



(12) 发明专利申请

(10) 申请公布号 CN 105027542 A

(43) 申请公布日 2015. 11. 04

(21) 申请号 201480010701. 1

(51) Int. Cl.

(22) 申请日 2014. 03. 07

H04M 11/00(2006. 01)

(30) 优先权数据

A63H 11/00(2006. 01)

2013-059420 2013. 03. 22 JP

G06F 3/16(2006. 01)

(85) PCT国际申请进入国家阶段日

G10L 15/10(2006. 01)

2015. 08. 26

H04M 1/00(2006. 01)

(86) PCT国际申请的申请数据

PCT/JP2014/001299 2014. 03. 07

(87) PCT国际申请的公布数据

W02014/147989 JA 2014. 09. 25

(71) 申请人 丰田自动车株式会社

地址 日本爱知县

(72) 发明人 林宏直 岩田拓也

(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 李洋 舒艳君

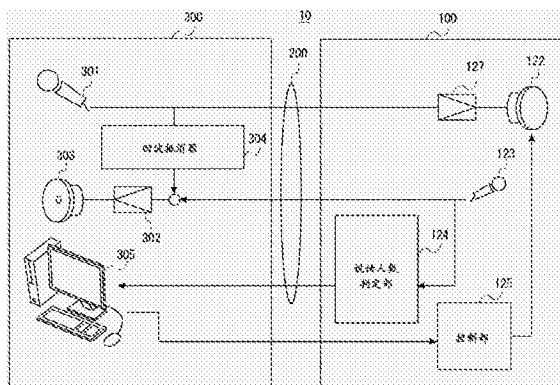
权利要求书2页 说明书16页 附图14页

(54) 发明名称

通讯系统以及机器人

(57) 摘要

本发明涉及通讯系统以及机器人, 通讯系统具备第一麦克风、第一扬声器、第二扬声器、第二麦克风、说话人数判定部以及控制部。第一麦克风获取周围的声音作为声音信号。第一扬声器设置于第一麦克风的附近, 并输出声音。第二扬声器将从第一麦克风获取到的声音信号作为声音输出。第二麦克风获取周围的声音作为声音信号, 并对第一扬声器输出。说话人数判定部根据第一麦克风获取到的声音信号, 判定第一麦克风周围的说话人的人数。控制部基于说话人数判定部的判定, 来控制第一扬声器的音量。



1. 一种通讯系统,其中,具备:
  - 第一麦克风,其获取周围的声音作为声音信号;
  - 第一扬声器,其设置于所述第一麦克风的附近,输出声音;
  - 第二扬声器,其从所述第一麦克风获取所述声音信号,输出所述声音信号作为声音;
  - 第二麦克风,其获取周围的声音作为声音信号,对所述第一扬声器输出获取到的所述声音信号;说话人数判定部,其根据所述第一麦克风获取到的声音信号,来判定说话人数,所述说话人数表示所述第一麦克风的周围的说话人的人数;以及  
控制部,其基于所述说话人数判定部的判定,来控制所述第一扬声器的音量。
2. 根据权利要求 1 所述的通讯系统,其中,
  - 所述控制部进行控制,以使得在所述说话人数小于阈值的情况下,与所述说话人数是阈值以上的情况相比较,减小所述第一扬声器的音量。
3. 根据权利要求 2 所述的通讯系统,其中,
  - 所述控制部进行控制,以使得在所述说话人数小于阈值的情况下,在比规定时间少的时间内,即便当所述说话人数变为阈值以上时,也保持减小所述第一扬声器的音量的状态。
4. 根据权利要求 2 所述的通讯系统,其中,
  - 所述控制部进行控制,以使得在所述说话人数是阈值以上的情况下,在比规定时间少的时间内,即便当所述说话人数变为小于阈值时,也保持增大所述第一扬声器的音量的状态。
5. 根据权利要求 1 所述的通讯系统,其中,
  - 所述控制部进行控制,以使得在所述说话人数在特定时间内减少得比规定数量大的情况下,减小所述第一扬声器的音量。
6. 根据权利要求 1 所述的通讯系统,其中,
  - 所述控制部进行控制,以使得在所述说话人数是阈值以下并且正在减少的情况下,减小所述第一扬声器的音量。
7. 根据权利要求 1 所述的通讯系统,其中,
  - 所述通讯系统还具备比较部,所述比较部输出对所述第一麦克风获取到的声音与所述第二麦克风获取到的声音的大小进行比较的比较结果,
  - 所述控制部基于所述比较部的比较结果来控制所述第一扬声器的音量。
8. 根据权利要求 1 所述的通讯系统,其中,
  - 所述控制部基于一定时间内的所述说话人数来控制所述第一扬声器的音量。
9. 根据权利要求 8 所述的通讯系统,其中,
  - 所述说话人数判定部对所述一定时间内说话的人数进行计数,作为一定时间内的所述说话人数。
10. 一种机器人,与人之间执行使用了声音的通讯,其中,具备:
  - 扬声器,其输出声音;
  - 麦克风,其获取所述机器人的周围的声音作为声音信号;
  - 说话人数判定部,其根据所述麦克风获取到的声音信号,来判定所述机器人的周围的说话人的人数;以及

控制部,其基于所述说话人数判定部的判定,来控制所述扬声器的音量。

11. 一种音量控制的方法,是通讯系统中的音量控制的方法,所述通讯系统设置有:

第一麦克风,其获取周围的声音作为声音信号;

第一扬声器,其设置于所述第一麦克风的附近,输出声音;

第二扬声器,其从所述第一麦克风获取所述声音信号,输出所述声音信号作为声音;以

及

第二麦克风,其获取周围的声音作为声音信号,对所述第一扬声器输出获取到的所述声音信号,

所述音量控制的方法包括:

根据所述第一麦克风获取到的声音信号,来判定所述第一麦克风的周围的说话人的人数的步骤;以及

基于所述判定,来控制所述第一扬声器的音量的步骤。

12. 一种控制程序,是机器人的控制程序,所述机器人具备输出声音的扬声器以及获取周围的声音作为声音信号的麦克风,其中,

使机器人执行:

根据所述麦克风获取到的声音信号,来判定所述机器人的周围的说话人的人数的步骤;以及

基于所述判定,来控制所述扬声器的音量的步骤。

## 通讯系统以及机器人

### 技术领域

[0001] 本发明涉及通讯系统、机器人、音量控制的方法以及机器人的控制程序。

### 背景技术

[0002] 近年来,进行了将机器人放置到有人的场所(例如会场),执行机器人与人的通讯这样的处理。例如,在专利文献1中,公开有检测在机器人周围的人的人数的通讯机器人的技术。该通讯机器人根据检测到的人的人数,使用扬声器对在机器人周围的人进行不同的提醒。

[0003] 专利文献1:日本特开2007-118129号公报

[0004] 在如上述那样执行机器人与人的通讯的情况下,存在机器人(在机器人是远程操作型的机器人的情况下对机器人进行远程操作的人)无法充分地识别机器人周边的人的交谈状态这样的问题点。即,机器人或者操作该机器人的人无法识别机器人周边的人现在是正在聊天还是正在安静地听某人演讲等。因此,可能产生例如在机器人周边的人正在安静地听某人演讲的时候,机器人执行发出大声的动作的情况。这样,机器人的举动可能扰乱周围的气氛。

### 发明内容

[0005] 本发明是为了解决这样的问题点而完成的,其目的在于,提供能够根据周围的环境来变更通讯音量的通讯系统、机器人、音量控制的方法以及机器人的控制程序。

[0006] 本发明的通讯系统具备第一麦克风、第一扬声器、第二扬声器、第二麦克风、说话人数判定部以及控制部。第一麦克风获取周围的声音作为声音信号。第一扬声器设置于上述第一麦克风的附近,输出声音。第二扬声器从上述第一麦克风获取上述声音信号,输出该声音信号作为声音。第二麦克风获取周围的声音作为声音信号,对上述第一扬声器输出获取到的该声音信号。说话人数判定部根据上述第一麦克风获取到的声音信号,来判定说话人数,上述说话人数表示上述第一麦克风的周围的说话人的人数。控制部基于上述说话人数判定部的判定,来控制上述第一扬声器的音量。通过这样的构成,通讯系统能够根据第一麦克风的周围的说话人的人数,来控制第一扬声器的音量。因此,通讯系统能够根据第一麦克风的周围的环境,来变更通讯的音量。

[0007] 也可以是,上述控制部进行控制,以使得在上述说话人数小于阈值的情况下,与上述说话人数是阈值以上的情况相比较,减小上述第一扬声器的音量。通过这样的构成,在第一麦克风的周围的说话人的人数小于阈值时,从位于其附近的第一扬声器输出的声音的音量变小。因此,通讯系统能够防止阻碍在第一扬声器的周围进行的演讲。

[0008] 也可以是,上述控制部进行控制,以使得在上述说话人数小于阈值的情况下,在比规定的时间少的时间段内,即便当上述说话人数为阈值以上时,也保持减小上述第一扬声器的音量的状态。通过这样的构成,控制部在第一麦克风的周围正在进行演讲的情况下,即便当较短的期间发出了多个人的声音时,也能够保持减小第一扬声器的音量的状态。由此,通

讯系统能够防止阻碍在第一扬声器的周围进行的演讲。

[0009] 也可以是,上述控制部进行控制,以使得在上述说话人数是阈值以上的情况下,在比规定的时间少的时间段内,即便当上述说话人数小于阈值时,也保持增大上述第一扬声器的音量的状态。通过这样的构成,控制部在第一麦克风的周围正在进行多个人的交谈的情况下,即便当较短的期间为安静的状态时,也能够将第一扬声器的音量保持原样。由此,通讯系统能够防止由于减小了第一扬声器的音量而导致阻碍使用第二扬声器以及第二麦克风说话的人的通讯。

[0010] 也可以是,上述控制部进行控制,以使得在上述说话人数在特定时间内减少得比规定数量大的情况下,减小上述第一扬声器的音量。通过这样的构成,控制部能够检测将第一麦克风的周围的状态从多个人正在聊天的状态变化到听众安静地等待演讲的状态这一情况,来减小第一扬声器的音量。由此,通讯系统能够防止阻碍在第一扬声器的周围进行的演讲。

[0011] 也可以是,上述控制部进行控制,以使得在上述说话人数是阈值以下并且正在减少的情况下,减小上述第一扬声器的音量。通过这样的构成,控制部能够检测第一麦克风的周围的状态为听众正在安静地等待演讲的状态这一情况,来减小第一扬声器的音量。由此,通讯系统能够防止阻碍在第一扬声器的周围进行的演讲。

[0012] 也可以是,上述通讯系统还具备比较部,上述比较部输出对上述第一麦克风获取到的声音与上述第二麦克风获取到的声音的大小进行比较的比较结果,上述控制部基于上述比较部的比较结果来控制上述第一扬声器的音量。由此,通讯系统能够使第一扬声器的音量与第一扬声器的周围的声音的大小一致。

[0013] 也可以是,上述控制部基于一定时间内的上述说话人数来控制上述第一扬声器的音量。由此,即使说话人数仅在较短的时间变化的情况下,控制部也能够不管其变化如何,均进行麦克周围的准确状态的判定。通讯系统能够根据第一麦克风的周围的环境更高精度地变更通讯的音量。

[0014] 也可以是,上述说话人数判定部对在上述一定时间内进行说话的人数进行计数,作为一定时间内的上述说话人数。由此,说话人数判定部能够在不同的人彼此在不同的时刻进行交谈的情况下,准确地把握说话人的合计人数。因此,控制部能够进行第一麦克风的周围的准确的状态的判定。由此,通讯系统能够根据第一麦克风的周围的环境更高精度地变更通讯的音量。

[0015] 本发明的机器人是与人之间执行使用了声音的通讯的机器人,具备扬声器、麦克风、说话人数判定部以及控制部。扬声器输出声音。麦克风获取上述机器人的周围的声音作为声音信号。说话人数判定部根据上述麦克风获取到的声音信号,来判定上述机器人的周围的说话人的人数。控制部基于上述说话人数判定部的判定结果,来控制上述扬声器的音量。通过这样的构成,机器人能够根据周围的说话人的人数控制扬声器的音量。因此,机器人能够根据周围的环境来变更通讯的音量。

[0016] 本发明的音量控制的方法是通讯系统中的音量控制的方法,上述通讯系统设置有:第一麦克风,其获取周围的声音作为声音信号;第一扬声器,其设置于上述第一麦克风的附近,输出声音;第二扬声器,其从上述第一麦克风获取上述声音信号,输出上述声音信号作为声音;以及第二麦克风,其获取周围的声音作为声音信号,对上述第一扬声器输出获

取到的上述声音信号。该音量控制的方法至少包括以下的步骤 (a) ~ (b)。

[0017] (a) 根据上述第一麦克风获取到的声音信号,来判定上述第一麦克风的周围的说话人的人数;以及

[0018] (b) 基于上述判定,来控制上述第一扬声器的音量。

[0019] 通过这样的构成,通讯系统能够根据第一麦克风的周围的说话人的人数,来控制第一扬声器的音量。因此,通讯系统能够根据第一麦克风的周围的环境来变更通讯的音量。

[0020] 本发明的机器人的控制程序是具备输出声音的扬声器、以及获取周围的声音作为声音信号的麦克风的机器人的控制程序。该机器人的控制程序使机器人至少执行以下的步骤 (a) ~ (b)。

[0021] (a) 根据上述麦克风获取到的声音信号,来判定上述机器人的周围的说话人的人数;以及

[0022] (b) 基于上述判定,来控制上述扬声器的音量。

[0023] 通过这样的构成,机器人能够根据周围的说话人的人数来控制扬声器的音量。因此,机器人能够根据周围的环境,来变更通讯的音量。

[0024] 通过本发明,能够提供能够根据周围的环境来变更通讯的音量的通讯系统、机器人、音量控制的方法以及机器人的控制程序。

## 附图说明

[0025] 图 1 是表示实施方式 1 的机器人的外观的例子的图。

[0026] 图 2 是表示实施方式 1 的机器人的头部的构成例的图。

[0027] 图 3 是表示实施方式 1 的机器人的内部构成的一个例子的框构成图。

[0028] 图 4 是表示具有实施方式 1 的机器人的通讯系统的一个例子的构成图。

[0029] 图 5 是表示在实施方式 1 的机器人的周边的人的说话声音的大小的一个例子的曲线图。

[0030] 图 6 是实施方式 1 的说话人数判定部判定出的说话人数的曲线图的例子。

[0031] 图 7 是表示实施方式 1 的控制终端的状态判定的处理的一个例子的流程图。

[0032] 图 8 是表示在实施方式 1 中,说话人数判定部判定出的说话人数以及移动平均说话人数的一个例子的曲线图。

[0033] 图 9 是相关技术的通讯系统的构成图。

[0034] 图 10 是表示实施方式 1 的机器人的处理的一个例子的框构成图。

[0035] 图 11 是表示实施方式 2 的机器人的内部构成的一个例子的框构成图。

[0036] 图 12 是表示具有实施方式 2 的机器人的通讯系统的一个例子的构成图。

[0037] 图 13 是表示实施方式 3 的机器人的内部构成的一个例子的框构成图。

[0038] 图 14 是表示具有实施方式 3 的机器人的通讯系统的一个例子的构成图。

## 具体实施方式

[0039] 实施方式 1

[0040] 以下,参照附图对本发明的实施方式 1 进行说明。图 1 是表示本实施方式 1 的机器人 100 的外观的例子的图。机器人 100 是模拟人的形状而制造出的远程操作机器人,具

备头部 101、躯干部 102、右臂部 103、左臂部 104、右腿部 105 以及左腿部 106。机器人 100 能通过远程操作使各部移动。

[0041] 图 2 是表示头部 101 的构成例的图。头部 101 具有右眼 111、左眼 112、口 113、右耳 114 以及左耳 115。右眼 111、左眼 112 分别设置于与人脸的右眼以及左眼对应的位置，右眼 111 作为拍摄装置发挥作用。机器人 100 能够通过该拍摄装置拍摄周围的环境，向进行远程操作的人（以下，记载为远程操作者）发送该拍摄数据。远程操作者通过将该拍摄数据作为影像来进行观察，能够视觉识别机器人 100 周围的环境。

[0042] 口 113 设置于与人脸的口对应的位置。在口 113 设置有扬声器。机器人 100 通过设置于口 113 的扬声器将从远程操作侧发送来的声音信号作为声音再生，能够使机器人 100 的周围的人听见。

[0043] 右耳 114 以及左耳 115 设置于与人脸的右耳以及左耳对应的位置。在右耳 114 设置有麦克风（以下，简略记载为麦克），机器人 100 能够通过该麦克获取周围的声响（例如机器人 100 周围的人的交谈），并将其作为电信号的声音数据发送给远程操作者。远程操作者通过利用扬声器将该声音数据作为声音来加以听取，从而能够确认机器人 100 的周围的声响。此外，所谓机器人 100 的周围的人例如是在机器人 100 的半径数米内的人。

[0044] 图 3 是表示机器人 100 的内部构成的一个例子的框构成图。机器人 100 具备拍摄装置 121、扬声器 122、麦克 123、说话人数判定部 124、控制部 125 以及可动部 126。拍摄装置 121 是设置于右眼 111 的上述的拍摄装置。扬声器 122 是设置于口 113 的上述的扬声器。麦克 123 是设置于右耳 114 的上述的麦克。

[0045] 说话人数判定部 124 基于从麦克 123 获取到的周围的声响，来判定在机器人 100 的周围说话的说话人有几个人。对该处理的详细后述。

[0046] 控制部 125 根据来自远程操作侧的控制信号来控制机器人 100 的各部，例如是由 CPU(Central Processing Unit:中央处理器)、存储器、其他电路构成的 IC(Integrated Circuit:集成电路)。控制部 125 例如能够根据控制信号来调整扬声器 122 的音量。可动部 126 是能够根据来自控制部 125 的控制来使头部 101 ~ 左腿部 106 移动的单元，例如是致动器。

[0047] 图 4 是表示具有具备以上所示的构成的机器人 100 的通讯系统的一个例子的构成图。在图 4 中，通讯系统 10 具备机器人 100、网络 200 以及远程操作侧 300，机器人 100 经由网络 200 与远程操作侧 300 连接。

[0048] 在图 4 中，作为机器人 100 的构成要素，除了扬声器 122(第一扬声器)、麦克 123(第一麦克风)、说话人数判定部 124、控制部 125 以外，还记载有放大器 127。放大器 127 对来自远程操作侧 300 的声音信号进行放大后供给至扬声器 122。此外，虽然机器人 100 具有图 1 ~ 图 3 所示的构成要素，但在图 4 中省略了它们的记载。

[0049] 网络 200 是有线或者无线的网络。远程操作侧 300 是机器人 100 的远程操作者执行该远程操作的地方，具备麦克 301(第二麦克风)、放大器 302、扬声器 303(第二扬声器)、回波抵消器 304 以及控制终端 305。

[0050] 以下，对于远程操作侧 300 的各部进行说明。麦克 301 经由网络 200 与扬声器 122 连接。通过远程操作者对麦克 301 输入自身的声音，从而该声音被转换为电信号，经由网络 200、放大器 127 被送出至扬声器 122。按照上述，扬声器 122 将该声音信号转换为声音并输

出。这样一来,远程操作者能够使自身的声音从距离较远的机器人 100 输出。

[0051] 放大器 302 经由网络 200 获取麦克 123 输出的声音信号,进行放大并输出至扬声器 303。扬声器 303 将该声音信号转换为声音并输出。这样一来,远程操作者能够听到机器人 100 周围的声响。

[0052] 回波抵消器 304 是消除由扬声器 122、麦克 123、麦克 301、扬声器 303 产生的回波的单元。回波抵消器 304 被设置在连接麦克 301 和扬声器 122 的连接线与连接放大器 302 和麦克 123 的连接线之间。

[0053] 通过远程操作者从麦克 301 输入自身的声音,从而该声音被从扬声器 122 输出。此时,由于该声音从麦克 123 输入,从而声音被从扬声器 303 输出。即,会产生回波。为了防止该回波产生,回波抵消器 304 从连接线检测输出至麦克 301 的远程操作者的声音信号,将使该声音信号取消的取消信号输出至连接放大器 302 和麦克 123 的连接线。由此,回波抵消器 304 能够防止回波产生。

[0054] 控制终端 305(控制部)获取说话人数判定部 124 输出的判定结果,并基于此执行机器人 100 的控制。例如控制终端 305 能够自动地控制机器人 100 的扬声器 122 的音量。以下,对该控制的详细进行说明。

[0055] 图 5 是表示在机器人 100 的周边的人的说话声音的大小的一个例子的曲线图。在图 5 中,以在机器人 100 的周边有 A 先生~F 先生这 6 个人作为前提。在图 5 中,横轴表示时刻,纵轴表示说话声音的大小。另外,在图 5 中初始时刻是  $t_0$ 。

[0056] 在图 5 中,A 先生在时刻  $t_0 \sim t_1$  不说话,在时刻  $t_1 \sim t_2$  用较小的声音说话,在时刻  $t_2 \sim t_3$  不说话,在时刻  $t_3 \sim t_4$  用较小的声音说话,在时刻  $t_4 \sim t_5$  不说话,在时刻  $t_5 \sim t_6$  用较大的声音说话,在时刻  $t_6$  以后不说话。B 先生在时刻  $t_0 \sim t_7$  不说话,在时刻  $t_7 \sim t_8$  用中等程度的声音说话,在时刻  $t_8 \sim t_9$  不说话,在时刻  $t_9$  以后用中等程度的声音说话。C 先生在时刻  $t_0 \sim t_{10}$  用较小的声音说话,在时刻  $t_{10} \sim t_{11}$  不说话,在时刻  $t_{11} \sim t_{12}$  用较小的声音说话,在时刻  $t_{12}$  以后用较大的声音说话。D 先生在时刻  $t_0 \sim t_{13}$  不说话,在时刻  $t_{13} \sim t_{14}$  用较大的声音说话,在时刻  $t_{14}$  以后不说话。E 先生在时刻  $t_0 \sim t_{15}$  不说话,在时刻  $t_{15} \sim t_{16}$  用较小的声音说话,在时刻  $t_{16}$  以后不说话。F 先生从时刻  $t_0$  以后不说话。这样,人的说话声音的特征在于,除了较大的声音、较小的声音混在以外,若从较长的跨度来看,一个人的声音并不连续而是断续的。麦克 123 获取这样的说话声音作为声音数据。

[0057] 麦克 123 获取到的声音数据经由网络 200、放大器 302 被输入至扬声器 303。并且,麦克 123 获取到的声音数据被输入至说话人数判定部 124。

[0058] 说话人数判定部 124 根据麦克 123 获取到的声音数据判定当前正在说话的人(说话人)的数量。这里,对于人的声音,人的声带的固有振动不同。因此,即使同时 2 个人以上的声音发出的情况下,说话人数判定部 124 也通过解析声音数据,来检测 2 个以上的不同的固有振动。由此,说话人数判定部 124 判定 2 个人以上正在发出声音。这样一来,说话人数判定部 124 判定说话人数,说话人数表示麦克 123 周围的说话人的人数。

[0059] 图 6 是说话人数判定部 124 判定出的说话人数的曲线图的例子。在图 6 中,横轴表示时刻,纵轴表示特定的时刻中的说话人数。在图 6 中,在初始时刻  $t_{20}$  中,麦克 123 的周边(即,机器人 100 的周边)的说话人数是 1 人。在时刻  $t_{21}$  中,说话人数减少到 0 人。在



时刻 t22 中,说话人数返回到 1 人。在时刻 t23 中,说话人数增加到 2 人。在时刻 t24 中,说话人数为 3 人,之后立刻增加到 4 人。在时刻 t25 中,说话人数减少到 2 人。在时刻 t26 中,说话人数为 1 人。在时刻 t27 中,说话人数返回到 2 人。在时刻 t28 中,说话人数一时为 1 人,但立刻返回到 2 人。在时刻 t29 中,说话人数为 1 人。在时刻 t30 中,说话人数返回到 2 人。在时刻 t31 中,说话人数为 3 人。在时刻 t32 中,说话人数为 2 人。

[0060] 说话人数判定部 124 如以上那样判定麦克 123 周围的说话人的人数。说话人数判定部 124 经由网络 200 向控制终端 305 输出。控制终端 305 基于其判定结果来控制扬声器 122 的音量。

[0061] 并且,说话人数判定部 124 基于说话人数判定部 124 判定出的现在麦克 123 周围的说话人数以及某一定时间内的过去的麦克 123 周围的说话人数,计算出周围的说话人数的单纯移动平均(以下,记载为移动平均说话人数)。移动平均说话人数的数据也作为说话人数判定部 124 判定的结果输出至控制终端 305。

[0062] 在以下的说明中,控制终端 305 通过基于程序自动地设定扬声器 122 的音量,输出控制扬声器 122 的音量的控制信号,来控制扬声器 122 的音量。其中,远程操作者也可以通过基于显示于控制终端 305 的判定结果来操作控制终端 305,从而控制扬声器 122 的音量。具体而言,将控制扬声器 122 的音量的控制信号从控制终端 305 输出至控制部 125,控制部 125 根据该控制信号来控制扬声器 122 的音量。由此,远程操作者能够控制扬声器 122 的音量。

[0063] 这里,说话人数判定部 124 在麦克 123 获取到的声音中,将发出一定等级以上的音量的声音的人计数为说话人数,发出小于一定等级的音量的声音的人不计数在说话人数中。“发出一定等级以上的音量的声音的人”是在通常的对人距离中与人进行交谈的人,“发出小于一定等级的音量的声音的人”表示用低声细语、耳语这样小的声音与人进行交谈的人。说话人数判定部 124 例如将音压等级的阈值设为 55dB,将发出其以上的音量的声音的人计数在说话人数中,从而仅将发出一定等级以上的音量的声音的人计数在说话人数中。这是因为相对于一般的交谈的音压等级约为 60dB,低声细语的音压等级是 30 ~ 40dB,通过将音压等级的阈值设为 55dB,能够仅将进行一般的交谈的人计数在说话人数中。此外,音压等级的阈值也可以是 50dB、40dB 等其他值。

[0064] 控制终端 305 基于被设定的阈值以及由说话人数判定部 124 计算出的移动平均说话人数,将周围的状态分类成以下的 3 个状态来进行判定。

[0065] a. 多个人正在聊天的状态

[0066] b. 比较安静地等待某人开始说话的状态(等待演讲的状态)

[0067] c. 特定的人进行演讲的状态或者马上就要进行演讲的状态

[0068] 以下,对于详细进行说明。

[0069] 在控制终端 305,在麦克 123 周围的说话人数中,作为第一阈值 R1 设定有 1.5,作为第二阈值 R1' 设定有 2.5,作为第三阈值 R2 设定有 3.5,作为第四阈值 R2' 设定有 4.5。这里第一阈值 R1 ~ 第四阈值 R2' 的大小关系是  $R1 < R1' < R2 < R2'$ 。

[0070] 并且,作为移动平均说话人数的一定时间内的倾斜度  $\alpha$ (微分值)的阈值设定有第五阈值 R3 以及第六阈值 R3'(R3 以及 R3' 是任意负数)。这里,第一阈值 R1、第二阈值 R1' 以及第五阈值 R3 是用于供控制终端 305 判定周围的状况是特定的人进行演讲的状态等还

是等待演讲的状态的阈值。另外,第三阈值 R2、第四阈值 R2' 以及第六阈值 R3' 是用于供控制终端 305 判定周围的状况是多个人正在聊天的状态还是等待演讲的状态的阈值。以下,对于具体的控制终端 305 的判定进行说明。

[0071] 如果说话人数判定部 124 计算出的移动平均说话人数是第三阈值 R2 以上(即,如果移动平均说话人数是 3.5 人以上),则作为原则,控制终端 305 将周围的状况判定为是多个人正在聊天的状态(状态 a)。是因为人数不是少数的人正在进行交谈,所以考虑为周围的状态不是正在进行演讲等的状态,而是正在进行聊天的状态。

[0072] 但是,即使移动平均说话人数是第三阈值 R2 以上,在移动平均说话人数小于第四阈值 R2', 并且移动平均说话人数的一定时间内的倾斜度  $\alpha$  (微分值) 小于第六阈值 R3' 的情况下,控制终端 305 也将周围的状况判定为是状态 b。换句话说,控制终端 305 将周围的状况判定为从多个人正在聊天的状态突然过渡到等待演讲的状态。这里第六阈值 R3' 是被视为在移动平均说话人数小于第四阈值 R2' 时,移动平均说话人数以倾斜度第六阈值 R3' 减少,从而在特定时间内(例如 3 秒以内)移动平均说话人数减少到小于第三阈值 R2 的值。

[0073] 如果说话人数判定部 124 计算出的移动平均说话人数是第一阈值 R1 以上并且小于第三阈值 R2,则作为原则,控制终端 305 将周围状况判定为是安静地等待某人开始说话的状态(状态 b)。周围的状态是少数人正在交谈的状态,虽然不是特定的人正在演讲的状态或者演讲马上就要进行的状态,但是也不是周围的人随意地正在聊天的状态。因此,周围的状态考虑为是周围的人等待演讲等,并且少数人正在说话的状态这是妥当的。

[0074] 但是,即使移动平均说话人数是第一阈值 R1 以上,在移动平均说话人数小于第二阈值 R1', 并且移动平均说话人数的一定时刻内的倾斜度  $\alpha$  小于第五阈值 R3 的情况下,控制终端 305 也将周围的状况判定为是状态 c。换句话说,控制终端 305 将周围的状况判定为从等待演讲的状态突然过渡到特定的人正在进行演讲的状态等。这里第五阈值 R3 是被视为在移动平均说话人数小于第二阈值 R1' 时,移动平均说话人数以倾斜度第五阈值 R3 减少,从而在特定时间内移动平均说话人数减少到小于第一阈值 R1 的值。

[0075] 如果说话人数判定部 124 计算出的移动平均说话人数小于第一阈值 R1,则控制终端 305 将周围的状况判定为是特定的人正在进行演讲的状态或者演讲马上就要进行的状态(状态 c)。这是因为在麦克 123 周围的说话人数是 1 人的情况下,考虑为在周围不进行交谈,而正在进行对于听众的祝辞、致辞、讲演等演讲。另外,在周围的说话人数是 0 人的情况下,考虑为是麦克 123 周围的人正在等待马上就要开始的演讲而不说话的状态。

[0076] 综上所述,控制终端 305 如下那样判定麦克 123 周围的状态是上述状态 a、b、c 中的哪一个状态。如果移动平均说话人数是第三阈值 R2 以上,则作为原则,控制终端 305 判定为麦克 123 周围的状态是状态 a。但是,控制终端 305 在移动平均说话人数小于第四阈值 R2' 并且是第三阈值 R2 以上且移动平均说话人数的倾斜度小于第六阈值 R3' 的情况下,判定为麦克 123 周围的状态是状态 b。如果移动平均说话人数小于第三阈值 R2 且是第一阈值 R1 以上,则作为原则,控制终端 305 判定为麦克 123 周围的状态是状态 b。但是,控制终端 305 在移动平均说话人数小于第二阈值 R1' 并且是第一阈值 R1 以上且移动平均说话人数的倾斜度小于第五阈值 R3 的情况下,判定为麦克 123 周围的状态是状态 c。如果移动平均说话人数小于第一阈值 R1,则控制终端 305 判定为麦克 123 周围的状态是状态 c。

[0077] 图 7 是表示某特定时刻中的控制终端 305 的状态判定的处理的一个例子的流程

图。以下,对于控制终端 305 的状态判定处理的例子进行说明。

[0078] 控制终端 305 首先判定移动平均说话人数是否小于第一阈值 R1 (步骤 S1)。如果移动平均说话人数小于第一阈值 R1 (步骤 S1 的是),则控制终端 305 将麦克 123 周围的状态判定为状态 c (步骤 S2)。

[0079] 如果移动平均说话人数是第一阈值 R1 以上 (步骤 S1 的否),则控制终端 305 判定移动平均说话人数是否是第三阈值 R2 以上 (步骤 S3)。

[0080] 在移动平均说话人数小于第三阈值 R2 的情况 (步骤 S3 的否)下,控制终端 305 进一步判定是否移动平均说话人数小于第二阈值 R1', 并且移动平均说话人数的倾斜度  $\alpha$  小于第五阈值 R3 (步骤 S4)。

[0081] 在移动平均说话人数小于第二阈值 R1', 并且移动平均说话人数的倾斜度  $\alpha$  小于第五阈值 R3 的情况 (步骤 S4 的是)下,控制终端 305 将麦克 123 周围的状态判定为状态 c (步骤 S2)。

[0082] 在移动平均说话人数是第二阈值 R1' 以上,或者移动平均说话人数的倾斜度  $\alpha$  是第五阈值 R3 以上的情况 (步骤 S4 的否)下,控制终端 305 将麦克 123 周围的状态判定为状态 b (步骤 S5)。

[0083] 返回步骤 S3 的判定处理继续说明。在移动平均说话人数是第三阈值 R2 以上的情况 (步骤 S3 的是)下,控制终端 305 进一步判定是否移动平均说话人数小于第三阈值 R2', 并且移动平均说话人数的倾斜度  $\alpha$  小于第六阈值 R3' (步骤 S6)。

[0084] 在移动平均说话人数小于第三阈值 R2', 并且移动平均说话人数的倾斜度  $\alpha$  小于第六阈值 R3' 的情况 (步骤 S6 的是)下,控制终端 305 将麦克 123 周围的状态判定为状态 b (步骤 S5)。

[0085] 在移动平均说话人数是第三阈值 R2' 以上,或者移动平均说话人数的倾斜度  $\alpha$  是第六阈值 R3' 以上的情况 (步骤 S6 的否)下,控制终端 305 将麦克 123 周围的状态判定为状态 a (步骤 S7)。

[0086] 此外,即使为说话人数判定部 124 计算出的移动平均说话人数与阈值 R1 或者 R2 的大小关系发生变化的状态,如果该状态不继续规定时间 T 以上,则控制终端 305 也不视为周围的状态发生了变化。即,即使在视为周围的状态在小于规定时间 T 内发生了变化的状态下,控制终端 305 也不判定为周围的状态发生了变化。这是为了避免即使实际上周围的状态没有发生变化,控制终端 305 也判定周围的状态发生了变化这一情况的措置。规定的时间 T 是视为麦克 123 周围的状态临时变化的程度较短的时间,例如 1 ~ 5 秒左右的时间 (以下也相同)。

[0087] 图 8 是表示说话人数判定部 124 判定出的说话人数以及移动平均说话人数的一个例子的曲线图。使用该图 8,对控制终端 305 控制扬声器 122 的音量的一个例子进行说明。此外,在图 8 中,横轴表示时刻,纵轴表示说话人数。图 8 的实线的曲线图是表示特定时刻中的麦克 123 周围的说话人数的经过的曲线图。实线的曲线图的数据作为说话人数判定部 124 判定的结果输出至控制终端 305。

[0088] 图 8 的虚线的曲线图是将作为对象的时刻之前的时刻的说话人数加入考虑而由说话人数判定部 124 计算出的移动平均说话人数的曲线图。该虚线的曲线图的数据也作为说话人数判定部 124 判定的结果输出至控制终端 305。

[0089] 在图 8 中,在期间 A(即,时刻  $t_{41} \sim t_{42}$ ) 中,移动平均说话人数超过第三阈值  $R_2$ 。另外,移动平均说话人数的倾斜度是第六阈值  $R_3'$  以上。因此,在上述的图 7 的处理流程中,控制终端 305 执行步骤  $S_1 - S_3 - S_6$  的判定,将周围的状况判定为是多个人正在聊天的状态(状态 a)(步骤  $S_7$ )。

[0090] 此时,控制终端 305 执行使扬声器 122 的音量为大音量(例如 70dB)的控制。即,控制终端 305 不减小扬声器 122 的音量。这是因为考虑由于周围的人在聊天中,即使不减小扬声器 122 的音量,也不会影响周围的人。相反,若控制终端 305 减小扬声器 122 的音量,则来自扬声器 122 的声音与周围的声音相比较过于小,可能机器人 100 的交谈对象(即,机器人 100 的远程操作者的交谈对象)变得听不见。因此,控制终端 305 在期间 A 中将最初设定的扬声器 122 的音量设为大音量。

[0091] 并且,在图 8 中,在时刻  $t_{43}$  中,为移动平均说话人数小于第一阈值  $R_1$ ,并且移动平均说话人数的倾斜度  $\alpha$  ( $\alpha < 0$ ) 小于第五阈值  $R_3$ 。根据以上内容,控制终端 305 执行步骤  $S_1$  的判定,判定为麦克 123 周围的状态从少数人正在进行交谈的状态(状态 b)过渡到特定的人正在演讲的状态等(状态 c)(步骤  $S_2$ )。基于该判定,控制终端 305 执行减小扬声器 122 的音量(例如使音量为 50dB)的控制。该状态被认为会立刻移至马上就要进行演讲的状态。因此,控制终端 305 预先进行降低扬声器的音量的控制,这是因为考虑了最好与周围立刻移至安静的环境相对应。

[0092] 在时刻  $t_{43}$  之后周围的移动平均说话人数也为第一阈值  $R_1$  以下。控制终端 305 从此处将周围的状况判定为是特定的人正在进行演讲的状态等(状态 c)。因此,控制终端 305 仍然控制为减小扬声器 122 的音量。

[0093] 如以上那样,控制终端 305 在时刻  $t_{43}$  以后的期间 B 中,执行减小扬声器 122 的音量的控制。

[0094] 此外,控制终端 305 在判定为麦克 123 周围的状态是状态 b 的情况下,执行使扬声器 122 的音量为中等程度(例如使音量为 60dB)的控制。这样,控制终端 305 按照麦克 123 周围的状态是状态 a、状态 b、状态 c 的顺序,较大地设定扬声器 122 的音量。

[0095] 此外,控制终端 305 也可以不基于由说话人数判定部 124 计算出的移动平均说话人数,而基于由说话人数判定部 124 判定出的当前麦克 123 周围的说话人数,执行图 7 所示的判定处理。此时,在图 8 的期间 B 内的时刻  $t_{45} \sim t_{46}$ 、时刻  $t_{47} \sim t_{48}$ ,麦克 123 周围的说话人数为 2 人(即,第一阈值  $R_1$  以上)。但是,时刻  $t_{45} \sim t_{46}$ 、时刻  $t_{47} \sim t_{48}$  的期间分别是小于规定时间  $T$  的期间。因此,控制终端 305 判定为特定的人正在进行演讲的状态(状态 c)正在继续。由此,控制终端 305 在期间 B 中控制为保持减小扬声器 122 的音量的状态,不返回原来的音量。

[0096] 但是,通过控制终端 305 基于移动平均说话人数执行图 7 所示的判定处理,即便麦克 123 周边的当前的说话人数短时间内发生了变动,也缓和该变动来判定状态。换句话说,能够防止控制终端 305 判定的麦克 123 周边的状态频繁地切换。因此,为了防止控制终端 305 对麦克 123 的音量的控制频繁地改变,优选控制终端 305 基于移动平均说话人数执行图 7 所示的判定处理。

[0097] 以上说明的实施方式 1 的通讯系统 10 能够根据周围的环境来变更音量。该通讯系统 10 至少具备麦克 123、扬声器 122、扬声器 303、麦克 301、说话人数判定部 124 以及控

制终端 305。麦克 123 获取周围的声音作为声音信号。扬声器 122 设置于麦克 123 的附近（例如距离麦克 123 数十 cm ~ 1m 左右的距离），并输出声音。扬声器 303 从麦克 123 获取声音信号，并将该声音信号作为声音输出。麦克 301 获取周围的声音作为声音信号，将获取到的该声音信号对扬声器 122 输出。说话人数判定部 124 根据麦克 123 获取到的声音信号，判定麦克 123 周围的说话人的人数（例如麦克 123 的半径数米内的说话人的人数）。控制终端 305 基于说话人数判定部 124 的判定，控制扬声器 122 的音量。通过这样的构成，通讯系统 10 能够根据麦克 123 周围的说话人的人数，来控制扬声器 122 的音量。因此，通讯系统 10 能够根据麦克 123 周围的环境变更通讯的音量。

[0098] 图 9 是表示以往的通讯系统的构成的构成图。图 9 的通讯系统 90 具备机器人 700、网络 800 以及远程操作侧 900，机器人 700 经由网络 800 与远程操作侧 900 连接。

[0099] 机器人 700 具备扬声器 701、麦克 702、放大器 703 以及控制部 704。扬声器 701、麦克 702、放大器 703 分别与图 4 的扬声器 122、麦克 123、放大器 127 对应。控制部 704 执行扬声器 122 以及麦克 123 的控制。

[0100] 远程操作侧 900 具备麦克 901、放大器 902、扬声器 903 以及回波抵消器 904。麦克 901、放大器 902、扬声器 903 以及回波抵消器 904 分别与图 4 的麦克 301、放大器 302、扬声器 303 以及回波抵消器 304 对应。

[0101] 在通讯系统 90 中，在远程操作侧 900 的远程操作者利用扬声器 903 听在分离地方的麦克 702 获取到的声音。另外，远程操作者通过对于麦克 901 发声，将自身的声音从在分离的地方的扬声器 701 输出。这样一来，远程操作者能够与在分离的地方的人（即，在机器人 700 的附近的人）进行交谈。

[0102] 但是，远程操作者无法准确地识别机器人 700 的周围的环境。即，远程操作者无法识别机器人 700 周边的人现在正在聊天，还是正在安静地听某人的演讲等。因此，例如在机器人周边的人正在安静地听某人的演讲的时候，若远程操作者利用麦克 901 与机器人 700 的附近的人打招呼，则较大的声音从扬声器 701 发出，可能会扰乱周围的气氛。

[0103] 与此相对，实施方式 1 的通讯系统 10 能够根据麦克 123 周围的环境变更通讯的音量。

[0104] 此外，控制终端 305 也可以在说话人数判定部 124 判定出的麦克 123 周围的说话人的人数小于阈值的情况下，控制为与该人数是阈值以上的情况相比较，减小扬声器 122 的音量（这里，所谓周围的说话人的人数既可以是说话人数判定部 124 判定出的当前麦克 123 周围的说话人数，也可以是说话人数判定部 124 计算出的移动平均说话人数。这点以下的说明也相同。）。通过这样的构成，在麦克 123 周围的说话人的人数小于阈值时，从位于其附近的扬声器 122 输出的声音的音量变小。因此，通讯系统 10 能够防止阻碍在扬声器 122 的周围进行的演讲这一情况。

[0105] 控制终端 305 也可以在说话人数判定部 124 判定出的麦克 123 周围的说话人的人数小于阈值的情况下，在比规定的时间 T 少的时间内该人数变为阈值以上时，控制为保持减小扬声器 122 的音量的状态。通过这样的构成，控制终端 305 能够在麦克 123 的周围正在进行演讲的情况下，在较短的期间内多个人的声音发出时，保持减小扬声器 122 的音量的状态。由此，通讯系统 10 能够防止阻碍在扬声器 122 的周围进行的演讲这一情况。

[0106] 控制终端 305 也可以在说话人数判定部 124 判定出的麦克 123 周围的说话人的人

数是阈值以上的情况下,在比规定的时间 T 少的时间内该人数小于阈值时,控制为保持增大扬声器 122 的音量的状态。通过这样的构成,控制终端 305 能够在麦克 123 的周围多个人正在进行交谈的情况下,在较短的期间内为安静的状态时,也将扬声器 122 的音量保持原样。由此,通讯系统 10 通过减小扬声器 122 的音量,能够防止阻碍使用扬声器 303 以及麦克 301 进行说话的人(远程操作者)的通讯这一情况。

[0107] 控制终端 305 也可以在说话人数判定部 124 判定出的麦克 123 周围的说话人的人数在特定时间内减少得比规定数量大的情况(换句话说倾斜度  $\alpha$  小于规定的倾斜度的情况)下,控制为减小扬声器 122 的音量。通过这样的构成,控制终端 305 能够检测将麦克 123 周围的状态从多个人正在聊天的状态变化到听众正在安静地等待演讲的状态这一情况,减小扬声器 122 的音量。由此,通讯系统 10 能够防止阻碍在扬声器 122 的周围进行的演讲。

[0108] 控制终端 305 也可以在说话人数判定部 124 判定出的麦克 123 周围的说话人的人数是阈值以下并且正在减少的情况下,控制为减小扬声器 122 的音量。通过这样的构成,控制终端 305 检测麦克 123 周围的状态是听众正在安静地等待演讲的状态这一情况,来减小扬声器 122 的音量。由此,通讯系统 10 能够防止阻碍在扬声器 122 的周围进行的演讲这一情况。

[0109] 此外,也可以代替控制终端 305,而由机器人 100 的控制部 125 执行图 7 所记载的处理流程。这里,机器人 100 是在与人之间执行使用例如声音的通讯的机器人,具备扬声器 122、麦克 123、说话人数判定部 124 以及控制部 125。

[0110] 图 10 是表示这种机器人的处理的一个例子的框构成图。扬声器 122 输出声音。麦克 123 获取机器人 100 的周围的声音作为声音信号。说话人数判定部 124 根据麦克 123 获取到的声音信号,判定机器人 100 的周围的说话人的人数。控制部 125 基于说话人数判定部 124 的判定结果,控制扬声器 122 的音量。通过这样的构成,机器人 100 能够根据周围的说话人的人数来控制扬声器 122 的音量。因此,机器人 100 能够根据周围的环境自主地变更通讯的音量。

[0111] 并且,在实施方式 1 中,也公开有通讯系统中的音量控制的方法。该通讯系统具有:麦克 123,其获取周围的声音作为声音信号;扬声器 122,其设置于麦克 123 的附近,输出声音;扬声器 303,其从麦克 123 获取声音信号,并将该声音信号作为声音输出;以及麦克 301,其获取周围的声音作为声音信号,将获取到的该声音信号对扬声器输出。该音量控制的方法至少包括以下的步骤(a)~(b)。

[0112] (a) 根据麦克 123 获取到的声音信号,来判定麦克 123 周围的说话人的人数;以及

[0113] (b) 基于上述判定,来控制扬声器 122 的音量。

[0114] 通过这样的构成,通讯系统能够根据麦克 123 周围的说话人的人数来控制扬声器 122 的音量。因此,通讯系统能够根据扬声器 122 的周围的环境变更通讯的音量。

[0115] 图 7 所示的流程图的处理方法只不过是一个例子。例如,在图 7 中,控制终端 305 也可以不是先执行步骤 S3 而如上述那样先执行步骤 S1 的判定。另外,第一阈值 R1、第二阈值 R1'、第三阈值 R2 以及第四阈值 R2' 的值只不过是具体例,只要满足  $R1 < R1' < R2 < R2'$  的大小关系即可,也可以是其他的值。第五阈值 R3 以及第六阈值 R3' 的值也能够基于第一阈值 R1 ~ 第四阈值 R2' 的值变更。

[0116] 此外,在上述的处理例中,控制终端 305 在麦克 123 周围的状态是状态 a 的情况

下,将扬声器 122 的音量设定为音量 70dB。这里,控制终端 305 既可以在麦克 123 周围的状态是状态 a 时不将扬声器 122 的音量一直设为音量 70dB,也可以以扬声器 122 的音量稍稍(例如 1、2dB 左右)变化的方式调整扬声器 122 的音量。麦克 123 周围的状态是状态 b、状态 c 时也相同。这样一来,控制终端 305 按照麦克 123 周围的状态是状态 a、状态 b、状态 c 的顺序,将扬声器 122 的音量较大地设定即可。

[0117] 此外,在麦克 123 周围的状态是状态 b 时,控制终端 305 也可以与状态 a 时相同地设定扬声器 122 的音量。

[0118] 此外,控制终端 305 判定上述状态 a、b、c 的方法也可以考虑其他方法。以下,依次对状态 a、b、c 进行说明。

[0119] 例如,在控制终端 305 检测到以下状态的情况下,控制终端 305 能够判定为麦克 123 周围的状态是状态 a。控制终端 305 在判定为在麦克 123 的周围中规定等级以上的音量的较大声音或者普通音量的声音的说话人有多人,并且该声音继续一定时间以上的情况下,能够判定为麦克 123 周围的状态是状态 a。这里所谓“规定等级以上的音量”例如是大声的音量亦即 70 ~ 80dB 左右的音量,所谓“普通的音量”例如是一般的说话声音的音量亦即 60dB 左右的音量(以下也相同)。而且所谓“一定时间”是指足以确定出麦克 123 周围的状态的时间,例如 7 ~ 10 秒左右的时间(以下也相同)。该时间为了不将在演讲中偶尔发生的数秒左右的笑声、说话声音判定为“多个人正在聊天”的状态而设定。所谓“说话人多”例如表示说话人有 4 人以上(上述的第三阈值 R2 以上)的情况。

[0120] 并且,控制终端 305 也可以在按照上述判定为麦克 123 周围的状态是状态 a 的情况下,在比规定时间 T 少的时间内不满足上述条件的情况下,也判定为麦克 123 周围的状态是状态 a。这是因为即使多个人正在聊天的状态下,也可能存在一瞬间成为没有交谈的安静状态的情况,所以若将该状态判定为状态 b 或者状态 c,则扬声器的音量未被基于现实的状态准确地控制。此外,规定的时间 T 按照上述例如是 1 秒 ~ 5 秒左右的时间。

[0121] 并且,控制终端 305 不只在某时刻中的麦克 123 周围的说话人数是第三阈值 R2 以上的情况下,在一定时间内的麦克 123 周围的个人的说话人数比规定数量多的情况下,也可以判定为麦克 123 周围的状态是状态 a。此外,一定时间的定义遵从上述内容。

[0122] 例如,控制终端 305 在 10 秒期间 5 人以上的说话人在麦克 123 周围的情况下,能够判定为麦克 123 周围的状态是状态 a。这里说话人数判定部 124 例如在 10 秒的期间最初 A 先生、B 先生、以及 C 先生进行交谈,接下来 A 先生、D 先生、以及 E 先生进行交谈的情况下,判定为 5 个说话人在麦克 123 的周围。说话人数判定部 124 按照上述检测不同的固有振动,来判定为进行交谈的是 A 先生 ~ E 先生中的哪一个。

[0123] 这是考虑了在多个人在麦克 123 的周围正在聊天的情况下,也并不是人一直说话,也存在说话中断的情况。在这样的中断有短时间(例如 1 ~ 3 秒)时,若控制终端 305 仅考虑该短时间的状态来判定麦克 123 周围的状态是状态 b 或者状态 c,则扬声器的音量未被基于现实的状态准确地控制。但是,控制终端 305 在即使有交谈的中断,一定时间内的麦克 123 周围的个人的说话人数也比规定数量多的情况下,判定为麦克 123 周围的状态是状态 a。由此,控制终端 305 能够进行麦克 123 周围的准确状态的判定。

[0124] 并且,该判定方法特别是在聊天中某说话人改变对象来说话的情况下,能够进行准确的状态的判定。例如,在 10 秒的期间最初 A 先生、B 先生、以及 C 先生进行交谈,接下来

A 先生、D 先生、以及 E 先生进行交谈的情况下,特定时刻中的说话人的人数是 3 人,即小于第三阈值 R2,但一定时间内的个人的说话人的人数是 5 人,即为第三阈值 R2 以上。这样的時候,控制终端 305 仅考虑该特定的时刻的状态来判定为麦克 123 周围的状态是状态 b,则扬声器的音量未被基于现实的状态准确地控制。然而,控制终端 305 在某特定的时刻中的说话人的人数少的情况下,在一定时间内的麦克 123 周围的个人的说话人数也比规定数量多的情况下,判定为麦克 123 周围的状态是状态 a。由此,控制终端 305 能够进行麦克 123 周围的准确状态的判定。

[0125] 另外,在控制终端 305 检测到以下的状态的情况下,控制终端 305 能够判定为麦克 123 周围的状态是状态 b。控制终端 305 也可以在麦克 123 的周围中发出规定等级以上的音量的较大的声音或者普通的音量的声音的说话人的数量在特定时间内减少了规定数量以上的情况下,判定为麦克 123 周围的状态是状态 b。这里“规定等级以上的音量”、“普通的音量”的例子如上所述。

[0126] 控制终端 305 在发出一般的说话声音的说话人的数量例如在 3 秒以内减少了 4 人以上的情况下,能够判定为麦克 123 周围的状态从状态 a 过渡至状态 b。这样,在发出规定水平以上的音量的较大声音或者普通音量的声音的说话人的数量在短时间急剧地减少的情况下,控制终端 305 能够判定为麦克 123 周围的状态是状态 b。该状态表示以演讲马上就要开始为契机从多个人在麦克 123 的周围正在聊天的状态移至正在聊天的人停止聊天状态的状态。

[0127] 并且,控制终端 305 在麦克 123 的周围中发出规定等级以上的音量的较大的声音或者普通的音量的声音的说话人的数量是一定以下的状态下,在一定时间内减少了规定数量以上说话人数的情况下,能够判定为麦克 123 周围的状态从状态 a 变为状态 b。例如控制终端 305 在发出一般的说话声音的说话人的数量在特定的时刻中是 10 人以下的情况下,在 10 秒以内减少了 4 人以上说话人的数量的情况下,能够判定为麦克 123 周围的状态从状态 a 变为状态 b。这样,控制终端 305 在发出规定等级以上的音量的较大的声音或者普通的音量的声音的说话人的数量是一定以下,且说话人的数量还继续减少的情况下,能够判定为麦克 123 周围的状态是状态 b。该状态表示以演讲马上就要开始为契机从正在进行聊天的人不多的状态移至正在聊天的人依次停止聊天状态的状态。

[0128] 并且,在控制终端 305 检测到以下状态的情况下,控制终端 305 能够判定为麦克 123 周围的状态是状态 c。控制终端 305 在麦克 123 的周围中在一定时间以上发出规定等级以上的音量的较大声音的说话人有少数(1 人~3 人左右)的情况下,能够判定为麦克 123 周围的状态是状态 c。这表示少数人在麦克 123 的周围正在进行演讲的状态。另外,进行演讲的人并不局限于 1 人,也可以是 2 人、3 人这样的多个并且少数的人。这是因为例如在会场中,可能存在进行了主持人和来宾的通话这样的情况。

[0129] 此时,说话人数判定部 124 也可以将发出小于上述一定等级的音量的声音的人作为“用小于一定等级的音量说话的说话人”计数为说话人数。控制终端 305 在说话人数判定部 124 判定为在麦克 123 的周围有用小于一定等级的音量说话的说话人的情况下,也能够判定为麦克 123 周围的状态是状态 c。所谓“小于一定等级的音量”例如是低声细语或者耳语这样的较小的声音的音量,指 30~40dB 左右的音量。这表示少数人正在麦克 123 的周围进行演讲而另一方面其他的人为了不打扰演讲而以较小的声音说话的状态。



[0130] 并且,控制终端 305 在比规定的时间 T 少的时间(例如不足 3 秒~5 秒的时间)内发出规定等级以上的音量的较大声音或者普通音量的声音的说话人有规定数量以上的情况下,也能够判定为麦克 123 周围的状态是状态 c。该处理是用于防止在演讲期间的时间内发生笑声、聊天的情况下,控制终端 305 仅检测到该时间内的状态就判定为是多个人正在聊天的状态,而用不适合现实的做法来控制扬声器 122 的音量的处理。

[0131] 并且,控制终端 305 在没有在比规定的时间 T 少的时间(例如 2 秒~小于 5 秒的时间)内发出规定等级以上的音量的较大声音的说话人的情况下,也能够判定为麦克 123 周围的状态是状态 c。该处理是用于防止正在进行演讲的人在说话的期间插入小停顿的情况下,控制终端 305 仅检测到该期间的状态就判定为是状态 b,用不适合现实的做法来控制扬声器 122 的音量的处理。

[0132] 并且,控制终端 305 在说话人的数量从有少数说话人的状态 a 的状态减少并变为少数(例如 2、3 人)的情况下,也能够判定为麦克 123 周围的状态是状态 c。这表示在麦克 123 周围的状态从听众正在聊天的状态变化到演讲马上就要开始的状态的情况下,少数人没注意到演讲开始正在进行交谈的状态。因此,控制终端 305 能够适当地反映现实的状态,并较小地控制扬声器 122 的音量。

[0133] 如以上那样,控制终端 305 也能够不基于特定时刻的说话人的人数,而基于说话人数判定部 124 判定出的某一定时间内的说话人数,来控制扬声器 122 的音量。由此,仅在短时间内说话人数变化的情况下,控制终端 305 也能够不管该变化如何,都进行麦克 123 周围的准确状态的判定。由此,通讯系统能够根据第一麦克风的周围的环境更高精度地变更通讯的音量。

[0134] 另外,说话人数判定部 124 也可以对在一定时间内进行交谈的个人的数量进行计数,作为一定时间内的麦克 123 周围的说话人数。由此,说话人数判定部 124 在不同的人彼此在不同的时刻正在进行交谈的情况下,能够准确地把握进行了交谈的合计的人数。因此,控制终端 305 能够进行麦克 123 周围的准确状态的判定。

[0135] 实施方式 2

[0136] 以下,参照附图对本发明的实施方式 2 进行说明。实施方式 2 的机器人的外观图以及头部的构成与实施方式 1 相同,所以省略说明。

[0137] 图 11 是表示实施方式 2 的机器人 100 的内部构成的一个例子的框构成图。图 11 的机器人 100 与图 3 的机器人 100 相比,新具备比较部 128。

[0138] 图 12 是表示具有机器人 100 的通讯系统的一个例子的构成图。在通讯系统 20 中,比较部 128 通过比较从麦克 301 输出的声音信号与从麦克 123 输出的声音信号,来比较扬声器 122 输出的声音与位于麦克 123 周围的声音的大小。换句话说,比较部 128 比较机器人 100 发出的声音的大小与在机器人 100 的周围的人发出的声音的大小。

[0139] 比较部 128 将该比较结果输出至控制终端 305,控制终端 305 显示该比较结果。远程操作者基于显示于控制终端 305 的比较结果,从控制终端 305 控制扬声器 122 的音量。由此,通讯系统 30 能够使扬声器 122 的音量与扬声器 122 周围的声音的大小一致。

[0140] 例如,比较部 128 判定为扬声器 122 输出的声音的音量是 65dB,从麦克 123 获取到的声音的音量是 55dB。此时,远程操作者判定为处于从机器人 100 输出的声音比机器人 100 的周围的声音大的状态(即,机器人 100 与周围的人相比较是吵闹的状态)。远程操作者通

过该判定,从控制终端 305 向控制部 125 输出控制信号,以使扬声器 122 的音量减小 10dB 左右。由此,比较部 128 能够使扬声器 122 的音量随着机器人周围的声音的音量减小。

[0141] 相反,比较部 128 判定为扬声器 122 输出的声音的音量是 50dB,从麦克 123 获取到的声音的音量是 60dB。此时,远程操作者判定为处于从机器人 100 输出的声音比机器人 100 周围的声音小的状态(即,机器人 100 的声音周围的人难以听见的状态)。远程操作者通过该判定,从控制终端 305 向控制部 125 输出控制信号,以使扬声器 122 的音量增大 10dB 左右。由此,比较部 128 能够使扬声器 122 的音量随着机器人周围的声音的音量变大。

[0142] 远程操作者因为位于远离机器人 100 的场所,所以无法直接听见机器人 100 发出的声音和机器人 100 周边的人发出的声音。因此,远程操作者可能错误设定机器人 100 发出的声音的音量。例如,可能存在机器人 100 发出的声音的音量与机器人 100 周边的人发出的声音的音量相比过大,妨碍机器人 100 周边的人的交谈、演讲的情况。与此相反,也可能存在机器人 100 发出的声音的音量与机器人 100 周边的人发出的声音的音量相比过小,妨碍机器人 100 周边的人经由机器人 100 与远程操作者进行交谈的情况。实施方式 2 的机器人 100 以及通讯系统 30 能够通过使机器人 100 发出的声音的音量与机器人 100 周边的人发出的声音的音量一致,来解决这样的课题。

[0143] 此外,控制部 125 也可以基于比较部 128 的比较结果,自动地控制扬声器 122 的音量。对于控制部 125 的具体的控制方法如上所述。换句话说控制部 125 基于比较部 128 的比较结果控制扬声器 122 的音量,以使扬声器 122 输出的声音的音量与从麦克 123 获取到的声音的音量为几乎相同程度的大小。

[0144] 实施方式 3

[0145] 以下,参照附图对本发明的实施方式 2 进行说明。实施方式 3 的机器人的外观图以及头部的构成与实施方式 1 相同,所以省略说明。

[0146] 图 13 是表示实施方式 3 的机器人 100 的内部构成的一个例子的框构成图。图 13 的机器人 100 与图 3 的机器人 100 相比,内部不具备说话人数判定部 124。

[0147] 图 14 是表示具有机器人 100 的通讯系统的一个例子的构成图。在通讯系统 30 中,说话人数判定部作为说话人数判定部 306 设置于远程操作侧 300。说话人数判定部 306 执行的处理与实施方式 1 的说话人数判定部 124 相同,省略说明。说话人数判定部 306 将判定结果输出至控制终端 305。

[0148] 控制终端 305 将控制信号输出至控制部 125,以便基于判定结果控制扬声器 122 的音量。控制终端 305 的扬声器 122 的音量控制的方法与实施方式 1 中说明的方法相同。控制部 125 基于控制信号控制扬声器 122 的音量。

[0149] 这样,在实施方式 3 中,说话人数判定部 306 不设置于机器人 100 而设置于远程操作侧 300。这是用于抑制在说话人数判定部 306 设置于机器人 100 的情况下,将该判定结果输出至控制终端 305 时由于网络 200 产生的延迟而导致的影响的构成。这里,说话人数判定部 306 使用从麦克 123 输出的声音信号被回波抵消器 304 回波取消之前的声音信号来执行判定。

[0150] 实施方式 2 的比较部 128 也与图 14 的说话人数判定部 306 相同,能够不设置于机器人 100 而设置于远程操作侧 300。

[0151] 此外,本发明并不局限于上述实施方式,能够在不脱离主旨的范围内适当地变更。

例如,实施方式 1 中的机器人 100 作为被远程操作者操作的远程操作机器人进行了说明,但也可以是自主地判定周围的状况并输出声音等的机器人。即使是这样的机器人,也能够通过与实施方式 1 相同的判定方法,控制扬声器的音量与环境一致。并且,即使代替机器人 100,具备相同的构成部件的计算机设置于通讯系统 10,通讯系统 10 也能够起到相同的效果。

[0152] 在实施方式 1 中,说话人数判定部 124 能够不计算周围的说话人数的移动平均,执行其他的滤波处理,来计算周围的说话人数的时间的过渡。控制部 125 基于该过渡所涉及的数据,与实施方式 1 相同地控制扬声器 122 的音量。

[0153] 实施方式 1 所示的处理流程能够作为控制方法之一,使计算机(例如个人计算机、移动终端、机器人)执行。例如,也可以将处理流程作为控制程序使计算机执行。该计算机具备输出声音的扬声器、以及获取周围的声音作为声音信号的麦克风。计算机根据麦克风获取到的声音信号判定机器人的周围的说话人的人数。接下来,计算机基于该判定控制上述扬声器的音量。如以上那样,计算机执行处理。对于实施方式 2、3 所示的处理,也能够相同地使计算机执行。

[0154] 程序能够使用各种类型的非暂时性计算机可读介质(non-transitory computer readable medium)来储存,并供给至计算机。非暂时性计算机可读介质包括有各种类型的有实体的记录介质(tangible storage medium)。非暂时性计算机可读介质的例子包括磁记录介质(例如软盘、磁带、硬盘驱动器)、光磁记录介质(例如光磁盘)、CD-ROM、CD-R、CD-R/W、半导体存储器(例如,掩膜 ROM、PROM(Programmable ROM)、EPROM(Erasable PROM)、闪存 ROM、RAM(Random Access Memory))。另外,程序也可以通过各种类型的暂时性计算机可读介质(transitory computer readable medium)供给至计算机。暂时性计算机可读介质的例子包括电信号、光信号、以及电磁波。暂时性计算机可读介质能够经由电线以及光纤等的有线通信路、或者无线通信路将程序供给至计算机。

[0155] 该申请主张于 2013 年 3 月 22 日提出的日本专利申请 2013-059420 的优先权,并在此引用其全部公开内容。

[0156] 附图标记说明

[0157] 10、20、30... 通讯系统;100... 机器人;101... 头部;102... 躯干部;103... 右臂部;104... 左臂部;105... 右腿部;106... 左腿部;111... 右眼;12... 左眼;113... 口;114... 右耳;115... 左耳;121... 拍摄装置;122... 扬声器;123... 麦克;124... 说话人数判定部;125... 控制部;126... 可动部;127... 放大器;128... 比较电路;200... 网络;300... 远程操作侧;301... 麦克;302... 放大器;303... 扬声器;304... 回波抵消器;305... 控制终端;306... 说话人数判定部。

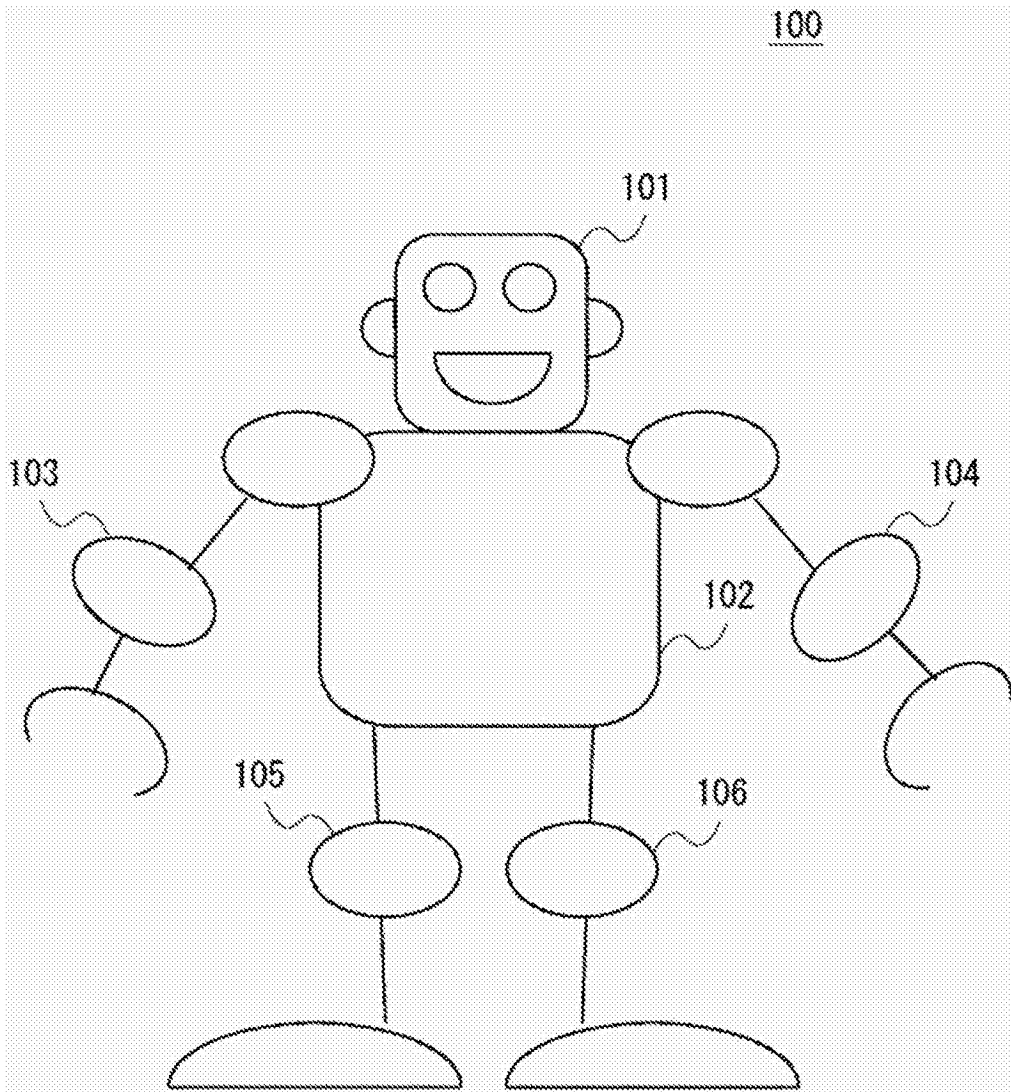


图 1

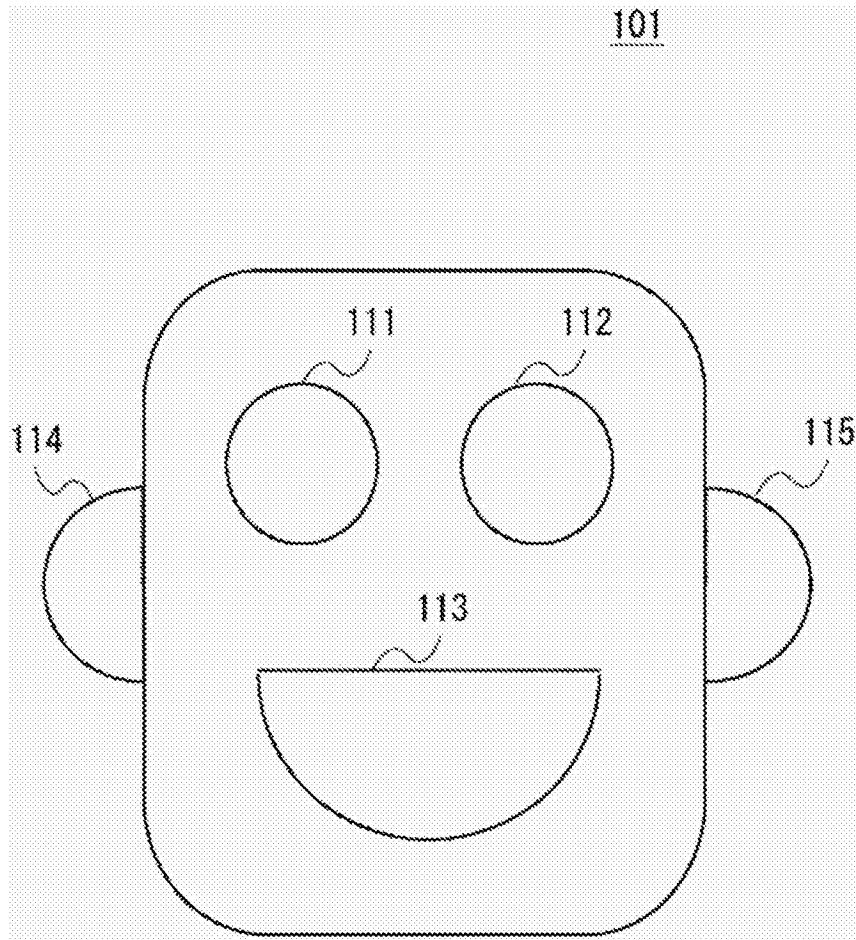


图 2

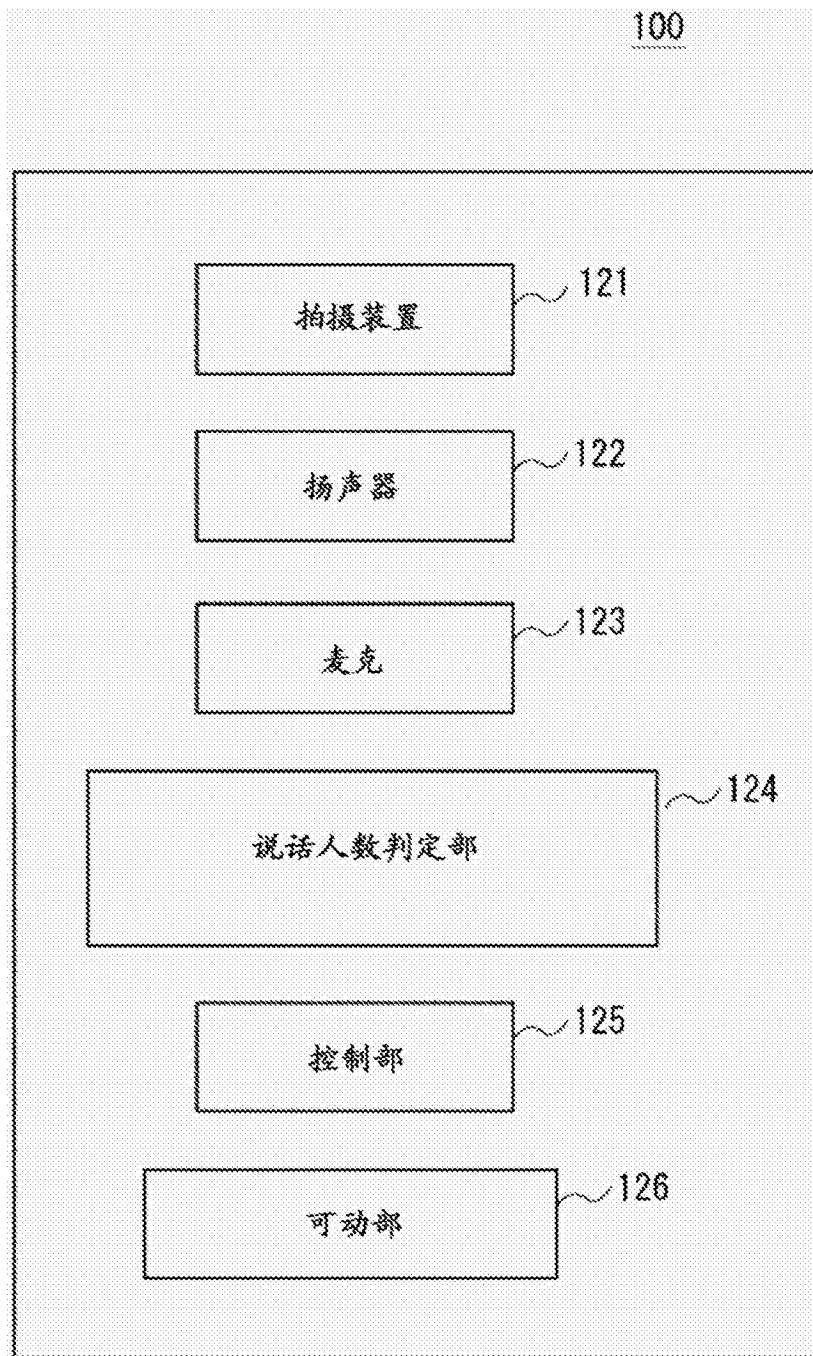


图 3

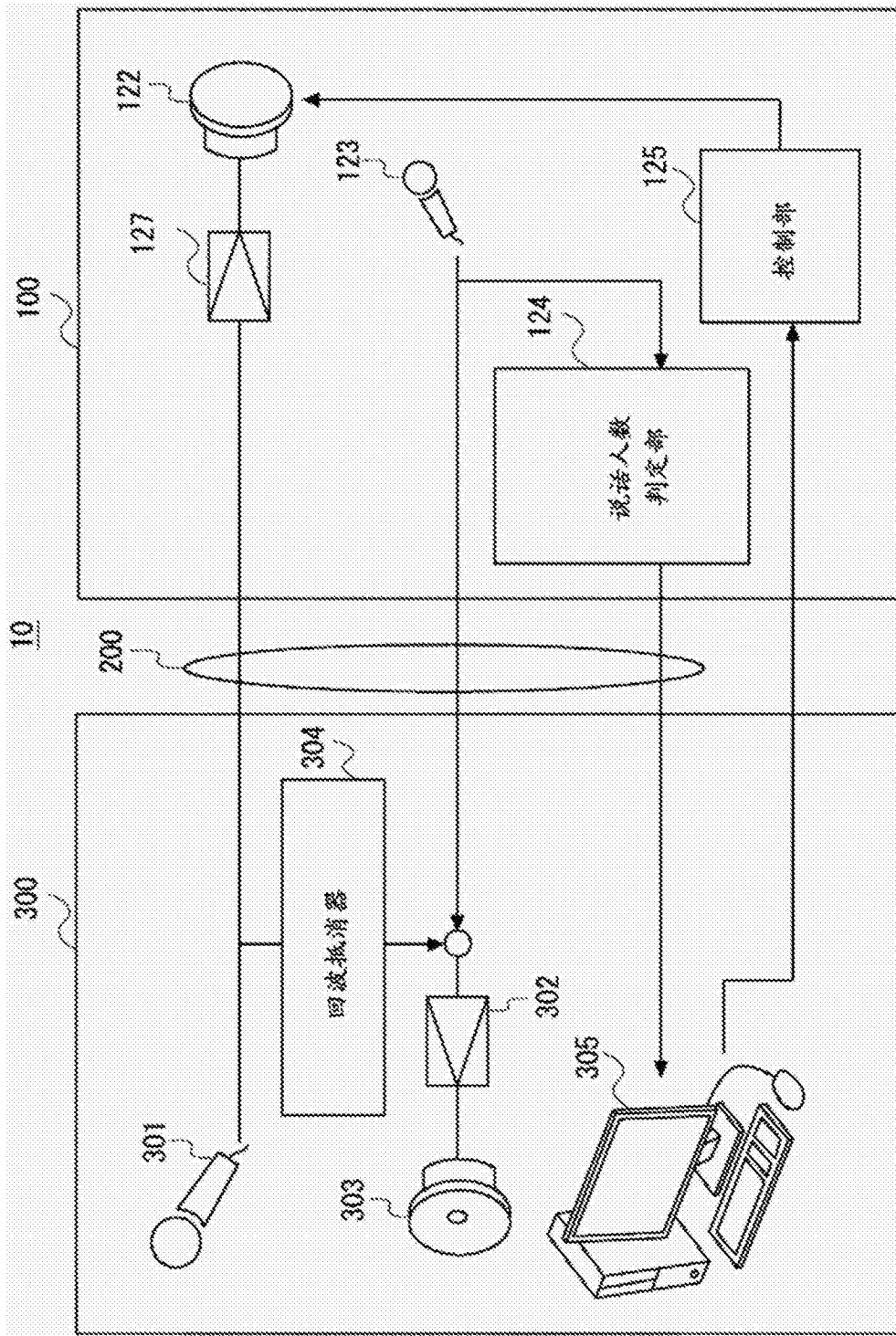


图 4

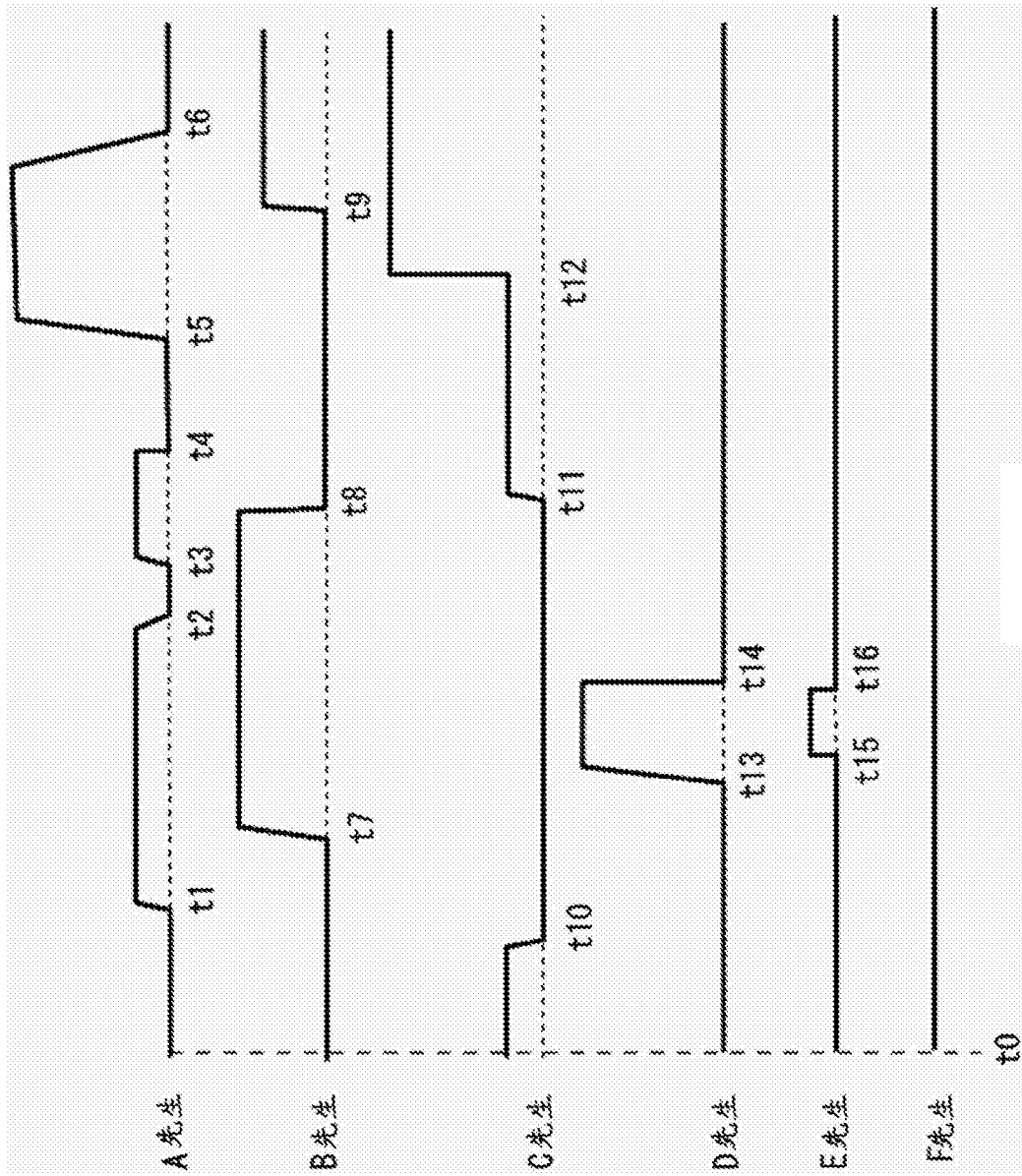


图 5



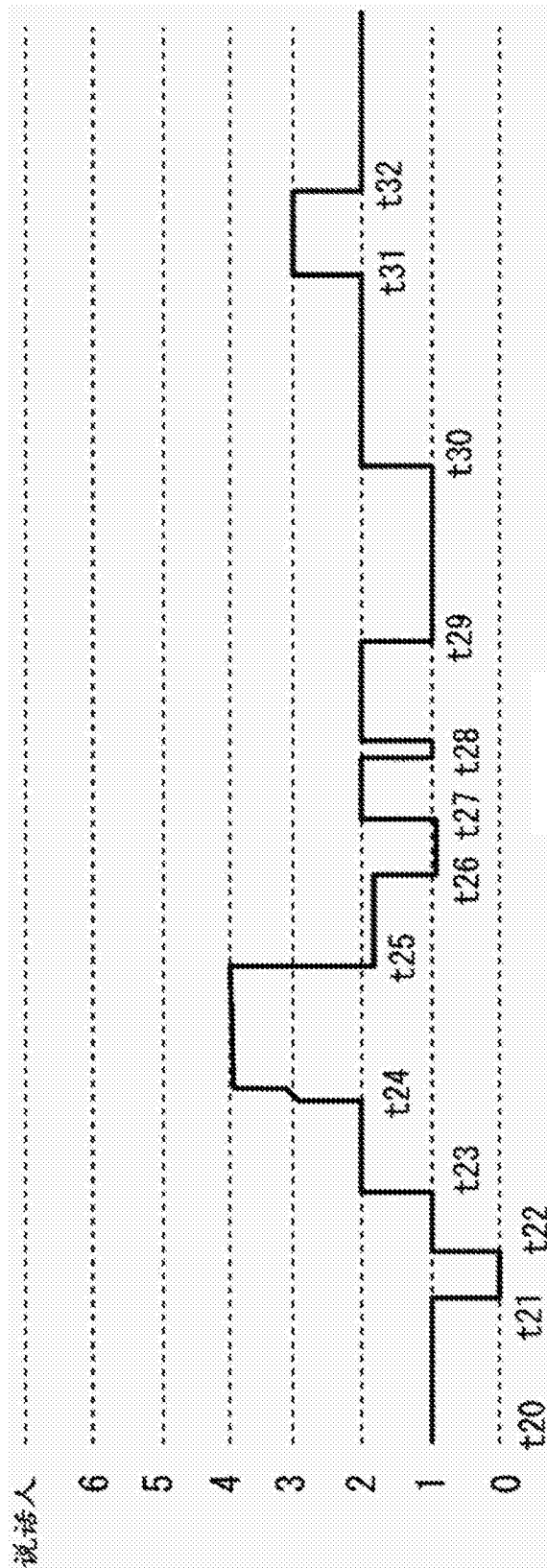


图 6

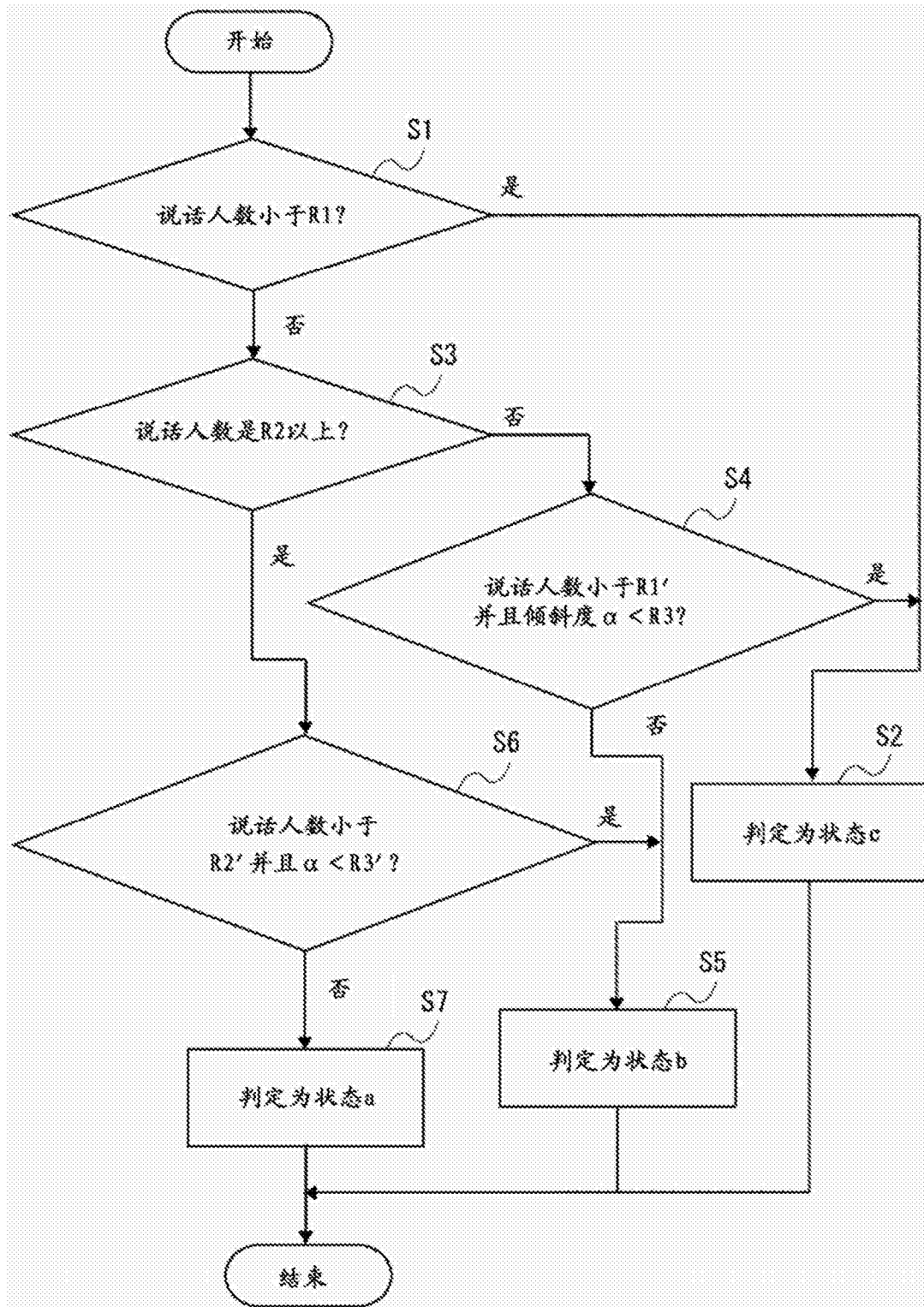


图 7

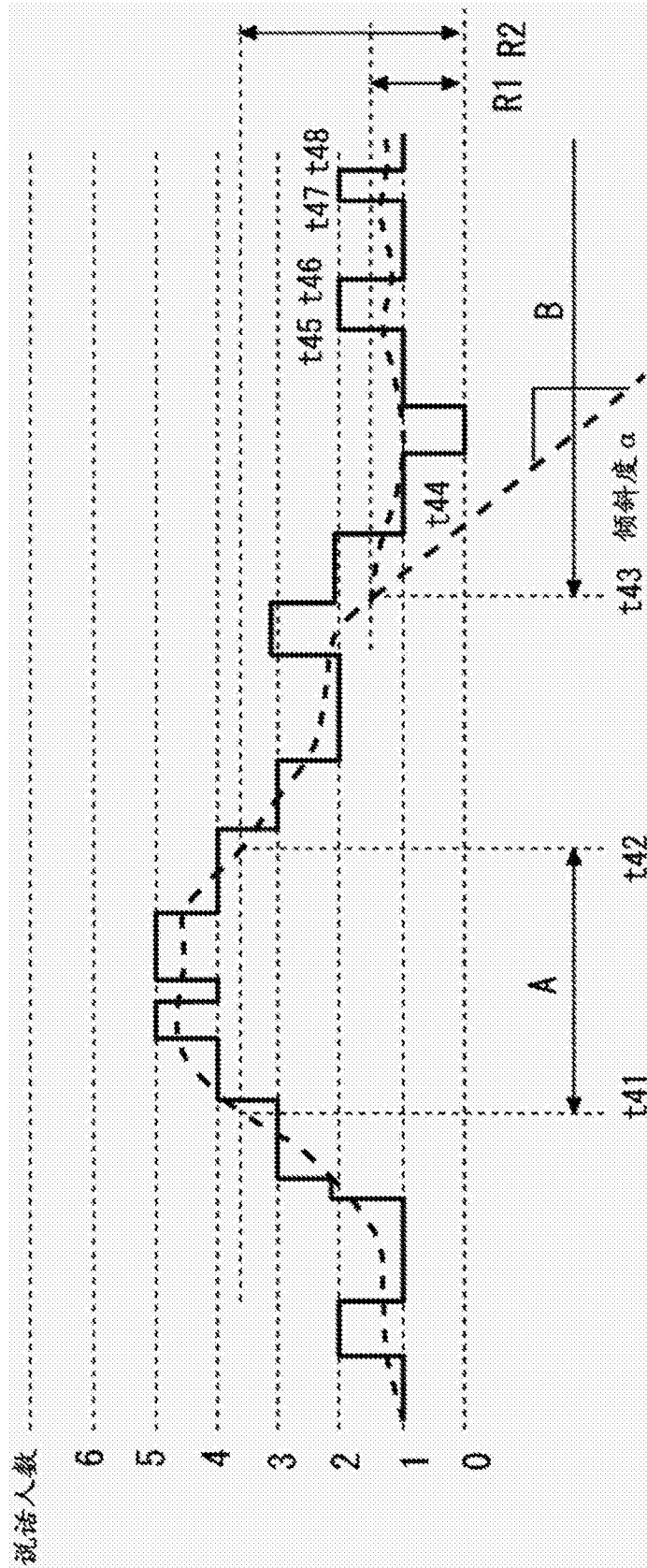


图 8

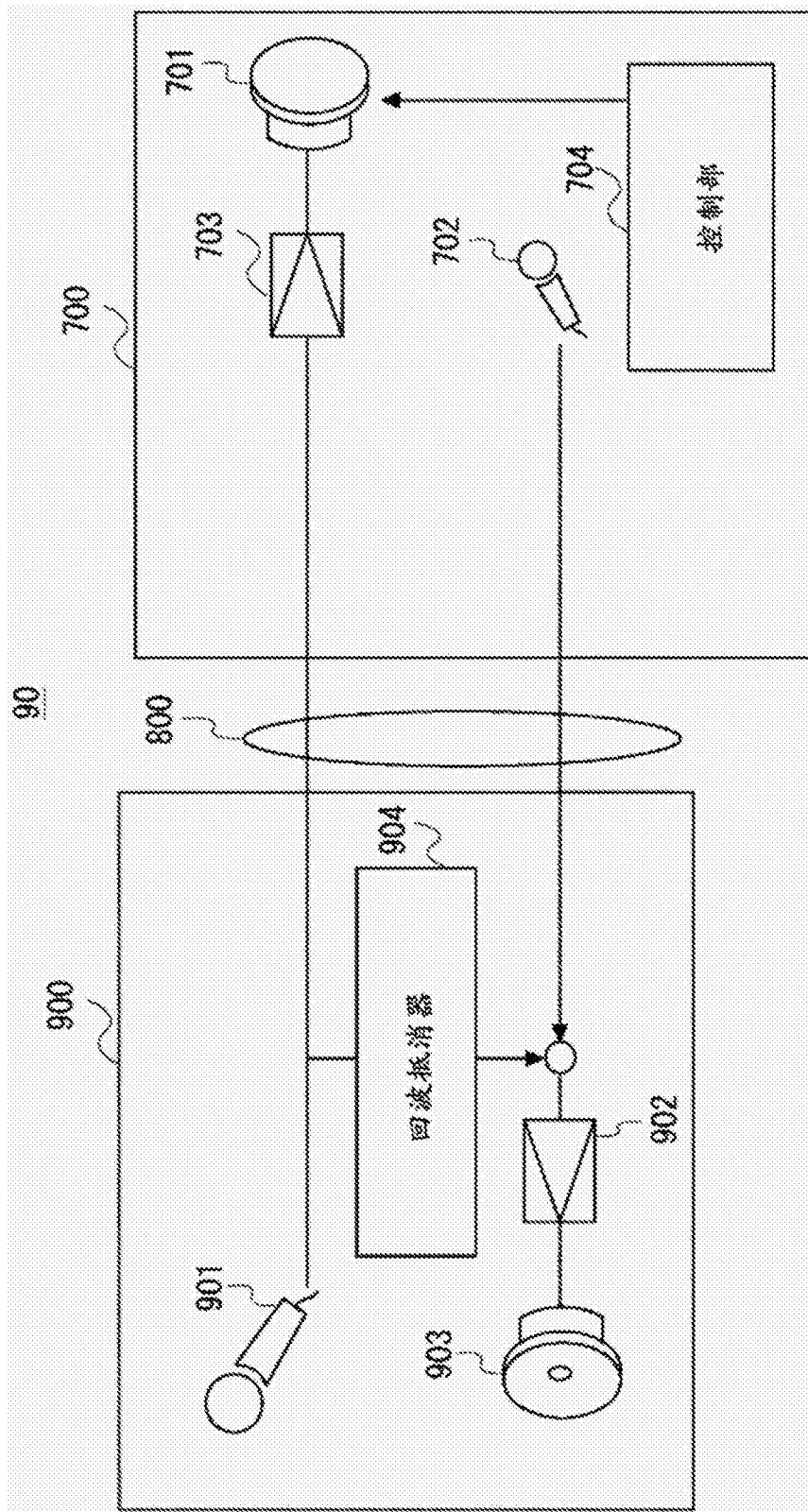


图 9

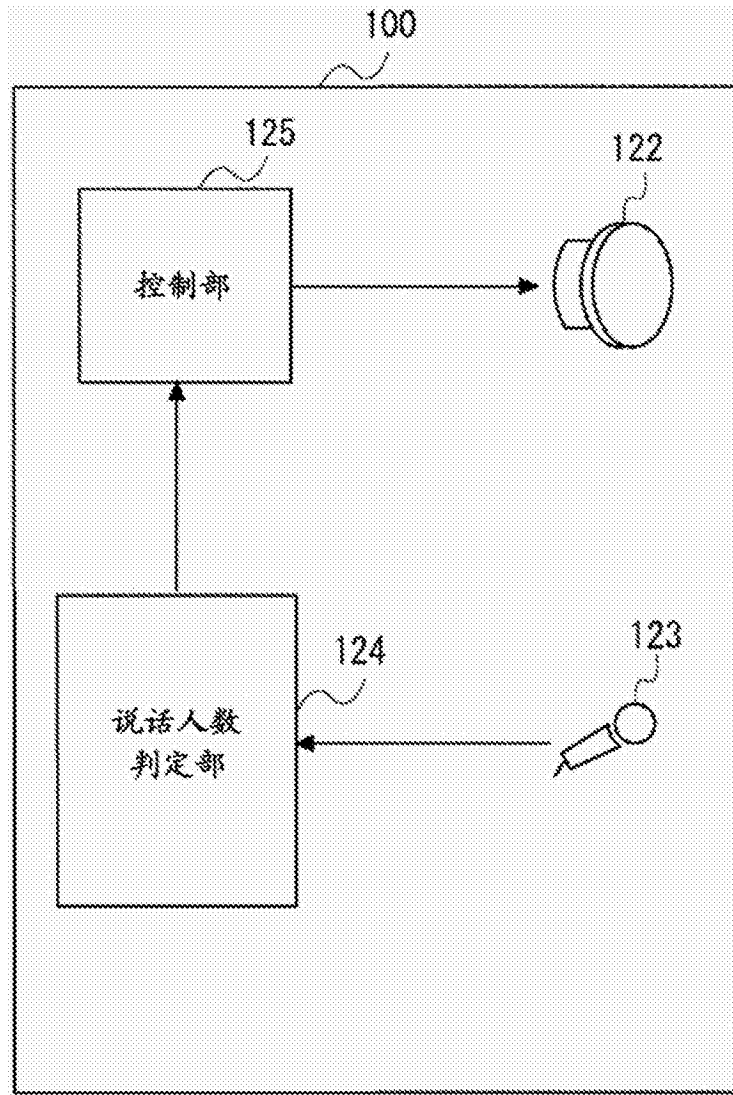


图 10

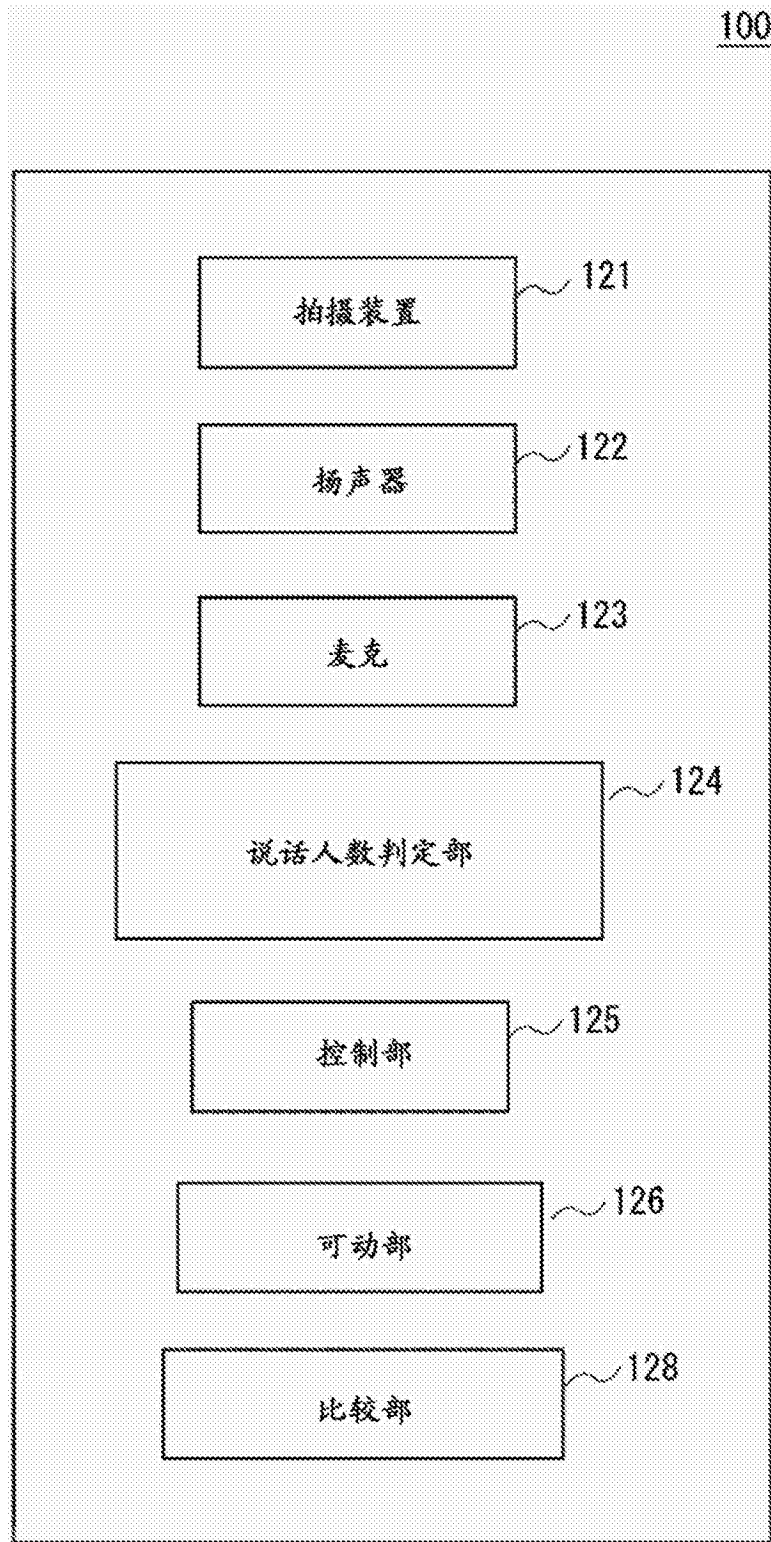


图 11

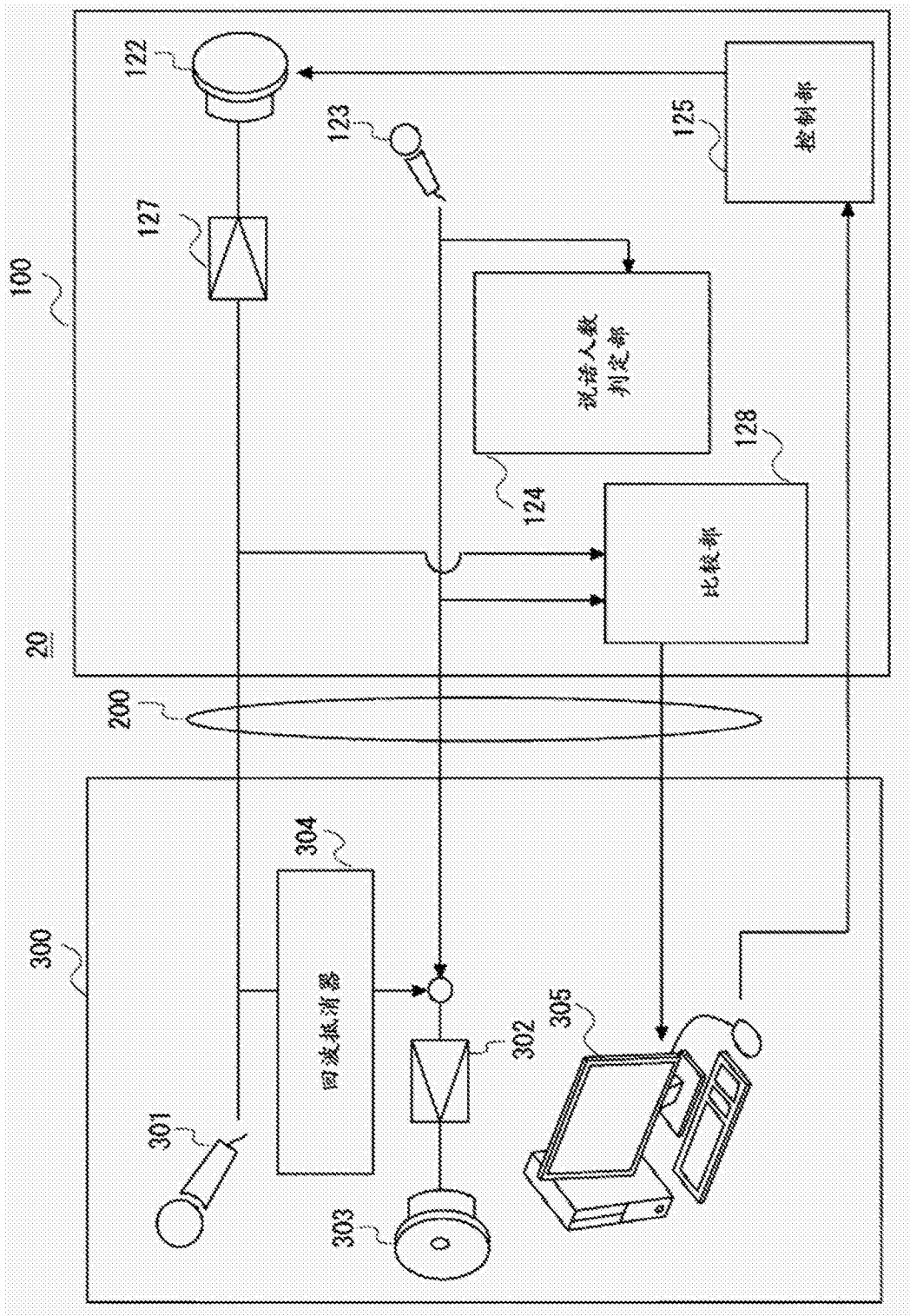


图 12

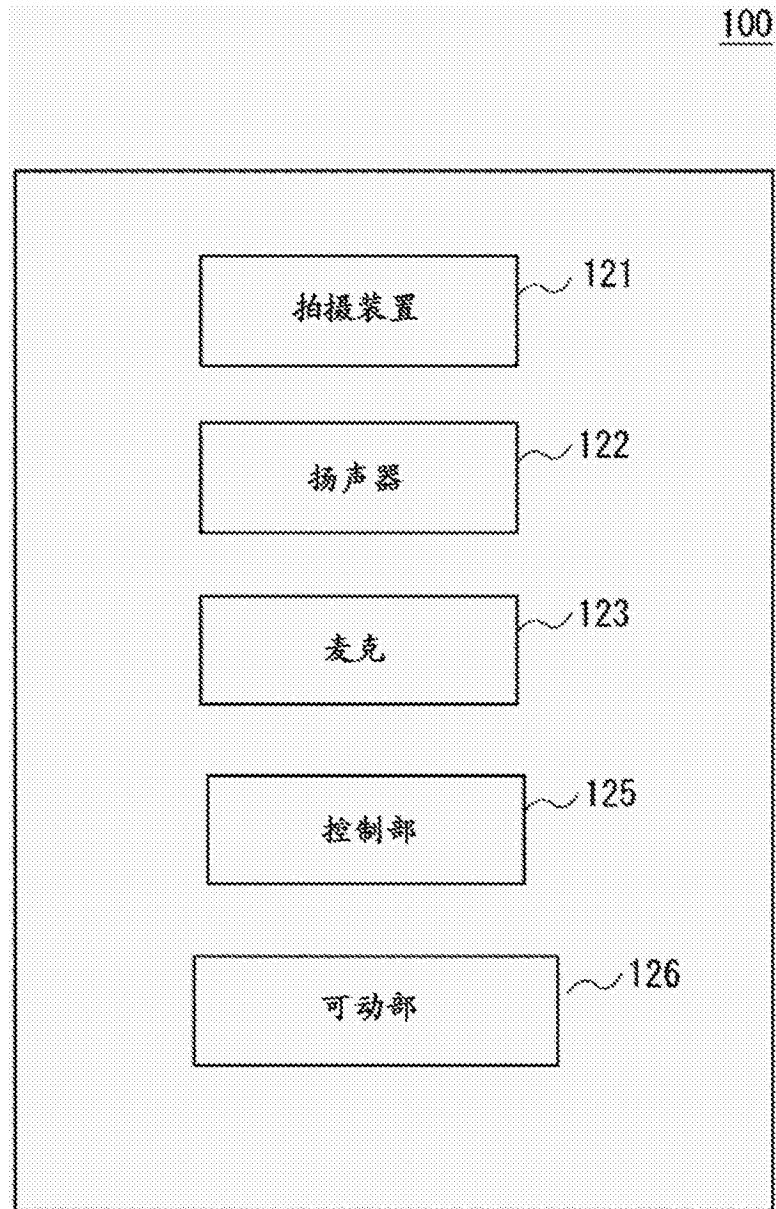


图 13





图 14