

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 10 月 24 日 (24.10.2019)



(10) 国际公布号
WO 2019/201304 A1

- (51) 国际专利分类号:
G10L 15/00 (2013.01) **G06K 9/00** (2006.01)
- (21) 国际申请号: PCT/CN2019/083261
- (22) 国际申请日: 2019 年 4 月 18 日 (18.04.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201810360826.7 2018 年 4 月 20 日 (20.04.2018) CN
- (71) 申请人: 比亚迪股份有限公司 (**BYD COMPANY LIMITED**) [CN/CN]; 中国广东省深圳市坪山新区比亚迪路 3009 号, Guangdong 518118 (CN)。
- (72) 发明人: 陈少贵 (**CHEN, Shaogui**); 中国广东省深圳市坪山新区比亚迪路 3009 号, Guangdong 518118 (CN)。王书克 (**WANG, Shuke**); 中国广东省深圳市坪山新区比亚迪路 3009 号,

Guangdong 518118 (CN)。武伦佳 (**WU, Lunjia**); 中国广东省深圳市坪山新区比亚迪路 3009 号, Guangdong 518118 (CN)。

(74) 代理人: 北京清亦华知识产权代理事务所 (普通合伙) (**TSINGYIHUA INTELLECTUAL PROPERTY LLC**); 中国北京市海淀区清华园清华大学照澜院商业楼 301 室, Beijing 100084 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) **Title:** FACE RECOGNITION-BASED VOICE PROCESSING METHOD, AND DEVICE

(54) 发明名称: 基于人脸识别的语音处理方法及其设备

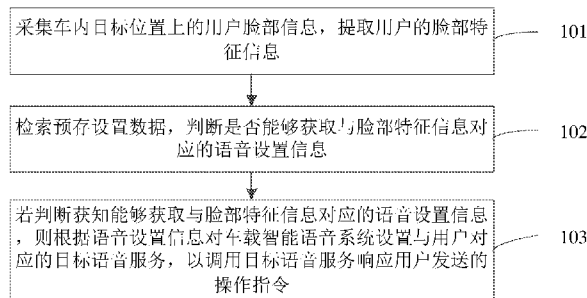


图 1

- 101 ACQUIRE USER FACE INFORMATION AT A TARGET POSITION IN A VEHICLE, AND EXTRACT FACIAL FEATURE INFORMATION OF THE USER
- 102 RETRIEVE PRE-STORED CONFIGURATION DATA, AND DETERMINE WHETHER VOICE CONFIGURATION INFORMATION CORRESPONDING TO THE FACIAL FEATURE INFORMATION CAN BE OBTAINED
- 103 IF IT IS DETERMINED THAT VOICE CONFIGURATION INFORMATION CORRESPONDING TO THE FACIAL FEATURE INFORMATION CAN BE OBTAINED, CONFIGURE A TARGET VOICE SERVICE CORRESPONDING TO THE USER ON A VEHICLE-MOUNTED INTELLIGENT VOICE SYSTEM ACCORDING TO THE VOICE CONFIGURATION INFORMATION, TO INVOKE THE TARGET VOICE SERVICE TO RESPOND TO AN OPERATION INSTRUCTION SENT BY THE USER

(57) **Abstract:** A face recognition-based voice processing method, and a device. The method comprises: acquiring user face information at a target position in a vehicle, and extracting facial feature information of the user (101); retrieving pre-stored configuration data, and determining whether voice configuration information corresponding to the facial feature information can be obtained (102); and if it is determined that voice configuration information corresponding to the facial feature information can be obtained, configuring a target voice service corresponding to the user on a vehicle-mounted intelligent voice system according to the voice configuration information, to invoke the target voice service to respond to an operation instruction sent by the user (103). Thus, different voice services are provided according to different users, thereby meeting personalized requirements of the users and improving the user experience in intelligence.

(57) **摘要:** 一种基于人脸识别的语音处理方法及其设备, 其中, 方法包括: 采集车内目标位置上的用户脸部信息, 提取用户的脸部特征信息 (101); 检索预存设置数据, 判断是否能够获取与脸部特征信息对应的语音设置信息 (102); 若判断获知能够获取与脸部特征信息对应的语音设置信息, 则根据语音设置信息对车载智能语音系统设置与用户对应的目标语音服务, 以调用目标语音服务响应用户发送的操作指令 (103)。由此, 基于不同的用户进行不同的语音服务, 满足了用户的个性化要求, 提高了用户的智能化体验。

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

基于人脸识别的语音处理方法及其设备

5 相关申请的交叉引用

本公开要求比亚迪股份有限公司于 2018 年 04 月 20 日提交的、公开名称为“基于人脸识别的语音处理方法及其设备”的、中国专利申请号“201810360826.7”的优先权。

技术领域

10 本公开涉及智能交互技术领域，尤其涉及一种基于人脸识别的语音处理方法及其设备。

背景技术

随着汽车消费的逐渐普及，以车代步成为了一种生活，越来越多的信息娱乐功能加入到车辆中，这大大地增加了司机及乘员的舒适性，车载智能语音识别便是其中的一个典型。
15 车载智能语音的应用解决了驾驶员操控车辆中控多媒体易分心导致事故的问题，大大地解放了驾驶员的双手。目前市场现有的车载智能语音功能丰富多样，能够导航设置、空调控制、音乐播放、蓝牙电话、车窗座椅控制等。对于智能语音识别功能本身，可以进行按键语音唤醒、语音唤醒设置、语音播报声音及方言设置等。

相关技术中，基于功能要求为用户提供智能化的车载智能语音服务，然而，这种仅仅
20 满足用户功能要求的服务方式，限制了用户与产品的距离，不利于提高用户和产品的粘性。

公开内容

本公开旨在至少在一定程度上解决相关技术中的技术问题之一。

为此，本公开的第一个目的在于提出一种基于人脸识别的语音处理方法，该方法基于
25 不同的用户进行不同的语音服务，满足了用户的个性化要求，提高了用户的智能化体验。
本公开的第二个目的在于提出一种车载智能语音系统。

本公开的第三个目的在于提出一种车机设备。

为达上述目的，本公开第一方面实施例提出了一种基于人脸识别的语音处理方法，包括以下步骤：采集车内目标位置上的用户脸部信息，提取所述用户的脸部特征信息；检索
30 预存设置数据，判断是否能够获取与所述脸部特征信息对应的语音设置信息；若判断获知能够获取与所述脸部特征信息对应的语音设置信息，则根据所述语音设置信息对车载智能语音系统设置与所述用户对应的目标语音服务，以调用所述目标语音服务响应所述用户发

送的操作指令。

为达上述目的，本公开第二方面实施例提出了一种车载智能语音系统，包括：提取模块，用于采集车内目标位置上的用户脸部信息，提取所述用户的脸部特征信息；判断模块，用于检索预存设置数据，判断是否能够获取与所述脸部特征信息对应的语音设置信息；处理模块，用于在判断获知能够获取与所述脸部特征信息对应的语音设置信息时，根据所述语音设置信息对车载智能语音系统设置与所述用户对应的目标语音服务，以调用所述目标语音服务响应所述用户发送的操作指令。

为达上述目的，本公开第三方面实施例提出了一种车机设备，包括：车载摄像头和中控多媒体设备，其中，所述车载摄像头与所述中控多媒体设备连接，其中，所述车载摄像头，用于采集车内目标位置上的用户脸部信息，并将所述用户脸部信息发送至所述中控多媒体设备；所述中控多媒体设备，用于根据提取所述用户的脸部特征信息，检索预存设置数据，判断是否能够获取与所述脸部特征信息对应的语音设置信息，若判断获知能够获取与所述脸部特征信息对应的语音设置信息，则将所述语音设置信息设置与所述用户对应的目标语音服务，以调用所述目标语音服务响应所述用户发送的操作指令。

本公开实施例提供的技术方案可以包括以下有益效果：

采集车内目标位置上的用户脸部信息，提取用户的脸部特征信息，检索预存设置数据，判断是否能够获取与脸部特征信息对应的语音设置信息，若判断获知能够获取与脸部特征信息对应的语音设置信息，则根据语音设置信息对车载智能语音系统设置与用户对应的目标语音服务，以调用目标语音服务响应用户发送的操作指令。由此，基于不同的用户进行不同的语音服务，满足了用户的个性化要求，提高了用户的智能化体验。

本公开附加的方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本公开的实践了解到。

附图说明

本公开上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

图 1 是根据本公开一个实施例的基于人脸识别的语音处理方法的流程图；

图 2 是根据本公开一个实施例的设置数据的存储示意图；

图 3 是根据本公开一个实施例的车载智能语音系统的结构示意图；

图 4 是根据本公开另一个实施例的车载智能语音系统的结构示意图；

图 5 是根据本公开另一个实施例的基于人脸识别的语音处理方法的流程图；

图 6 是根据本公开又一个实施例的基于人脸识别的语音处理方法的流程图；

图 7 是根据本公开一个实施例的车载智能语音系统的结构示意图。

具体实施方式

下面详细描述本公开的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本公开，而不能理解为对本公开的限制。

为了解决现有技术中，车载端的语音服务仅仅着眼于功能服务本身实施对应功能，导致用户和产品具有距离感的技术问题，本公开实施例提供了一种基于人脸识别的语音处理方法，可以通过人脸识别技术准确识别用户，针对不同的用户提供不同的语音服务，满足了用户的个性化需求，提高了用户和产品的粘性。

图 1 是根据本公开一个实施例的基于人脸识别的语音处理方法的流程图，如图 1 所示，该方法包括：

步骤 101，采集车内目标位置上的用户脸部信息，提取用户的脸部特征信息。

其中，目标位置根据应用场景的不同，可以对应于车辆内的不同位置，在一些可能的实施例中，为了优先给驾驶员提供较优的语音服务，目标位置为车辆的驾驶座位置；在一些可能的实施例中，为了避免分散驾驶员的注意力，由副驾驶进行语音服务的控制，因而，目标位置为车辆的副驾驶位置。

另外，用户的面部特征信息包括但不限于用户的五官特征信息、面部轮廓信息等。

需要说明的是，在不同的应用场景下，采集车内目标位置上用户脸部信息的方式不同，示例说明如下：

第一种示例：

在本示例中，按照预设周期采集车内目标位置上的用户脸部信息。

其中，预设周期可以是根据系统处理速度确定的，系统处理速度越高，对应的预设周期对应的时长越短。

在实际执行过程中，可以调用安装在车内前端的摄像头采集目标位置上的用户脸部信息，其中，安装在车内前端的摄像头可以根据目标位置上用户脸部的位置调整角度和高度，以保证采集的用户脸部信息尽量完整，提高识别的精确度。

第二种示例：

在本示例中，检测是否满足预设的采集触发条件，若检测满足预设的采集触发条件，则采集车内目标位置上的用户脸部信息。

其中，预设的采集触发条件包括但不限于车辆由停车状态到启动状态等驾驶员有可能发生变化的条件，当检测到满足预设的采集触发条件后，才进行用户脸部信息的采集，提

高了采集的有效性，避免了重复采集一个用户面部信息，减轻了系统处理压力。

在实际执行过程中，可能由于光线不足等原因导致无法直接获取到有效的用户的脸部信息，因而，在本公开的一个实施例中，在采集车内目标位置上的用户脸部信息之前，还可以检测车内环境的光线亮度，比如，通过光线传感器检测车内环境的光线亮度，将光线亮度与预设阈值进行比较，其中，预设阈值与拍照装置有关，对应于拍照装置可以采集到脸部信息的最小光线亮度值。

进而，如果比较获知光线亮度小于预设阈值，则表明当前环境光线不足，可能难以采集到清晰的用户的脸部信息，从而，对目标位置启动辅助照明设备，比如，打开车辆内部的照明系统等。

步骤 102，检索预存设置数据，判断是否能够获取与脸部特征信息对应的语音设置信息。

其中，语音设备信息包括播报音量、播报声音、播报方言中的一种或者多种信息组合。

可以理解，为了拉近产品和用户的距离，在本公开的实施例中，在用户启动语音服务时，针对不同的用户提供不同的语音服务，比如，如图 2 所示，针对不同的用户对应存储了不同的语音设置信息，在本实施例中，预先设置并存储设置数据，该设置数据中存储有脸部特征信息和语音设置信息的对应关系，从而，在获取到用户的脸部特征信息后，查询该对应关系以匹配出对应的语音设置信息。

需要说明的是，根据应用场景的不同，设置预存设置数据的方式不同，作为一种可能的实现方式，为用户提供语音设置信息界面，接收用户在该界面中根据个人喜好设定的语音设置信息，其中，在用户输入语音设置信息时，可以同步采集用户的面部图像以提取用户的脸部特征信息，或者，由用户上传照片以便于根据用户上传的照片提取脸部特征信息，进而，将脸部特征信息与对应的语音设置信息对应存储。

步骤 103，若判断获知能够获取与脸部特征信息对应的语音设置信息，则根据语音设置信息对车载智能语音系统设置与用户对应的目标语音服务，以调用目标语音服务响应用户发送的操作指令。

其中，用户发送的操作指令的方式可以为文字输入形式、语音输入形式等任意可以被识别的形式。

在本公开的实施例中，如果判断获知能够获取到与脸部特征信息对应的语音设置信息，则根据语音设置信息对车载智能语音系统设置与用户对应的目标语音服务，以调用目标语音服务响应用户发送的操作指令，以实现为用户提供个性化的语音服务。

在本实施例中，可以通过车载智能语音系统的屏幕获取用户通过操作面板输入的操作指令，比如，获取用户对预设的选项卡的触控操作，或者，文字输入操作等，或者，通过

车载智能语音系统的麦克风获取用户语音输入的操作指令。

进而，对用户语音输入的操作指令进行语音识别，调用本地的语音合成指令或者获取云端反馈的响应信息，也就是说，对用户输入的操作指令可以采用本地的语音识别并进行语音合成与该操作指令对应的语音合成指令，该语音合成指令为针对操作指令的响应信息，或者，可以将语音识别后的语音信息发送至云端，以供云端根据预先存储的语音信息和响

5 应消息的对应关系，匹配并反馈对应的响应消息，通过车载智能语音系统输出给用户，以基于对应的目标语音服务为用户提供该响应消息对应的语音服务。

举例而言，在当前用户为用户 1 时，如果获取到与用户 1 的脸部特征对应的语音设置信息为“粤语、音量 5”，则根据语音设置信息对车载智能语音系统设置与用户对应的目标语音服务为“粤语、音量 5”，则当用户 1 发出的操作指令为“帮我导航去##”，则将该指令发送

10 至云端后获取到云端反馈的对应的导航信息，车载智能语音服务以“粤语、音量 5”根据导航信息为用户进行语音导航。

在本公开的一个实施例中，当判断获知不存在与脸部特征信息对应的语音设置信息，则调用车载智能语音系统预先根据默认设置信息设置的标准语音服务响应用户发送的操作

15 指令，以满足用户的语音服务需求。

当然，为了进一步提高基于人脸识别的语音处理方法的灵活性，也可以根据应用需求随时更改或者添加与语音设置信息。

在本公开的一个实施例中，接收用户输入的包含语音设置信息的设置指令，比如，接收用户输入的“我要制定属于我的语音服务”，则查询预存设置数据，判断是否包含用户的脸部特征信息，若判断获知包含用户的脸部特征信息，则删除当前存储的与脸部特征信息

20 对应的语音设置信息，并存储语音设置信息与脸部特征信息的对应关系，若判断获知不包含用户的脸部特征信息，则在设置数据中存储语音设置信息与用户的脸部特征信息的对应关系。由此，通过对预存设置数据的更新和修改，一方面，适应新用户的个性化语音服务的提供，另一方面，适应已经注册的用户的喜好的变化。

基于以上描述，在实际执行过程中，实现基于人脸识别的语音处理方法的模块可以根据应用场景的不同，以不同的形式实现，下面示例说明：

25

在本示例中，由包括车机设备 100 和云端 200 组成对应的车载智能语音系统，其中，车机设备 100 与云端 200 通过无线网络实现通讯，将车机设备 100 采集的相关信息上传至云端 200 进行识别，并将云端 200 识别出的语音内容下下发至车机设备 100，当然，根据

30 应用需要，也可以下发至第三方设备，其中无线网络可以是 2G、3G、4G 或者未来更新的移动通讯技术或者车机连接 WIFI，通讯链路可以是公网也可以是专网。

图 3 是根据本公开一个实施例的车载智能语音系统的结构示意图，如图 3 所示，车机

设备 100 包括车载摄像头 110 和中控多媒体设备 120。

其中，车载摄像头 110，用于采集车内目标位置上的用户脸部信息，并将用户脸部信息发送至中控多媒体设备。

其中，车载摄像头 110 作为车机设备 100 的用户图像采集设备，用于采集用户的脸部信息传输至中控多媒体设备 120，摄像头安装于车内，要求可清晰地拍摄到用户脸部。作为一种可能的实现方式，摄像头可采用高分辨率红外摄像头，车内安装有红外灯，以便在晚上或者光照条件差的情况下仍可以清晰地采集用户脸部信息。

中控多媒体设备 120，用于根据提取用户的脸部特征信息，检索预存设置数据，判断是否能够获取与脸部特征信息对应的语音设置信息，若判断获知能够获取与脸部特征信息对应的语音设置信息，则将语音设置信息设置与用户对应的目标语音服务，以调用目标语音服务响应用户发送的操作指令。

在本公开的实施例中，中控多媒体设备 120 接收车载摄像头 110 采集的用户的脸部信息，经过图像处理，提取用户面部特征信息，对比本地存储的特征数据识别不同用户。

作为一种可能的实现方式，中控多媒体设备 120 带有显示屏，提供用户交互界面，用户交互界面用于显示语音服务的相关文字信息，或者，故障提示信息等。

在本公开的一个实施例中，参照图 4，车机设备 100 还可包括麦克风 130，其中，麦克风 130 与中控多媒体设备 120 连接，其中，

麦克风 130，用于接收用户输入的包含语音设置信息的设置指令。

其中，麦克风 130 作为车机设备 100 的语音接收设备，用于接收用户输入的语音设置信息等，将语音设置信息的音频信号转换成电平信号，为语音识别提供信号源，麦克风 130 带有滤波功能，以消除环境噪声。

中控多媒体设备 120，用于查询预存设置数据，判断是否包含用户的脸部特征信息，若判断获知包含用户的脸部特征信息，则删除当前存储的与脸部特征信息对应的语音设置信息，并存储语音设置信息与脸部特征信息的对应关系，若判断获知不包含用户的脸部特征信息，则在设置数据中存储语音设置信息与用户的脸部特征信息的对应关系。

在本公开的一个实施例中，继续参照图 4，车机设备 100 还可包括扬声器 140，其中，扬声器 140 与中控多媒体设备 120 连接，其中，

扬声器 140 用于根据语音设置信息对用户语音输入的操作指令进行语音响应。

扬声器 140 作为车机设备 100 的发声设备，用于发出中控多媒体设备 120 的语音合成指令，将合成语音指令的电平信号转换为音频信息，为智能语音系统与用户交互输出的一部分。

为了更加清楚的说明本公开实施例的基于人脸识别的语音处理方法，下面以该方法结

合上述系统的应用为例进行说明，其中，为了说明的方便，分别集中在预存设置数据侧和用户个性化语音服务识别侧进行描述。

在本示例中，目标位置为驾驶位置。

首先，集中在预存设置数据侧描述。

5 图 5 是根据本公开另一个实施例的基于人脸识别的语音处理方法的流程图，其中，如图 5 所示，该方法包括：

步骤 201，启动中控多媒体设备。

其中，启动完成的条件是多媒体各项功能可以正常工作，可以调用车载摄像头、麦克风、扬声器等外设。

10 步骤 202，中控多媒体启动后，调用车载摄像头，采集驾驶员位置上的驾驶员脸部信息，提取驾驶员的脸部特征信息。

步骤 203，用户设置智能语音系统的设置选项，对语音识别系统按照个人的喜好进行个性化设置，确定设置数据。

15 用户设置智能语音系统的设置选项时，可以是使用智能语音来进行语音设置，也可以是通过中控多媒体设备触摸显示屏上的设置界面进行设置。

步骤 204，用户保存智能语音系统设置选项的内容，中控多媒体设备提取当前用户的面部脸部特征信息与智能语音系统设置数据形成列表，对比已存储的列表中有用户面部特征信息相同，则判断为同一用户，只更新设置数据，如果没有相同，则判断为新用户，将新用户的面部特征信息和设置数据存储添加到本地列表中。

20 其中，在本实施例中，当中控多媒体采集不到驾驶员脸部特征信息时，例如驾驶座无人、摄像头被遮挡，存储设置数据时，智能语音系统的设置数据列表中，用户脸部特征信息空缺，之后所有无法确定用户脸部特征信息的设置数据都存储在这一用户信息空缺栏中。

下面其次集中在用户个性化语音服务识别侧进行描述。

25 图 6 是根据本公开又一个实施例的基于人脸识别的语音处理方法的流程图，如图 6 所示，该方法包括：

步骤 301，启动中控多媒体设备。

其中，启动完成的条件是多媒体各项功能可以正常工作，可以调用车载摄像头、麦克风、扬声器等外设。

30 步骤 302，中控多媒体启动后，调用车载摄像头，采集驾驶员位置上的驾驶员脸部信息，提取驾驶员的脸部特征信息。

步骤 303，中控多媒体提取当前驾驶员脸部特征信息，检索对比本地已有用户设置信息列表，如果有用户信息相同，则提取对应的智能语音系统设置数据，如果没有用户信息，

则设置智能语音系统的设置数据为默认设置。

步骤 304, 根据提取的用户设置数据, 中控多媒体主动设置智能语音系统的设置选项, 完成对不同用户不同系统设置的自适应。

其中, 中控多媒体提取用户脸部特征信息, 经检索对比, 如果不在已有的设置数据中, 5 则判断该用户为新用户, 智能语音系统的设置选项初始化为默认值。如果在已有的设置数据中, 则提取设置信息, 用户调用车载智能语音系统时, 可以用户自己设定的唤醒词, 智能语音功能被唤醒时应答用户设定的称呼, 包括播报音量、播报声音、播报方言等, 都是符合用户自己设定的需求。

由此通过用户人脸识别、设置数据、设置数据提取, 形成了自动的对不同用户存储不 10 同设置数据、设置不同的目标语音服务, 达到了智能语音系统设置的智能化与自动化目的。

综上所述, 本公开实施例的基于人脸识别的语音处理方法, 采集车内目标位置上的用户脸部信息, 提取用户的脸部特征信息, 检索预存设置数据, 判断是否能够获取与脸部特征信息对应的语音设置信息, 若判断获知能够获取与脸部特征信息对应的语音设置信息, 则根据语音设置信息对车载智能语音系统设置与用户对应的目标语音服务, 以调用目标语音 15 服务响应用户发送的操作指令。由此, 基于不同的用户进行不同的语音服务, 满足了用户的个性化要求, 提高了用户的智能化体验。

为了实现上述实施例, 本公开还提出了一种车载智能语音系统, 图 7 是根据本公开一个实施例的车载智能语音系统的结构示意图, 如图 7 所示, 该系统包括提取模块 1100、判 20 断模块 1200 和处理模块 1300。

其中, 提取模块 1100, 用于采集车内目标位置上的用户脸部信息, 提取用户的脸部特征信息。

在本公开的一个实施例中, 提取模块 1100 具体用于按照预设周期采集车内目标位置上的用户脸部信息; 或者, 检测是否满足预设的采集触发条件, 若检测满足预设的采集触发 25 条件, 则采集车内目标位置上的用户脸部信息。

判断模块 1200, 用于检索预存设置数据, 判断是否能够获取与脸部特征信息对应的语音设置信息。

处理模块 1300, 用于在判断获知能够获取与脸部特征信息对应的语音设置信息时, 根据语音设置信息对车载智能语音系统设置与用户对应的目标语音服务, 以调用目标语音服 30 务响应用户发送的操作指令。

在本公开的一个实施例中, 处理模块 1300, 还用于在判断获知不存在与脸部特征信息对应的语音设置信息时, 调用车载智能语音系统预先根据默认设置信息设置的标准语音服

务响应用户发送的操作指令。

综上所述，本公开实施例的车载智能语音系统，采集车内目标位置上的用户脸部信息，提取用户的脸部特征信息，检索预存设置数据，判断是否能够获取与脸部特征信息对应的语音设置信息，若判断获知能够获取与脸部特征信息对应的语音设置信息，则根据语音设置信息对车载智能语音系统设置与用户对应的目标语音服务，以调用目标语音服务响应用户发送的操作指令。由此，基于不同的用户进行不同的语音服务，满足了用户的个性化要求，提高了用户的智能化体验。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本公开的至少一个实施例或示例中。在本说明书中，对上述术语的示意性表述不必须针对的是相同的实施例或示例。而且，描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外，在不相互矛盾的情况下，本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

此外，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本公开的描述中，“多个”的含义是至少两个，例如两个，三个等，除非另有明确具体的限定。

流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为，表示包括一个或更多个用于实现定制逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分，并且本公开的优选实施方式的范围包括另外的实现，其中可以不按所示出或讨论的顺序，包括根据所涉及的功能按基本同时的方式或按相反的顺序，来执行功能，这应被本公开的实施例所属技术领域的技术人员所理解。

在流程图中表示或在此以其他方式描述的逻辑和/或步骤，例如，可以被认为用于实现逻辑功能的可执行指令的定序列表，可以具体实现在任何计算机可读介质中，以供指令执行系统、装置或设备（如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统）使用，或结合这些指令执行系统、装置或设备而使用。就本说明书而言，“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。计算机可读介质的更具体的示例（非穷尽性列表）包括以下：具有一个或多个布线的电连接部（电子装置），便携式计算机盘盒（磁装置），随机存取存储器（RAM），只读存储器（ROM），可擦除可编程只读存储器（EPROM 或闪速存储器），光纤装置，以及

便携式光盘只读存储器（CDROM）。另外，计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质，因为可以例如通过对纸或其他介质进行光学扫描，接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得所述程序，然后将其存储在计算机存储器中。

5 应当理解，本公开的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中，多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。如，如果用硬件来实现和在另一实施方式中一样，可用本领域公知的下列技术中的任一项或他们的组合来实现：具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路，具有合适的组合逻辑门电路的专用集成电路，可编程门阵列（PGA），现场可编程门阵列（FPGA）等。

10 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，该程序在执行时，包括方法实施例的步骤之一或其组合。

15 此外，在本公开各个实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用时，也可以存储在一个计算机可读存储介质中。

20 上述提到的存储介质可以是只读存储器，磁盘或光盘等。尽管上面已经示出和描述了本公开的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本公开的限制，本领域的普通技术人员在本公开的范围内可以对上述实施例进行变化、修改、替换和变型。

权利要求书

1、一种基于人脸识别的语音处理方法，其特征在于，包括以下步骤：

采集车内目标位置上的用户脸部信息，提取所述用户的脸部特征信息；

5 检索预存设置数据，判断是否能够获取与所述脸部特征信息对应的语音设置信息；

若判断获知能够获取与所述脸部特征信息对应的语音设置信息，则根据所述语音设置信息对车载智能语音系统设置与所述用户对应的目标语音服务，以调用所述目标语音服务响应所述用户发送的操作指令。

2、如权利要求 1 所述的方法，其特征在于，在所述判断是否能够获取与所述脸部特征信息对应的语音设置信息之后，还包括：

若判断获知不存在与所述脸部特征信息对应的语音设置信息，则调用所述车载智能语音系统预先根据默认设置信息设置的标准语音服务响应所述用户发送的操作指令。

3、如权利要求 1 或 2 所述的方法，其特征在于，在所述提取所述用户的脸部特征信息之后，还包括：

15 接收所述用户输入的包含语音设置信息的设置指令；

查询预存设置数据，判断是否包含所述用户的脸部特征信息；

若判断获知包含所述用户的脸部特征信息，则删除当前存储的与所述脸部特征信息对应的语音设置信息，并存储所述语音设置信息与所述脸部特征信息的对应关系；

若判断获知不包含所述用户的脸部特征信息，则在所述设置数据中存储所述语音设置信息与所述用户的脸部特征信息的对应关系。

4、如权利要求 1-3 任一所述的方法，其特征在于，所述语音设置信息，包括：

播报音量、播报声音、播报方言中的一种或者多种信息组合。

5、如权利要求 1-4 任一所述的方法，其特征在于，在所述采集车内目标位置上的用户脸部信息之前，还包括：

25 检测车内环境的光线亮度，将所述光线亮度与预设阈值进行比较；

若比较获知所述光线亮度小于预设阈值，则对所述目标位置启动辅助照明设备。

6、如权利要求 1-5 任一所述的方法，其特征在于，所述采集车内目标位置上的用户脸部信息，包括：

按照预设周期采集车内目标位置上的用户脸部信息；或者，

30 检测是否满足预设的采集触发条件，若检测满足预设的采集触发条件，则采集车内目标位置上的用户脸部信息。

7、如权利要求 1-6 任一所述的方法，其特征在于，还包括：

通过所述车载智能语音系统的屏幕获取所述用户通过操作面板输入的操作指令；
或者，

通过所述车载智能语音系统的麦克风获取所述用户语音输入的操作指令。

8、如权利要求 7 所述的方法，其特征在于，还包括：

5 对所述用户语音输入的操作指令进行语音识别；

调用本地的语音合成指令或者获取云端反馈的响应信息，通过所述车载智能语音系统
输出给所述用户。

9、一种车载智能语音系统，其特征在于，包括：

提取模块，用于采集车内目标位置上的用户脸部信息，提取所述用户的脸部特征信息；

10 判断模块，用于检索预存设置数据，判断是否能够获取与所述脸部特征信息对应的语音
设置信息；

处理模块，用于在判断获知能够获取与所述脸部特征信息对应的语音设置信息时，根据
所述语音设置信息对车载智能语音系统设置与所述用户对应的目标语音服务，以调用所
述目标语音服务响应所述用户发送的操作指令。

15 10、如权利要求的 9 所述的车载智能语音系统，其特征在于，所述处理模块，还用于：
在判断获知不存在与所述脸部特征信息对应的语音设置信息时，调用所述车载智能语
音系统预先根据默认设置信息设置的标准语音服务响应所述用户发送的操作指令。

11、如权利要求的 9 或 10 所述的车载智能语音系统，其特征在于，所述提取模块，具
体用于：

20 按照预设周期采集车内目标位置上的用户脸部信息；或者，

检测是否满足预设的采集触发条件，若检测满足预设的采集触发条件，则采集车内目
标位置上的用户脸部信息。

12、一种车机设备，其特征在于，包括：车载摄像头和中控多媒体设备，其中，所述
车载摄像头与所述中控多媒体设备连接，其中，

25 所述车载摄像头，用于采集车内目标位置上的用户脸部信息，并将所述用户脸部信息
发送至所述中控多媒体设备；

所述中控多媒体设备，用于根据提取所述用户的脸部特征信息，检索预存设置数据，
判断是否能够获取与所述脸部特征信息对应的语音设置信息，若判断获知能够获取与所述
脸部特征信息对应的语音设置信息，则将所述语音设置信息设置与所述用户对应的目标语
音服务，以调用所述目标语音服务响应所述用户发送的操作指令。

30 13、如权利要求 12 所述的车机设备，其特征在于，还包括：

麦克风，其中，所述麦克风与所述中控多媒体设备连接，其中，

所述麦克风，用于接收所述用户输入的包含语音设置信息的设置指令；

所述中控多媒体设备，用于查询预存设置数据，判断是否包含所述用户的脸部特征信息，若判断获知包含所述用户的脸部特征信息，则删除当前存储的与所述脸部特征信息对应的语音设置信息，并存储所述语音设置信息与所述脸部特征信息的对应关系，若判断获
5 知不包含所述用户的脸部特征信息，则在所述设置数据中存储所述语音设置信息与所述用户的脸部特征信息的对应关系。

14、如权利要求 12 或 13 所述的车机设备，其特征在于，还包括：

扬声器，其中，所述扬声器与所述中控多媒体设备连接，其中，

所述扬声器，用于根据所述语音设置信息对所述用户语音输入的操作指令进行语音响
10 应。

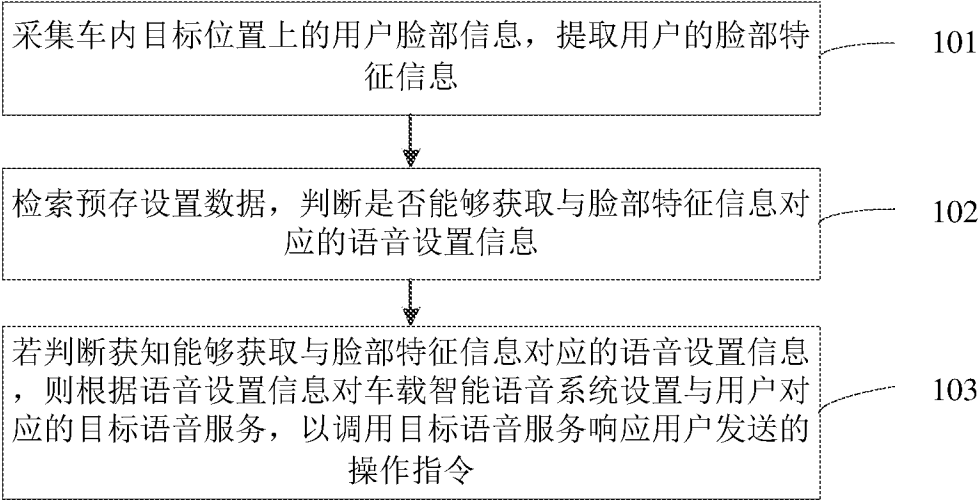


图 1

智能语音系统设置	
用户1	用户1设置数据
用户2	用户2设置数据
用户3	用户3设置数据

图 2

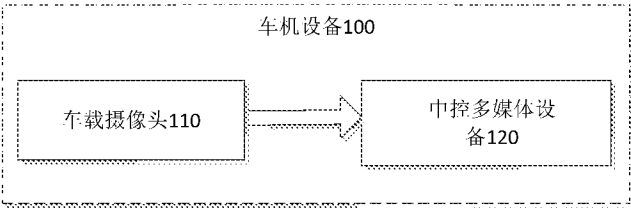


图 3

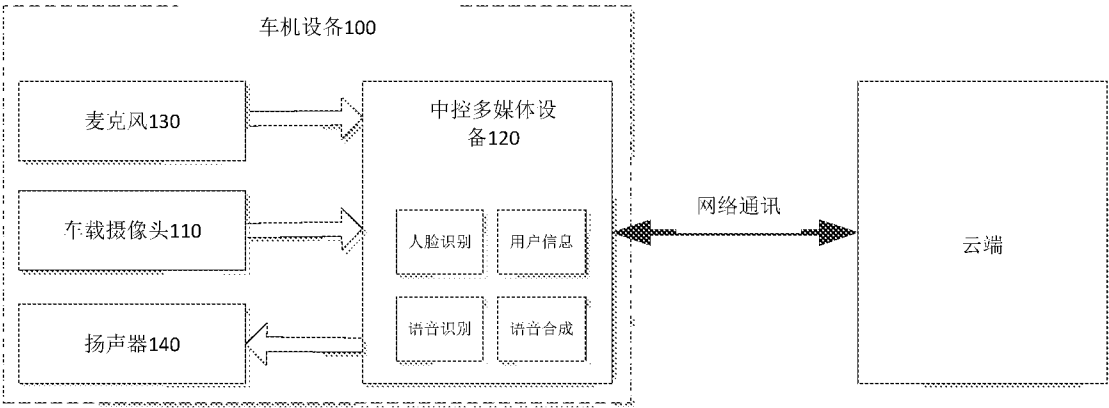


图 4

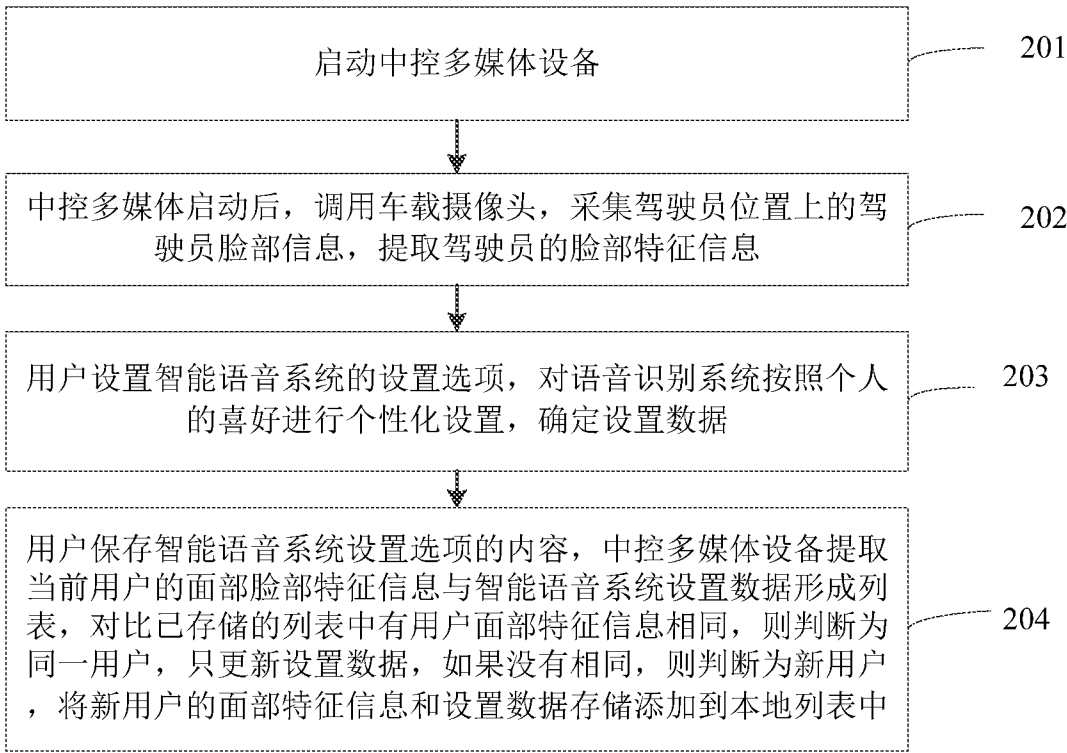


图 5

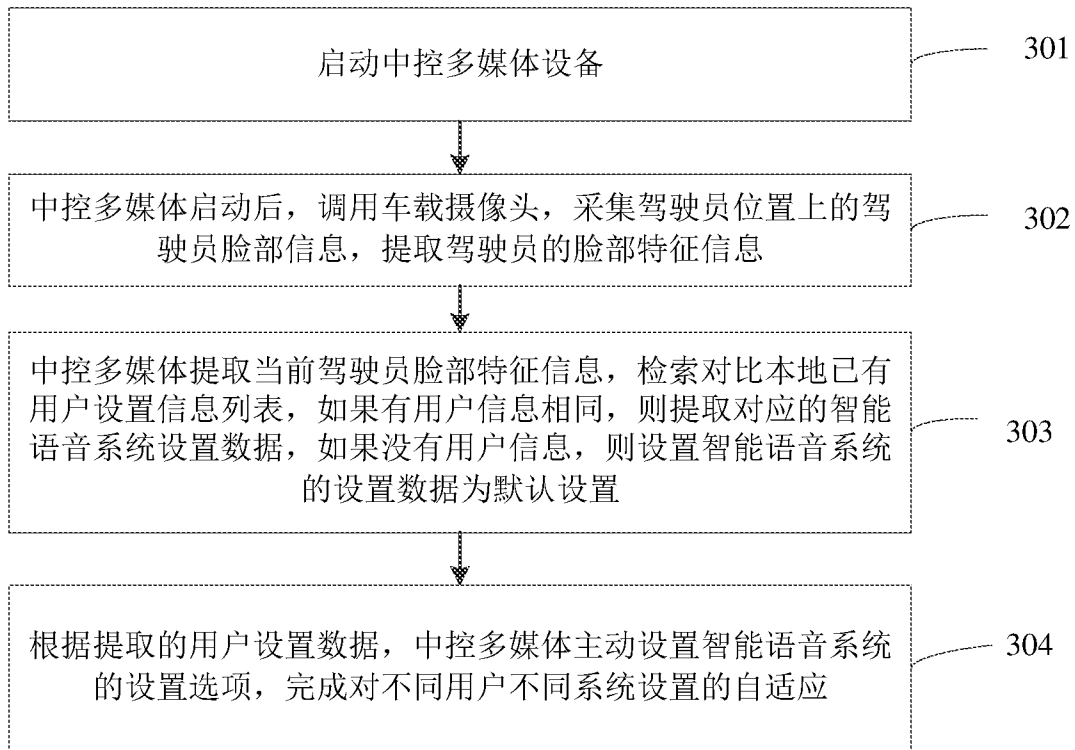


图 6

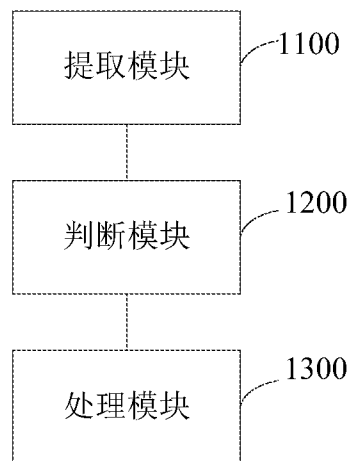


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/083261

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/00(2013.01)i; G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L; G06K9

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, VEN: 人脸, 识别, 语音, 车载, 设置, 配置, 个性, 特定, 专用, FACE, RECOGNITION, SPEECH, PERSONALITY

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2017247000 A1 (AUTOCONNECT HOLDINGS LLC.) 31 August 2017 (2017-08-31) figures 8 and 20, and description, paragraphs [0002]-[0022], [0228]-[0247] and [0372]-[0426]	1-4, 7-10, 12-14
Y	US 2017247000 A1 (AUTOCONNECT HOLDINGS LLC.) 31 August 2017 (2017-08-31) figures 8 and 20, and description, paragraphs [0002]-[0022], [0228]-[0247] and [0372]-[0426]	5-6, 11
Y	CN 105825167 A (VIVO MOBILE COMMUNICATION CO., LTD.) 03 August 2016 (2016-08-03) figure 2, and description, paragraphs [0057]-[0067]	5-6, 11
Y	CN 104795067 A (HUAWEI TECHNOLOGIES CO., LTD.) 22 July 2015 (2015-07-22) figures 1-6, and description, paragraphs [0069]-[0166]	1-14
Y	CN 107919117 A (DELPHI ELECTRONIC (SUZHOU) CO., LTD.) 17 April 2018 (2018-04-17) figures 1-2, and description, paragraphs [0038]-[0043]	1-14
Y	CN 107122692 A (HU, JIANJIA) 01 September 2017 (2017-09-01) figures 1-5, and description, paragraphs [0012]-[0040]	1-14



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

12 June 2019

Date of mailing of the international search report

08 July 2019

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/083261

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107391122 A (GREE ELECTRIC APPLIANCES INC. OF ZHUHAI) 24 November 2017 (2017-11-24) entire document	1-14
A	CN 106650601 A (WUHAN HUINENG ROBOT TECHNOLOGY CO., LTD.) 10 May 2017 (2017-05-10) entire document	1-14
A	CN 206877691 U (GREE ELECTRIC APPLIANCES INC. OF ZHUHAI) 12 January 2018 (2018-01-12) entire document	1-14

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/083261

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2017247000	A1	31 August 2017	None			
CN	105825167	A	03 August 2016	None			
CN	104795067	A	22 July 2015	US	9990924	B2	05 June 2018
				US	2017263251	A1	14 September 2017
				US	2015206533	A1	23 July 2015
				US	9583101	B2	28 February 2017
				US	2018247650	A1	30 August 2018
CN	107919117	A	17 April 2018	None			
CN	107122692	A	01 September 2017	None			
CN	107391122	A	24 November 2017	None			
CN	106650601	A	10 May 2017	None			
CN	206877691	U	12 January 2018	None			

国际检索报告

国际申请号

PCT/CN2019/083261

A. 主题的分类

G10L 15/00(2013.01)i; G06K 9/00(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L; G06K9

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, CNTXT, CNKI, VEN:人脸, 识别, 语音, 车载, 设置, 配置, 个性, 特定, 专用, FACE, RECOGNITION, SPEECH, PERSONALITY

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	US 2017247000 A1 (AUTOCONNECT HOLDINGS LLC.) 2017年 8月 31日 (2017 - 08 - 31) 图8、20, 说明书第【0002】-【0022】、【0228-247】、【0372-426】段	1-4, 7-10, 12-14
Y	US 2017247000 A1 (AUTOCONNECT HOLDINGS LLC.) 2017年 8月 31日 (2017 - 08 - 31) 图8、20, 说明书第【0002-0022】、【0228】-【0247】、【0372】-【0426】段	5-6, 11
Y	CN 105825167 A (维沃移动通信有限公司) 2016年 8月 3日 (2016 - 08 - 03) 图2, 说明书第【0057】-【0067】段	5-6, 11
Y	CN 104795067 A (华为技术有限公司) 2015年 7月 22日 (2015 - 07 - 22) 图1-6, 说明书第【0069】-【0166】段	1-14
Y	CN 107919117 A (德尔福电子苏州有限公司) 2018年 4月 17日 (2018 - 04 - 17) 图1-2, 说明书第【0038】-【0043】段	1-14
Y	CN 107122692 A (胡渐佳) 2017年 9月 1日 (2017 - 09 - 01) 图1-5, 说明书第【0012】-【0040】段	1-14
A	CN 107391122 A (珠海格力电器股份有限公司) 2017年 11月 24日 (2017 - 11 - 24) 全文	1-14
A	CN 106650601 A (武汉慧能机器人科技有限公司) 2017年 5月 10日 (2017 - 05 - 10) 全文	1-14

☒ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2019年 6月 12日

国际检索报告邮寄日期

2019年 7月 8日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

王晓萍

传真号 (86-10)62019451

电话号码 62085739

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 206877691 U (珠海格力电器股份有限公司) 2018年 1月 12日 (2018 - 01 - 12) 全文	1-14

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/083261

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
US	2017247000	A1	2017年 8月 31日	无			
CN	105825167	A	2016年 8月 3日	无			
CN	104795067	A	2015年 7月 22日	US	9990924	B2	2018年 6月 5日
				US	2017263251	A1	2017年 9月 14日
				US	2015206533	A1	2015年 7月 23日
				US	9583101	B2	2017年 2月 28日
				US	2018247650	A1	2018年 8月 30日
CN	107919117	A	2018年 4月 17日	无			
CN	107122692	A	2017年 9月 1日	无			
CN	107391122	A	2017年 11月 24日	无			
CN	106650601	A	2017年 5月 10日	无			
CN	206877691	U	2018年 1月 12日	无			

表 PCT/ISA/210 (同族专利附件) (2015年1月)