



(12) 发明专利申请

(10) 申请公布号 CN 103038818 A

(43) 申请公布日 2013. 04. 10

(21) 申请号 201180037444. 7

(51) Int. Cl.

(22) 申请日 2011. 06. 24

G10L 15/30(2013. 01)

(30) 优先权数据

61/358, 407 2010. 06. 24 US

(85) PCT申请进入国家阶段日

2013. 01. 29

(86) PCT申请的申请数据

PCT/US2011/041729 2011. 06. 24

(87) PCT申请的公布数据

W02011/163538 EN 2011. 12. 29

(71) 申请人 本田技研工业株式会社

地址 日本东京都

(72) 发明人 R·黄 P·瓦吉菲纳扎里

S·扬玛莫图

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 王茂华

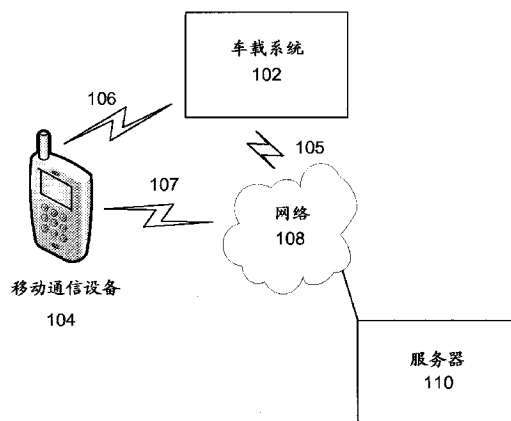
权利要求书 3 页 说明书 9 页 附图 8 页

(54) 发明名称

在车载语音识别系统与车外语音识别系统之间的通信系统和方法

(57) 摘要

一种基于车辆的系统和方法,其用于接收语音输入并且确定是否使用车载资源或车辆外部的资源来执行语音识别分析。



1. 一种在车辆中的基于计算机的方法,包括步骤:
接收第一语音输入表示;
确定对所述第一语音输入表示的语音识别处理应该由所述车辆中的基于计算机的语音识别系统来执行,还是使用所述车辆外部的语音识别系统来执行;
响应于所述语音识别处理应该在所述车辆中执行的所述确定,对于所述第一语音输入表示,执行所述车辆中的所述基于计算机的语音识别;以及
响应于所述语音识别处理应该在所述车辆的外部执行的所述确定,向外部系统发送所述第一语音输入表示。
2. 根据权利要求1所述的基于计算机的方法,其中所述确定的步骤包括步骤:
确定所述第一语音输入表示的第一环境;以及
基于所述第一环境,确定对所述第一语音输入表示的语音识别处理应该由所述车辆中的基于计算机的语音识别系统来执行,还是使用所述车辆外部的语音识别系统来执行。
3. 根据权利要求2所述的基于计算机的方法,其中确定所述第一环境使用由第一用户在所述车辆中按下的车辆键的类型和/或在所述车辆中的输出设备上显示的信息中的至少一个。
4. 根据权利要求3所述的基于计算机的方法,其中在所述输出设备上显示的所述信息是音频信息、视频信息和/或导航信息中的至少一个。
5. 根据权利要求1所述的基于计算机的方法,其中所述执行所述车辆中的所述基于计算机的语音识别的步骤生成第一语音识别结果,所述方法还包括步骤:
确定用于表示所述第一语音识别结果正确的可能性的所述车辆中执行的语音识别处理的置信度值;以及
如果所述置信度值低于第一阈值,则确定对所述第一语音输入表示的所述语音识别处理应该使用所述车辆外部的语音识别系统来执行。
6. 根据权利要求1所述的基于计算机的方法,还包括步骤:
当所述确定是要使用外部语音识别系统时,确定所述车辆与所述外部系统之间的无线传输连接是否超过最小阈值;以及
如果所述无线传输连接的质量低于所述最小阈值,则确定所述语音识别处理应该使用所述车辆中的语音识别系统来执行。
7. 根据权利要求1所述的基于计算机的方法,其中从所述车辆的第一用户接收所述第一语音输入表示,并且所述方法还包括步骤:
向所述外部系统发送关于所述第一用户和/或所述车辆中的至少一个的简档信息。
8. 根据权利要求1所述的基于计算机的方法,其中所述确定步骤包括步骤:
确定所述第一语音输入表示的第一特性;以及
基于所述第一特性,确定对所述第一语音输入表示的语音识别处理应该由所述车辆中的基于计算机的语音识别系统来执行,还是使用所述车辆外部的语音识别系统来执行。
9. 根据权利要求8所述的基于计算机的方法,其中所述第一特性是下列中的至少一个:语音类型、频率、速度、韵律、可理解性、音调、音高、大小、方向、音量和/或持续时间。
10. 根据权利要求8所述的基于计算机的方法,其中所述执行所述车辆中的基于计算机的语音识别的步骤生成第一语音识别结果,所述方法还包括步骤:

确定用于表示所述第一语音识别结果正确的可能性的所述车辆中执行的语音识别处理的置信度值；以及

如果所述置信度值低于第一阈值，则确定对所述第一语音输入表示的所述语音识别处理应该使用所述车辆外部的语音识别系统来执行。

11. 根据权利要求 8 所述的基于计算机的方法，还包括步骤：

当所述确定是要使用外部语音识别系统时，确定所述车辆与所述外部系统之间的无线传输连接是否超过最小阈值；以及

如果所述无线传输连接的质量低于所述最小阈值，则确定所述语音识别处理应该使用对所述第一语音输入的所述表示的所述车辆中的语音识别系统来执行。

12. 根据权利要求 8 所述的基于计算机的方法，其中从所述车辆的第一用户接收所述第一语音输入表示，所述方法还包括步骤：

向所述外部系统发送关于所述第一用户和 / 或所述车辆中的至少一个的简档信息。

13. 一种车辆中的系统，包括：

麦克风，用于从第一用户接收第一语音输入表示；

语音识别定位单元，用于确定对所述第一语音输入表示的语音识别处理应该由所述车辆中的基于计算机的语音识别系统来执行，还是使用所述车辆外部的语音识别系统来执行；

所述语音识别系统，用于响应于所述语音识别处理应该在所述车辆中执行的所述确定，对于所述第一语音输入表示，执行所述车辆中的所述基于计算机的语音识别；以及

通信单元，用于响应于所述语音识别处理应该在所述车辆的外部执行的所述确定，向外部系统发送所述第一语音输入表示。

14. 根据权利要求 13 所述的系统，还包括：

环境单元，用于确定所述第一语音输入表示的第一环境；

其中所述语音识别定位单元基于所述第一环境，确定对所述第一语音输入表示的语音识别处理应该由所述车辆中的基于计算机的语音识别系统来执行，还是使用所述车辆外部的语音识别系统来执行。

15. 根据权利要求 13 所述的系统，其中所述语音识别系统生成第一语音识别结果，所述系统还包括：

置信度值单元，用于确定用于表示所述第一语音识别结果正确的可能性的所述车辆中执行的语音识别处理的置信度值；

其中如果所述置信度值低于第一阈值，则所述语音识别定位单元确定所述语音识别处理应该使用所述车辆外部的语音识别系统来执行。

16. 根据权利要求 13 所述的系统，

其中当所述确定是要使用外部语音识别系统时，所述通信单元确定所述车辆与所述外部系统之间的无线传输连接是否超过最小阈值；以及

其中如果所述无线传输连接的质量低于所述最小阈值，则所述语音识别定位单元确定对所述第一语音输入表示的所述语音识别处理应该使用所述车辆内的所述语音识别系统来执行。

17. 根据权利要求 13 所述的系统，其中所述通信单元向所述外部系统发送关于所述第

一用户和所述车辆中的至少一个的简档信息。

18. 一个或多个具有计算机可执行指令的非瞬态计算机可读存储介质,所述指令当被执行时使得至少一个处理器执行以下步骤:

从所述车辆的第一用户接收第一语音输入表示;

确定对所述第一语音输入表示的语音识别处理应该由所述车辆中的基于计算机的语音识别系统来执行,还是由所述车辆外部的语音识别系统来执行;

响应于所述语音识别处理应该在所述车辆中执行的所述确定,对于所述第一语音输入表示,执行所述车辆中的所述基于计算机的语音识别;以及

响应于所述语音识别处理应该在所述车辆的外部执行的所述确定,向外部系统发送所述第一语音输入表示。

19. 根据权利要求 18 所述的非瞬态计算机可读存储介质,还包括:

确定所述第一语音输入表示的第一环境;以及

基于所述第一环境,确定对所述第一语音输入表示的语音识别处理应该由所述车辆中的基于计算机的语音识别系统来执行,还是使用所述车辆外部的语音识别系统来执行。

20. 根据权利要求 18 所述的非瞬态计算机可读存储介质,其中所述执行所述车辆中的所述基于计算机的语音识别的步骤生成第一语音识别结果,所述方法还包括:

确定用于表示所述第一语音识别结果正确的可能性的所述车辆中执行的语音识别处理的置信度值;以及

如果所述置信度值低于第一阈值,则确定对所述第一语音输入表示的所述语音识别处理应该使用所述车辆外部的语音识别系统来执行。

21. 根据权利要求 18 所述的非瞬态计算机可读存储介质,还包括步骤:

当所述确定是要使用外部语音识别系统时,确定所述车辆与所述外部系统之间的无线传输连接是否超过最小阈值;以及

如果所述无线传输连接的质量低于所述最小阈值,则确定所述语音识别处理应该使用所述车辆中的语音识别系统来执行。

22. 根据权利要求 18 所述的非瞬态计算机可读存储介质,还包括步骤:

向外部系统发送关于所述第一用户和 / 或所述车辆中的至少一个的简档信息。

在车载语音识别系统与车外语音识别系统之间的通信系统 和方法

[0001] 相关申请

[0002] 本发明涉及并且要求 2010 年 6 月 24 日递交的美国临时专利申请 61/358,407 的优先权,通过引用的方式将该申请整体并入本文。

技术领域

[0003] 实施方式涉及车辆通信系统和方法的领域。

背景技术

[0004] 常规车辆采用可听(例如语音)命令来实现各种车辆特征。常见问题在于由于车辆内的有限的处理能力,车载语音识别系统不如车外语音识别系统那么健壮并且可能导致不准确的识别结果。但是,采用车外语音识别系统可能导致长的响应时间。所需要的是用于车辆的语音识别的系统和方法,当其负责车辆内的对时间敏感并且对环境敏感的语音识别时提供健壮的语音识别。

发明内容

[0005] 一种车辆内的基于计算机的方法包括步骤:接收第一语音输入;确定对所述第一语音输入的语音识别处理应该由所述车辆中的基于计算机的语音识别系统执行还是使用所述车辆外部的语音识别系统执行;响应于所述语音识别处理将在所述车辆中执行的确定,对于所述第一语音输入,执行所述车辆内的基于计算机的语音识别;以及响应于所述语音识别处理应该在所述车辆的外部执行的确定,向外部系统发送所述第一语音输入。

[0006] 在说明书中描述的特征和优点不是包括一切的,并且具体而言,考虑到附图、说明书和权利要求,许多附加特征和优点将对于本领域的普通技术人员清楚。此外,应该注意到,说明书中使用的措词原则上为了可读性和介绍的目的选择,而不是选择以界定或限制创造性主题。

附图说明

[0007] 图 1 是实施方式可以在其中操作的环境。

[0008] 图 2 是根据一个实施方式的移动通信设备的更详细的说明。

[0009] 图 3 是根据一个实施方式的车载系统的更详细的说明。

[0010] 图 4 是根据一个实施方式的服务器设备的更详细的说明。

[0011] 图 5 是用于说明根据一个实施方式的语音识别系统的操作的流程图。

[0012] 图 6 是用于说明根据一个实施方式的步骤 504 的操作的流程图。

[0013] 图 7 是用于说明根据一个实施方式的步骤 504 的操作的流程图。

[0014] 图 8 是根据一个实施方式响应于对于信息的用户命令/请求的系统的说明。

具体实施方式

[0015] 现在参考附图来描述实施方式,其中在附图中相同的附图标记指示相同的或功能上类似的元件。并且在附图中,每个附图标记的最左边的数字对应于在其中首次使用该附图标记的附图。

[0016] 在说明书中对于“一个实施方式”或对于“实施方式”的参考意味着结合实施方式描述的具体特征、结构或特性被包括在至少一个实施方式中。在说明书在各种地方的出现的短语“在一个实施方式中”或“实施方式”没有必要全部涉及相同的实施方式。

[0017] 根据计算机内存中的数据比特上的操作的算法和符号表示给出下文的详细描述的一些部分。这些算法描述和表示是数据处理领域的熟练技术人员用于向本领域的其他技术人员最有效地传达他们的工作的手段。在本文中以及通常,算法被认为是一种导致所希望的结果的自一致的步骤(指令)序列。该步骤是需要物理量的物理操作的步骤。这些物理量通常但不是必须采取能够被进行存储、传递、组合、比较以及其他操作的光、电、磁信号的形式。有时候主要为了普通使用的原因而将这些信号称为比特、值、元素、符号、字符、项、数量等是方便的。此外,在不失一般性的前提下将需要物理量的物理操作或变换或物理量的表示的特定步骤的配置称为模块或代码设备也是方便的。

[0018] 但是,全部这些以及类似的术语将要与合适的物理量相关联并且仅仅是应用于这些量的便捷标记。要理解,除非从下文的讨论中显而易见明确声明不是这样,否则贯穿说明书,利用诸如“处理”或“计算”或“算”或“确定”或“显示”之类的术语的讨论涉及这样一种计算机系统或类似的电子计算设备(如特定计算设备)的动作和处理,其中该计算机系统或类似的电子计算设备用于操作被表示为计算机系统的存储器、寄存器或其他此类的信息存储、传输或显示设备中的物理、电子量。

[0019] 本文的实施方式的特定方面包括以算法的形式描述的过程处理步骤和指令。应该注意到,可以将本文的实施方式的过程步骤和指令实现在软件、固件或硬件中,并且当实现在软件中时,可以被下载到由各种操作系统使用的不同平台上或者从该不同平台操作该过程步骤和指令。实施方式还可以在计算系统上执行的计算机程序产品中。

[0020] 本文的实施方式还涉及用于执行操作的装置。该装置可以具体被构造为用于例如具体计算机的目的或者可以包括被存储在计算机中的计算机程序选择性地激活或重配置的通用计算机。该计算机程序可以被存储在计算机可读存储介质中,例如但不限于任意类型的盘片,包括软盘、光盘、CD-ROM、磁光盘、只读存储器(ROM)、随机访问存储器(RAM)、EPROM、EEPROM、磁或光卡、专用集成电路(ASIC)或适用于存储电子信息并且被耦合到计算机系统总线的任意类型的介质。存储器可以包括以上任意和/或可以存储信息/数据/程序的其他设备。此外,说明书中涉及的计算机可以包括单个处理器或者可以是使用多个处理器设计以增加计算能力的架构。

[0021] 本文给出的算法和显示不是固有地涉及任意具体计算机或其他装置。各种通用系统可以与根据本文的教导的程序一起使用,或者可以证明构造更专用的装置来执行该方法步骤是方便的。用于各种这些系统的结构将从下文的说明变清楚。另外,没有参考任意具体的编程语言来描述本文的实施方式。应该认识到,可以是由各种编程语言来实现本文所述的实施方式的教导,并且提供对于具体语言的下文的任意参考,以便本文的实施方式的允许和最佳模式的公开。

[0022] 另外,原则上为了易读性和介绍性的目的选择说明书中使用的措词,而不是为了界定或限制发明性的主题来选择在说明书中使用的措词。因此,本文的实施方式的公开适用于说明而不是限制。

[0023] 图 1 说明了用于各种实施方式的示例性操作系统 100。操作系统 100 可以包括车载系统 102、可选择的无线移动通信设备 104、用于车载系统 102 与网络 108 之间的通信的可选择的通信链路 105、用于车载系统 102 与无线移动通信设备 104 之间的通信的可选择的短距离通信链路 106、无线移动通信设备 104 与网络 108 之间的可选择的无线网络连接通信链路 107 以及处理设备,如被连接到网络 108 的服务器 110。本文所述的通信链路可以直接地或间接地连接两个设备。

[0024] 车载系统 102 和无线移动通信设备 104 可以经由使用例如蓝牙 (Bluetooth®) 技术的短距离通信技术或其他短距离通信技术 (例如通用串行总线 (USB)) 的短距离通信链路 106 来彼此通信。车载系统 102 和无线移动通信设备 104 可以经由短距离通信链路 106 彼此连接或配对。

[0025] 移动通信设备 (MCD) 102 的示例例如包括蜂窝电话、个人数字助理 (PDA)、智能电话、膝上电脑或者具有处理器、通信能力并且易于运输的其他设备。在普通意义上,MCD 104 应用可以是车辆特征和交互的较大套件的一部分。应用的示例包括可用于 iPhone™ (可从加利福尼亚库比蒂诺的 Apple Computer 商业上获得) 的应用或者用于运行 Android™ (可从加利福尼亚山景城的 Google Inc. 商业可得的) 操作系统的电话的应用。

[0026] 在备选实施方式中,不利用移动通信设备 104,并且改为使用嵌入在车辆中的通信设备,例如车辆嵌入式电话、无线网络卡或其他设备 (例如支持 Wi-Fi 的设备)。为了易于讨论,本文的描述关于使用移动通信设备 104 的实施方式来描述实施方式的操作。但是,这不适用于限制实施方式的范围,并且想到其他实施方式使用车载系统 102 与网络 108 之间的其他通信系统进行操作,如上所述。

[0027] 车载系统 102 可以向无线移动通信设备 104 发送信息。无线移动通信设备 104 可以经由短距离通信链路 106 向车载系统 102 发送信息。无线移动通信设备 104 可以存储从车载系统 102 接收的信息,和 / 或经由网络 108 向远程处理设备 (例如服务器 110) 提供该信息。

[0028] 在一些实施方式中,车载系统 102 可以向无线移动通信设备 104 提供信息。无线移动通信设备 104 可以使用该信息来获得来自网络 108 和 / 或服务器 110 的附加信息。也可以响应于从车载系统 102 提供关于无线移动通信设备 104 上的提示的信息,获得该附加信息。

[0029] 网络 108 可以包括无线通信网络,例如蜂窝电话网络以及一个或多个其他网络 (例如互联网、公共交换电话网 (PSTN)、分组交换网络、帧中继网络、光纤网络和 / 或其他类型的网络)。

[0030] 图 2 是无线移动通信设备 104 的示例性实施方式的功能方框图。无线移动通信设备 104 可以包括处理器 202、内存 204、短距离通信模块 206、无线网络连接通信模块 208、输入设备 210、输出设备 212 和存储器设备 218。

[0031] 处理器 202 可以包括用于解释并且执行存储在有形介质中的指令的一个或多个常规处理器,有形介质例如内存 204 或存储器 218 (可以是介质卡)、闪存 (flash RAM) 或其

他有形介质。如本文所使用的,内存 204 可以包括随机访问存储器 (RAM) 或另一类型的动态存储器设备和 / 或只读存储器 (ROM) 或另一类型的静态存储器设备,用于存储由处理器 202 执行的该信息和指令。RAM 或另一类型的动态存储器设备可以存储指令以及临时变量或在由处理器 202 执行指令期间使用的其他中间信息。ROM 或另一类型的静态存储器设备可以存储用于处理器 202 的静态信息和指令。

[0032] 短距离通信模块 206 可以包括用于经由诸如蓝牙的无线短距离技术或诸如通用串行总线 (USB) 的其他短距离通信技术来发送和接收信息的第一收发器。无线网络连接通信模块 208 可以包括用于经由无线通信网络 (包括但不限于蜂窝电话网络) 来发送和接收信息的第二收发器。

[0033] 输入设备 210 可以包括硬键和 / 或软键、麦克风、触摸屏和 / 或其他输入设备。输出设备 212 可以包括扬声器和 / 或显示器。

[0034] 图 3 是车载系统 102 的一个示例性实施方式的功能方框图。车载系统 102 可以包括处理器 302、内存 304、短距离通信模块 306、输入设备 210、输出设备 312、卫星通信模块 316、存储器设备 318 和语音识别模块 320。

[0035] 处理器 302 可以包括用于解释并且执行存储在有形介质中的指令的一个或多个常规处理器,有形介质例如内存 304、介质卡、闪存 (flash RAM)、或其他有形介质。内存 304 可以包括随机访问存储器 (RAM) 或另一类型的动态存储器设备和 / 或只读存储器 (ROM) 或另一类型的静态存储器设备,用于存储由处理器 302 执行的该信息和指令。RAM 或另一类型的动态存储器设备可以存储指令以及临时变量或在由处理器 302 执行指令期间使用的其他信息。如上文所述,ROM 或另一类型的静态存储器设备可以存储用于处理器 302 的静态信息和指令。

[0036] 短距离通信模块 306 可以包括用于经由诸如蓝牙的无线短距离技术或诸如通用串行总线 (USB)、Wi-Fi 等的其他短距离通信技术来发送和接收信息的第一收发器。

[0037] 输入设备 310 可以包括硬键和 / 或软键、麦克风、触摸屏和 / 或其他输入设备。输出设备 312 可以包括扬声器和 / 或显示器。

[0038] 在一些实施方式中,车载系统 102 的组件可以包括用于指导包括车载系统 120 的车辆的用户到目的地的导航单元。即,车载系统 102 的至少一些组件可以作为导航单元一起运行。当车载系统 102 包括导航单元时,车载系统 102 可以包括用于与全球定位卫星通信的卫星通信模块 316。该导航单元还可以包括兴趣点 (“POI”) 数据库。该数据库可以包括关于车辆的用户可能想要去旅行的任意地方的信息。可以由类别 (例如餐馆、购物、急救服务等) 来安排 POI。

[0039] 语音识别模块 320 可以包括独立的处理器、内存和 / 或存储器,或者可以使用车辆中的处理器 302、内存 304 和存储器 318 中的一个或多个。语音识别模块 320 可以存储一个或多个用户的语音简档 (voice profile),以允许其识别由一个或多个用户说出的指令、请求和 / 或查询。

[0040] 图 4 是服务器 110 的示例性实施方式的功能方框图。服务器 110 可以包括处理器 402、内存 404、存储器设备 406、网络连接通信模块 408、输入设备 210、输出设备 412、存储器设备 418 和语音识别模块 420。

[0041] 处理器 402 可以包括用于解释并且执行存储在有形介质中的指令的一个或多个

常规处理器,有形介质例如内存 404、介质卡、闪存 (flash RAM) 或其他有形介质。如本文所使用的,内存 404 可以包括随机接入存储器 (RAM) 或另一类型的动态存储器设备以及只读存储器 (ROM) 或另一类型的静态存储器设备。RAM 或另一类型的动态存储器设备可以存储指令以及临时变量或在由处理器 402 执行指令期间使用的其他中间信息。ROM 或另一类型的静态存储器设备可以存储用于处理器 402 的静态信息和指令。

[0042] 语音识别模块 420 可以包括独立的处理器、内存和 / 或存储器,或者可以使用处理器 402、内存 404 和存储器 414 中的一个或多个。语音识别模块 420 可以存储一个或多个用户的语音简档,以允许其识别由一个或多个用户说出的指令、请求和 / 或查询。

[0043] 网络连接通信模块 408 可以包括用于经由例如网络 108 或者其他网络 (如广域网、局域网、互联网等) 的有线的或无线的通信网络来发送和接收信息的收发器。

[0044] 图 5 是用于说明根据一个实施方式的语音识别系统的操作的流程图。由输入设备 310 接收 502 语音命令 (VC)。一个示例是用户例如通过按下车辆中的硬 / 软键并且随后说出命令、请求、查询等来发起语音命令。为了容易讨论,命令被用作示例并且包括任意这样的输入。语音命令可以被转换成并且作为数字表示来存储或者作为模拟表示来存储。为了容易讨论,本文将讨论数字表示。车辆确定语音识别是否应该由车辆的车载来完成,或者语音识别是否应该由车外的语音识别模块 420 语音识别语音识别来完成 (例如在服务器 110、820 中)。车载或车外语音识别系统 320、420 的示例是可从坐落于美国纽约州阿蒙克的 IBM 公司商业上获得的 ViaVoice、可从坐落于马萨诸塞州伯灵顿的 Nuance 通信商业上获得的 Nuance 语音识别、以及可从坐落于华盛顿州贝尔维的 VoiceBox 技术商业上获得的 Voice Box 语音识别。

[0045] 处理器 302 从输入设备 310 接收输入信号并且识别请求的环境。例如,环境可以基于按下的硬 / 软键 (电话呼叫硬 / 软键或全球命令硬 / 软键) 和 / 或基于在输出设备 312 上显示的信息 (例如导航、音频、视频)。车载系统 102 确定 504 是否应该使用车载语音识别模块 320 或外部语音识别模块 420 来完成语音识别。图 6 和 7 是如何完成这一确定的示例。

[0046] 图 6 是用于说明根据一个实施方式的步骤 504 的操作的流程图。在该实施方式中,所确定的环境可用于确定 604 是否应该使用车载语音识别模块 320 或外部语音识别模块 420 来执行语音识别 (VR)。例如在语音控制与显示器屏幕上的信息相关联的车辆中,如果显示器屏幕显示播放音频盘片,例如 CD 或 DVD,则车载系统确定语音命令的环境与板上音频系统相关并且因此应该被在板上分析,因为该分析将更加快速、更加可靠并且准确性通常高。

[0047] 如果 604 初始决定是在板上执行语音识别,则语音命令 (VC) 被输入 606 到车载语音识别模块 320。车载语音识别模块 320 使用语音识别技术分析语音命令。可以使用许多可能的语音识别技术。确定置信度 (confidence score) 的一个示例在 Jiang 的“Confidence measures for speech recognition :A survey”(2004 年 Speech Communication) 一文中阐述了,该文可从置信度 www.sciencedirect.com 获得,通过参考的方式将其整体并入本文。

[0048] 如果 608 由语音识别模块 320 确定的置信度大于第一阈值 (T1),则过程继续到步骤 510,在步骤 510 中,可以将语音识别分析的结果存储 510 在车辆中在与车辆的用户或驾驶员相关联的语音简档中 (并且可选择地被发送到服务器 110 以便存储在位于车辆外的语

音简档中)。此后,命令可以被执行 534,或者可以向用户提供所请求的信息。

[0049] 如果 608 由语音识别模块 320 确定的置信度不超过第一阈值 (T1),则这表明车载语音识别模块不产生满意的结果,并且在该实施方式中,向外部语音识别模块 420 发送该语音命令,其中外部语音识别模块 420 可以具有更大的处理功率并且因此能够以更高的置信度确定该语音命令的意义。该过程然后可以继续到下文所述的步骤 526。如下文参考步骤 720 所述,可以与语音命令一起发送附加信息。

[0050] 应该理解,第一阈值 (T1) 可以是许多阈值中的一个,其中每个单独的阈值可以具有它自己的结果。例如,置信度可能太低以至于系统将需要用户重复命令而不执行任何其他动作。

[0051] 图 7 是用于说明根据一个实施方式的步骤 504 的操作的流程图。图 7 中所示的步骤不是全部需要的并且不需要按照所示次序来完成。它们可以与图 6 中所示的步骤相结合。车载系统 102 确定 504 应该使用车载语音识别模块 320 还是外部语音识别模块 420 完成语音识别。车载系统 102 接收语音命令 (步骤 502)。车载系统 102 确定命令和 / 或语音的特性,例如语音类型、频率、速度、韵律、可理解性、音调、音高、大小、方向、音量和 / 或持续时间。例如,该命令是否小于大约 5 秒钟,该语音是女性还是男性的语音,或者说话者说话是否快速。在图 7 中所示的实施方式中,使用大小作为示例并且在下文将进一步描述大小。如果 702 命令的大小过大,例如其超过第一大小因子 S1,则车载系统 102 继续到步骤 708。具有长的持续时间的命令的一个示例是用户想要请求长文本消息的发送的命令。如果 702 大小不太大,则车载系统 102 使用车载语音识别模块 320 来确定置信度 (CS)。如果 704 置信度不超过第一阈值 (T1),则车载系统 102 继续到步骤 708。

[0052] 如果置信度超过第一阈值 (T1),则知道具有足够的信心来继续该命令 / 请求。车载系统然后确定 706 是否应该在车辆内还是在车辆外部执行所确定的命令 / 请求。例如在一些实施方式中,应该在车辆内解释特定命令 / 请求,例如音频控制命令、外围控制命令、环境命令 (例如内部温度、风扇速度的修改)、静态数据、与急救相关的请求 (例如请求呼叫 911 或者用于关于最近的医院的信息的请求)、**HandsFreeLink®** (HFL) 命令和地址条目。应该被外部执行的命令 / 请求的示例包括获取新闻或其他动态数据、发送电子邮件或文本消息、购买活动的门票、进行预订、搜索兴趣点、或者获取天气信息或实时交通信息。

[0053] 如果 706 确定命令 / 请求是应该在车辆中执行的命令 / 请求,则过程继续到如本文所述的步骤 510。

[0054] 如果 706 确定命令 / 请求不是应该在板上执行的命令,例如获得实时交通信息或者关于餐馆的营业时间的细节,则车载系统 102 确定 708 是否存在与车载系统 102 与服务器 110 之间的通信链路 105、106、107 中的任意一个相关的任意连接性问题。如果 708 存在通信问题,例如不能够准确地发送和接收数据,则车载系统 102 确定 710 是否可以在车辆板上执行所识别的命令。如果可以,则继续到步骤 510。如果 710 不可以在板上完成所识别的命令或者如果命令未知,则如果命令未知 (例如由于太大或者具有低的信心因子),则车载系统 102 与用户通信 712,以要求用户重复该语音命令。备选地,可以发送 712 用于指示连接问题阻止命令被当前执行的错误消息。在备选实施方式中,当连接性问题不再存在时,车载系统 102 存储该语音命令并且向服务器 110 发送请求。

[0055] 如果 708 不存在连接性问题,则车载系统发送语音命令并且可以提供 720 可以由

外部语音识别模块 420 使用的附加的与用户简档相关的信息和 / 或车辆信息。下文更详细地描述在发送 720 中可以包括的信息的类型。该过程然后继续到步骤 526。

[0056] 如果基于步骤 504 中的决定,语音识别将要在服务器中执行,则外部语音识别模块 420 执行 526 接收命令的语音识别分析。服务器中的外部语音识别模块 420 可以是更强大的分析引擎 / 模块,因为处理器 402 可能更强大,在内存 404 和存储器 414 中可能存在更多可用数据并且可以利用来自其他网络的附加资源。另外,可以使用与该车辆相关联的多个用户简档辅助准确地解释该用户的命令 / 请求,如下文所述的。如果 528 由外部语音识别模块 420 确定的置信度超过第二阈值 (T2,其可以与第一阈值 T1 不同),则可以将语音识别的结果存储 530 在服务器 110 中,并且可以向该车辆发送该结果,其中该结果也可以被存储 510 在该车辆中。车辆然后可以执行 534 所识别的命令 (或者发送所请求的信息)。存储在服务器和 / 或车辆中的结果可用于向语音识别模块 320、420 提供反馈以改进未来的性能。

[0057] 如果 528 由外部语音识别模块 420 确定的置信度不超过第二阈值 (T2),则可以请求 540 用户重复该语音命令并且该过程在步骤 502 处重复开始。

[0058] 在另一个实施方式中,向车载语音识别模块 320 和外部语音识别模块 420 都发送语音命令。由车载系统 102 收集结果,并且向传送通信该结果。在语音识别的结果不同的情况中,则可以比较用于每个情况的置信度的值,并且可以使用具有更高的置信度的结果。备选地,可以使用默认值,其中使用外部语音识别模块 420 的结果,因为假定外部语音识别模块 420 具有更多用于分析语音命令的资源并且因此有可能更准确。在来自车载语音识别模块 320 的结果被传送给用户或者在车载系统 102 接收到来自外部语音识别模块 320 的结果之前执行该命令的情况中,车载系统可以忽略后一个结果,传送该后一个结果给用户或者可以执行后一个命令。

[0059] 如上所述,如果 708 不存在连接性问题,则车载系统发送语音命令并且可以提供 720 可以由外部语音识别模块 420 使用的附加的与用户简档相关的信息和 / 或车辆信息。当向外部语音识别模块 420 发送语音命令时,当前实施方式提供附加信息,其中与当该附加信息对于外部语音识别模块 420 不可用的情况相比,该附加信息可以被该外部语音识别模块 420 使用以识别说话者并且更准确地识别语音命令。

[0060] 在一个实施方式中,外部语音识别模块 420 可以包括用于车辆的用户简档。在备选实施方式中,可以存在多个与该车辆相关联的用户简档,其中每个用户简档辅助外部语音识别模块 420 执行语音命令的语音识别。简档可以例如包括信息,例如用户的口音、发音、句子结构、话音训练数据、关于用户的意向的信息 (例车辆的一个用户可能频繁地在商店喝咖啡而另一个用户典型地去另一个商店喝咖啡)。

[0061] 在各种实施方式中,车载系统 102 通过 (a) 提供可用于辅助识别正确的用户简档的信息 720 和 / 或 (b) 提供关于车辆模型和状态的附加信息 720 来辅助外部语音识别模块 420,该附加信息可用于在该语音命令的同时提供车辆的声学模型。可以提供 720 的用于辅助外部语音识别模块 420 识别正确的用户简档的信息包括:密钥卡识别 (当车辆具有多个密钥卡时)、当与车辆配对时的蜂窝电话号码或电话号码、车辆 / 导航单元识别号码、用户信息 (例如面部特征 (使用面部识别) 和 / 或当驾驶时驾驶员对座位施加的重量 (在具有不同重量的驾驶员之间可以进行区分))、来自车辆中的 GPS 的车辆位置信息——该位置可

用于辅助识别具有与车辆相关联的简档的驾驶员的习惯。例如驾驶员 1 工作在加利福尼亚州的托伦斯并且驾驶员 2 工作在加利福尼亚州的阿纳海姆,当车辆的位置处于 / 靠近托伦斯时,外部语音识别模块 420 可能能够使用位置信息来更准确地确定驾驶员 1 是说话者并且因此当执行语音识别时应该使用与驾驶员 1 相关联的简档。可以包括的附加信息的示例是:如果使用屏幕来说话或选择时的用户的名称、车辆气候控制数据、速度数据、当日时间、刹车形式、任意其他与车辆相关 / 由车辆生成的数据、车辆与正前方的车辆之间的距离或者可能有助于外部语音识别模块 420 区分两个驾驶员的意向并且因此更准确地识别将有助于识别语音命令的具体用户的其他车辆信息。除了语音命令之外,另外发送该信息,并且其因此是对于语音命令的固有特性(例如语音命令的频率)的增补。

[0062] 可以被外部语音识别模块 420 使用的附加信息涉及车辆的音响效果。所提供 720 的信息可以包括车辆的模型、当尝试识别语音命令时允许基于车辆来使用预先确定的声学模型的车辆识别号码 / 导航识别号码。另外,速度、偏航、车辆位置(马路是否嘈杂?)、风扇的状态(高、低、关闭)、乘客的数量、窗户的位置和 / 或当做出语音命令时可以被外部语音识别模块 420 使用以调整声学模型以更准确地反应在车辆中出现的条件的任意其他与车辆相关的 / 由车辆生成的数据。该信息 720 还可以被车载语音识别模块 320 使用。

[0063] 图 8 是根据一个实施方式用于响应于对于信息的用户命令 / 请求的系统的说明。在该实施方式中,可以使用上述车辆语音识别 / 控制系统和方法来辅助用户获取请求信息或者执行请求命令。图 8 是根据一个实施方式的系统的说明。该系统包括车辆 802,车辆 802 包括板载语音识别(VR)系统 804(类似于板是语音识别模块 320)。上文描述了该 VR 系统的示例。

[0064] 可选择地,车辆 802 经由通信链路被无线耦合到导航服务器 810。导航服务器 810 可以可选择地包括或者直接或间接被耦合到交通数据库 812、兴趣点数据库 814 和 / 或地图数据库 816。

[0065] 车辆 802 和 / 或导航服务器 810 被耦合到语音识别服务器 820。语音识别服务器 820 可以被耦合到其他数据库、互连网络和 / 或耦合到程序 / 应用。在一个实施方式中,用户可以对于用户希望使用语音识别服务器 820 的应用 824、网站 822、数据库 826 设置偏好。示例包括 ESPN、Yelp、flixster、Open Table、Google、CNN、Yahoo 等。可以预先设置这些偏好或者可以使用语音识别服务器 820 设置 / 修改这些偏好。

[0066] 在一个实施方式中,用户请求信息或动作。例如,用户可能请求“下午七点在好的日本餐厅预订四个人的位置”。将该口头请求与附加信息 720 一起从车辆 802 发送到导航服务器 810 和 / 或语音识别服务器 820。所发送的附加信息 720 可以包括用户信息、车辆信息、速度、位置、当日时间、计划路径信息以及根据本文所示实施方式发送的信息。

[0067] 语音识别服务器 820 接收语音信息 / 命令和附加信息并且执行如上所述常规语音识别,以收集信息并且 / 或者执行所请求的命令。通过发送位置、速度、当日时间和 / 或计划路径信息,车辆使得语音识别服务器 820 或车辆 802 能够负责车辆的改变位置和行驶的方向。

[0068] 例如,往南行驶的车辆中的用户将优选识别在行驶方向中的餐厅而不是需要原路返回的餐厅。在一个实施方式中,可以由用户选择该偏好。在一个实施方式中,语音识别服务器 820 或车辆还可以基于例如被选择的路径,确定在具体时间上的估计位置。例如在以

上示例中,如果用户已选择目的地并且车辆具有路径,车辆然后可以估计在预订时间(例如 7:00pm)附加车辆的位置,以辅助语音识别服务器确定选择哪个餐厅。在另一个示例中,语音识别服务器 820 和 / 或车辆 802 可以使用路径信息和位置信息来选择信息,例如靠近计划路径的餐厅。

[0069] 在对用户的口头请求执行语音识别之后,语音识别服务器 820 识别一个或多个信息源并且连接到这些源。在一个实施方式中,语音识别服务器 820 使用上文所述的用户偏好信息。该语音识别系统可以连接到互联网网站 822、应用 824 和 / 或数据库 826 以获得该信息。在上述示例中,语音识别系统可以连接到 Yelp 网站 (www.yelp.com) 或类似的网站,以获得视情况与餐厅质量、营业时间、可用性相关的信息。语音识别服务器 820 或在该示例中的外部网站可以视情况连接到多个合适的网站,以便获得信息 / 执行由车辆 802 中的用户说的命令。例如在以上示例中,在识别合适的餐厅之后,外部语音识别服务器 820 可以连接到预订网站 / 服务(如 Open Table, www.opentable.com) 以进行晚餐预订。

[0070] 在获取请求信息之后,语音识别服务器 820 向车辆 802 回发该信息,并且该信息 / 命令被提供给用户。可以由车辆 802 中的扬声器口头地提供该信息,可以在车辆中的屏幕上向用户显示该信息,可以发送文本消息或者可以使用另一个通信方法。在向用户提供信息时,可以高亮和 / 或灰化特定信息,以提供关于该选择的信息。例如,如果用户请求关于该区域中的好的日本餐厅的信息,则车辆可以显示多个餐厅,同时高亮当前开门的那些餐厅或者灰化当前关门的那些餐厅。

[0071] 虽然已经说明并且描述了具体实施方式和应用,但是应该理解实施方式不限于本文所述的精确构造和组件,并且在不脱离实施方式的精神和范围的前提下,可以在本文的实施方式的方法和装置的配置、操作和细节中做出各种修改、改变和变形。

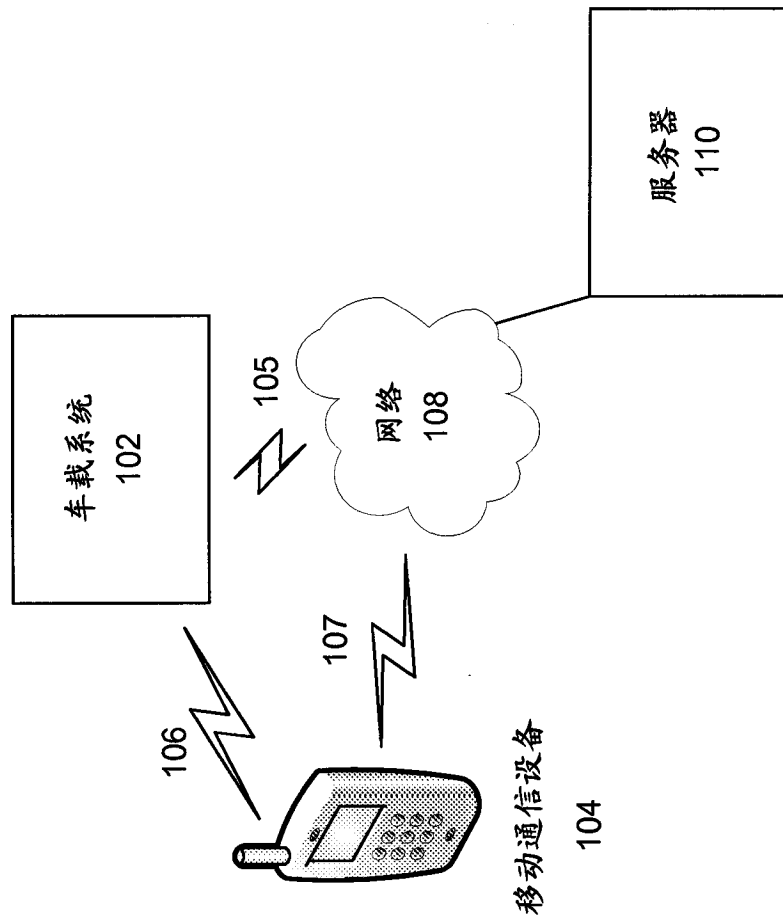


图 1

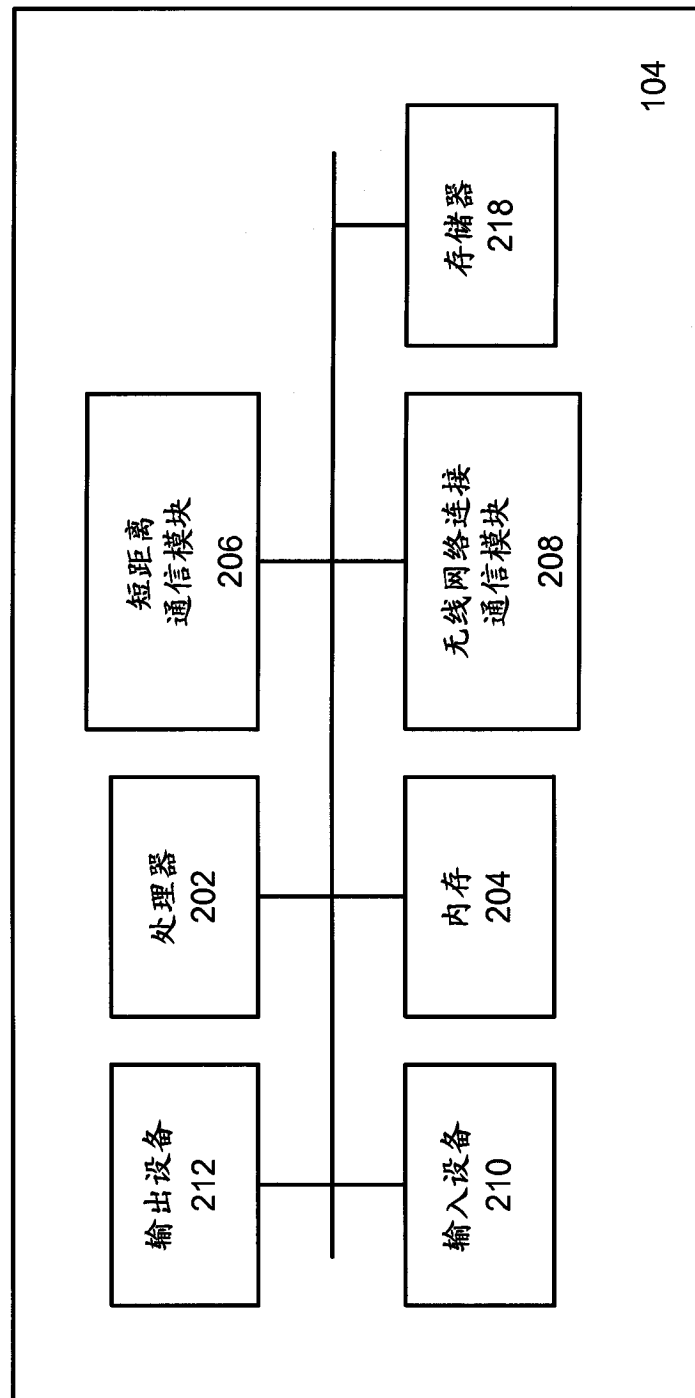


图 2

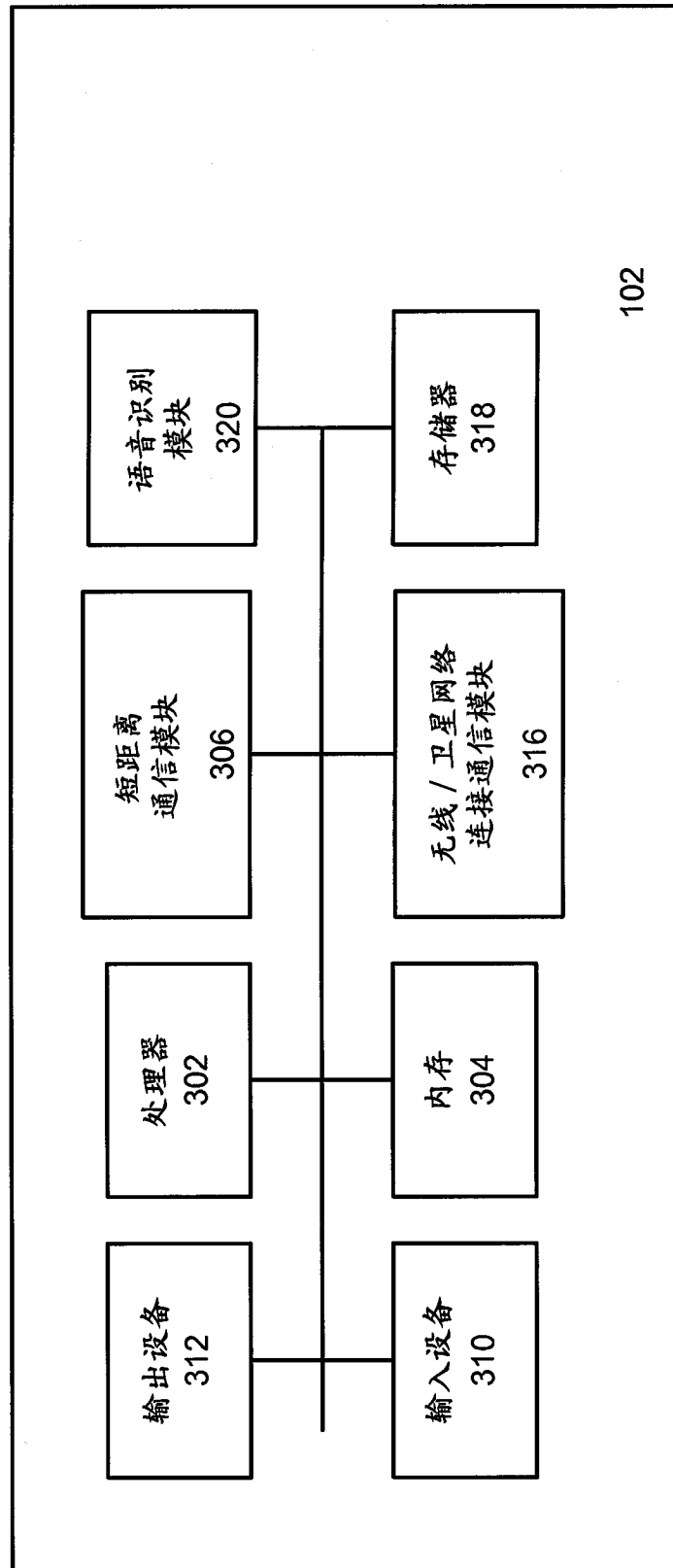


图 3

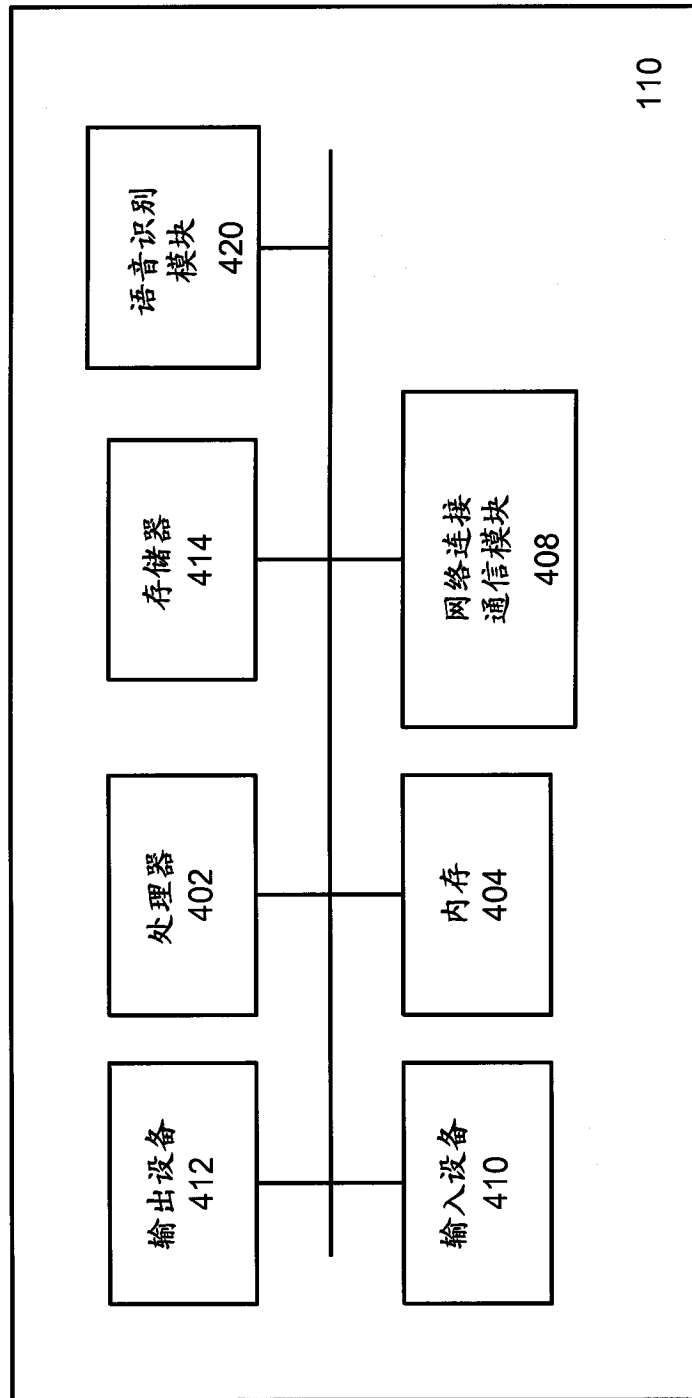


图 4

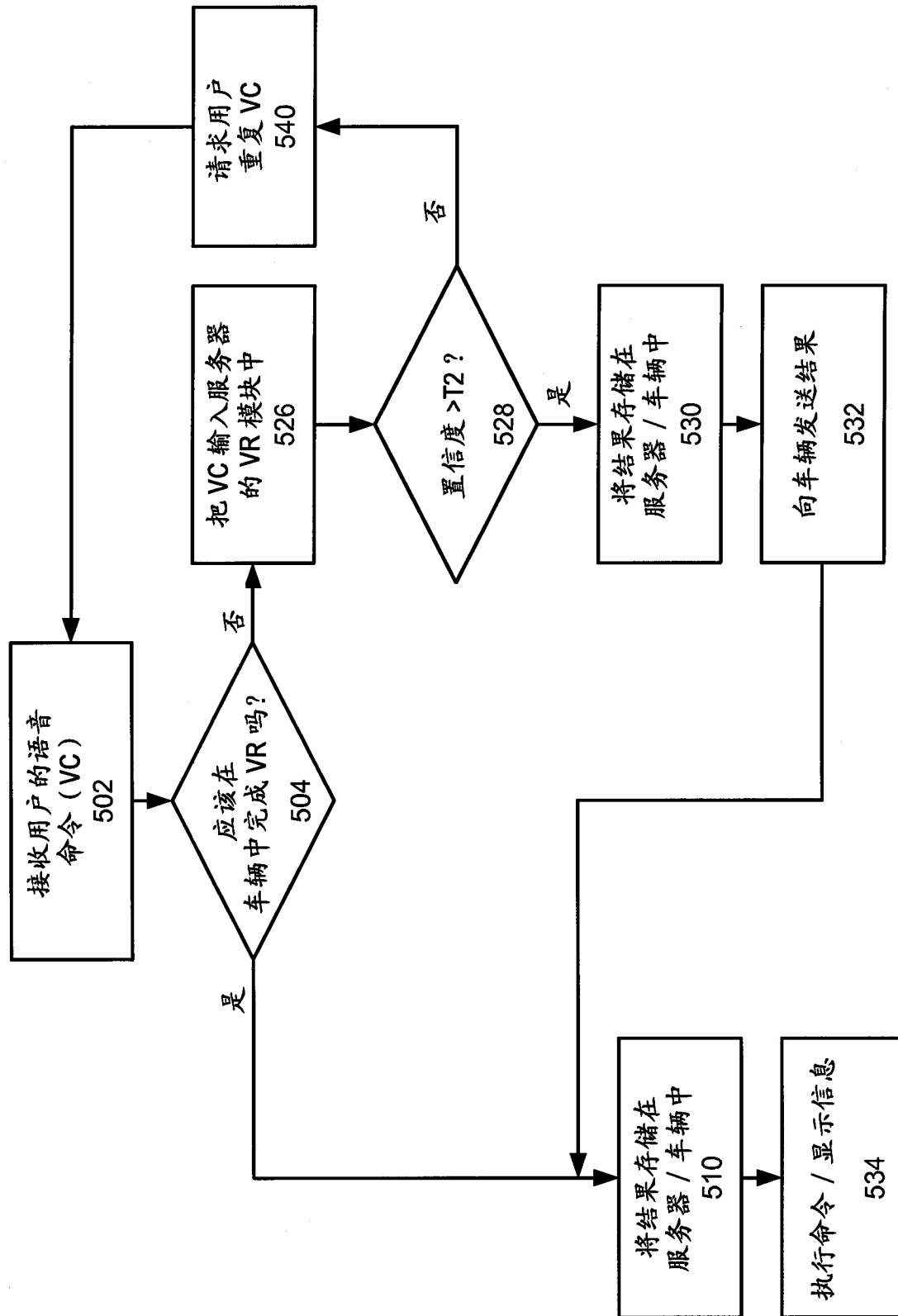


图 5

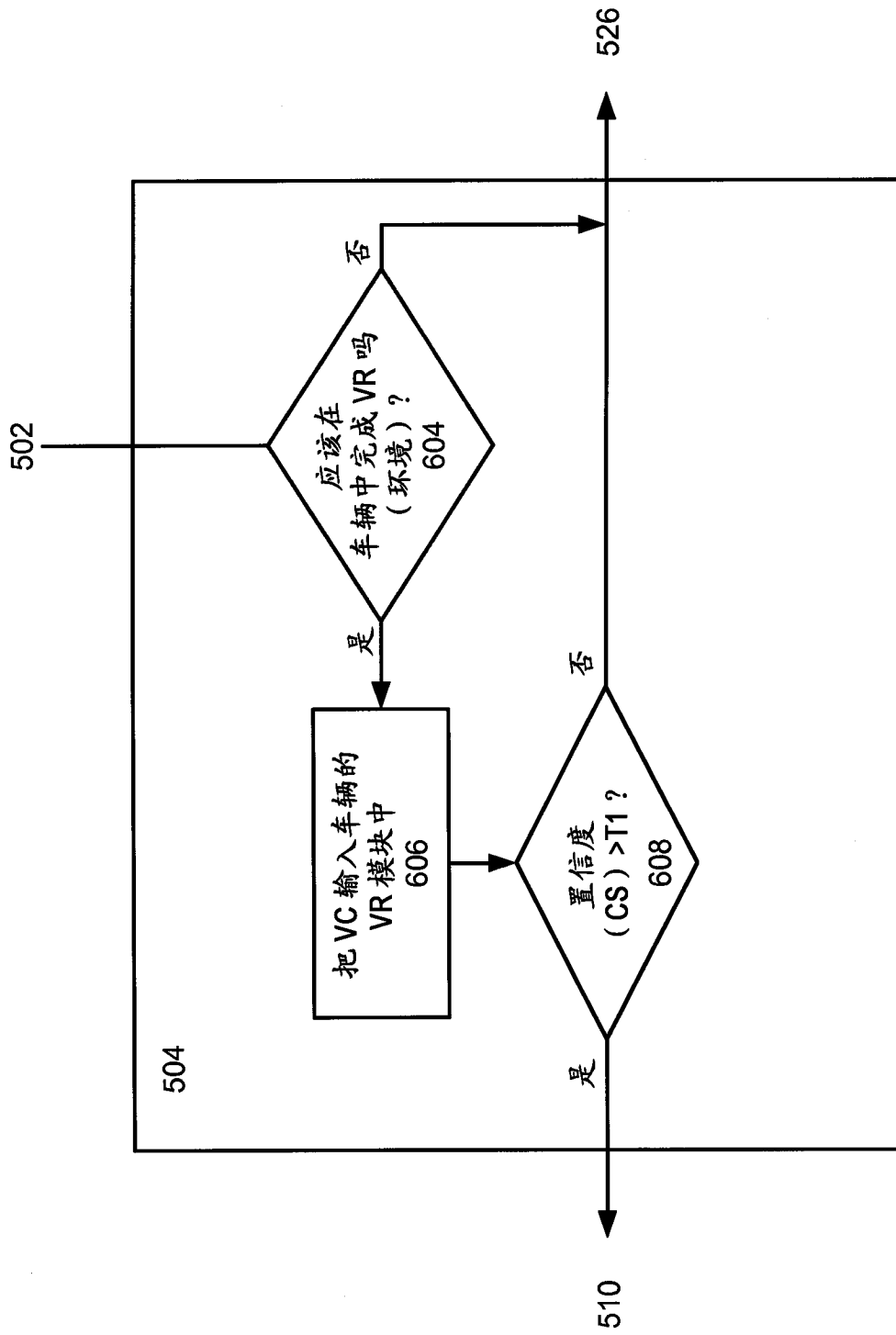


图 6

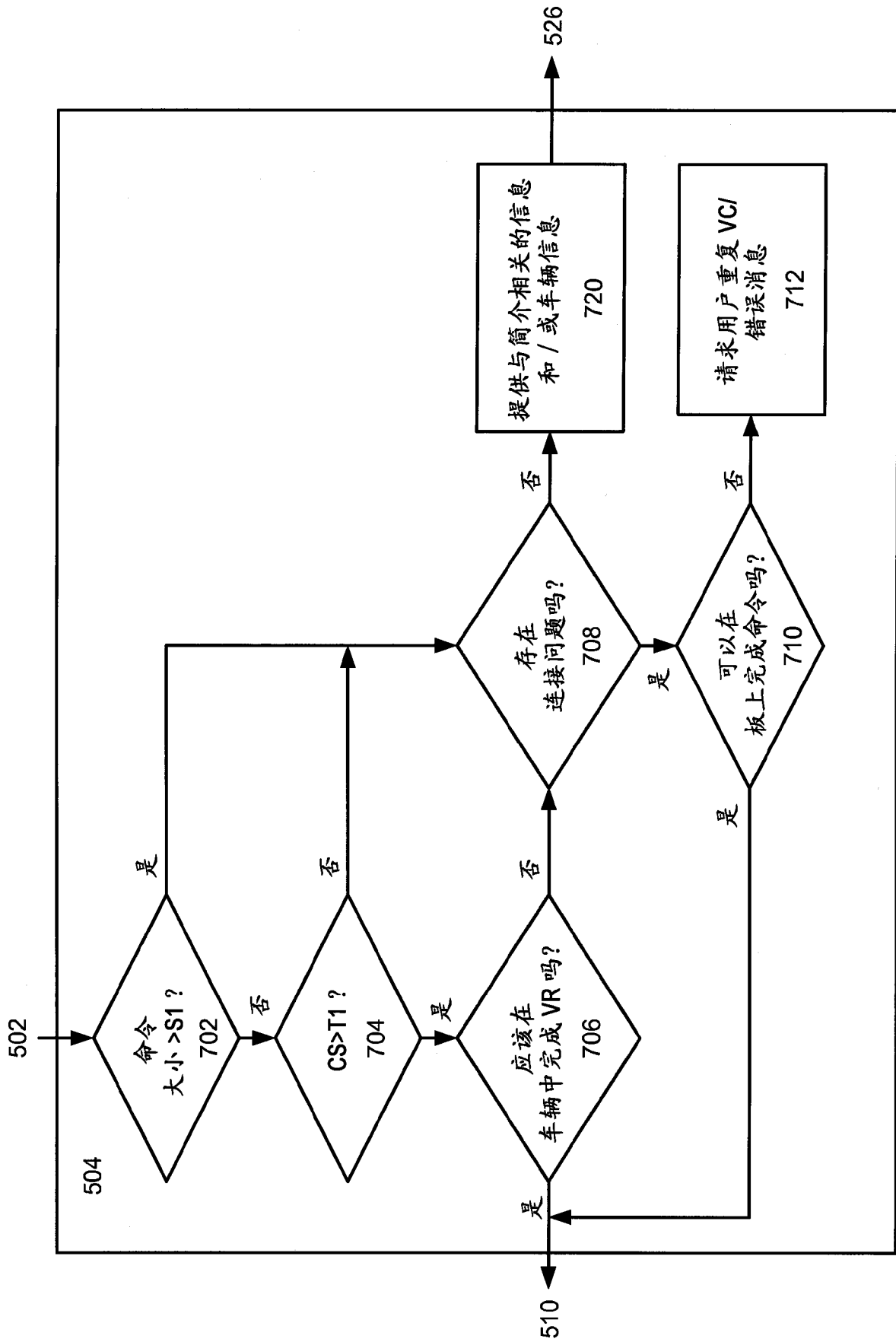


图 7

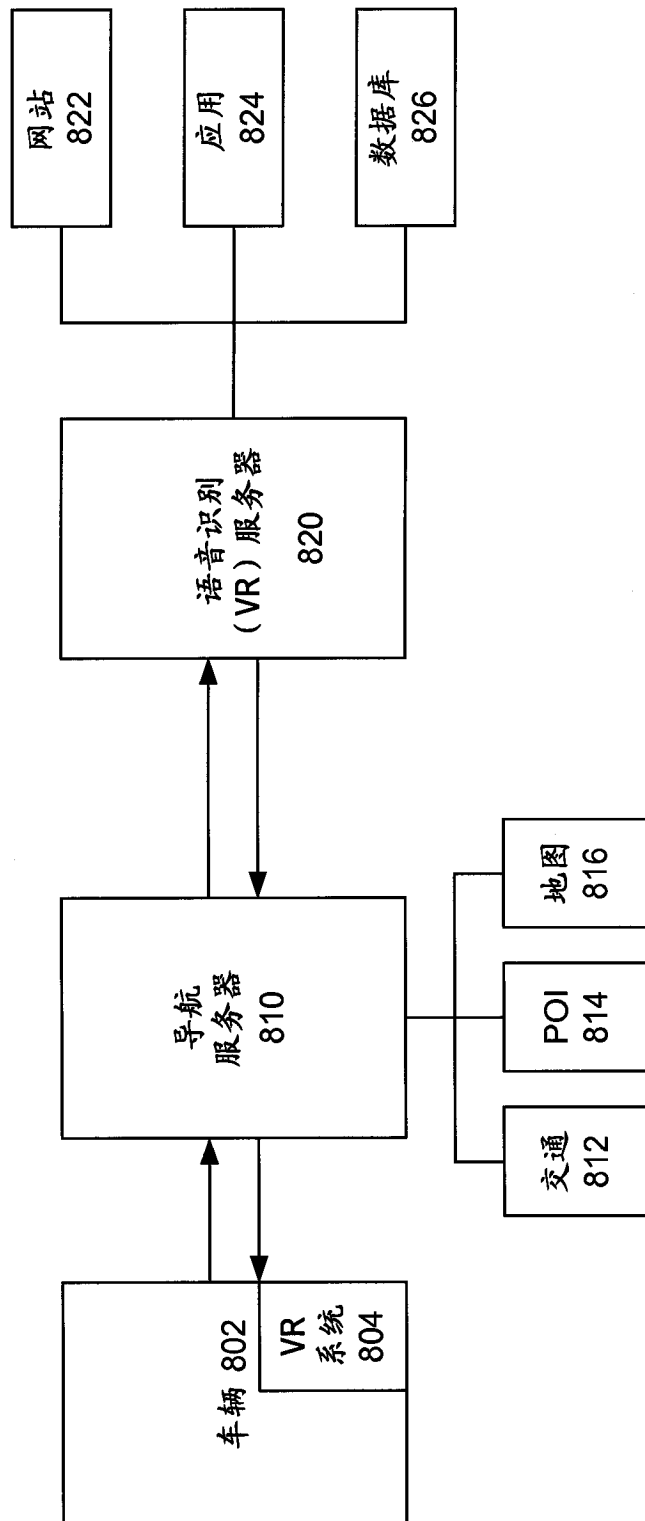


图 8