(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 108010523 B
(45)授权公告日 2023.05.09

(21)申请号 201710825776.0

(22)申请日 2017.09.14

(65)同一申请的已公布的文献号
申请公布号 CN 108010523 A

(43)申请公布日 2018.05.08

(30)优先权数据
2017-012338 2017.01.26 JP
2017-145693 2017.07.27 JP
62/416220 2016.11.02 US

(73)专利权人 松下电器(美国)知识产权公司
地址 美国加利福尼亚州

(72)发明人 西川由理 山上胜义

(74)专利代理机构 北京市中咨律师事务所 11247
专利代理师 段承恩 徐健

(51)Int.Cl.
G10L 15/22(2006.01)
G10L 15/26(2006.01)
G10L 15/30(2013.01)

(56)对比文件
CN 103714088 A,2014.04.09
EP 2680165 A1,2014.01.01
US 6871179 B1,2005.03.22

审查员 朱思韦

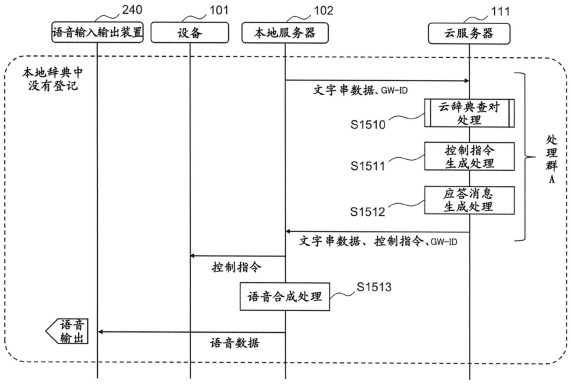
权利要求书2页 说明书28页 附图28页

(54)发明名称

信息处理方法以及记录介质

(57)摘要

本公开涉及信息处理方法以及记录介质。信息处理方法包括:(A)取得表示用户的语音的第1语音信息;(B)在判断为从所述第1语音信息生成的第1文字串信息与第1数据库的任一个文字串信息都不相符时,将所述第1文字串信息输出到服务器;(C)从所述服务器取得第1语义信息和与所述第1语义信息对应的控制指令中的至少一方;(D)基于所述第1语义信息和所述控制指令中的至少一方,进行指示使所述至少一个设备执行预定工作;(E)将从第2文字串信息生成的第2语音信息输出给扬声器。所述第1语义信息在所述服务器上的第2数据库中关联于与所述第1文字串信息相符的文字串信息。所述第2文字串信息在所述第1数据库中关联于所述第1语义信息。



CN 108010523 B

1.一种信息处理方法,是由通过与用户的对话对至少一个设备进行控制的处理器来执行的信息处理方法,包括:

取得表示从麦克风输入的所述用户的语音的第1语音信息;

参照使多个文字串信息与多个语义信息关联的第1数据库,在判断为从所述第1语音信息生成的第1文字串信息与所述第1数据库的所述多个文字串信息中的任一个都不相符的情况下,将所述第1文字串信息经由网络输出到服务器;

经由所述网络从所述服务器取得第1语义信息和与所述第1语义信息对应的控制指令中的至少一方,所述第1语义信息在所述服务器上的第2数据库中关联于与所述第1文字串信息相符的文字串信息;

基于所述第1语义信息和所述控制指令中的至少一方,进行指示使所述至少一个设备执行预定工作;

经由所述网络从所述服务器取得在所述第2数据库中与所述第1语义信息关联的一个以上的文字串信息;

通过对所述一个以上的文字串信息与所述第1数据库的所述多个文字串信息进行查对,从所述一个以上的文字串信息中确定第2文字串信息;

将从所述第2文字串信息生成的第2语音信息输出给扬声器,所述第2文字串信息在所述第1数据库中关联于所述第1语义信息。

2.一种信息处理方法,是由通过与用户的对话对至少一个设备进行控制的处理器来执行的信息处理方法,包括:

取得表示从麦克风输入的所述用户的语音的第1语音信息;

参照使多个文字串信息与多个语义信息关联的第1数据库,在判断为从所述第1语音信息生成的第1文字串信息与所述第1数据库的所述多个文字串信息中的任一个都不相符的情况下,将所述第1文字串信息经由网络输出到服务器;

经由所述网络从所述服务器取得第1语义信息和与所述第1语义信息对应的控制指令中的至少一方,所述第1语义信息在所述服务器上的第2数据库中关联于与所述第1文字串信息相符的文字串信息;

基于所述第1语义信息和所述控制指令中的至少一方,进行指示使所述至少一个设备执行预定工作;

通过对所述第1语义信息与所述第1数据库的所述多个语义信息进行查对,从所述第1数据库的所述多个文字串信息中确定第2文字串信息;

将从所述第2文字串信息生成的第2语音信息输出给扬声器,所述第2文字串信息在所述第1数据库中关联于所述第1语义信息。

3.根据权利要求1或2所述的信息处理方法,还包括:

在取得所述第1语义信息后,基于所述第1语义信息生成所述控制指令。

4.根据权利要求1或2所述的信息处理方法,还包括:

在取得所述第1语音信息后,从所述第1语音信息生成所述第1文字串信息。

5.根据权利要求1或2所述的信息处理方法,

所述网络为互联网,

所述第1数据库保存于能够不经由所述互联网而与所述至少一个设备通信的本地服务

器上。

6.根据权利要求1或2所述的信息处理方法，

在判断为所述第1文字串信息与所述第1数据库的所述多个文字串信息中的一个相符的情况下，不与所述服务器进行通信，进行指示使所述至少一个设备执行所述预定工作。

7.一种记录有程序的非瞬时性记录介质，

所述程序使权利要求1或2所述的信息处理方法由所述处理器执行。

8.一种信息处理方法，是由通过与用户的对话对至少一个设备进行控制的处理器来执行的信息处理方法，包括：

(A)取得表示从麦克风输入的所述用户的语音的第1语音信息；

(B)在所述第1语音信息包括第1短语的情况下，

(b1)在取得了所述第1语音信息后经过了第1处理时间的时间点，对所述至少一个设备输出用于使其执行预定工作的控制指令，

(b2)经由互联网从服务器取得在第2数据库中与所述第1短语关联的一个以上的文字串信息，

(b3)通过对所述一个以上的文字串信息与第1数据库的多个文字串信息进行查对，从所述一个以上的文字串信息中确定与所述第1短语不同的第2短语，所述第2短语登记于所述处理器不经由所述互联网就能够通信的所述第1数据库，所述第1短语没有登记于所述第1数据库，并且登记于所述服务器上的所述第2数据库，并且，

(b4)将用于对所述用户提示所述第2短语的第2语音信息输出到扬声器；(C)在所述第1语音信息包括所述第2短语的情况下，

(c1)在取得了所述第1语音信息后经过了第2处理时间的时间点，对所述至少一个设备输出用于使其执行所述预定工作的所述控制指令，所述第2处理时间比所述第1处理时间短。

9.根据权利要求8所述的信息处理方法，

所述第1处理时间包含所述处理器经由所述互联网与所述服务器通信的通信时间，

所述第2处理时间不包含所述通信时间。

10.一种记录有程序的非瞬时性记录介质，

所述程序使权利要求8或9所述的信息处理方法由所述处理器执行。

信息处理方法以及记录介质

技术领域

[0001] 本公开涉及信息处理方法以及记录介质。

背景技术

[0002] 近年来,能够使用语音识别进行控制的以家电为代表的设备正受到关注。这些设备存在如下问题:家电终端等本地侧装置的存储容量受限,所以能够登记的单词数少,因此用户必须记住有限的说话短语。于是,最近,可在云服务器(cloud server)上控制的语音对话正引起关注。云服务器的存储容量大,因此,能够构建具有丰富词汇的辞典,而且能够进行辞典的频繁的更新升级,所以具有能够对应于用户的各种表达方式进行语音对话这一优点。另一方面,云服务器与设备之间的通信时间往返需要500ms～数秒左右,语音对话中会产生用户能够识别程度的延迟这一点成为课题。

[0003] 例如,专利文献1公开了语音识别技术的一例。专利文献1的装置以及程序使用语音指令对与消费电子产品关联的设备进行语音控制。该装置以及程序通过从作为中心发挥功能的语音输入对应装置向终端装置发送本地侧的终端装置的辞典短缺的、与用户特有的表达对应的同义词,使终端装置的识别率提高。

[0004] 现有技术文献

[0005] 专利文献1:日本特开2014-106523号公报

发明内容

[0006] 发明所要解决的问题

[0007] 本公开提供提高语音识别反应(response)的信息处理方法以及记录介质。

[0008] 用于解决问题的技术方案

[0009] 本公开的一个技术方案涉及的信息处理方法由通过与用户的对话对至少一个设备进行控制的处理器来执行。在本技术方案中,信息处理方法包括:(A)取得表示从麦克风输入的所述用户的语音的第1语音信息;(B)参照使多个文字串信息与多个语义(意思)信息关联的第1数据库,在判断为从所述第1语音信息生成的第1文字串信息与所述第1数据库的所述多个文字串信息中的任一个都不相符的情况下,将所述第1文字串信息经由网络输出到服务器;(C)经由所述网络从所述服务器取得第1语义信息和与所述第1语义信息对应的控制指令中的至少一方,所述第1语义信息在所述服务器上的第2数据库中关联于与所述第1文字串信息相符的文字串信息;(D)基于所述第1语义信息和所述控制指令中的至少一方,进行指示使所述至少一个设备执行预定工作;(E)将从第2文字串信息生成的第2语音信息输出给扬声器,所述第2文字串信息在所述第1数据库中关联于所述第1语义信息。

[0010] 本公开的一个技术方案涉及的信息处理方法由服务器上的第2处理器来执行,所述第2处理器能够经由网络与通过与用户的对话对至少一个设备进行控制的第1处理器通信。在本技术方案中,信息处理方法包括:(A)在从第1语音信息生成的第1文字串信息与由所述第1处理器参照的第1数据库内的任一个文字串信息都不相符的情况下,经由所述网络

从所述第1处理器取得所述第1文字串信息，所述第1语音信息表示从麦克风输入的所述用户的语音；(B)将第1语义信息和与所述第1语义信息对应的控制指令中的至少一方经由所述网络输出到所述第1处理器，所述第1语义信息关联于第2数据库内的多个文字串信息中的与所述第1文字串信息相符的一个文字串信息；(C)将在所述第1数据库中关联于所述第1语义信息的第2文字串信息、和从所述第2文字串信息生成的第2语音信息中的至少一方经由所述网络输出到所述第1处理器，所述第2语音信息是被输出给扬声器的信息。

[0011] 发明效果

[0012] 根据本公开的信息处理方法，能够提高语音识别反应。

附图说明

[0013] 图1A是表示具备实施方式涉及的语音处理装置的语音对话代理系统被配置的环境的一例的图，是表示具备语音对话代理系统的信息管理系统所提供的服务的全貌的图。

[0014] 图1B是表示图1A的数据中心运营公司相当于设备制造商的例子的图。

[0015] 图1C是表示图1A的数据中心运营公司相当于设备制造商以及管理公司的双方或者某一方的例子的图。

[0016] 图2是表示实施方式涉及的语音对话代理系统的构成的概略图。

[0017] 图3是表示实施方式涉及的语音输入输出装置的硬件结构的一例的图。

[0018] 图4是表示实施方式涉及的设备的硬件结构的一例的图。

[0019] 图5是表示实施方式涉及的本地服务器的硬件结构的一例的图。

[0020] 图6是表示实施方式涉及的云服务器的硬件结构的一例的图。

[0021] 图7是表示实施方式涉及的语音输入输出装置的系统结构的一例的图。

[0022] 图8是表示实施方式涉及的设备的系统结构的一例的图。

[0023] 图9是表示实施方式涉及的本地服务器的系统结构的一例的图。

[0024] 图10是表示实施方式涉及的云服务器的系统结构的一例的图。

[0025] 图11是实施方式涉及的云辞典DB的具体例。

[0026] 图12是由实施方式涉及的语音对话代理系统进行的推荐说话内容的通信处理的时序图。

[0027] 图13是由实施方式涉及的语音对话代理系统进行的推荐说话内容的通信处理的时序图。

[0028] 图14是实施方式涉及的云服务器上的云辞典查对处理的流程图。

[0029] 图15是表示实施方式涉及的语音对话代理系统中的各种信息的传输的图。

[0030] 图16是与由变形例1涉及的语音对话代理系统进行的推荐说话内容的通信处理中的处理群A有关的时序图。

[0031] 图17是变形例1涉及的云服务器上的云辞典查对处理的流程图。

[0032] 图18是表示变形例1涉及的语音对话代理系统中的各种信息的传输的图。

[0033] 图19是变形例1涉及的本地服务器上的文字串查对处理的流程图。

[0034] 图20是与由变形例2涉及的语音对话代理系统进行的推荐说话内容的通信处理中的处理群A有关的时序图。

[0035] 图21是变形例2涉及的云服务器上的云辞典查对处理的流程图。

[0036] 图22是表示变形例2涉及的语音对话代理系统中的各种信息的传输的图。

[0037] 图23是变形例2涉及的本地服务器上的文字串查对处理的流程图。

[0038] 图24是表示实施方式以及变形例涉及的语音对话代理系统能够适用的、服务的类型1(本公司数据中心型云服务)中的信息管理系统所提供的服务的全貌的图。

[0039] 图25是表示实施方式以及变形例涉及的语音对话代理系统能够适用的、服务的类型2(IaaS利用型云服务)中的信息管理系统所提供的服务的全貌的图。

[0040] 图26是表示实施方式以及变形例涉及的语音对话代理系统能够适用的、服务的类型3(PaaS利用型云服务)中的信息管理系统所提供的服务的全貌的图。

[0041] 图27是表示实施方式以及变形例涉及的语音对话代理系统能够适用的、服务的类型4(SaaS利用型云服务)中的信息管理系统所提供的服务的全貌的图。

[0042] 标号说明

[0043] 101、101a、101b:设备

[0044] 102:本地服务器

[0045] 111:云服务器

[0046] 240:语音输入输出装置

[0047] 300:语音输入输出装置的处理电路

[0048] 301:语音输入输出装置的声音收集电路

[0049] 302:语音输入输出装置的语音输出电路

[0050] 303:语音输入输出装置的通信电路

[0051] 310:语音输入输出装置的CPU

[0052] 320:语音输入输出装置的存储器

[0053] 330:语音输入输出装置的总线

[0054] 341:语音输入输出装置的设备ID

[0055] 342:语音输入输出装置的程序

[0056] 410:设备的输入输出电路

[0057] 430:设备的CPU

[0058] 440:设备的存储器

[0059] 441:设备的设备ID

[0060] 442:设备的程序

[0061] 450:设备的通信电路

[0062] 460:设备的总线

[0063] 470:设备的处理电路

[0064] 530:本地服务器的CPU

[0065] 540:本地服务器的存储器

[0066] 541:本地服务器的网关ID

[0067] 542:本地服务器的程序

[0068] 551:本地服务器的第一通信电路

[0069] 552:本地服务器的第二通信电路

[0070] 560:本地服务器的总线

[0071] 570:本地服务器的处理电路

[0072] 580:本地服务器的声学模型DB

[0073] 581:本地服务器的语言模型DB

[0074] 582:本地服务器的语音片段DB

[0075] 583:本地服务器的韵律控制DB

[0076] 584:本地服务器的本地辞典DB

[0077] 585:本地服务器的响应生成DB

[0078] 650:云服务器的通信电路

[0079] 670:云服务器的处理电路

[0080] 671:云服务器的CPU

[0081] 672:云服务器的存储器

[0082] 680:云服务器的总线

[0083] 690:云服务器的云辞典DB

[0084] 691:云服务器的应答生成DB

[0085] 700:语音输入输出装置的声音收集部

[0086] 710:语音输入输出装置的语音检测部

[0087] 720:语音输入输出装置的语音区间截取部

[0088] 730:语音输入输出装置的通信部

[0089] 740:语音输入输出装置的语音输出部

[0090] 800:设备的通信部

[0091] 810:设备的设备控制部

[0092] 900:本地服务器的通信部

[0093] 910:本地服务器的接收数据解析部

[0094] 920:本地服务器的语音识别部

[0095] 930:本地服务器的本地辞典查对部

[0096] 940:本地服务器的应答生成部

[0097] 950:本地服务器的语音合成部

[0098] 960:本地服务器的发送数据生成部

[0099] 1000:云服务器的通信部

[0100] 1020:云服务器的云辞典查对部

[0101] 1030:云服务器的应答生成部

具体实施方式

[0102] [成为本公开的技术基础的见解]

[0103] 在如专利文献1所公开的现有技术中,本发明人发现了会产生以下的问题。上述专利文献1的装置以及程序在本地侧的装置中学习同义词。因此,存在本地侧的装置随着尽管存储容量有限,但仍然学习同义词,会导致其存储区域的规模增大这一问题。为了解决这种问题,本发明人研究了以下的改善方案。

[0104] 本公开的一个技术方案涉及的语音处理装置,具备:取得部,其取得通过语音识别

处理所获得的识别文字信息;存储部,其存储具有第1辞典中的、至少将文字信息与任务(task)信息关联的信息的第1辞典信息;查对部,其基于所述第1辞典信息,使用所述第1辞典所登记的文字信息及任务信息中的至少一方、和根据与所述第1辞典不同的第2辞典及所述识别文字信息所确定的文字信息及任务信息中的至少一方,确定与所述识别文字信息对应的文字信息及任务信息中的至少一方;以及输出部,其输出关于由所述查对部所确定的与所述识别文字信息对应的文字信息及任务信息中的至少一方的提示信息;所述提示信息包括与推荐文字信息关联的信息,所述推荐文字信息是被登记于所述第1辞典以及所述第2辞典这两方的文字信息,与所述推荐文字信息对应的任务信息对应于与所述识别文字信息对应的任务信息,所述推荐文字信息与所述识别文字信息不同。

[0105] 在上述技术方案中,输出包括与推荐文字信息关联的信息的提示信息。与推荐文字信息对应的任务信息对应于识别文字信息的任务信息。再者,推荐文字信息被登记于第1辞典以及第2辞典这两方。例如,在识别文字信息仅登记于某一方辞典的情况下,通过输出提示信息,推荐文字信息被推荐。由此,只要按照推荐文字信息来发出语音,则在与对应于从该语音识别出的文字信息的任务信息有关的处理中,无需进行在第1辞典与第2辞典之间的信息的查对。也就是说,在具有第1辞典的装置与具有第2辞典的装置之间的信息的收发降低。由此,与任务信息有关的处理速度提高。也就是说,例如,在用户说出仅登记于一方辞典的说话短语的情况下,通过将登记于另一方辞典并且进行与该说话短语同样处理的说话短语推荐给用户,从而用户通过语音来进行使用另一方辞典的设备控制时的反应提高。此外,第1辞典信息也可以是第1辞典本身。

[0106] 在上述一个技术方案涉及的语音处理装置中,例如,也可以为,所述存储部存储所述第2辞典,所述查对部在所述第2辞典中确定与所述识别文字信息对应的任务信息、和对应于与所述识别文字信息对应的任务信息并且不同于所述识别文字信息的其他文字信息,所述推荐文字信息包括所述其他文字信息,所述提示信息包括与所述识别文字信息对应的任务信息、和与所述推荐文字信息关联的信息。

[0107] 在上述技术方案中,确定并输出第2辞典中的与识别文字信息对应的任务信息、和与包括第2辞典中的不同于识别文字信息的其他文字信息的推荐文字信息关联的信息。例如,在识别文字信息没有登记于第1辞典,但登记于第2辞典的情况下,语音处理装置使用第2辞典,确定上述任务信息以及推荐文字信息。由此,上述任务信息以及推荐文字信息的确定处理能够仅在存储第2辞典的语音处理装置中实施,因此,能够提高处理速度。

[0108] 在上述一个技术方案涉及的语音处理装置中,例如,也可以为,所述其他文字信息是也登记于所述第1辞典的文字信息。

[0109] 在上述一个技术方案涉及的语音处理装置中,例如,也可以为,所述其他文字信息被确定有多个,多个所述其他文字信息的一部分是也登记于所述第1辞典的文字信息。

[0110] 在上述技术方案中,上述多个其他文字信息可以包括登记于第1辞典的文字信息以及没有登记于第1辞典的文字信息。因此,通过对上述多个其他文字信息与第1辞典进行查对,能够提取登记于第1辞典的文字信息。此外,语音处理装置只要提取任务信息对应于识别文字信息的文字信息,将提取到的文字信息作为提示信息输出即可,无需判别提取到的文字信息登记于第1辞典和第2辞典中的哪一方。由此,语音处理装置的通用性提高。

[0111] 在上述一个技术方案涉及的语音处理装置中,例如,也可以为,所述输出部包括发

送所述提示信息的通信部。

[0112] 在上述技术方案中，语音处理装置以通信方式输出提示信息。由此，语音处理装置能够对位于远处的装置输出提示信息。

[0113] 上述一个技术方案涉及的语音处理装置例如也可以为，还具备接收根据所述第2辞典及所述识别文字信息所确定的任务信息的通信部，所述第1辞典信息是所述第1辞典，所述查对部在所述第1辞典中，将与接收到的任务信息对应的文字信息确定作为所述推荐文字信息。

[0114] 在上述技术方案中，语音处理装置在即使作为根据第2辞典及识别文字信息所确定的文字信息及任务信息中的至少一方，仅能够取得与识别文字信息对应的任务信息的情况下，也能够使用所取得的任务信息，取得并输出推荐文字信息。由此，根据第2辞典及识别文字信息来确定文字信息及任务信息中的至少一方的处理变得简易。

[0115] 上述一个技术方案涉及的语音处理装置例如也可以为，还具备接收根据所述第2辞典及所述识别文字信息所确定的文字信息的通信部，所述第1辞典信息是所述第1辞典，所述查对部从接收到的文字信息中，将登记于所述第1辞典的文字信息确定作为所述推荐文字信息。

[0116] 在上述技术方案中，语音处理装置在即使作为根据第2辞典及识别文字信息所确定的文字信息及任务信息中的至少一方，仅能够取得根据第2辞典及识别文字信息所确定的文字信息的情况下，也能够使用所取得的信息，取得并输出推荐文字信息。由此，根据第2辞典及识别文字信息来确定文字信息及任务信息中的至少一方的处理变得简易。

[0117] 在上述一个技术方案涉及的语音处理装置中，例如，也可以为，所述输出部包括使提示装置提示所述提示信息的提示控制部。

[0118] 在上述技术方案中，语音处理装置能够使另外的提示装置提示提示信息，让用户认识。

[0119] 在上述一个技术方案涉及的语音处理装置中，也可以为，所述任务信息包括与所述文字信息的语义有关的语义信息及用于控制设备的工作的控制信息中的至少一方，所述语义信息与所述控制信息相关联，所述文字信息与所述语义信息及所述控制信息中的至少一方关联。

[0120] 在上述技术方案中，文字信息与语义信息及控制信息中的至少一方对应，由此，基于文字信息的控制变得顺畅。

[0121] 本公开的一个技术方案涉及的语音处理方法包括：取得通过语音识别处理所获得的识别文字信息；基于具有第1辞典中的、至少将文字信息与任务信息关联的信息的第1辞典信息，使用所述第1辞典所登记的文字信息及任务信息中的至少一方、和根据与所述第1辞典不同的第2辞典以及所述识别文字信息所确定的文字信息及任务信息中的至少一方，确定与所述识别文字信息对应的文字信息及任务信息中的至少一方，输出关于与所述识别文字信息对应的文字信息及任务信息中的至少一方的提示信息；所述提示信息包括与推荐文字信息关联的信息，所述推荐文字信息是登记于所述第1辞典以及所述第2辞典这两方的文字信息，与所述识别文字信息对应的任务信息对应于与所述推荐文字信息对应的任务信息，所述推荐文字信息与所述识别文字信息不同。

[0122] 本公开的一个技术方案涉及的程序使计算机执行：取得通过语音识别处理所获得

的识别文字信息;基于具有第1辞典中的、至少将文字信息与任务信息关联的信息的第1辞典信息,使用所述第1辞典所登记的文字信息及任务信息中的至少一方、和根据与所述第1辞典不同的第2辞典以及所述识别文字信息所确定的文字信息及任务信息中的至少一方,确定与所述识别文字信息对应的文字信息及任务信息中的至少一方,输出关于与所述识别文字信息对应的文字信息及任务信息中的至少一方的提示信息;所述提示信息包括与推荐文字信息关联的信息,所述推荐文字信息是登记于所述第1辞典以及所述第2辞典这两方的文字信息,与所述识别文字信息对应的任务信息对应于与所述推荐文字信息对应的任务信息,所述推荐文字信息与所述识别文字信息不同。

[0123] 此外,这些总括性的或者具体的技术方案可以通过系统、方法、集成电路、计算机程序或者计算机可读取的CD-ROM等记录介质来实现,也可以通过系统、方法、集成电路、计算机程序以及记录介质的任意组合来实现。

[0124] 以下,参照附图,具体地对实施方式进行说明。此外,以下说明的实施方式均表示本公开的技术的一个具体例。在以下的实施方式中表示的数值、形状、构成要素、步骤、步骤的顺序等为一例,并非旨在限定本公开。另外,对于以下的实施方式中的构成要素中的、没有记载在表示最上位概念的独立权利要求中的构成要素,作为任意的构成要素进行说明。另外,在所有的实施方式中,也可以组合各自的内容。

[0125] 此外,在本公开中,“A及B中的至少一方”意思与A和/或B相同。

[0126] [实施方式]

[0127] [1.提供的服务的全貌]

[0128] 首先,参照图1A～图1C,对具备实施方式涉及的语音处理装置的语音对话代理系统1被配置的信息管理系统所提供的服务的全貌进行说明。图1A是表示具备实施方式涉及的语音处理装置的语音对话代理系统被配置的环境的一例的图,是表示具备语音对话代理系统的信息管理系统所提供的服务的全貌的图。图1B是表示图1A的数据中心运营公司相当于设备制造商的例子的图。图1C是表示图1A的数据中心运营公司相当于设备制造商以及管理公司的双方或者某一方的例子的图。此外,语音处理装置既可以是后述的家庭网关(也称为本地服务器)102,也可以是云服务器111,还可以是包括家庭网关102以及云服务器111。

[0129] 如图1A所示,信息管理系统4000具备组4100、数据中心运营公司4110以及服务提供商4120。组4100例如是企业、团体或者家庭等,其规模不限。组4100具备包括第1设备101a和第2设备101b的多个设备101、以及家庭网关102。多个设备101例如是家电设备。多个设备101既可以包括例如智能手机、个人计算机(PC)或者电视等能够与互联网等通信网连接的设备,也可以包括例如照明、洗衣机或者冰箱等其自身不能与互联网等通信网连接的设备。多个设备101也可以包括尽管其自身不能与互联网等通信网连接但能经由家庭网关102与互联网等通信网连接的设备。另外,用户5100使用组4100内的多个设备101。

[0130] 数据中心运营公司4110具备云服务器111。云服务器111是经由互联网等通信网与各种装置协作的虚拟服务器。云服务器111主要对难以由通常的数据库管理工具等处理的海量数据(大数据)等进行管理。数据中心运营公司4110进行数据的管理、云服务器111的管理以及对进行这些管理的数据中心的运营等。关于数据中心运营公司4110所进行的工作的详细情况,将在后面进行说明。下面,说明作为通信网而使用了互联网的情况,但通信网不限定于互联网。

[0131] 在此，数据中心运营公司4110不限于仅进行数据的管理或者云服务器111的管理的公司。例如，如图1B所示，在开发或制造多个设备101之中的一个设备的设备制造商进行数据的管理或者云服务器111的管理等的情况下，设备制造商相当于数据中心运营公司4110。另外，数据中心运营公司4110不限于一个公司。例如，如图1C所示，在设备制造商以及管理公司共同或者分担进行数据的管理或者云服务器111的管理的情况下，双方或者某一方相当于数据中心运营公司4110。

[0132] 服务提供商4120具备服务器121。这里所说的服务器121，其规模不限，例如也包括个人用PC内的存储器等。另外，也有服务提供商4120不具备服务器121的情况。

[0133] 此外，在上述的信息管理系统4000中，家庭网关102不是必须的。例如，在云服务器111进行所有的数据管理的情况下等，无需家庭网关102。另外，也存在如家庭内的所有设备101都连接于互联网的情况这样，不存在其自身无法与互联网连接的设备的情况。

[0134] 接着，说明信息管理系统4000中的信息的传输。首先，组4100的第1设备101a或者第2设备101b分别将各自的日志信息发送到数据中心运营公司4110的云服务器111。云服务器111累积第1设备101a以及第2设备101b的日志信息(图1A的箭头131)。在此，日志信息是指表示多个设备101的例如运行状况以及工作日期时间等的信息。例如，日志信息可以包括电视的视听历史记录、录像机的录像预约信息、洗衣机的运行日期时间、洗涤物的量、冰箱的开闭日期时间、以及冰箱的开闭次数等，但不限于这些信息，也可以包括能够从各种设备101取得的各种信息。此外，日志信息也可以经由互联网从多个设备101本身直接提供到云服务器111。另外，日志信息也可以从多个设备101暂时累积到家庭网关102，再从家庭网关102提供到云服务器111。

[0135] 接着，数据中心运营公司4110的云服务器111将所累积的日志信息以一定的单位提供给服务提供商4120。在此，所谓一定的单位，既可以是能够对数据中心运营公司4110所累积的信息进行整理以提供给服务提供商4120的单位，也可以是服务提供商4120所要求的单位。另外，虽然设为以一定的单位提供日志信息，但是也可以不以一定的单位提供，还可以是提供的信息量根据状况而变化。日志信息根据需要保存在服务提供商4120所保有的服务器121中(图1A的箭头132)。

[0136] 而且，服务提供商4120将日志信息整理为与向用户提供的服务相适的信息，并提供给用户。被提供信息的用户既可以是使用多个设备101的用户5100，也可以是外部的用户5200。作为向用户5100、用户5200提供信息的方法，例如也可以从服务提供商4120直接向用户5100、用户5200提供信息(图1A的箭头133、134)。另外，作为向用户5100提供信息的方法，例如也可以是再次经由数据中心运营公司4110的云服务器111向用户5100提供信息的方法(图1A的箭头135、136)。另外，也可以为，数据中心运营公司4110的云服务器111将日志信息整理为与向用户提供的服务相适的信息，并提供给服务提供商4120。此外，用户5100与用户5200既可以是不同的，也可以是相同的。

[0137] [2-1.实施方式涉及的语音对话代理系统的构成]

[0138] 以下，说明实施方式涉及的语音对话代理系统1的构成。语音对话代理系统1是在用户说出仅登记于云侧辞典的说话短语的情况下，将进行同样处理的登记于本地侧辞典的说话短语推荐给用户的系统。此时，语音对话代理系统1适当地向用户推荐能够在本地侧的装置中高速应答的说话短语。由此，语音对话代理系统1提高用户进行设备控制时的反应。

[0139] 首先，关于语音对话代理系统1的构成，依次说明语音对话代理系统的构成、语音输入输出装置的硬件结构、设备的硬件结构、本地服务器的硬件结构、云服务器的硬件结构、语音输入输出装置的功能块、设备的功能块、本地服务器的功能块、以及云服务器的功能块。接下来，关于语音对话代理系统1的工作，依次说明推荐能够在终端侧也就是说本地侧高速应答的说话短语的处理的时序、以及由语音对话代理系统1进行的云辞典查对处理的流程。

[0140] 参照图2，说明实施方式涉及的语音对话代理系统1的构成。图2是表示实施方式涉及的语音对话代理系统1的构成的概略图。语音对话代理系统1包括语音输入输出装置240、多个设备101、本地服务器102、信息通信网络220和云服务器111。本地服务器102是家庭网关的一例。信息通信网络220例如是互联网，是通信网的一例。在本实施方式中，多个设备101由电视243、空调244以及冰箱245构成。另外，构成多个设备101的设备不限定于电视243、空调244以及冰箱245，可以是任意设备。语音输入输出装置240、多个设备101以及本地服务器102配置于组4100。在此，既可以是本地服务器102构成语音处理装置，也可以是云服务器111构成语音处理装置，还可以是本地服务器102以及云服务器111一起构成语音处理装置。

[0141] 在图2所示的例子中，人即用户5100位于语音对话代理系统1所配置的组4100内。另外，设用户5100为相对于语音对话代理系统1的说话者。

[0142] 语音输入输出装置240是取得组4100内的语音的声音收集部的一例，也是向组4100内输出语音的语音输出部的一例。组4100是语音输入输出装置240能够通过语音向用户提供信息的空间。语音输入输出装置240对组4100内的用户5100的语音进行识别，根据基于识别出的语音输入的用户5100的指示，由语音输入输出装置240提示语音信息，并且控制设备101。更具体而言，语音输入输出装置240显示遵从基于语音输入的用户5100的指示的内容，或回答用户5100的提问，或控制设备101。

[0143] 另外，在此，对于语音输入输出装置240、多个设备101以及本地服务器102之间的连接，能够使用有线或者无线方式的连接。对于无线方式的连接，能够应用各种无线通信。例如，既可以应用Wi-Fi(注册商标)(Wireless Fidelity)等无线局域网(LAN:Local Area Network)，也可以应用Bluetooth(注册商标)、ZigBee(注册商标)等近距离无线通信。

[0144] 另外，语音输入输出装置240、设备101以及本地服务器102中的至少一部分也可以被一体化。例如，也可以为，对语音输入输出装置240装载本地服务器102的功能，语音输入输出装置240自身作为与云服务器111通信的本地终端来发挥功能。或者，也可以为，语音输入输出装置240被装载于多个设备101的各设备、或者多个设备101中的一方。在后者的情况下，装载了语音输入输出装置240的设备101也可以控制其他设备101。或者，也可以为，语音输入输出装置240的功能和本地服务器102的功能中的至少本地服务器102的功能被装载于多个设备101的各设备、或者多个设备101中的一方。在前者的情况下，也可以为，各设备101自身作为与云服务器111通信的本地终端来发挥功能，在后者的情况下，也可以为，经由作为装载了本地服务器102的功能的本地终端的一个设备101，其他设备101与云服务器111进行通信。

[0145] 再者，从硬件结构的观点，对语音输入输出装置240、设备101、本地服务器102以及云服务器111进行说明。图3表示实施方式涉及的语音输入输出装置240的硬件结构的一例。

如图3所示，语音输入输出装置240具有处理电路300、声音收集电路301、语音输出电路302以及通信电路303。处理电路300、声音收集电路301、语音输出电路302以及通信电路303通过总线330互相连接，能够在彼此之间进行数据以及命令的授受。

[0146] 处理电路300能够通过CPU(Central Processing Unit)310和保存有设备ID341以及计算机程序342的存储器320的组合来实现。CPU310控制语音输入输出装置240的工作，但也可以还控制经由本地服务器102所连接的各设备101的工作。在该情况下，处理电路300将各设备101的控制命令经由本地服务器102进行发送，但也可以直接对各设备101进行发送。CPU310执行部署于存储器320的计算机程序342所记述的命令群。由此，CPU310能够实现各种功能。计算机程序342中记述有用于实现后述的语音输入输出装置240的工作的命令群。上述的计算机程序342也可以预先保存于作为产品的语音输入输出装置240的存储器320。或者，也可以为，计算机程序342记录于CD-ROM等记录介质并作为产品在市场上流通、亦或通过互联网等电通信线路来传输，通过记录介质或者电通信线路所取得的计算机程序342被保存于存储器320。

[0147] 或者，处理电路300也可以通过专用的硬件来实现，该专用的硬件构成为实现以下说明的工作。此外，设备ID341是对设备101唯一地赋予的标识符。设备ID341既可以由设备101的制造商独自地赋予，或者也可以为原则上在网络上唯一地分配的物理地址(所谓的MAC(Media Access Control)地址)。

[0148] 此外，在图3中，在保存有计算机程序342的存储器320中保存了设备ID341。然而这是处理电路300的构成的一例。例如，也可以为，计算机程序342保存于RAM(Random Access Memory)或者ROM(Read Only Memory)，设备ID341保存于闪存。

[0149] 声音收集电路301收集用户的语音并生成模拟语音信号，将该模拟语音信号变换为数字数据发送给总线330。

[0150] 语音输出电路302将通过总线330接收到的数字数据变换为模拟语音信号，输出该模拟语音信号。

[0151] 通信电路303是经由有线通信或者无线通信与其他设备(例如本地服务器102)进行通信的电路。虽然不进行限定，但在本实施方式中，通信电路303经由网络与其他设备进行通信，例如经由遵循以太网(注册商标)标准的网络等有线局域网来进行通信。通信电路303将由处理电路300所生成的日志信息以及ID信息发送给本地服务器102。另外，通信电路303将从本地服务器102接收到的信号通过总线330发送给处理电路300。

[0152] 除了图示的构成要素以外，语音输入输出装置240还可以包括用于实现语音输入输出装置240所要求的功能的其他构成要素。

[0153] 图4表示实施方式涉及的设备101的硬件结构的一例。图2所示的电视243、空调244以及冰箱245是设备101的一例。如图4所示，设备101具有输入输出电路410、通信电路450和处理电路470。输入输出电路410、通信电路450以及处理电路470通过总线460互相连接，能够在彼此之间进行数据以及命令的授受。

[0154] 处理电路470能够通过CPU430和保存有设备ID441以及计算机程序442的存储器440的组合来实现。CPU430控制设备101的工作。CPU430执行部署于存储器440的计算机程序442所记述的命令群，能够实现各种功能。计算机程序442中记述有用于实现设备101的工作的命令群。上述的计算机程序442也可以预先保存于作为产品的设备101的存储器440。或

者，也可以为，计算机程序442记录于CD-ROM等记录介质并作为产品在市场上流通、亦或通过互联网等电通信线路来传输，通过记录介质或者电通信线路所取得的计算机程序442被保存于存储器440。

[0155] 或者，处理电路470也可以通过专用的硬件来实现，该专用的硬件构成为实现以下说明的工作。此外，设备ID441是对设备101唯一地赋予的标识符。设备ID441既可以由设备101的制造商独自地赋予，或者也可以为原则上在网络上唯一地分配的物理地址(所谓的MAC地址)。

[0156] 此外，在图4中，在保存有计算机程序442的存储器440中保存了设备ID441。然而这是处理电路470的构成的一例。例如，也可以为，计算机程序442保存于RAM或者ROM，设备ID441保存于闪存。

[0157] 输入输出电路410输出处理电路470进行了处理后的结果。另外，输入输出电路410将被输入的模拟信号变换为数字数据发送给总线330。

[0158] 通信电路450是经由有线通信或者无线通信与其他装置(例如本地服务器102)进行通信的电路。虽然不进行限定，但在本实施方式中，通信电路450经由网络与其他装置进行通信，例如经由遵循以太网(注册商标)标准的网络等有线局域网来进行通信。通信电路450将由处理电路470所生成的日志信息以及ID信息发送给本地服务器102。另外，通信电路450将从本地服务器102接收到的信号通过总线460发送给处理电路470。

[0159] 除了图示的构成要素以外，设备101还可以包括用于实现设备101所要求的功能的其他构成要素。

[0160] 图5表示本地服务器102的硬件结构的一例。本地服务器102构成语音输入输出装置240、设备101以及信息通信网络220之间的网关。如图5所示，本地服务器102具备第一通信电路551、第二通信电路552、处理电路570、声学模型DB(数据库:DataBase)580、语言模型DB581、语音片段DB582、韵律控制DB583、本地辞典DB584和应答生成DB585来作为构成要素。这些构成要素通过总线560互相连接，能够在彼此之间进行数据以及命令的授受。

[0161] 处理电路570与声学模型DB580、语言模型DB581、语音片段DB582、韵律控制DB583、本地辞典DB584以及应答生成DB585连接，能够进行对保存于这些DB的管理信息的取得以及编辑。此外，在本实施方式中，声学模型DB580、语言模型DB581、语音片段DB582、韵律控制DB583、本地辞典DB584以及应答生成DB585是本地服务器102内部的构成要素，但也可以设置在本地服务器102的外部。在该情况下，各DB以及本地服务器102的构成要素之间的连接手段除了总线560，还可以包括互联网线路、有线或者无线局域网等通信线路。

[0162] 第一通信电路551是经由有线通信或者无线通信与其他装置(例如语音输入输出装置240以及设备101)进行通信的电路。虽然不进行限定，但在本实施方式中，第一通信电路551经由网络与其他设备进行通信，例如经由遵循以太网(注册商标)标准的网络等有线局域网来进行通信。第一通信电路551将由处理电路570所生成的日志信息以及ID信息发送给语音输入输出装置240以及设备101。另外，第一通信电路551将从语音输入输出装置240以及设备101接收到的信号通过总线560发送给处理电路570。

[0163] 第二通信电路552是经由有线通信或者无线通信与云服务器111进行通信的电路。第二通信电路552经由有线通信或者无线通信连接于通信网，进而经由通信网与云服务器111进行通信。在本实施方式中，通信网是信息通信网络220。第二通信电路552例如经由遵

循以太网(注册商标)标准的网络等有线局域网来进行通信。第二通信电路552在与云服务器111之间,收发各种信息。

[0164] 处理电路570能够通过CPU530和保存有能够唯一地辨识的网关ID(以下也称为GW-ID)541以及计算机程序542的存储器540的组合来实现。CPU530控制本地服务器102的工作,但也可以还控制语音输入输出装置240以及设备101的工作。网关ID541是对本地服务器102唯一地赋予的标识符。网关ID541既可以由本地服务器102的制造商独自地赋予,或者也可以为原则上在网络上唯一地分配的物理地址(所谓的MAC地址)。CPU530执行部署于存储器540的计算机程序542所记述的命令群,能够实现各种功能。计算机程序542中记述有用于实现本地服务器102的工作的命令群。上述的计算机程序542也可以预先保存于作为产品的本地服务器102的存储器540。或者,也可以为,计算机程序542记录于CD-ROM等记录介质并作为产品在市场上流通、亦或通过互联网等电通信线路来传输,通过记录介质或者电通信线路所取得的计算机程序542被保存于存储器540。

[0165] 或者,处理电路570也可以通过专用的硬件来实现,该专用的硬件构成为实现以下说明的工作。除了图示的构成要素以外,本地服务器102还可以包括用于实现本地服务器102所要求的功能的其他构成要素。

[0166] 此外,在图5中,在保存有计算机程序542的存储器540中保存了网关ID541。然而这是处理电路570的构成的一例。例如,也可以为,计算机程序542保存于RAM或者ROM,网关ID541保存于闪存。

[0167] 声学模型DB580登记有包括语音的波形等频率模式以及与语音对应的文字串等的各种声学模型。语言模型DB581登记有包括单词及其排列方式等的各种语言模型。语音片段DB582登记有以音素等作为单位并且表现了语音的特征的各种语音片段。韵律控制DB583登记有用于控制文字串的韵律的各种信息。本地辞典DB584将各种文字串和与文字串分别对应的语义标签关联并进行了登记。文字串由单词、文节(在日语中,由自然发音对文进行划分而得到的最小单位)等的短语等构成。语义标签指的是表示某个文字串的语义的逻辑表现。例如,在存在多个文字串的语义相同的文字串的情况下,对这些多个文字串共通设定同一语义标签。例如,语义标签将任务对象的名称、向任务对象的任务内容等表示作为关键字。例如,参照图11,示出了文字串和与文字串对应的语义标签的组合的例子。应答生成DB585将各种语义标签、和与各种语义标签对应的设备101的控制指令关联并进行了登记。应答生成DB585将与控制指令等对应的应答消息的文字串也就是说文本信息,和语义标签以及控制指令关联并进行了登记。

[0168] 图6表示云服务器111的硬件结构的一例。如图6所示,云服务器111具备通信电路650、处理电路670、云辞典DB690和应答生成DB691来作为构成要素。这些构成要素通过总线680互相连接,能够在彼此之间进行数据以及命令的授受。

[0169] 处理电路670具有CPU671和保存有程序673的存储器672。CPU671控制云服务器111的工作。CPU671执行部署于存储器672的计算机程序673所记述的命令群。由此,CPU671能够实现各种功能。计算机程序673中记述有云服务器111用于实现后述的工作的命令群。上述的计算机程序673可以记录于CD-ROM等记录介质并作为产品在市场上流通、亦或通过互联网等电通信线路来传输。具备图6所示的硬件的装置(例如PC)通过读入该计算机程序673,可以作为本实施方式的云服务器111来发挥功能。

[0170] 处理电路670与云辞典DB690和应答生成DB691连接，能够进行对保存于这些DB的管理信息的取得以及编辑。此外，在本实施方式中，云辞典DB690以及应答生成DB691是云服务器111内部的构成要素，但也可以设置在云服务器111的外部。在该情况下，各DB以及云服务器111的构成要素之间的连接手段除了总线680，还可以包括互联网线路、有线或者无线局域网等通信线路。

[0171] 通信电路650是经由有线通信或者无线通信与其他通信设备(例如本地服务器102)进行通信的电路。通信电路650经由有线通信或者无线通信连接于通信网，进而经由通信网与其他装置(例如本地服务器102)进行通信。在本实施方式中，通信网是信息通信网络220。通信电路650例如经由遵循以太网(注册商标)标准的网络等有线局域网来进行通信。

[0172] 云辞典DB690与本地辞典DB584同样地，将各种文字串和与文字串分别对应的语义标签关联并进行了登记。文字串由单词、文节等构成。云辞典DB690登记有远远多于本地辞典DB584的文字串及语义标签的组合。再者，云辞典DB690也登记有本地对应信息，该本地对应信息是所登记的文字串是否为登记于本地辞典DB584的文字串的信息。在具有多个本地服务器102的情况下，云辞典DB690也可以登记与本地服务器102各自的网关ID对应的本地对应信息。例如，参照图11，示出了文字串、与文字串对应的语义标签、和相对于文字串的本地对应信息的组合的例子。应答生成DB691与本地服务器102的应答生成DB585具有同样的构成。

[0173] 接下来，从系统结构的观点，对语音输入输出装置240、设备101、本地服务器102以及云服务器111进行说明。图7是表示语音输入输出装置240的系统结构的一例的框图。如图7所示，语音输入输出装置240具备声音收集部700、语音检测部710、语音区间截取部720、通信部730和语音输出部740。

[0174] 声音收集部700对应于图3的声音收集电路301。声音收集部700收集用户的语音并生成模拟语音信号，将所生成的模拟语音信号变换为数字数据，根据变换出的数字数据生成语音信号。

[0175] 语音检测部710以及语音区间截取部720通过图3的处理电路300来实现。执行计算机程序342的CPU310在某个时间点例如作为语音检测部710发挥功能，在不同的另一时间点作为语音区间截取部720发挥功能。此外，这两个构成要素中的至少一方也可以通过DSP(Digital Signal Processor，数字信号处理器)等专用的进行处理的硬件来实现。

[0176] 语音检测部710判定是否检测出了语音。例如，在检测出的语音的等级小于等于预定值的情况下，语音检测部710判断为没有检测出语音。语音区间截取部720从所取得的语音信号中检测语音存在的区间。例如该区间是时间区间。

[0177] 通信部730对应于图3的通信电路303。通信部730经由网络等有线通信或者无线通信与语音输入输出装置240的其他装置(例如本地服务器102)进行通信。通信部730例如经由遵循以太网(注册商标)标准的网络等有线局域网来进行通信。通信部730将语音区间截取部720检测出的语音区间的语音信号发送给其他装置。另外，通信部730将从其他装置接收到的语音信号递送给语音输出部740。

[0178] 语音输出部740对应于图3的语音输出电路302。语音输出部740将通信部730接收到的语音信号变换为模拟语音信号，输出该模拟语音信号。

[0179] 图8是表示设备101的系统结构的一例的框图。如图8所示，设备101具备通信部800

和设备控制部810。

[0180] 通信部800对应于图4的通信电路450。通信部800经由网络等有线通信或者无线通信与设备101的其他装置(例如本地服务器102)进行通信。通信部800例如经由遵循以太网(注册商标)标准的网络等有线局域网来进行通信。

[0181] 设备控制部810对应于图4的输入输出电路410以及处理电路470。设备控制部810读取通信部800接收到的控制数据,控制设备101的工作。另外,设备控制部810控制设备101的工作控制方面的处理结果的输出。例如,设备控制部810实施通信部800所接收到的控制数据的由处理电路470进行的读入以及处理、输入输出电路410的输入输出控制等。

[0182] 图9是表示本地服务器102的系统结构的一例的框图。如图9所示,本地服务器102具备通信部900、接收数据解析部910、语音识别部920、本地辞典查对部930、应答生成部940、语音合成部950和发送数据生成部960。

[0183] 通信部900对应于图5的第一通信电路551以及第二通信电路552。通信部900经由网络等有线通信或者无线通信与本地服务器102的其他装置(例如语音输入输出装置240以及设备101)进行通信。通信部900另外也经由有线通信或者无线通信与信息通信网络220等通信网连接,进而经由通信网与云服务器111进行通信。通信部900例如经由遵循以太网(注册商标)标准的网络等有线局域网来进行通信。通信部900将从其他装置以及云服务器111等接收到的数据递送给接收数据解析部910。另外,通信部900将发送数据生成部960所生成的数据发送给其他装置以及云服务器111等。

[0184] 接收数据解析部910对应于图5的处理电路570。接收数据解析部910解析通信部900所接收到的数据的类别。另外,接收数据解析部910根据对接收到的数据的类别进行解析而得到的结果,判断是在本地服务器102内部进行进一步的处理,还是应该对其他装置发送数据。在前者的情况下,接收数据解析部910将接收到的数据递送给语音识别部920等。在后者的情况下,接收数据解析部910决定接下来应该发送的装置、和应该对该装置发送的数据的组合。

[0185] 语音识别部920通过图5的处理电路570、声学模型DB580和语言模型DB581来实现。语音识别部920根据语音信号来变换为文字串数据。具体而言,语音识别部920从声学模型DB580取得预先登记的声学模型的信息,根据声学模型和语音数据的频率特性,将语音数据变换为音素数据。再者,语音识别部920从语言模型DB581取得预先登记的语言模型的信息,根据语言模型和音素数据的排列方式,将音素数据变换为特定的文字串数据。语音识别部920将变换出的文字串数据交给本地辞典查对部930。

[0186] 本地辞典查对部930通过图5的处理电路570和本地辞典DB584来实现。本地辞典查对部930根据文字串数据来变换为语义标签。所谓语义标签,具体而言,是指出成为控制对象的设备以及任务内容等的关键字。本地辞典查对部930通过对接收到的文字串数据与本地辞典DB584进行查对,提取与该文字串数据相符的语义标签。此外,在本地辞典DB584中,将单词等文字串和与文字串对应的语义标签关联并进行了保存。通过在本地辞典DB584内搜索与所接收到的文字串一致的文字串,提取与接收到的文字串相符的、也就是说相适的语义标签。

[0187] 应答生成部940通过图5的处理电路570和应答生成DB585来实现。应答生成部940对从本地辞典查对部930接收的语义标签与应答生成DB585进行查对,基于与语义标签对应

的控制指令，生成控制成为控制对象的设备101的控制信号。再者，应答生成部940基于查对结果，生成应该提供给用户5100的文本信息的文字串数据。

[0188] 语音合成部950通过图5的处理电路570、语音片段DB582和韵律控制DB583来实现。语音合成部950根据文字串数据来变换为语音信号。具体而言，语音合成部950从语音片段DB582以及韵律控制DB583分别取得预先登记的语音片段模型以及韵律控制模型的信息，根据语音片段模型、韵律控制模型以及文字串数据，将文字串数据变换为特定的语音信号。

[0189] 发送数据生成部960对应于图5的处理电路570。发送数据生成部960根据接收数据解析部910所决定的接下来应该发送的装置以及应该对该装置发送的数据的组合，生成发送数据。

[0190] 图10是表示云服务器111的系统结构的一例的框图。如图10所示，云服务器111具备通信部1000、云辞典查对部1020和应答生成部1030。

[0191] 通信部1000对应于图6的通信电路650。通信部1000经由网络等有线通信或者无线通信，连接于信息通信网络220等通信网，进而经由通信网，与其他装置(例如本地服务器102)进行通信。通信部1000例如经由遵循以太网(注册商标)标准的网络等有线局域网来进行通信。

[0192] 云辞典查对部1020通过图6的处理电路670和云辞典DB690来实现。云辞典查对部1020根据文字串数据来变换为语义标签，进而查对文字串的同义词是否登记于本地辞典DB584。文字串的同义词指的是语义标签共通的文字串。具体而言，云辞典查对部1020通过对接收到的文字串数据与云辞典DB690进行查对，提取与该文字串数据相符也就是说相适的语义标签。再者，云辞典查对部1020使用提取到的语义标签，与云辞典DB690进行查对，由此提取被给予同一语义标签的其他文字串。再者，云辞典查对部1020将提取到的文字串中的登记于本地辞典DB584的文字串输出，将该文字串数据和与该文字串数据对应的也就是说相适的语义标签交给应答生成部1030。

[0193] 应答生成部1030通过图6的处理电路670和应答生成DB691来实现。应答生成部1030在应答生成DB691中查对所接收到的语义标签，基于与语义标签对应的控制指令，生成控制成为控制对象的设备101的控制信号。再者，应答生成部1030基于查对结果，生成应该提供给用户5100的文本信息的文字串数据。

[0194] 图11是表示云辞典DB690的具体例的图。在云辞典DB690中，将单词等文字串、语义标签以及本地对应信息彼此关联并进行了保持。本地对应信息在图11中是被记载为有无登记于本地辞典DB的信息，是对于文字串及语义标签的组合的每一个，文字串是否登记于本地辞典DB584这一信息。此外，在本地辞典DB584中，将文字串及语义标签彼此关联并进行了保持。

[0195] [2-2-1.实施方式涉及的语音对话代理系统的工作]

[0196] 接下来，关于语音对话代理系统1的工作，说明推荐能够在终端侧也就是说本地服务器102高速应答的说话短语的处理的流程。图12以及图13表示由语音对话代理系统1进行的、推荐能够在本地侧高速应答的说话短语的处理的一系列的时序。该时序在用户5100通过语音开始对语音输入输出装置240进行某些指示时开始。

[0197] 用户5100通过语音对语音输入输出装置240输入指示时，在步骤S1501中，语音输入输出装置240取得用户5100的语音数据。语音输入输出装置240的通信电路303将取得的

语音数据发送给本地服务器102。本地服务器102接收这些数据。

[0198] 接下来，在步骤S1502中，本地服务器102从语音输入输出装置240接收语音数据，进行语音数据的语音识别处理。语音识别处理指的是通过本地服务器102所具有的语音识别部920识别用户的语音的处理。具体而言，本地服务器102保持有登记于声学模型DB580以及语言模型DB581的声学模型以及语言模型的信息。用户5100对语音输入输出装置240输入语音时，本地服务器102的CPU530从用户5100的语音中提取频率特性，从声学模型DB580所保持的声学模型中提取与提取到的频率特性对应的音素数据。接着，CPU530通过对提取到的音素数据的排列方式最接近于语言模型DB581所保持的语言模型的哪个文字串数据进行查对，将音素数据变换为特定的文字串数据。其结果，语音数据被变换为文字串数据。

[0199] 接下来，在步骤S1503中，本地服务器102进行文字串数据的本地辞典查对处理。本地辞典查对处理指的是通过本地服务器102所具有的本地辞典查对部930将文字串数据变换为语义标签的处理。具体而言，本地服务器102保持有登记于本地辞典DB584的辞典的信息。本地服务器102的CPU530对在步骤S1502中变换出的文字串数据与本地辞典DB584进行查对，输出与该文字串数据对应的语义标签。此外，在该文字串数据没有登记于本地辞典DB584的情况下，CPU530不将该文字串数据变换为语义标签。

[0200] 在接下来的步骤S1504中，本地服务器102判定本地辞典DB584中是否登记有与文字串数据一致的数据。在登记了的情况下(步骤S1504:是)，本地服务器102的本地辞典查对部930输出与文字串数据对应的特定的语义标签，前进至处理群B的步骤S1520。处理群B是本地辞典DB584中登记有从语音数据变换出的文字串数据的情况下的处理，如后所述，包括步骤S1520～S1522的处理。另一方面，在没有登记的情况下(步骤S1504:否)，本地服务器102的本地辞典查对部930输出表示没有与文字串数据对应的语义标签这一情况的错误。本地服务器102将文字串数据及其网关ID进行组合并发送给云服务器111，前进至处理群A的步骤S1510。处理群A是本地辞典DB584中没有登记从语音数据变换出的文字串数据的情况下的处理，如后所述，包括步骤S1510～S1512的处理。

[0201] 在处理群B的步骤S1520中，本地服务器102进行控制指令生成处理。控制指令生成处理指的是通过本地服务器102所具有的应答生成部940，根据语义标签生成控制指令的处理。具体而言，本地服务器102保持有登记于应答生成DB585的控制指令的信息。本地服务器102的CPU530对在步骤S1503中变换出的语义标签与应答生成DB585进行查对，输出与语义标签对应的控制指令，发送给对应的设备101。

[0202] 接下来，在步骤S1521中，本地服务器102进行应答消息生成处理。应答消息生成处理指的是通过本地服务器102所具有的应答生成部940生成应答消息的处理。具体而言，本地服务器102保持有登记于应答生成DB585的应答消息的信息。本地服务器102的CPU530对在步骤S1503中变换出的语义标签与应答生成DB585进行查对，输出与控制指令对应的应答消息那样的与语义标签对应的应答消息。例如，在语义标签为图11所示的“heater_on”的情况下，CPU530输出保存于应答生成DB585的应答消息“打开暖气设备”。

[0203] 进而，在步骤S1522中，本地服务器102进行语音合成处理。语音合成处理指的是本地服务器102所具有的语音合成部950将应答消息变换为语音数据的处理。具体而言，本地服务器102保持有登记于语音片段DB582的语音片段的信息、和登记于韵律控制DB583的韵律信息。本地服务器102的CPU530读入登记于语音片段DB582的语音片段的信息、和登记于

韵律控制DB的韵律信息，根据应答消息的文字串数据来变换为特定的语音数据。本地服务器102将在步骤S1522中变换出的语音数据发送给语音输入输出装置240。

[0204] 另外，如图13所示，关于处理群A，在步骤S1510中，云服务器111进行从本地服务器102接收到的文字串数据的云辞典查对处理。云辞典查对处理指的是通过云服务器111所具有的云辞典查对部1020将文字串变换为语义标签的处理。具体而言，云服务器111保持有登记于云辞典DB690的辞典的信息。云服务器111的CPU671对在步骤S1502中变换出的文字串数据与云辞典DB690进行查对，输出与该文字串数据对应的语义标签。如此，云辞典DB690中不仅登记有本地辞典DB584所登记的文字串数据，而且登记有本地辞典DB584没有登记的各种文字串数据。将在后面说明云辞典查对处理的详细情况。

[0205] 接下来，在步骤S1511中，云服务器111进行控制指令生成处理。控制指令生成处理指的是通过云服务器111所具有的应答生成部1030，根据语义标签生成控制指令的处理。具体而言，云服务器111保持有登记于应答生成DB691的控制指令的信息。云服务器111的CPU671对在步骤S1510中变换出的语义标签与应答生成DB691进行查对，输出与语义标签对应的控制指令。

[0206] 进而，在步骤S1512中，云服务器111进行应答消息生成处理。应答消息生成处理指的是通过云服务器111所具有的应答生成部1030，根据语义标签生成应答消息的处理。具体而言，云服务器111保持有登记于应答生成DB691的应答消息的信息。云服务器111的CPU671对在步骤S1510中变换出的语义标签与应答生成DB691进行查对，输出与语义标签等对应的应答消息。在步骤S1512中生成的应答消息包括后述的推荐消息，但也可以还包括如在步骤S1521中生成的那样的与控制指令对应的消息。

[0207] 云服务器111将在步骤S1511中生成的控制指令、和在步骤S1512中生成的应答消息与成为对象的本地服务器102的网关ID一起发送给该本地服务器102。本地服务器102将接收到的控制指令发送给设备101。

[0208] 接下来，本地服务器102在步骤S1513中进行语音合成处理。语音合成处理指的是本地服务器102所具有的语音合成部950将应答消息变换为语音数据的处理，与步骤S1522的处理是同样的。本地服务器102的CPU530根据应答消息的文字串数据来变换为特定的语音数据。本地服务器102将在步骤S1513中变换出的语音数据发送给语音输入输出装置240。此外，在本地服务器102从云服务器111接收的应答消息没有包括与控制指令对应的消息的情况下，本地服务器102也可以对控制指令与应答生成DB585进行查对并取得与控制指令对应的消息，进行所取得的消息的语音合成处理。

[0209] 在此，参照图14以及图15，说明步骤S1510的云辞典查对处理的详细情况。图14是步骤S1510的云辞典查对处理的流程图。图15是表示实施方式涉及的语音对话代理系统1中的各种信息的传输的图。

[0210] 在步骤S1410中，云服务器111从本地服务器102接收文字串数据。

[0211] 接下来，在步骤S1420中，云服务器111进行将文字串数据变换为语义标签的处理。具体而言，云服务器111的CPU671对文字串数据与云辞典DB690进行查对，输出与文字串数据对应的语义标签。

[0212] 进而，在步骤S1430中，云服务器111判定云辞典DB690中是否登记有被给予了与在步骤S1420中输出的语义标签相同的语义标签的其他文字串。该其他文字串是与云服务器

111从本地服务器102接收到的文字串不同的文字串。

[0213] 在步骤S1430中的判定的结果为登记了的情况下(步骤S1430:是),在步骤S1440中,云服务器111判定语义标签相同的该其他文字串中,是否有登记于本地辞典DB584的文字串。另一方面,在没有登记的情况下(步骤S1430:否),云服务器111进行对步骤S1420中的语义标签的输出,结束云辞典查对处理。

[0214] 在步骤S1440中的判定的结果为登记了的情况下(步骤S1440:是),在步骤S1450中,云服务器111将登记于本地辞典DB584的文字串的列表作为推荐对象进行输出。另一方面,在没有登记的情况下(步骤S1440:否),云服务器111进行对步骤S1420中的语义标签的输出,结束云辞典查对处理。

[0215] 例如,云服务器111在步骤S1410中接收“冷得发抖”这一文字串数据。该文字串数据是因为图12的步骤S1503的本地辞典查对处理的结果被判定为没有登记于本地服务器102的本地辞典DB584,所以被发送到云服务器111的数据。

[0216] 云服务器111在步骤S1420中对文字串“冷得发抖”与图11所示的云辞典DB690的文字串列表即“文字串”列进行查对。其结果,云服务器111将文字串“冷得发抖”变换为与其对应的语义标签<heater_on>。此时,云服务器111既可以从云辞典DB690中提取与文字串“冷得发抖”完全一致的文字串,也可以从云辞典DB690中提取与文字串“冷得发抖”为同义词的文字串,还可以从云辞典DB690中提取与文字串“冷得发抖”的一部分文字串、例如“发抖”一致的文字串。而且,云服务器111承认与所提取到的文字串对应的语义标签是文字串数据“冷得发抖”的语义标签。

[0217] 进而,云服务器111在步骤S1430中判定云辞典DB690中是否登记有被给予了语义标签<heater_on>的其他文字串。具体而言,云服务器111查对图11所示的云辞典DB690的“语义标签”列,判定为文字串“暖气设备”、“保暖”以及“太冷”被赋予了相同的语义标签<heater_on>。

[0218] 接下来,云服务器111在步骤S1440中判定本地辞典DB584中登记有文字串“暖气设备”、“保暖”以及“太冷”中的哪个文字串。云服务器111查对图11所示的云辞典DB690的“有无登记于本地辞典DB”列,判定为本地辞典DB584中也登记有文字串“暖气设备”以及“保暖”。

[0219] 之后,云服务器111在步骤S1450中将文字串“暖气设备”以及“保暖”作为推荐对象进行输出。在此,推荐对象是推荐文字信息的一例。如此,云服务器111在云辞典查对处理中,输出与从本地服务器102接收到的文字串数据对应的语义标签,输出与该语义标签对应并且登记于本地辞典DB584的文字串的列表。

[0220] 云服务器111在图13的步骤S1512的应答消息生成处理中,生成包括将文字串“暖气设备”和/或“保暖”作为说话短语进行推荐的推荐消息的应答消息。具体而言,云服务器111例如生成“若从下次起说‘暖气设备’或者‘保暖’,会更快呦”这一推荐消息。在此,推荐消息是推荐文字信息的一例。云服务器111将生成的应答消息,和与文字串“冷得发抖”的语义标签对应的控制指令<command_1>以及网关ID一起发送给本地服务器102。本地服务器102在步骤S1513的语音合成处理中,将接收到的应答消息“若从下次起说‘暖气设备’或者‘保暖’,会更快呦”变换为语音数据,发送给语音输入输出装置240。

[0221] 以上,实施方式涉及的语音对话代理系统1在用户说出仅登记于云侧的辞典的说

话短语的情况下，将能够进行同样处理的登记于本地侧的辞典的说话短语推荐给用户，由此，能够实现用户进行设备控制时的反应提高。而且，在实施方式中，在云侧生成推荐该说话短语的推荐消息。

[0222] 此外，在实施方式中，云服务器111也可以不具备应答生成DB691。在该情况下，云服务器111也可以在处理群A的处理中，输出与从本地服务器102接收的文字串对应的语义标签、和与该语义标签对应并且登记于本地辞典DB584的文字串的列表，发送给本地服务器102。本地服务器102也可以对接收到的语义标签与应答生成DB585进行查对，生成控制指令，根据接收到的文字串的列表来生成包括推荐消息的应答消息。

[0223] [2-2-2.语音对话代理系统的工作的变形例1]

[0224] 参照图16～图19，说明语音对话代理系统1的工作中的处理群A的处理的变形例1。对于本变形例，以与实施方式的不同之处为中心进行说明。此外，图16是与由变形例1涉及的语音对话代理系统1进行的推荐说话内容的通信处理中的处理群A有关的时序图。图17是变形例1涉及的云服务器111上的云辞典查对处理的流程图。图18是表示变形例1涉及的语音对话代理系统1中的各种信息的传输的图。图19是变形例1涉及的本地服务器102上的文字串查对处理的流程图。

[0225] 参照图16，在处理群A的步骤S15101中，与图13的步骤S1510的处理同样地，云服务器111进行从本地服务器102接收到的文字串数据的云辞典查对处理，输出与该文字串数据对应的语义标签。

[0226] 在此，参照图17以及图18，在本变形例的云辞典查对处理中，云服务器111仅进行图14所示的步骤S1410以及S1420的处理。具体而言，云服务器111在步骤S1410以及S1420中，将从本地服务器102接收到的文字串数据与云辞典DB690进行查对，输出与该文字串数据对应的语义标签。例如，如图18所示，云服务器111接收文字串数据“冷得发抖”，作为与其对应的语义标签而输出语义标签<heater_on>。由此，云服务器111在云辞典查对处理中，仅输出与从本地服务器102接收到的文字串数据对应的语义标签。

[0227] 返回到图16，在接着步骤S15101的步骤S1511中，云服务器111对在步骤S15101中输出的语义标签与应答生成DB691进行查对，输出与语义标签对应的控制指令。云服务器111将控制指令与成为对象的本地服务器102的网关ID一起发送给该本地服务器102。此外，云服务器111也可以除了控制指令、或者取代控制指令，将在步骤S15101中输出的语义标签发送给本地服务器102。在云服务器111不发送控制指令的情况下，本地服务器102也可以基于从云服务器111接收的语义标签，生成控制指令。

[0228] 之后，在步骤S15131中，本地服务器102进行基于控制指令的文字串查对处理。文字串查对处理是将与控制指令对应的语义标签与本地辞典DB584进行查对，并将包含于本地辞典DB584并且与控制指令对应的文字串作为推荐对象进行输出的处理。具体而言，本地服务器102的应答生成部940对控制指令与应答生成DB585进行查对，输出与控制指令对应的语义标签。进而，本地服务器102的本地辞典查对部930对被输出的语义标签与本地辞典DB584进行查对，将与语义标签对应的文字串作为推荐对象进行输出。之后，与实施方式中的由云服务器111进行的推荐消息的生成同样地，应答生成部940生成对推荐对象的文字串进行推荐的推荐消息。另外，应答生成部940也可以对控制指令与应答生成DB585进行查对，还生成与控制指令对应的消息。如此，本地服务器102生成包括推荐消息及与控制指令对应

的消息中的至少推荐消息的应答消息。

[0229] 更具体而言，参照图18以及图19，如下说明步骤S15131中的文字串查对处理。首先，在步骤S1610中，本地服务器102从云服务器111接收与语义标签对应的控制指令。例如，如图18所示，本地服务器102接收与语义标签<heater_on>对应的控制指令<command_1>。

[0230] 接下来，在步骤S1620中，本地服务器102判定本地辞典DB584中是否登记有与控制指令对应的文字串。具体而言，本地服务器102的CPU530对控制指令与应答生成DB585进行查对，输出与控制指令对应的语义标签。进而，CPU530对输出的语义标签与本地辞典DB584进行查对，判定本地辞典DB584中是否登记有与语义标签对应的文字串。

[0231] 在步骤S1620中的判定的结果为登记有文字串的情况下(步骤S1620：是)，在步骤S1630中，本地服务器102输出与语义标签对应的文字串的列表。例如，如图18所示，本地服务器102输出与控制指令<command_1>对应的文字串“暖气设备”及“保暖”中的至少一方。被输出的文字串也可以大于等于两个。如此，本地服务器102输出与控制指令对应并且登记于本地辞典DB584的文字串的列表。此外，本地服务器102也可以基于输出的文字串的列表，生成推荐消息。再者，本地服务器102也可以对控制指令与应答生成DB585进行查对，生成与控制指令对应的消息。

[0232] 在步骤S1620中的判定的结果为没有登记文字串的情况下(步骤S1620：否)，本地服务器102结束文字串查对处理。该状况可以包括控制指令没有登记于应答生成DB585的情况、和与控制指令对应的语义标签没有登记于本地辞典DB584的情况。在这种情况下，本地服务器102停止对设备101的控制，可以不生成推荐消息，或者也不生成与控制指令对应的消息。或者，本地服务器102也可以对用户的语音不适当进行提示。

[0233] 返回到图16，在接着步骤S15131的步骤S1513中，本地服务器102进行语音合成处理。本地服务器102的CPU530将应答消息的文字串变换为特定的语音数据，发送给语音输入输出装置240。

[0234] 以上，变形例1涉及的语音对话代理系统1在用户说出仅登记于云侧的辞典的说话短语的情况下，在本地侧生成推荐能够进行同样处理的登记于本地侧的辞典的说话短语的推荐消息。由此，在云服务器111中，无需进行用于生成推荐消息的处理。这种云服务器111也可以仅具有将从本地服务器102接收的文字串数据变换为控制指令并发送给本地服务器102的功能，能够应用通用的云服务器。

[0235] [2-2-3.语音对话代理系统的工作的变形例2]

[0236] 参照图20～图23，说明语音对话代理系统1的工作中的处理群A的处理的变形例2。对于本变形例，以与实施方式的不同之处为中心进行说明。此外，图20是与由变形例2涉及的语音对话代理系统1进行的推荐说话内容的通信处理中的处理群A有关的时序图。图21是变形例2涉及的云服务器111上的云辞典查对处理的流程图。图22是表示变形例2涉及的语音对话代理系统1中的各种信息的传输的图。图23是变形例2涉及的本地服务器102上的文字串查对处理的流程图。

[0237] 参照图20，在处理群A的步骤S15102中，与图13的步骤S1510的处理同样地，云服务器111进行从本地服务器102接收到的文字串数据的云辞典查对处理，输出与该文字串数据对应的语义标签。

[0238] 在此，参照图21以及图22，在本变形例的云辞典查对处理中，云服务器111进行图

14所示的步骤S1410、S1420以及S1430的处理。具体而言，云服务器111在步骤S1410以及S1420中，将从本地服务器102接收到的文字串数据与云辞典DB690进行查对，输出与该文字串数据对应的语义标签。例如，如图22所示，云服务器111接收文字串数据“冷得发抖”，输出与其对应的语义标签<heater_on>。进而，在步骤S1430中，云服务器111判定云辞典DB690中是否登记有被给予了与在步骤S1420中输出的语义标签相同的语义标签的其他文字串。

[0239] 在步骤S1430中的判定的结果是登记了的情况下(步骤S1430：是)，在步骤S14502中，云服务器111将云辞典DB690所登记的文字串的列表作为推荐对象进行输出。在没有登记的情况下(步骤S1430：否)，云服务器111进行对步骤S1420中的语义标签的输出，结束云辞典查对处理。如此，在本变形例中，不对与语义标签对应并且登记于云辞典DB690的所有文字串进行是否登记于本地辞典DB584的判定，而将该所有文字串作为推荐对象输出。例如，如图22所示，云服务器111输出与语义标签<heater_on>对应的文字串“暖气设备”、“保暖”以及“太冷”等。

[0240] 返回到图20，在接着步骤S15102的步骤S1511中，云服务器111对在步骤S15102中输出的语义标签与应答生成DB691进行查对，输出与语义标签对应的控制指令。另外，云服务器111对控制指令与应答生成DB691进行查对，输出与控制指令对应的应答消息。在步骤S1511中生成的应答消息可以包括与控制指令对应的消息，但不包括推荐消息。例如，如图22所示，云服务器111输出与语义标签<heater_on>对应的控制指令<command_1>。

[0241] 云服务器111将在步骤S15102中输出的文字串列表，和在步骤S1511中生成的控制指令与网关ID一起发送给本地服务器102。此外，云服务器111也可以除了控制指令、或者取代控制指令，将在步骤S15102中输出的语义标签发送给本地服务器102。例如，在云服务器111不发送控制指令或者不具有生成控制指令的功能的情况下，本地服务器102也可以基于从云服务器111接收的语义标签，生成控制指令。

[0242] 接下来，在步骤S15132中，本地服务器102进行基于从云服务器111接收到的文字串列表的文字串查对处理。文字串查对处理是对文字串列表所包含的文字串与本地辞典DB584进行查对，并将文字串列表以及本地辞典DB584两方所包含的文字串作为推荐对象进行输出的处理。具体而言，本地服务器102的本地辞典查对部930对文字串列表与本地辞典DB584进行查对，输出推荐对象的文字串。进而，本地服务器102的应答生成部940生成对推荐对象的文字串进行推荐的推荐消息来作为应答消息。另外，应答生成部940对从云服务器111接收到的控制指令与应答生成DB585进行查对，将与控制指令对应的消息作为应答消息进行输出。

[0243] 更具体而言，参照图22以及图23，如下说明步骤S15132中的文字串查对处理。首先，在步骤S1710中，本地服务器102从云服务器111接收文字串列表。例如，如图22所示，本地服务器102接收包括“暖气设备”、“保暖”以及“太冷”等的文字串列表。

[0244] 接下来，在步骤S1720中，本地服务器102判定文字串列表的文字串是否登记于本地辞典DB584。具体而言，本地服务器102的CPU530对文字串列表与本地辞典DB584进行查对，判定本地辞典DB584中是否登记有与文字串列表的文字串相同的文字串。

[0245] 在步骤S1720中的判定的结果为登记有相同的文字串的情况下(步骤S1720：是)，在步骤S1730中，本地服务器102输出本地辞典DB584所登记的文字串的列表。例如，如图22所示，本地服务器102从文字串“暖气设备”、“保暖”以及“太冷”中输出文字串“暖气设备”

和/或“保暖”。被输出的文字串也可以大于等于一个。进而，本地服务器102基于输出的文字串的列表，生成推荐消息。例如，生成“若从下次起说‘暖气设备’或者‘保暖’，会更快呦”这一推荐消息。另外，本地服务器102也可以对控制指令与应答生成DB585进行查对，生成与控制指令对应的消息。另一方面，在步骤S1720中的判定的结果为没有登记相同的文字串的情况下(步骤S1720:否)，本地服务器102结束文字串查对处理。在这种情况下，本地服务器102也可以停止对设备101的控制，对用户的语音不适当进行提示。

[0246] 返回到图20，在接着步骤S15132的步骤S1513中，本地服务器102进行语音合成处理。本地服务器102的CPU530根据包括推荐消息以及与控制指令对应的消息的应答消息的文字串来变换为特定的语音数据，发送给语音输入输出装置240。

[0247] 以上，变形例2涉及的语音对话代理系统1在用户说出仅登记于云侧的辞典的说话短语的情况下，在本地侧生成推荐能够进行同样处理的登记于本地侧的辞典的说话短语的推荐消息。再者，将能够与用户的说话短语进行同样处理的登记于云侧的辞典的说话短语全部发送到本地侧。在本地侧，从接收到的说话短语中，输出与登记于本地侧的辞典的说话短语相同的说话短语，并进行推荐。由此，在云服务器111中，无需进行对语义标签与从本地侧接收到的说话短语相同的说话短语、和登记于本地侧的辞典的说话短语的查对，也无需进行用于生成推荐消息的处理。这种云侧的辞典也可以不包括与本地侧的辞典有关的信息。

[0248] [3.效果等]

[0249] 作为本公开的实施方式涉及的语音处理装置的一个技术方案的云服务器111具备:作为取得部的通信部1000、作为存储部的云辞典DB690、作为查对部的云辞典查对部1020、和作为输出部的应答生成部1030。通信部1000取得通过语音识别处理所获得的识别文字信息。云辞典DB690存储具有本地辞典DB584的第1辞典中的、至少将文字信息与任务信息关联的信息的第1辞典信息。云辞典查对部1020基于第1辞典信息，使用第1辞典所登记的文字信息及任务信息中的至少一方、和根据与第1辞典不同的云辞典查对部1020的第2辞典及识别文字信息所确定的文字信息及任务信息中的至少一方，而且云辞典查对部1020确定与识别文字信息对应的文字信息及任务信息中的至少一方。应答生成部1030输出关于与识别文字信息对应的文字信息及任务信息中的至少一方的提示信息。提示信息包括与推荐文字信息关联的信息。推荐文字信息是登记于第1辞典以及第2辞典这两方的文字信息，与推荐文字信息对应的任务信息对应于与识别文字信息对应的任务信息，再者，推荐文字信息与识别文字信息不同。

[0250] 此外，第1辞典信息是与本地辞典DB584所登记的第1辞典有关的信息，包括将第1辞典的文字信息与任务信息关联的信息。例如，第1辞典信息也可以包括与云辞典DB690所登记的第2辞典和本地辞典DB584所登记的第1辞典的对应关系有关的信息。例如，如图11所示，第1辞典信息也可以包括与第2辞典的文字串及语义标签、和它们的有无向本地辞典DB584的登记之间的对应关系有关的信息。另外，第1辞典信息也可以包括第1辞典的全部内容。此外，任务信息也可以包括控制指令及语义标签中的至少一方。例如，提示信息也可以包括推荐消息、识别文字信息的任务信息以及推荐对象的文字串中的至少一个来作为与推荐文字信息关联的信息。

[0251] 在上述的构成中，输出包括与推荐文字信息关联的信息的提示信息。与推荐文字

信息对应的任务信息对应于识别文字信息的任务信息。再者，推荐文字信息登记于第1辞典以及第2辞典这两方。例如，在识别文字信息没有登记于本地辞典DB584的第1辞典，但登记于云辞典DB690的第2辞典的情况下，通过云辞典查对部1020中的查对，可确定与识别文字信息对应的文字信息及任务信息中的至少一方。从所确定出的文字信息中，选出其任务信息与识别文字信息的任务信息对应的文字信息，进而，从选出的文字信息中，选出第1辞典以及第2辞典均登记有的文字信息。该文字信息是登记于本地辞典DB584的第1辞典并且任务信息与识别文字信息对应的推荐文字信息。通过推荐这种推荐文字信息，用户今后能够发出使用登记于本地辞典DB584的文字串的指令。由此，相对于用户的指令的处理能够尽量在本地侧进行，因此处理速度会提高。也就是说，在用户说出仅登记于云侧的辞典的说话短语的情况下，通过将进行同样处理的登记于本地侧的辞典的说话短语推荐给用户，用户通过语音进行设备控制时的反应会提高。

[0252] 在实施方式涉及的语音处理装置的一个技术方案的云服务器111中，云辞典DB690存储第2辞典。云辞典查对部1020在云辞典DB690的第2辞典中确定与识别文字信息对应的任务信息、和对应于与识别文字信息对应的任务信息并且不同于识别文字信息的其他文字信息。此外，推荐文字信息包括上述其他文字信息。提示信息包括与识别文字信息对应的任务信息、和与推荐文字信息关联的信息。

[0253] 在上述的构成中，云服务器111在云辞典DB690中，确定并输出与识别文字信息对应的任务信息、和与包括识别文字信息的其他文字信息的推荐文字信息关联的信息。例如，在识别文字信息没有登记于本地辞典DB584的第1辞典，但登记于云辞典DB690的第2辞典的情况下，云服务器111使用云辞典DB690，确定上述任务信息以及推荐文字信息。由此，能够仅在云服务器111侧实施上述任务信息以及推荐文字信息的确定处理，因此能够提高处理速度。再者，本地服务器102使用从云服务器111接收的上述任务信息以及推荐文字信息，能够在本地服务器102侧实施设备101的控制以及向用户的推荐文字信息的提示。

[0254] 再者，在实施方式涉及的语音处理装置的一个技术方案的云服务器111中，用云辞典DB690的第2辞典确定的其他文字信息是也登记于本地辞典DB584的第1辞典的文字信息。在上述的构成中，上述其他文字信息是登记于云辞典DB690的第2辞典以及本地辞典DB584的第1辞典这两方的文字信息。

[0255] 另外，在变形例2涉及的语音处理装置的一个技术方案的云服务器111中，用云辞典DB690的第2辞典确定的其他文字信息被确定有多个，多个其他文字信息的一部分是也登记于本地辞典DB584的第1辞典的文字信息。在上述的构成中，上述多个其他文字信息可以包括登记于本地辞典DB584的第1辞典的文字信息以及没有登记于第1辞典的文字信息。例如，本地服务器102在从云服务器111接收到上述多个其他文字信息时，通过对上述多个其他文字信息与本地辞典DB584的第1辞典进行查对，能够提取登记于本地辞典DB584的文字信息。在该情况下，云服务器111只要提取任务信息对应于识别文字信息的文字信息，将提取到的文字信息作为提示信息输出即可，无需判别提取到的文字信息是否登记于云辞典DB690的第2辞典及本地辞典DB584的第1辞典这两方。由此，能够使用通用的云服务器111。

[0256] 在变形例1涉及的语音处理装置的一个技术方案的云服务器111中，云辞典查对部1020在云辞典DB690的第2辞典中，确定与识别文字信息对应的任务信息，提示信息包括由云辞典查对部1020确定的任务信息作为与推荐文字信息关联的信息。在上述的构成中，云

服务器111输出在云辞典DB690中确定的与识别文字信息对应的任务信息即可，无需提取任务信息对应于识别文字信息的文字信息等。由此，能够使用通用的云服务器111。

[0257] 实施方式涉及的语音处理装置的一个技术方案的云服务器111包括发送提示信息的通信部1000作为输出部。在上述的构成中，云服务器111以通信方式发送提示信息。由此，云服务器111可以配置在远离本地服务器102的位置。本地服务器102能够不受云服务器111的影响，配置于各种设施。

[0258] 作为本公开的实施方式涉及的语音处理装置的另一个技术方案的本地服务器102具备：作为取得部的语音识别部920、作为存储部的本地辞典DB584、作为查对部的本地辞典查对部930、和作为输出部的应答生成部940以及语音合成部950。语音识别部920取得通过语音识别处理所获得的识别文字信息。本地辞典DB584存储具有本地辞典DB584的第1辞典中的、至少将文字信息与任务信息关联的信息的第1辞典信息。本地辞典查对部930基于第1辞典信息，使用第1辞典所登记的文字信息及任务信息中的至少一方、和根据与第1辞典不同的云辞典DB690的第2辞典及识别文字信息所确定的文字信息及任务信息中的至少一方，而且本地辞典查对部930确定与识别文字信息对应的文字信息及任务信息中的至少一方。应答生成部940以及语音合成部950输出关于与识别文字信息对应的文字信息及任务信息中的至少一方的提示信息。提示信息包括与推荐文字信息关联的信息。推荐文字信息是被登记于第1辞典以及第2辞典这两方的文字信息，与推荐文字信息对应的任务信息对应于与识别文字信息对应的任务信息，再者，推荐文字信息与识别文字信息不同。此外，第1辞典信息也可以是本地辞典DB584所登记的第1辞典。任务信息可以包括控制指令及语义标签中的至少一方。例如，提示信息可以包括包含推荐信息的应答信息来作为与推荐文字信息关联的信息。

[0259] 在上述的构成中，输出包括与推荐文字信息关联的信息的提示信息。与推荐文字信息对应的任务信息对应于识别文字信息的任务信息。再者，推荐文字信息被登记于第1辞典以及第2辞典这两方。例如，在识别文字信息没有登记于本地辞典DB584的第1辞典，但登记于云辞典DB690的第2辞典的情况下，本地服务器102输出包括与推荐文字信息关联的信息的提示信息。这种推荐文字信息是虽然与识别文字信息不同，但任务信息与识别文字信息对应并且登记于第1辞典以及第2辞典这两方的文字信息，也就是说，是登记于本地辞典DB584并且任务信息与识别文字信息对应的文字信息。通过推荐这种推荐文字信息，用户今后能够发出使用登记于本地辞典DB584的文字串的指令。由此，相对于用户的指令的处理能够尽量在本地侧进行，因此处理速度会提高。

[0260] 在实施方式涉及的语音处理装置的另一个技术方案的本地服务器102中，本地辞典查对部930在本地辞典DB584的第1辞典中，确定与识别文字信息对应的任务信息。在上述的构成中，本地服务器102通过确定与识别文字信息对应的任务信息，能够进行对连接于本地服务器102的设备101的控制。

[0261] 变形例1涉及的语音处理装置的另一个技术方案的本地服务器102还具备通信部900，通信部900接收根据云辞典DB690的第2辞典以及识别文字信息所确定的任务信息。第1辞典信息是本地辞典DB584的第1辞典。本地辞典查对部930在本地辞典DB584的第1辞典中，将与接收到的任务信息对应的文字信息确定作为推荐文字信息。在上述的构成中，本地服务器102例如即使在仅能够从云服务器111取得与识别文字信息对应的任务信息的情况下，

也能够使用取得的任务信息，取得并输出推荐文字信息。由此，云服务器111将与识别文字信息对应的任务信息作为推荐文字信息输出即可，无需判别与该任务信息对应的文字信息是否登记于云辞典DB690的第2辞典以及本地辞典DB584的第1辞典这两方。由此，能够使用通用的云服务器111。

[0262] 变形例2涉及的语音处理装置的另一个技术方案的本地服务器102还具备通信部900，通信部900接收根据云辞典DB690的第2辞典以及识别文字信息所确定的文字信息。第1辞典信息是本地辞典DB584的第1辞典。本地辞典查对部930从接收到的文字信息中，将在本地辞典DB584的第1辞典中所登记的文字信息确定作为推荐文字信息。此外，所接收的文字信息可以是包括一个以上的文字串的文字信息。在上述的构成中，例如，云服务器111输出推荐文字信息即可，无需判别推荐文字信息是否登记于云辞典DB690的第2辞典以及本地辞典DB584的第1辞典这两方。由此，能够使用通用的云服务器111。

[0263] 实施方式涉及的语音处理装置的另一个技术方案的本地服务器102包括作为使另外的提示装置提示提示信息的提示控制部的发送数据生成部960。在上述的构成中，本地服务器102能够使设备101等另外的装置提示例如基于从云服务器111接收的信息的提示信息，让用户认识。

[0264] 实施方式的又一个技术方案涉及的语音处理装置具备彼此之间收发信息的作为本地装置的本地服务器102以及作为云装置的云服务器111。本地服务器102包括：语音识别部920，其取得通过语音识别处理所获得的识别文字信息；作为第1存储部的本地辞典DB584，其存储将文字信息与任务信息关联的第1辞典；作为第1查对部的本地辞典查对部930；和作为第1输出部的应答生成部940以及语音合成部950。云服务器111包括：作为第2存储部的云辞典DB690，其存储将文字信息与任务信息关联的第2辞典；作为第2查对部的云辞典查对部1020；和作为第2输出部的应答生成部1030。云辞典查对部1020对本地辞典DB584的第1辞典所登记的文字信息及任务信息中的至少一方、与根据云辞典DB690的第2辞典以及识别文字信息所确定的文字信息及任务信息中的至少一方进行查对，确定与识别文字信息对应的文字信息以及任务信息中的至少一方。应答生成部1030将关于与识别文字信息对应的文字信息及任务信息中的至少一方的提示信息输出给本地服务器102。此外，提示信息包括与推荐文字信息关联的信息。推荐文字信息是登记于第1辞典以及第2辞典这两方的文字信息，与推荐文字信息对应的任务信息对应于与识别文字信息对应的任务信息，推荐文字信息与识别文字信息不同。本地辞典查对部930对从云服务器111接收的提示信息、与第1辞典所登记的文字信息及任务信息中的至少一方进行查对。应答生成部940以及语音合成部950将与推荐文字信息关联的信息作为语音等的消息进行输出。

[0265] 根据上述的构成，能获得与由实施方式涉及的语音处理装置的一个技术方案的云服务器111以及本地服务器102所实现的效果同样的效果。尤其是，在用户说出仅登记于云侧的云辞典DB690的说话短语的情况下，通过将进行同样处理的登记于本地侧的本地辞典DB584的说话短语推荐给用户，从而用户通过语音进行设备控制时的反应会提高。

[0266] 在实施方式以及变形例涉及的语音处理装置的各种技术方案的云服务器111以及本地服务器102中，任务信息包括与文字信息的语义有关的语义信息及用于控制设备的工作的控制信息中的至少一方，语义信息与控制信息相关联，文字信息与语义信息及控制信息中的至少一方关联。此外，也可以对语义相似的同义的文字信息赋予共通的语义信息。例

如,语义信息可以是语义标签,控制信息也可以是控制指令。根据上述的构成,文字信息与语义信息及控制信息中的至少一方对应,由此,基于文字信息的控制变得顺畅。另外,语义信息对于语义相似的文字信息是共通的,再者,控制信息与共通化的语义信息对应。由此,与语义相似的文字信息有关的任务信息被统一。因此,任务信息的变化减少,由此,云服务器111以及本地服务器102中的基于任务信息的处理的速度提高。

[0267] 实施方式的一个技术方案涉及的语音处理方法包括:取得通过语音识别处理所获得的识别文字信息;基于具有第1辞典中的、至少将文字信息与任务信息关联的信息的第1辞典信息,使用第1辞典所登记的文字信息及任务信息中的至少一方、和根据与第1辞典不同的第2辞典以及识别文字信息所确定的文字信息及任务信息中的至少一方,确定与识别文字信息对应的文字信息及任务信息中的至少一方,输出关于与识别文字信息对应的文字信息及任务信息中的至少一方的提示信息。此外,提示信息包括与推荐文字信息关联的信息,推荐文字信息是登记于第1辞典以及第2辞典这两方的文字信息,与识别文字信息对应的任务信息对应于与推荐文字信息对应的任务信息,推荐文字信息与识别文字信息不同。

[0268] 根据上述的语音处理方法,能获得与由实施方式涉及的语音处理装置实现的效果同样的效果。此外,上述方法也可以通过MPU(Micro Processing Unit)、CPU、处理器、LSI(Large Scale Integration:大规模集成电路)等电路、IC卡(Integrated Circuit Card)或者单体的模块等来实现。

[0269] 另外,实施方式以及变形例中的处理也可以通过软件程序或者由软件程序而成的数字信号来实现。例如,实施方式中的处理通过如下的程序来实现。

[0270] 也就是说,程序是使计算机执行以下功能的程序:取得通过语音识别处理所获得的识别文字信息;基于具有第1辞典中的、至少将文字信息与任务信息关联的信息的第1辞典信息,使用第1辞典所登记的文字信息及任务信息中的至少一方、和根据与第1辞典不同的第2辞典以及识别文字信息所确定的文字信息及任务信息中的至少一方,确定与识别文字信息对应的文字信息及任务信息中的至少一方,输出关于与识别文字信息对应的文字信息及任务信息中的至少一方的提示信息。此外,提示信息包括与推荐文字信息关联的信息,推荐文字信息是登记于第1辞典以及第2辞典两方的文字信息,与识别文字信息对应的任务信息对应于与推荐文字信息对应的任务信息,再者,推荐文字信息与识别文字信息不同。

[0271] [其他]

[0272] 以上,作为在本申请中公开的技术的例示对实施方式以及变形例涉及的语音处理装置等进行了说明,但本公开不限定于实施方式以及变形例。本公开中的技术也能够应用于适当进行了变更、替换、附加、省略等的实施方式的变形例或者其他实施方式。另外,也可以将实施方式以及变形例中说明的各构成要素进行组合,作为新的实施方式或者变形例。

[0273] 如上所述,本公开的总括性的或者具体的技术方案也可以通过系统、方法、集成电路、计算机程序或者计算机可读取的CD-ROM等记录介质来实现。另外,本公开的总括性的或者具体的技术方案也可以通过系统、方法、集成电路、计算机程序以及记录介质的任意组合来实现。

[0274] 例如,上述实施方式以及变形例涉及的语音处理装置所包含的各处理部典型地被实现为作为集成电路的LSI。它们可以个别地单片化,或者包括一部分或全部地单片化。

[0275] 另外,集成电路化不限于LSI,也可以通过专用电路或通用处理器实现。也可以利

用能够在LSI制造后编程的FPGA(Field Programmable Gate Array:现场可编程门阵列)、能够重构LSI内部的电路单元的连接或设定的可重构处理器。

[0276] 此外,在上述实施方式以及变形例中,各构成要素既可以用专用的硬件构成,也可以通过执行适合于各构成要素的软件程序来实现。各构成要素也可以通过CPU或者处理器等程序执行部将记录于硬盘或者半导体存储器等记录介质中的软件程序读出并执行来实现。

[0277] 再者,本公开的技术可以是上述程序,或者是记录有上述程序的非瞬时的计算机可读取的记录介质。另外,毫无疑问能够经由互联网等传输介质来使上述程序流通。

[0278] 另外,在上述使用的序数、数量等数字全部是为了具体地对本公开的技术进行说明而示例的,本公开不限于所示例的数字。另外,构成要素间的连接关系是为了具体地对本公开的技术进行说明而示例的,实现本公开的功能的连接关系不限定于此。

[0279] 另外,框图中的功能块的分割为一例,可以将多个功能块作为一个功能块来实现,将一个功能块分割为多个,或将一部分功能转移到其他功能块。另外,也可以是单一的硬件或软件以并行或时分方式处理具有类似功能的多个功能块的功能。

[0280] 以上,基于实施方式以及变形例说明了一个技术方案涉及的语音处理装置等,但本公开不限定于该实施方式以及变形例。只要不偏离本公开的宗旨,将本领域技术人员想到的各种变形应用于实施方式而得到的方式、和将不同的实施方式中的构成要素组合而构建的方式也可以包含在一个技术方案的范围内。

[0281] 此外,只要与语音对话代理系统和用户的对话有关,就能够适用本公开。例如,在用户使用语音对话代理系统使家电设备等工作的情况下是有效的。例如,假定在用户使对应于语音操作的微波炉或者烤箱工作时,进行了“加热”这一指示的情况。此时,语音对话代理系统能够向用户反问“加热几分钟?”或者“加热为多少度?”等具体的指示。能够对此进行回话的用户(针对反问,代理系统受理指示的用户)仅为当初指示了“加热”的用户。

[0282] 除此之外,本公开也能够适用于针对用户的抽象指示,语音对话代理系统反问具体的内容的工作。另外,语音对话代理系统向用户反问的内容也可以是工作执行的确认等。

[0283] 此外,在上述技术方案中,来自用户的语音的输入也可以通过系统或者各家电设备所具备的麦克风来进行。另外,从语音对话代理系统对用户的反问也可以从系统或者各家电设备所具备的扬声器等传达给用户。

[0284] 在本公开中,“预定工作”例如也可以是经由扬声器对用户输出语音的工作。即,在本公开中,成为控制对象的“设备”也可以是语音输入输出装置(例如扬声器)。

[0285] 在本公开中,“处理器”、“麦克风”和/或“扬声器”例如也可以内置于成为控制对象的“设备”。

[0286] 此外,在上述技术方案中所说明的技术,例如能够在以下的云服务的类型中实现。但是,实现在上述技术方案中所说明的技术的云服务的类型不限于这些。

[0287] 以下,依次对利用了服务的类型1(本公司数据中心型云服务)的信息管理系统所提供的服务的全貌、利用了服务的类型2(IaaS利用型云服务)的信息管理系统所提供的服务的全貌、利用了服务的类型3(PaaS利用型云服务)的信息管理系统所提供的服务的全貌、利用了服务的类型4(SaaS利用型云服务)的信息管理系统所提供的服务的全貌进行说明。

[0288] [服务的类型1:本公司数据中心型云服务]

[0289] 图24是表示实施方式以及变形例涉及的语音对话代理系统能够适用的、服务的类型1(本公司数据中心型云服务)中的信息管理系统所提供的服务的全貌的图。如图24所示，在本类型中，服务提供商4120从组4100取得信息，向用户提供服务。在本类型中，服务提供商4120具有数据中心运营公司的功能。即，服务提供商4120保有管理大数据(big data)的云服务器111。因此，不存在数据中心运营公司。

[0290] 在本类型中，服务提供商4120对数据中心(云服务器)4203进行运营及管理。另外，服务提供商4120管理操作系统(OS)4202以及应用4201。服务提供商4120使用服务提供商4120所管理的OS4202以及应用4201来提供服务(箭头204)。

[0291] [服务的类型2:IaaS利用型云服务]

[0292] 图25是表示实施方式以及变形例涉及的语音对话代理系统能够适用的、服务的类型2(IaaS利用型云服务)中的信息管理系统所提供的服务的全貌的图。在此，IaaS是基础设施即服务(Infrastructure as a Service)的简称，是将用于构建及运行计算机系统的基础本身作为经由互联网的服务来提供的云服务提供模型。

[0293] 如图25所示，在本类型中，数据中心运营公司4110对数据中心(云服务器)4203进行运营及管理。另外，服务提供商4120管理OS4202以及应用4201。服务提供商4120使用服务提供商4120所管理的OS4202以及应用4201来提供服务(箭头204)。

[0294] [服务的类型3:PaaS利用型云服务]

[0295] 图26是表示实施方式以及变形例涉及的语音对话代理系统能够适用的、服务的类型3(PaaS利用型云服务)中的信息管理系统所提供的服务的全貌的图。在此，PaaS是平台即服务(Platform as a Service)的简称，是将成为用于构建及运行软件的根基的平台作为经由互联网的服务来提供的云服务提供模型。

[0296] 如图26所示，在本类型中，数据中心运营公司4110管理OS4202，对数据中心(云服务器)4203进行运营及管理。另外，服务提供商4120管理应用4201。服务提供商4120使用数据中心运营公司4110所管理的OS4202以及服务提供商4120所管理的应用4201来提供服务(箭头204)。

[0297] [服务的类型4:SaaS利用型云服务]

[0298] 图27是表示实施方式以及变形例涉及的语音对话代理系统能够适用的、服务的类型4(SaaS利用型云服务)中的信息管理系统所提供的服务的全貌的图。在此，SaaS是软件即服务(Software as a Service)的简称。SaaS利用型云服务例如是具有下述功能的云服务提供模型，即未保有数据中心(云服务器)的公司或个人等利用者能够经由互联网等网络使用保有数据中心(云服务器)的平台提供商提供的应用。

[0299] 如图27所示，在本类型中，数据中心运营公司4110管理应用4201，管理OS4202，对数据中心(云服务器)4203进行运营及管理。另外，服务提供商4120使用数据中心运营公司4110所管理的OS4202以及应用4201来提供服务(箭头204)。

[0300] 以上，不论在哪一种云服务的类型中，都是服务提供商4120提供服务。另外，例如，服务提供商或者数据中心运营公司既可以自行开发OS、应用或者大数据的数据库等，另外也可以向第三方订购。

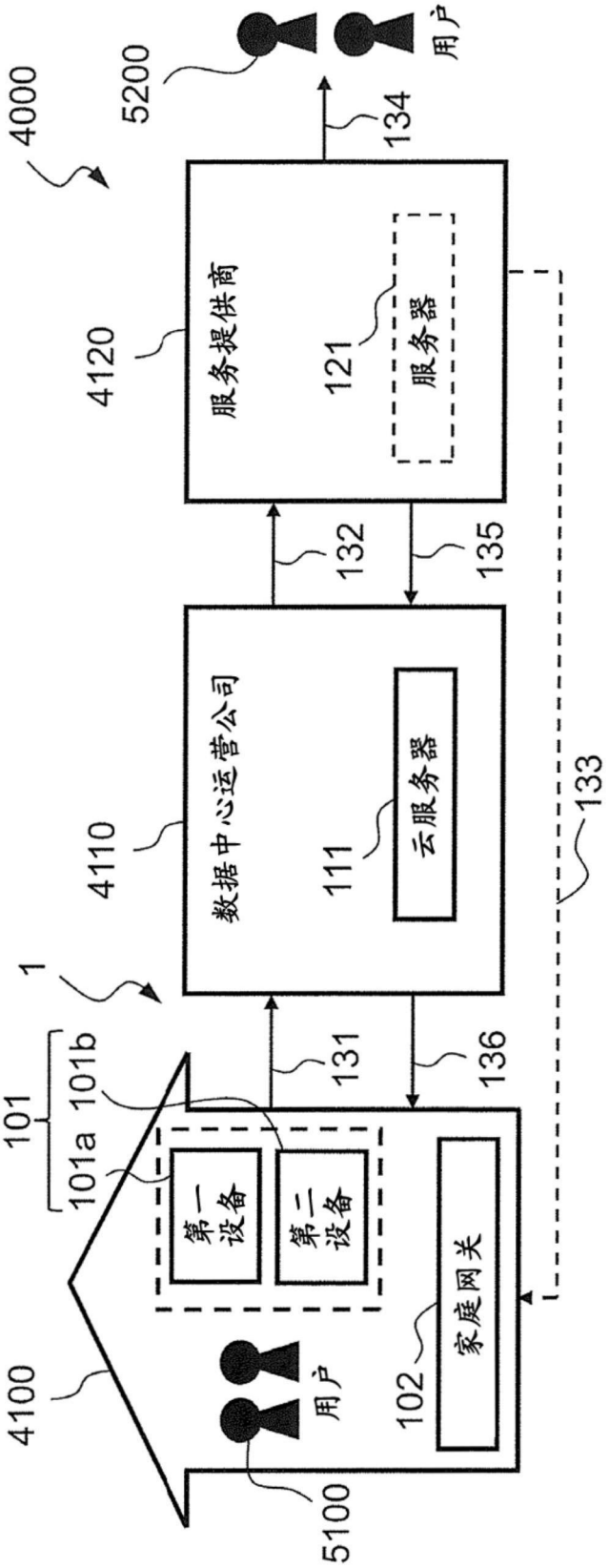

4000
5200
用户
134
4120
服务提供商
121
服务器
132
135
4110
数据中心运营公司
111
云服务器
133
1
131
136
101
101a
101b
第一设备
第二设备
4100
用户
102
家庭网关
5100

图1A

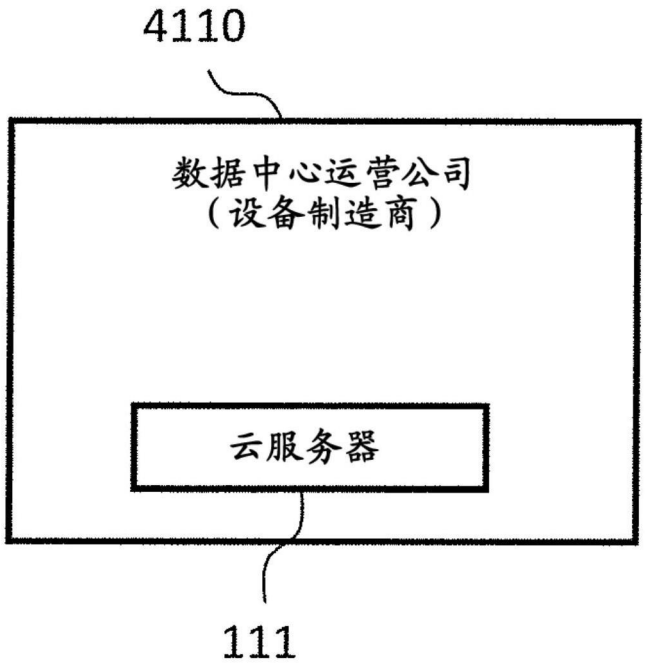

4110
数据中心运营公司
（设备制造商）
云服务器
111

图1B

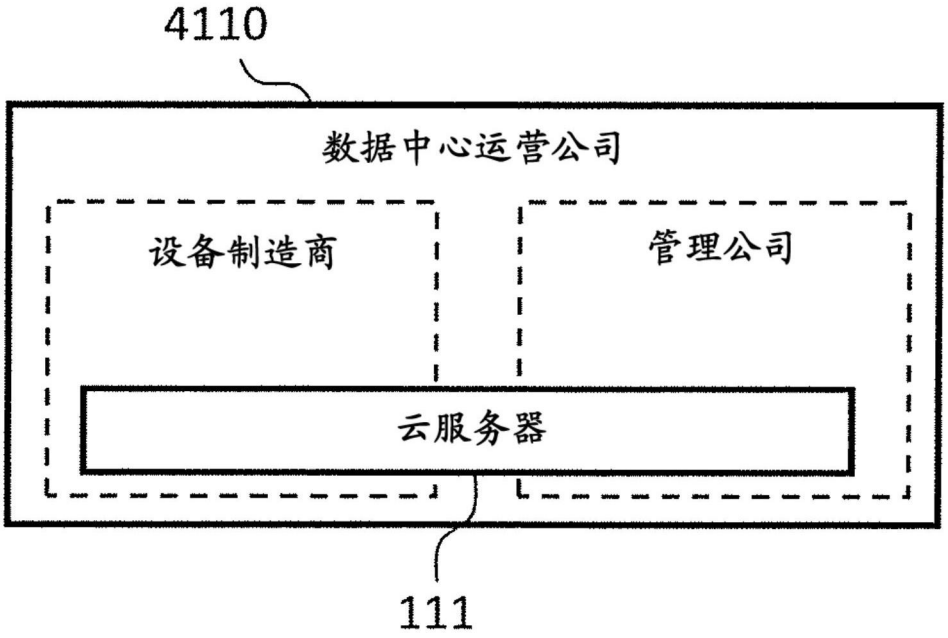

4110
数据中心运营公司
设备制造商
管理公司
云服务器
111

图1C

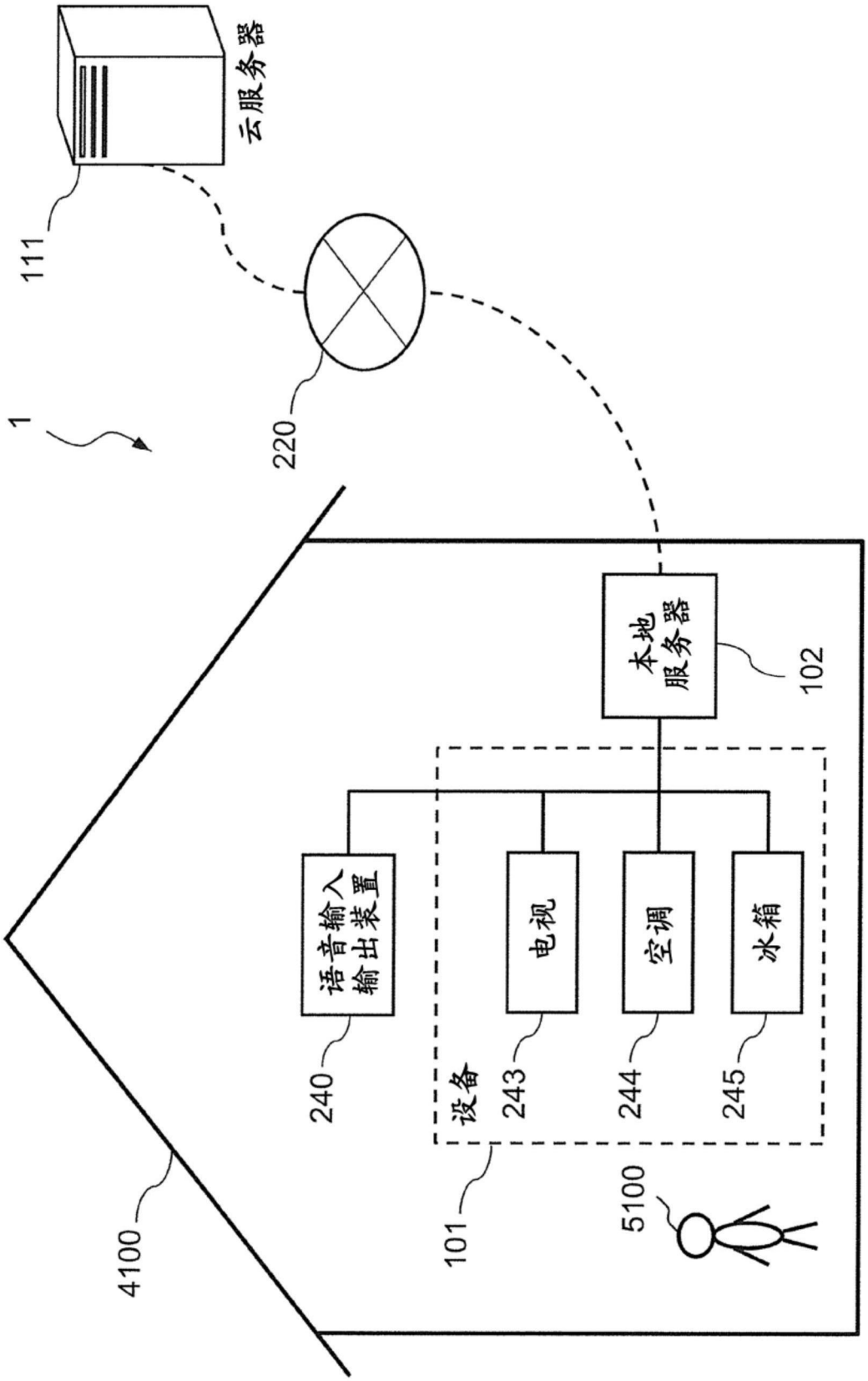

云服务器
111
1
220
本地服务器
102
语音输入输出装置
240
电视
243
空调
244
冰箱
245
设备
101
5100
4100

图2

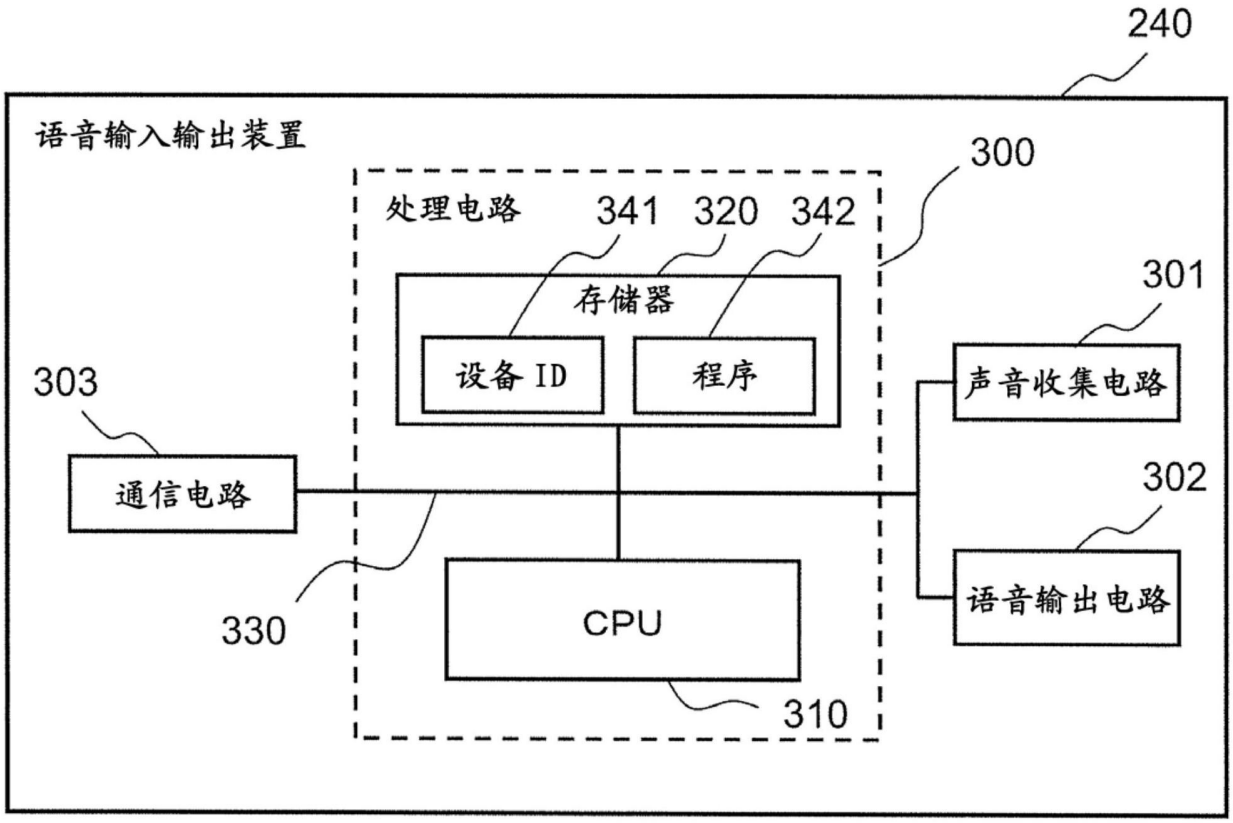

240
语音输入输出装置
300
处理电路
341
320
342
存储器
设备 ID
程序
301
声音收集电路
303
通信电路
302
语音输出电路
330
CPU
310

图3

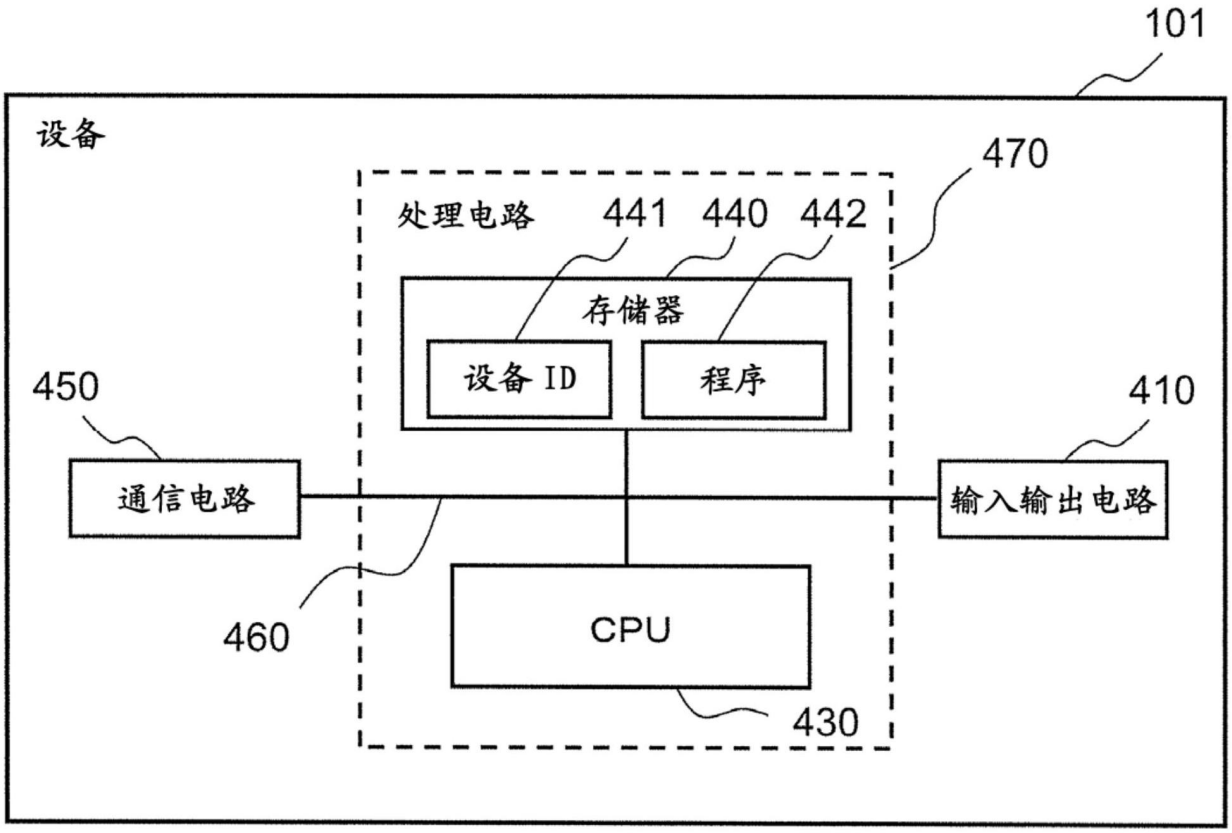

101
设备
470
处理电路
441
440
442
存储器
设备 ID
程序
450
410
通信电路
输入输出电路
460
CPU
430

图4

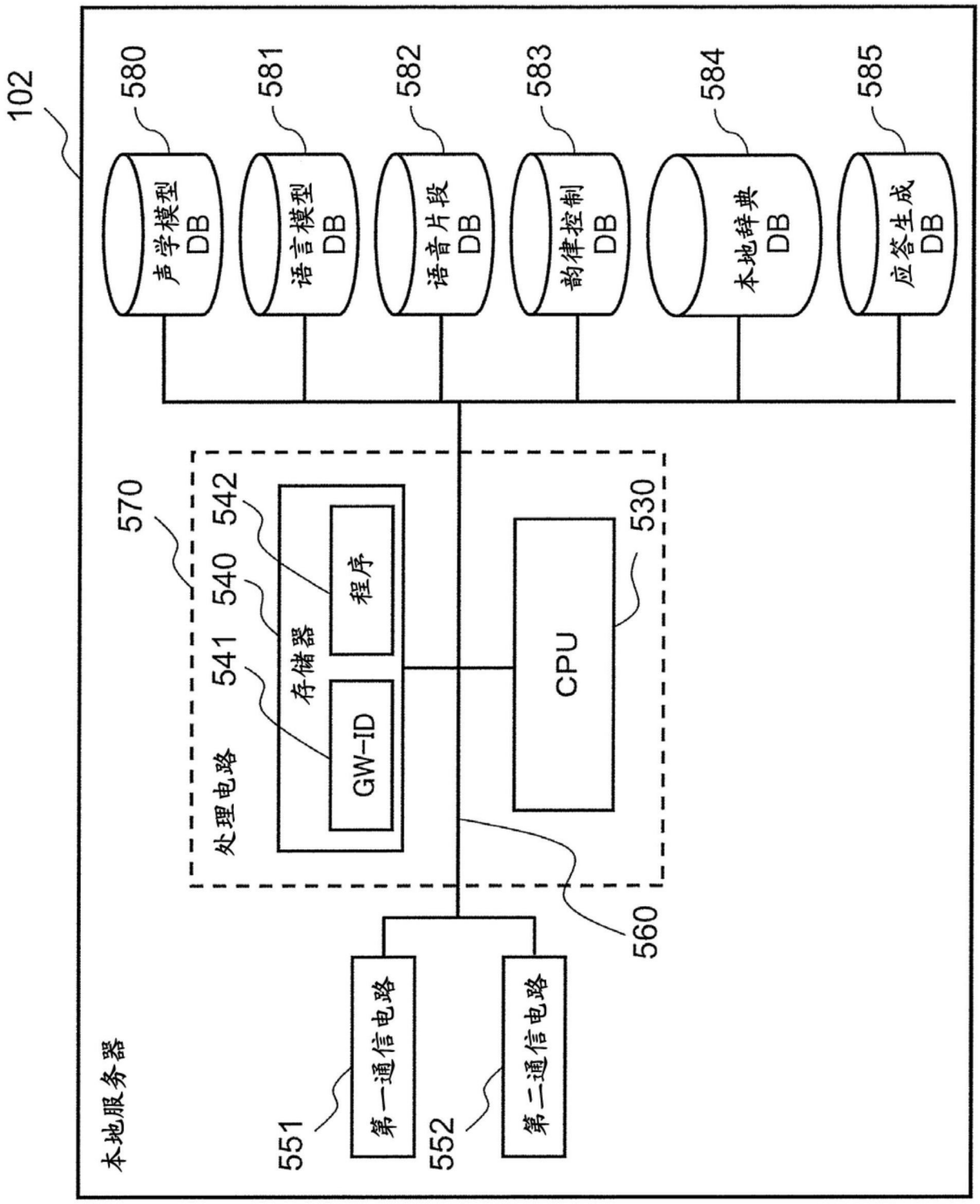

102
580
声学模型 DB
581
语言模型 DB
582
语音片段 DB
583
韵律控制 DB
584
本地辞典 DB
585
应答生成 DB
570
处理电路
542
程序
540
存储器
541
GW-ID
530
CPU
560
551
第一通信电路
552
第二通信电路
本地服务器

图5

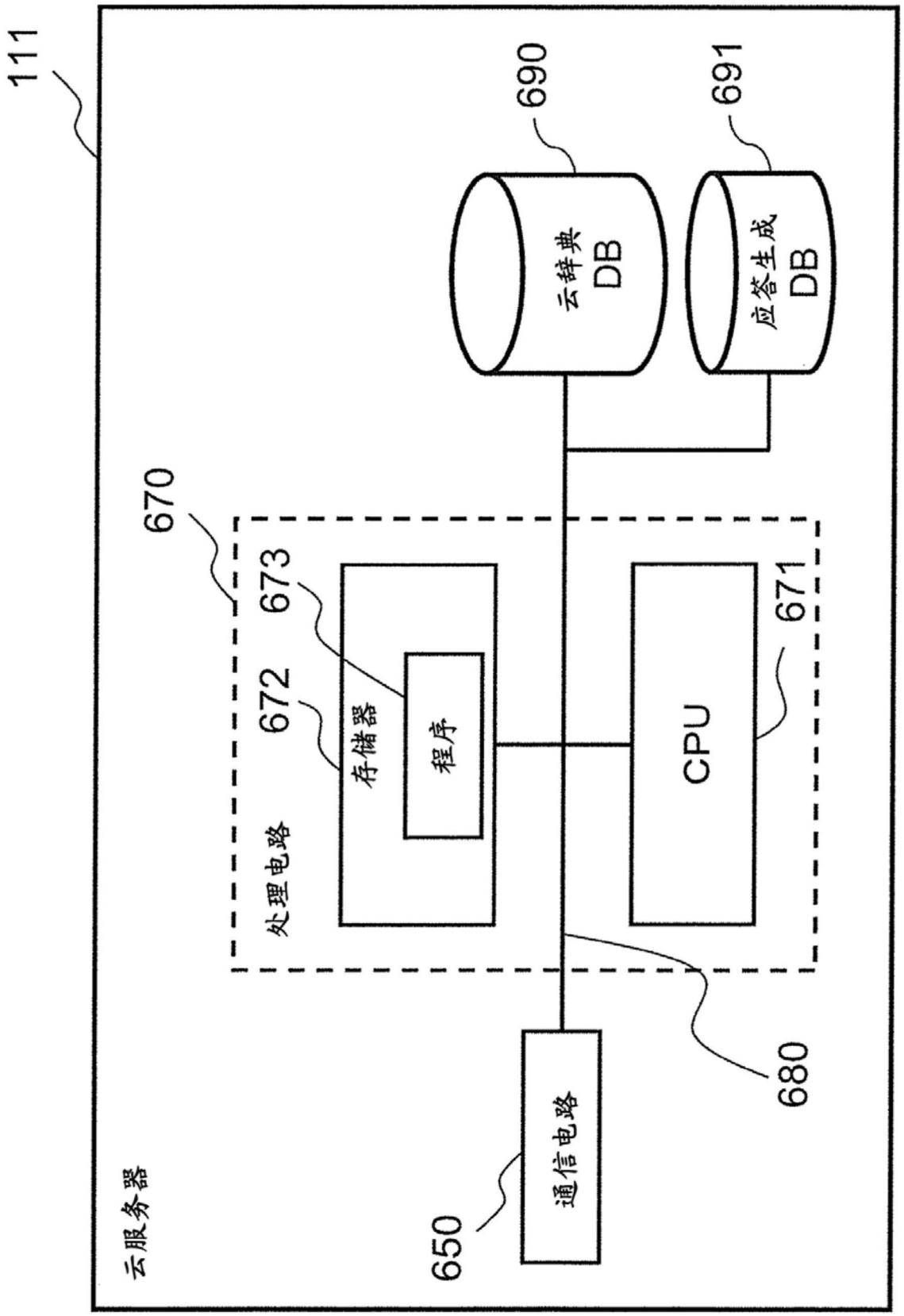

111
690
云辞典
DB
691
应答生成
DB
670
673
672
存储器
程序
处理电路
CPU
671
680
650
通信电路
云服务器

图6

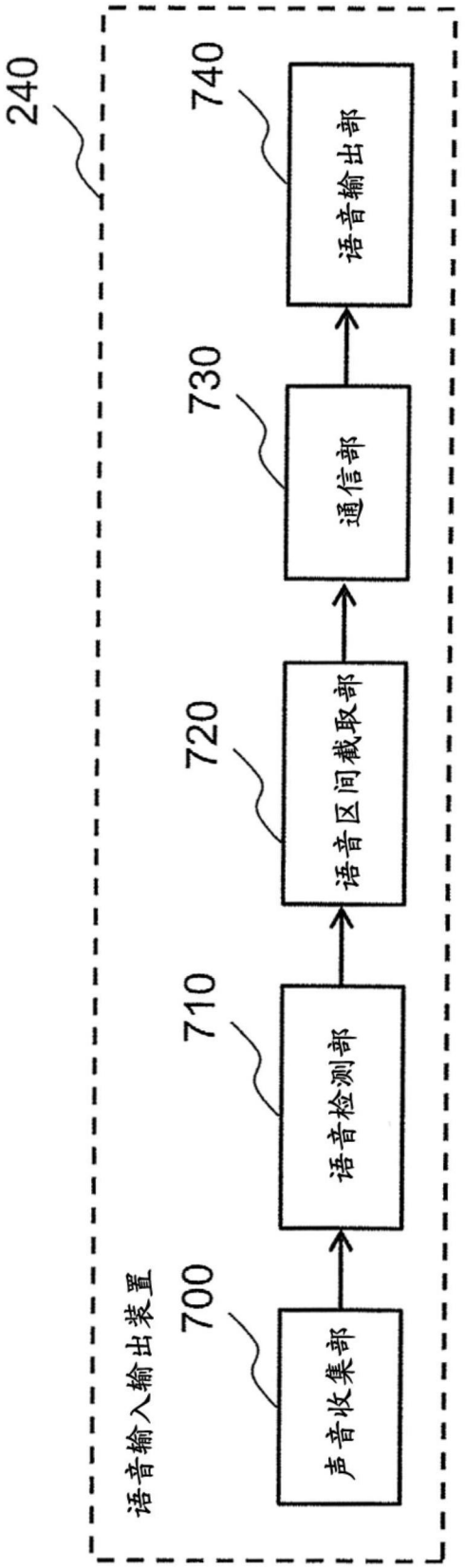

240
语音输入输出装置
700
声音收集部
710
语音检测部
720
语音区间截取部
730
通信部
740
语音输出部

图7

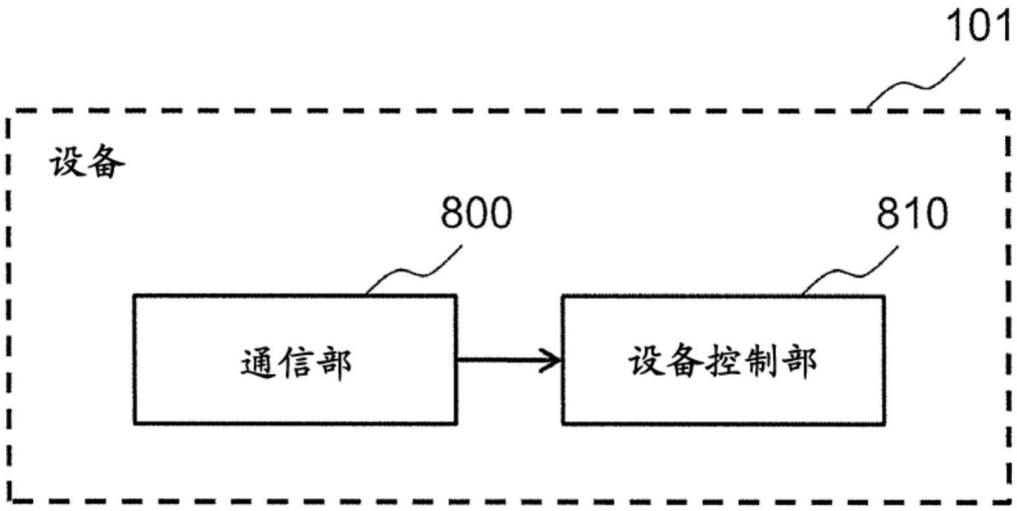


图8

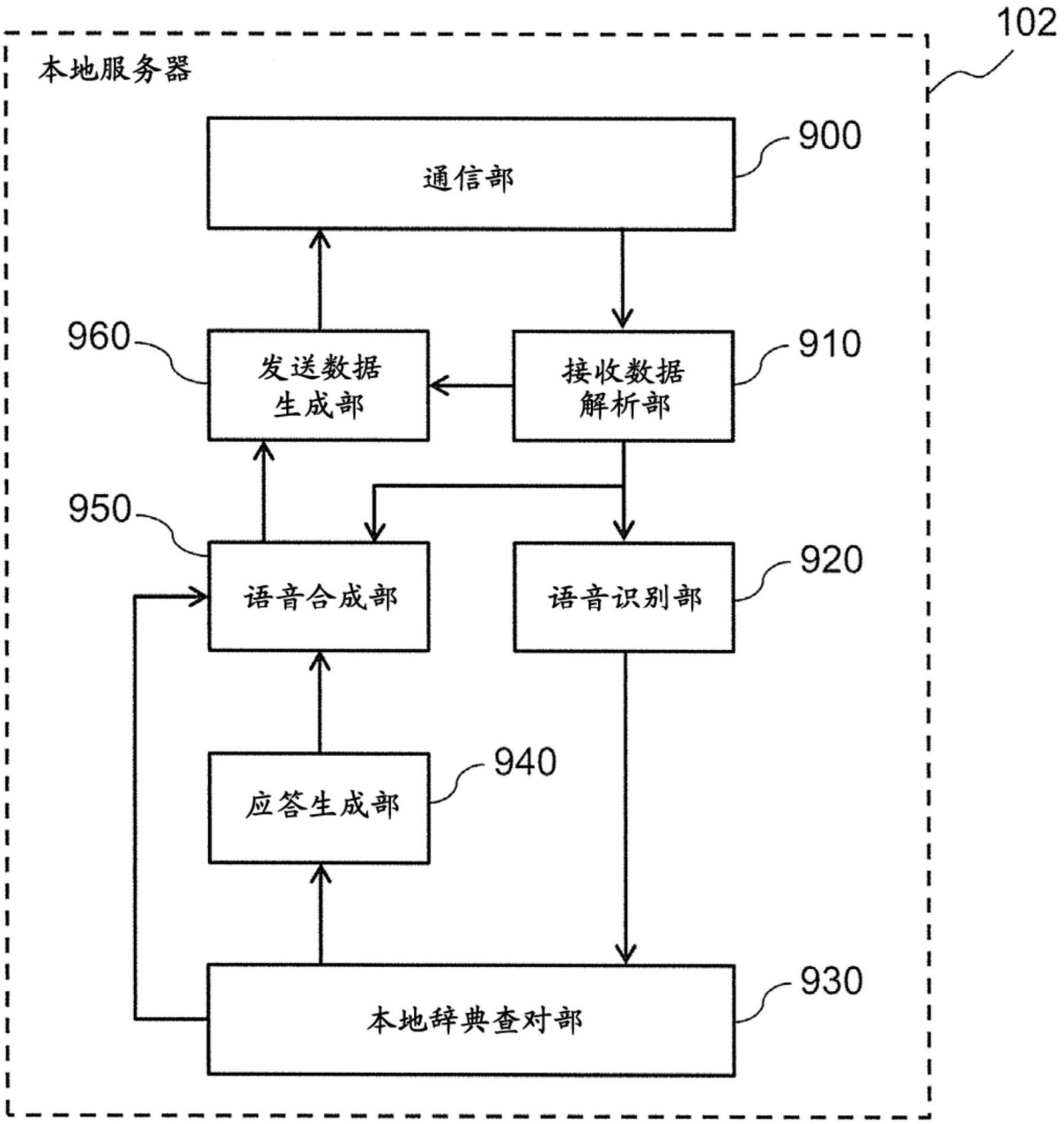

102
本地服务器
900
通信部
960
发送数据
生成部
910
接收数据
解析部
950
语音合成部
920
语音识别部
940
应答生成部
930
本地辞典查对部

图9

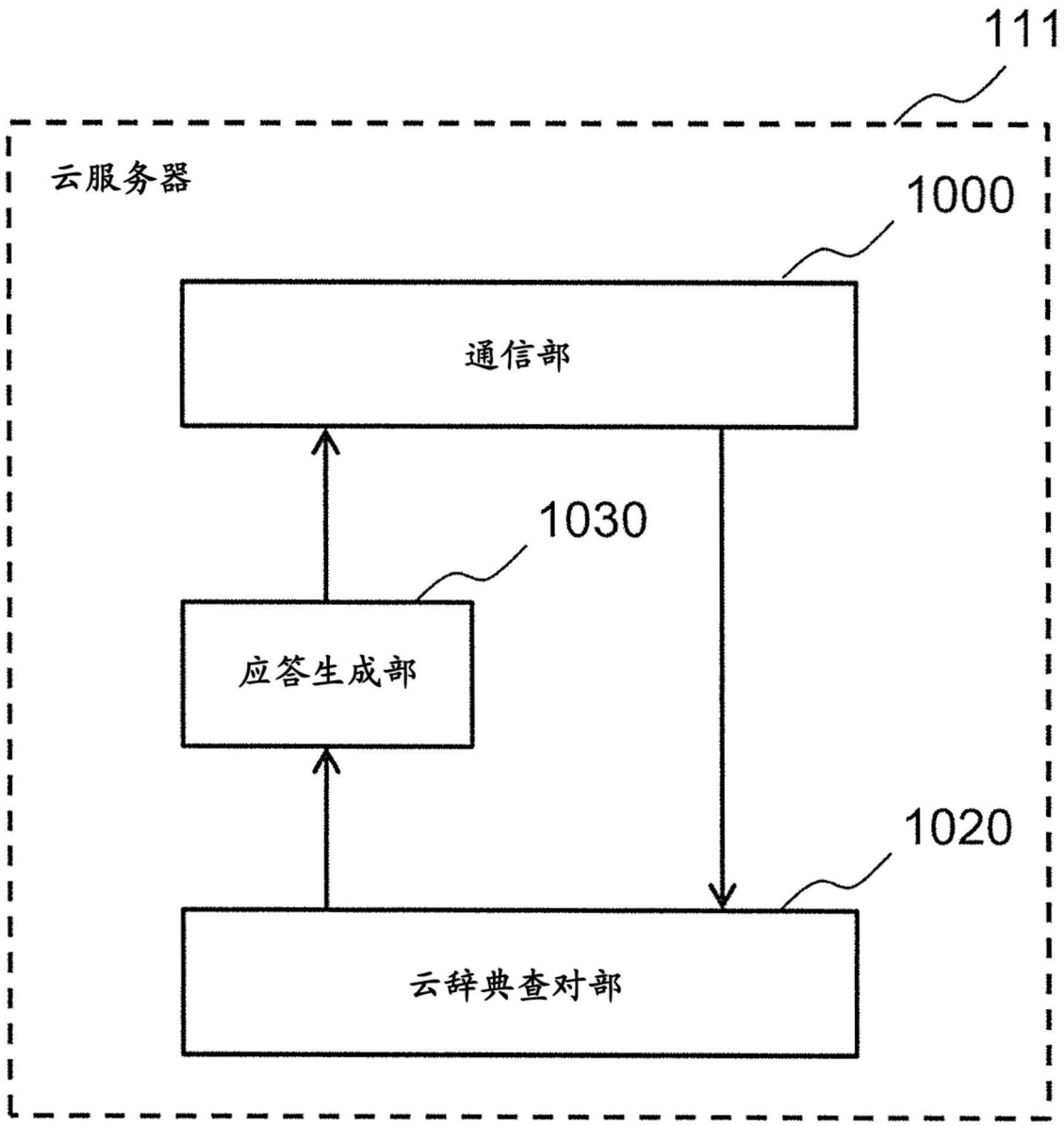


图10

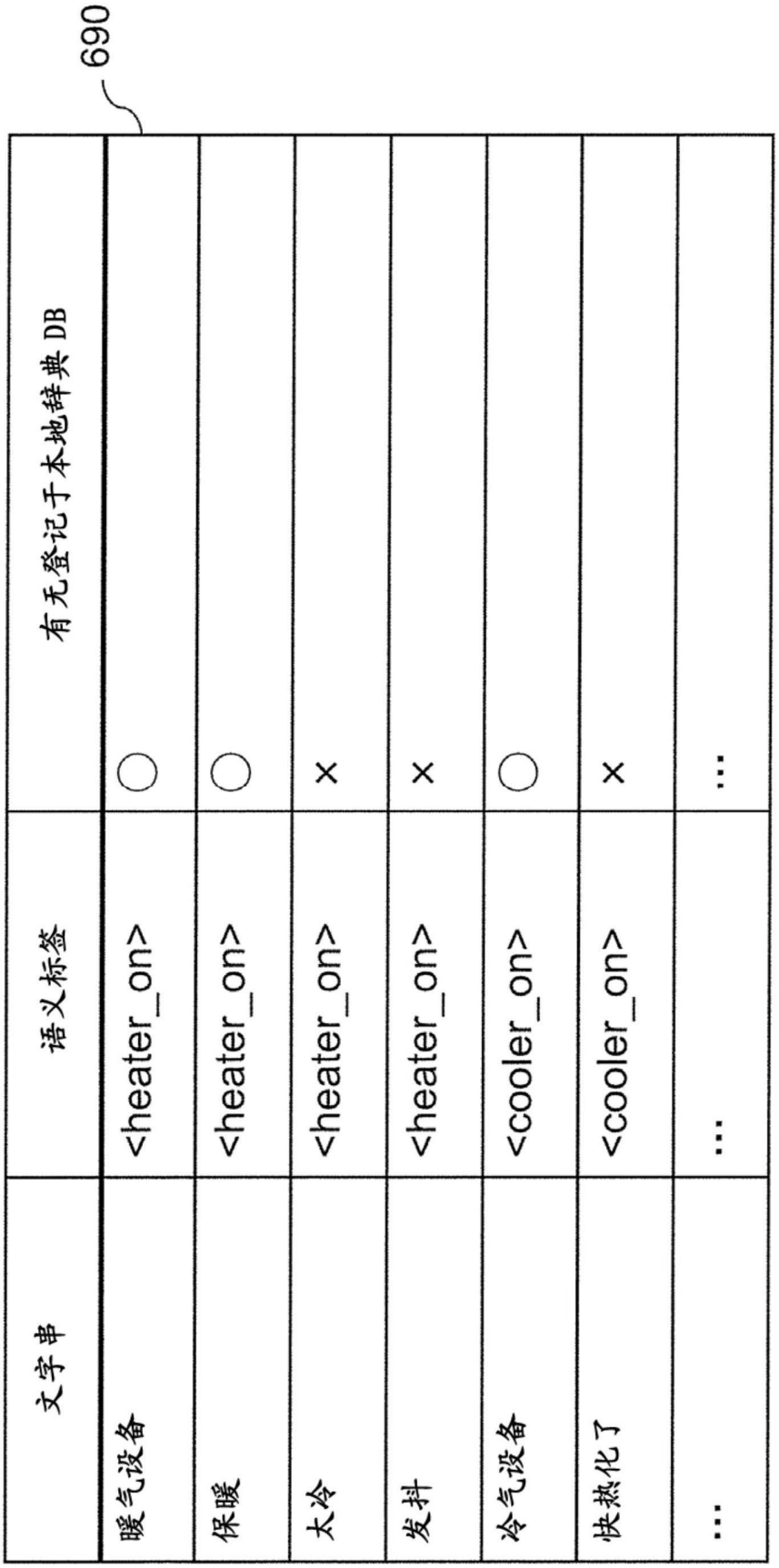

690
文字串
语义标签
有无登记于本地辞典DB
暖气设备
<heater_on>
○
保暖
<heater_on>
○
太冷
<heater_on>
×
发抖
<heater_on>
×
冷气设备
<cooler_on>
○
快热化了
<cooler_on>
×
…
…
…

图11

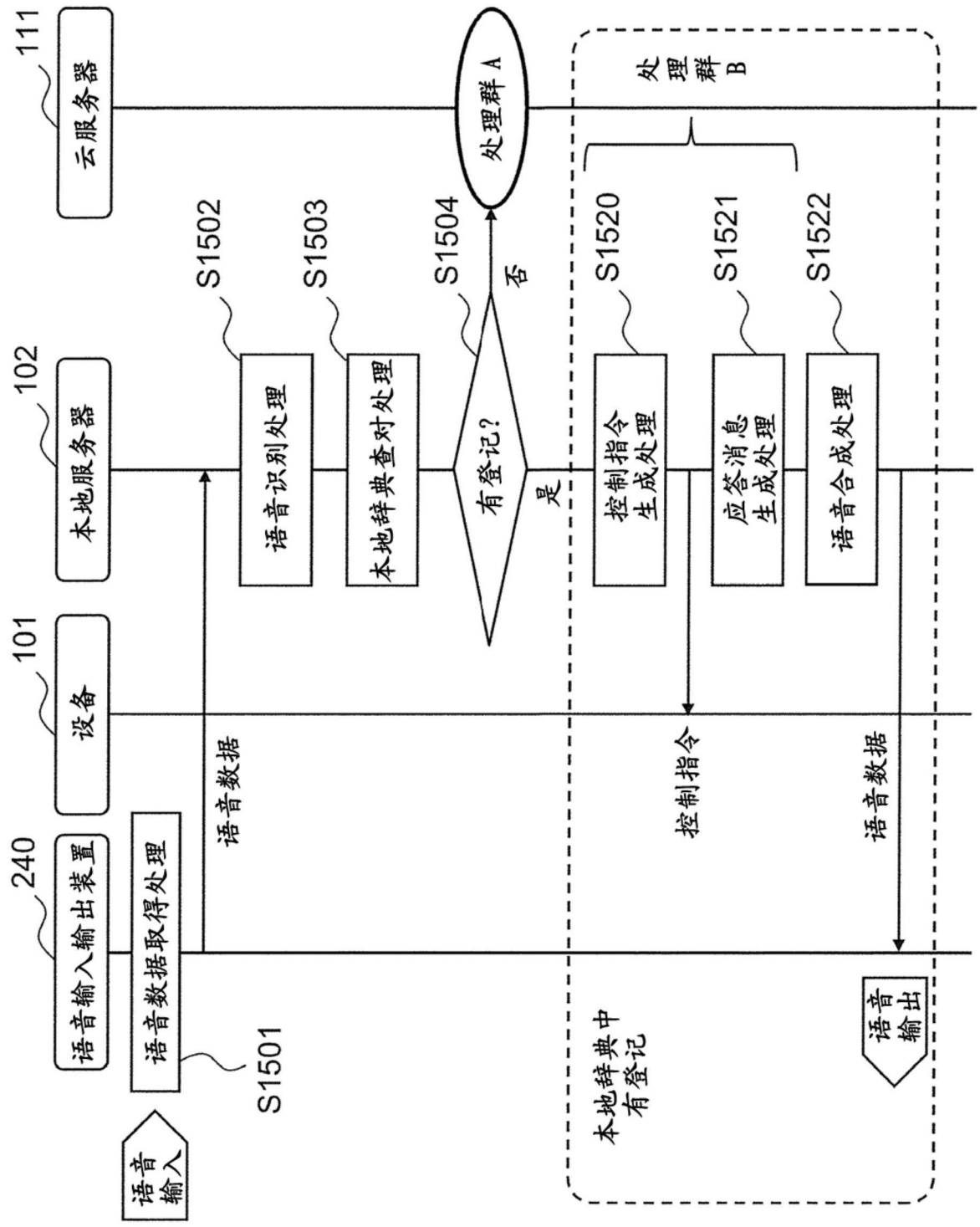

111
云服务器
102
本地服务器
101
设备
240
语音输入输出装置
语音输入
语音数据取得处理
S1501
语音数据
S1502
语音识别处理
S1503
本地辞典查对处理
S1504
有登记?
否
是
处理群A
处理群B
S1520
控制指令生成处理
S1521
应答消息生成处理
S1522
语音合成处理
控制指令
语音数据
本地辞典中有登记
语音输出

图12

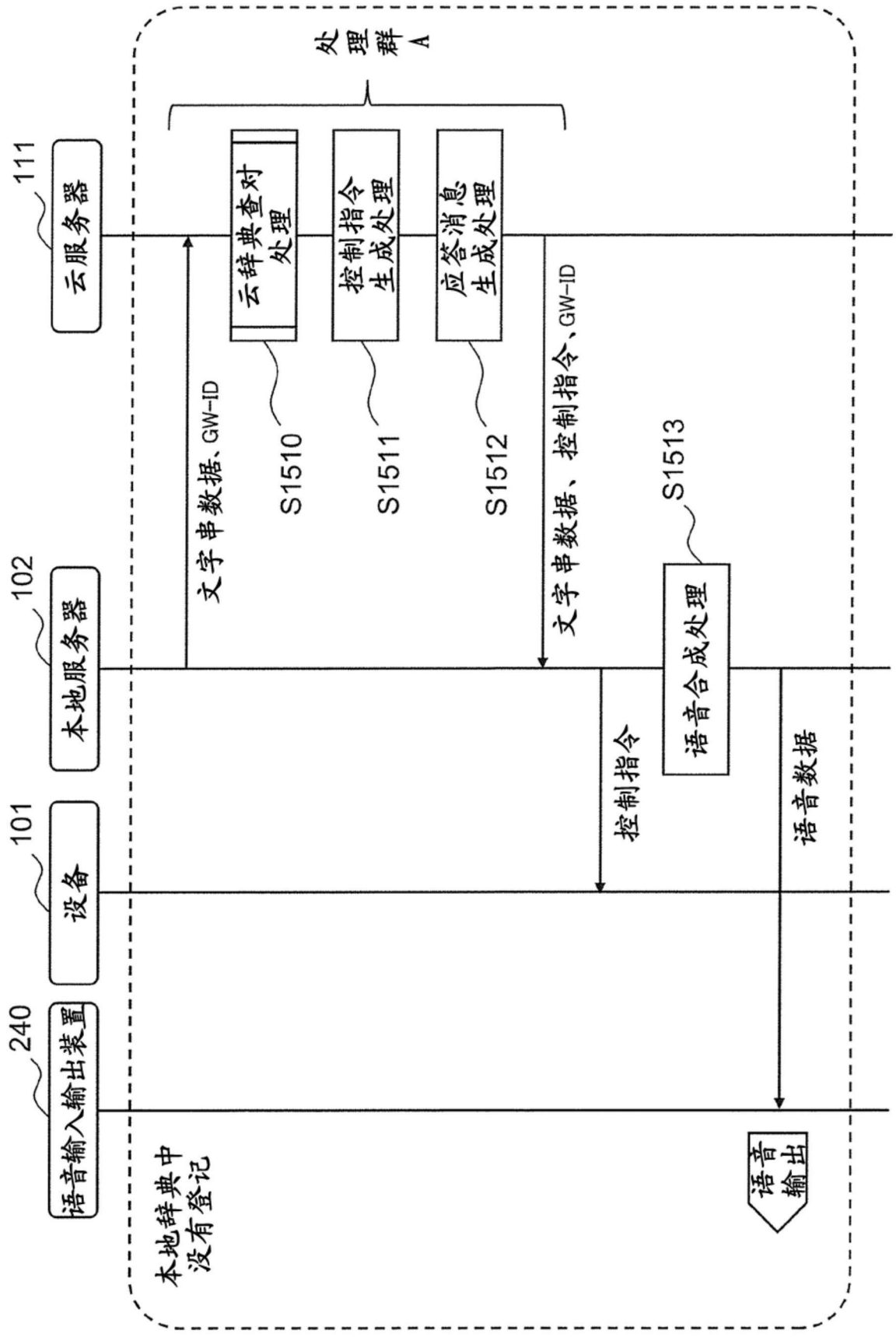

处理群A
111
云服务器
云辞典查对处理
控制指令生成处理
应答消息生成处理
S1510
S1511
S1512
文字串数据、GW-ID
文字串数据、控制指令、GW-ID
S1513
102
本地服务器
语音合成处理
101
设备
控制指令
语音数据
240
语音输入输出装置
本地辞典中没有登记
语音输出

图13

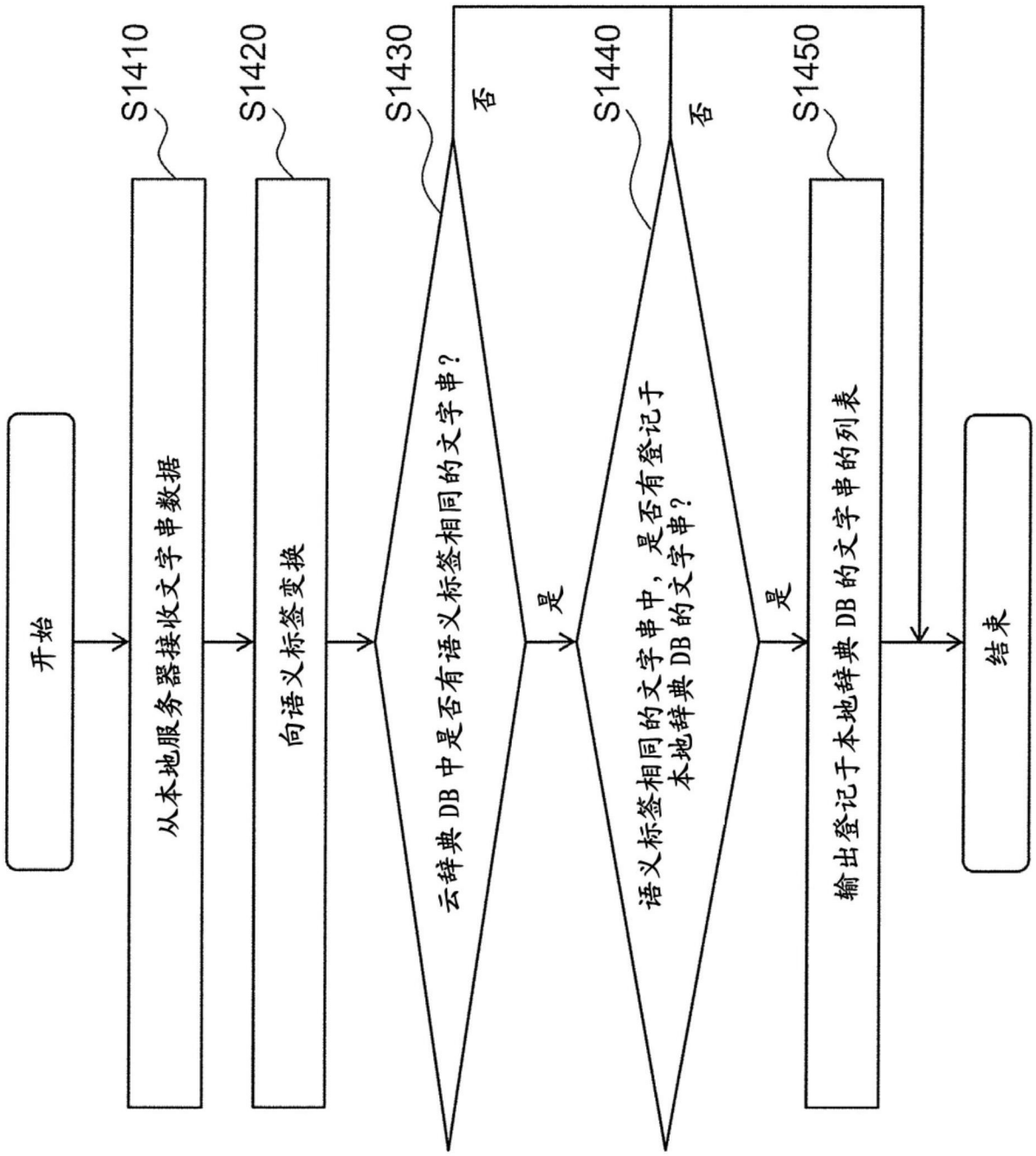

开始
S1410
从本地服务器接收文字串数据
S1420
向语义标签变换
S1430
云辞典 DB 中是否有语义标签相同的文字串？
否
是
S1440
语义标签相同的文字串中，是否有登记于本地辞典 DB 的文字串？
否
是
S1450
输出登记于本地辞典 DB 的文字串的列表
结束

图14

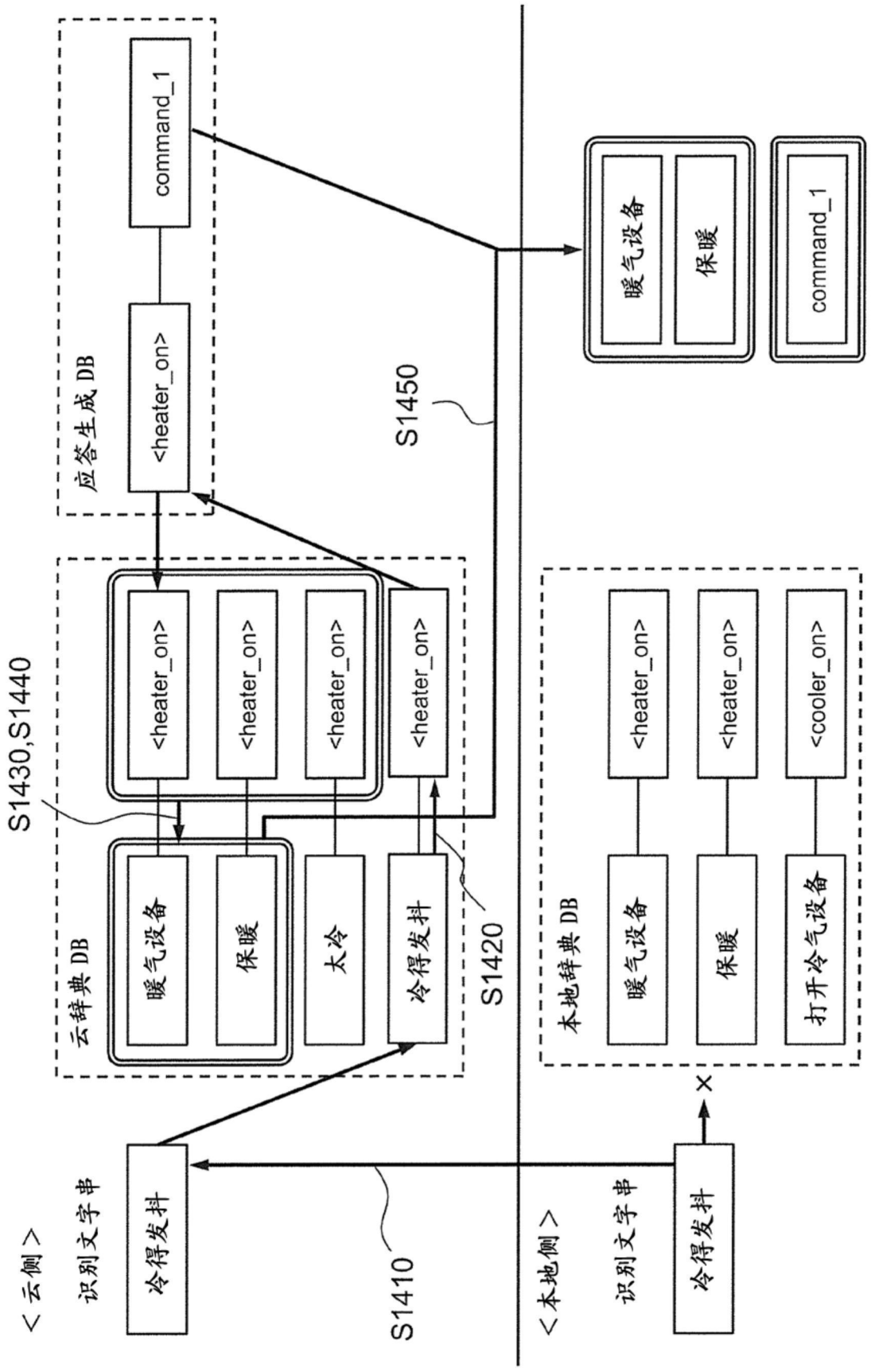

<云侧>
识别文字串
冷得发抖
S1410
云辞典 DB
暖气设备
保暖
太冷
冷得发抖
<heater_on>
<heater_on>
<heater_on>
<heater_on>
S1420
S1430,S1440
应答生成 DB
<heater_on>
command_1
S1450
<本地侧>
识别文字串
冷得发抖
本地辞典 DB
暖气设备
保暖
打开冷气设备
<heater_on>
<heater_on>
<cooler_on>
暖气设备
保暖
command_1

图15

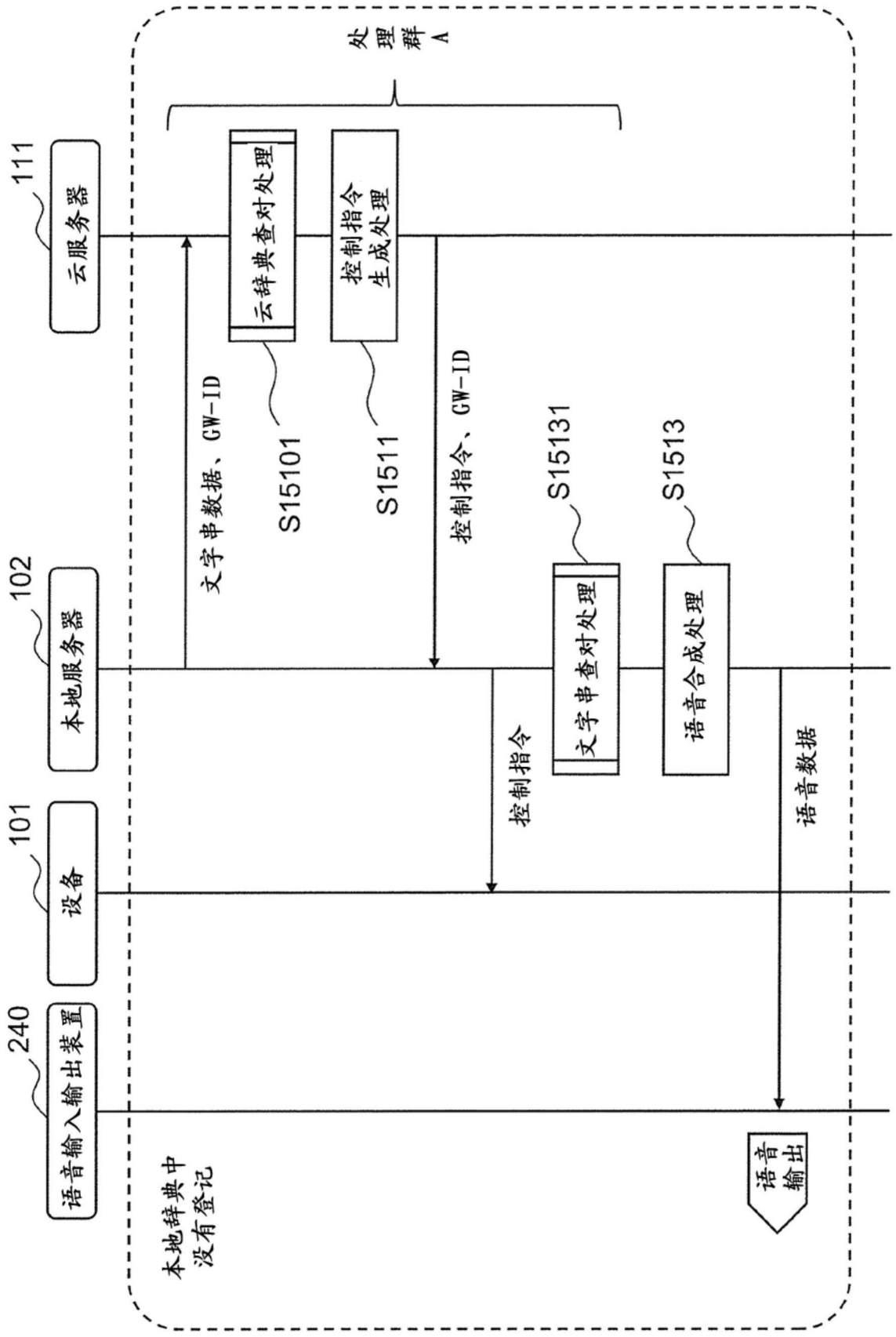

处理群A
111
云服务器
102
本地服务器
101
设备
240
语音输入输出装置
云辞典查对处理
控制指令生成处理
文字串数据、GW-ID
控制指令、GW-ID
S15101
S1511
S15131
S1513
文字串查对处理
语音合成处理
控制指令
语音数据
语音输出
本地辞典中没有登记

图16

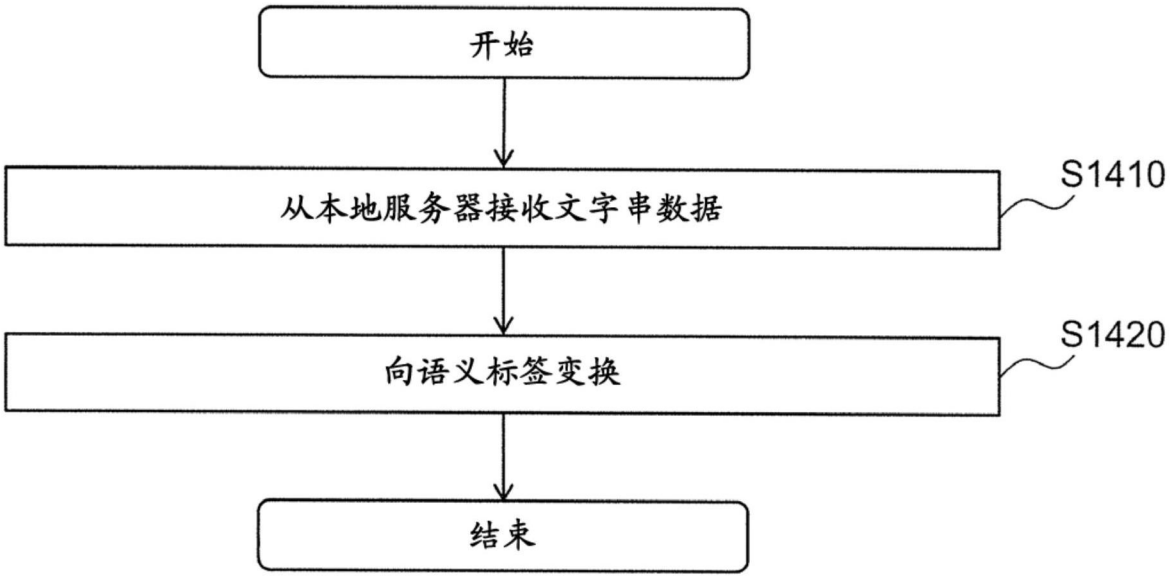


图17

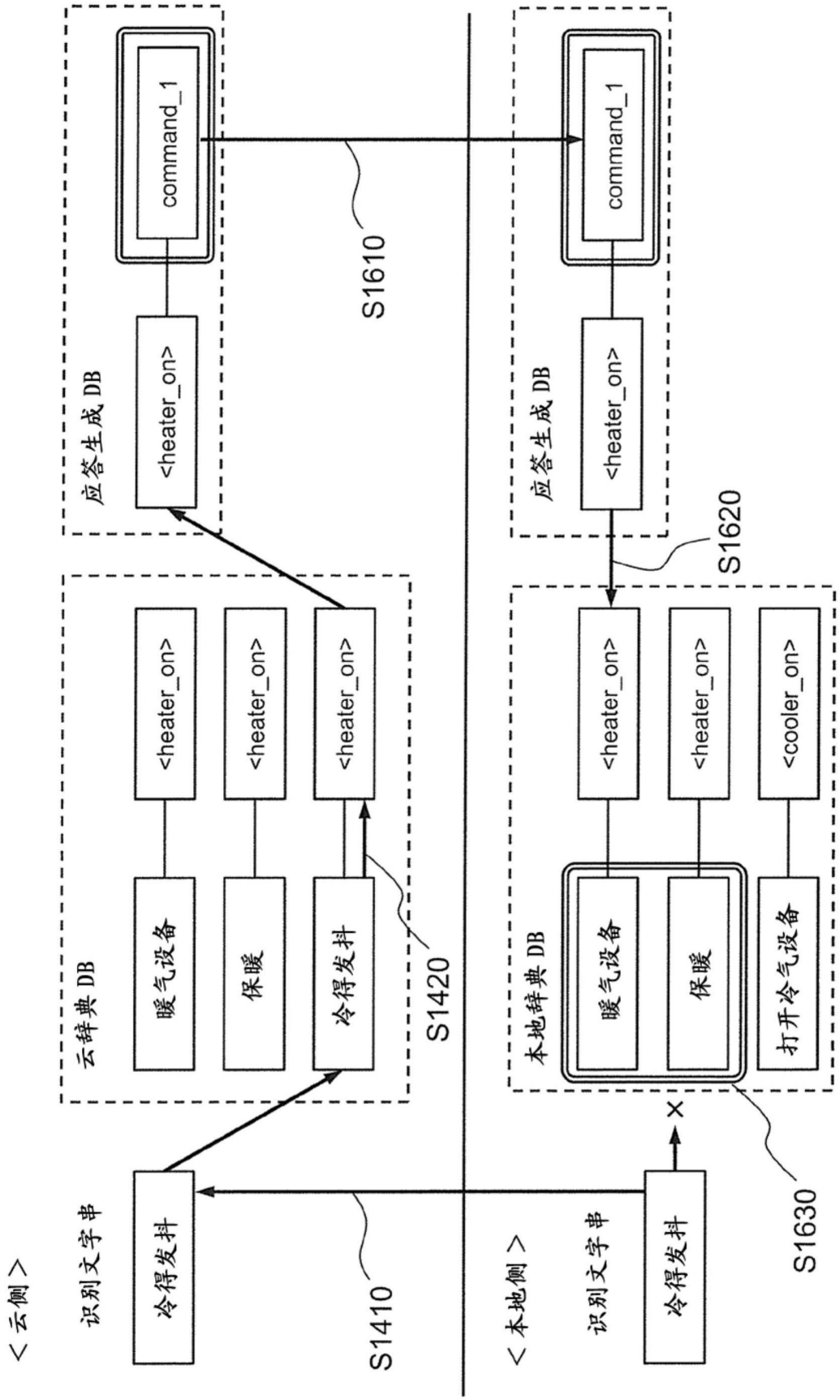

<云侧>
识别文字串
冷得发抖
S1410
云辞典 DB
暖气设备
保暖
冷得发抖
<heater_on>
<heater_on>
<heater_on>
S1420
应答生成 DB
<heater_on>
command_1
S1610
<本地侧>
识别文字串
冷得发抖
X
本地辞典 DB
暖气设备
保暖
打开冷气设备
<heater_on>
<heater_on>
<cooler_on>
S1620
S1630
应答生成 DB
<heater_on>
command_1

图18

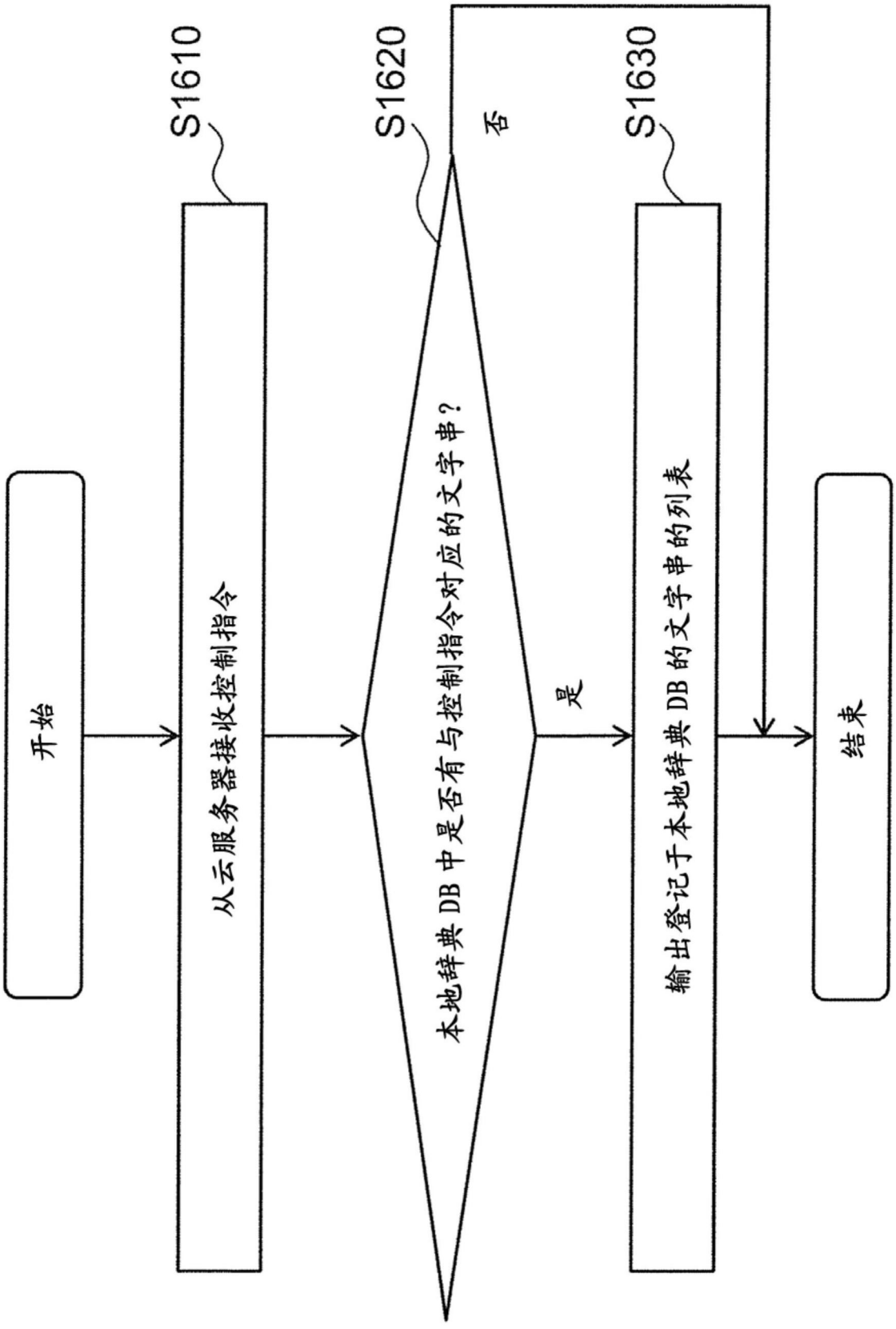

开始
从云服务器接收控制指令
S1610
本地辞典 DB 中是否有与控制指令对应的文字串？
S1620
否
是
输出登记于本地辞典 DB 的文字串的列表
S1630
结束

图19

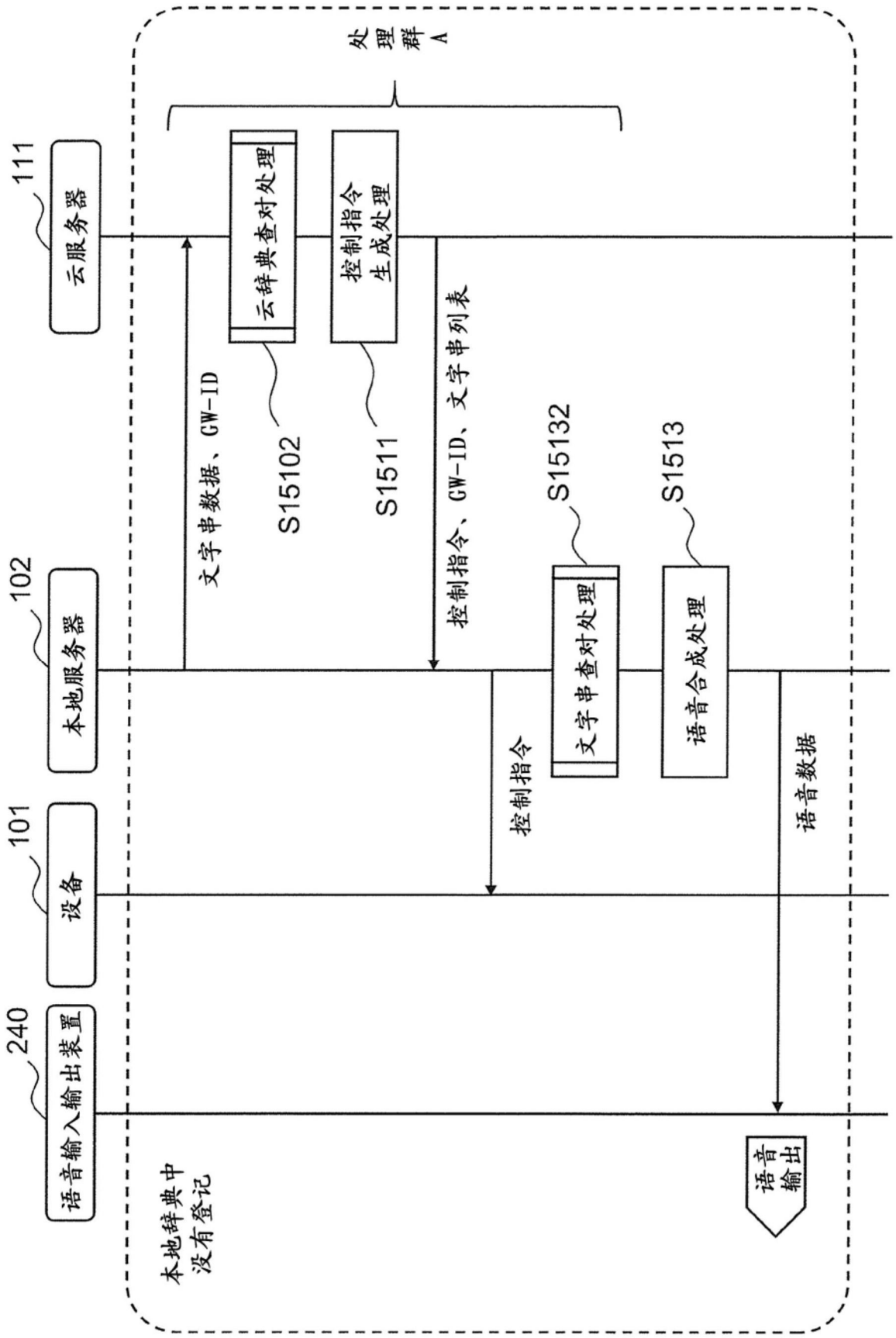

处理群A
111
云服务器
云辞典查对处理
控制指令生成处理
文字串数据、GW-ID
S15102
S1511
控制指令、GW-ID、文字串列表
S15132
S1513
102
本地服务器
文字串查对处理
语音合成处理
控制指令
语音数据
101
设备
240
语音输入输出装置
本地辞典中没有登记
语音输出

图20

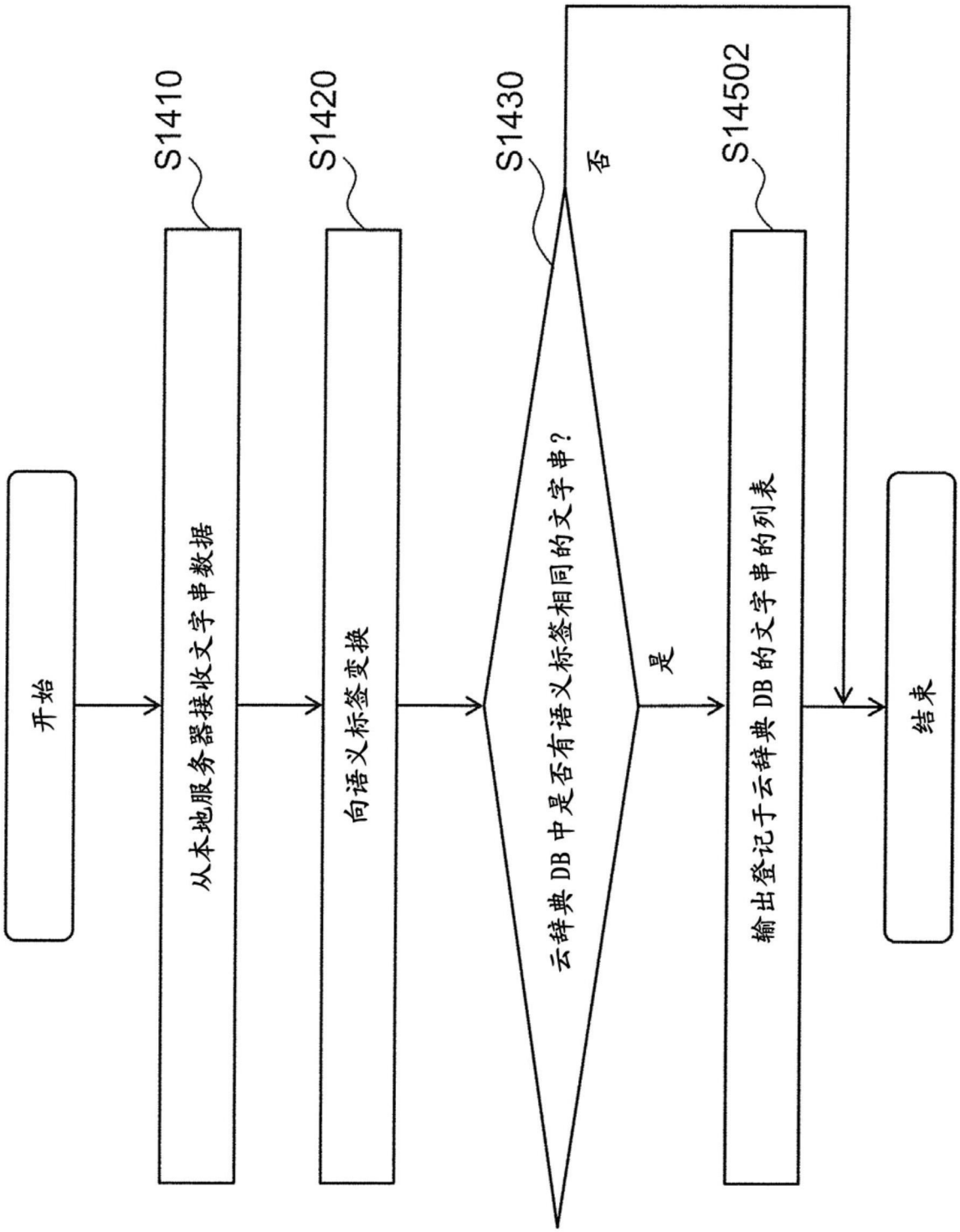

开始
S1410
从本地服务器接收文字串数据
S1420
向语义标签变换
S1430
云辞典 DB 中是否有语义标签相同的文字串？
否
是
S14502
输出登记于云辞典 DB 的文字串的列表
结束

图21

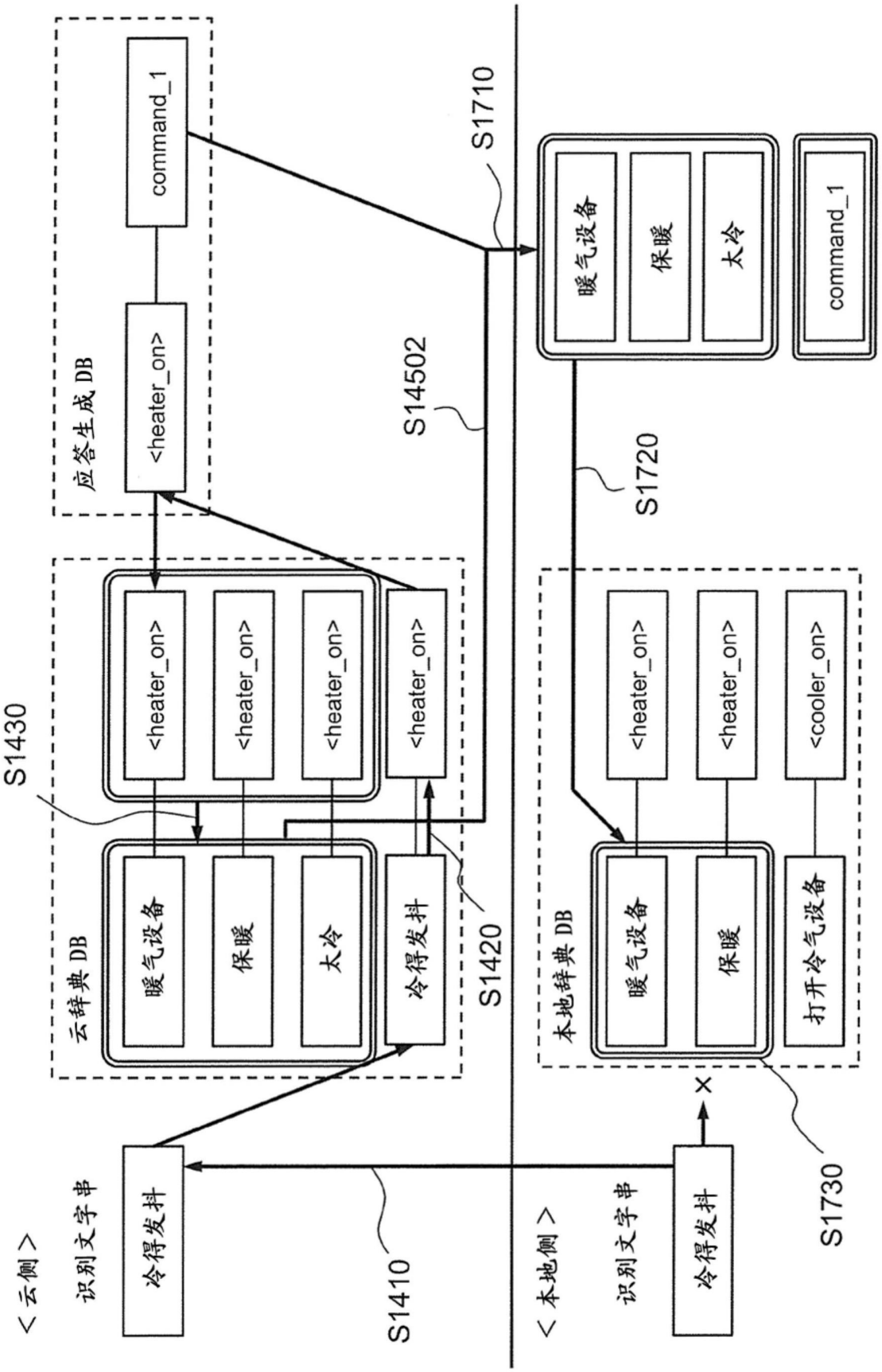

<云侧>
识别文字串
冷得发抖
S1410
云辞典 DB
暖气设备
保暖
太冷
冷得发抖
<heater_on>
<heater_on>
<heater_on>
<heater_on>
S1420
S1430
应答生成 DB
<heater_on>
command_1
S1710
S14502
S1720
<本地侧>
识别文字串
冷得发抖
本地辞典 DB
暖气设备
保暖
打开冷气设备
<heater_on>
<heater_on>
<cooler_on>
S1730
暖气设备
保暖
太冷
command_1

图22

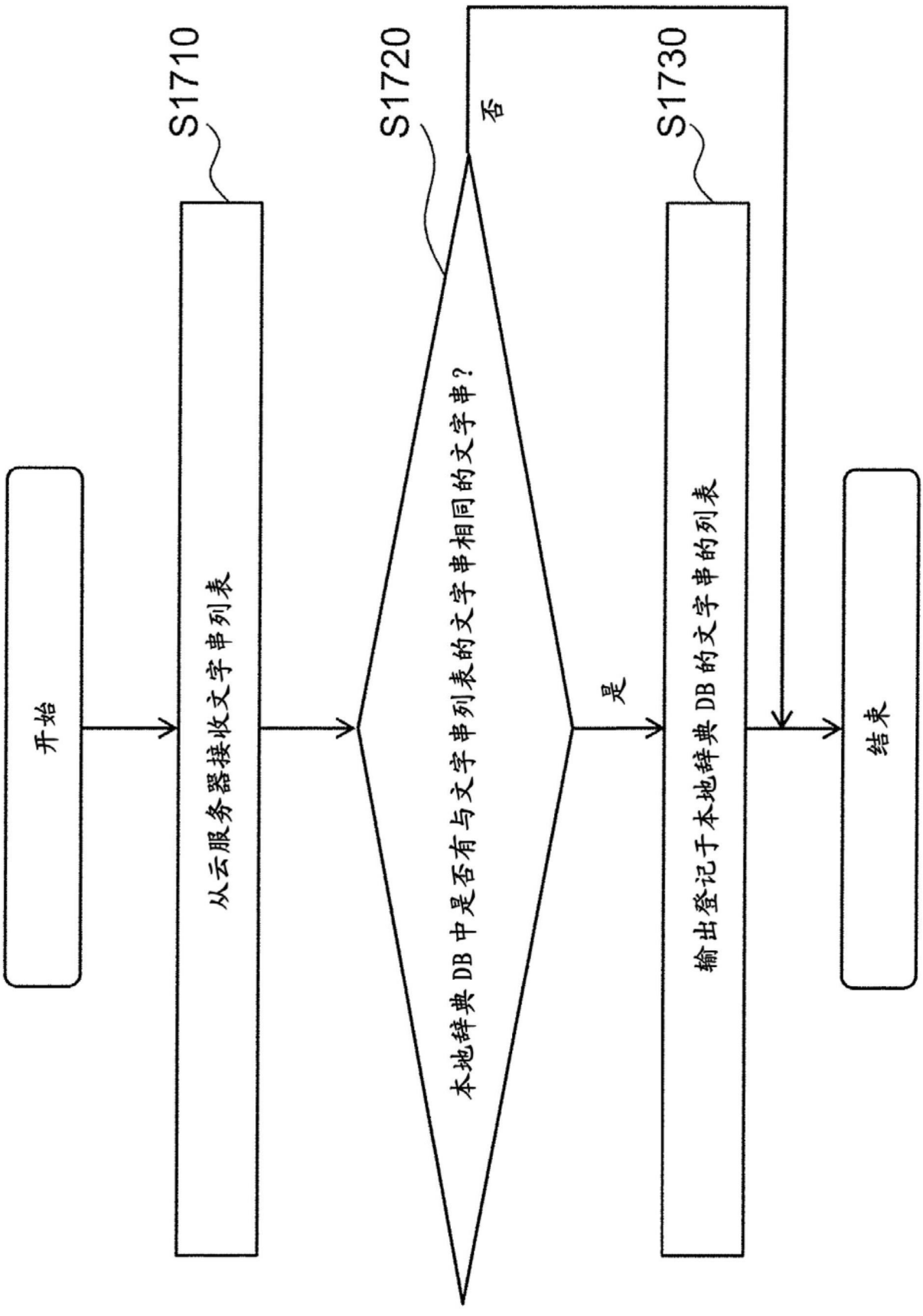

开始
从云服务器接收文字串列表
S1710
本地辞典 DB 中是否有与文字串列表的文字串相同的文字串？
S1720
是
否
输出登记于本地辞典 DB 的文字串的列表
S1730
结束

图23

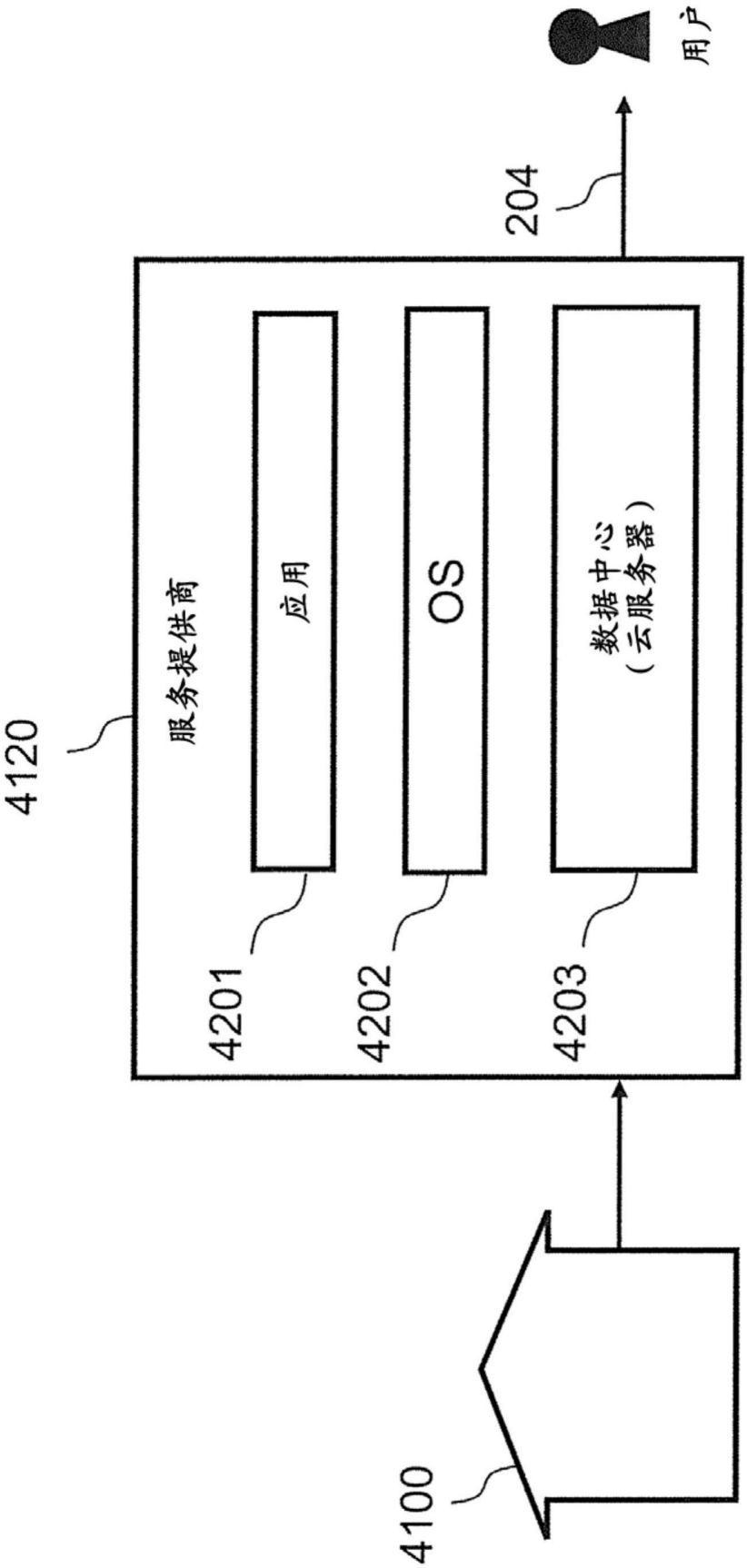


图24

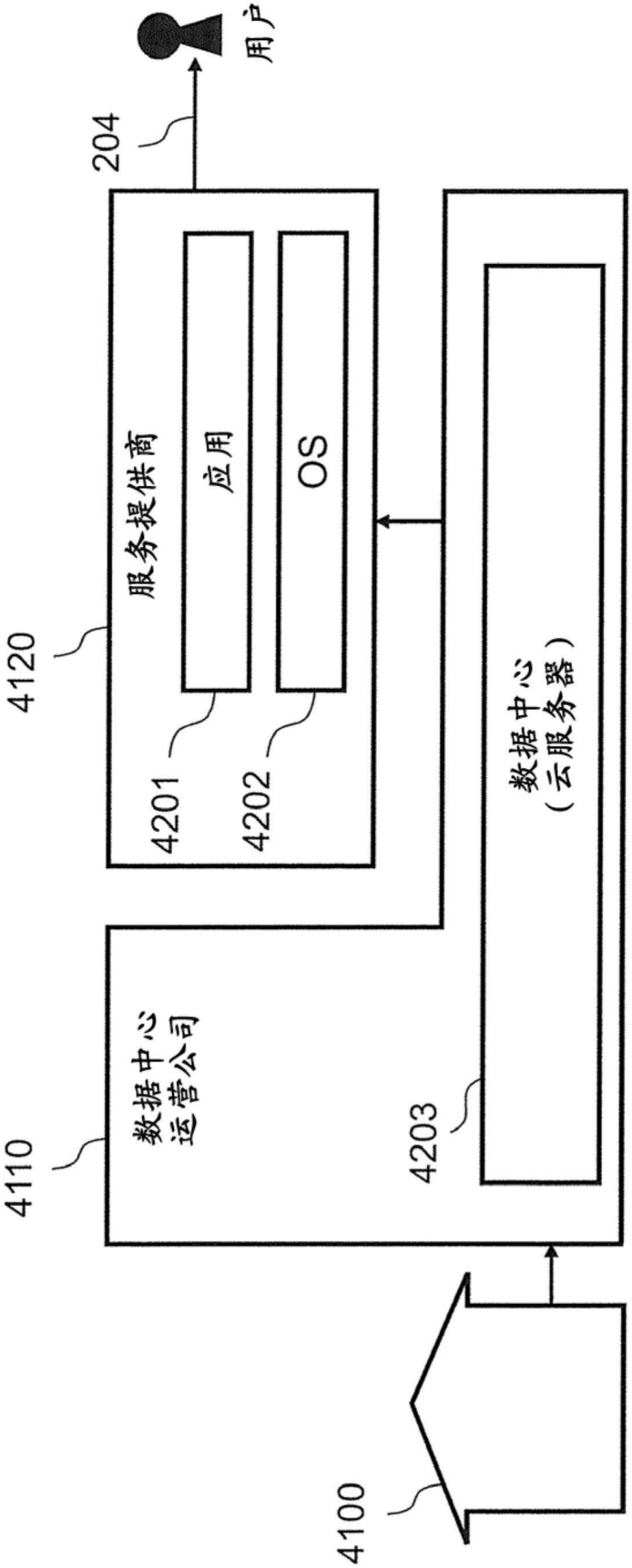

用户
204
服务提供商
应用
OS
4120
4201
4202
数据中心
运营公司
4110
数据中心
（云服务器）
4203
4100

图25

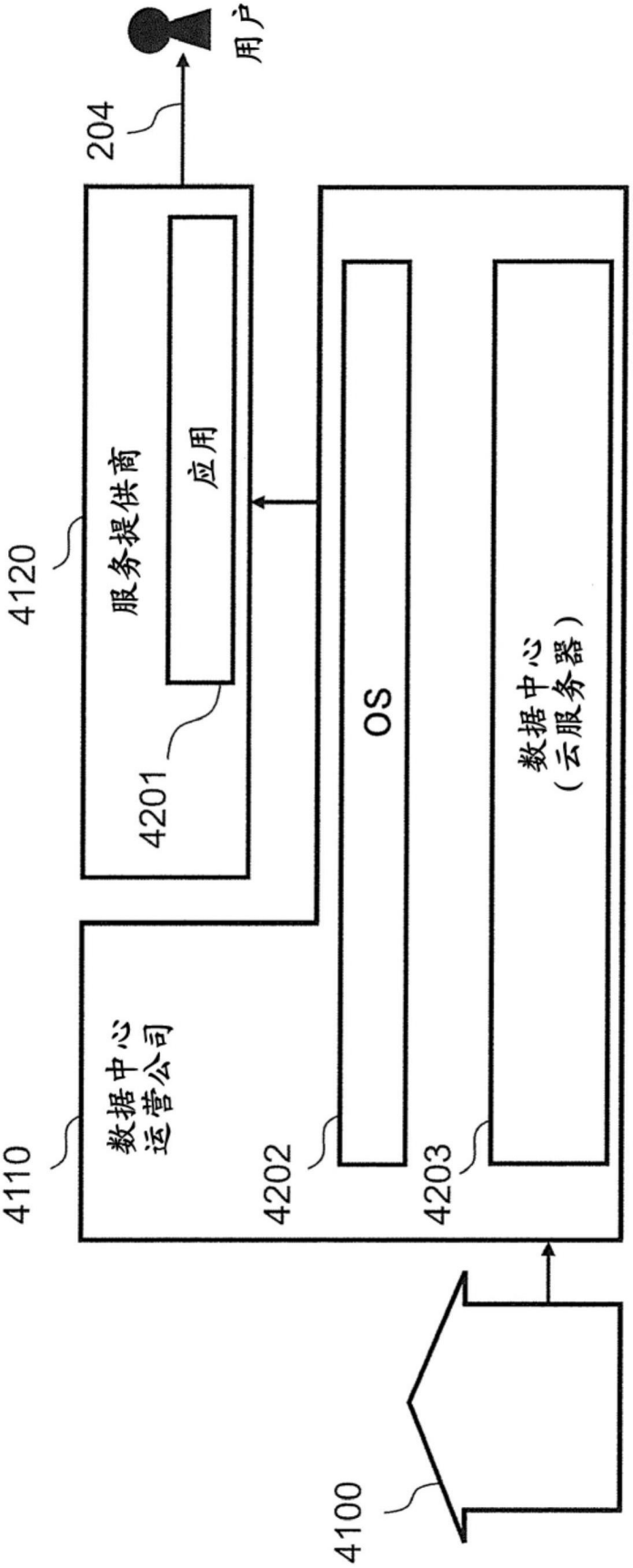

用户
204
4120
服务提供商
应用
4201
OS
数据中心
(云服务器)
4110
数据中心
运营公司
4202
4203
4100

图26

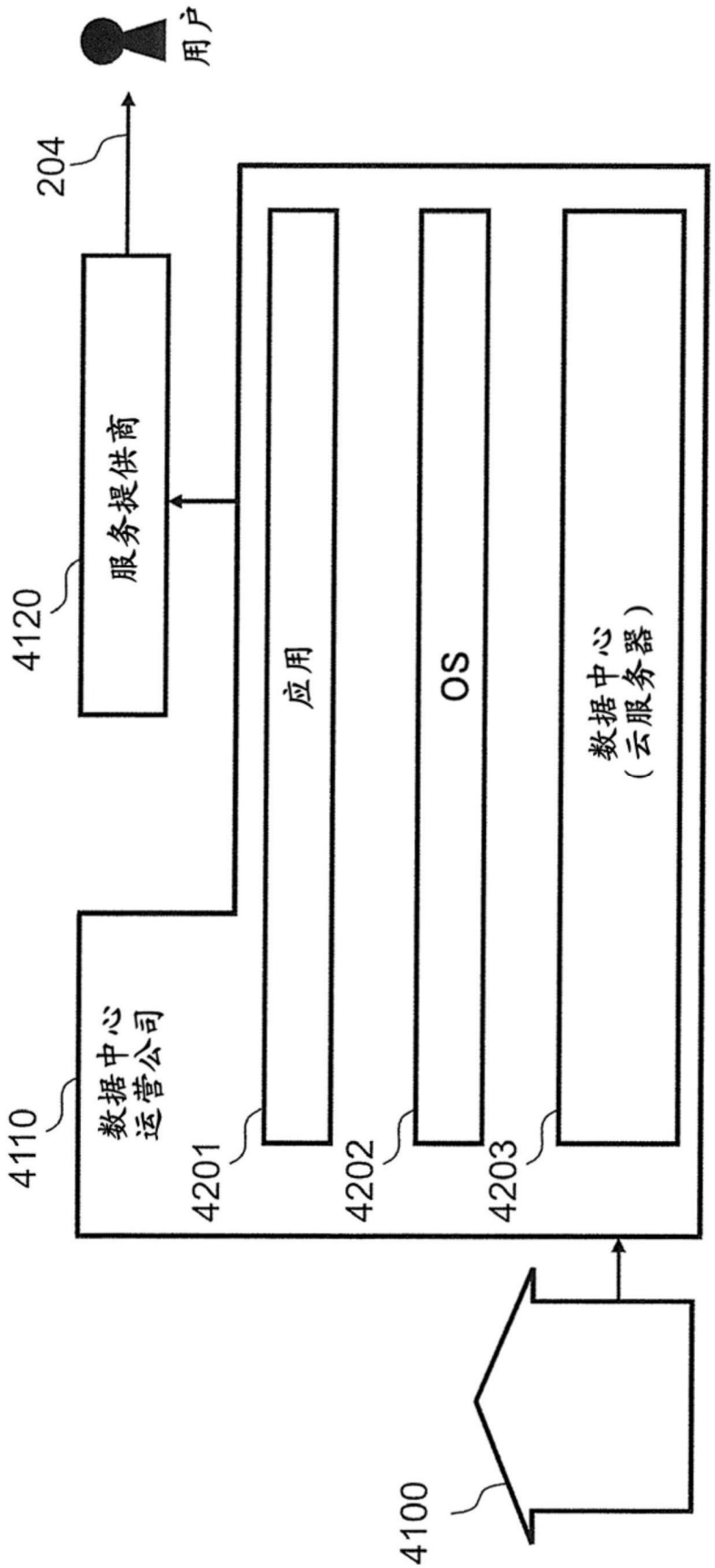

用户
204
服务提供商
4120
应用
OS
数据中心
（云服务器）
数据中心
运营公司
4110
4201
4202
4203
4100

图27