



(12) 发明专利申请

(10) 申请公布号 CN 105609109 A

(43) 申请公布日 2016. 05. 25

(21) 申请号 201510677501. 8

(22) 申请日 2015. 10. 16

(30) 优先权数据

14/515933 2014. 10. 16 US

(71) 申请人 通用汽车有限责任公司

地址 美国密执安州

(72) 发明人 J. L. 霍尔德伦 G. 塔尔瓦 X. 赵

M. J. 赫格

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 陈国慧 董均华

(51) Int. Cl.

G10L 15/28(2013. 01)

G10L 15/30(2013. 01)

G10L 15/22(2006. 01)

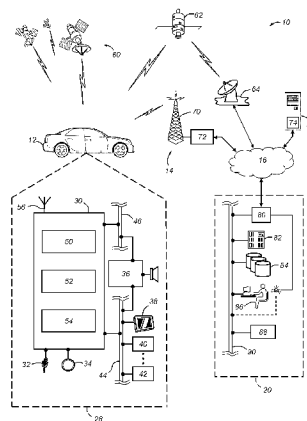
权利要求书1页 说明书13页 附图4页

(54) 发明名称

混合式自动语音识别

(57) 摘要

本发明提供混合式自动语音识别。一种向自动语音识别(ASR)系统提供车辆中接收到的语音的系统和方法,包括:在车辆处从车辆乘员接收语音;将接收到的语音提供至远程定位ASR系统和基于车辆的ASR系统;以及此后对于由基于车辆的ASR系统处理的语音确定置信级;当所确定的置信级在预定置信阈值之上时在车辆中呈现来自基于车辆的ASR系统的结果;当所确定的置信级不在预定置信阈值之上时在车辆中呈现来自远程定位ASR系统的结果。



1. 一种向自动语音识别(ASR)系统提供车辆中接收到的语音的方法,包括以下步骤:
 - (a)在车辆处从车辆乘员接收语音;
 - (b)将接收到的语音提供至远程定位ASR系统和基于车辆的ASR系统;以及此后
 - (c)对于由所述基于车辆的ASR系统处理的语音确定置信级;
 - (d)当所确定的置信级在预定置信阈值之上时,在车辆中呈现来自所述基于车辆的ASR系统的结果;
 - (e)当所确定的置信级不在预定置信阈值之上时,在车辆中呈现来自所述远程定位ASR系统的结果。
2. 根据权利要求1所述的方法,还包括以下步骤:将对于由所述基于车辆的ASR系统处理的语音确定的置信级与所述远程定位ASR系统的置信级进行比较,并且如果两个置信级均在距预定置信阈值的预定范围内,则呈现来自所述基于车辆的ASR系统和所述远程定位ASR系统两者的结果。
3. 根据权利要求1所述的方法,还包括以下步骤:确定在所述基于车辆的ASR系统处接收到的语音的上下文。
4. 根据权利要求3所述的方法,还包括以下步骤:在所述基于车辆的ASR系统处存储上下文分类器。
5. 根据权利要求4所述的方法,其中,所述上下文分类器还包括基于规则的分类器。
6. 根据权利要求4所述的方法,其中,所述上下文分类器还包括基于统计的分类器。
7. 根据权利要求1所述的方法,还包括以下步骤:在车辆中呈现来自所述基于车辆的ASR系统的多个结果。
8. 根据权利要求1所述的方法,还包括以下步骤:确定来自所述远程定位ASR系统的语音识别结果在预定量的时间期满之前已经到达。
9. 根据权利要求8所述的结果,还包括以下步骤:响应于来自所述远程定位ASR系统的语音识别结果在预定量的时间期满之前的到达,而允许在车辆中呈现语音识别结果。
10. 一种向自动语音识别(ASR)系统提供车辆中接收到的语音的方法,包括以下步骤:
 - (a)在车辆处从车辆乘员接收语音;
 - (b)在继续语音识别处理之前向接收到的语音应用上下文分类器;
 - (c)从所述上下文分类器的输出确定接收到的语音与基于车辆的语音处理相关联;以及
 - (d)基于步骤(c)将接收到的语音发送至所述基于车辆的ASR系统而不是远程定位ASR系统。

混合式自动语音识别

技术领域

[0001] 本发明涉及语音识别,并且更具体地涉及在本地以及在远程位置处执行的语音识别。

背景技术

[0002] 车辆乘员当操作车辆时使用自动语音识别(ASR)系统,以口头地传达各种命令或消息。当车辆乘员说话时,位于车辆处的话筒可以接收该语音,将语音转换为电信号,并且将信号传递至ASR系统,该ASR系统使用信号来确定接收到的语音的内容。ASR系统可以位于车辆中能够使用存储在车上的语法本地地实施语音识别的位置处。然而,也能够将接收到的语音无线地发送至远程定位ASR系统,在该处可以使用许多语法来确定语音的内容。

[0003] 在位于车辆上或者远程位置处的ASR系统处执行语音识别可以导致一些权衡。例如,在车辆处接收到并且使用车辆ASR系统处理的语音可以比如果接收到的语音发送至车外更快地开始语音识别。但是当与远程定位ASR系统相比时,存储在车辆处并且由车辆ASR系统使用的语法可能其内容上受限,或者车辆ASR系统的处理能力可能受限。与之相对,将接收到的语音无线地发送至远程定位ASR系统可能遭受与接收到的语音的无线发送以及有关接收到的语音的语音分析结果的无线接收相关的发送延迟。当车辆可以访问任一位置处的ASR系统时,将车辆中接收到的语音选择性地传达至车辆ASR系统、远程定位ASR系统或者两者可能增加响应时间。

发明内容

[0004] 根据一个实施例,一种方法包括向自动语音识别(ASR)系统提供车辆中接收到的语音。该方法包括在车辆处从车辆乘员接收语音;将接收到的语音提供至远程定位ASR系统和基于车辆的ASR系统;以及此后确定对于由基于车辆的ASR系统处理的语音的置信级;当所确定的置信级在预定置信阈值之上时在车辆中呈现来自基于车辆的ASR系统的结果;当所确定的置信级不在预定置信阈值之上时在车辆中呈现来自远程定位ASR系统的结果。

[0005] 根据另一实施例,一种方法包括向自动语音识别(ASR)系统提供车辆中接收到的语音。该方法包括在车辆处从车辆乘员接收语音;在继续语音识别处理之前对接收到的语音应用上下文分类器;从上下文分类器的输出确定接收到的语音与基于车辆的语音处理相关联;以及基于步骤(c)将接收到的语音发送至基于车辆的ASR系统而不是远程定位ASR系统。

[0006] 1.一种向自动语音识别(ASR)系统提供车辆中接收到的语音的方法,包括以下步骤:

- (a)在车辆处从车辆乘员接收语音;
- (b)将接收到的语音提供至远程定位ASR系统和基于车辆的ASR系统;以及此后
- (c)对于由所述基于车辆的ASR系统处理的语音确定置信级;
- (d)当所确定的置信级在预定置信阈值之上时,在车辆中呈现来自所述基于车辆的ASR

系统的结果；

(e)当所确定的置信级不在预定置信阈值之上时,在车辆中呈现来自所述远程定位ASR系统的结果。

[0007] 2.根据方案1所述的方法,还包括以下步骤:将对于由所述基于车辆的ASR系统处理的语音确定的置信级与所述远程定位ASR系统的置信级进行比较,并且如果两个置信级均在距预定置信阈值的预定范围内,则呈现来自所述基于车辆的ASR系统和所述远程定位ASR系统两者的结果。

[0008] 3.根据方案1所述的方法,还包括以下步骤:确定在所述基于车辆的ASR系统处接收到的语音的上下文。

[0009] 4.根据方案3所述的方法,还包括以下步骤:在所述基于车辆的ASR系统处存储上下文分类器。

[0010] 5.根据方案4所述的方法,其中,所述上下文分类器还包括基于规则的分类器。

[0011] 6.根据方案4所述的方法,其中,所述上下文分类器还包括基于统计的分类器。

[0012] 7.根据方案1所述的方法,还包括以下步骤:在车辆中呈现来自所述基于车辆的ASR系统的多个结果。

[0013] 8.根据方案1所述的方法,还包括以下步骤:确定来自所述远程定位ASR系统的语音识别结果在预定量的时间期满之前已经到达。

[0014] 9.根据方案8所述的结果,还包括以下步骤:响应于来自所述远程定位ASR系统的语音识别结果在预定量的时间期满之前的到达,而允许在车辆中呈现语音识别结果。

[0015] 10.根据方案1所述的方法,还包括以下步骤:将接收到的语音同时向基于远程的ASR系统和所述基于车辆的ASR系统提供。

[0016] 11.一种向自动语音识别(ASR)系统提供车辆中接收到的语音的方法,包括以下步骤:

(a)在车辆处从车辆乘员接收语音;

(b)在继续语音识别处理之前向接收到的语音应用上下文分类器;

(c)从所述上下文分类器的输出确定接收到的语音与基于车辆的语音处理相关联;以

及

(d)基于步骤(c)将接收到的语音发送至所述基于车辆的ASR系统而不是远程定位ASR系统。

[0017] 12.根据方案11所述的方法,还包括以下步骤:在所述基于车辆的ASR系统处存储所述上下文分类器。

[0018] 13.根据方案12所述的方法,其中,所述上下文分类器还包括基于规则的分类器。

[0019] 14.根据方案12所述的方法,其中,所述上下文分类器还包括基于统计的分类器。

[0020] 15.根据方案12所述的方法,还包括以下步骤:在车辆中呈现来自所述基于车辆的ASR系统的多个结果。

[0021] 16.根据方案11所述的方法,还包括以下步骤:接收来自所述远程定位ASR系统的语音识别结果并且经由音频系统在车辆中呈现这些结果。

附图说明

[0022] 以下将结合附图描述本发明的一个或多个实施例,其中相同的附图标记表示相同的元件,并且在附图中:

图1是示出了能够采用本文公开的方法的通信系统的实施例的方框图;以及

图2是示出了自动语音识别(ASR)系统的实施例的方框图;

图3是示出了向ASR系统提供车辆中接收到的语音的方法的实施例的流程图;以及

图4是示出了向ASR系统提供车辆中接收到的语音的方法的另一实施例的流程图。

具体实施方式

[0023] 以下描述的系统和方法通过向位于车辆处的自动语音识别(ASR)系统、位于远离车辆的位置的ASR系统或两者选择性地提供在车辆处接收到的语音,来提高语音识别结果返回至车辆乘员的速度。在一个实施方式中,在车辆处从车辆乘员接收到的语音可以同时向车辆处的ASR系统和远程ASR系统提供。车辆处的ASR系统可以在接收到的语音也正向远程定位ASR系统发送的同时,开始处理所接收到的语音。

[0024] 在过去,通过将接收到的语音提供至位于车辆处的ASR系统并且然后等待语音识别输出,来处理接收到的语音。如果来自车辆ASR系统的输出不令人满意,则车辆将然后将接收到的语音发送至远程ASR系统。通过交替地将接收到的语音提供至基于车辆的ASR系统并且随后提供至基于远程的ASR系统,由于减少了从车辆的无线通信的消耗,所以能够以降低的成本获得语音识别结果。然而,当车辆ASR系统无法满意地分析接收到的语音时,车辆乘员可能已经经历了在发出语音与车辆ASR系统确定其不能识别接收到的语音的时刻之间的延迟。

[0025] 当车辆ASR系统产生低于预定可接受概率或置信阈值的语音识别结果时,同时向车辆ASR系统和远程定位ASR系统提供接收到的语音导致更快的语音识别结果。在该情况下,当车辆ASR系统结果不可接受时,已经启动了远程ASR系统来产生对于所接收语音的语音识别结果。因此,由远程定位ASR系统产生的语音识别结果能够比如果车辆在确定了车辆处的语音识别不可接受之后才等待启动这样的处理的情形显著地更向前。通过在将语音提供至车辆ASR系统的同时向远程定位ASR系统发送接收到的语音,可以与车辆确定其语音识别结果不可接受同时或者紧随其后,已经产生了并在车辆处接收了远程语音识别结果。

[0026] 也可以通过分析接收到的语音的上下文并且使用该上下文来确定是否使用车辆ASR系统执行语音识别,或者将接收到的语音发送至远程定位ASR系统,来改进语音识别处理。车辆可以使用车辆ASR系统的预处理部分来识别关键词和/或统计地分析接收到的语音以识别接收到的语音的上下文。基于所确定的上下文,车辆可以确定接收到的语音将在车辆处更高效地处理,还是应该将接收到的语音无线地发送至远程定位的ASR。

[0027] 参照图1,示出了操作环境,包括移动车辆通信系统10并且可以用于实施本文公开的方法。通信系统10通常包括车辆12、一个或多个无线载波系统14、陆地通信网络16、计算机18以及呼叫中心20。应该理解的是,所公开的方法可以与许多不同系统一起使用,并且并非特别限于在此所示的操作环境。此外,系统10的架构、构成、设置和操作及其单个部件是在本领域中通常已知的。因此,以下段落简单地提供了一种这样的通信系统10的简单概要,然而在此未示出的其他系统也可以采用所公开的方法。

[0028] 车辆12在图示实施例中示出为客车,但是应该意识到也可以使用任何其他车辆,

包括摩托车、卡车、多功能越野车(SUV)、旅行车(RV)、船舶、飞机等。一些车辆电子器件28在图1中整体示出,并且包括远程信息处理单元30、话筒32,一个或多个按钮或其他控制输入部34、音频系统36、视频显示器38和GPS模块40以及许多车辆系统模块(VSM)42。这些装置中的一些可以直接连接至远程信息处理单元,例如像话筒32和按钮34,而其他装置使用诸如通信总线44或娱乐总线46等一个或多个网络连接件而间接地连接。适当的网络连接件的示例包括控制器局域网(CAN)、面向媒体的系统传输(MOST)、本地互联网(LIN)、本地局域网(LAN)、以及其他适当的连接件,诸如以太网或符合已知ISO、SAE和IEEE标准和规范的其他连接件,仅列举少数。

[0029] 远程信息处理单元30可以是安装在车辆中并且启用通过无线载波系统14并且经由无线网络的无线语音和/或数据通信的OEM安装(嵌入式)或配件市场装置。这使得车辆能够与呼叫中心20、其他远程信息处理单元或一些其他实体或装置通信。远程信息处理单元优选地使用无线电传输以建立与无线载波系统14的通信信道(语音信道和/或数据信道),使得能够通过信道发送和接收语音和/或数据传输。通过提供语音和数据通信两者,远程信息处理单元30使得车辆能够提供许多不同的服务,包括关于导航、电话、紧急援助、诊断、信息娱乐等的服务。数据可以经由数据连接、诸如经由通过数据信道的数据包传输而发送,或者使用本领域已知技术经由语音信道而发送。对于包括语音通信(例如采用在呼叫中心20处的现场顾问或语音响应单元)和数据通信(例如向呼叫中心20提供GPS位置数据或车辆诊断数据)两者的组合服务,系统可以采用通过语音信道的单个呼叫,并且根据需要通过语音信道在语音和数据传输之间切换,并且这可以使用本领域技术人员已知的技术来完成。

[0030] 根据一个实施例,远程信息处理单元30采用了根据GSM或CDMA标准的蜂窝通信,并且因此包括用于像免提呼叫那样的语音通信的标准蜂窝芯片集50、用于数据传输的无线调制解调器、电子处理装置52、一个或多个数字存储器装置54以及双频天线56。应该意识到,调制解调器可以通过存储在远程信息处理单元中并且由处理器52执行的软件来实施,或者其可以是位于远程信息处理单元30的内部或外部的单独硬件部件。调制解调器可以使用许多不同的标准或协议来操作,诸如EVDO、CDMA、GPRS和EDGE。车辆与其他联网装置之间的无线网络也可以使用远程信息处理单元30执行。为此目的,远程信息处理单元30可以配置为根据一个或多个无线协议无线地通信,诸如IEEE 802.11协议、WiMAX或蓝牙中的任意一种。当用于诸如TCP/IP等数据包交换通信时,远程信息处理单元可以采用静态IP地址配置,或者可以设置为从网络上的其它装置诸如路由器或者从网络地址服务器自动地接收分配的IP地址。

[0031] 处理器52可以是能够处理电子指令的任何类型的装置,包括微处理器、微控制器、主处理器、控制器、车辆通信处理器、以及专用集成电路(ASIC)。其可以是仅用于远程信息处理单元30的专用处理器,或者可以与其他车辆系统共用。处理器52执行各种类型的数字化存储指令,诸如存储在存储器54中的软件或固件程序,其使得远程信息处理单元能够提供各类服务。例如,处理器52可以执行程序或处理数据以执行本文所讨论的方法的至少一部分。

[0032] 远程信息处理单元30可以用于提供多种范围的车辆服务,包括向和/或从车辆的无线通信。这些服务包括:逐向指引和结合基于GPS的车辆导航模块40而提供的其他导航相

关服务;安全气囊展开通知和其他紧急或路边辅助相关服务,结合诸如车身控制模块(未示出)等一个或多个碰撞传感器接口模块而提供;使用一个或多个诊断模块的诊断报告;以及信息娱乐相关服务,其中音乐、网页、电影、电视节目、视频游戏和/或其他信息由信息娱乐模块(未示出)下载并且存储用于当前或稍后回放。以上列出的服务绝非是远程信息处理单元30所有能力的穷举列表,而仅仅是远程信息处理单元能够提供一些服务的枚举。此外,应该理解的是,至少一些前述模块可以以保存在远程信息处理单元30的内部或外部的软件指令的形式实施,它们可以是位于远程信息处理单元30的内部或外部的硬件部件,或者它们可以相互或者与位于遍布车辆的位置的其他系统集成和/或共用,仅列举几种可能性。在模块实施为位于远程信息处理单元30的外部的VSM42的情况下,它们可以利用车辆总线44来与远程信息处理单元交换数据和命令。

[0033] GPS模块40从GPS卫星的集群60接收无线电信号。模块40从这些信号能够确定用于向车辆驾驶员提供导航和其他位置相关服务的车辆位置。导航信息可以呈现在显示器38(或车辆内的其他显示器)上或者可以口头地呈现,诸如当提供逐向导航时所完成。导航服务可以使用专用车内导航模块(其可以是GPS模块40的一部分)来提供,或者可以经由远程信息处理单元30完成一些或所有导航服务,其中将位置信息发送至远程位置,用于向车辆提供导航地图、地图注释(感兴趣地点、餐馆等)、路线计算等等。可以将位置信息提供至呼叫中心20或其他远程计算机系统,诸如计算机18,为了诸如车队管理等其他目的。此外,可以经由远程信息处理单元30从呼叫中心20将新的或更新的地图数据下载至GPS模块40。

[0034] 除了音频系统36和GPS模块40之外,车辆12可以包括形式为电子硬件部件的其他车辆系统模块(VSM)42,其位于遍布车辆的位置并且通常从一个或多个传感器接收输入,并且使用感测到的输入来执行诊断、监控、控制、报告和/或其他功能。每个VSM42优选地由通信总线44连接至其他VSM,也连接至远程信息处理单元30,并且可以编程以运行车辆系统和子系统诊断测试。作为示例,一个VSM42可以是控制诸如燃料点火和点火定时等发动机运转的各个方面的发动机控制模块(ECM),另一个VSM42可以是调节车辆动力传动系的一个或多个部件的运转的动力传动系控制模块,并且另一个VSM42可以是管理位于遍布车辆的位置的各种电气部件(像车辆的电动车门锁和前灯)的车身控制模块。根据一个实施例,发动机控制模块装备有车载诊断(OBD)特征,其提供无数个实时数据,诸如从包括车辆排放传感器的各个传感器接收到的数据,并且提供允许技工快速识别并修理车辆内的故障的一系列标准化诊断故障代码(DTC)。如本领域技术人员意识到的那样,上述VSM仅是可以用于车辆12的一些模块的示例,因为大量其他模块也是可能的。

[0035] 车辆电子部件28还包括大量车辆用户接口,其向车辆乘员提供提供和/或接收信息的机构,包括话筒32、按钮34、音频系统36和视频显示器38。当在本文中使用术语“车辆用户接口”广义地包括任何适当形式的电子装置,包括硬件和软件部件,其位于车辆上并且使得车辆用户能够通过车辆的部件或与车辆的部件通信。话筒32向远程信息处理单元提供音频输入,以使得驾驶员或其他乘员能够提供语音命令并且经由无线载波系统14执行免提呼叫。为此目的,其可以采用本领域已知的人机接口(HMI)技术连接至车载自动语音处理单元。按钮34允许用户手动地输入至远程信息处理单元30中,以发起无线电话呼叫并且提供其他数据、响应或控制输入。可以使用单独的按钮来向呼叫中心20发起与常规服务辅助呼叫相对的紧急呼叫。音频系统36向车辆乘员提供音频输出,并且可以是专用独立系统或

者主车辆音频系统的一部分。根据在此所示的具体实施例,音频系统36操作性地联接至车辆总线44和娱乐总线46,并且能够提供AM、FM和卫星无线电、CD、DVD和其他多媒体功能。该功能可以结合或独立于如上所述的信息娱乐模块来提供。视频显示器38优选地是图形显示器,诸如仪表板上的触摸屏或者从挡风玻璃折回的平视显示器,并且可以用于提供许多输入和输出功能。也可以采用各种其他车辆用户接口,因为图1的接口仅是一个具体实施方式的示例。

[0036] 无线载波系统14优选地是蜂窝电话系统,包括多个蜂窝塔(仅示出一个)、一个或多个移动交换中心(MSC)72、以及需要将无线载波系统14与陆地网络16连接的任何其他联网部件。每个蜂窝塔70包括发送和接收天线以及基站,其中基站从不同的蜂窝塔直接地或者经由诸如基站控制器等中间设备连接至MSC72。蜂窝系统14可以实施任何适当的通信技术,包括例如像AMPS等模拟技术、或者诸如CDMA(例如CDMA 2000)或GSM/GPRS等较新的数字技术。如本领域技术人员将意识到的那样,各种蜂窝塔/基站/MSC设置是可能的,并且可以与无线系统14一起使用。例如,基站和蜂窝塔可以共同位于相同地点处或者它们可以相互远程地定位,每个基站可以负责单个蜂窝塔或者单个基站可以服务多个蜂窝塔,并且多个基站可以联接至单个MSC,仅列出少数可能的设置。

[0037] 除了使用无线载波系统14之外,也可以使用卫星通信形式的不同无线载波系统来提供与车辆的双向或单向通信。这可以使用一个或多个通信卫星62和上行链路发送站64而完成。单向通信可以例如是卫星无线电服务,其中由发送站64接收节目内容(新闻、音乐等),打包用于上传,并且然后发送至卫星62,卫星向订户广播节目。双向通信可以例如是使用卫星62在车辆12和基站64之间中继电话通信的卫星电话服务。如果使用,则除了无线载波系统14之外或者替代无线载波系统14,可以采用该卫星电话。

[0038] 陆地网络16可以是传统的路基远程通信网络,其连接至一个或多个固定电话并且将无线载波系统14连接至呼叫中心20。例如,陆地网络16可以包括公用交换电话网络(PSTN),例如用于提供硬连线电话、数据包交换通信和互联网基础设施。可以通过使用标准的有线网络、光纤或其他光学网络、电缆网络、电力线、其他无线网络诸如无线局域网(WLAN)、或提供宽带无线接入(BWA)的网络或其任意组合,来实施陆地网络16的一个或多个区段。此外,呼叫中心20无需经由陆地网络16连接,而是可以包括无线电话设备以使其能够与无线网络诸如无线载波系统14直接通信。

[0039] 计算机18可以是经由诸如互联网等专用或公用网络可访问的大量计算机中的一个。每个这种计算机18可以用于一个或多个目的,诸如可由车辆经由远程信息处理单元30和无线载波14访问的网页服务器。计算机18示出为操作远程定位的自动语音识别(ASR)系统74。以下将更详细地讨论远程定位ASR系统74的部件和功能。其他这些可访问的计算机18可以例如是:服务中心计算机,其中诊断信息和其他车辆数据可以从车辆经由远程信息处理单元30上传;客户端计算机,由车辆所有者或其他订户使用以用于访问或接收车辆数据或者设置或配置订户偏好或控制车辆功能等目的;或者第三方资料库,向其或从其提供车辆数据或其他信息,不论通过与车辆12还是呼叫中心20或两者通信。计算机18也可以用于提供诸如DNS服务等互联网连接性,或者作为使用DHCP或其他适当协议来向车辆12分配IP地址的网络地址服务器使用。

[0040] 呼叫中心20设计为向车辆电子部件28提供大量不同的系统后端功能,并且根据在

此所示的示例性实施例,通常包括一个或多个交换器80、服务器82、数据库84、现场顾问86、以及自动语音应答系统(VRS)88,所有这些是本领域已知的。这些各种呼叫中心部件优选地经由有线或无线局域网90相互联接。交换器80可以是专用分支交换(PBX)交换器,传送输入信号以使得语音传输通常由常规电话发送至现场顾问86或者使用VoIP发送至自动语音应答系统88。现场顾问电话也可以使用VoIP,如图1中虚线所示。VoIP和通过交换器80的其他数据通信经由连接在交换器80和网络90之间的调制解调器(未示出)实施。数据传输经由调制解调器传递至服务器82和/或数据库84。数据库84可以存储账号信息,诸如订户认证信息、车辆标识符、概貌记录、行为模式和其他相关订户信息。数据传输也可以由诸如802.11x、GPRS等无线系统进行。尽管图示实施例已经描述为将结合使用现场顾问86的人工呼叫中心20使用,但是将意识到呼叫中心可以替代地采用VRS88作为自动顾问,或者可以使用VRS 88与现场顾问86的组合。

[0041] 现在转向图2,示出了对可以用于启用当前公开方法的自动语音识别(ASR)系统210的例示性架构。一般而言,车辆乘员与ASR系统声音互动以用于一个或多个以下基本目的:训练系统理解车辆乘员的特定语音;存储离散的语音,诸如口语名签或口语控制词语,像数字或关键词;或者为了任何适当的的目的而识别车辆乘员的语音,诸如语音拨号、菜单导航、转录、服务请求、车辆装置或装置功能控制等等。

[0042] ASR系统210示出在车辆12中。然而,ASR系统210中包括的元件以及关于ASR系统210讨论的概念也可以在位于计算机18处的远程定位ASR系统74中找到,其中存在一些差异。例如,当与ASR系统210相比时,远程定位ASR系统74可以包括多种复杂处理能力和语言模型以及更多的最新语言模型。当使用远程定位ASR系统74时,车辆12可以对在车辆12处经由话筒32接收到的语音分包,并且通过无线载波系统14将语音无线地发送至远程定位ASR系统74。在输出结果之后,远程定位ASR系统74可以对语音识别结果分包并且将其无线地发送至车辆12。尽管远程定位ASR系统74示出为在计算机18中,但是其也可以位于系统74的其他地方,诸如在呼叫中心20的服务器82和数据库84中。在如何执行远程定位ASR系统的一个示例中,Google™提供了可以与由Droid™无线移动装置使用的Android™软件一起使用的应用程序设计接口(API)。如关于通信系统10所示,远程定位ASR系统74可以在计算机18、呼叫中心的服务器82/数据库84、或者位于远离车辆12的位置的其他基于计算机的服务器设施处实施。

[0043] 通常,ASR从人类语音提取声音数据,将声音数据与所存储的子字数据比较和对比,选择可以与其他所选子字连接的适当的子字,并且将连接的子字或词语输出以用于后处理,诸如口述或转录、地址簿拨号、存储至存储器、训练ASR模型或改写参数等等。

[0044] ASR系统对本领域人员而言是大体已知的,并且图2仅示出了一个特定的例示性ASR系统210。系统210包括诸如远程信息处理话筒32等用于接收语音的装置、以及声音接口33,声音接口33诸如是具有模数转换器以将语音数字化为声音数据的远程信息处理单元30的声卡。系统210还包括诸如远程信息处理存储器54等用于存储声音数据并存储语音识别软件和数据库的存储器、以及诸如远程信息处理器52等用于处理声音数据的处理器。处理器与存储器一起并且结合以下模块工作:用于将语音的声音数据流语法分析为诸如声音特征等参数表示的一个或多个前端处理器或预处理器软件模块212;用于将声音特征解码以产生对应于输入语音话语的数字子字或字输出数据的一个或多个解码器软件模块214;以

及用于使用来自解码器模块214的输出数据用于任何适当目的的一个或多个后处理器软件模块216。

[0045] 系统210也可以从任何其他适当的音频源31接收语音,其可以与采用以实线示出的预处理软件模块212直接通信,或者经由声音接口33而与其间接通信。音频源31可以包括例如诸如语音邮件系统等音频电话源、或者任何种类的其他电话服务。

[0046] 一个或多个模块可以用作对解码器模块214的输入。首先,语法和/或词典模型218可以提供规则,其管理哪个字词可以逻辑地跟在其他字词之后以形成有效的句子。广义地来讲,语法能够在任何给定时刻在任何给定ASR模式下限定系统210所预期的词汇的全域。例如,如果系统210处于用于训练命令的训练模式,则语法模块218可以包括系统210所已知并且由其使用的所有命令。在另一示例中,如果系统210处于主菜单模式,则有效的语法模型218可以包括由系统210预期的所有主菜单命令,诸如呼叫、拨号、退出、删除、目录等等。其次,声音模型220辅助选择对应于来自预处理器模块212的输入的最可能的子字或字词。第三,字词模型222和语句/语言模型224在将所选择的子字或字词放入到字词或语句的上下文中提供了规则、句式和/或语义。此外,语句/语言模型224可以在任何给定时刻在任何给定ASR模式下限定系统210所预期的语句全域和/或可以提供规则等,该规则管理哪些句子可以逻辑地跟在其他语句之后以形成有效的扩展语音。

[0047] 首先,从人类语音提取声音数据,其中车辆乘员向话筒32说话,话筒将话语转换为电信号并且将这些信号传达至声音接口33。话筒32中的声音响应元件将乘员的语音话语作为空气压力的变化捕捉,并且将话语转换为诸如直流电流或电压等模拟电信号的对应变化。声音接口33接收模拟电信号,首先对其取样以使得在离散的时刻捕捉模拟信号的值,并且然后将其量化以使得在每个取样时刻将模拟信号的幅度转换为数字语音数据的连续流。换言之,声音接口33将模拟电信号转换为数字电信号。数字数据是二进制位,其在远程信息处理存储器54中缓冲并且然后由远程信息处理器52处理,或者当它们最初由处理器52接收时可以实时地被处理。

[0048] 其次,预处理器模块212将数字语音数据的连续流变换为声音参数的离散序列。更具体地,处理器52执行预处理器模块212以将数字语音数据分段为例如10—30ms时长的重叠语音或声音帧。帧对应于声音子字,诸如音节、半音节、音素、双音素、音位等等。预处理器模块212还执行语音分析以从每个帧内从诸如随时间变化的特征向量等乘员语音提取声音参数。乘员语音内的话语可以表示为这些特征向量的序列。例如并且如本领域技术人员已知的那样,特征向量可以被提取并且可以包括例如音高、能量曲线、频谱属性和/或倒谱系数,其可以通过执行帧的傅立叶变换并且使用余弦变换而去关联声音频谱而获得。将覆盖了语音特定长度的声音帧和对应参数连接为待解码的语音的未知测试图形。

[0049] 预处理模块212还可以存储可以由基于规则的分类器或基于统计的分类器实施的上下文分类器。上下文分类器可以应用于从车辆乘员的接收到的语音识别的文本,并且用于识别该语音的会话上下文。通常而言,上下文分类器不涉及理解接收到的语音的精确内容而是理解语音上下文。例如,基于规则的分类器可以访问每个与字词列表相关联的多个存储的上下文。这些上下文和它们相关联的字词可以存储在语法模块218或者可由ASR 210访问的任何其他存储器位置中。当使用基于规则的分类器时,ASR系统210可以识别接收到的语音中的匹配与上下文相关联的一个或多个字词的一个或多个字词。当ASR系统210检测

到匹配字词时,ASR系统210可以确定与该字词关联的上下文。例如,基于规则的分类器可以对接收到的语音进行语法分析并且识别语音中“地址”和“方向”的存在。ASR系统210可以使用基于规则的分类器来确定所识别字词是否与上下文相关联。在该示例中,字词“地址”和“方向”可以与车辆导航上下文相关联。这些检测到的字词的存在可以然后使得基于规则的分类器将“导航”上下文分配至接收到的语音。在不同的示例中,ASR系统210可以检测字词“电子邮件”或“文本”并且确定那些字词与口述上下文相关联。

[0050] 基于统计的分类器可以在接收到的语音中识别单个字词或字词的组合,并且然后识别所提取的字词与特定上下文相关联的统计可能性。基于统计的分类器可以以各种方式实施。在一个示例中,基于统计的分类器可以分析所识别的文本,并且将其分类为上下文的预定集合,其指示了潜在的用户意图,诸如导航路线请求、感兴趣点、电话呼叫、或者电子邮件口述上下文。基于统计的分类器可以通过使用诸如支持向量机、信息论、基于熵测量的方法或神经网络等图形分类技术来注释所识别文本,并且使用这些技术分配对应的置信值。基于统计的分类器可以包括贝叶斯分类器、N元语法模型、和递归训练模型,列举少数。基于统计的分类器可以在一段时间内被训练以听取接收到的语音中的特定字词或字词的组合,并且然后在接收到的语音之后执行的一些动作之后,学习该动作的上下文。然后可以使用基于统计的分类器的训练来预测未来接收到的语音的上下文。在一个示例中,基于统计的分类器可以分析接收到的语音中包括的字词,并且然后得知作为分析字词的结果而已经使用了车辆12的GPS模块40。基于统计的分类器可以然后将“导航”上下文与分析得到的声音参数相关联。当基于统计的分类器搜集到字词或字词串以及与它们相关联的上下文时,基于统计的分类器能够将它们与未来提取的字词进行比较以确定可能的上下文。因此,当基于统计的分类器从接收到的语音提取字词,并且将它们与之前提取的字词或字词串以及它们相关联的上下文进行比较时,基于统计的分类器能够识别当前与过去参数之间的相似性。当存在相似性时,基于统计的分类器能够推断与过去的字词或字词组合相关联的上下文统计地可能适用于当前字词。

[0051] 第三,处理器执行解码器模块214以处理每个测试图形的输入特征向量。解码器模块214也称为识别引擎或分类器,并且使用所存储的语音的已知参考图形。类似于测试图形,参考图形定义为相关声音帧和对应参数的连接。解码器模块214将待识别的子字测试图形的声音特征向量与所存储的子字参考图形进行比较和对比,评估它们之间的差异或相似性的程度,并且最终使用判定逻辑来选择最佳匹配子字作为识别的子字。通常来讲,最佳匹配子字是对应于所存储已知参考图形的与通过本领域技术人员已知的用于分析和识别子字的各种技术中的任意一种所确定的测试图形具有最小相异性或者最高可能性是该测试图形的子字。这样的技术可以包括动态时间扭曲分类器、人工智能技术、神经网络、自由音位识别器、和/或诸如隐藏马尔可夫模型(HMM)引擎等概率图形匹配器。

[0052] HMM引擎对本领域技术人员已知用于产生声音输入的多语音识别模型假设。在最终经由对语音的特征分析而识别和选择代表了声音输入的最可能正确解码的识别输出时考虑该假设。更具体地,HMM引擎生成形式为子字词模型假设的“N-最佳”列表的统计模型,其根据HMM计算得到的置信值或者诸如通过应用贝叶斯定理而给定一个或另一个子字的观察到的声音数据序列的概率而排序。

[0053] 贝叶斯HMM方法识别对于声音特征向量的给定观察序列的与最可能话语或子字序

列对应的最佳假设,并且其置信值可以取决于各种因素,包括与输入声音数据相关联的声音信噪比。HMM还可以包括称为对角高斯混合的统计分布,其对于每个子字的每个观察到的特征向量产生可能性得分,该得分可以用于对假设的N-最佳列表重新排序。HMM引擎也可以识别并选择模型可能性得分最高的那个子字。

[0054] 以类似的方式,可以连接用于子字序列的单个HMM以建立单个或多个字词HMM。此后,可以生成并进一步评价单个或多个字词参考图形的N-最佳列表和相关参数值。

[0055] 在一个示例中,语音识别解码器214使用适当的声音模型、语法和算法来处理特征向量,以生成参考图形的N-最佳列表。当在本文中使用时,术语参考图形可以与模型、波形、模版、富信号模型、标本、假设或其他类型的参考互换。参考图形可以包括一系列代表了一个或多个字词或子字的特征向量,并且可以基于特定的说话者、说话风格和听觉环境条件。本领域技术人员将认识到,可以通过对ASR系统的适当的参考图形训练来生成参考图形并将其存储在存储器中。本领域技术人员还将认识到,可以操纵所存储的参考图形,其中基于参考图形训练与ASR系统实际使用之间的语音输入信号的差异来改变参考图形的参数值。例如,基于来自不同车辆乘员或不同声音条件的有限量训练数据,可以改变用于一个车辆乘员或某些声音条件训练的一组参考图形,并将其保存为用于不同车辆乘员或不同声音条件的另一组参考图形。换言之,参考图形无需是固定的,而是可以在语音识别期间调节。

[0056] 使用词汇表中的语法和任何适当的解码器算法及声音模型,处理器从存储器访问解释了测试图形的数个参考图形。例如,处理器可以生成N-最佳词汇表结果和参考图形的列表和对应的参数值,并且存储至存储器。示例性的参数值可以包括在词汇表的N-最佳列表中的每个参考图形的置信得分、以及相关关联的分段时长、可能性得分、信噪比(SNR)值等等。词汇表的N-最佳列表可以通过递减参数值的大小而排序。例如,具有最高置信得分的词汇表参考图形是第一个最佳参考图形等等。一旦建立了所识别子字的串,它们就可以用于利用来自字词模型222的输入来构造字词,并且采用来自语言模型224的输入来构造语句。

[0057] 最后,后处理器软件模块216从解码器模块214接收输出数据用于任何适当的目的。在一个示例中,后处理器软件模块216可以从单个或多个字词参考图形的N-最佳列表识别或选择其中一个参考图形作为所识别语音。在另一示例中,可以使用后处理器模块216来将声音数据转换为文本或数位以用于与ASR系统的其他方面或其他车辆系统一起使用。在又一示例中,可以使用后处理器模块216来向解码器214或预处理器212提供训练反馈。更具体地,后处理器216可以用于训练用于解码器模块214的声音模型,或者用于训练用于预处理器模块212的改写参数。

[0058] 方法或其部分可以在计算机程序产品中实施,计算机程序产品以计算机可读介质体现,并且包括可由一个或多个系统的一个或多个计算机的一个或多个处理器使用以使得系统实施一个或多个方法步骤的指令。计算机程序产品可以包括一个或多个软件程序,由以下构成:源代码、目标代码、可执行代码或其他格式的程序指令;一个或多个固件程序;或者硬件描述语言(HDL)文档;以及任何程序相关数据。数据可以包括数据结构、查找表、或任何其他适当格式的数据。程序指令可以包括程序模块、例行程序、程序、对象、部件等等。计算机程序可以在一个计算机或者相互通信的多个计算机上执行。

[0059] 程序可以在计算机可读介质上体现,其可以是非瞬时的并且可以包括一个或多个存储装置、制造物品等等。示例性的计算机可读介质包括计算机系统存储器,例如RAM(随机

存取存储器)、ROM(只读存储器);半导体存储器,例如EPROM(可擦除可编程ROM)、EEPROM(电可擦除可编程ROM)、闪存;磁盘或光盘或带等。计算机可读介质也可以包括计算机至计算机连接,例如,当数据通过网络或其它通信连接(有线、无线或其组合)传输或提供时。以上示例的任意组合也包括在计算机可读介质的范围内。因此应该理解的是,该方法可以至少部分地由能够执行对应于所公开方法的一个或多个步骤的指令的任何电子物品和/或装置来执行。

[0060] 现在转向图3,示出了向ASR系统提供车辆12中接收到的语音的方法300。方法300开始于步骤310,其中在车辆12处从车辆乘员接收语音。位于车辆12中的人可以通过向车辆12的话筒32中说话而与如上所讨论的ASR系统210互动。话筒32通信地链接至处理装置52,其可以使用ASR系统210开始对接收到的语音执行语音识别分析。由车辆乘员向ASR系统210提供的语音可以涉及大量上下文并且包括大范围的词汇表。在一个情形中,车辆乘员可能发出关于车辆功能的语音,其可以由车辆处的ASR系统容易地理解。可以训练车辆ASR系统以识别诸如通常随着车辆部件行驶而出现的“方向”和“感兴趣点”等字词或命令。然而,车辆乘员也可以对于关于非车辆上下文的语音请求语音识别。例如,车辆乘员可以依靠ASR系统来口述电子邮件消息。电子邮件消息的内容可以涉及许多上下文的任何一个(或多个)。方法300前进至步骤320。

[0061] 在步骤320处,将接收到的语音向远程定位ASR系统74和ASR系统210同时提供。与此同时,处理装置52开始处理接收到的语音,车辆远程信息处理单元30能够将接收到的语音的完整内容从车辆12无线地发送至远程定位ASR系统74,而与语音内容无关。当ASR系统210正在识别接收到的语音的内容时,其也正在通过无线载波系统14和陆地网络16从车辆远程信息处理单元30向远程定位ASR系统所处的计算机18被无线地发送。方法300前进至步骤330。

[0062] 在步骤330处,对于由基于车辆的ASR系统210处理的语音确定置信级。ASR系统210能够输出词汇表结果的N-最佳列表作为识别语音,并且向每个词汇表结果分配百分比形式的置信值。在一个示例中,ASR系统210能够分析接收到的语音,并且输出代表了语音可能解释的三个词汇表结果,并且具有42%、45%和47%的置信值。置信值可以代表ASR系统210已经正确地解释了所接收到的语音的置信级。方法300前进至步骤340。

[0063] 在步骤340处,当所确定的置信级在预定置信阈值之上时,在车辆12中呈现来自基于车辆的ASR系统210的结果。作为对每个词汇表结果产生置信值的一部分,ASR系统210可以将那些值与预定置信阈值进行比较。例如,预定置信阈值可以设定在40%。具有在该值之上的置信值的结果可以呈现给车辆乘员。使用以上示例值,ASR系统210能够以它们的47%、45%和42%的置信值的顺序从最高到最低输出语音的可能解释。

[0064] 然而,ASR系统210可以确定来自基于车辆ASR系统210的语音识别结果的所确定置信级低于预定置信阈值。在该情况下,处理装置52可以确定其是否已经接收到来自远程定位ASR系统74的语音识别结果。如果没有,则处理装置52可以选择为了语音识别结果而等待预定量的时间,在该时间之后处理装置52可以播放接收到的语音无法理解的提示。另一方面,如果处理装置52确定来自远程定位ASR系统74的语音识别结果已经到达或者在预定量的时间期满之前到达,则处理装置52可以确定结果是否可接受。例如,处理装置52可以将来自远程定位ASR系统74的结果与预定置信阈值进行比较。如果来自远程定位ASR系统74的结

果在预定置信阈值之上,则处理装置52可以经由音频系统36向车辆乘员可听见地播放它们。否则,如果来自ASR系统210和远程定位ASR系统74的结果均在预定阈值之下,则处理装置52可以拒绝来自两者的结果。在一个实施方式中,来自ASR系统210和远程定位ASR系统74两者的结果处在预定阈值稍微之上,诸如不多于百分之二十之上,处理装置52可以呈现来自ASR系统210和远程定位ASR系统74的结果。

[0065] 转向图4,示出了向ASR系统提供车辆12中接收到的语音的方法400。方法400开始于步骤410,其中在车辆12处从车辆乘员接收语音。该步骤可以如以上关于图3和步骤310描述的那样完成。方法400前进至步骤420。

[0066] 在步骤420处,在继续语音识别处理之前将上下文分类器应用于接收到的语音。车辆12处的ASR系统210能够使用其预处理模块212来识别接收到的语音的上下文。上下文分类器可以以不同方式实施,诸如通过使用基于规则的分类器或基于统计的分类器。如以上讨论的那样,上下文分类器能够识别包括在接收到的语音中的指示了对于语音的可识别上下文的关键词。或者在另一示例中,上下文分类器可以作用于所识别的文本,并且将其分类为用户意向的预定集合,在此称为上下文类别。为了执行统计分类,可以使用许多技术,诸如支持向量机、神经网络、N元语法模型,列举少数。上下文通常涉及由车辆乘员执行的任务。如以上讨论的那样,上下文的示例可以包括“导航”,其涉及至少部分地使用GPS模块40向车辆乘员提供逐向指引。“口述”可以是当车辆乘员通过与语音识别服务和消息客户端互动来发送电子邮件或SMS消息时的上下文。一旦上下文与接收到的语音相关联,方法400就前进至步骤430。

[0067] 在步骤430处,从上下文分类器的输出确定接收到的语音与基于车辆的语音处理相关联。当与远程地处理与一些上下文相关联的语音相比时,该上下文可以在位于车辆12处的ASR系统210处更高效地被处理。使用以上示例,ASR系统210可以具有语法和声音模型,其被调谐以响应于“导航”上下文以及其他车内通话中的通信。除了“导航”之外,其他车辆相关上下文也是可能的,诸如“车辆诊断”、“交通”或“感兴趣点”。方法400前进至步骤440。

[0068] 在步骤440处,基于确定接收到的语音的上下文是车辆相关的,而将接收到的语音发送至基于车辆的ASR系统而不是远程定位ASR系统。当从上下文分类器识别车辆相关上下文时,车辆12处的处理装置52能够确定将车辆12处的ASR系统210优化以如以上讨论的那样处理语音。然而,当车辆12处的ASR系统210确定了接收到的语音的上下文是非车辆相关的时,ASR系统210能够引导车辆远程信息处理单元30以将语音无线地发送至远程定位ASR系统74以用于远程语音处理。当车辆乘员在口述电子邮件消息时,这可能发生。车辆远程信息处理单元30可以然后接收来自车辆12处的远程语音处理的结果,并且经由音频系统36向车辆乘员呈现结果。方法400随后结束。在这样的方法中,其中可以从上下文分类器的输出确定接收到的语音与基于车辆的ASR不相关,方法可以替代地将语音发送至远程定位ASR系统,或者可以将其发送至车辆和远程定位ASR系统两者,如结合图3所讨论的那样。

[0069] 应该理解的是,前述是本发明的一个或多个实施例的描述。本发明不限于本文公开的具体实施例,而是仅由所附权利要求限定。此外,前述说明书中包含的陈述涉及具体实施例,并且不应解释为对本发明的范围或者权利要求中使用术语的定义的限制,除非以上明确地定义了术语或短语。各个其他实施例以及对所公开实施例的各种改变和修改将对本领域技术人员而言是显然的。所有这些其他实施例、改变和修改意在落入所附权利要求的

范围内。

[0070] 当在该说明书和权利要求中使用术语“举例”、“例如”、“比如”、“诸如”和“等”以及动词“包括”、“具有”、“包含”和它们的其他动词形式，当结合一个或多个部件的列表或其他项目使用时，每个应该解释为开放式的，意味着列举不应视作排除其他额外的部件或项目。其他术语应该使用它们最广泛的合理含义来解释，除非它们在需要不同解释的上下文中使用。

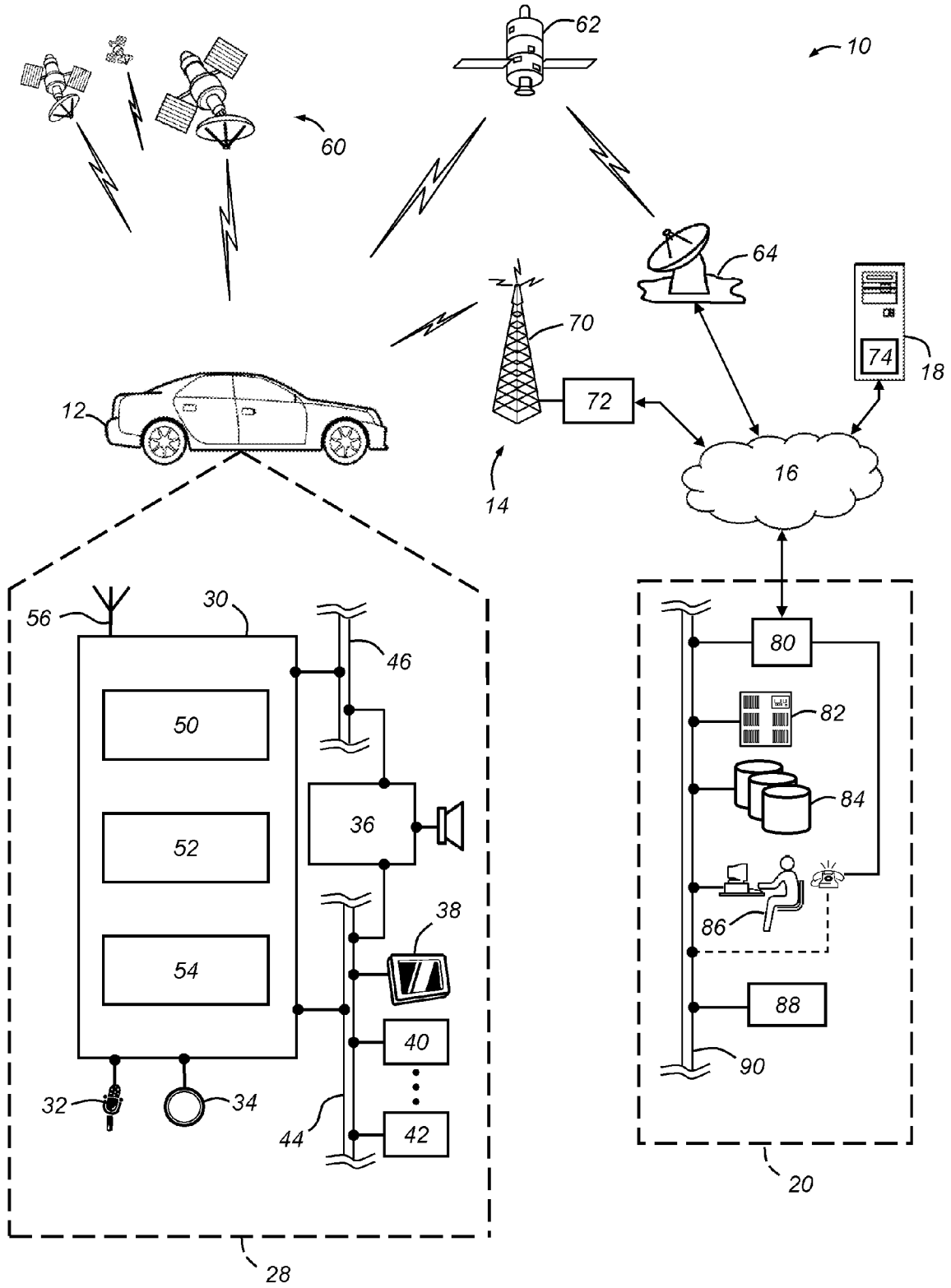


图 1

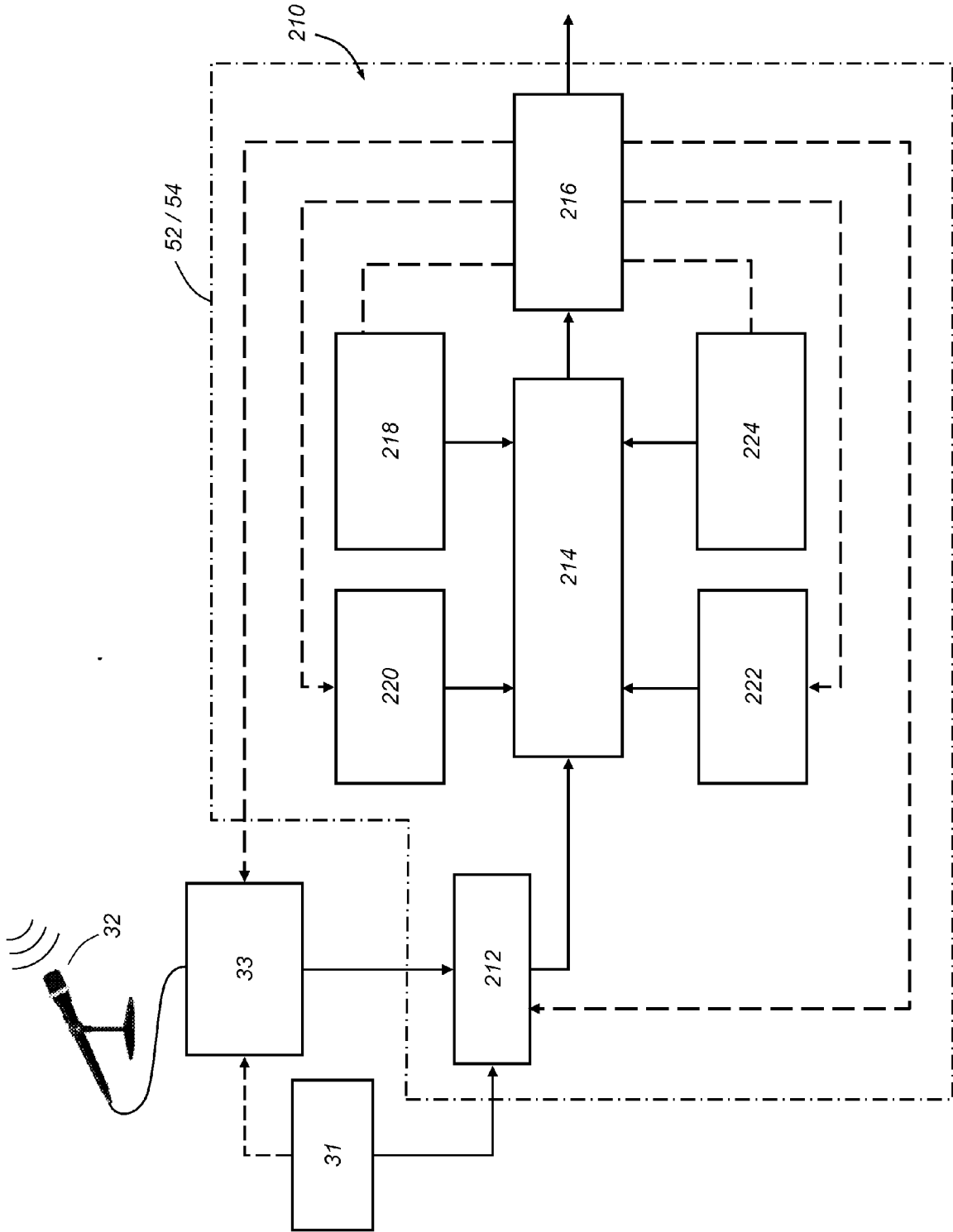


图 2

300

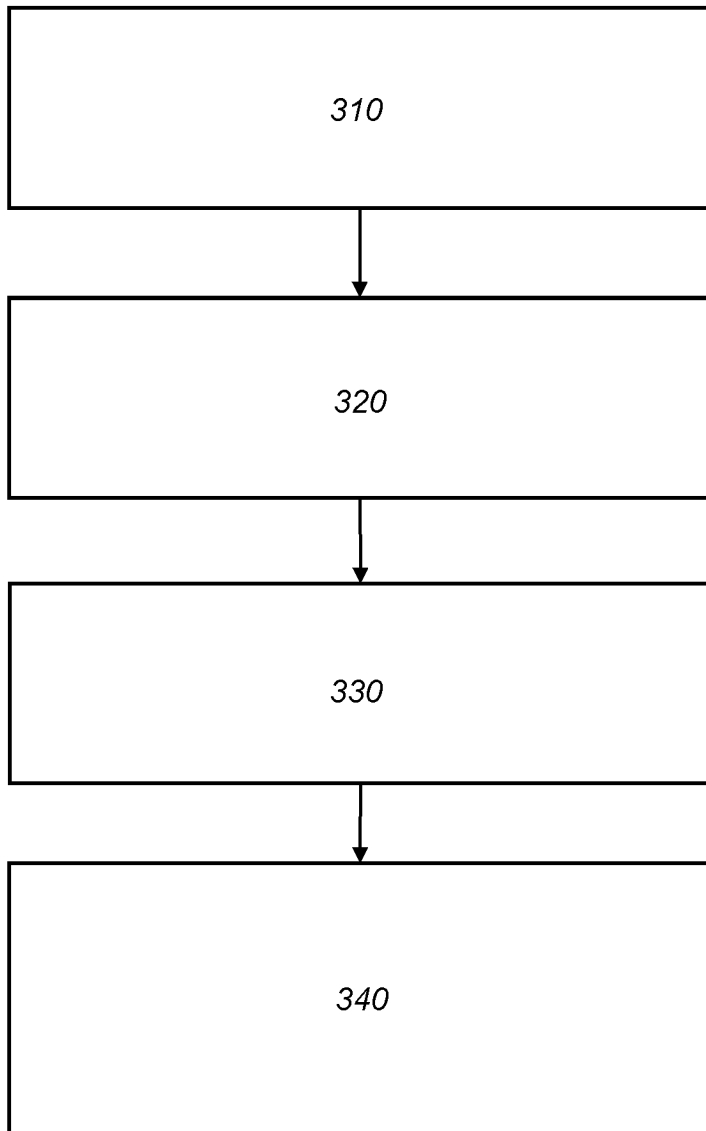


图 3

400

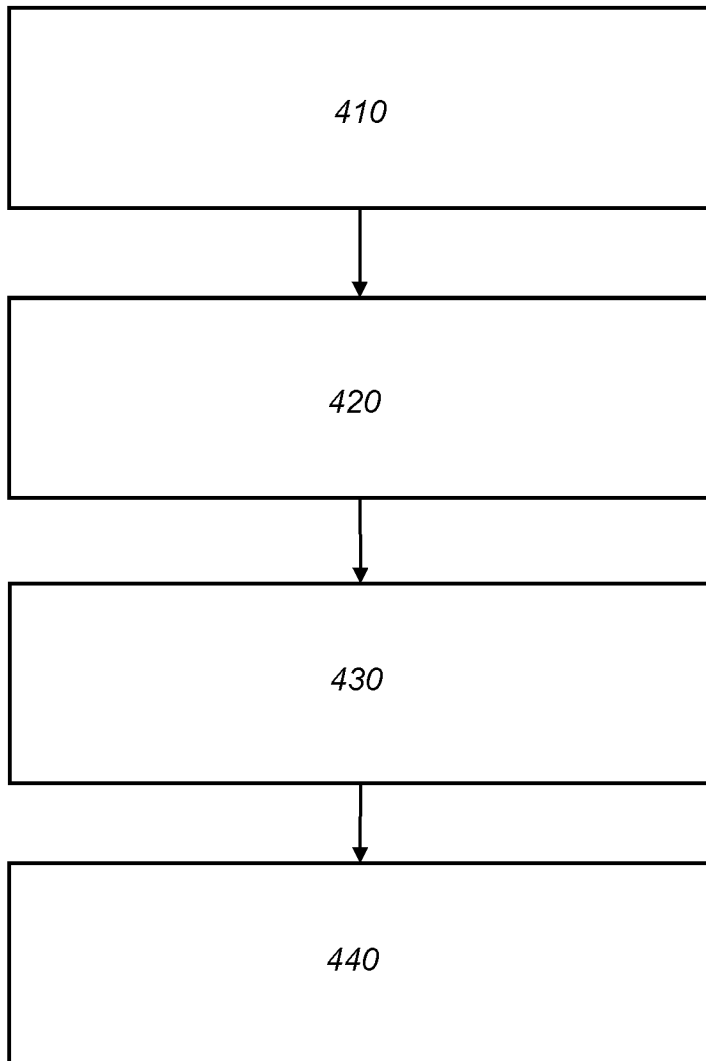


图 4