



(12) 发明专利

(10) 授权公告号 CN 114764022 B

(45) 授权公告日 2024.06.14

(21) 申请号 202111642752.4

(22) 申请日 2021.12.29

(65) 同一申请的已公布的文献号
申请公布号 CN 114764022 A

(43) 申请公布日 2022.07.19

(30) 优先权数据
17/248,196 2021.01.13 US

(73) 专利权人 百度(美国)有限责任公司
地址 美国加利福尼亚州桑尼维尔波尔多道
1195

(72) 发明人 孙宏艺 许珂诚 罗琦 林泽俊
王为 雷诺兹卫斯理 胡江滔
缪景皓

(74) 专利代理机构 北京清亦华知识产权代理事务
所(普通合伙) 11201
专利代理师 杜月

(51) Int.Cl.
G01H 11/06 (2006.01)
G01S 5/20 (2006.01)
G01S 5/22 (2006.01)
G06N 20/00 (2019.01)

(56) 对比文件
CN 110800031 A, 2020.02.14
审查员 吕新强

权利要求书2页 说明书9页 附图7页

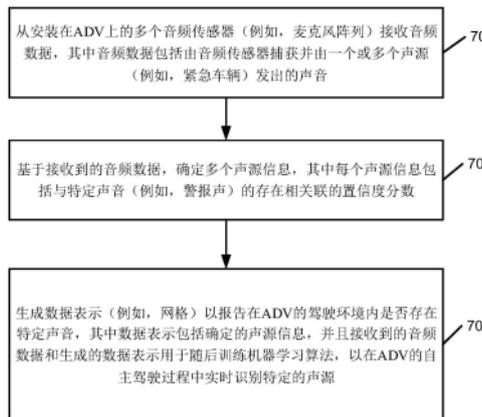
(54) 发明名称

用于自动驾驶车辆的声源检测和定位的系统和方法

(57) 摘要

公开了用于利用自动驾驶车辆(ADV)进行声源检测和定位的系统和方法。方法包括从安装在ADV上的多个音频传感器接收音频数据。音频数据包括由音频传感器捕获并由一个或多个声源发出的声音。基于接收到的音频数据,方法进一步包括确定多个声源信息。每个声源信息包括与特定声音的存在相关联的置信度分数。方法进一步包括生成数据表示以报告在ADV的驾驶环境内是否存在特定声音。数据表示包括确定的声源信息。接收到的音频数据和生成的数据表示用于随后训练机器学习算法,以在ADV的自动驾驶过程中实时识别特定的声源。

700



1. 一种用于在自动驾驶车辆 (ADV) 在驾驶环境内运行时利用ADV进行声源检测和定位的方法,所述方法包括:

从安装在ADV上的多个音频传感器接收音频数据,所述音频数据包括由多个音频传感器捕获并由一个或多个声源发出的声音;

基于接收到的音频数据,利用多个音频传感器执行声源定位以确定多个声源信息,每个声源信息包括与特定声音的存在相关联的置信度分数,并且进一步包括声源相对于相应音频传感器的方向、声源与相应音频传感器之间的距离、捕获的声音的相对位置、捕获的声音的绝对位置、捕获的声音的接近/离开状态、或与当前时间戳相关联的捕获的声音的强度中的至少一项;以及

生成数据表示以报告ADV的驾驶环境中是否存在特定声音,所述数据表示包括共同覆盖所述ADV的驾驶环境的多个区域的网格,每个区域对应于所述多个音频传感器中的音频传感器并报告指示区域内是否存在特定声音的结果的向量,所述结果的向量包括区域标识 (ID) 和一个声源信息;

其中接收到的音频数据和生成的数据表示用于随后训练机器学习算法以在ADV的自动驾驶期间实时识别特定声源。

2. 根据权利要求1所述的方法,其中每个区域被配置为部分地覆盖所述驾驶环境内的特定尺寸。

3. 根据权利要求1所述的方法,其中,所述声源为应急车辆,所述特定声音为警报声。

4. 根据权利要求1所述的方法,其中,所述置信度分数在值0~1的范围内。

5. 根据权利要求1所述的方法,其中,所述网格的中心表示所述ADV的位置。

6. 一种具有存储在其中的指令的非暂时性机器可读介质,所述指令在由处理器执行时使所述处理器执行如权利要求1至5中任一项所述的方法的操作。

7. 一种用于声源检测和定位的系统,包括:

处理器;以及

耦接到处理器且存储指令的存储器,当指令被处理器执行时,使处理器执行如权利要求1至5中任一项所述的方法的操作。

8. 一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时使得所述处理器执行如权利要求1至5中任一项所述的方法。

9. 一种用于声源检测和定位的系统,包括:

传感器系统,包括安装在ADV上的多个音频传感器,用于接收音频数据,所述音频数据包括由多个音频传感器捕获并由一个或多个声源发出的声音;

声源检测和定位模块,用于基于接收到的音频数据,利用多个音频传感器执行声源定位以确定多个声源信息,每个声源信息包括与特定声音的存在相关联的置信度分数,并且进一步包括声源相对于相应音频传感器的方向、声源与相应音频传感器之间的距离、捕获的声音的相对位置、捕获的声音的绝对位置、捕获的声音的接近/离开状态、或与当前时间戳相关联的捕获的声音的强度中的至少一项;以及生成数据表示以报告ADV的驾驶环境中是否存在特定声音,所述数据表示包括共同覆盖所述ADV的驾驶环境的多个区域的网格,每个区域对应于所述多个音频传感器中的音频传感器并报告指示区域内是否存在特定声音的结果的向量,所述结果的向量包括区域标识 (ID) 和一个声源信息,

其中接收到的音频数据和生成的数据表示用于随后训练机器学习算法以在ADV的自主驾驶期间实时识别特定声源。

10. 根据权利要求9所述的系统, 其中, 每个区域被配置为部分地覆盖所述驾驶环境内的特定尺寸。

11. 根据权利要求9所述的系统, 其中, 所述声源为应急车辆, 所述特定声音为警报声。

12. 根据权利要求9所述的系统, 其中, 所述置信度分数在值0~1的范围内。

13. 根据权利要求9所述的系统, 其中, 所述网格的中心表示所述ADV的位置。

用于自动驾驶车辆的声源检测和定位的系统和方法

技术领域

[0001] 本公开的实施例一般涉及操作自动驾驶车辆(autonomous driving vehicle, ADV)。更特别地,本公开的实施例涉及用于自动驾驶车辆的声源检测和定位。

背景技术

[0002] 以自主模式(例如,无人驾驶)操作的车辆可以减轻乘坐者,尤其是驾驶员的一些驾驶相关责任。当以自主模式操作时,车辆可以使用车载传感器导航到各种位置,从而允许车辆以最小的人机交互或者在没有任何乘客的一些情况下行驶。

[0003] 运动规划和控制是自动驾驶中的关键操作。然而,传统的运动规划操作主要从其曲率和速度来估计完成给定路径的难度,而没有考虑不同车辆类型的特征差异。相同的运动规划和控制被应用于所有类型的车辆,在某些情况下可能不准确并且不流畅。

[0004] 此外,运动规划和控制的操作通常需要感知周围的障碍物或物体。它们还需要实时检测特定的声源(例如,应急车辆)并对声源做出实时响应。虽然对于这种声源的检测和定位存在不同的解决方案,但是这些解决方案中的几个单独步骤是耗时的并且没有考虑到每一个线索。

发明内容

[0005] 第一方面,提供一种用于在自动驾驶车辆(ADV)在驾驶环境内运行时利用ADV进行声源检测和定位的方法,所述方法包括:

[0006] 从安装在ADV上的多个音频传感器接收音频数据,所述音频数据包括由多个音频传感器捕获并由一个或多个声源发出的声音;

[0007] 基于接收到的音频数据,确定多个声源信息,每个声源信息包括与特定声音的存在相关联的置信度分数;以及

[0008] 生成数据表示以报告ADV的驾驶环境中是否存在特定声音,所述数据表示包括所确定的多个声源信息;

[0009] 其中接收到的音频数据和生成的数据表示用于随后训练机器学习算法以在ADV的自动驾驶期间实时识别特定声源。

[0010] 第二方面,提供一种具有存储在其中的指令的非暂时性机器可读介质,所述指令在由处理器执行时使所述处理器执行如第一方面所述的方法的操作。

[0011] 第三方面,提供一种用于声源检测和定位的系统,包括:

[0012] 处理器;以及

[0013] 耦接到处理器且存储指令的存储器,当指令被处理器执行时,使处理器执行如第一方面所述的方法的操作。

[0014] 第四方面,提供一种计算机程序产品,包括计算机程序,,所述计算机程序在被处理器执行时使得所述处理器执行如第一方面所述的方法。

[0015] 第五方面,提供一种用于声源检测和定位的系统,包括:

[0016] 传感器系统,包括安装在ADV上的多个音频传感器,用于接收音频数据,所述音频数据包括由多个音频传感器捕获并由一个或多个声源发出的声音;

[0017] 声源检测和定位模块,用于基于接收到的音频数据,确定多个声源信息,每个声源信息包括与特定声音的存在相关联的置信度分数;以及生成数据表示以报告ADV的驾驶环境中是否存在特定声音,所述数据表示包括所确定的多个声源信息,

[0018] 其中接收到的音频数据和生成的数据表示用于随后训练机器学习算法以在ADV的自动驾驶期间实时识别特定声源。

[0019] 通过本公开,可以实现对一个或多个声源的检测、定位和跟踪,提供驾驶安全性。

附图说明

[0020] 本公开的实施例通过示例的方式示出并且不限于附图中的图,在附图中相同的附图标记表示相似的元件。

[0021] 图1是示出根据一个实施例的联网系统的框图。

[0022] 图2是示出根据一个实施例的自动驾驶车辆(ADV)的示例的框图。

[0023] 图3A-3B是示出根据一个实施例的与自动驾驶车辆一起使用的自动驾驶系统的示例的框图。

[0024] 图4A-4B是示出根据一个实施例的用于声源检测和定位的系统的框图。

[0025] 图5是示出根据一个实施例的使用用于声源检测和定位的系统的示例驾驶场景的图。

[0026] 图6A-6B是示出根据一个实施例的训练机器学习算法/模型的系统的框图。

[0027] 图7是根据一个实施例的利用ADV的用于声源检测和定位的示例方法的流程图。

具体实施方式

[0028] 将参考以下讨论的细节描述本公开的各个实施例和方面,并且附图将示出各个实施例。以下描述和附图是本公开的说明并且不应被解释为限制本公开。描述了许多具体细节以提供对本公开的各个实施例的全面理解。然而,在某些情况下,为了提供对本公开的实施例的简要讨论,没有描述公知或常规的细节。

[0029] 说明书中对“一个实施例”或“实施例”的引用意味着结合该实施例描述的特定特征、结构或特性可包括在本公开的至少一个实施例中。在说明书中的各个地方出现的短语“在一个实施例中”不一定都指同一实施例。

[0030] 本申请的方面涉及一次性声音检测(例如,警报声音检测)和定位算法,其可以执行对一个或多个声源(其可被称为“只听一次”(YOHO))的检测、定位和跟踪。

[0031] 根据一个方面,描述了一种在自动驾驶车辆(ADV)在驾驶环境内运行时利用ADV进行声源检测和定位的计算机实现方法。方法包括从安装在ADV上的多个音频传感器接收音频数据。音频数据包括由音频传感器捕获并由一个或多个声源发出的声音。基于接收到的音频数据,方法进一步包括确定多个声源信息。每个声源信息包括与特定声音的存在相关联的置信度分数。方法进一步包括生成数据表示以报告在ADV的驾驶环境内是否存在特定声音。数据表示包括确定的声源信息。接收到的音频数据和生成的数据表示用于随后训练机器学习算法,以在ADV的自动驾驶过程中实时识别特定的声源。

[0032] 其他方面涉及一种非暂时性机器可读介质和一种用于声源检测和定位的系统,其执行与上述方面类似的操作,并在下面更详细地描述。

[0033] 图1是示出根据本公开的一个实施例的自动驾驶网络配置的框图。参考图1,网络配置100包括自动驾驶车辆(ADV) 101,自动驾驶车辆101可以通过网络102通信地耦接到一个或多个服务器103-104。尽管示出一个ADV,但是多个ADV可以通过网络102彼此耦接和/或耦接到服务器103-104。网络102可以是任何类型的网络,诸如局域网(LAN)、诸如因特网的广域网(WAN)、蜂窝网络、卫星网络、或其组合,有线或无线。服务器103-104可以是任何类型的服务器或服务器群集,诸如Web或云服务器、应用服务器、后端服务器或其组合。服务器103-104可以是数据分析服务器、内容服务器、交通信息服务器、地图和兴趣点(map and point of interest, MPOI)服务器或位置服务器等。

[0034] ADV指的是能够被配置为处于自主模式的车辆,在自主模式中,车辆在很少或没有驾驶员输入的情况下导航通过环境。这种ADV可包括具有一个或多个传感器的传感器系统,传感器被配置为检测关于车辆操作于其中的环境的信息。车辆及其相关联的控制器使用检测的信息导航通过环境。ADV 101可以以手动模式、全自主模式或部分自主模式操作。

[0035] 在一个实施例中,ADV 101包括但不限于自动驾驶系统(ADS) 110、车辆控制系统111、无线通信系统112、用户接口系统113和传感器系统115。ADV 101进一步可包括在普通车辆中包括的某些常见组件,例如引擎、车轮、转向盘、变速器等,这些组件可以被车辆控制系统111和/或ADS 110使用各种通信信号和/或命令,诸如,例如加速信号或命令、减速信号或命令、转向信号或命令、制动信号或命令等控制。

[0036] 组件110-115可以经由互连、总线、网络或其组合彼此通信地耦接。例如,组件110-115可以经由控制器局域网(CAN)总线彼此通信地耦接。CAN总线是一种车辆总线标准,其被设计为允许微控制器和设备在没有主机的应用中彼此通信。它是基于消息的协议,最初被设计用于车辆内的多路电气布线,但是也用于许多其它环境中。

[0037] 现在参考图2,在一个实施例中,传感器系统115包括但不限于一个或多个相机211、全球定位系统(GPS)单元212、惯性测量单元(IMU) 213、雷达单元214以及光检测和测距(light detection and range, LIDAR)单元215。GPS系统212可包括可操作以提供关于ADV的位置的信息的收发器。IMU单元213可基于惯性加速度感测ADV的位置和朝向变化。雷达单元214可以表示使用无线电信号感测ADV的局部环境内的对象的系统。在一些实施例中,除了感测对象之外,雷达单元214还可以额外地感测对象的速度和/或航向。LIDAR单元215可以使用激光感测ADV所处的环境中的对象。LIDAR单元215可包括一个或多个激光源、激光扫描器以及一个或多个检测器,以及其他系统组件。相机211可包括一个或多个设备以捕获ADV周围的环境的图像。相机211可以是静态相机和/或摄影机。相机可以是机械可移动的,例如通过将相机安装在旋转和/或倾斜的平台上。

[0038] 传感器系统115还可包括其它传感器,诸如声纳传感器、红外传感器、转向传感器、油门传感器、制动传感器和音频传感器(例如麦克风)。音频传感器可以被配置为捕获来自ADV周围的环境的声音。转向传感器可以被配置为感测转向盘、车辆的车轮或其组合的转向角。油门传感器和制动传感器分别感测车辆的油门位置和制动位置。在一些情况下,油门传感器和制动传感器可以集成为集成的油门/制动传感器。

[0039] 在一个实施例中,车辆控制系统111包括但不限于转向单元201、油门单元202(也

称为加速单元)和制动单元203。转向单元201用于调节车辆的方向或航向。油门单元202用于控制马达或引擎的速度,马达或引擎的速度进而控制车辆的速度和加速度。制动单元203通过提供摩擦力以使车辆的车轮或轮胎变慢来使车辆减速。注意,图2所示的组件可以以硬件、软件或其组合实现。

[0040] 返回参考图1,无线通信系统112允许ADV 101和外部系统,诸如设备、传感器、其他车辆等之间的通信。例如,无线通信系统112可以直接或经由通信网络与一个或多个设备无线通信,诸如经由网络102与服务器103-104通信。无线通信系统112可以使用任何蜂窝通信网络或无线局域网(WLAN),例如使用WiFi与另一组件或系统通信。无线通信系统112可以例如使用红外链路、蓝牙等直接与设备(例如,乘客的移动设备、显示设备、车辆101内的扬声器)通信。用户接口系统113可以是在车辆101内实现的外围设备的部分,包括例如键盘、触摸屏显示设备、麦克风和扬声器等。

[0041] ADV 101的一些或所有功能可以由ADS 110控制或管理,尤其是当以自动驾驶模式操作时。ADS 110包括必要的硬件(例如,处理器、存储器、存储)和软件(例如,操作系统、规划和路由程序),以从传感器系统115、控制系统111、无线通信系统112和/或用户接口系统113接收信息,处理接收的信息,规划从起点到目的地的路线或路径,然后基于规划和控制信息驾驶车辆101。可替换地,ADS 110可以与车辆控制系统111集成。

[0042] 例如,作为乘客的用户可以例如经由用户接口指定行程的起始位置和目的地。ADS 110获得行程相关数据。例如,感知和规划系统110可以从MPOI服务器获得位置和路线信息,MPOI服务器可以是服务器103-104的部分。位置服务器提供位置服务,以及MPOI服务器提供地图服务和某些位置的POI。可替换地,这种位置和MPOI信息可本地缓存在ADS 110的永久存储设备中。

[0043] 当ADV 101沿着路线移动时,ADS 110还可以从交通信息系统或服务器(traffic information system or server,TIS)获得实时交通信息。注意,服务器103-104可以由第三方实体操作。可替换地,服务器103-104的功能可以与ADS 110集成。基于实时交通信息、MPOI信息和位置信息,以及由传感器系统115检测或感测的实时局部环境数据(例如,障碍物、对象、附近车辆),ADS 110可规划最佳路线,并例如经由控制系统111根据规划路线驾驶车辆101,以安全且高效地到达指定目的地。

[0044] 服务器103可以是为各种客户端执行数据分析服务的数据分析系统。在一个实施例中,数据分析系统103包括机器学习引擎122。如下文更详细描述,机器学习引擎122针对各种目的生成或训练一组规则、算法和/或预测模型124,诸如用于运动规划和控制的一个或多个声源(例如,紧急车辆)的检测和定位。然后将算法124上传到ADV,以在自动驾驶期间实时使用。

[0045] 图3A和3B是示出根据一个实施例的与ADV一起使用的自动驾驶系统的示例的框图。系统300可以被实现为图1的ADV101的一部分,包括但不限于ADS110、控制系统111和传感器系统115。参考图3A-3B,ADS 110包括但不限于定位模块301、感知模块302、预测模块303、决策模块304、规划模块305、控制模块306、路由模块307和声源检测和定位308。在一些实施例中,ADS 110还可以包括机器学习引擎122,尽管机器学习引擎122可以安装在远程服务器上,诸如服务器103(如前所述)。

[0046] 模块301-308(和引擎122,在一些实施例中)中的一些或全部可以用软件、硬件或

其组合来实现。例如,这些模块可以安装在永久存储设备352中,加载到存储器351中,并且由一个或多个处理器(未示出)执行。注意,这些模块中的一些或所有模块可以通信地耦接到图2的车辆控制系统111的一些或所有模块或与其集成。模块301-308(和引擎122,在一个实施例中)中的一些模块可以集成在一起作为集成模块。

[0047] 定位模块301确定ADV 300的当前位置(例如,利用GPS单元212)并管理与用户的行程或路线有关的任何数据。定位模块301(也称为地图和路线模块)管理与用户的行程或路线有关的任何数据。用户可以例如经由用户接口登录并指定行程的起始位置和目的地。定位模块301与ADV 300的其他组件交流诸如地图和路线信息311,以获得与行程有关的数据。例如,定位模块301可以从位置服务器和地图与POI(MPOI)服务器获得位置和路线信息。位置服务器提供位置服务,以及MPOI服务器提供地图服务和某些位置的POI,它们可作为地图和路线数据311的部分被缓存。当ADV 300沿着路线移动时,定位模块301还可以从交通信息系统或服务器获得实时交通信息。

[0048] 基于由传感器系统115提供的传感器数据和由定位模块301获得的定位信息,由感知模块302确定对周围环境的感知。感知信息可以表示普通驾驶员将感知到的驾驶员正在驾驶的周围车辆的情况。感知可包括车道配置、交通灯信号、例如以对象形式的另一车辆、行人、建筑物、人行横道或其它交通相关标志(例如,停止标志、让步标志)等的相对位置。车道配置包括描述一个或多个车道的信息,诸如,例如车道的形状(例如,直的或弯曲的)、车道的宽度、道路中的车道数量、单向或双向车道、合并或分离车道、驶出车道等。

[0049] 感知模块302可包括计算机视觉系统或计算机视觉系统的功能,以处理和分析由一个或多个相机捕获的图像,以便识别ADV的环境中的对象和/或特征。对象可包括交通信号、车行道边界、其他车辆、行人和/或障碍物等。计算机视觉系统可使用对象识别算法、视频跟踪和其它计算机视觉技术。在一些实施例中,计算机视觉系统可以映射环境、跟踪对象、以及估计对象的速度等。感知模块302还可基于由诸如雷达和/或LIDAR的其他传感器提供的其他传感器数据检测对象。

[0050] 对于每个对象,预测模块303预测对象将在环境下表现什么。鉴于一组地图/路线信息311和交通规则312,基于感知在时间点的驾驶环境的感知数据执行预测。例如,如果对象是在相反方向上的车辆并且当前驾驶环境包括十字路口,则预测模块303将预测车辆将可能直线向前移动还是转弯。如果感知数据指示十字路口没有交通灯,则预测模块303可以预测车辆在进入十字路口之前可能必须完全停止。如果感知数据指示车辆当前处于仅左转弯车道或仅右转弯车道,则预测模块303可预测车辆将更可能分别进行左转弯或右转弯。

[0051] 对于每个对象,决策模块304作出关于如何处理该对象的决策。例如,对于特定对象(例如,在交叉路线中的另一车辆)以及描述该对象的元数据(例如,速度、方向、转向角),决策模块304决定如何遇到该对象(例如,超车、让行、停止、通过)。决策模块304可根据一组规则,诸如交通规则或驾驶规则312作出这些决策,这组规则可存储于永久存储设备352中。

[0052] 路由模块307被配置为提供从起点到目的地点的一条或多条路线或路径。对于例如从用户接收到的从起始位置到目的地位置的给定行程,路由模块307获得路线和地图信息311并且确定从起始位置用以到达目的地位置的所有可能的路线或路径。路由模块307可以为其确定的从起始位置用以到达目的地位置的每条路线生成地形图形式的参考线。参考线是指没有来自诸如其它车辆、障碍物或交通状况的其它者的任何干扰的理想路线或路

径。也就是说,如果在道路上没有其它车辆、行人或障碍物,则ADV应该精确地或紧密地跟随参考线。然后,地形图被提供给决策模块304和/或规划模块305。决策模块304和/或规划模块305检查所有可能的路线,以根据由其它模块提供的其它数据(诸如来自定位模块301的交通状况、由感知模块302感知的驾驶环境和由预测模块303预测的交通状况)选择和修改最佳路线中的一个。取决于在时间点的特定驾驶环境,用于控制ADV的实际路径或路线可以接近或不同的于由路由模块307提供的参考线。

[0053] 基于对于每个感知的对象的决策,规划模块305使用由路由模块307提供的参考线作为基础规划用于ADV的路径或路线以及驾驶参数(例如,距离、速度和/或转向角)。也就是说,对于给定对象,决策模块304决策对该对象做什么,而规划模块305确定如何做。例如,对于给定对象,决策模块304可以决策经过对象,而规划模块305可以确定是在对象的左侧还是右侧经过。规划和控制数据由规划模块305生成,包括描述车辆300在下一个移动周期(例如,下一个路线/路径段)中将如何移动的信息。例如,规划和控制数据可指示车辆300以30英里每小时(mph)的速度移动10米,然后以25mph的速度改变到右车道。

[0054] 基于规划和控制数据,控制模块306通过根据由规划和控制数据限定的路线或路径向车辆控制系统111发送适当的命令或信号控制和驾驶ADV。规划和控制数据包括足够的信息以沿着路径或路线在不同的时间点使用适当的车辆设置或驾驶参数(例如,油门、制动、转向命令)将车辆从路线或路径的第一点驾驶到第二点。

[0055] 在一个实施例中,规划阶段在多个规划周期(也称为驾驶周期,诸如在100毫秒(ms)的每个时间间隔内)内执行。对于每个规划周期或驾驶周期,将基于规划和控制数据发出一个或多个控制命令。也就是说,对于每100ms,规划模块305规划下一个路线段或路径段,例如包括目标位置和ADV到达目标位置所需的时间。可替换地,规划模块305还可指定具体速度、方向和/或转向角等。在一个实施例中,规划模块305为下一预定的时间段,诸如5秒规划路线段或路径段。对于每个规划周期,规划模块305基于在先前周期中规划的目标位置规划用于当前周期(例如,下一个5秒)的目标位置。然后控制模块306基于当前周期的规划和控制数据生成一个或多个控制命令(例如,油门、制动、转向控制命令)。

[0056] 注意,决策模块304和规划模块305可以集成为集成模块。决策模块304/规划模块305可包括导航系统或导航系统的功能,以确定用于ADV的驾驶路径。例如,导航系统可以确定一系列速度和方向航向,以影响ADV沿着基本上避开感知的障碍物的路径的移动,同时大体使ADV沿着通向最终目的地的基于车行道的路径前进。目的地可以根据经由用户接口系统113的用户输入而被设置。当ADV在操作时,导航系统可以动态地更新驾驶路径。导航系统可以并入来自GPS系统的数据和一个或多个地图,以便确定用于ADV的驾驶路径。

[0057] 继续参考图3A以及图4A-4B(其是示出根据一个实施例的用于声源检测和定位的系统的框图),声源检测和定位模块308可以接收由传感器系统115的音频传感器411(例如,麦克风或麦克风阵列)提供的音频数据313。音频数据313可以包括记录或捕获的当ADV沿着路线操作时在不同时间点由一个或多个声源(或对象)发出的声音。如图3A中所示,音频数据313可以存储在永久存储设备352中。在一个实施例中,声源(或对象)可以是紧急车辆,诸如救护车、警车、消防车等。

[0058] 基于接收或输入的音频数据313,声源检测和定位模块308可确定并产生各种声源信息以报告在ADV的驾驶环境内是否存在特定声音(例如,警报声)。例如,模块308可以确定

特定声音的存在的置信度分数(或可能性)。在一个实施例中,可以基于捕获的声音的多普勒效应来确定置信度分数。每个置信度分数可以在特定范围内(例如,值0~1),其中0表示最低置信度或可能性,1表示最高置信度或可能性。需要注意的是,虽然在此示例中使用了0~1的范围,但这不是限制性的,可以使用任何数字范围来表示置信度分数。

[0059] 利用音频传感器411使用或执行声源定位,模块308可以确定每个声源相对于对应的音频传感器411(即,捕获由声源产生的声音的音频传感器411)的方向,声源和相应音频传感器之间的距离,以及基于安装在ADV上的每个音频传感器411的位置的每个可能检测到的声音(或声源)的相对位置。声源定位和检测模块308能够确定(或计算)声音(或声源)的精确或绝对位置、声音的接近/离开状态、和/或当前时间戳的声音强度。这样的声源定位方法或技术对于本领域技术人员来说是众所周知的,因此为简洁起见,在此不再赘述。在一个实施例中,模块308可以产生声源信息(例如,置信度分数、声源方向、声源和音频传感器之间的距离、每个可能检测到的声音的相对位置、声音的精确位置、声音的接近/离开状态、声音的强度等)作为数据表示314(例如,网格)的一部分。如图3A中所示,数据表示314可以存储在永久存储设备352中。

[0060] 现在转向图4B,数据表示314可以包括覆盖驾驶环境的多个区域(表示为框)。数据表示314用于报告在数据表示内是否存在特定声音(例如,警报声)。在图4B所示的示例中,数据表示314具有固定大小的M个区域,其中M是大于0的整数。每个区域(或区域的子集)可以对应于特定的音频传感器411以部分地覆盖ADV的驾驶环境。数据表示314可以被可视化为自上而下的视图,其中数据表示314的中心表示ADV(例如,图1的ADV101)的位置。在一些实施例中,数据表示314的每个区域被配置为覆盖驾驶环境中的特定尺寸,诸如5x5平方米(m²)或10x10m²,但这不是限制性的。在一个实施例中,数据表示314的每个区域314报告结果向量421以指示特定声音是否存在于此区域中。每个向量421可以包括但不限于区域标识符(ID)和声源信息(如前所述)。

[0061] 图5是示出根据一个实施例的使用用于声源检测和定位的系统的示例驾驶场景的图。参考图5,ADV 101可以包括安装在ADV周围(例如左侧、右侧、前部和后部)的多个音频传感器411(例如,麦克风阵列),但是音频传感器的数量没有限制。当ADV 101沿着路线行驶时,一个或多个音频传感器411可以分别检测从声源509-510(例如,警车、救护车、消防车等)产生的多种声音。作为示例,假设后部音频传感器411完全检测到从声源509产生的声音,而剩余音频传感器中的每一个仅部分地检测到或未能检测到声音。在此场景下,与后部音频传感器411相关联的数据表示区域的置信度分数将更接近于值1,而与剩余传感器相关联的置信度分数将更接近于值0。此外,使用声源定位(如前所述),可以确定声源509相对于后部音频传感器411的方向,声源509和后部音频传感器411之间的距离,以及由声源509产生的声音的相对位置。还可以确定声源509产生的声音的精确或绝对位置、声音的接近/离开状态、和/或当前时间戳处的声音强度。

[0062] 同时,假设右侧(或乘客侧)音频传感器411完全检测到从声源510产生的声音,而剩余音频传感器中的每一个仅部分检测到或未能检测到声音。在此场景下,与右侧音频传感器411相关联的数据表示区域的置信度分数将更接近于值1,而与剩余传感器相关联的置信度分数将更接近于值0。此外,使用声源定位(如前所述),可以确定声源510相对于右侧音频传感器411的方向,声源510和右侧音频传感器411之间的距离,以及声源510产生的声音

的相对位置。还可以确定声源510产生的声音的精确或绝对位置、声音的接近/离开状态、和/或当前时间戳处的声音强度。

[0063] 图6A-6B是图示根据一个实施例的训练机器学习算法/模型的系统的框图。参考图6A,输入音频数据313和输出数据表示314(由声源检测和定位模块308产生)可以提供给机器学习122以生成或训练一组规则、算法和/或预测模型124,用于针对运动规划和控制的一个或多个声源(例如紧急车辆)的检测和定位。算法/模型124(例如,深度神经网络、卷积深度神经网络、深度信念网络和/或循环神经网络)可以概括并找到输入音频数据313和音频标签(例如,指示特定声音诸如警报声的标签)之间的关系以生成输出610。在一个实施例中,算法/模型124可以存储在永久存储设备352或远程服务器(例如,服务器103)上。

[0064] 参考图6B,输出610可以被表示为数据表示(例如,网格),其中置信分数和/或声源信息(未示出)填充在数据表示的某些区域中。在图6B的示例中,特定声音(例如警报声)可能存在于置信度分数为0.7和0.9(可能是高斯分布)的区域中,ADV的前部朝北朝向虚线箭头。因此,可以确定警报声一个来自ADV的左侧以及另一个来自ADV的右后方。

[0065] 在一个实施例中,算法/模型124然后可以被上传到ADV上以在自动驾驶过程中实时使用。即,模型124产生的输出610可以被ADS 110用来生成通信信号和/或命令以在检测到特定声音时相应地操作ADV。例如,当检测到一个或多个警报器声音时,ADS 110可以基于来自数据表示(输出610)中每个区域的声源信息来控制ADV停止或靠边停在ADV的驾驶环境内的安全位置。

[0066] 图7是根据一个实施例的利用ADV的用于声源检测和定位的示例方法的流程图。方法或过程700可由处理逻辑执行,处理逻辑可包括软件、硬件或其组合。例如,过程700可由图1的ADS 110执行。

[0067] 参考图7,在块701,处理逻辑从安装在ADV上的多个音频传感器(例如,麦克风电阵列)接收音频数据。音频数据包括由音频传感器捕获并由一个或多个声源(例如,紧急车辆)发出的声音。在块702,基于接收到的音频数据,处理逻辑确定多个声源信息。每个声源信息包括与特定声音(例如,警报声)的存在相关联的置信度分数。在块703,处理逻辑生成数据表示(例如,网格)以报告在ADV的驾驶环境内是否存在特定声音。数据表示包括确定的声源信息。接收到的音频数据和生成的数据表示用于随后训练机器学习算法,以在ADV的自动驾驶过程中实时识别特定的声源。

[0068] 注意,如上所示和所述的一些或所有组件可以用软件、硬件或其组合实现。例如,这些组件可以被实现为安装并存储在永久存储设备中的软件,软件可以由处理器(未示出)加载并执行在存储器中以执行贯穿本申请所述的过程或操作。可替换地,这些组件可被实现为被编程或嵌入到专用硬件中的可执行代码,专用硬件诸如集成电路(例如,专用IC或ASIC)、数字信号处理器(DSP)或现场可编程门阵列(FPGA),可执行代码可经由来自应用的对应驱动器和/或操作系统访问。此外,这些组件可以被实现为处理器或处理器内核中的特定硬件逻辑,作为经由一个或多个特定指令软件组件可访问的指令集的一部分。

[0069] 已经关于对计算机存储器内的数据位的操作的算法和符号表示呈现了前述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技术人员用来最有效地将他们的工作实质传达给本领域的其他技术人员的方式。算法在这里并且通常被认为是导致期望结果的自洽操作序列。这些操作是需要对物理量进行物理操纵的那些操作。

[0070] 然而,应当记住的是,所有这些和类似的术语都与适当的物理量相关联,并且仅仅是应用于这些量的方便的标记。除非特别声明,否则从以上讨论中显而易见的是,应当理解的是,在整个说明书中,使用诸如所附权利要求书中所阐述的术语的讨论指的是计算机系统或类似电子计算设备的动作和过程,所述计算机系统或类似电子计算设备将计算机系统的寄存器和存储器内的表示为物理(电子)量的数据操纵和变换成计算机系统存储器或寄存器或其它这样的信息存储、传输或显示设备内的类似地表示为物理量的其它数据。

[0071] 本公开的实施例还涉及用于执行本文的操作的装置。这种计算机程序存储在非瞬态计算机可读介质中。机器可读介质包括用于以机器(例如,计算机)可读的形式存储信息的任何机制。例如,机器可读(例如,计算机可读)介质包括机器(例如,计算机)可读存储介质(例如,只读存储器(“ROM”)、随机存取存储器(“RAM”)、磁盘存储介质、光存储介质、闪存设备)。

[0072] 在前述附图中描述的过程或方法可以由包括硬件(例如,电路、专用逻辑等)、软件(例如,体现在非暂时性计算机可读介质上)或两者的组合的处理逻辑来执行。尽管以上根据一些顺序操作描述了过程或方法,但是应当理解的是,可以以不同的顺序执行所述的一些操作。此外,一些操作可以并行地而不是顺序地执行。

[0073] 本公开的实施例不是参考任何特别编程语言描述的。将了解,可使用各种编程语言实施如本文所述的本公开的实施例的教导。

[0074] 在前述说明书中,已经参考本公开的具体示例性实施例描述了本公开的实施例。显然,在不背离如所附权利要求书中所阐述的本公开的更宽的精神和范围的情况下,可以对其进行各种修改。因此,说明书和附图应被认为是说明性的而不是限制性的。

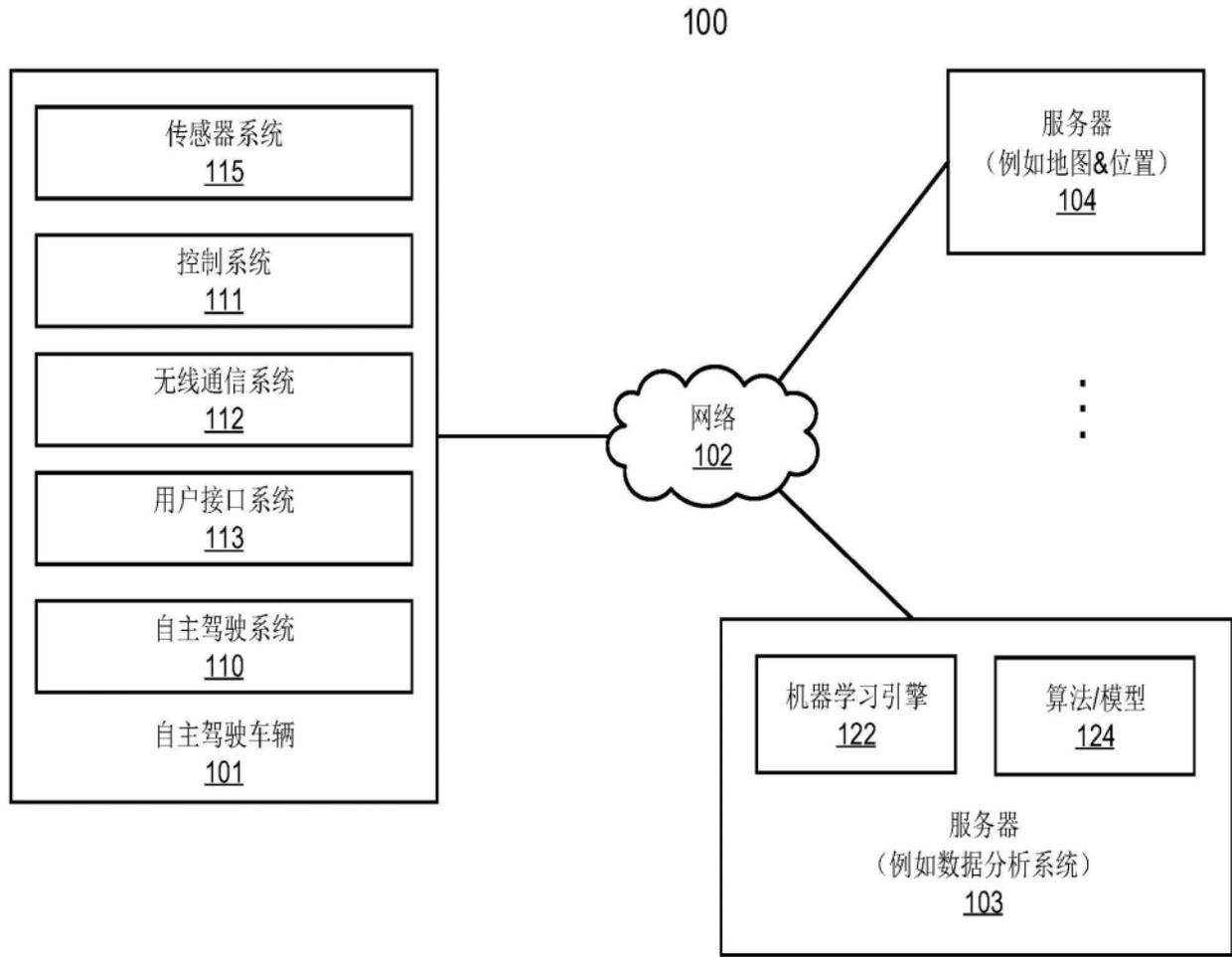


图1

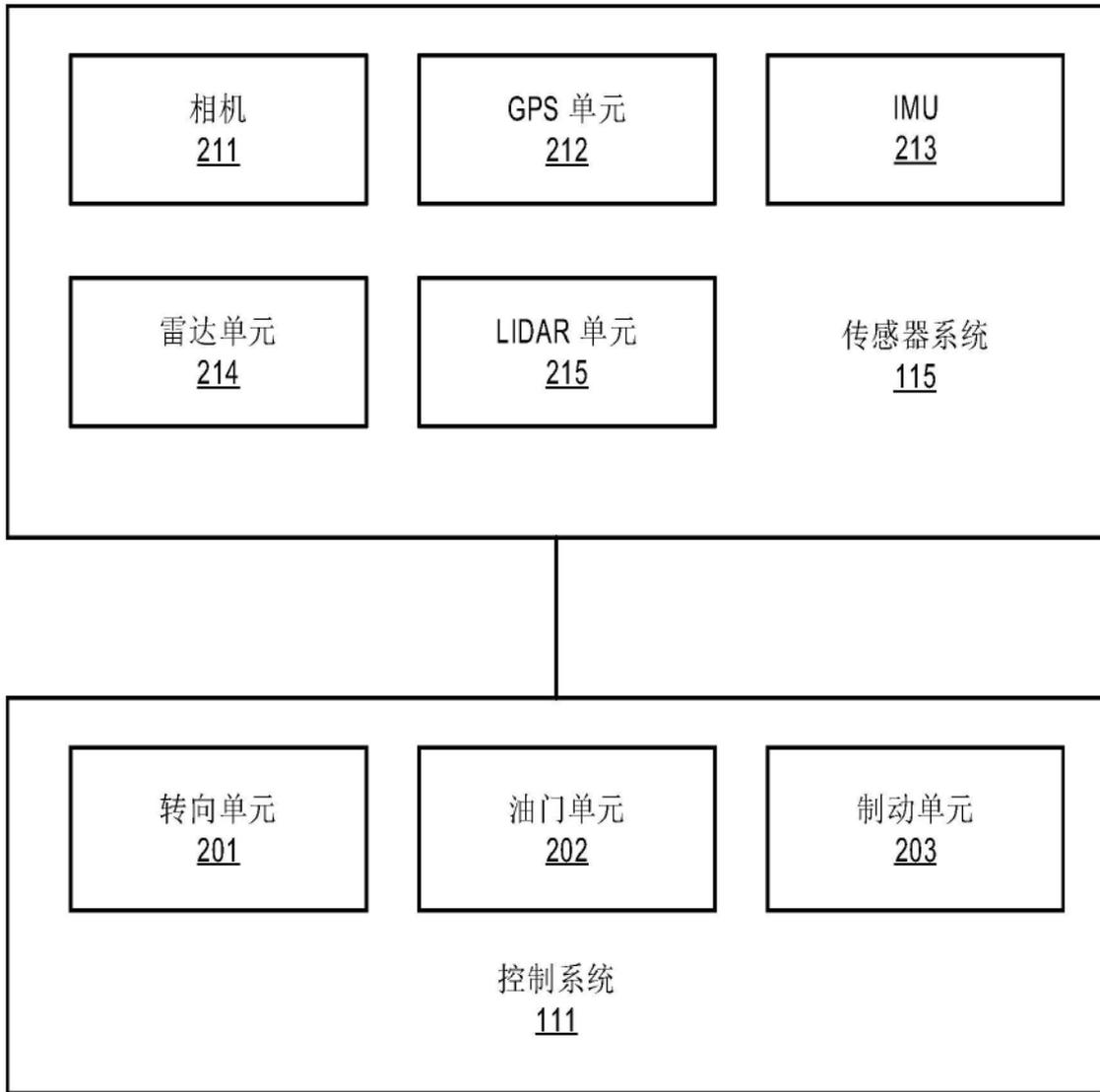


图2

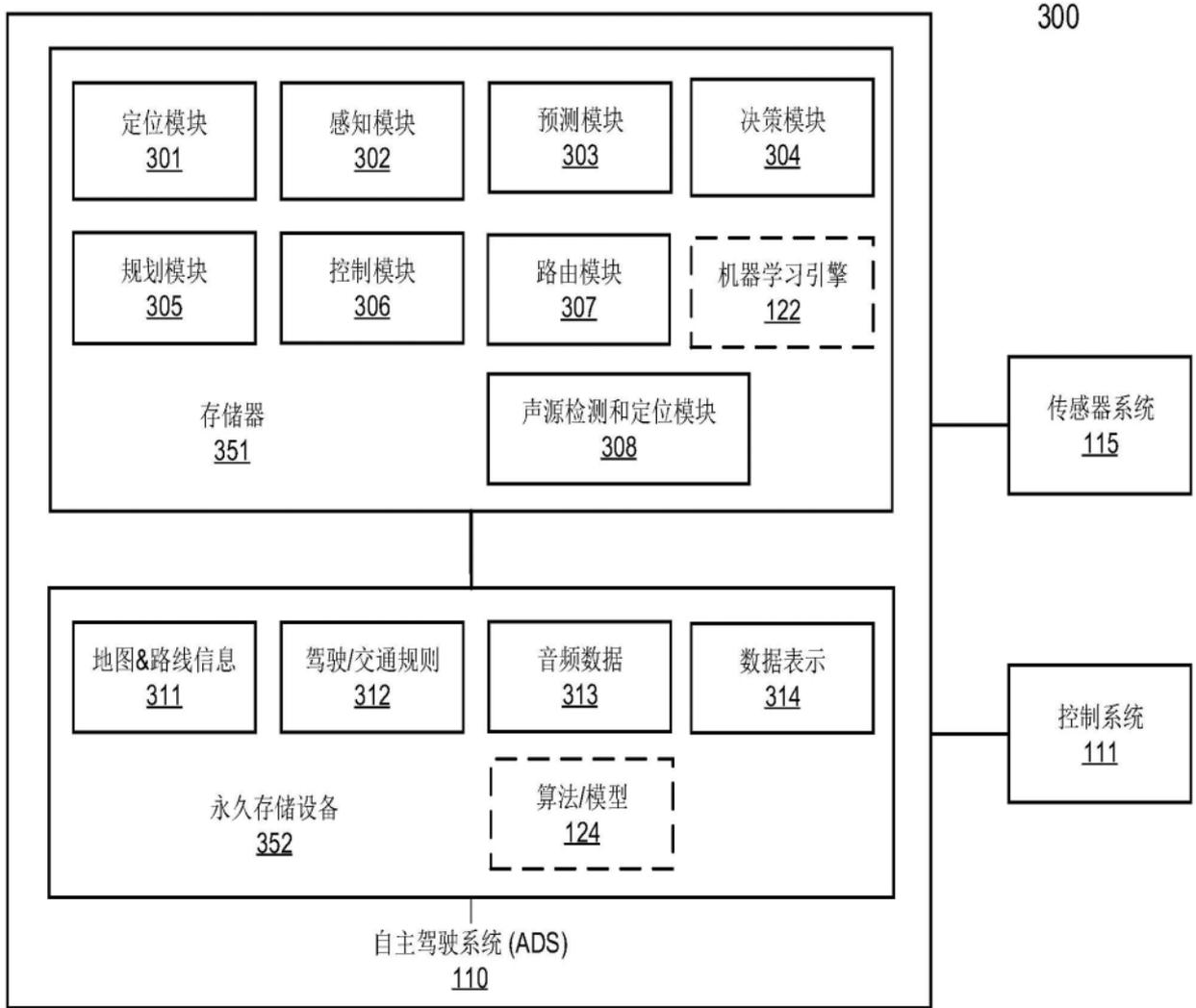


图3A

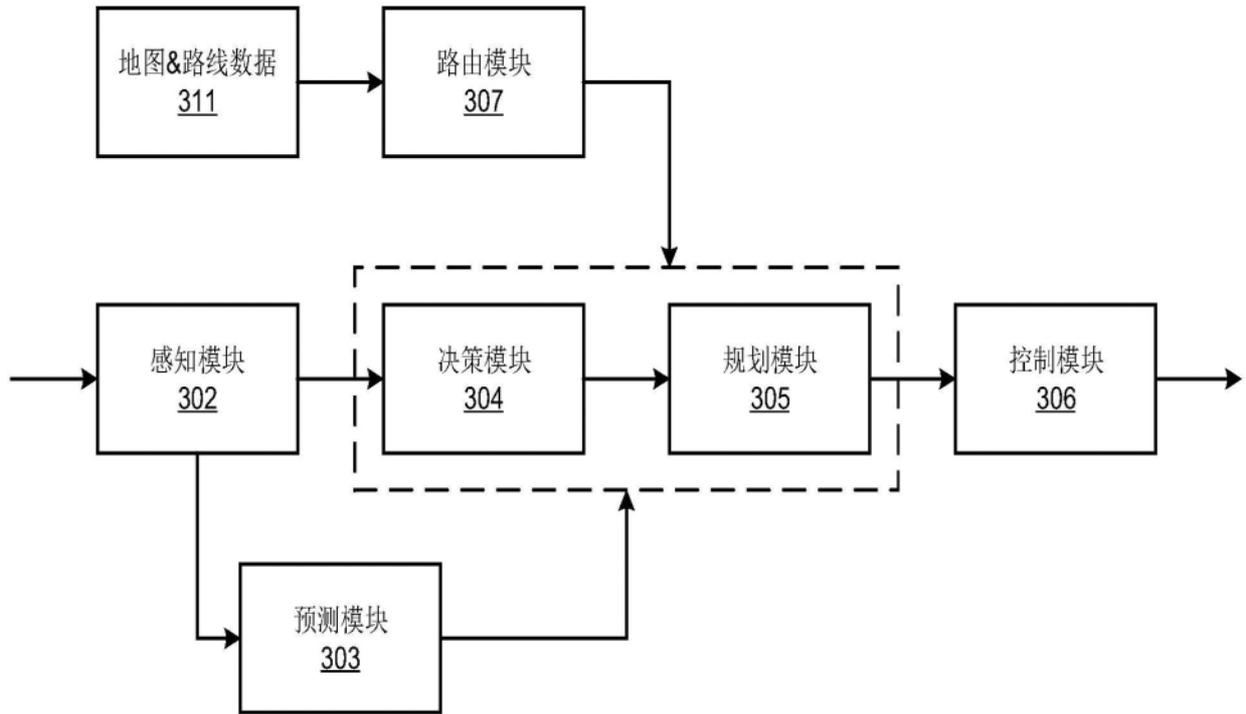


图3B

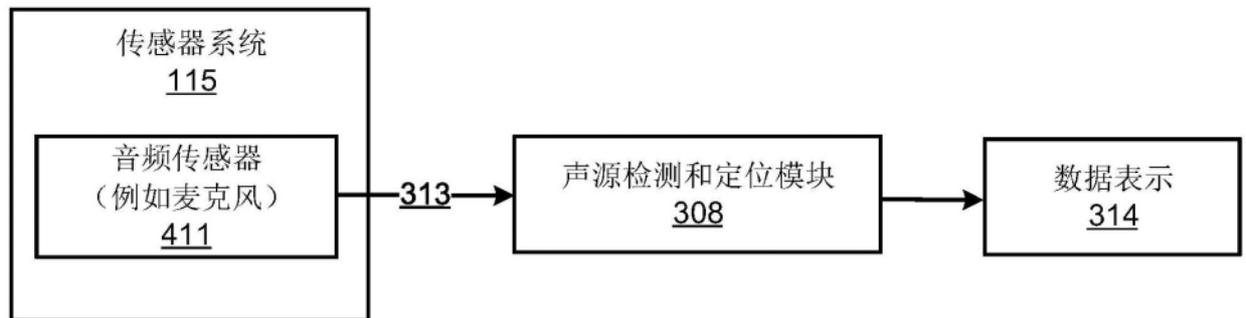


图4A

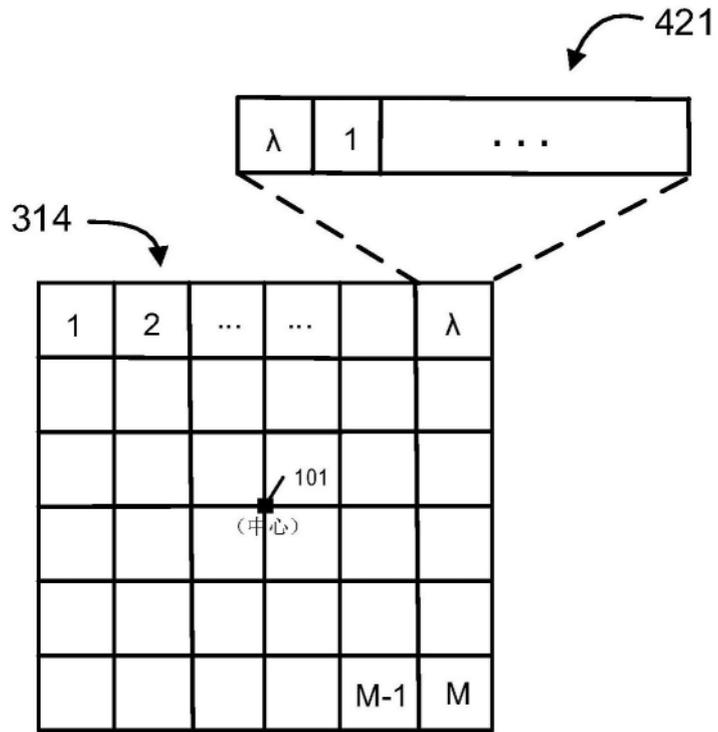


图4B

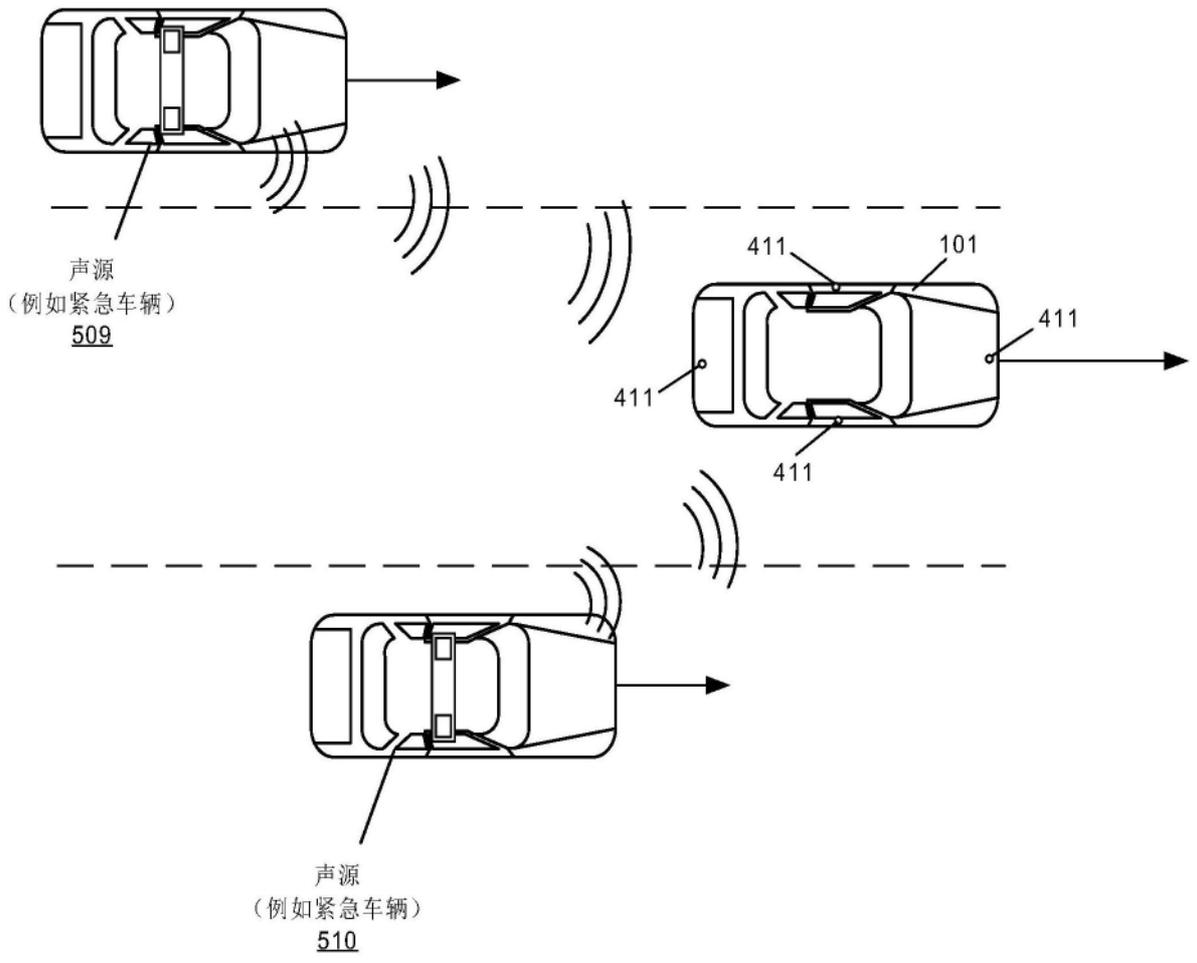


图5

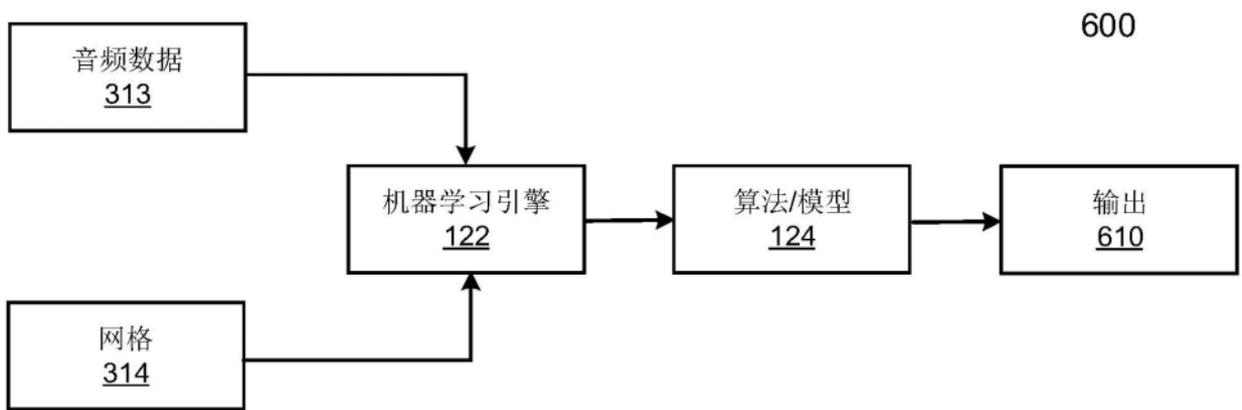


图6A

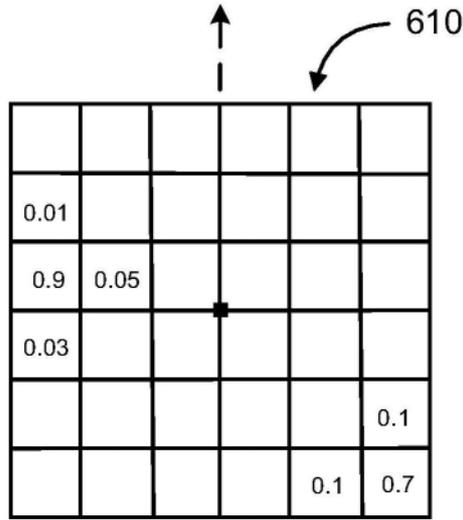


图6B

700

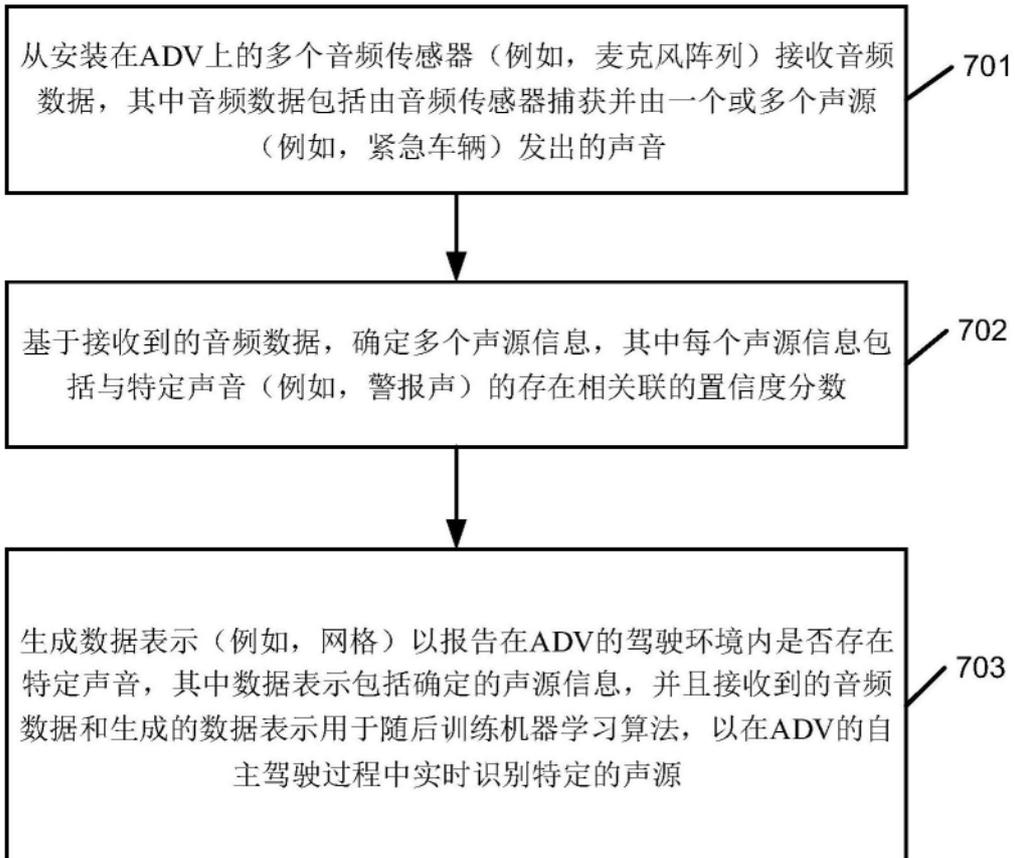


图7