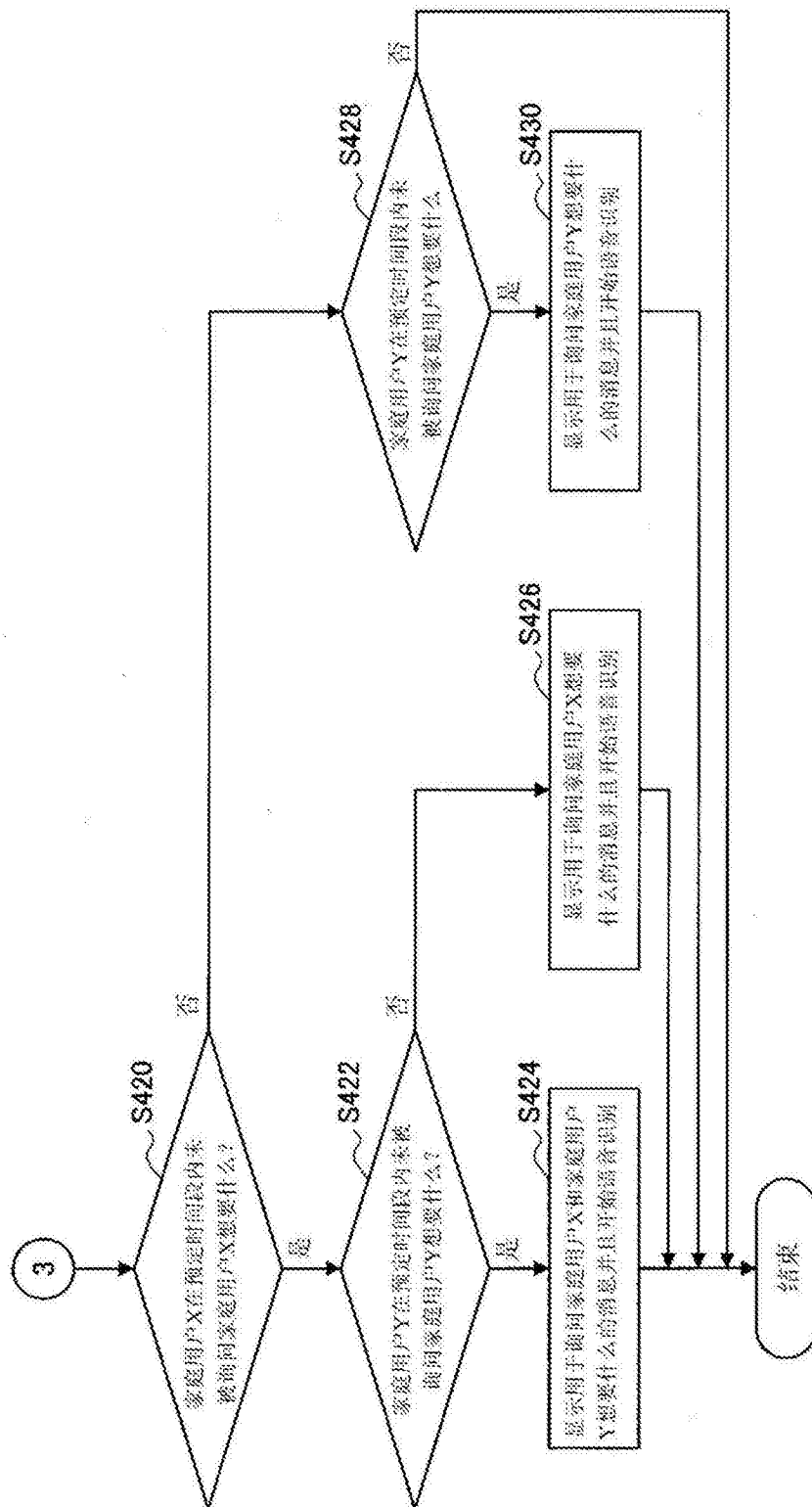


图12

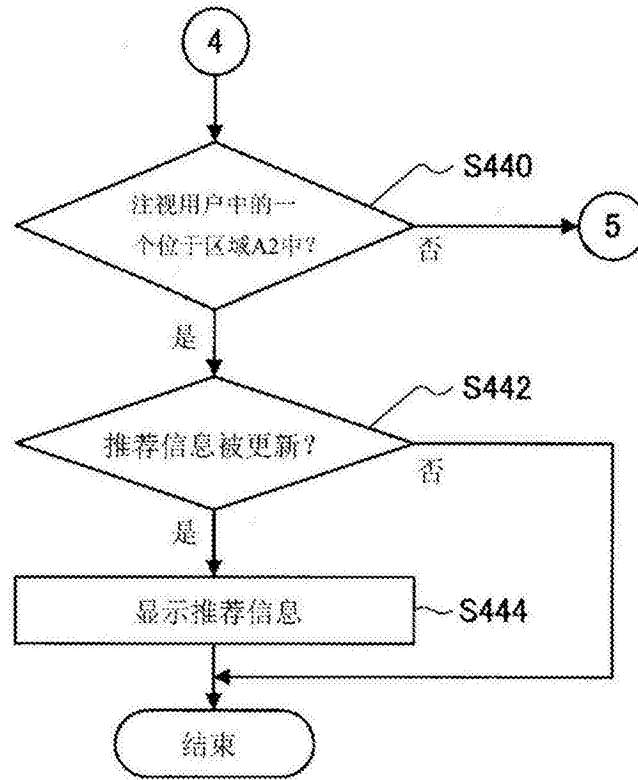


图13

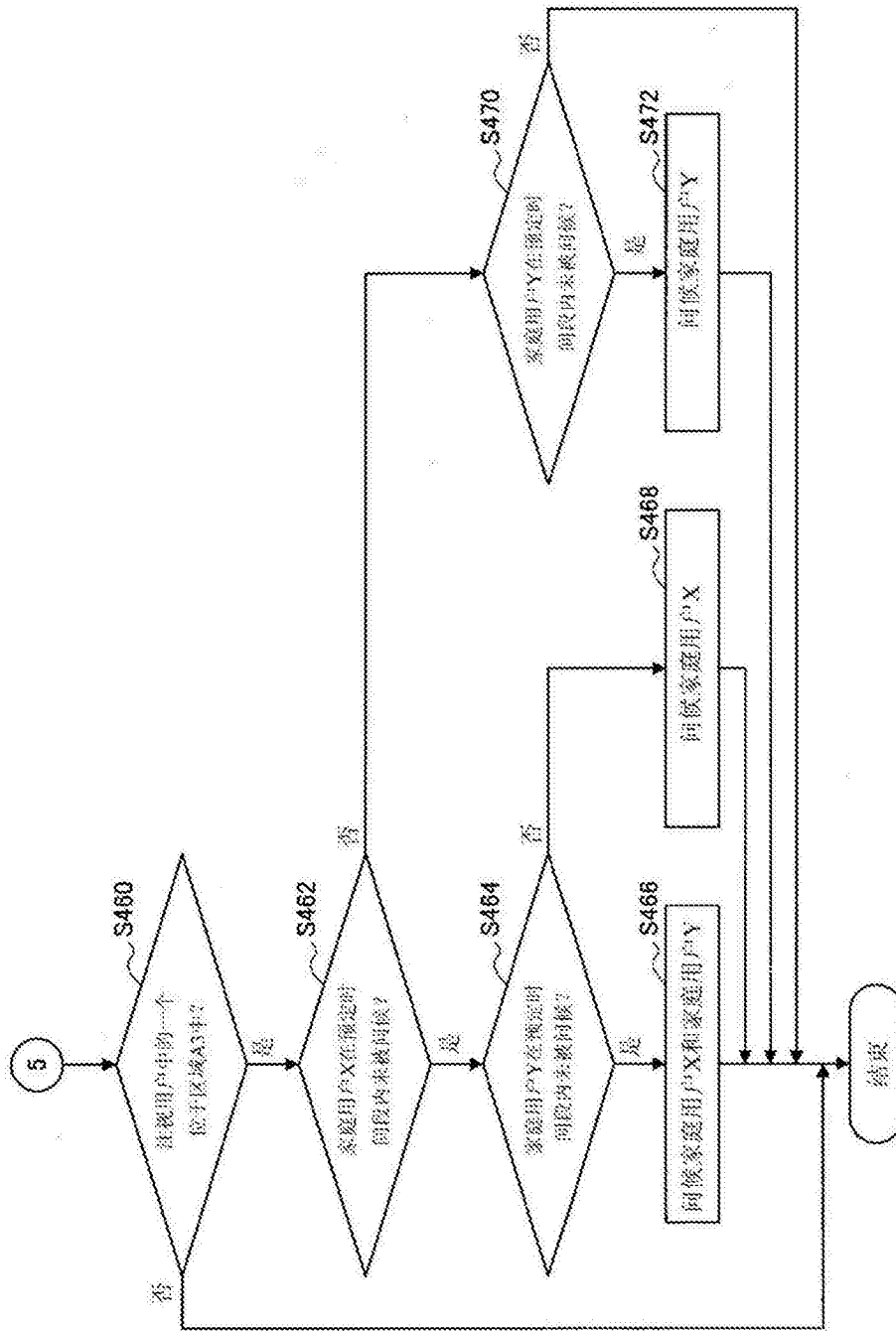


图14

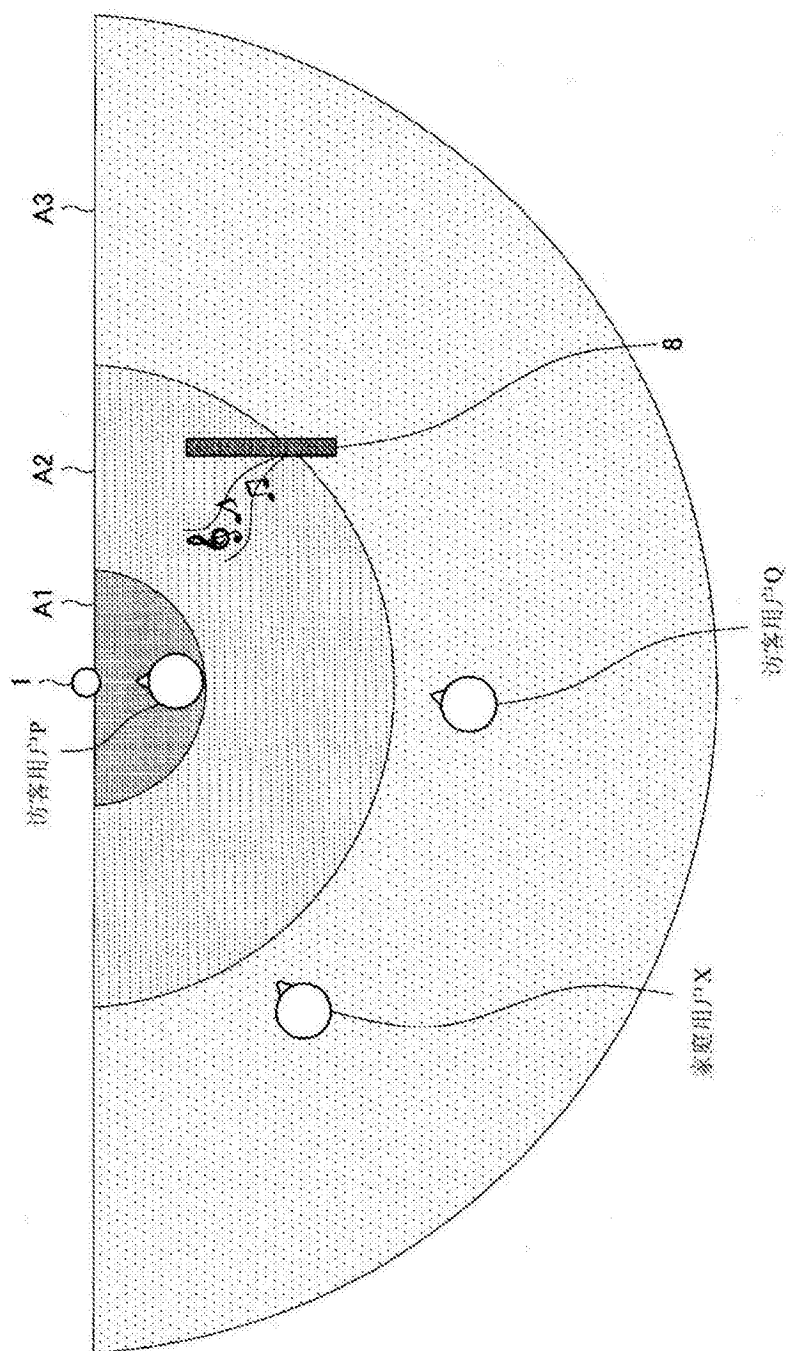


图15

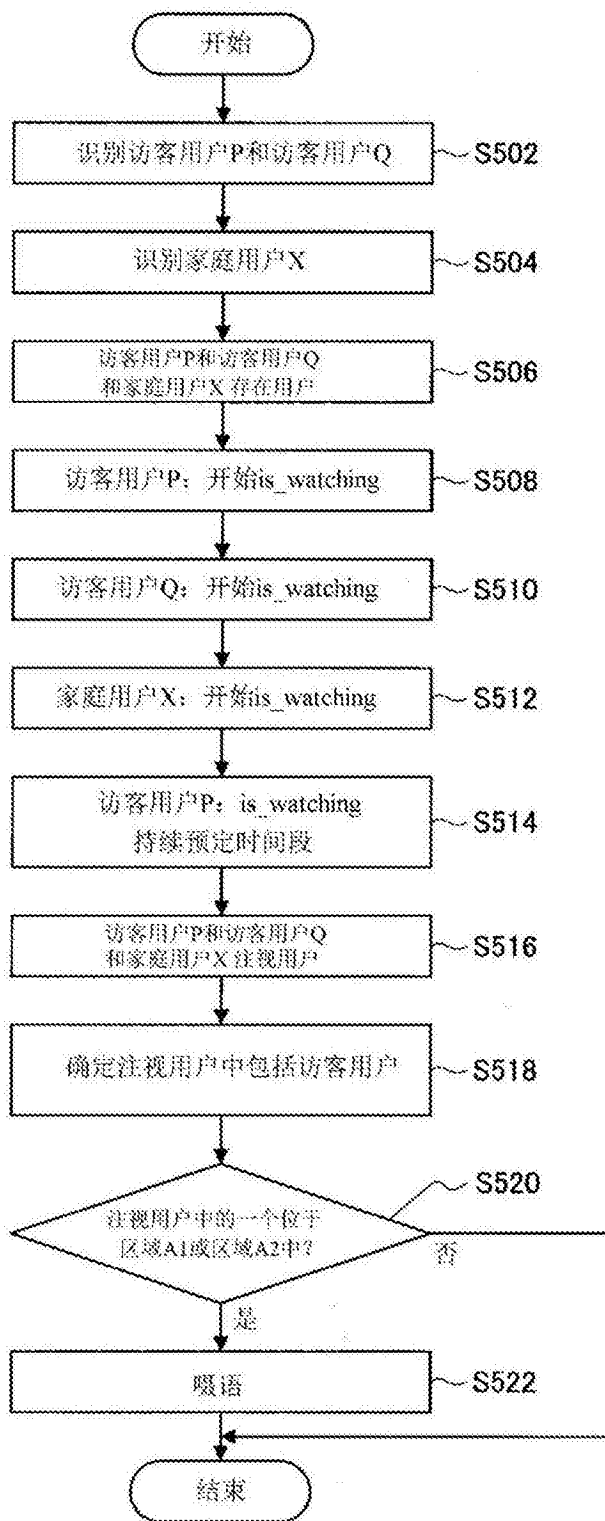


图16

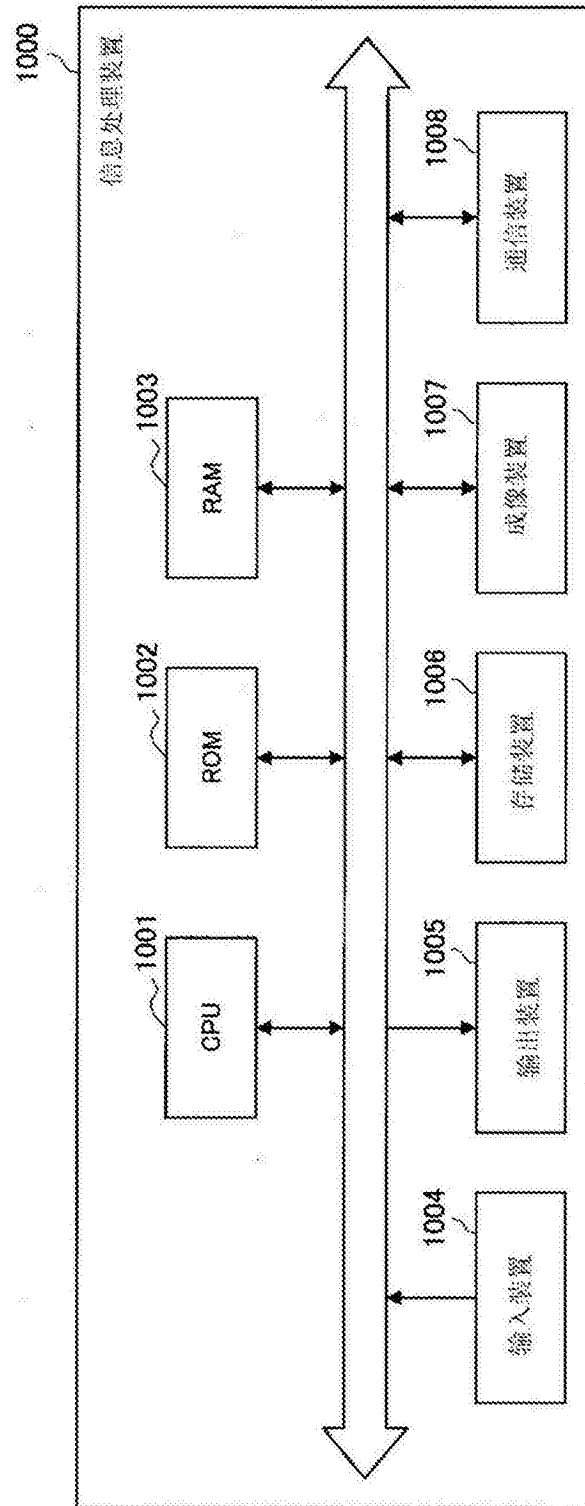


图17