



(12) 发明专利

(10) 授权公告号 CN 109937447 B

(45) 授权公告日 2023.03.10

(21) 申请号 201780069660.7

小窪浩明

(22) 申请日 2017.11.14

(74) 专利代理机构 永新专利商标代理有限公司

(65) 同一申请的已公布的文献号

72002

申请公布号 CN 109937447 A

专利代理师 徐殿军

(43) 申请公布日 2019.06.25

(51) Int.Cl.

(30) 优先权数据

G10L 15/32 (2013.01)

2016-222723 2016.11.15 JP

G10L 15/00 (2013.01)

G10L 15/10 (2006.01)

(85) PCT国际申请进入国家阶段日

2019.05.10

(56) 对比文件

US 2010004930 A1, 2010.01.07

(86) PCT国际申请的申请数据

PCT/JP2017/040998 2017.11.14

US 2010004930 A1, 2010.01.07

US 2014163977 A1, 2014.06.12

(87) PCT国际申请的公布数据

W02018/092786 JA 2018.05.24

CN 1882015 A, 2006.12.20

US 9117452 B1, 2015.08.25

CN 105374357 A, 2016.03.02

CN 101207584 A, 2008.06.25

(73) 专利权人 歌乐株式会社

地址 日本埼玉县

审查员 薛彦杰

(72) 发明人 本间健 张睿 松本卓也

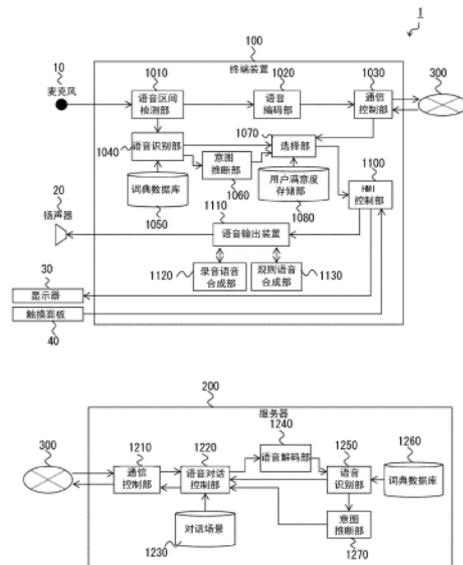
权利要求书3页 说明书25页 附图9页

(54) 发明名称

语音识别装置、语音识别系统

(57) 摘要

语音识别装置具备：语音检测部，检测从用户输入的语音；信息提供部，使用基于第一语音识别部对所述语音的识别结果的第一语音识别信息、或基于与所述第一语音识别部不同的第二语音识别部对所述语音的识别结果的第二语音识别信息中的某一方的语音识别信息，进行向所述用户的信息提供；以及选择部，基于从所述语音被输入起的经过时间，选择所述第一语音识别信息或所述第二语音识别信息中的某一方作为所述信息提供部所使用的语音识别信息，并且使由所述信息提供部进行的所述信息提供的方法变化。



1. 一种语音识别装置,具备:

语音检测部,检测从用户输入的语音;

信息提供部,使用第一语音识别信息或第二语音识别信息中的某一方的语音识别信息来进行向所述用户的信息提供,该第一语音识别信息基于第一语音识别部对所述语音的识别结果,该第二语音识别信息基于与所述第一语音识别部不同的第二语音识别部对所述语音的识别结果;以及

选择部,基于从所述语音被输入起的经过时间,选择所述第一语音识别信息或所述第二语音识别信息中的某一方作为所述信息提供部所使用的语音识别信息,并且使由所述信息提供部进行的所述信息提供的方法变化,

所述选择部基于从所述语音被输入起的经过时间,计算第一用户满意度与第二用户满意度,该第一用户满意度表示所述用户对使用了所述第一语音识别信息的情况下的所述信息提供的满意程度的预测值,该第二用户满意度表示所述用户对使用了所述第二语音识别信息的情况下的所述信息提供的满意程度的预测值,

比较所述第一用户满意度与所述第二用户满意度,基于该比较结果选择所述第一语音识别信息或所述第二语音识别信息中的某一方。

2. 如权利要求1所述的语音识别装置,其中,

在先取得所述第一语音识别信息、尚未取得所述第二语音识别信息的情况下,所述选择部,

测量与从所述语音被输入起到获得所述第一语音识别信息为止的经过时间相关的第一经过时间,并且预测与从所述语音被输入起到获得所述第二语音识别信息为止的经过时间相关的第二经过时间,

基于测量出的所述第一经过时间,计算所述第一用户满意度,

基于预测出的所述第二经过时间,计算所述第二用户满意度,

比较计算出的所述第一用户满意度与所述第二用户满意度,基于该比较结果决定是否选择所述第一语音识别信息。

3. 如权利要求1所述的语音识别装置,其中

在取得完毕所述第一语音识别信息、未取得所述第二语音识别信息的情况下,所述选择部,

测量与从所述语音被输入起到当前为止的经过时间相关的第三经过时间,

基于测量出的所述第三经过时间,计算所述第一用户满意度以及所述第二用户满意度,

比较计算出的所述第一用户满意度与所述第二用户满意度,基于该比较结果决定是否选择所述第一语音识别信息。

4. 如权利要求1所述的语音识别装置,其中,

在先取得所述第一语音识别信息、后取得所述第二语音识别信息的情况下,所述选择部,

测量与从所述语音被输入起到获得所述第二语音识别信息为止的经过时间相关的第二经过时间,

基于测量出的所述第二经过时间,计算所述第一用户满意度以及所述第二用户满意

度，

比较所述第一用户满意度与所述第二用户满意度，基于该比较结果，选择所述第一语音识别信息或所述第二语音识别信息中的某一方。

5. 如权利要求1所述的语音识别装置，其中，

所述选择部进一步基于根据所述语音的属性预先确定的多个域中的与所述第一语音识别信息及所述第二语音识别信息分别对应的第一域及第二域、以及分别对应于所述第一语音识别信息及所述第二语音识别信息而求出的第一推断正确率及第二推断正确率的至少一个，计算所述第一用户满意度以及所述第二用户满意度。

6. 如权利要求5所述的语音识别装置，其中，

所述第一语音识别部以及所述第二语音识别部的至少一个使用多个词典数据的某一个来进行所述语音的识别，

所述选择部基于所述第一语音识别部以及所述第二语音识别部的至少一个在所述语音的识别中使用的词典数据，推断所述第一域以及所述第二域的至少一个。

7. 如权利要求5所述的语音识别装置，其中，

所述第一语音识别信息以及所述第二语音识别信息的至少一个包含意图推断信息，该意图推断信息表示针对所述语音的、所述用户的意图的推断结果，

所述选择部基于所述意图推断信息，推断所述第一域以及所述第二域的至少一个。

8. 如权利要求5所述的语音识别装置，其中，

所述选择部基于过去的所述第一域以及所述第二域的推断历史，推断所述第一域以及所述第二域。

9. 如权利要求5所述的语音识别装置，其中，

所述选择部基于所述第一域及所述第二域、针对所述第一语音识别信息的可靠度及针对所述第二语音识别信息可靠度、以及从所述语音被输入起的经过时间的至少一个，决定所述第一推断正确率以及所述第二推断正确率。

10. 如权利要求1所述的语音识别装置，其中

所述第一语音识别信息以及所述第二语音识别信息分别包含意图推断信息，该意图推断信息表示针对所述语音的、所述用户的意图的推断结果，

所述选择部选择所述第一语音识别信息或所述第二语音识别信息中的某一方所含的所述意图推断信息。

11. 如权利要求1所述的语音识别装置，其中，

所述选择部选择采用动作、确认动作、以及舍弃动作中的某一个动作，并对应于所选择的动作使所述信息提供的方法变化，

所述采用动作是，将基于所述第一语音识别信息或所述第二语音识别信息中的某一个的输入操作作为所述用户的输入操作来采用；

所述确认动作是，将基于所述第一语音识别信息或所述第二语音识别信息中的某一个的输入操作在所述用户确认之后作为所述用户的输入操作来采用，

所述舍弃动作是，不采用基于所述第一语音识别信息的输入操作以及基于所述第二语音识别信息的输入操作的任一个而是进行舍弃。

12. 一种语音识别系统，具备终端装置与服务器，其中，

所述终端装置具备：

语音检测部，检测从用户输入的语音；

第一语音识别部，执行用于识别所述语音的语音识别处理，并输出基于所述语音的识别结果的第一语音识别信息；

第一通信控制部，将基于所述语音的语音信息发送至所述服务器，并接收从所述服务器发送的第二语音识别信息；

信息提供部，使用所述第一语音识别信息或所述第二语音识别信息中的某一方进行向所述用户的信息提供；以及

选择部，基于从所述语音被输入起的经过时间，选择所述第一语音识别信息或所述第二语音识别信息中的某一方，并且使由所述信息提供部进行的所述信息提供的方法变化，

所述服务器具备：

第二通信控制部，接收从所述终端装置发送的所述语音信息，并将所述第二语音识别信息发送至所述终端装置；以及

第二语音识别部，基于所述语音信息执行用于识别所述语音的语音识别处理，并输出基于所述语音的识别结果的所述第二语音识别信息，

所述选择部基于从所述语音被输入起的经过时间，计算第一用户满意度与第二用户满意度，该第一用户满意度表示所述用户对使用了所述第一语音识别信息的情况下的所述信息提供的满意程度的预测值，该第二用户满意度表示所述用户对使用了所述第二语音识别信息的情况下的所述信息提供的满意程度的预测值，

比较所述第一用户满意度与所述第二用户满意度，基于该比较结果选择所述第一语音识别信息或所述第二语音识别信息中的某一方。

## 语音识别装置、语音识别系统

### 技术领域

[0001] 本发明涉及语音识别装置以及语音识别系统。

### 背景技术

[0002] 近年,在智能手机、汽车导航中,实现了使用语音识别引擎的语音输入功能。在专利文献1中公开了一种语音识别系统,其经由网络将车载装置与服务器连接,并使用分别搭载于车载装置与服务器的语音识别处理部,在车载装置与服务器中分散地进行语音识别处理。

[0003] 现有技术文献

[0004] 专利文献

[0005] 专利文献1:日本特开2013-88477号公报

### 发明内容

[0006] 发明要解决的课题

[0007] 专利文献1所公开的语音识别系统例如在像菜单画面那样从预先确定的选择项中用语音选择该项目的输入模式中,选择车载装置的语音识别处理部,在用语音输入任意的文本的输入模式中,选择服务器的语音识别处理部。由此分配了由哪个语音识别处理部来执行语音识别处理。然而,在这种语音识别处理的分配方法中,有时即使是输入了用车载装置的语音识别处理部就足够能识别的语音的情况,也会选择服务器的语音识别处理部,在响应上花费时间。因而,不一定能够实现对用户而言舒适的语音输入功能。

[0008] 用于解决课题的手段

[0009] 本发明的语音识别装置具备:语音检测部,检测从用户输入的语音;信息提供部,使用第一语音识别信息或第二语音识别信息中的某一方的语音识别信息来进行向所述用户的提供信息,该第一语音识别信息基于第一语音识别部对所述语音的识别结果,该第二语音识别信息基于与所述第一语音识别部不同的第二语音识别部对所述语音的识别结果;以及选择部,基于从所述语音被输入起的经过时间,选择所述第一语音识别信息或所述第二语音识别信息中的某一方作为所述信息提供部所使用语音识别信息,并且使由所述信息提供部进行的所述信息提供的方法变化。

[0010] 本发明的语音识别系统具备终端装置与服务器,其中,所述终端装置具备:语音检测部,检测从用户输入的语音;第一语音识别部,执行用于识别所述语音的语音识别处理,并输出基于所述语音的识别结果的第一语音识别信息;第一通信控制部,将基于所述语音的语音信息发送至所述服务器,并接收从所述服务器发送的第二语音识别信息;信息提供部,使用所述第一语音识别信息或所述第二语音识别信息中的某一方,进行向所述用户的提供信息;以及选择部,基于从所述语音被输入起的经过时间,选择所述第一语音识别信息或所述第二语音识别信息中的某一方,并且使由所述信息提供部进行的所述信息提供的方法变化,所述服务器具备:第二通信控制部,接收从所述终端装置发送的所述语音信息,并

将所述第二语音识别信息发送至所述终端装置；以及第二语音识别部，基于所述语音信息执行用于识别所述语音的语音识别处理，并输出基于所述语音的识别结果的所述第二语音识别信息。

[0011] 发明效果

[0012] 根据本发明，能够实现对用户而言舒适的语音输入功能。

### 附图说明

[0013] 图1是表示本发明的一实施方式的语音识别系统的构成的图。

[0014] 图2是表示在采用动作、确认动作、舍弃动作的各个中所显示的画面的例子的图。

[0015] 图3是在选择部中执行的处理的流程图。

[0016] 图4是表示结果为第一个新到达的判断例程的流程图。

[0017] 图5是表示结果为第二个新到达的判断例程的流程图。

[0018] 图6是表示结果为未到达时的判断例程的流程图。

[0019] 图7是表示结果为一个到达完毕的判断例程的流程图。

[0020] 图8是表示用于计算用户满意度的函数的例子的图。

[0021] 图9是表示用于计算用户满意度的函数的另一个例子的图。

### 具体实施方式

[0022] 以下，基于附图对本发明的实施方式进行说明。另外，在以下说明的实施方式中，以汽车导航系统为代表的车载终端为例，对本发明的语音识别装置进行说明。但是，本发明也能够应用于具有语音输入功能的其他装置、例如移动电话、智能手机、机器人等。

[0023] 图1是表示本发明的一实施方式的语音识别系统1的构成的图。图1所示的语音识别系统1通过经由通信网络300连接终端装置100与服务器200来构成。

[0024] 终端装置100例如被搭载于车辆，具备语音区间检测部1010、语音编码部1020、通信控制部1030、语音识别部1040、词典数据库1050、意图推断部1060、选择部1070、用户满意度存储部1080、HMI控制部1100、语音输出装置1110、录音语音合成部1120以及规则语音合成部1130。另外，终端装置100例如具备未图示的CPU、存储器、记录介质等，通过在CPU中执行存储于存储器、记录介质的规定的程序，能够实现语音区间检测部1010、语音编码部1020、通信控制部1030、语音识别部1040、意图推断部1060、选择部1070、HMI控制部1100、语音输出装置1110、录音语音合成部1120以及规则语音合成部1130的各功能。另外，能够使用记录介质内的规定的存储区域来实现词典数据库1050以及用户满意度存储部1080的各功能。

[0025] 语音区间检测部1010检测通过麦克风10从用户输入的语音。语音区间检测部1010例如从自麦克风10输入的语音信号提取用户实际说出的部分，对该部分的信号波形进行采样并语音数据化，从而检测用户的语音。

[0026] 语音编码部1020将由语音区间检测部1010检测到的用户的语音转换为用于向服务器200发送的数据形式。语音编码部1020例如通过对从语音区间检测部1010输出的语音数据进行规定的压缩处理来减少数据量，并生成适合向服务器200的发送的语音信息。

[0027] 通信控制部1030将基于由语音区间检测部1010检测到的用户的语音的语音信息、

即由语音编码部1020生成的语音信息发送至服务器200。另外,若通过在接收到语音信息的服务器200中像后述那样执行语音识别处理来从服务器200发送语音识别信息,则接收该语音识别信息并输出到选择部1070。通信控制部1030能够与使用移动电话网、因特网等公共通信线路构成的通信网络300连接,经由通信网络300与服务器200之间进行通信。

[0028] 语音识别部1040是终端装置100所具备的语音识别引擎,执行用于识别由语音区间检测部1010检测到的用户的语音的语音识别处理。在语音识别部1040执行的语音识别处理中,使用词典数据库1050。在词典数据库1050中,预先登记有多个词典数据。语音识别部1040通过使用该多个词典数据的某一个来确定与用户的语音对应的文本。具体而言,例如语音识别部1040根据使用终端装置100向用户提供的目录(content)的内容(例如,导航信息、音乐播放等)、搭载有终端装置100的车辆的状态等,从登记在词典数据库1050的多个词典数据中选择与用户说出的可能性高的语音的种类对应的词典数据。然后,使用选择的词典数据检索与从语音区间检测部1010输入的语音数据对应的文本,从而确定与用户的语音对应的文本。语音识别部1040的语音识别结果、即与由语音识别部1040确定的用户的语音对应的文本被输出至意图推断部1060以及选择部1070。

[0029] 意图推断部1060基于由语音识别部1040获得的语音识别结果来推断针对从麦克风10输入的语音的用户的意图。例如作为语音识别结果获得的文本是提示设施的检索的文本的话,则意图推断部1060推断为用户的意图是“设施检索”。同样地,能够推断例如“电话呼叫”、“音频播放”等用户的意图。表示意图推断部1060对用户的意图的推断结果的意图推断信息与语音识别部1040的语音识别结果(文本)一起作为终端装置100中的基于语音识别结果的语音识别信息被输入至选择部1070。以下,将该终端装置100中的基于语音识别结果的语音识别信息称为“终端语音识别信息”。另外,将由通信控制部1030接收的服务器200中的基于语音识别结果的语音识别信息称为“服务器语音识别信息”。另外,意图推断部1060并不是终端装置100中必须的构成,也可以省略。在该情况下,在终端装置100中不进行用户的意图推断。

[0030] 选择部1070选择从语音识别部1040以及意图推断部1060输入的终端语音识别信息和从通信控制部1030输入的服务器语音识别信息的某一方。此时,选择部1070基于与存储于用户满意度存储部1080的用户的满意度相关的数据,针对终端语音识别信息与服务器语音识别信息分别计算用户的满意度,并基于该计算结果选择某一方的语音识别信息。然后,将选择的语音识别信息的文本向HMI控制部1100输出,并且使用选择的语音识别信息控制HMI控制部1100的动作。另外,关于选择部1070的具体的语音识别信息的选择方法,以后进行详细说明。

[0031] HMI控制部1100基于选择部1070的控制,使用从选择部1070输出的文本控制进行向用户的信息提供的HMI(Human Machine Interface,人机界面)。具体而言,例如,对语音输出装置1110进行与文本相应的语音的输出指令、或对与终端装置100连接的显示器30输出与文本相应的图像、文字的显示指令,从而控制HMI并进行向用户的信息提供。

[0032] 语音输出装置1110根据来自HMI控制部1100的语音输出指令向扬声器20输出语音信号。语音输出装置1110能够使用预先保存于录音语音合成部1120的录音语音、由规则语音合成部1130生成的合成语音来生成任意的语音信号。若从HMI控制部1100输出语音输出指令,则在语音输出装置1110中生成与该语音输出指令相应的语音信号,并输出至扬声器

20。

[0033] 触摸面板40是与显示器30一体地构成的操作部件,检测用户用手触摸显示器30的画面时的操作信息并输出至HMI控制部1100。HMI控制部1100基于来自触摸面板40的操作信息对显示于显示器30的画面进行控制。

[0034] 服务器200设置于远离终端装置100的位置,具备通信控制部1210、语音对话控制部1220、对话场景(scenario)1230、语音解码部1240、语音识别部1250、词典数据库1260以及意图推断部1270。另外,服务器200例如具备未图示的CPU、存储器、记录介质等,通过在CPU中执行存储于存储器、记录介质的规定的程序,能够实现通信控制部1210、语音对话控制部1220、语音解码部1240、语音识别部1250以及意图推断部1270的各功能。另外,能够使用记录介质内的规定的存储区域来实现对话场景1230以及词典数据库1260的各功能。

[0035] 通信控制部1210接收从终端装置100发送的语音信息。另外,若语音识别部1250基于接收到的语音信息执行语音识别处理,且基于该结果的语音识别信息从语音识别部1250经由语音对话控制部1220被输入,则将该语音识别信息作为服务器语音识别信息发送至终端装置100。通信控制部1210能够与通信网络300连接,经由通信网络300与终端装置100之间进行通信。

[0036] 语音对话控制部1220使用对话场景1230进行终端装置100通过显示器30、扬声器20向用户提示的画面、语音引导等的控制。在对话场景1230中,例如组合记录有在各情景中应该向用户输出的画面、语音的信息、以及与用户可能对这些信息说出的语音对应的文本、意图。并且,作为与用户的语音相应的转变目标的信息,还记录有接下来应该输出的画面、语音的信息。语音对话控制部1220能够基于这些信息,经由通信控制部1210对终端装置100的动作进行控制。

[0037] 语音解码部1240基于从终端装置100发送的语音信息将用户的语音复原。语音解码部1240例如通过在终端装置100的语音编码部1020中将压缩后的语音数据解冻,从而经由麦克风10将已输入至终端装置100的用户的语音复原。

[0038] 语音识别部1250是服务器200所具备的语音识别引擎,执行用于识别基于从终端装置100接收到的语音信息的用户的语音的语音识别处理。在语音识别部1250执行的语音识别处理中,使用词典数据库1260。语音识别部1250通过使用词典数据库1260来检索与从语音解码部1240输入的语音数据对应的文本,从而确定与用户的语音对应的文本。另外,也可以与终端装置100的词典数据库1050相同,在词典数据库1260中预先登记多个词典数据,语音识别部1250使用该多个词典数据的某一个来确定与用户的语音对应的文本。语音识别部1250的语音识别结果、即与由语音识别部1250确定的用户的语音对应的文本被输出至意图推断部1270,并且经由语音对话控制部1220被输出至通信控制部1210。

[0039] 意图推断部1270基于由语音识别部1250获得的语音识别结果,来推断用户对输入至终端装置100的语音的意图。意图推断部1270能够通过终端装置100中的意图推断部1060相同的方法来推断用户的意图。表示意图推断部1270对用户的意图的推断结果的意图推断信息与语音识别部1250的语音识别结果(文本)一起作为服务器语音识别信息经由语音对话控制部1220输出至通信控制部1210。通信控制部1210将这样输入的服务器语音识别信息发送至终端装置100。另外,意图推断部1270在服务器200中不是必须的构成,也可以省略。在该情况下,在服务器200中不进行用户的意图推断。

[0040] 接下来,对由终端装置100进行的向用户的信息提供进行说明。语音识别系统1中的终端装置100像前述那样,在选择部1070中,针对终端语音识别信息与服务器语音识别信息分别计算用户的满意度,并基于该计算结果选择某一方的语音识别信息。然后,以使使用了选择的语音识别信息时的用户满意度成为最大的方式对HMI控制部1100的动作进行控制,并进行向用户的信息提供。例如,通过以进行采用动作、确认动作、舍弃动作这三种动作的某一个方式对HMI控制部1100进行控制,从而使向用户提供信息的方法变化,并使用户满意度达到最大。

[0041] 图2是表示在采用动作、确认动作、舍弃动作的各个中显示于显示器30的画面的例子的图。此时,假设用户例如说出了“横滨国际机场”这一语音。并且,假设用户对该语音的意图为“将横滨国际机场设定为目的地”这一输入操作。另外,语音识别系统1以识别用户说出的上述的语音并根据该识别结果如用户的意图那样进行信息提供的方式进行动作。但是,有时由于用户的语音所含的噪声较大等理由,在终端装置100的语音识别部1040、服务器200的语音识别部1250中,在语音识别中产生错误。因此,在语音识别系统1中,终端装置100进行上述三种动作的某一个,从而根据状况,在对用户取得了确认的基础上决定用户的输入操作的内容、或不采用获得的语音识别结果而暂时舍弃。

[0042] 图2(a)的画面2010是在对上述的语音进行了采用动作的情况下显示于显示器30的画面的例子。在该情况下,选择部1070判断为,语音识别系统1基于终端装置100或服务器200的语音识别结果确定的“将横滨国际机场设定为目的地”这一输入操作是如用户的意图那样的可靠性高。因此,将其作为用户的输入操作而立即采用,并反映在在终端装置100向用户提供的导航信息的处理中。

[0043] 图2(b)的画面2020以及图2(c)的画面2030是对上述的语音进行了确认动作的情况下显示于显示器30的画面的例子。在该情况下,选择部1070使用户确认语音识别系统1基于终端装置100或服务器200的语音识别结果确定的输入操作是否是如用户的意图那样。在图2(b)的画面2020中,仅显示有一个语音识别系统1识别出的输入操作的候补。用户用“是”与“否”这两个选择对该候补是不是如意图那样进行确认。另一方面,在图2(c)的画面2030中,显示有多个语音识别系统1识别出的输入操作的候补。用户通过从这些候补中选择某一个,来进行与意图一致的输入操作的确认。在用户的确认后,选择部1070将确认后的输入操作作为用户的输入操作而采用。

[0044] 图2(d)的画面2040是对上述的语音进行了舍弃动作的情况下显示于显示器30的画面的例子。在该情况下,选择部1070判断为语音识别系统1基于终端装置100或服务器200的语音识别结果确定的输入操作的候补不正确,不采用任何一个候补而进行舍弃。然后,通过促使用户再次发言,来谋求重新进行语音识别系统1中的语音识别,以获得正确的输入操作的候补。

[0045] 接下来,对终端装置100的选择部1070中的语音识别信息的选择方法进行说明。图3是在选择部1070中执行的处理的流程图。图3的流程图所示的处理在用户说出的语音从麦克风10输入至终端装置100时开始。另外,可以通过用户按压设于终端装置100的未图示的操作按钮,判断为用户开始发言,也可以设为不按压操作按钮就能够检测用户的发言。

[0046] 在步骤3010中,选择部1070待机至来自终端装置100的语音识别部1040的语音识别结果即终端语音识别信息、或来自服务器200的语音识别部1250的语音识别结果即服务

器语音识别信息到达、或者从处理的开始或前次的语音识别结果的到达起经过一定时间为止。另外,在此等待经过一定时间的理由是因为,在未获得语音识别结果的情况下定期地预测在该时刻的用户满意度,并根据该值来决定接下来的动作。语音识别结果到达、或经过一定时间后,解除待机而使处理进入接下来的步骤3020。

[0047] 在步骤3020中,选择部1070对在步骤3010中解除了待机的原因是语音识别结果的到达与一定时间的经过的哪个进行判断。其结果,在因语音识别结果的到达解除了待机的情况下使处理进入步骤3030,在因一定时间的经过解除了待机的情况下使处理进入步骤3060。

[0048] 在步骤3030中,选择部1070判定已到达的语音识别结果是否是第一个。在是第一个的情况下、即是最初到达的语音识别结果的情况下,进入步骤3040,在是第二个的情况下、即是第二个到达的语音识别结果的情况下,进入步骤3050。另外,在该步骤3030的处理中,与已到达的语音识别结果的输出源是终端装置100(语音识别部1040)与服务器200(语音识别部1250)的哪个无关。但是,通常由于通信的延迟等,来自终端装置100的语音识别结果先到达,来自服务器200的语音识别结果后到达。

[0049] 在步骤3040中,选择部1070进入结果为第一个新到达的判断例程。在此,按照图4所示的流程图,执行用于判断是使用最初到达的语音识别结果、还是等待接下来的语音识别结果的处理。另外,关于图4的流程图的详细情况以后进行说明。

[0050] 在步骤3050中,选择部1070进入结果为第二个新到达的判断例程。在此,按照图5所示的流程图,执行用于判断使用最初到达的语音识别结果与第二个到达的语音识别结果的哪个的处理。另外,关于图5的流程图的详细情况以后进行说明。

[0051] 在步骤3060中,选择部1070判定语音识别结果是否为未到达。在未到达的情况下、即没有从终端装置100(语音识别部1040)与服务器200(语音识别部1250)的任何一个接收到语音识别结果的情况下,进入步骤3070,在第一个语音识别结果到达完毕的情况下、即从终端装置100(语音识别部1040)与服务器200(语音识别部1250)的某一方接收完毕语音识别结果的情况下,进入步骤3080。

[0052] 在步骤3070中,选择部1070进入结果为未到达时的判断例程。在此,按照图6所示的流程图,执行用于判断是否等待至语音识别结果到达为止的处理。另外,关于图6的流程图的详细情况以后进行说明。

[0053] 在步骤3080中,选择部1070进入结果为第一个到达完毕的判断例程。在此,按照图7所示的流程图,执行用于判断是使用最初到达的语音识别结果、还是等待接下来的语音识别结果的处理。另外,关于图7的流程图的详细情况以后进行说明。

[0054] 在执行步骤3040、3050、3070或3080的某一个后,选择部1070执行步骤3090。在步骤3090中,选择部1070判断是否等待接下来的语音识别结果。在此,确认规定的标志状态。其结果,在标志建立的情况下,判断为等待接下来的语音识别结果,并返回至步骤3010而待机。另一方面,在标志未建立的情况下,使处理进入步骤3100。另外,步骤3090的判定中使用的标志,在步骤3040、3070或3080中满足规定的条件时被设置。关于该条件的详细情况,以后参照图4、6、7进行说明。

[0055] 在步骤3100中,选择部1070将进行由步骤3040、3050、3070、3080的某一个决定的动作的指令输出至HMI控制部1100。按照该指令,HMI控制部1100进行HMI的控制,使用在图2

中说明的那样的画面显示进行向用户的信息提供。即,在采用动作的情况下,如图2(a)的画面2010那样的画面显示于显示器30,在确认动作的情况下,如图2(b)的画面2020、图2(c)的画面2030那样的画面显示于显示器30,在舍弃动作的情况下,如图2(d)的画面2040那样的画面显示于显示器30。另外,终端装置100执行与这些画面对应的处理。

[0056] 一旦执行了步骤3100,选择部1070就结束图3的流程图所示的处理,并待机至接下来的语音输入为止。

[0057] 接下来,依次对图3的步骤3040、3050、3070、3080的各处理的详细情况进行说明。首先,参照图4的流程图,对在步骤3040中执行的结果为第一个新到达的判断例程进行说明。

[0058] 在步骤4010中,选择部1070接收已到达的语音识别结果,并对该语音识别结果设定识别符号i1。该识别符号i1由表示语音识别结果的符号即“i”和表示第一个语音识别结果的数字即“1”组合而构成。以下,将设定有识别符号i1的第一个语音识别结果称为“语音识别结果i1”。

[0059] 在步骤4020中,选择部1070判断语音识别结果i1的输出源是终端装置100与服务器200的哪个。在输出源为服务器200的情况下、即语音识别结果i1为服务器语音识别信息的情况下,使处理进入步骤4030,在输出源为终端装置100的情况下、即语音识别结果i1为终端语音识别信息的情况下,使处理进入步骤4040。

[0060] 在步骤4030中,选择部1070设定第一个语音识别结果的输出源k1以及第二个语音识别结果的输出源k2。在此,对第一个语音识别结果即语音识别结果i1的输出源k1设定表示服务器200的“服务器”,对未接收的第二个语音识别结果的输出源k2设定表示终端装置100的“终端”。

[0061] 在步骤4040中,选择部1070设定第一个语音识别结果的输出源k1以及第二个语音识别结果的输出源k2。在此,与步骤4030相反,对第一个语音识别结果即语音识别结果i1的输出源k1设定表示终端装置100的“终端”,对未接收的第二个语音识别结果的输出源k2设定表示服务器200的“服务器”。

[0062] 一旦在步骤4030或4040中设定了识别符号k1、k2,在步骤4050中,选择部1070推断与语音识别结果i1对应的域d1。域是指表示输入到用户的语音的属性的信息。在终端装置100中,例如根据由对应的输入操作执行的任务的难度、向用户提供的信息的种类等将语音分类为多个属性,对应于该多个属性而预先设定有多个域。例如,对于在提供导航信息时输入的语音,设定有“住址检索”、“设施名检索”、“导航命令”这三种域。选择部1070推断在终端装置100中设定的这些域中的哪个域是与语音识别结果i1对应的域,并基于该推断结果设定域d1。此时,也可以根据语音识别部1040或1250使用词典数据库1050或1230中预先登记的多个词典中的哪个词典进行了语音识别来推断域d1。另外,在作为语音识别结果i1取得的终端语音识别信息或服务器语音识别信息中包含表示意图推断部1060或1270的用户的意图的推断结果的意图推断信息的情况下,也可以基于该意图推断信息推断域d1。另外,在无法推断与语音识别结果i1对应的域的情况下,将域d1设定为“未知”。

[0063] 在步骤4060中,选择部1070推断语音识别结果i1的可靠度c1。可靠度是指表示语音识别结果的可靠性的指标,能够使用各种指标。

[0064] 在步骤4070中,选择部1070将从用户结束发言而终端装置100中被输入语音起到

当前为止的、即在终端装置100中获得语音识别结果i1为止的经过时间作为经过时间t1进行测量。另外,只要与从终端装置100中被输入语音起到获得语音识别结果i1为止的经过时间相关,则也可以把其他时间指标作为经过时间t1进行测量。例如,可以从用户已开始发言的时刻起开始经过时间t1的测量,也可以预测从接收语音识别结果i1起到画面显示于显示器30为止的时间,并将该时间添加到经过时间t1。

[0065] 在步骤4080中,选择部1070计算语音识别结果i1的推断正确率p1。在此,基于在步骤4030或4040中设定的输出源k1、在步骤4050中推断出的域d1、在步骤4060中推断出的可靠度c1、以及在步骤4070测量出的经过时间t1,使用预先设定的函数Fp来计算推断正确率p1。推断正确率p1是指,表示语音识别结果i1以何种程度的概率与用户的语音一致的值。即,语音识别结果i1与用户的语音一致的概率不仅根据可靠度c1,有时也根据域d1、输出源k1而发生变化。另外,有时由经验可知,根据在终端装置100的语音识别部1040、服务器200的语音识别部1250中的语音识别处理所需的时间,语音识别结果i1与用户的语音一致的概率发生变化。因此,通过使用这样的经验上的知识等来预先调查k1、d1、c1、t1的各变量与推断正确率p1的相关关系,并基于该结果设定函数Fp。然后,通过对该函数Fp输入在步骤4030~4070的各处理中设定的k1、d1、c1、t1的各变量的值,能够计算推断正确率p1。其中,在步骤4080中,k1、d1、c1、t1的变量未必需要全部使用。通过使用任意的一个以上的变量,能够进行推断正确率p1的计算。

[0066] 在步骤4090中,选择部1070计算对使用了语音识别结果i1的情况下的信息提供的用户满意度。在此,针对HMI控制部1100进行向用户的信息提供时的前述的三种动作、即采用动作、确认动作、舍弃动作的各个,计算对语音识别结果i1的用户满意度。以下,将采用动作中的用户满意度表示为Sa1,将确认动作中的用户满意度表示为Sc1,将舍弃动作中的用户满意度表示为Sr1。

[0067] 上述的用户满意度Sa1、Sc1、Sr1能够基于在步骤4050中推断出的域d1、在步骤4070中测量出的经过时间t1、以及在步骤4080中计算出的推断正确率p1而分别进行计算。例如,通过将d1、t1以及p1的变量代入按每个动作的种类而预先设定的函数Fsa、Fsc、Fsr,能够分别计算用户满意度Sa1、Sc1、Sr1。这些函数Fsa、Fsc、Fsr使用经验上的知识等预先设定,并作为与用户的满意度相关的数据储存于用户满意度存储部1080。其中,在步骤4090中,d1、t1、p1的变量未必需要全部使用。通过使用任意的一个以上的变量,能够进行用户满意度Sa1、Sc1、Sr1的计算。

[0068] 图8是表示用于计算用户满意度Sa1、Sc1、Sr1的函数Fsa、Fsc、Fsr的例子的图。图8(a)所示的图表8010、8011、8012是用于计算采用动作中的用户满意度Sa1的函数Fsa的例子。图表8010作为函数Fsa的例子,示出了推断正确率p1为90%时的经过时间t1与用户满意度Sa1的关系。图表8011作为函数Fsa的例子,示出了推断正确率p1为80%时的经过时间t1与用户满意度Sa1的关系。图表8012作为函数Fsa的例子,示出了推断正确率p1为60%时的经过时间t1与用户满意度Sa1的关系。在这些图表中,用户满意度Sa1的最低值为1、最高值为5。

[0069] 图8(b)所示的图表8020、8021、8022是用于计算确认动作中的用户满意度Sc1的函数Fsc的例子。图表8020作为函数Fsc的例子,示出了推断正确率p1为90%时的经过时间t1与用户满意度Sc1的关系。图表8021作为函数Fsc的例子,示出了推断正确率p1为80%时的

经过时间 $t_1$ 与用户满意度 $Sc_1$ 的关系。图表8022作为函数 $F_{sc}$ 的例子,示出了推断正确率 $p_1$ 为60%时的经过时间 $t_1$ 与用户满意度 $Sc_1$ 的关系。另外,图8(b)所示的图表8023是用于计算舍弃动作中的用户满意度 $Sr_1$ 的函数 $F_{sr}$ 的例子,示出了经过时间 $t_1$ 与用户满意度 $Sr_1$ 的关系。在这些图表中,用户满意度 $Sc_1$ 、 $Sr_1$ 的最低值为1、最高值为5。

[0070] 另外,图8(a)以及图8(b)的各图表示出了域 $d_1$ 例如为“导航命令”的情况下的例子。

[0071] 作为用户满意度的一般的特性,已知从用户发言起到终端装置100进行某种动作为止的时间越长则用户满意度越低。因此,在图8(a)以及图8(b)的各图表中,随着经过时间 $t_1$ 增加,用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 分别逐渐地降低。

[0072] 在此,在图8(a)以及图8(b)的各图表中可知,在经过时间 $t_1$ 的值相对较小时,采用动作中的用户满意度 $Sa_1$ 最高,确认动作中的用户满意度 $Sc_1$ 第二高,舍弃动作中的用户满意度 $Sr_1$ 最低。另外,可知,随着经过时间 $t_1$ 的值增大,采用动作中的用户满意度 $Sa_1$ 大幅降低,另一方面,确认动作中的用户满意度 $Sc_1$ 、舍弃动作中的用户满意度 $Sr_1$ 不太降低。进而,可知与经过时间 $t_1$ 的增加对应的采用动作中的用户满意度 $Sa_1$ 的降低在推断正确率 $p_1$ 的值越大时越显著。关于这种由向用户的信息提供方法的不同而导致的用户满意度的变化趋势的差异,以下进行说明。

[0073] 一般来说,若推断正确率 $p_1$ 的值足够高,则在尽可能早的时机进行采用动作而采用正确的语音识别结果的用户满意度更高。另一方面,若在推断正确率 $p_1$ 的值较低的情况下进行采用动作,则采用错误的语音识别结果、在终端装置100中进行与用户的意图不同的输入操作的风险升高。由此,在推断正确率 $p_1$ 的值较低的情况下,通过进行确认动作而向用户提示语音识别结果表示的一个或多个选择项,相比进行了采用动作的情况更有可能能够执行如用户的意图那样的输入操作,因此平均来看用户满意度提高。另外,在经过时间 $t_1$ 增长、并且推断正确率 $p_1$ 较低的情况下,若通过进行采用动作而采用错误的语音识别结果,则用户满意度变得极低。因而,在这种情况下也预测为,相比采用动作,进行确认动作能够获得更高的用户满意度。

[0074] 而且,在进行了舍弃动作的情况下,总体上用户满意度降低。但是,在经过时间 $t_1$ 较长的情况下,通过进行舍弃动作而促使用户再次发言,能够对用户而言容易理解地传递语音识别未能顺利进行。另外,在这种情况下,能够期待用户比前次更清楚的发言,因此能够获得正确的语音识别结果的可能性提高。这样,在经过时间 $t_1$ 较长的情况下,有时相比采用动作、确认动作,进行舍弃动作的用户满意度更高。

[0075] 图9是表示在域 $d_1$ 与图8不同的情况下,用于计算用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 的函数 $F_{sa}$ 、 $F_{sc}$ 、 $F_{sr}$ 的例子。图9(a)所示的图表9010、9011、9012是用于计算采用动作中的用户满意度 $Sa_1$ 的函数 $F_{sa}$ 的例子。图表9010作为函数 $F_{sa}$ 的例子,示出了推断正确率 $p_1$ 为90%时的经过时间 $t_1$ 与用户满意度 $Sa_1$ 的关系。图表9011作为函数 $F_{sa}$ 的例子,示出了推断正确率 $p_1$ 为80%时的经过时间 $t_1$ 与用户满意度 $Sa_1$ 的关系。图表9012作为函数 $F_{sa}$ 的例子,示出了推断正确率 $p_1$ 为60%时的经过时间 $t_1$ 与用户满意度 $Sa_1$ 的关系。在这些图表中,用户满意度 $Sa_1$ 的最低值为1、最高值为5。

[0076] 图9(b)所示的图表9020、9021、9022是用于计算确认动作中的用户满意度 $Sc_1$ 的函数 $F_{sc}$ 的例子。图表9020作为函数 $F_{sc}$ 的例,示出了推断正确率 $p_1$ 为90%时的经过时间 $t_1$ 与

用户满意度 $Sc1$ 的关系。图表9021作为函数 $Fsc$ 的例子,示出了推断正确率 $p1$ 为80%时的经过时间 $t1$ 与用户满意度 $Sc1$ 的关系。图表9022作为函数 $Fsc$ 的例子,示出了推断正确率 $p1$ 为60%时的经过时间 $t1$ 与用户满意度 $Sc1$ 的关系。另外,图9(b)所示的图表9023是用于计算舍弃动作中的用户满意度 $Sr1$ 的函数 $Fsr$ 的例子,示出了经过时间 $t1$ 与用户满意度 $Sr1$ 的关系。在这些图表中,用户满意度 $Sc1$ 、 $Sr1$ 的最低值为1、最高值为5。

[0077] 另外,图9(a)以及图9(b)的各图表示出了域 $d1$ 例如为“住址检索”或“设施名检索”的情况下的例子。在这些图表中,与图8(a)、图8(b)所示的域 $d1$ 为“导航命令”的情况下的各图表相比,虽然经过时间 $t1$ 增长,但用户满意度 $Sa1$ 、 $Sc1$ 的降低幅度变小。关于这种由域 $d1$ 的不同而导致的用户满意度的变化趋势的差异,以下进行说明。

[0078] 一般来说,像域为“导航命令”的情况那样,在用户发出的语音的种类被相对限制的情况下,即使在不具备与服务器的连接功能而仅通过车辆的终端装置进行语音识别的以往的语音识别装置中也能够以相对较短的响应时间获得正确的语音识别结果。因此,对迄今为止使用这样的以往的语音识别装置的用户而言,对经过时间 $t1$ 增长的抵触感较强,用户满意度开始降低的经过时间 $t1$ 缩短。另一方面,像域为“设施名检索”、“住址检索”的情况那样,在用户发出的语音的种类涉及多个方面的情况下,在以往的语音识别装置中响应时间花费较长、或若用户不准确地说出设施名、住址名则无法进行语音识别。因此,在这种以往中语音识别为困难的状况的情况下,即使经过时间 $t1$ 增长,用户满意度的降低程度也相对减少。

[0079] 图8以及图9的各图表所示的用户满意度的函数 $Fsa$ 、 $Fsc$ 、 $Fsr$ 是考虑如以上说明那样的由域 $d1$ 的不同而导致的用户满意度的变化趋势的差异而在终端装置100中预先设定的函数。而且,也考虑了前述那样的由向用户的信息提供方法的不同而导致的用户满意度的变化趋势的差异。另外,在函数 $Fsa$ 、 $Fsc$ 、 $Fsr$ 的设计中,例如能够应用基于实际进行对受检者的实验而得的结果来决定各函数的方法、基于规定的感官评价来决定各函数的方法、按照规定的设计策略来决定各函数的方法等。

[0080] 另外,在上述的说明中,与是语音识别引擎的种类、即终端装置100的语音识别部1040或服务器200的语音识别部1250的哪个的语音识别结果无关,使用相同的函数 $Fsa$ 、 $Fsc$ 、 $Fsr$ 来求出用户满意度 $Sa1$ 、 $Sc1$ 、 $Sr1$ 。但是,由于根据语音识别引擎与域 $d1$ 的组合,推断正确率 $p1$ 大幅改变,因此与其相应地用户满意度也可能大幅改变。因此,也可以根据语音识别引擎的种类而在用户满意度存储部1080中预先设定各自的函数 $Fsa$ 、 $Fsc$ 、 $Fsr$ 的数据,区分使用它们而求出用户满意度 $Sa1$ 、 $Sc1$ 、 $Sr1$ 。

[0081] 若返回到图4的说明,则在步骤4095中,选择部1070推断第二个引擎的域、即与未接收的第二个语音识别结果对应的域 $d2'$ 。在此,通常的话,认为域 $d2'$ 与第一个引擎的域、即与语音识别结果 $i1$ 对应的域 $d1$ 相同的可能性较高。因而,推断为域 $d2'$ 与域 $d1$ 相同,并原样地设定与域 $d1$ 相同的内容即可。

[0082] 在步骤4100中,选择部1070预测从用户结束发言而终端装置100中被输入语音起到获得未接收的第二个语音识别结果为止的经过时间 $t2$ 。在此,基于在步骤4030或4040设定的输出源 $k2$ 、在步骤4095推断出的域 $d2'$ 、在步骤4060推断出的可靠度 $c1$ 、以及表示终端装置100与服务器200之间的通信状态的通信状态 $cs$ ,使用预先设定的函数 $Ft$ 来计算经过时间 $t2$ 的预测值。即,由于在输出源 $k2$ 为终端装置100的情况与出源 $k2$ 为服务器200的情况下,

至第二个语音识别结果到达为止的时间发生改变,因此认为经过时间 $t_2$ 发生变化。另外,由于根据与第二个语音识别结果对应的域 $d_2'$ 、语音识别所需的时间也不同,因此认为经过时间 $t_2$ 发生变化。进而,在输出源 $k_2$ 为服务器200的情况下,通信状态 $cs$ 也成为较大地左右经过时间 $t_2$ 的重要因素。除此之外,关于语音识别结果 $i_1$ 的可靠度 $c_1$ ,也存在能够用于预测由其它的语音识别引擎进行的语音识别的困难性的可能性。因此,预先调查 $k_2$ 、 $d_2'$ 、 $c_1$ 、 $cs$ 的各变量与经过时间 $t_2$ 的相关关系,并基于该结果来设定函数 $F_t$ 。然后,通过对该函数 $F_t$ 输入在先前的各处理中设定的 $k_2$ 、 $d_2'$ 、 $c_1$ 、 $cs$ 的各变量的值,能够推断经过时间 $t_2$ 。另外,函数 $F_t$ 可以由数式表示,也可以作为基于预先收集到的数据而设定的每个变量的表值。另外,与前述的经过时间 $t_1$ 相同,只要与从终端装置100被输入语音起到获得第二个语音识别结果为止的经过时间相关,则也可以将其他时间指标作为经过时间 $t_2$ 进行预测。其中,在步骤4100中, $k_2$ 、 $d_2'$ 、 $c_1$ 、 $cs$ 的变量未必需要全部使用。通过使用任意的一个以上的变量,能够进行经过时间 $t_2$ 的推断。

[0083] 在步骤4110中,选择部1070计算第二个语音识别结果的推断正确率 $p_2$ 。在此,使用在当前时刻能够利用的全部信息来计算推断正确率 $p_2$ 。即,基于在步骤4030或4040中设定的输出源 $k_1$ 以及 $k_2$ 、在步骤4095中推断出的域 $d_2'$ 、在步骤4060中推断出的可靠度 $c_1$ 、以及在步骤4100中预测到的经过时间 $t_2$ ,使用预先设定的函数 $F_{p2}$ 来计算推断正确率 $p_2$ 。另外,函数 $F_{p2}$ 与在步骤4080中用于计算语音识别结果 $i_1$ 的推断正确率 $p_1$ 的函数 $F_p$ 相同,能够预先设定。其中,在步骤4110中, $k_1$ 、 $k_2$ 、 $d_2'$ 、 $c_1$ 、 $t_2$ 的变量未必需要全部使用。通过使用任意的一个以上的变量,能够进行推断正确率 $p_2$ 的计算。

[0084] 在步骤4120中,选择部1070计算对使用了第二个语音识别结果的情况下的信息提供的用户满意度。在此,针对采用动作、确认动作的各个,通过与步骤4090相同的方法来计算对第二个语音识别结果的用户满意度 $Sa_2'$ 、 $Sc_2'$ 。即,基于在步骤4095中推断出的域 $d_2'$ 、在步骤4100中预测出的经过时间 $t_2$ 、以及在步骤4110中计算出的推断正确率 $p_2$ ,并将这些变量代入前述的函数 $F_{sa}$ 、 $F_{sc}$ 中,从而分别计算用户满意度 $Sa_2'$ 、 $Sc_2'$ 。其中,在步骤4120中, $d_2'$ 、 $t_2$ 、 $p_2$ 的变量未必需要全部使用。另外,在此,也可以不对舍弃动作中的用户满意度 $Sr_2'$ 进行计算。其理由是因为,在特意等待第二个语音识别结果之后判断为最好舍弃该语音识别结果的情况在现实中不存在。

[0085] 在步骤4130以后的处理中,选择部1070将对在步骤4090中计算出的语音识别结果 $i_1$ 的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 与对在步骤4120中计算出的第二个语音识别结果的用户满意度 $Sa_2'$ 、 $Sc_2'$ 进行比较。基于该比较结果,选择部1070判断是选择语音识别结果 $i_1$ 作为由HMI控制部1100进行的向用户的信息提供中所使用的语音识别结果、还是选择不选择语音识别结果 $i_1$ 而等待至接收第二个语音识别结果为止。另外,在选择语音识别结果 $i_1$ 的情况下,判断使用采用动作、确认动作、舍弃动作中的哪个来进行向用户的信息提供。

[0086] 在步骤4130中,选择部1070判定相比使用语音识别结果 $i_1$ ,是否等待第二个语音识别结果的用户满意度更高。具体而言,选择部1070将对语音识别结果 $i_1$ 的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 的最大值与对第二个语音识别结果的用户满意度 $Sa_2'$ 、 $Sc_2'$ 的最大值进行比较。其结果,若用户满意度 $Sa_2'$ 、 $Sc_2'$ 的最大值比用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 的最大值大,则判断为等待第二个语音识别结果的用户满意度更高,并使处理进入步骤4170。另一方面,若用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 的最大值为用户满意度 $Sa_2'$ 、 $Sc_2'$ 的最大值以上,则判断为不等待第

二个语音识别结果而使用语音识别结果i1的用户满意度更高,并使处理进入步骤4140。

[0087] 在步骤4140中,选择部1070将基于语音识别结果i1的采用动作中的用户满意度Sa1与确认动作中的用户满意度Sc1以及舍弃动作中的用户满意度Sr1进行比较。其结果,若Sa1为Sc1以上且Sr1以上,则使处理进入步骤4180,在不是的情况下,使处理进入步骤4150。

[0088] 在步骤4150中,选择部1070将基于语音识别结果i1的确认动作中的用户满意度Sc1与舍弃动作中的用户满意度Sr1进行比较。其结果,若Sc1为Sr1以上,则使处理进入步骤4160,若为不是的情况、即Sc1小于Sr1,则使处理进入步骤4190。

[0089] 在步骤4160中,选择部1070选择舍弃动作作为基于语音识别结果i1的向用户提供信息的方法。之后,选择部1070结束图4的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,进行舍弃动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(d)的画面2040那样的画面显示于显示器30,催促用户再次的发言。

[0090] 在步骤4170中,选择部1070决定为等待第二个语音识别结果,并进行建立前述的标志的处理。之后,选择部1070结束图4的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为肯定之后,返回至步骤3010,待机至第二个语音识别结果被接收、或经过一定时间为止。

[0091] 在步骤4180中,选择部1070选择采用动作作为基于语音识别结果i1的向用户提供信息的方法。之后,选择部1070结束图4的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果i1进行采用动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(a)的画面2010那样的画面显示于显示器30,并移至与识别出的输入操作相应的处理。

[0092] 在步骤4190中,选择部1070选择确认动作作为基于语音识别结果i1的向用户提供信息的方法。之后,选择部1070结束图4的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果i1进行确认动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(b)的画面2020或图2(c)的画面2030那样的画面显示于显示器30,在等待用户的确认之后,移至与经确认的输入操作相应的处理。

[0093] 如以上说明那样,在步骤4130以后的处理中,将对语音识别结果i1的用户满意度Sa1、Sc1、Sr1与对第二个语音识别结果的用户满意度Sa2'、Sc2' 进行比较,基于该比较结果选择用户满意度变得最高那样的动作。关于该点,以下列举具体例而进一步进行说明。

[0094] 另外,在以下说明的具体例中,作为前提条件,假设最初来自终端装置100的语音识别结果到达选择部1070( $k1 = \text{“终端”}$ ),且此时的经过时间 $t1$ 为3秒。另外,假设语音识别结果i1为“日本武道馆”,且与该识别结果对应的输入操作为设施名检索。因此,域 $d1$ 的推断结果为“设施名检索”,针对其的推断正确率 $p1$ 的计算结果为80%。

[0095] 在步骤4100、4110中,基于针对语音识别结果i1的上述的各信息,分别计算与未接收的第二个语音识别结果( $k2 = \text{“服务器”}$ )相关的经过时间 $t2$ 与推断正确率 $p2$ 。其结果,作为前提条件,经过时间 $t2$ 被预测为7秒,推断正确率 $p2$ 为90%。另外,在域 $d2'$ 中设定与域 $d1$ 相同的“设施名检索”。

[0096] 基于上述的前提条件,使用最初的语音识别结果或第二个语音识别结果进行了由

采用动作、确认动作、舍弃动作的各个进行的对用户的信息提供的情况下的用户满意度,例如可像以下那样求出。另外,在以下的各用户满意度的计算中,使用图9所示的函数 $F_{sa}$ 、 $F_{sc}$ 、 $F_{sr}$ 。

[0097] 根据图9的图表9011、9021、9023,使用了最初的语音识别结果情况下的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 可分别像以下那样求出。

[0098] 采用动作: $Sa_1=4.0$

[0099] 确认动作: $Sc_1=3.8$

[0100] 舍弃动作: $Sr_1=1.8$

[0101] 另外,根据图9的图表9010、9020、9023,使用了第二个语音识别结果情况下的用户满意度 $Sa_2'$ 、 $Sc_2'$ 可分别像以下那样求出。

[0102] 采用动作: $Sa_2'=4.2$

[0103] 确认动作: $Sc_2'=4.3$

[0104] 若比较上述的各用户满意度,则使用第二个语音识别结果而进行了确认动作的情况下的用户满意度 $Sc_2'$ 最高。因此,在该具体例中,不使用最初的语音识别结果而等待第二个语音识别结果。

[0105] 另外,在上述的具体例中,虽然对用户满意度存储部1080中预先存储的函数 $F_{sa}$ 、 $F_{sc}$ 、 $F_{sr}$ 的数据与推断正确率 $p_1$ 、 $p_2$ 分别一致的情况进行了说明,但即使在不一致情况下也用能够进行用户满意度的计算。例如,推断正确率 $p_1$ 、 $p_2$ 为85%的情况下的采用动作中的用户满意度 $Sa_1$ 、 $Sa_2'$ ,可通过从图9的图表9010、9011分别读取用户满意度并计算这些用户满意度的平均值来求出。同样,确认动作中的用户满意度 $Sc_1$ 、 $Sc_2'$ 可通过从图9的图表9020、9021分别读取用户满意度并计算这些用户满意度的平均值来求出。即,在与推断正确率 $p_1$ 、 $p_2$ 一致的用户满意度的数据未存储于用户满意度存储部1080的情况下,通过对其前后的用户满意度的数据进行线性插补,能够求出用户满意度。另外,例如在推断正确率为100%的情况等、在前后不存在用户满意度的数据而无法进行线性插补的情况下,也能够使用最接近其推断正确率的用户满意度的数据来求出用户满意度。

[0106] 接下来,参照图5的流程图,对在步骤3050中执行的结果为第二个新到达的判断例程进行说明。

[0107] 在步骤5010中,选择部1070接收第二个到达的语音识别结果,并对该语音识别结果设定识别符号 $i_2$ 。以下,将设定有识别符号 $i_2$ 的第二个语音识别结果称为“语音识别结果 $i_2$ ”。

[0108] 在步骤5020中,选择部1070推断与语音识别结果 $i_2$ 对应的域 $d_2$ 。另外,域 $d_2$ 的推断方法与图4的步骤4050中的域 $d_1$ 的推断方法相同。

[0109] 在步骤5030中,选择部1070推断语音识别结果 $i_2$ 的可靠度 $c_2$ 。

[0110] 在步骤5040中,选择部1070测量从用户结束发言而终端装置100中被输入语音起到当前为止的、即在终端装置100中获得语音识别结果 $i_2$ 为止的经过时间 $t_2$ 。另外,与前述的经过时间 $t_1$ 相同,只要与从终端装置100中被输入语音起到获得语音识别结果 $i_2$ 为止的经过时间相关,则也可以将其他时间指标作为经过时间 $t_2$ 进行测量。

[0111] 在步骤5050中,选择部1070计算语音识别结果 $i_2$ 的推断正确率 $p_2$ 。在此,与图4的步骤4080中计算出的推断正确率 $p_1$ 相同,基于在步骤4030或4040中设定的输出源 $k_2$ 、在步

骤5020中推断出的域 $d_2$ 、在步骤5030中推断出的可靠度 $c_2$ 、在步骤5040中测量出的经过时间 $t_2$ ，使用预先设定的函数 $F_p$ 来计算推断正确率 $p_2$ 。其中，在步骤5050中， $k_2$ 、 $d_2$ 、 $c_2$ 、 $t_2$ 的变量未必需要全部使用。通过使用任意的一个以上的变量，能够进行推断正确率 $p_2$ 的计算。

[0112] 在步骤5060中，选择部1070判断当前的模式为最大满意度模式与最大正确率模式的哪个。其结果，在当前的模式为最大满意度模式的情况下，使处理进入步骤5160，在为最大正确率模式的情况下，使处理进入步骤5070。另外，最大满意度模式是指，以使用户满意度达到最大的方式选择语音识别结果以及HMI控制部1100的动作而进行向用户的信息提供的模式，最大正确率模式是指，使用推断正确率 $p_1$ 或 $p_2$ 达到最大的语音识别结果而进行向用户的信息提供的模式。

[0113] 在当前的模式为最大正确率模式的情况下，在步骤5070中，选择部1070将语音识别结果 $i_1$ 的推断正确率 $p_1$ 与语音识别结果 $i_2$ 的推断正确率 $p_2$ 进行比较。其结果，在推断正确率 $p_1$ 更高的情况下，使处理进入步骤5080，在推断正确率 $p_2$ 更高的情况、或推断正确率 $p_1$ 与推断正确率 $p_2$ 相同的情况下，使处理进入步骤5090。

[0114] 在步骤5080中，选择部1070将语音识别结果 $i_1$ 、与其对应的推断正确率 $p_1$ 、域 $d_1$ 以及可靠度 $c_1$ 分别设定为最大似然结果。然后，为了作为在接下来的步骤5100中计算用户满意度时的变量进行参照，分别设定为 $i=i_1$ 、 $p=p_1$ 、 $d=d_1$ 、 $c=c_1$ 。

[0115] 在步骤5090中，选择部1070将语音识别结果 $i_2$ 、与其对应的推断正确率 $p_2$ 、域 $d_2$ 以及可靠度 $c_2$ 分别设定为最大似然结果。然后，为了作为在接下来的步骤5100中计算用户满意度时的变量进行参照，分别设定为 $i=i_2$ 、 $p=p_2$ 、 $d=d_2$ 、 $c=c_2$ 。

[0116] 在步骤5100中，选择部1070计算对使用了在步骤5080或5090中设定为最大似然结果的语音识别结果 $i$  ( $i=i_1$ 或 $i_2$ ) 的情况下的信息提供的用户满意度 $S_a$ 、 $S_c$ 、 $S_r$ 。在此，与在图4的步骤4090中计算出的用户满意度 $S_{a1}$ 、 $S_{c1}$ 、 $S_{r1}$ 相同，能够通过将在步骤5090中设定的 $d$ 、 $t$ 以及 $p$ 的变量代入函数 $F_{sa}$ 、 $F_{sc}$ 、 $F_{sr}$ ，来分别计算使用了最大似然结果的情况下的基于采用动作、确认动作、舍弃动作的用户满意度 $S_a$ 、 $S_c$ 、 $S_r$ 。其中，在步骤5100中， $d$ 、 $t$ 、 $p$ 的变量未必需要全部使用。通过使用任意的一个以上的变量，能够进行用户满意度 $S_a$ 、 $S_c$ 、 $S_r$ 的计算。

[0117] 在步骤5110中，选择部1070将基于语音识别结果 $i$ 的采用动作中的用户满意度 $S_a$ 与确认动作中的用户满意度 $S_c$ 以及舍弃动作中的用户满意度 $S_r$ 进行比较。其结果，若 $S_a$ 为 $S_c$ 以上且 $S_r$ 以上，则使处理进入步骤5120，在不是的情况下，使处理进入步骤5130。

[0118] 在步骤5120中，选择部1070选择采用动作作为基于语音识别结果 $i$ 的向用户提供信息的方法。之后，选择部1070结束图5的流程图所示的处理，使处理进入图3的步骤3090。在该情况下，在步骤3090被判定为否定之后，在步骤3100中，基于语音识别结果 $i$ 进行采用动作的指令被从选择部1070输出至HMI控制部1100。其结果，如图2(a)的画面2010那样的画面显示于显示器30，移至与识别到的输入操作相应的处理。

[0119] 在步骤5130中，选择部1070将基于语音识别结果 $i$ 的确认动作中的用户满意度 $S_c$ 与舍弃动作中的用户满意度 $S_r$ 进行比较。其结果，若 $S_c$ 为 $S_r$ 以上，则使处理进入步骤5140，若为不是的情况、即 $S_c$ 小于 $S_r$ ，则使处理进入步骤5150。

[0120] 在步骤5140中，选择部1070选择确认动作作为基于语音识别结果 $i$ 的向用户提供信息的方法。之后，选择部1070结束图5的流程图所示的处理，使处理进入图3的步骤3090。

在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果*i*进行确认动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(b)的画面2020或图2(c)的画面2030那样的画面显示于显示器30,在待用户的确认之后,移至与经确认的输入操作相应的处理。

[0121] 在步骤5150中,选择部1070选择舍弃动作作为基于语音识别结果*i*的向用户提供信息的方法。之后,选择部1070结束图5的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,进行舍弃动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(d)的画面2040那样的画面显示于显示器30,催促用户再次的发言。

[0122] 如以上说明那样,在最大正确率模式下执行的步骤5070~5150的处理中,针对最初接收到的语音识别结果*i*<sub>1</sub>与第二个接收到的语音识别结果*i*<sub>2</sub>中的、推断正确率更高的语音识别结果,从采用动作、确认动作、舍弃动作的中选择并执行用户满意度变得最高的动作。

[0123] 在当前的模式为最大满意度模式的情况下,在步骤5160中,选择部1070计算对使用了最初接收到的语音识别结果*i*<sub>1</sub>的情况下的信息提供的用户满意度*Sa*<sub>1</sub>、*Sc*<sub>1</sub>、*Sr*<sub>1</sub>。在此,与图4的步骤4090相同,能够通过将在步骤4050中推断出的域*d*<sub>1</sub>、在步骤5040中测量出的经过时间*t*<sub>2</sub>、以及在步骤4080中计算出的推断正确率*p*<sub>1</sub>的变量代入函数*Fsa*、*Fsc*、*Fsr*,来分别计算使用了语音识别结果*i*<sub>1</sub>的情况下的基于采用动作、确认动作、舍弃动作的用户满意度*Sa*<sub>1</sub>、*Sc*<sub>1</sub>、*Sr*<sub>1</sub>。其中,在步骤5160中,*d*<sub>1</sub>、*t*<sub>2</sub>、*p*<sub>1</sub>的变量未必需要全部使用。通过使用任意的一个以上的变量,能够进行用户满意度*Sa*<sub>1</sub>、*Sc*<sub>1</sub>、*Sr*<sub>1</sub>的计算。

[0124] 在步骤5170中,选择部1070计算对使用了第二个接收到的语音识别结果*i*<sub>2</sub>的情况下的信息提供的用户满意度*Sa*<sub>2</sub>、*Sc*<sub>2</sub>。在此,与步骤5160相同,能够通过将在步骤5020中推断出的域*d*<sub>2</sub>、在步骤5040中测量出的经过时间*t*<sub>2</sub>、以及在步骤5050中计算出的推断正确率*p*<sub>2</sub>的变量代入函数*Fsa*、*Fsc*、*Fsr*,来分别计算使用了语音识别结果*i*<sub>2</sub>的情况下的基于采用动作、确认动作、舍弃动作的用户满意度*Sa*<sub>2</sub>、*Sc*<sub>2</sub>。其中,在步骤5170中,*d*<sub>2</sub>、*t*<sub>2</sub>、*p*<sub>2</sub>的变量未必需要全部使用。通过使用任意的一个以上的变量,能够进行用户满意度*Sa*<sub>2</sub>、*Sc*<sub>2</sub>的计算。

[0125] 在步骤5180中,选择部1070判定基于语音识别结果*i*<sub>1</sub>的采用动作中的用户满意度*Sa*<sub>1</sub>是否在步骤5160以及5170中计算出的全部的用户满意度之中为最大。其结果,在*Sa*<sub>1</sub>为最大的情况下,使处理进入步骤5190,在不是的情况下,使处理进入步骤5200。

[0126] 在步骤5190中,选择部1070选择采用动作作为基于语音识别结果*i*<sub>1</sub>的向用户提供信息的方法。之后,选择部1070结束图5的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果*i*<sub>1</sub>进行采用动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(a)的画面2010那样的画面显示于显示器30,移至与识别到的输入操作相应的处理。

[0127] 在步骤5200中,选择部1070判定基于语音识别结果*i*<sub>2</sub>的采用动作中的用户满意度*Sa*<sub>2</sub>是否在步骤5160以及5170中计算出的全部的用户满意度之中为最大。其结果,在*Sa*<sub>2</sub>为最大的情况下,使处理进入步骤5210,在不是的情况下,使处理进入步骤5220。

[0128] 在步骤5210中,选择部1070选择采用动作作为基于语音识别结果*i*<sub>2</sub>的向用户提供信息的方法。之后,选择部1070结束图5的流程图所示的处理,使处理进入图3的步骤3090。

在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果i2进行采用动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(a)的画面2010那样的画面显示于显示器30,移至与识别到的输入操作相应的处理。

[0129] 在步骤5220中,选择部1070判定基于语音识别结果i1的确认动作中的用户满意度Sc1是否在步骤5160以及5170中计算出的全部的用户满意度之中为最大。其结果,在Sc1为最大的情况下,使处理进入步骤5230,在不是的情况下,使处理进入步骤5240。

[0130] 在步骤5230中,选择部1070选择确认动作作为基于语音识别结果i1的向用户提供信息的方法。之后,选择部1070结束图5的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果i1进行确认动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(b)的画面2020或图2(c)的画面2030那样的画面显示于显示器30,在等待用户的确认之后,移至与经确认的输入操作相应的处理。

[0131] 在步骤5240中,选择部1070判定基于语音识别结果i2的确认动作中的用户满意度Sc2是否在步骤5160以及5170中计算出的全部的用户满意度之中为最大。其结果,在Sc2为最大的情况下,使处理进入步骤5250,在不是的情况下,使处理进入步骤5260。

[0132] 在步骤5250中,选择部1070选择确认动作作为基于语音识别结果i2的向用户提供信息的方法。之后,选择部1070结束图5的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果i2进行确认动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(b)的画面2020或图2(c)的画面2030那样的画面显示于显示器30,在待用户的确认之后,移至与经确认的输入操作相应的处理。

[0133] 在步骤5260中,选择部1070选择舍弃动作作为基于语音识别结果i1、i2的向用户提供信息的方法。之后,选择部1070结束图5的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,进行舍弃动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(d)的画面2040那样的画面显示于显示器30,催促用户再次的发言。

[0134] 如以上说明那样,在最大满意度模式下执行的步骤5160~5260的处理中,针对最初接收到的语音识别结果i1与第二个接收到的语音识别结果i2,分别计算采用动作、确认动作、舍弃动作的满意度,选择执行用户满意度变得最高的语音识别结果与动作的组合。

[0135] 接下来,参照图6的流程图对在步骤3070中执行的结果为未到达时的判断例程进行说明。

[0136] 在步骤6010中,选择部1070将与未到达的语音识别结果对应的域d0设定为“未知”。即,在该时刻没有获得任何语音识别结果,因此域为未知而将域d0设定为“未知”。

[0137] 在步骤6020中,选择部1070作为针对未到达的语音识别结果的临时的推断正确率p0,将规定的正确率设定为例如50%。即,在该时刻没有获得任何语音识别结果,因此推断正确率为未知而设定临时的推断正确率p0。

[0138] 在步骤6030中,选择部1070测量从用户结束发言而终端装置100中被输入语音起到当前为止的经过时间t0。

[0139] 在步骤6040中,选择部1070计算在当前时刻进行了某种动作的情况下的用户满意

度 $Sa_0$ 、 $Sc_0$ 、 $Sr_0$ 。在此，与在图4的步骤4090中计算出的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 相同，能够通过将在步骤6010中设定的域 $d_0$ 、在步骤6030中测量出的经过时间 $t_0$ 、以及在步骤6020设定的推断正确率 $p_0$ 的变量代入函数 $F_{sa}$ 、 $F_{sc}$ 、 $F_{sr}$ ，来分别计算当前时刻的基于采用动作、确认动作、舍弃动作的用户满意度 $Sa_0$ 、 $Sc_0$ 、 $Sr_0$ 。其中，在步骤6040中， $d_0$ 、 $t_0$ 、 $p_0$ 的变量未必需要全部使用。通过使用任意的一个以上的变量，能够进行用户满意度 $Sa_0$ 、 $Sc_0$ 、 $Sr_0$ 的计算。

[0140] 在步骤6050中，选择部1070将舍弃动作中的用户满意度 $Sr_0$ 与采用动作中的用户满意度 $Sa_0$ 以及确认动作中的用户满意度 $Sc_0$ 进行比较。其结果，若 $Sr_0$ 为 $Sa_0$ 以上且 $Sc_0$ 以上，则使处理进入步骤6060，在不是的情况下，使处理进入步骤6070。

[0141] 在步骤6060中，选择部1070选择舍弃动作。之后，选择部1070结束图6的流程图所示的处理，使处理进入图3的步骤3090。在该情况下，在步骤3090被判定为否定之后，在步骤3100中，进行舍弃动作的指令被从选择部1070输出至HMI控制部1100。其结果，如图2(d)的画面2040那样的画面显示于显示器30，催促用户再次的发言。

[0142] 在步骤6070中，选择部1070决定等待至语音识别结果到达为止，并进行建立前述的标志的处理。之后，选择部1070结束图6的流程图所示的处理，使处理进入图3的步骤3090。在该情况下，在步骤3090被判定为肯定之后，返回至步骤3010，待机至接收语音识别结果、或再次经过一定时间为止。

[0143] 在以上说明的步骤6010~6070的处理中，决定是否与语音识别结果为未到达无关地进行舍弃动作而催促用户再次的发言。因此，若在该处理中决定为轻易地进行舍弃动作，则用户必须反复进行多次发言，大幅损害了语音识别系统1的使用的便利性。为了避免这样的情况，在步骤6010中设定域 $d_0$ 时，优选的是设定语音识别为最困难的域、例如“未知”。另外，在步骤6020中设定临时的推断正确率 $p_0$ 时，优选的是设定进行采用动作、确认动作所需的最低限度的推断正确率、例如50%。

[0144] 接下来，参照图7的流程图对在步骤3080中执行的结果为一个到达完毕的判断例程进行说明。

[0145] 在步骤7010中，选择部1070将从用户结束发言而终端装置100中被输入语音起到当前为止的经过时间作为经过时间 $t_3$ 进行测量。另外，与前述的经过时间 $t_1$ 相同，只要是与从终端装置100中被输入语音起到当前为止的经过时间相关，则也可以将其他时间指标作为经过时间 $t_3$ 进行测量。

[0146] 在步骤7020中，选择部1070判定在步骤7010中测量出的经过时间 $t_3$ 是否为在图4的步骤4100中预测出的经过时间 $t_2$ 以下。在 $t_3$ 为 $t_2$ 以下的情况下，由于在图4的步骤4130中预先得知等待第二个语音识别结果的用户满意度更高，因此使处理进入步骤7030。另一方面，在 $t_3$ 大于 $t_2$ 的情况下，由于在当前时刻进行某种动作的用户满意度更高可能更高，因此使处理进入步骤7040。

[0147] 在步骤7030中，选择部1070决定等待第二个语音识别结果，并进行建立前述的标志的处理。之后，选择部1070结束图7的流程图所示的处理，使处理进入图3的步骤3090。在该情况下，在步骤3090被判定为肯定之后，返回至步骤3010，待机至接收第二个语音识别结果、或再次经过一定时间为止。

[0148] 在步骤7040中，选择部1070判断经过时间 $t_3$ 是否大幅偏离期待值。例如，能够求出经过时间 $t_3$ 与经过时间 $t_2$ 的预测值的差分，在该差分超过了规定的阈值的情况下，判断为

经过时间 $t_3$ 大幅偏离期待值。另外,在预先得知经过时间 $t_2$ 可取的值的范围的情况下,也可以通过将经过时间 $t_3$ 与该范围进行比较来判断是否经过时间 $t_3$ 大幅偏离期待值。例如,在经过时间 $t_2$ 可取的值的范围以正态分布等概率分布预先存储于终端装置100的情况下,若经过时间 $t_3$ 落在该概率分布的规定区域、例如从上限值至5%的区域内,则能够判断为经过时间 $t_3$ 大幅偏离期待值。其结果,在判断为经过时间 $t_3$ 大幅偏离期待值的情况下,使处理进入步骤7100,在不是的情况下,使处理进入步骤7045。

[0149] 在步骤7045中,选择部1070推断第二个引擎的域、即与未接收的第二个语音识别结果对应的域 $d_2'$ 。在此,与图4的步骤4095相同,推断为域 $d_2'$ 与域 $d_1$ 相同,原样地设定与域 $d_1$ 相同的内容。

[0150] 在步骤7050中,选择部1070计算第二个语音识别结果的推断正确率 $p_2$ 。在此,通过与图4的步骤4110相同的方法,计算推断正确率 $p_2$ 。即,基于在步骤4030或4040中设定的输出源 $k_1$ 以及 $k_2$ 、在步骤7045中推断出的域 $d_2'$ 、在步骤4060中推断出的可靠度 $c_1$ 、以及在步骤7010中测量出的经过时间 $t_3$ ,使用预先设定的函数 $F_{p2}$ 来计算推断正确率 $p_2$ 。另外,在此,与图4的步骤4110不同,使用当前时刻下的经过时间 $t_3$ 。其中,在步骤7050中, $k_1$ 、 $k_2$ 、 $d_2'$ 、 $c_1$ 、 $t_3$ 的变量未必需要全部使用。通过使用任意的一个以上的变量,能够进行推断正确率 $p_2$ 的计算。

[0151] 在步骤7060中,选择部1070计算对使用了第二个语音识别结果的情况下的信息提供的用户满意度。在此,与图4的步骤4120相同,针对采用动作、确认动作的各个计算对第二个语音识别结果的用户满意度 $Sa_2'$ 、 $Sc_2'$ 。即,基于在步骤7045中推断出的域 $d_2'$ 、在步骤7010中测量出的经过时间 $t_3$ 、以及在步骤7050计算出的推断正确率 $p_2$ ,并将这些变量代入函数 $F_{sa}$ 、 $F_{sc}$ ,从而分别计算用户满意度 $Sa_2'$ 、 $Sc_2'$ 。另外,在此,也与步骤7050相同,使用当前时刻下的经过时间 $t_3$ 。其中,在步骤7060中, $d_2'$ 、 $t_3$ 、 $p_2$ 的变量未必需要全部使用。

[0152] 在步骤7070中,选择部1070计算对使用了接收完毕的语音识别结果 $i_1$ 的情况下的信息提供的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 。在此,与图4的步骤4090相同,能够通过将在步骤4050中推断出的域 $d_1$ 、在步骤7010中测量出的经过时间 $t_3$ 、以及在步骤4080中计算出的推断正确率 $p_1$ 的变量代入函数 $F_{sa}$ 、 $F_{sc}$ 、 $F_{sr}$ ,来分别计算使用了语音识别结果 $i_1$ 的情况下的基于采用动作、确认动作、舍弃动作的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 。其中,在步骤7070中, $d_1$ 、 $t_3$ 、 $p_1$ 的变量未必需要全部使用。通过使用任意的一个以上的变量,能够进行用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 的计算。

[0153] 在步骤7080以后的处理中,选择部1070将在步骤7070中计算出的对接收完毕的语音识别结果 $i_1$ 的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 与在步骤7060中计算出的对第二个语音识别结果的用户满意度 $Sa_2'$ 、 $Sc_2'$ 进行比较。基于该比较结果,选择部1070与图4的步骤4130以后的处理相同,判断是选择语音识别结果 $i_1$ 作为用于由HMI控制部1100进行的向用户的信息提供中使用的语音识别结果、还是不选择语音识别结果 $i_1$ 而等待至接收第二个语音识别结果为止。另外,在选择语音识别结果 $i_1$ 的情况下,判断使用采用动作、确认动作、舍弃动作中的哪个来进行向用户的信息提供。

[0154] 在步骤7080中,选择部1070判定相比使用接收完毕的语音识别结果 $i_1$ 是否等待第二个语音识别结果的用户满意度更高。具体而言,选择部1070与图4的步骤4130相同,将对语音识别结果 $i_1$ 的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 的最大值与对第二个语音识别结果的用户满意

度Sa2'、Sc2'的最大值进行比较。其结果,若用户满意度Sa2'、Sc2'的最大值大于用户满意度Sa1、Sc1、Sr1的最大值,则判断为等待第二个语音识别结果的用户满意度更高,使处理进入步骤7090。另一方面,若用户满意度Sa1、Sc1、Sr1的最大值为用户满意度Sa2'、Sc2'的最大值以上,则判断为不等待第二个语音识别结果而使用当前时刻下的接收完毕的语音识别结果i1的用户满意度更高,并使处理进入步骤7100。

[0155] 在步骤7090中,选择部1070决定等待第二个语音识别结果,并进行建立前述的标志的处理。之后,选择部1070结束图7的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为肯定之后,返回至步骤3010,待机至接收第二个语音识别结果、或再次经过一定时间为止。

[0156] 在步骤7100中,选择部1070将基于语音识别结果i1的采用动作中的用户满意度Sa1与确认动作中的用户满意度Sc1以及舍弃动作中的用户满意度Sr1进行比较。其结果,若Sa1为Sc1以上且Sr1以上,则使处理进入步骤7110,在不是的情况下,使处理进入步骤7120。

[0157] 在步骤7110中,选择部1070选择采用动作作为基于语音识别结果i1的向用户提供信息的方法。之后,选择部1070结束图7的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果i1进行采用动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(a)的画面2010那样的画面显示于显示器30,移至与识别到的输入操作相应的处理。

[0158] 在步骤7120中,选择部1070将基于语音识别结果i1的确认动作中的用户满意度Sc1与舍弃动作中的用户满意度Sr1进行比较。其结果,若Sc1为Sr1以上,则使处理进入步骤7130,若为不是的情况、即Sc1小于Sr1,则使处理进入步骤7140。

[0159] 在步骤7130中,选择部1070选择确认动作作为基于语音识别结果i1的向用户提供信息的方法。之后,选择部1070结束图7的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,基于语音识别结果i1进行确认动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(b)的画面2020或图2(c)的画面2030那样的画面显示于显示器30,在待用户的确认之后,移至与经确认的输入操作相应的处理。

[0160] 在步骤7140中,选择部1070选择舍弃动作作为基于语音识别结果i1的向用户提供信息的方法。之后,选择部1070结束图7的流程图所示的处理,使处理进入图3的步骤3090。在该情况下,在步骤3090被判定为否定之后,在步骤3100中,进行舍弃动作的指令被从选择部1070输出至HMI控制部1100。其结果,如图2(d)的画面2040那样的画面显示于显示器30,对用户促使再次的发言。

[0161] 如以上说明那样,在步骤7080以后的处理中,将对接收完毕的语音识别结果i1的用户满意度Sa1、Sc1、Sr1与对第二个语音识别结果的用户满意度Sa2'、Sc2'进行比较,基于该比较结果选择用户满意度达到最高那样的动作。

[0162] 根据以上说明的本发明的一实施方式,起到以下的作用效果。

[0163] (1) 作为语音识别装置的终端装置100具备:检测从用户输入的语音的语音检测部即语音区间检测部1010;使用基于语音识别部1040的语音的识别结果的终端语音识别信息、或基于与终端语音识别部不同的语音识别部1250的语音的识别结果的服务器语音识别信息的某一方的语音识别信息来进行对用户的信息提供的信息提供部即HMI控制部1100;

以及选择部1070。选择部1070通过执行图3~图7的流程图所示的处理,基于从语音被输入起的经过时间,选择这些语音识别信息的某一方作为HMI控制部1100所使用的语音识别信息,并且使由HMI控制部1100进行的信息提供的方法变化。由于像这样设置,因此能够实现对用户而言舒适的语音输入功能。

[0164] (2) 选择部1070在步骤4090以及4120、或步骤5160以及5170、或步骤7070以及7060中,基于从语音被输入起的经过时间 $t_1$ 、 $t_2$ 或 $t_3$ ,计算表示用户对使用了最初的语音识别信息的情况下的信息提供的满意程度的预测值的用户满意度、以及表示用户对使用了第二个语音识别信息的情况下的信息提供的满意程度的预测值的用户满意度。然后,在步骤4130、或步骤5180、5200、5220以及5240、或步骤7080中,比较这些用户满意度,并基于该比较结果选择最初的语音识别信息或第二个语音识别信息的某一方。由于像这样设置,因此能够适当地选择某一个语音识别信息。

[0165] (3) 在先取得最初的语音识别信息、尚未取得第二个语音识别信息的情况下,选择部1070按照图4的流程图,执行结果为第一个新到达的判断例程。在该处理中,选择部1070测量与从语音被输入起到获得最初的语音识别信息为止的经过时间相关的经过时间 $t_1$ (步骤4070),并且预测与从语音被输入起到获得第二个语音识别信息为止的经过时间相关的经过时间 $t_2$ (步骤4100)。然后,基于测量出的经过时间 $t_1$ ,计算对使用了最初的语音识别信息的情况下的信息提供的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ (步骤4090),基于预测到的经过时间 $t_2$ ,计算对使用了第二个语音识别信息的情况下的信息提供的用户满意度 $Sa_2'$ 、 $Sc_2'$ (步骤4120)。之后,比较计算出的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 与用户满意度 $Sa_2'$ 、 $Sc_2'$ (步骤4130),基于该比较结果,决定是否选择最初的语音识别信息(步骤4140~4170)。由于像这样设置,因此能考虑够用户的满意度而适当地进行是选择最初的语音识别信息、还是等待第二个语音识别信息的判断。

[0166] (4) 在取得完毕最初的语音识别信息、未取得第二个语音识别信息的情况下,选择部1070按照图7的流程图,执行结果为一个到达完毕的判断例程。在该处理中,选择部1070测量与从语音被输入起到当前为止的经过时间相关的经过时间 $t_3$ (步骤7010),基于测量出的经过时间 $t_3$ 计算对使用了取得完毕的语音识别信息的情况下的信息提供的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 、以及对使用了第二个语音识别信息的情况下的信息提供的用户满意度 $Sa_2'$ 、 $Sc_2'$ (步骤7070、7060)。之后,比较计算出的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 与用户满意度 $Sa_2'$ 、 $Sc_2'$ (步骤7080),基于该比较结果,决定是否选择取得完毕的语音识别信息(步骤7090~7140)。由于像这样设置,因此能够考虑用户的满意度而适当地进行是选择取得完毕的语音识别信息、还是等待第二个语音识别信息的判断。

[0167] (5) 在先取得最初的语音识别信息、后取得第二个语音识别信息的情况下,选择部1070按照图5的流程图,执行结果为第二个新到达的判断例程。在该处理中,选择部1070测量与从语音被输入起到获得第二个语音识别信息为止的经过时间相关的经过时间 $t_2$ (步骤5040),基于测量出的经过时间 $t_2$ ,计算对使用了最初的语音识别信息的情况下的信息提供的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 、以及对使用了第二个语音识别信息的情况下的信息提供的用户满意度 $Sa_2$ 、 $Sc_2$ (步骤5160、5170)。之后,比较计算出的用户满意度 $Sa_1$ 、 $Sc_1$ 、 $Sr_1$ 与用户满意度 $Sa_2$ 、 $Sc_2$ (步骤5180、5200、5220、5240),基于该比较结果,选择最初的语音识别信息或第二个语音识别信息的某一方(步骤5190、5210、5230、5250)。由于像这样设置,因此能够考

考虑用户的满意度而适当地进行选择最初的语音识别信息与第二个语音识别信息的哪个的判断。

[0168] (6) 选择部1070在步骤4090以及4120、或步骤5160以及5170、或步骤7070以及7060中,进一步基于根据语音的属性而预先确定的多个域中的与最初的语音识别信息及第二个语音识别信息分别对应的域d1及域d2、以及分别对应于最初的语音识别信息及第二个语音识别信息而求出来的推断正确率p1及推断正确率p2的至少一个,计算对最初的语音识别信息的用户满意度以及对第二个语音识别信息的用户满意度。由于像这样设置,因此能够适当地计算用户满意度。

[0169] (7) 语音识别部1040以及语音识别部1250的至少一个使用在词典数据库1050、1260中预先登记的多个词典数据的某一个来进行语音的识别。也可以设为,选择部1070在步骤4050以及5020中,基于语音识别部1040以及语音识别部1250的至少一个在语音的识别中使用的词典数据来推断域d1以及域d2的至少一个。这样,能够容易地进行域d1、d2的推断。

[0170] (8) 最初的语音识别信息以及第二个语音识别信息的至少一个还能够包含表示针对语音的用户的意图的推断结果的意图推断信息。在该情况下,也可以设为,选择部1070在步骤4050以及5020中,基于该意图推断信息,推断域d1以及域d2的至少一个。这样,能够容易地进行域d1、d2的推断。

[0171] (9) 选择部1070在步骤4080以及5050中,基于域d1以及域d2、针对最初的语音识别信息的可靠度c1以及针对第二个语音识别信息的可靠度c2、以及从语音被输入起的经过时间t1以及t2的至少一个来决定推断正确率p1以及推断正确率p2。由于像这样设置,因此能够准确地决定推断正确率p1以及推断正确率p2。

[0172] (10) 选择部1070在步骤4180、5120、5190、5210以及7110、或步骤4190、5140、5230、5250以及7130、或步骤4160、5150、5260、6060以及7140中,选择采用基于最初的语音识别信息或第二个语音识别信息的某一个的输入操作作为用户的输入操作的采用动作、在用户的确认后采用基于最初的语音识别信息或第二个语音识别信息的某一个的输入操作作为用户的输入操作的确认动作、以及不采用基于最初的语音识别信息的输入操作以及基于第二个语音识别信息的输入操作的任一个而进行舍弃的舍弃动作中的某一个动作,并对应于与选择的动作使由HMI控制部1100进行的信息提供的方法变化。由于像这样设置,因此能够根据状况以适当的方法进行向用户的信息提供。

[0173] (11) 语音识别系统1具备终端装置100与服务器200。终端装置100具备:检测从用户输入的语音的语音检测部即语音区间检测部1010;执行用于识别检测到的语音的语音识别处理,并输出基于该语音的识别结果的终端语音识别信息的语音识别部1040;将基于检测到的语音的语音信息发送至服务器200,并接收从服务器200发送的服务器语音识别信息的通信控制部1030;使用终端语音识别信息或服务器语音识别信息的某一方来进行向用户的信息提供的信息提供部即HMI控制部1100;以及基于从语音被输入起的经过时间,选择这些语音识别信息的某一方,并且使由HMI控制部1100进行的信息提供的方法变化的选择部1070。服务器200具备:接收从终端装置100发送的语音信息,并将服务器语音识别信息发送至终端装置100的通信控制部1210;以及执行用于基于接收到的语音信息识别语音的语音识别处理,并输出基于该语音的识别结果的服务器语音识别信息的语音识别部1250。由于

像这样设置,因此能够通过语音识别系统1,实现对用户而言舒适的语音输入功能。

[0174] <第一变形例>

[0175] 接下来,对本发明的第一变形例进行说明。在前述的实施方式中,设为能够通过决定变量来唯一地决定到获得推断正确率 $p_1$ 以及 $p_2$ 、用户满意度、未接收的第二个语音识别结果为止的经过时间 $t_2$ 的预测值等的推断值而进行了说明。但是,实际上,在这些推断值中存在偏差,并且,在为了决定这些推断值的而使用的变量中也存在偏差。因此,在第一变形例中,对通过将用于求出这些推断值的函数、变量以概率分布来呈现,从而考虑各种偏差而求出最可靠的推断值的例子进行说明。另外,在以下的说明中,对在图4的流程图所示的结果为第一个新到达的判断例程中,基于概率分布求出各推断值的情况进行说明,在其他判断例程中进行求出的情况也相同。

[0176] 首先,在图4的步骤4080中,选择部1070代替实施方式中说明的函数 $F_p$ 而使用概率密度函数 $p_p$ 来计算语音识别结果 $i_1$ 的推断正确率 $p_1$ 。概率密度函数 $p_p$ 是表示推断正确率 $p_1$ 可取的值的概率密度的函数,可根据 $k_1$ 、 $d_1$ 、 $c_1$ 、 $t_1$ 的各变量来定义该函数形式。若使用概率密度函数 $p_p$ ,则最准确的推断正确率 $p_1$ 作为概率密度函数 $p_p$ 的期待值例如可由以下的式(1)求出。

$$[0177] \quad p_1 = E[p_p] = \int p'_1 \cdot p_p(p'_1, d_1, c_1, t_1) \cdot dp'_1 \quad (1)$$

[0178] 进而,在步骤4090中,选择部1070也能够针对对使用了语音识别结果 $i_1$ 情况下的信息提供的每个动作的用户满意度 $S_{a1}$ 、 $S_{c1}$ 、 $S_{r1}$ 求出概率的值。例如,代替与采用动作、确认动作、舍弃动作的各个对应的前述的函数 $F_{sa}$ 、 $F_{sc}$ 、 $F_{sr}$ ,可使用概率密度函数 $p_{sa}$ 、 $p_{sc}$ 、 $p_{sr}$ ,由以下的式(2)~(4)求出用户满意度 $S_{a1}$ 、 $S_{c1}$ 、 $S_{r1}$ 。

$$[0179] \quad S_{a1} = E[p_{sa}] = \iint S'_{a1} \cdot p_{sa}(S'_{a1}, d_1, c_1, p_1) \cdot p'_1 \cdot p_p(p'_1, d_1, c_1, t_1) \cdot dp'_1 dS'_{a1} \quad (2)$$

$$[0180] \quad S_{c1} = E[p_{sc}] = \iint S'_{c1} \cdot p_{sc}(S'_{c1}, d_1, c_1, p_1) \cdot p'_1 \cdot p_p(p'_1, d_1, c_1, t_1) \cdot dp'_1 dS'_{c1} \quad (3)$$

$$[0181] \quad S_{r1} = E[p_{sr}] = \iint S'_{r1} \cdot p_{sr}(S'_{r1}, d_1, c_1, p_1) \cdot p'_1 \cdot p_p(p'_1, d_1, c_1, t_1) \cdot dp'_1 dS'_{r1} \quad (4)$$

[0182] 如上述的式(2)~(4)那样,通过不唯一地决定推断正确率 $p_1$ 而是表示为概率分布,能够基于推断正确率 $p_1$ 的偏差来计算更可靠的用户满意度。

[0183] 另外,以上说明的那样的概率的推断值的计算方法在求解第二个引擎的各推断值、即与第二个语音识别结果相关的各推断值时,能够发挥更高的效果。其理由是因为,与第二个语音识别结果相关的各推断值需要基于最初的语音识别结果来求出。另外,与第二个语音识别结果相关的推断值为前述的经过时间 $t_2$ 、推断正确率 $p_2$ 、域 $d_2'$ 等。

[0184] 首先,在图4的步骤4095中,选择部1070推断与第二个语音识别结果对应的域。在实施方式中,该域推断为与最初的语音识别结果 $i_1$ 所对应的域 $d_1$ 相同的域 $d_2'$ 。但是,严格来说,也认为与第二个语音识别结果对应的域与域 $d_1$ 不同。因此,在第一变形例中,将与第二个语音识别结果对应的域设为 $d_2$ ,由以下的式(5)表示某状况下的域 $d_2$ 的概率 $P_d$ 。

$$[0185] \quad P_d(d_2 | d_1, c_1, t_1, p_1) \quad (5)$$

[0186] 接下来,在步骤4100中,选择部1070预测从用户结束发言而终端装置100被输入语音起到获得未接收的第二个语音识别结果为止的经过时间 $t_2$ 。若用概率密度函数 $p_t$ 表示该

经过时间 $t_2$ 的概率,则经过时间 $t_2$ 的期待值由以下的式(6)表示。

$$[0187] \quad t_2 = E[p_t] = \int t'_2 \cdot p_t(t'_2, d_2, k_2, c_1, c_s) \cdot dt'_2 \quad (6)$$

[0188] 在此,上述的式(6)中的域 $d_2$ 的概率 $P_d$ 由前述的式(5)表示。另外,式(5)中的推断正确率 $p_1$ 由前述的式(1)表示。因而,将式(6)变形,经过时间 $t_2$ 的期待值由以下的式(7)表示。其中,在式(7)中, $D$ 表示域 $d_2$ 可取的域的全集合。

$$[0189] \quad t_2 = E[p_t] = \sum_{d'_2 \in D} \{ \int \int P_d(d'_2 | d_1, c_1, t_1, p_1) \cdot t'_2, d_2, k_2, c_1, c_s)$$

$$[0190] \quad \cdot p'_1 \cdot p_p(p'_1, d_1, c_1, t_1) \cdot dp'_1 dt'_2 \} \quad (7)$$

[0191] 另外,关于第二个语音识别结果的推断正确率 $p_2$ 、对使用了第二个语音识别结果的情况下的信息提供的用户满意度,能够基于上述的经过时间 $t_2$ 、域 $d_2$ ,以与最初的语音识别结果 $i_1$ 的推断正确率 $p_1$ 、用户满意度相同的方式进行计算。

[0192] 在第一变形例中,如以上说明那样,考虑各推断值的概率上的偏差而计算用户满意度、到达时间 $t_2$ 等的推断值。由此,即使在不确定性较多的情景下,也能够进行可用性(usability)较高的信息提供。

[0193] <第二变形例>

[0194] 接下来,对本发明的第二变形例进行说明。在前述的实施方式中,对使用由终端装置100获得的语音识别结果或由服务器200获得的语音识别结果的某一个进行采用动作、确认动作、舍弃动作的某一个,从而进行向用户的信息提供的例子进行了说明。与此相对,在第二变形例中,对使用由终端装置100获得的意图推断结果或由服务器200获得的意图推断结果的某一个来进行向用户的信息提供的例子进行说明。另外,像实施方式所说明的那样,在终端装置100中通过意图推断部1060进行用户的意图推断,在服务器200中通过意图推断部1270进行用户的意图推断。设为在输入至选择部1070的终端语音识别信息与服务器语音识别信息中包含各自的意图推断结果。

[0195] 另外,在意图推断部1060以及1270中推断的用户的意图是指,用户说出的内容表示意图对终端装置100的何种输入操作。例如,在终端装置100向用户提供导航信息的情况下,设施名检索、自家路线检索、电话呼叫、地图放大、地图缩小等被推断为用户的意图。而且,在设施检索中,也可以推断设施检索中的查询等。在意图推断部1060以及1270中,通过使用了预先确定的规则、统计方法等的公知的方法,能够根据语音识别结果推断这些用户的意图。

[0196] 在图4的步骤4010、图5的步骤5010中,选择部1070能够代替语音识别结果 $i_1$ 、 $i_2$ 而使用由意图推断部1060或1270推断出的用户的意图。在这些意图中,也可以赋予表示意图推断的可靠性的可靠度。而且,作为基于语音识别结果的可靠度与意图推断结果的可靠度这两方的可靠度,例如也可以使用将它们相乘或相加而得的可靠度。若这样的话,则能够考虑语音识别的可靠性与意图推断的可靠性这两方而进行处理。

[0197] 根据以上说明的本发明的第二变形例,终端语音识别信息以及服务器语音识别信息分别包含表示针对语音的用户意图的推断结果的意图推断信息。选择部1070选择这些语音识别信息的某一方所含的意图推断信息。由于像这样设置,因此能够进行考虑了用户的意图的信息提供。

[0198] <第三变形例>

[0199] 接下来,对本发明的第三变形例进行说明。在前述的实施方式中,对基于对用户不

久之前刚说出的语音的语音识别结果、意图推断结果而进行域的推断的例子进行了说明。但是,实际上,以高频度利用的域按每个用户存在一定的趋势。另外,用户在一系列的输入操作中所利用的域大致一定,突然说出符合不同的域的语音的情况较少。因此,也可以考虑这一点而基于过去的域的推断历史来推断这次的域。在该情况下,终端装置100将过去的域的推断历史作为与用户满意度相关的数据事先存储于用户满意度存储部1080。由此,在图4的步骤4050、图5的步骤5020中,选择部1070能够基于过去的域的推断历史,推断与最初的语音识别结果i1对应的域d1、与第二个语音识别结果i2对应的域d2。

[0200] 根据以上说明的本发明的第三变形例,选择部1070在步骤4050、5020中,基于过去的域d1以及域d2的推断历史,推断域d1以及域d2。由于像这样设置,因此能够考虑用户的趋势、一系列的输入操作而更加准确地进行域的推断。

[0201] 另外,在以上说明的本发明的实施方式以及变形例中,对在终端装置100所具备的语音识别部1040与服务器200所具备的语音识别部1250中分别进行语音识别,并在选择部1070中选择这些语音识别结果的例子进行了说明。但是,本发明并不限于这种方式。例如,也可以设为,终端装置100能够与多个服务器连接,终端装置100取得并选择由这些多个服务器分别获得的语音识别结果。在该情况下,终端装置100也可以不具备语音识别部1040。或者,也可以设为,终端装置100具备多个语音识别部,并选择由这些多个语音识别部分别获得的语音识别结果。

[0202] 以上说明的实施方式、各种变化例只是一个例子,只要不损害发明的特征,本发明并不限于这些内容。本发明并不限于上述的实施方式及变形例,能够在不脱离本发明的主旨的范围内进行各种变更。

[0203] 下面的优先权基础申请的公开内容作为引用文本而编入于此。

[0204] 日本专利申请2016年第222723号(2016年11月15日申请)

[0205] 附图标记说明

[0206] 1:语音识别系统

[0207] 100:终端装置

[0208] 200:服务器

[0209] 1010:语音区间检测部

[0210] 1020:语音编码部

[0211] 1030:通信控制部

[0212] 1040:语音识别部

[0213] 1050:词典数据库

[0214] 1060:意图推断部

[0215] 1070:选择部

[0216] 1080:用户满意度存储部

[0217] 1100:HMI控制部

[0218] 1110:语音输出装置

[0219] 1120:录音语音合成部

[0220] 1130:规则语音合成部

[0221] 1210:通信控制部

- [0222] 1220:语音对话控制部
- [0223] 1230:对话场景
- [0224] 1240:语音解码部
- [0225] 1250:语音识别部
- [0226] 1260:词典数据库
- [0227] 1270:意图推断部。

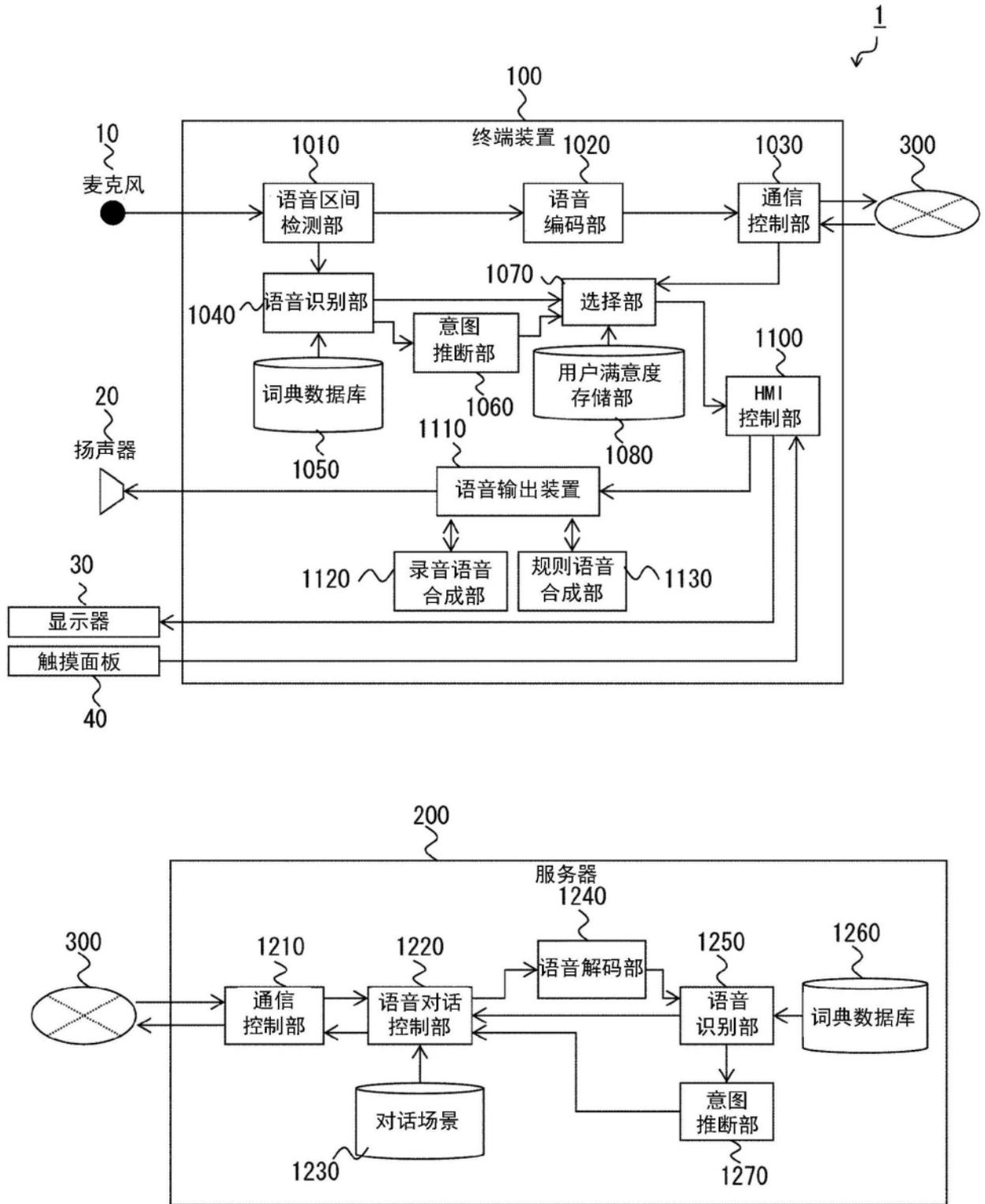
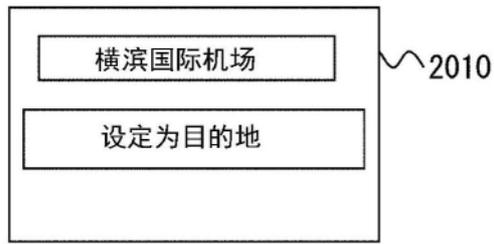
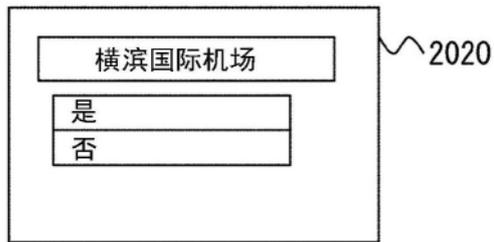


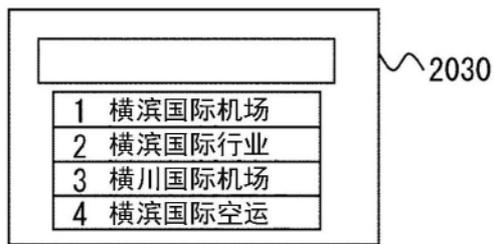
图1



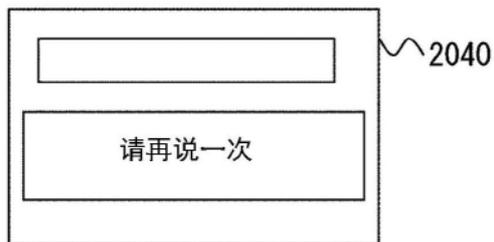
(a)



(b)



(c)



(d)

图2

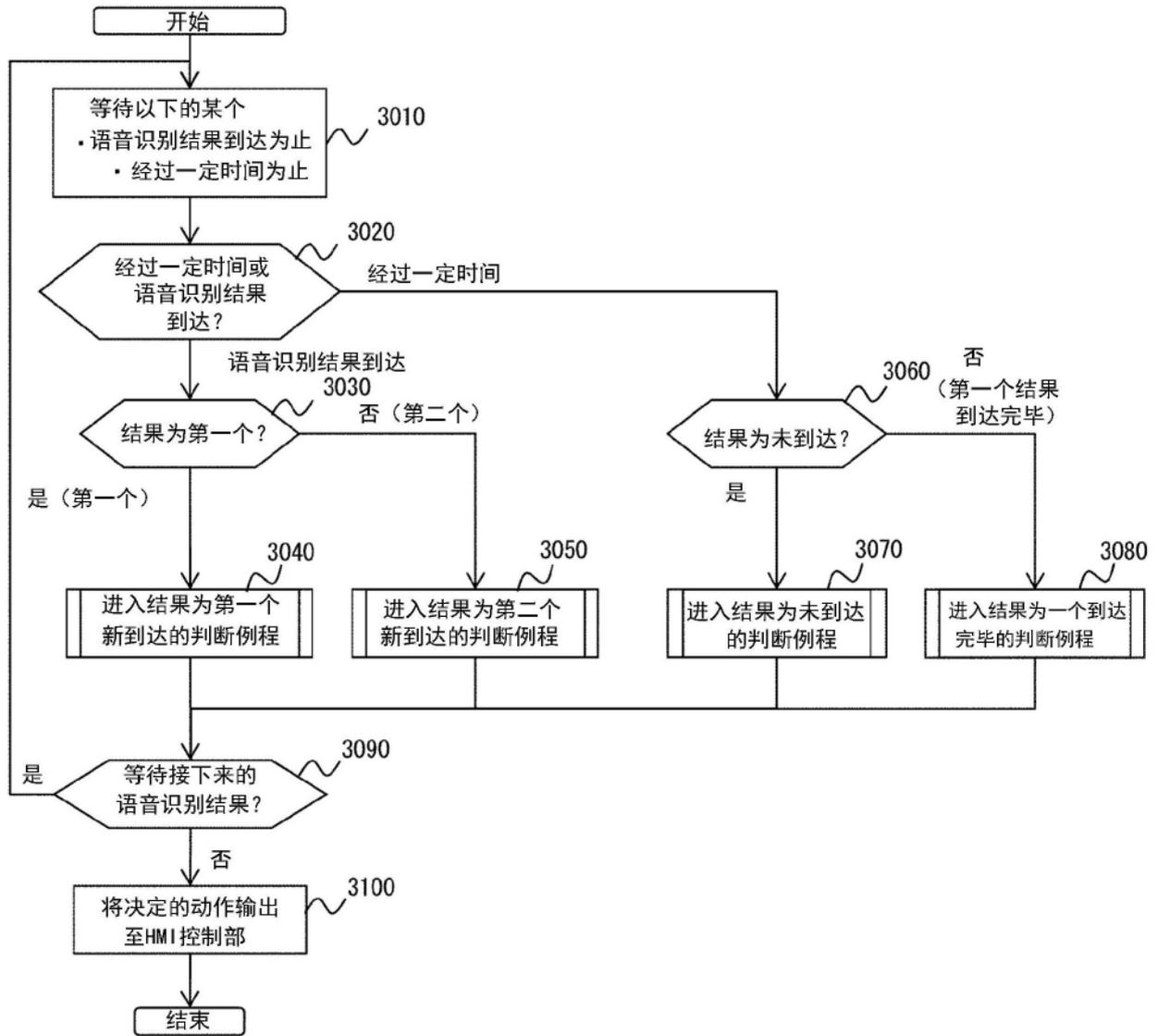


图3

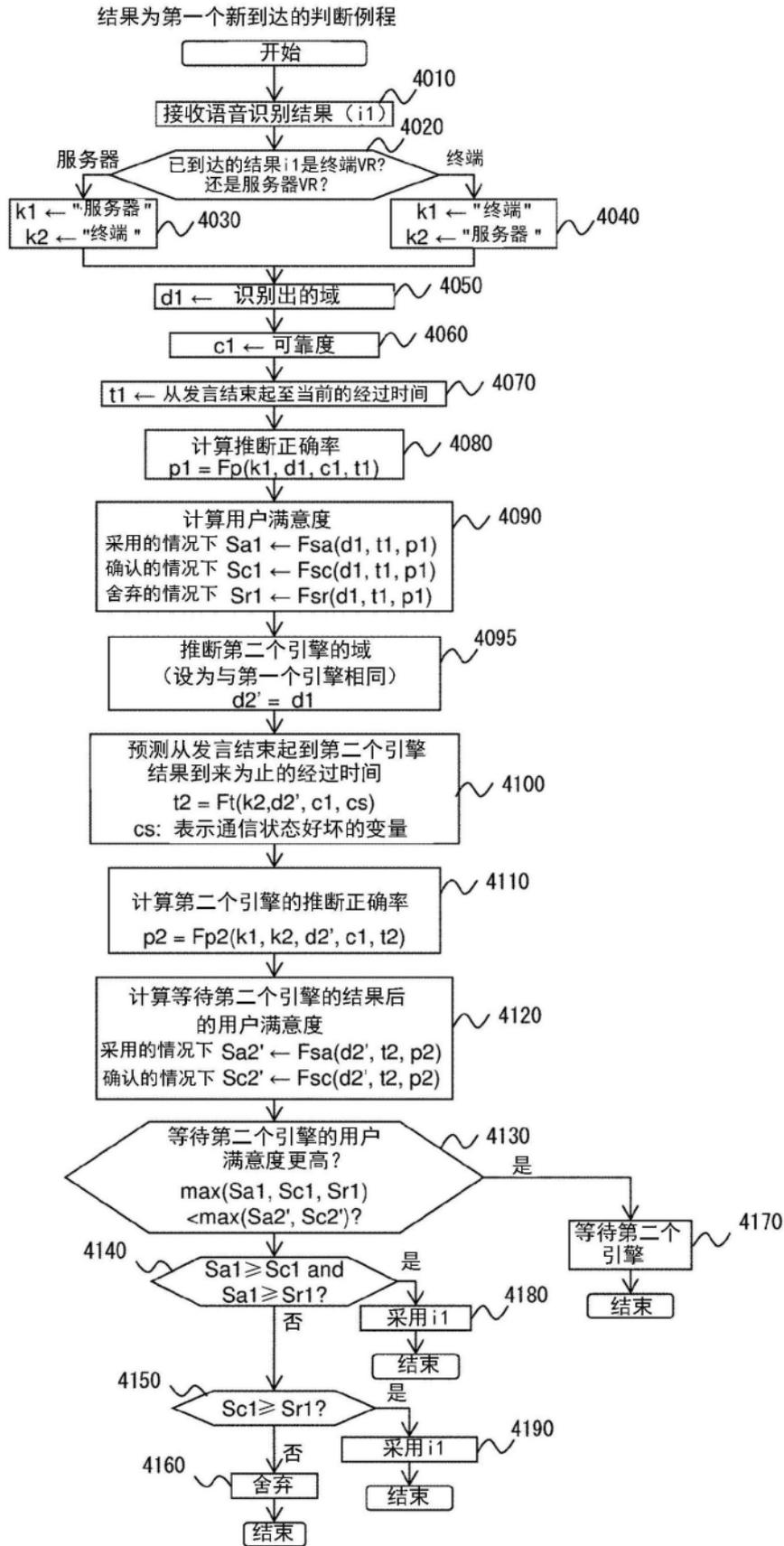


图4

结果为第二个新到达的判断例程

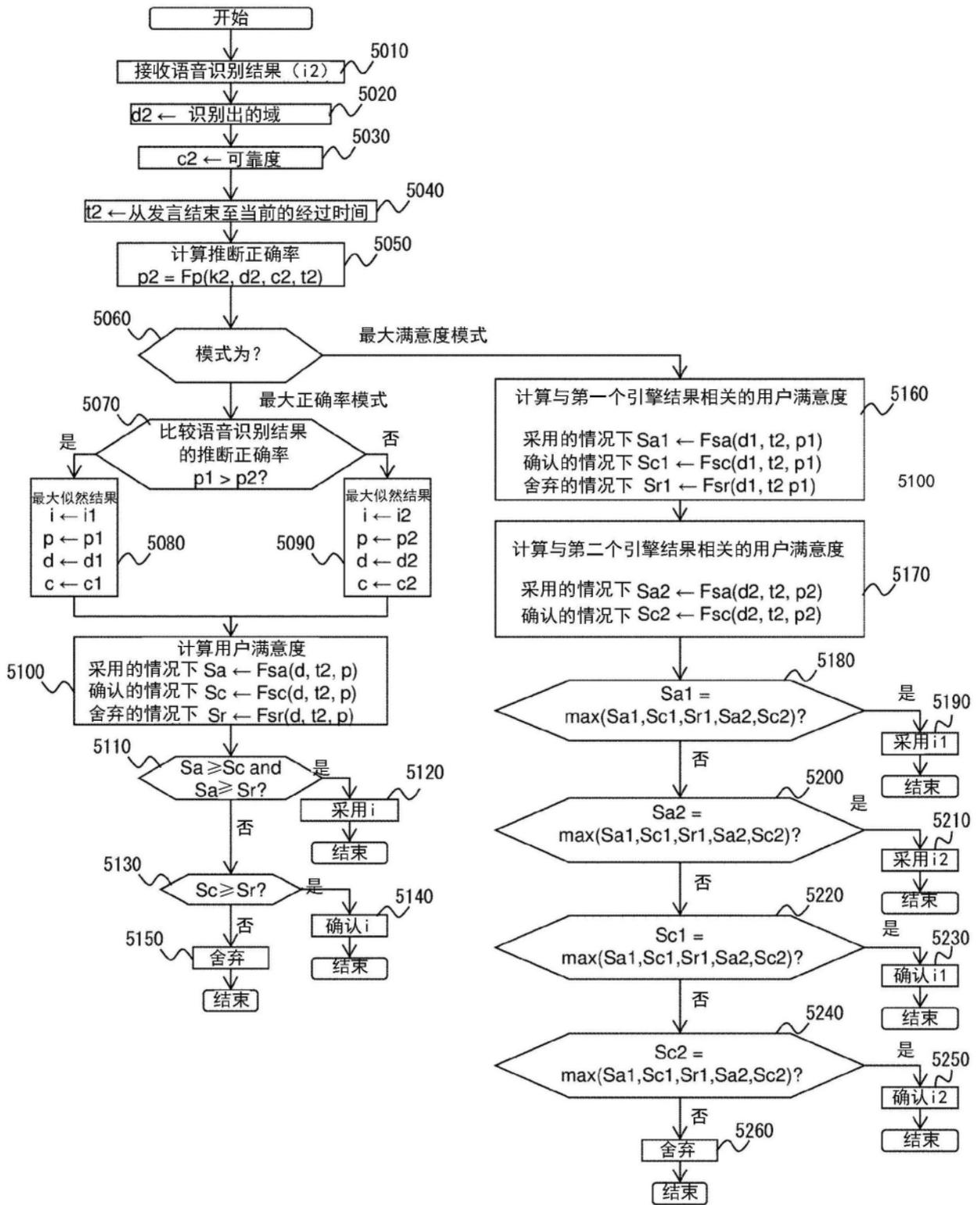


图5

结果为未到达时的判断例程

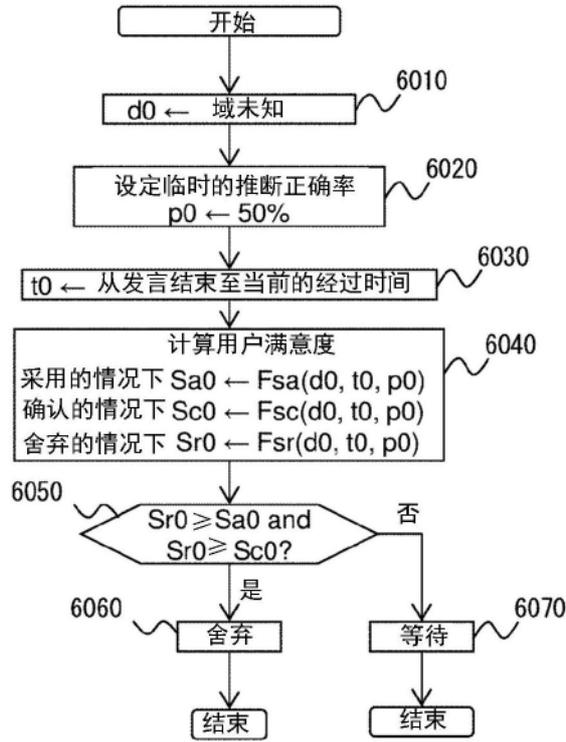


图6

结果为一个到达完毕的判断例程

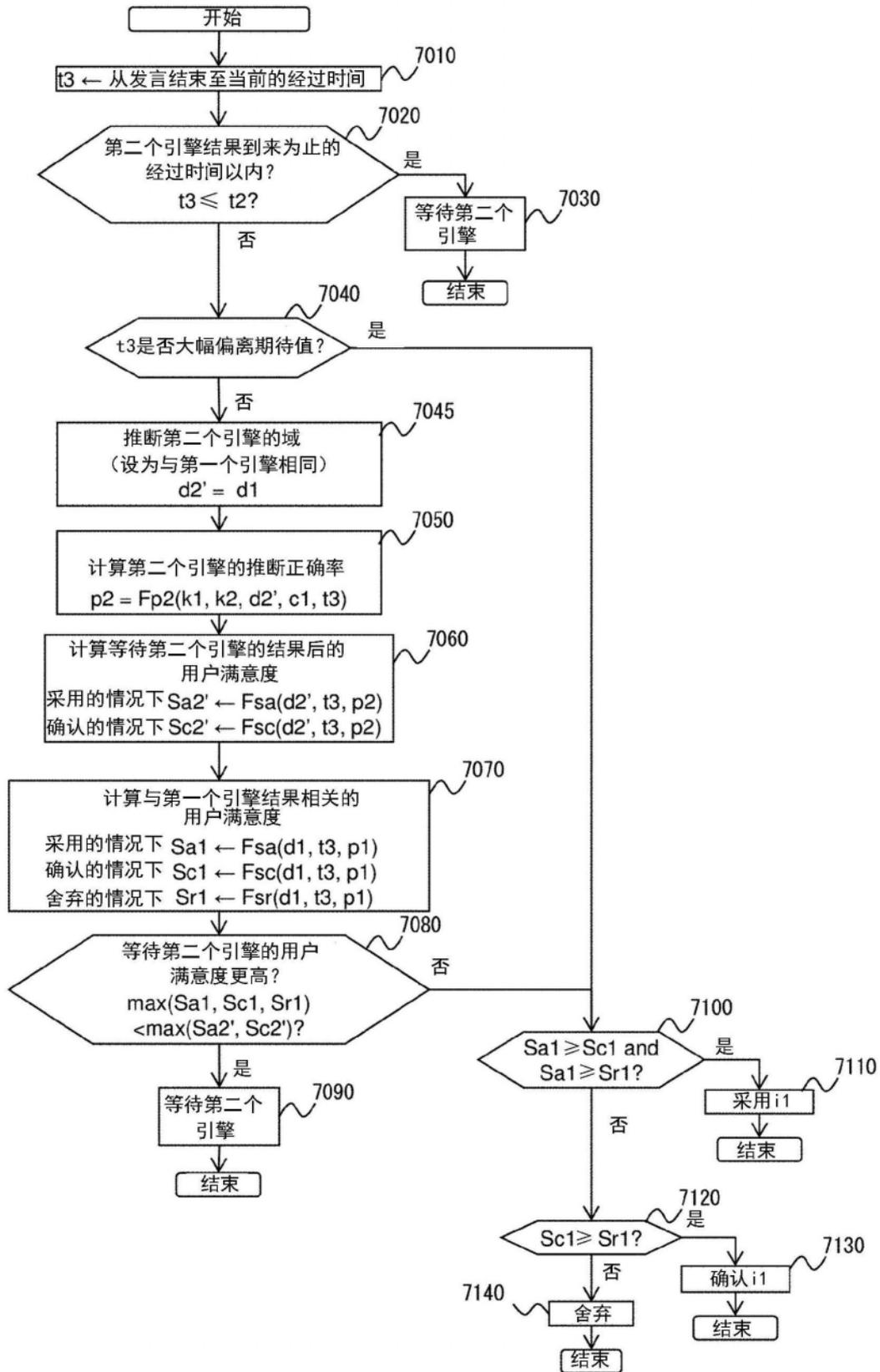


图7

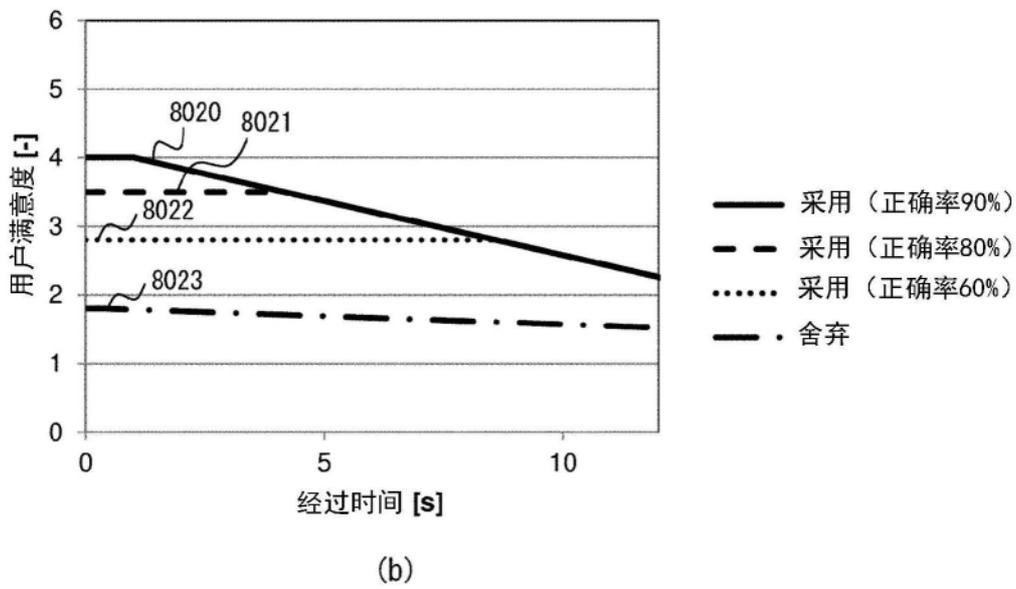
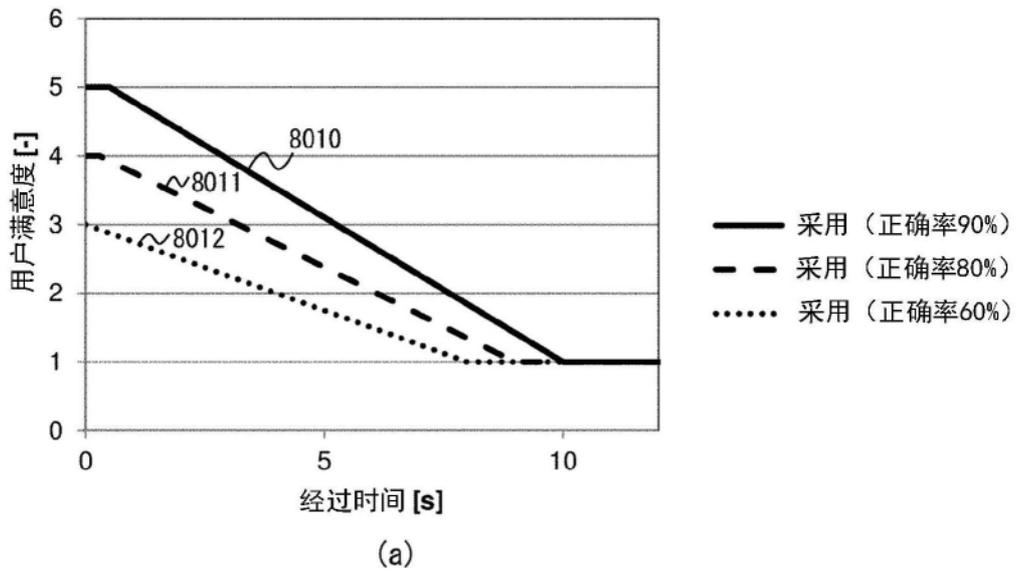


图8

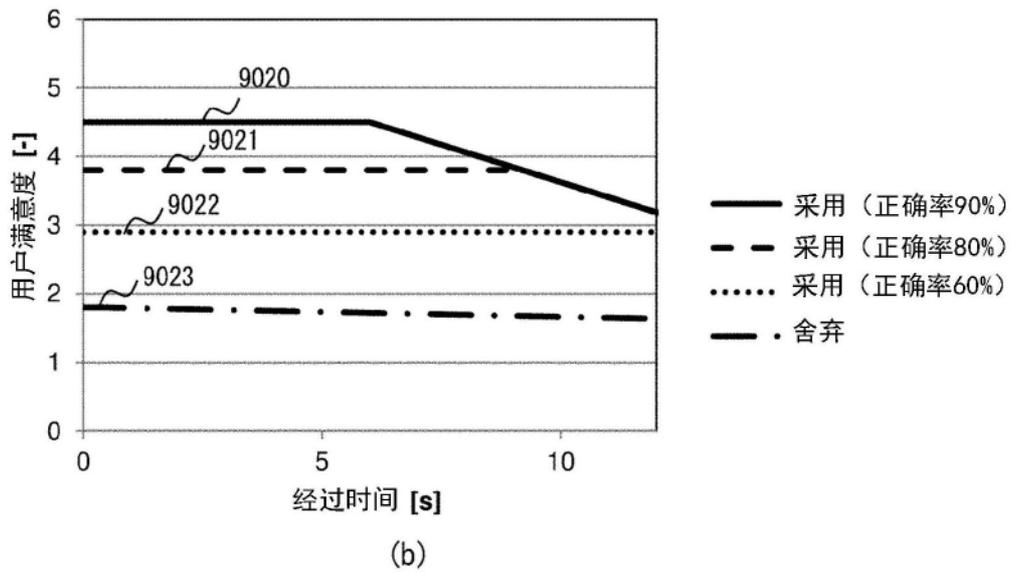
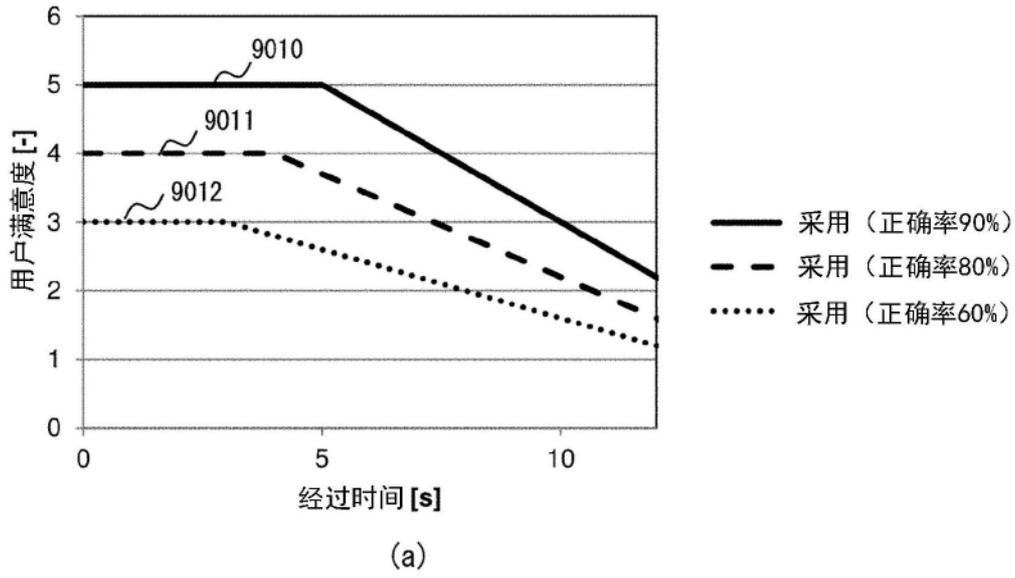


图9