

(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 6 月 6 日 (06.06.2019)



(10) 国际公布号
WO 2019/105376 A1

- (51) 国际专利分类号:
G06F 3/01 (2006.01)
- (21) 国际申请号: PCT/CN2018/117864
- (22) 国际申请日: 2018 年 11 月 28 日 (28.11.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201711237225.9 2017 年 11 月 30 日 (30.11.2017) CN
- (71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。
- (72) 发明人: 刘永霞 (LIU, Yongxia); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35

层, Guangdong 518057 (CN)。史波 (SHI, Bo); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。易科 (YI, Ke); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。李瑞 (LI, Rui); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。刘杉 (LIU, Shan); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

- (74) 代理人: 北京德琦知识产权代理有限公司 (DEQI INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区知春路 1 号学院国际大厦 7 层, Beijing 100083 (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,

(54) Title: GESTURE RECOGNITION METHOD, TERMINAL AND STORAGE MEDIUM

(54) 发明名称: 手势识别方法、终端及存储介质

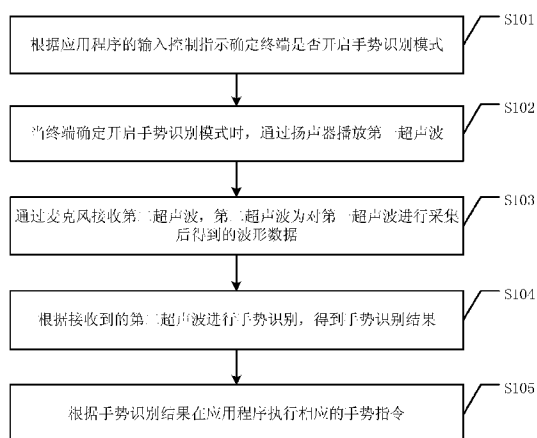


图 1b

S101 DETERMINE, ACCORDING TO AN INPUT CONTROL INSTRUCTION OF AN APPLICATION PROGRAM, WHETHER A TERMINAL STARTS A GESTURE RECOGNITION MODE
S102 WHEN IT IS DETERMINED THAT THE TERMINAL STARTS THE GESTURE RECOGNITION MODE, PLAY A FIRST ULTRASONIC WAVE BY MEANS OF A SPEAKER
S103 RECEIVE A SECOND ULTRASONIC WAVE BY MEANS OF A MICROPHONE, THE SECOND ULTRASONIC WAVE BEING WAVEFORM DATA OBTAINED UPON ACQUISITION OF THE FIRST ULTRASONIC WAVE
S104 PERFORM GESTURE RECOGNITION ACCORDING TO THE RECEIVED SECOND ULTRASONIC WAVE, TO OBTAIN A GESTURE RECOGNITION RESULT
S105 EXECUTE A CORRESPONDING GESTURE INSTRUCTION ON THE APPLICATION PROGRAM ACCORDING TO THE GESTURE RECOGNITION RESULT

(57) Abstract: The embodiments of the present application provide a gesture recognition method. The method is applied to a terminal, which terminal is provided with a speaker and a microphone and has an application program installed thereon. The method comprises: determining, according to an input control instruction of the application program, whether the terminal starts a gesture recognition mode; when it is determined that the terminal starts the gesture recognition mode, playing a first ultrasonic wave by means of the speaker; receiving a second ultrasonic wave by means of the microphone, the second ultrasonic wave being waveform data obtained upon acquisition of the first ultrasonic wave; performing gesture recognition according to the received second ultrasonic wave, to obtain a gesture recognition result; and executing a corresponding gesture instruction on the application program according to the gesture recognition result.

(57) 摘要: 本申请实例提供一种手势识别方法, 所述方法应用于终端, 所述终端中配置有扬声器和麦克风, 所述终端上安装有应用程序, 所述方法包括: 根据所述应用程序的输入控制指示确定所述终端是否开启手势识别模式; 当所述终端确定开启手势识别模式时, 通过所述扬声器播放第一超声波; 通过所述麦克风接收第二超声波, 所述第二超声波为对所述第一超声波进行采集后得到的波形数据; 根据接收到的所述第二超声波进行手势识别, 得到手势识别结果; 根据所述手势识别结果在所述应用程序执行相应的手势指令。

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, 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))。

手势识别方法、终端及存储介质

本申请要求于 2017 年 11 月 30 日提交中国专利局、申请号为 201711237225.9、发明名称为“手势识别方法和终端”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本申请涉及计算机技术领域，尤其涉及一种手势识别方法、终端及存储介质。

背景

10 用户可以使用手机浏览器来获取网络资源，例如访问网页。用户在手机浏览器上的输入方式主要是接触式输入，即用户需要在手机终端的触摸屏或者物理按键上进行手动输入。

用户还可以使用非接触的输入方式来进行输入，例如手势识别，可以方便用户的使用。其中，手势是指在人的意识支配下，人手作出的各类动作，如手指弯曲、伸展和手在空间的运动等，可以是执行某项任务，
15 也可以是与人的交流，以表达某种含义或意图。基于手势识别的三维交互输入技术，常用的有基于数据手套的和基于视觉（如摄像机）的手势识别。

技术内容

本申请实例提供一种手势识别方法，所述方法应用于终端，所述终端中配置有扬声器和麦克风，所述终端上安装有应用程序，所述方法包括：
20

根据所述应用程序的输入控制指示确定所述终端是否开启手势识别模式；

当所述终端确定开启手势识别模式时，通过所述扬声器播放第一超声波；

通过所述麦克风接收第二超声波，所述第二超声波为对所述第一超声波进行采集后得到的波形数据；

- 5 根据接收到的所述第二超声波进行手势识别，得到手势识别结果；
根据所述手势识别结果在所述应用程序执行相应的手势指令。

本申请实例还提供一种终端，所述终端中配置有扬声器和麦克风，所述终端上安装有应用程序，所述终端包括：

- 模式确定模块，用于根据所述应用程序的输入控制指示确定所述终端是否开启手势识别模式；

超声波发送模块，用于当所述终端确定开启手势识别模式时，通过所述扬声器播放第一超声波；

超声波采集模块，用于通过所述麦克风接收第二超声波，所述第二超声波为对所述第一超声波进行采集后得到的波形数据；

- 15 手势识别模块，用于根据接收到的所述第二超声波进行手势识别，
得到手势识别结果；

指令执行模块，用于根据所述手势识别结果在所述应用程序执行相应的手势指令。

- 本申请实例还提供一种终端，所述终端包括：扬声器、麦克风、处理器和存储器；

所述处理器和所述扬声器、所述麦克风、所述存储器进行相互的通信；

所述存储器用于存储指令；

所述扬声器，用于在所述处理器的控制下播放第一超声波；

- 25 所述麦克风，用于在所述处理器的控制下接收第二超声波；

所述处理器用于执行所述存储器中的所述指令，执行如前述第一方

面中任一项所述的方法。

本申请实例提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有指令，当其在计算机上运行时，使得计算机执行上述各方面所述的方法。

5 附图简要说明

为了更清楚地说明本申请实例中的技术方案，下面将对实例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实例，对于本领域的技术人员来讲，还可以根据这些附图获得其他的附图。

- 10 图 1a 为本申请实例涉及的一种系统构架示意图；
图 1b 为本申请实例提供的一种手势识别方法的流程方框示意图；
图 1c 为本申请实例涉及的一种系统构架示意图；
图 2 为本申请实例提供的超声波发送的流程示意图；
图 3 为本申请实例提供的超声波接收的流程示意图；
15 图 4 为本申请实例提供的一维手势识别处理的流程示意图；
图 5 为本申请实例提供的二维手势识别处理的流程示意图；
图 6 为本申请实例提供的手势识别动作管理的流程示意图；
图 7-a 为本申请实例提供的手势识别的动作开启的示意图；
图 7-b 为本申请实例提供的手势识别的动作结束的示意图；
20 图 8 为本申请实例提供的手势识别的基本操作的示意图；
图 9-a 为本申请实例提供的一种终端的组成结构示意图；
图 9-b 为本申请实例提供的一种模式确定模块的组成结构示意图；
图 9-c 为本申请实例提供的另一种终端的组成结构示意图；
图 9-d 为本申请实例提供的另一种终端的组成结构示意图；
25 图 10 为本申请实例提供的手势识别方法应用于终端的组成结构示意图；

图 11 为本申请实例提供另一种终端的组成结构示意图。

实施方式

本申请实例提供了一种手势识别方法和终端，用于降低在终端的应用程序上的手势识别的复杂度，降低对终端的计算性能的要求。

5 为使得本申请的目的、特征、优点能够更加的明显和易懂，下面将结合本申请实例中的附图，对本申请实例中的技术方案进行清楚、完整地描述，显然，下面所描述的实例仅仅是本申请一部分实例，而非全部实例。基于本申请中的实例，本领域的技术人员所获得的所有其他实例，都属于本申请保护的范围。

10 本申请的说明书和权利要求书及上述附图中的术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，以便包含一系列单元的过程、方法、系统、产品或设备不必限于那些单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它单元。

15 以下分别进行详细说明。

接触式输入方式中，在有些场合不方便或者无法进行接触式输入时，例如，在厨房做菜，需要翻食谱时不方便进行翻页切换等操作，又如在开车的时候，不方便进行手机操作，用户均无法进行输入。

20 基于图像识别的方式中，终端可以基于用户所作出的手势的图像来识别用户的手势，然而图像识别算法复杂、计算量大、功耗高、容易受到光线的影响，对终端设备的性能要求高，因此，该方式无法实现在终端上的精确识别。

基于上述缺陷，本申请提出一种手势识别方法，在一些实例中，该方法具体可以应用于终端对用户的手势识别场景中。请参阅图 1a 所示，
25 本申请实例提供的手势识别方法应用于终端 101，终端 101 中配置有扬声器 102 和麦克风 103，终端 101 上还安装有应用程序 104，例如终端

上可安装的应用程序包括：浏览器 APP，或者办公软件 APP，或者游戏 APP 等，此处不做限定。

图 1b 是根据本申请一个示例性实例示出的一种手势识别方法的流程示意图。

5 如图 1b 所示，该手势识别方法，包括如下步骤：

S101、根据应用程序的输入控制指示确定终端是否开启手势识别模式。

在本申请实例中，终端上安装有应用程序，用户可以使用该应用程序，例如用户可以使用浏览器 APP 查询网页。输入控制指示具体可以设置在应用程序的设置菜单上，也可以设置在该应用程序的配置文件中，
10 例如输入控制指示可以设置在安装配置文件中，在应用程序安装到终端上时由用户来配置该输入控制指示。又如，在应用程序的输入控制指示菜单上可以设置手势识别模式的开启页面，若用户选择启用手势识别模式，则可以确定该终端开启了手势识别模式，若用户没有开启手势识别
15 模式，则可以确定该终端关闭了手势识别模式。

需要说明的是，在本申请实例中，当终端确定开启手势识别模式时，才触发执行后续步骤，否则结束本申请实例中的手势识别方法。本申请实例中通过应用程序的输入控制指示可以控制终端是否启用手势识别模式。

20 在本申请的一些实例中，步骤 S101 根据应用程序的输入控制指示确定终端是否开启手势识别模式，包括：

根据应用程序的配置文件确定应用程序的输入控制指示是否为默认开启手势识别模式；

检测应用程序在终端上是否运行成功；

25 在应用程序运行成功时，若终端默认开启手势识别模式，确定终端开启了手势识别模式。

其中，本申请实例中可以配置应用程序的输入控制指示为默认开启手势识别模式，则可以在应用程序在终端上运行起来后就自动打开手势识别模式，从而可以省去用户重新设置是否开启的麻烦，简化用户的操作。使得用户在运行某个应用程序时，就可以自动启用手势识别模式。

- 5 在本申请的一些实例中，步骤 S101 根据应用程序的输入控制指示确定终端是否开启手势识别模式，包括：

根据应用程序的配置文件确定应用程序的输入控制指示为第一二维手势开启手势识别模式，或者第二二维手势关闭手势识别模式；

当检测到第一二维手势时，确定终端开启了手势识别模式；

- 10 当检测到第二二维手势时，确定终端关闭了手势识别模式。

其中，本申请实例中手势识别模式的开启和关闭还可以由预设的一些手势动作来决定，例如可以预设一些二维手势，通过这些预设的二维手势由用户来控制终端是否开启手势识别模式，从而满足用户对终端的实时控制需求。举例说明，根据应用程序的配置文件确定应用程序的输入控制指示可以为第一二维手势开启手势识别模式，则终端可以通过扬声器发射超声波，通过麦克风采集超声波，从而识别出用户是否做了预设的某一个手势动作（例如第一二维手势），只有该预设的手势动作被检测出时可以开启手势识别模式。又如，根据应用程序的配置文件确定应用程序的输入控制指示可以为第二二维手势关闭手势识别模式，则终端可以通过扬声器发射超声波，通过麦克风采集超声波，从而识别出用户是否做了预设的某一个手势动作（例如第二二维手势），只有该预设的手势动作被检测出时就可以关闭手势识别模式。

在本申请的一些实例中，除了执行前述的方法步骤之外，本申请实例提供的手势识别方法还可以包括如下步骤：

- 25 当扬声器播放训练超声波时，通过麦克风录制指定手势动作的波形数据；

根据指定手势动作的波形数据建立手势波形模板库；

根据手势波形模板库确定指定手势动作对应的手势识别结果。

本申请实例中，用户操作不同类型的终端进行手势识别时，由于不同的终端机型对用户手势动作接收到的波形数据存在差异，为了提升手势动作的匹配准确度，本申请实例中建立手势波形模板库。例如为了解决终端的计算能力不足问题，还可以建立基于云端服务器的手势波形模板库。也即，在一些实例中，本申请实例提供的手势识别方法还可以应用于如图 1c 所示的系统构架中。

如图 1c 所示，该系统构架包括终端 101 和云端服务器 105，两者通过一个或多个互联网 106 进行交互。

用户通过终端 101 录制了指定手势动作的波形数据，云端服务器 105 基于上述交互就可以建立手势波形模板库，采用机器学习的方法进行手势识别，提升手势匹配的准确度。在一些实例中，云端服务器 105 可以包括两个组成部分：一部分是离线手势波形模板库训练，一部分是在线手势识别。

终端 101 可以将接收到的第二超声波提取到超声波数据，然后将该超声波数据发送给云端服务器 105，该云端服务器 105 可以采用在线云端识别的方式通过手势波形模板库获取到手势识别结果，然后云端服务器 105 将该手势识别结果发送给终端 101，使得不同终端类型都可以获取到适用于该终端类型的手势识别结果，例如，通过终端的配置文件可以获取到终端的类型信息和终端的配置信息，例如终端的系统类型、操作系统版本、终端采用的处理器个数、内容空间大小，从而根据终端的类型信息和终端的配置信息进入为这种终端所设置的手势波形模板库，从而提高手势匹配的准确度。

S102、当终端确定开启手势识别模式时，通过扬声器播放第一超声波。

在本申请实例中，终端内置有扬声器，终端可以首先生成超声波信号，然后通过扬声器播放出超声波，将扬声器播放出的超声波定义为“第一超声波”。该第一超声波从终端所在的位置发射出去之后，该第一超声波碰到目标障碍物会向终端所在的位置进行反射。其中，目标障碍物主要是用户的人手，例如用户的单个或多个手指，或者用户的一个或两个手掌等。

在本申请实例中，终端可以使用内置的超声波发生器来产生超声波，然后通过终端内置的扬声器来发送超声波。超声波发生器通过机械振动的方式产生超声波，通常以纵波的方式在弹性介质内传播，是一种能量的传播形式。超声波的波长短，方向性好。超声波发生器产生的超声波频率高于 20000 赫兹，方向性好，穿透能力强，易于获得较集中的声能。

在本申请的一些实例中，通过扬声器播放第一超声波之后，本申请实例提供的手势识别方法还包括：

15 检测扬声器播放的第一超声波的信号波形是否存在异常；

若第一超声波的信号波形存在异常，重新执行前述步骤 S102：通过扬声器播放第一超声波；

若第一超声波的信号波形不存在异常，触发执行如下步骤 S103：通过麦克风接收第二超声波。

20 其中，在扬声器播放第一超声波之后，可以对播放数据进行异常检测，例如从时域、频域等多维度对第一超声波的信号波形进行检测。举例说明如下，终端可以检测播放的数据有没有丢帧，例如判断采样到的频率点数是否与扬声器在 1 秒钟内播放的频率点数相同。又如，可以检测终端内置的扬声器是否存在播放异常。又如，终端可以在频域查看信号的能量是不是正常，比如播放的超声波信号为 20khz，判断接收到的
25 超声波能量是不是 20khz。又如，终端在播放超声波时，还可以用快节

奏的音乐播放检测有没有丢帧。通过前述对第一超声波的信号波形的异常判断，可以提高接收超声波的成功率。

S103、通过麦克风接收第二超声波，第二超声波为对第一超声波进行采集后得到的波形数据。

5 在本申请实例中，扬声器播出第一超声波之后，可以通过终端内置的麦克风来接收声波信号，将麦克风接收到的声波信号定义为“第二超声波”。其中终端内置的麦克风可以是一个或者多个，此处不做限定。通常从终端的周围环境中接收到的第二超声波是混合信号，混合有通过目标障碍物反射后的超声波，以及扬声器发射的超声波，以及从终端的
10 附近发射出的超声波。

在本申请的一些实例中，步骤 S103 通过麦克风接收第二超声波之后，本申请实例提供的手势识别方法还包括：

检测接收到的第二超声波的信号波形是否存在异常；

若第二超声波的信号波形存在异常，丢弃接收到的第二超声波；

15 若第二超声波的信号波形不存在异常，触发执行如下步骤：根据接收到的第二超声波进行手势识别。

其中，终端采集到第二超声波之后，还可以判断录制的信号波形是否存在异常，例如接收到的信号波形是否完整，即是否存在丢帧的情况。例如可以从频域、时域等多维度进行检测，否则影响后面的手势识别结果。检测的方法有多种，例如判断接收到的信号波形能量是否等于播放的信号波形能量。又如每秒录制的频率点数是不是和采样频率相同。又如信号的时频图能量有没有杂质。通过前述对第二超声波的异常判断，只有在接收到的第二超声波不存在异常时才执行手势识别，从而提高手势识别的准确率。

25 在本申请的一些实例中，步骤 S103 通过麦克风接收第二超声波，包括：

按照预设的录制参数控制麦克风录制终端的周围环境中的声波信号；

将录制到的声波信号存储到终端的录制缓冲区中，其中，录制缓冲区中存储的声波信号为第二超声波。

5 其中，终端预设的录制参数可以有多种，例如单声道或者多声道。终端可以根据需求选择单声道、多声道等方式录制。又如录制参数还可以包括采样频率，该采样频率和超声波的发送频率相同。在采集到声波信号之后，终端可以将该声波信号存储到录制缓冲区中。录制缓冲区的大小可以按照如下方式设置，若录制缓冲区设置的太小，有可能会导
10 丢帧，一般设置 2 倍以上的最小缓冲区的大小。

S104、根据接收到的第二超声波进行手势识别，得到手势识别结果。

在本申请实例中，通过前述步骤 S103 采集到第二超声波之后，可以对接收到的第二超声波进行手势识别，由于扬声器发射出的第一超声波会被用户的人手所阻挡，麦克风会采集到人手阻挡后的超声波，因此
15 通过对麦克风接收到的第二超声波的接收位置、波形能量等分析，就可以识别出用户所做的手势，得到手势识别结果。

在本申请的一些实例中，除了执行前述的方法步骤之外，步骤 S104 根据接收到的第二超声波进行手势识别，得到手势识别结果之后，本申请实例提供的手势识别方法还可以包括如下步骤：

20 通过所述应用程序的显示界面显示手势识别成功的提示信息。

其中，本申请实例中终端的应用程序的显示界面还可以实现与用户的实时互动，通过步骤 S104 据接收到的第二超声波进行手势识别，得到手势识别结果，说明当前的手势识别成功，为了避免用户重复做多个相同的动作，终端的应用程序的显示界面还可以显示提示信息，例如通
25 过显示界面上显示的一个动画告诉用户当前手势识别程序，或者通过文字或者声音来提示用户，此处不做限定。

S105、根据手势识别结果在应用程序执行相应的手势指令。

在本申请实例中，得到手势识别结果之后，终端可以响应用户所在的手势，可以根据手势识别结果执行相应的手势指令，例如响应用户的手势指令来操作终端的应用程序。例如手势识别结果为画圆圈，终端在应用程序上执行手势识别结果对应的手势指令：打开应用程序或者对应用程序的某个菜单进行操作等。其中，手势识别结果与手势指令的对应关系可以根据用户在终端上的预配置列表来确定。举例说明，终端对用户的手势进行识别得到手势识别结果，终端可以在浏览器应用程序上执行用户的手势指令，例如在浏览器的搜索框中输入文字。

10 在本申请的一些实例中，步骤 S102 通过扬声器播放第一超声波，包括：

通过扬声器播放 N 个频率的第一超声波，N 为大于或等于 1 的正整数；

在这种实现场景下，步骤 S103 通过麦克风接收第二超声波，包括：
15 通过麦克风从终端的周围环境中采集到 N 个频率的第二超声波。

其中，终端可以发射一个或多个频率的超声波，例如发送多个不同频率的超声波，这些超声波之间可以间隔相同的距离。终端通过麦克风可以分别采集扬声器播放的一个或多个频率的超声波。

在本申请的一些实例中，步骤 S104 根据接收到的第二超声波进行
20 手势识别，得到手势识别结果，包括：

当 N 的取值为 1 时，根据第二超声波的波形能量进行一维度的手势识别，得到第一手势识别结果，第一手势识别结果包括：手势靠近终端，或者远离终端。

其中，终端可以通过扬声器发射一个超声波，则终端可以通过麦克风采集到该超声波，终端对接收到的一个超声波可以进行粗粒度的一维
25 手势识别，例如通过多普勒的波形能量变化得到第一手势识别结果，该

第一手势识别结果包括：手势靠近终端，或者远离终端。

在本申请的另一些实例中，步骤 S104 根据接收到的第二超声波进行手势识别，得到手势识别结果，包括：

当 N 的取值大于或等于 2 时，对 N 个第二超声波分别计算出手势的
5 移动距离；

从 N 个手势的移动距离中排除出异常的移动距离，对于保留的移动距离进行一维度的手势识别，得到第一手势识别结果，第一手势识别结果包括：手势靠近终端，或者远离终端，或者相对于终端向左移动，或者相对于终端向右移动，或者相对于终端向上移动，或者相对于终端向
10 下移动。

其中，终端可以通过扬声器发射多个超声波，则终端可以通过麦克风分别采集到各个超声波，终端对接收到的多个超声波可以进行细粒度的一维手势识别。例如，采用 N 个频率的超声波相位变化计算距离，反应比较灵敏。发射 17500HZ-23000HZ 中间的 N 个超声波，相邻两个超
15 声波之间间隔相同的距离，采样频率 48000Hz，可以采样 512 个点实时计算距离变化，反应时间在 10.7ms。在声音传输的过程中会有些多径效应，还有些背景静态物体的干扰，为了减少多径效应和背景的干扰，可以用动态信号减去背景的干扰，用 N 个频率计算 N 个距离，根据最小二乘法，计算偏差，把偏差大的频率距离去掉，保留偏差小的距离。通过
20 排除异常的距离，可以提升距离的计算精度。

在本申请的一些实例中，步骤 S103 通过麦克风接收第二超声波，包括：

通过终端的两个麦克风分别采集到第二超声波。

在这种实现场景下，步骤 S104 根据接收到的第二超声波进行手势
25 识别，得到手势识别结果，包括：

根据两个麦克风分别接收到的第二超声波计算手势的相对位置和

初始位置;

根据计算出的相对位置和初始位置进行二维度的手势识别,得到第二手势识别结果,第二手势识别结果包括:手势变化的二维坐标。

其中,为了提高手势识别的精度,终端内可以设置至少两个麦克风来分别采集第二超声波。则对于每个麦克风接收到的第二超声波,都可以计算出手势的相对位置和初始位置。其中,手势的相对位置是指基于第二超声波的相位测量得到手势相对于终端的位置,手势的初始位置是指用户手势在初次识别时的位置。通过前述计算出的手势的相对位置和初始位置,可以进行二维的手势识别,得到手势的精确坐标,提高手势识别的准确度,防止误操作。

通过前述实例对本申请的举例说明可知,终端内设置有扬声器和麦克风,且该终端上安装有应用程序。首先根据应用程序的输入控制指示确定终端是否开启手势识别模式,在该终端确定开启手势识别模式时,通过扬声器播放第一超声波,然后通过麦克风接收第二超声波,第二超声波为对第一超声波进行采集后得到的波形数据,接下来根据接收到的第二超声波进行手势识别,得到手势识别结果,最后根据该手势识别结果在应用程序执行相应的手势指令。本申请实例中,使用终端内置的扬声器和麦克风就可以实现手势识别,即通过超声波的发送与检测就可以识别出用户的动作,降低了手势识别的复杂度,降低了对终端的计算性能的要求,并且实现了对用户在应用程序上的手势指令控制。进一步的,通过对第二超声波进行计算,得出手势的移动距离或手势的二维变化,提高了手势识别的精确率。

为便于更好的理解和实施本申请实例的上述方案,下面举例相应的应用场景,例如,以图 1a 所示的应用场景,来进行具体说明。

本申请实例可以应用到手机浏览器、游戏应用程序(Application, APP)中,本申请实例可以实现高精度、低延迟、无需任何外设、非接

触式的一维和二维的手势识别。本申请实例基于终端的扬声器和麦克风进行超声波手势检测，不需要任何外部设备。举例说明如下，本申请实例中用户不需要手动的操作终端的触摸屏，也不需要从键盘输入任何指令，用户只需要作出手势动作，本申请实例中的终端就可以通过超声波的发送与接收来检测出用户所作出的手势动作，从而在终端的应用程序上执行该手势动作对应的手势指令，例如用户通过做手势来操作游戏应用程序，可以控制游戏应用程序中的角色移动，游戏应用程序的开启与关闭等。

接下来以在浏览器 APP 的浏览场景中的手势识别为例，用户可以操作基本的上、下、左、右等常见的浏览场景。

请参阅图 2 所示，对超声波发送过程进行举例说明。

201、设置超声波的发送频率和采样频率。例如发送 N ($N \geq 1$) 个频率的超声波，采样频率大于等于两倍的发射频率。

202、设置超声波播放方式。

其中，终端可以采用静态和流式等多种播放方式，可以根据用户的需求设置。

超声波的播放要求比较严格，如果出现播放异常或者丢帧的情况，会导致计算异常或者错误，所以超声波的播放流程数据获取流程需要提前准备好，播放流程逻辑需要保证播放的波形数据没有丢帧，即需要播放的波形数据是完整的。

203、检测扬声器播放的声波信号。

其中，终端可以检测播放的数据是否异常，可以从时域、频域等多维度检测。检测有多种方式，举例说明，播放的数据检测有没有丢帧，是不是 1 秒钟播放采样频率个数。又如检测播放模块有没有报异常。又如在频域上查看信号的能量是不是正常，比如播放的频率是 20khz，接收的信号频率是不是 20khz。另外，播放过程还可以用快节奏的音乐

播放检测有没有丢帧。

如图 3 所示，接下来对超声波接收进行举例说明。

301、设置麦克风的录制参数。

其中，可以根据需求选择单声道、多声道等方式录制，采样频率为
5 超声波的发送频率的两倍。根据奈奎斯特采样定律，采样频率和播放频率比较，采样频率可以是播放频率的 2 倍。如果录制缓冲区设置的太小，有可能会

导致丢帧，一般设置 2 倍以上的最小缓冲区的大小。

302、启动录制线程。

其中，从麦克风读取数据，要防止播放的数据录制线程没有录制完
10 全，从而导致出现丢帧的情况。录制线程的逻辑尽量轻量，保证一秒钟接收采样频率个点数。

303、检测麦克风的录制数据。

其中，录制的信号波形是否完整，有无丢帧，需要从频域、时域等多
维度进行检测，否则影响后面的手势识别结果。检测的方法有多种，
15 比如接收到的信号能量是不是播放的频率的能量。又如，每秒录制的点数是不是和采样频率相同。又如，可以通过信号的时频图能量有没有杂质。

本申请实例中，超声波处理过程可以具体是一维手势识别处理，或者二维手势识别处理，接下来分别进行举例说明。

20 首先对一维手势识别处理进行说明，请参阅图 4 所示，一维手势识别处理可实现两种方案。

其中一种粗粒度的手势识别，主要包括：

401、获取超声波的波形能量。采用多普勒原理，手势靠近远离的过程中，在频谱上可以看到波形能量的明显变化。

25 402、对于波形的变化的速度、加速度等多维特征进行建模，得到一维的手势变化。主要是根据多普勒效应，手势靠近远离的过程中，波

形的变化趋势不同，根据波形变化的方向和加速度进行计算。举例说明如下，发射 20000Hz 的超声波，采样频率 44100Hz，每次可以采样 4096 个点进行计算。

另一种细粒度的手势识别，主要包括：

5 403、采用 N 个频率的超声波相位变化计算距离，反应比较灵敏。发射 17500HZ-23000HZ 中间的 N 个超声波，相邻两个超声波之间间隔相同的距离，采样频率 48000Hz，可以采样 512 个点实时计算距离变化，反应时间在 10.7ms，也即，该实例中，手势识别的精确度可以达到 ms 级别。

10 404、为了计算更准确，针对 N 个频率相位变化计算的距离，通过算法排除异常的距离，提升距离的计算精度。声音传输的过程中会有些多径效应，还有些背景静态物体的干扰，为了减少多径效应和背景的干扰，会减去背景的干扰，用 N 个频率计算 N 个距离，根据最小二乘法，计算偏差，把偏差大的频率距离去掉，保留偏差小的距离。

15 接下来对二维手势识别处理进行说明，请参阅图 5 所示，主要包括如下过程：

501、利用两个麦克风分别采集超声波。

502、判断两个麦克风的超声波计算是否完成。

其中，在手势移动的过程中，两个麦克风对波形产生的影响不同，

20 因此可以得到两个方向的变化。

二维手势识别处理主要解决手势动作的二维坐标的识别，从而实现二维图形（画圈、画正方形、画三角形等）或者写汉字的识别。利用扬声器和 2 个麦克风，可以采样两个声道的数据。

对于每个麦克风声道都执行步骤 503 和步骤 504。

25 503、基于一维手势识别，计算出手势移动的相对距离。

其中，基于一维手势识别的相位可以计算出手势移动的相对距离。

504、基于信号计算得到手势的初始位置。

其中，基于麦克风接收到的信号可以计算出手势的粗粒度的初始位置。初始位置是指接收到超声波时检测出的手势的初始位置。

505、结合两个方向上的坐标，得到二维的手势识别。

5 在两个麦克风声道的初始位置和相对位置计算完成之后，可以结合 x 和 y 两个方向上的坐标的计算二维的手势识别，可以得到手势识别的精确坐标 (x, y) ，通过初始位置以及相对位置的变化，可以得到实时的手势位置。

如图 6 所示，接下来对手势识别的管理进行举例说明。通过对手势
10 识别行动作管理，避免误操作和不必要的打扰。主要包括：

601、根据用户的二维手势判断手势识别的功能的开启和关闭。

用户可以根据需求设置自己喜欢的手势，建议采用复杂的二维手势，比如画圈、画正方形、画三角形之类的。如图 7-a 所示，手势识别
15 的动作开启可以是比划一个圆形，如图 7-b 所示，手势识别的动作关闭可以是比划一个正方形。

602、根据用户的一维手势执行手势识别的操作功能。

其中，用户可以设置手势识别的动作，可以满足基本的操作需求。
如图 8 所示，手势识别基本操作可以包括上下左右等基本动作。

需要说明的是，在本申请实例中，为了解决计算负担太重，会影响
20 超声波的播放和录制，进而影响手势识别的准确度，本申请实例中可以采用高效的编程语言，比如计算逻辑用 c 语言来实现，例如采用指针运算，优化处理器的消耗高的逻辑操作，尽量减少内存拷贝、输入输出操作，减少计算时间。因此，本申请实例中一维的手势识别计算的耗时可以减少到 ms 级别。

25 在手势识别场景中，手势识别的准确性和最快速的响应时间是最重要的，本申请实例中可以防止用户的误操作，手势识别模式的开启和结

束可以采用二维手势动作的识别方式。为了更快更精确的识别一个动作，比如靠近、远离、向左、向右，本申请实例提出一维手势的识别方式，其中，基于超声波相位的方式可以达到毫秒（mm）级别的精度。在手机等终端的硬件受限的情况下，本申请实例可以尽可能的提高计算性能，节省计算的时间。

需要说明的是，对于前述的各方法实例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请并不受所描述的动作顺序的限制，因为依据本申请，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实例均属于优选实例，所涉及的动作和模块并不一定是本申请所必须的。

为便于更好的实施本申请实例的上述方案，下面还提供用于实施上述方案的相关装置。

请参阅图 9-a 所示，本申请实例提供一种终端 900，所述终端中配置有扬声器和麦克风，所述终端上安装有应用程序，所述终端可以包括：模式确定模块 901、超声波发送模块 902、超声波采集模块 903、手势识别模块 904 和指令执行模块 905，其中，

模式确定模块 901，用于根据所述应用程序的输入控制指示确定所述终端是否开启手势识别模式；

超声波发送模块 902，用于当所述终端确定开启手势识别模式时，通过所述扬声器播放第一超声波；

超声波采集模块 903，用于通过所述麦克风接收第二超声波，所述第二超声波为对所述第一超声波进行采集后得到的波形数据；

手势识别模块 904，用于根据接收到的所述第二超声波进行手势识别，得到手势识别结果；

指令执行模块 905，用于根据所述手势识别结果在所述应用程序执

行相应的手势指令。

在本申请的一些实例中，所述超声波发送模块 902，具体用于通过所述扬声器播放 N 个频率的第一超声波，所述 N 为大于或等于 1 的正整数；

- 5 所述超声波采集模块，用于通过所述麦克风从所述终端的周围环境中采集到 N 个频率的第二超声波。

进一步的，在本申请的一些实例中，所述手势识别模块 904，具体用于当所述 N 的取值为 1 时，根据所述第二超声波的波形能量进行一维度的手势识别，得到第一手势识别结果，所述第一手势识别结果包括：
10 手势靠近所述终端，或者远离所述终端。

在本申请的一些实例中，请参阅图 9-b 所示，模式确定模块 901，包括：

控制指示解析模块 9011，用于根据所述应用程序的配置文件确定所述应用程序的输入控制指示是否为默认开启手势识别模式；

- 15 程序检测模块 9012，用于检测所述应用程序在所述终端上是否运行成功；

模式开启模块 9013，用于在所述应用程序运行成功时，若所述终端默认开启手势识别模式，确定所述终端开启了手势识别模式。

- 在本申请的一些实例中，模式确定模块 901，具体用于根据所述应用程序的配置文件确定所述应用程序的输入控制指示为第一二维手势开启手势识别模式，或者第二二维手势关闭手势识别模式；当检测到第一二维手势时，确定所述终端开启了手势识别模式；当检测到第二二维手势时，确定所述终端关闭了手势识别模式。
20

- 在本申请的一些实例中，请参阅图 9-c 所示，所述终端 900，还包括：模板库建立模块 906 和手势动作匹配模块 907，其中，
25

超声波采集模块 903，用于当所述扬声器播放训练超声波时，通过

所述麦克风录制指定手势动作的波形数据；

模板库建立模块 906，用于根据所述指定手势动作的波形数据建立手势波形模板库；

5 手势动作匹配模块 907，用于根据所述手势波形模板库确定所述指定手势动作对应的手势识别结果。

在本申请的一些实例中，请参阅图 9-d 所示，所述终端 900，还包括：

显示模块 908，用于所述手势识别模块 904 根据接收到的所述第二超声波进行手势识别，得到手势识别结果之后，通过所述应用程序的显示界面显示手势识别成功的提示信息。

在本申请的一些实例中，所述手势识别模块 904，包括：

距离计算模块，用于当所述 N 的取值大于或等于 2 时，对 N 个所述第二超声波分别计算出手势的移动距离；

15 一维识别模块，用于从 N 个所述手势的移动距离中排除出异常的移动距离，对于保留的移动距离进行一维度的手势识别，得到第一手势识别结果，所述第一手势识别结果包括：手势靠近所述终端，或者远离所述终端，或者相对于所述终端向左移动，或者相对于所述终端向右移动，或者相对于所述终端向上移动，或者相对于所述终端向下移动。

20 在本申请的一些实例中，所述超声波采集模块 903，具体用于通过所述终端的两个麦克风分别采集到第二超声波；

所述手势识别模块 904，包括：

位置计算模块，用于根据所述两个麦克风分别接收到的第二超声波计算手势的相对位置和初始位置；

25 二维识别模块，用于根据计算出的相对位置和初始位置进行二维度的手势识别，得到第二手势识别结果，所述第二手势识别结果包括：手势变化的二维坐标。

在本申请的一些实例中，所述超声波采集模块 903，包括：

声波录制模块，用于按照预设的录制参数控制所述麦克风录制所述终端的周围环境中的声波信号；

信号存储模块，用于将录制到的所述声波信号存储到所述终端的录制缓冲区中，其中，所述录制缓冲区中存储的声波信号为所述第二超声波。

通过以上对本申请实例的描述可知，终端内设置有扬声器和麦克风，且该终端上安装有应用程序。首先根据应用程序的输入控制指示确定终端是否开启手势识别模式，在该终端确定开启手势识别模式时，通过扬声器播放第一超声波，然后通过麦克风接收第二超声波，第二超声波为对第一超声波进行采集后得到的波形数据，接下来根据接收到的第二超声波进行手势识别，得到手势识别结果，最后根据该手势识别结果在应用程序执行相应的手势指令。本申请实例中，使用终端内置的扬声器和麦克风就可以实现手势识别，即通过超声波的发送与检测就可以识别出用户的动作，降低了手势识别的复杂度，降低了对终端的计算性能的要求，并且实现了对用户在应用程序上的手势指令控制。进一步的，通过对第二超声波进行计算，得出手势的移动距离或手势的二维变化，提高了手势识别的精确率。

本申请实例还提供了一种终端，如图 10 所示，为了便于说明，仅示出了与本申请实例相关的部分，具体技术细节未揭示的，请参照本申请实例方法部分。该终端可以为包括手机、平板电脑、PDA（Personal Digital Assistant，个人数字助理）、POS（Point of Sales，销售终端）、车载电脑等任意终端设备，以终端为手机为例：

图 10 示出的是与本申请实例提供的终端相关的手机的部分结构的框图。参考图 10，手机包括：射频（Radio Frequency，RF）电路 1010、存储器 1020、输入单元 1030、显示单元 1040、传感器 1050、音频电路

1060、无线保真（wireless fidelity，WiFi）模块 1070、处理器 1080、以及电源 1090 等部件。本领域技术人员可以理解，图 10 中示出的手机结构并不构成对手机的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。

5 下面结合图 10 对手机的各个构成部件进行具体的介绍：

RF 电路 1010 可用于收发信息或通话过程中，信号的接收和发送，特别地，将基站的下行信息接收后，给处理器 1080 处理；另外，将设计上行的数据发送给基站。通常，RF 电路 1010 包括但不限于天线、至少一个放大器、收发信机、耦合器、低噪声放大器（Low Noise Amplifier，
10 LNA）、双工器等。此外，RF 电路 1010 还可以通过无线通信与网络和其他设备通信。上述无线通信可以使用任一通信标准或协议，包括但不限于全球移动通讯系统（Global System of Mobile communication，GSM）、通用分组无线服务（General Packet Radio Service，GPRS）、码分多址（Code Division Multiple Access，CDMA）、宽带码分多址
15 （Wideband Code Division Multiple Access，WCDMA）、长期演进（Long Term Evolution，LTE）、电子邮件、短消息服务（Short Messaging Service，SMS）等。

存储器 1020 可用于存储软件程序以及模块，处理器 1080 通过运行存储在存储器 1020 的软件程序以及模块，从而执行手机的各种功能应用以及数据处理。存储器 1020 可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序（比如声音播放功能、图像播放功能等）等；存储数据区可存储根据手机的使用所创建的数据（比如音频数据、电话本等）等。此外，存储器 1020
20 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。
25

输入单元 1030 可用于接收输入的数字或字符信息，以及产生与手

机的用户设置以及功能控制有关的键信号输入。具体地，输入单元 1030 可包括触控面板 1031 以及其他输入设备 1032。触控面板 1031，也称为触摸屏，可收集用户在其上或附近的触摸操作（比如用户使用手指、触笔等任何适合的物体或附件在触控面板 1031 上或在触控面板 1031 附近的操作），并根据预先设定的程式驱动相应的连接装置。在一些实例中，触控面板 1031 可包括触摸检测装置和触摸控制器两个部分。其中，触摸检测装置检测用户的触摸方位，并检测触摸操作带来的信号，将信号传送给触摸控制器；触摸控制器从触摸检测装置上接收触摸信息，并将它转换成触点坐标，再送给处理器 1080，并能接收处理器 1080 发来的命令并加以执行。此外，可以采用电阻式、电容式、红外线以及表面声波等多种类型实现触控面板 1031。除了触控面板 1031，输入单元 1030 还可以包括其他输入设备 1032。具体地，其他输入设备 1032 可以包括但不限于物理键盘、功能键（比如音量控制按键、开关按键等）、轨迹球、鼠标、操作杆等中的一种或多种。

显示单元 1040 可用于显示由用户输入的信息或提供给用户的信息以及手机的各种菜单。显示单元 1040 可包括显示面板 1041，在一些实例中，可以采用液晶显示器（Liquid Crystal Display, LCD）、有机发光二极管（Organic Light-Emitting Diode, OLED）等形式来配置显示面板 1041。进一步的，触控面板 1031 可覆盖显示面板 1041，当触控面板 1031 检测到在其上或附近的触摸操作后，传送给处理器 1080 以确定触摸事件的类型，随后处理器 1080 根据触摸事件的类型在显示面板 1041 上提供相应的视觉输出。虽然在图 10 中，触控面板 1031 与显示面板 1041 是作为两个独立的部件来实现手机的输入和输入功能，但是在某些实例中，可以将触控面板 1031 与显示面板 1041 集成而实现手机的输入和输出功能。

手机还可包括至少一种传感器 1050，比如光传感器、运动传感器以

及其他传感器。具体地，光传感器可包括环境光传感器及接近传感器，其中，环境光传感器可根据环境光线的明暗来调节显示面板 1041 的亮度，接近传感器可在手机移动到耳边时，关闭显示面板 1041 和/或背光。作为运动传感器的一种，加速计传感器可检测各个方向上（一般为三轴）加速度的大小，静止时可检测出重力的大小及方向，可用于识别手机姿态的应用（比如横竖屏切换、相关游戏、磁力计姿态校准）、振动识别相关功能（比如计步器、敲击）等；至于手机还可配置的陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器，在此不再赘述。

音频电路 1060、扬声器 1061，传声器 1062、麦克风 1063 可提供用户与手机之间的音频接口。音频电路 1060 可将接收到的音频数据转换后的电信号，传输到扬声器 1061，由扬声器 1061 转换为声音信号输出；另一方面，传声器 1062 将收集的声音信号转换为电信号，由音频电路 1060 接收后转换为音频数据，再将音频数据输出处理器 1080 处理后，经 RF 电路 1010 以发送给比如另一手机，或者将音频数据输出至存储器 1020 以便进一步处理。扬声器 1061 可以生成超声波信号，然后麦克风 1063 可以从手机周围采集超声波信号。

WiFi 属于短距离无线传输技术，手机通过 WiFi 模块 1070 可以帮助用户收发电子邮件、浏览网页和访问流式媒体等，它为用户提供了无线的宽带互联网访问。虽然图 10 示出了 WiFi 模块 1070，但是可以理解的是，其并不属于手机的必须构成，完全可以根据需要在不改变申请的本质的范围内而省略。

处理器 1080 是手机的控制中心，利用各种接口和线路连接整个手机的各个部分，通过运行或执行存储在存储器 1020 内的软件程序和/或模块，以及调用存储在存储器 1020 内的数据，执行手机的各种功能和处理数据，从而对手机进行整体监控。在一些实例中，处理器 1080 可包括一个或多个处理单元；优选的，处理器 1080 可集成应用处理器和

调制解调处理器，其中，应用处理器主要处理操作系统、用户界面和应用程序等，调制解调处理器主要处理无线通信。可以理解的是，上述调制解调处理器也可以不集成到处理器 1080 中。

手机还包括给各个部件供电的电源 1090（比如电池），优选的，电源可以通过电源管理系统与处理器 1080 逻辑相连，从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。

尽管未示出，手机还可以包括摄像头、蓝牙模块等，在此不再赘述。

在本申请实例中，该终端所包括的处理器 1080 还具有控制执行以上由终端执行的手势识别方法流程。例如，处理器 1080 对超声波的识别过程可以参阅前述实例中的描述。

接下来介绍本申请实例提供的另一种终端，请参阅图 11 所示，终端 1100 包括：

扬声器 1101、麦克风 1102、处理器 1103 和存储器 1104（其中终端 1100 中的处理器 1103 的数量可以一个或多个，图 11 中以一个处理器为例）。在本申请的一些实例中，扬声器 1101、麦克风 1102、处理器 1103 和存储器 1104 可通过总线或其它方式连接，其中，图 11 中以通过总线连接为例。

存储器 1104 可以包括只读存储器和随机存取存储器，并向处理器 1103 提供指令和数据。存储器 1104 的一部分还可以包括非易失性随机存取存储器（英文全称：Non-Volatile Random Access Memory，英文缩写：NVRAM）。存储器 1104 存储有操作系统和操作指令、可执行模块或者数据结构，或者它们的子集，或者它们的扩展集，其中，操作指令可包括各种操作指令，用于实现各种操作。操作系统可包括各种系统程序，用于实现各种基础业务以及处理基于硬件的任务。

处理器 1103 控制终端的操作，处理器 1103 还可以称为中央处理单元（英文全称：Central Processing Unit，英文简称：CPU）。具体的应用

中，终端的各个组件通过总线系统耦合在一起，其中总线系统除包括数据总线之外，还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见，在图中将各种总线都称为总线系统。

上述本申请实例揭示的方法可以应用于处理器 1103 中，或者由处理器 1103 实现。处理器 1103 可以是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器 1103 中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器 1103 可以是通用处理器、数字信号处理器（英文全称：digital signal processing，英文缩写：DSP）、专用集成电路（英文全称：Application Specific Integrated Circuit，英文缩写：ASIC）、现场可编程门阵列（英文全称：Field-Programmable Gate Array，英文缩写：FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本申请实例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本申请实例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器 1104，处理器 1103 读取存储器 1104 中的信息，结合其硬件完成上述方法的步骤。

所述扬声器 1101，用于在所述处理器的控制下播放第一超声波；

所述麦克风 1102，用于在所述处理器的控制下接收第二超声波；

本申请实例中，处理器 1103，用于执行所述存储器中的所述指令，执行如前述实例中的方法。

另外需说明的是，以上所描述的装置实例仅仅是示意性的，其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为

单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实例方案的目的。另外，本申请提供的装置实例附图中，模块之间的连接关系表示它们之间具有通信连接，具体
5 可以实现为一条或多条通信总线或信号线。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件的方式来实现，当然也可以通过专用硬件包括专用集成电路、专用 CPU、专用存储器、专用元器件等来
10 实现。一般情况下，凡由计算机程序完成的功能都可以很容易地用相应的硬件来实现，而且，用来实现同一功能的具体硬件结构也可以是多种多样的，例如模拟电路、数字电路或专用电路等。但是，对本申请而言更多情况下软件程序实现是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在可读取的存储介质中，如计
15 算机的软盘、U 盘、移动硬盘、只读存储器（ROM, Read-Only Memory）、随机存取存储器（RAM, Random Access Memory）、磁碟或者光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实例所述的方法。

20 综上所述，以上实例仅用以说明本申请的技术方案，而非对其限制；尽管参照上述实例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对上述各实例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实例技术方案的精神和范围。

权利要求书

1、一种手势识别方法，所述方法应用于终端，所述终端中配置有扬声器和麦克风，所述终端上安装有应用程序，所述方法包括：

5 根据所述应用程序的输入控制指示确定所述终端是否开启手势识别模式；

当所述终端确定开启手势识别模式时，通过所述扬声器播放第一超声波；

通过所述麦克风接收第二超声波，所述第二超声波为对所述第一超声波进行采集后得到的波形数据；

10 根据接收到的所述第二超声波进行手势识别，得到手势识别结果；
根据所述手势识别结果在所述应用程序执行相应的手势指令。

2、根据权利要求 1 所述的方法，其中，所述根据所述应用程序的输入控制指示确定所述终端是否开启手势识别模式，包括：

15 根据所述应用程序的配置文件确定所述应用程序的输入控制指示是否为默认开启手势识别模式；

检测所述应用程序在所述终端上是否运行成功；

在所述应用程序运行成功时，若所述终端默认开启手势识别模式，确定所述终端开启了手势识别模式。

20 3、根据权利要求 1 所述的方法，其中，所述根据所述应用程序的输入控制指示确定所述终端是否开启手势识别模式，包括：

根据所述应用程序的配置文件确定所述应用程序的输入控制指示为第一二维手势开启手势识别模式，或者第二二维手势关闭手势识别模式；

当检测到第一二维手势时，确定所述终端开启了手势识别模式；

25 当检测到第二二维手势时，确定所述终端关闭了手势识别模式。

4、根据权利要求 1 所述的方法，其中，所述方法还包括：

当所述扬声器播放训练超声波时,通过所述麦克风录制指定手势动作的波形数据;

根据所述指定手势动作的波形数据建立手势波形模板库;

根据所述手势波形模板库确定所述指定手势动作对应的手势识别

5 结果。

5、根据权利要求 1 所述的方法,其中,所述根据接收到的所述第二超声波进行手势识别,得到手势识别结果之后,所述方法还包括:

通过所述应用程序的显示界面显示手势识别成功的提示信息。

6、根据权利要求 1 所述的方法,其中,所述通过所述扬声器播放
10 第一超声波,包括:

通过所述扬声器播放 N 个频率的第一超声波,所述 N 为大于或等于 1 的正整数;

所述根据接收到的所述第二超声波进行手势识别,得到手势识别结果,包括:

15 当所述 N 的取值大于或等于 2 时,对 N 个所述第二超声波分别计算出手势的移动距离;

从 N 个所述手势的移动距离中排除出异常的移动距离,对于保留的移动距离进行一维度的手势识别,得到第一手势识别结果,所述第一手势识别结果包括:手势靠近所述终端,或者远离所述终端,或者相对于所述终端向左移动,或者相对于所述终端向右移动,或者相对于所述终端向上移动,或者相对于所述终端向下移动。
20

7、根据权利要求 1 所述的方法,其中,所述通过所述麦克风接收第二超声波,包括:

通过所述终端的两个麦克风分别采集到第二超声波;

25 所述根据接收到的所述第二超声波进行手势识别,得到手势识别结果,包括:

根据所述两个麦克风分别接收到的第二超声波计算手势的相对位置和初始位置；

根据计算出的相对位置和初始位置进行二维度的手势识别，得到第二手势识别结果，所述第二手势识别结果包括：手势变化的二维坐标。

- 5 8、一种终端，所述终端中配置有扬声器和麦克风，所述终端上安装有应用程序，所述终端包括：

模式确定模块，用于根据所述应用程序的输入控制指示确定所述终端是否开启手势识别模式；

- 10 超声波发送模块，用于当所述终端确定开启手势识别模式时，通过所述扬声器播放第一超声波；

超声波采集模块，用于通过所述麦克风接收第二超声波，所述第二超声波为对所述第一超声波进行采集后得到的波形数据；

手势识别模块，用于根据接收到的所述第二超声波进行手势识别，得到手势识别结果；

- 15 指令执行模块，用于根据所述手势识别结果在所述应用程序执行相应的手势指令。

9、一种终端，所述终端包括：扬声器、麦克风、处理器和存储器；所述处理器和所述扬声器、所述麦克风、所述存储器进行相互的通信；

- 20 所述存储器用于存储指令；

所述扬声器，用于在所述处理器的控制下播放第一超声波；

所述麦克风，用于在所述处理器的控制下接收第二超声波；

所述处理器，用于执行所述存储器中的所述指令，执行如权利要求1至7中任一项所述的方法。

- 25 10、一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行如权利要求1-7任意一项所述的方法。

1/8

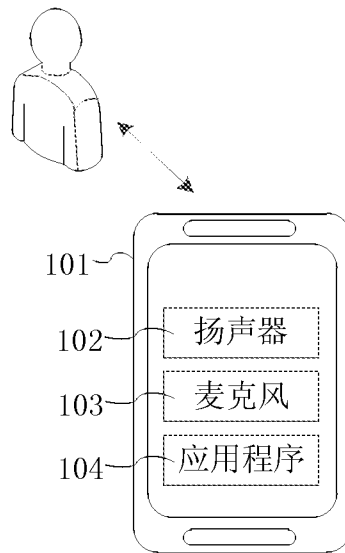


图 1a

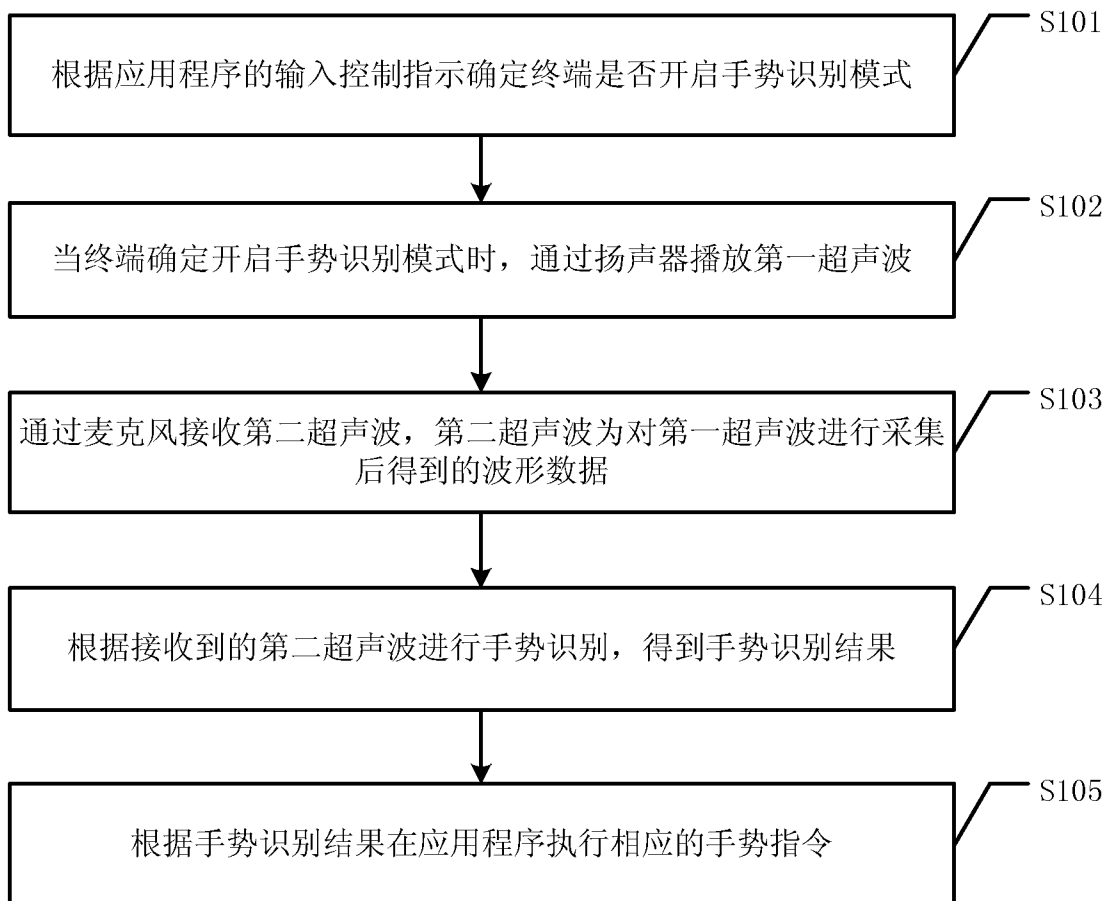


图 1b

2/8

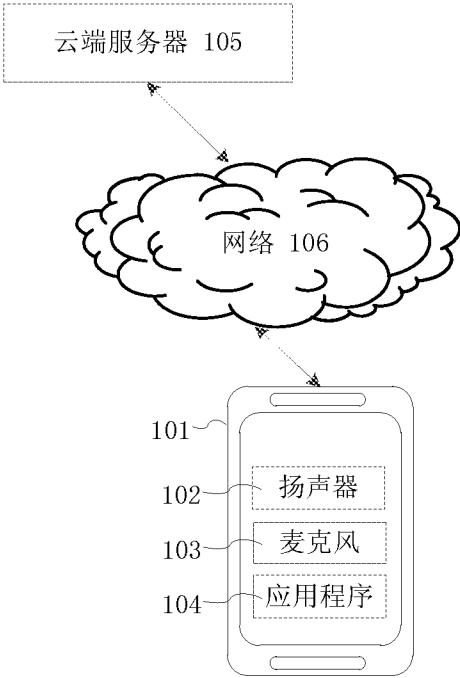


图 1c

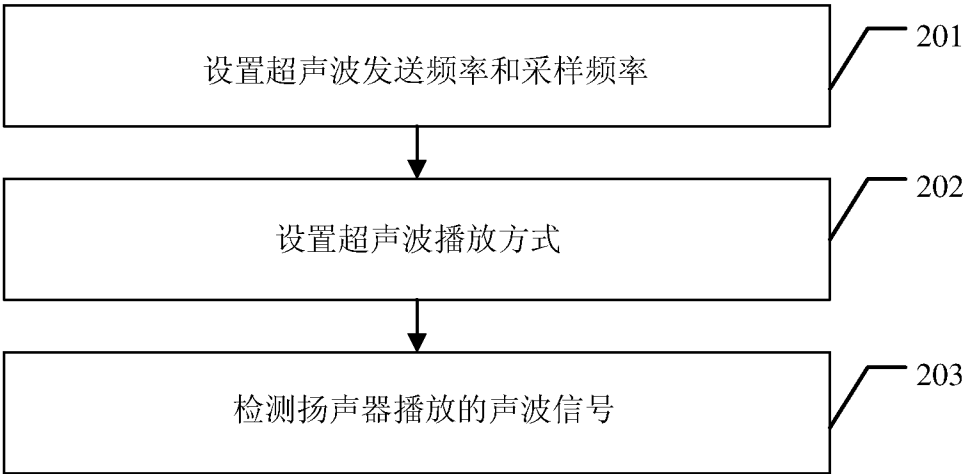


图 2

3/8

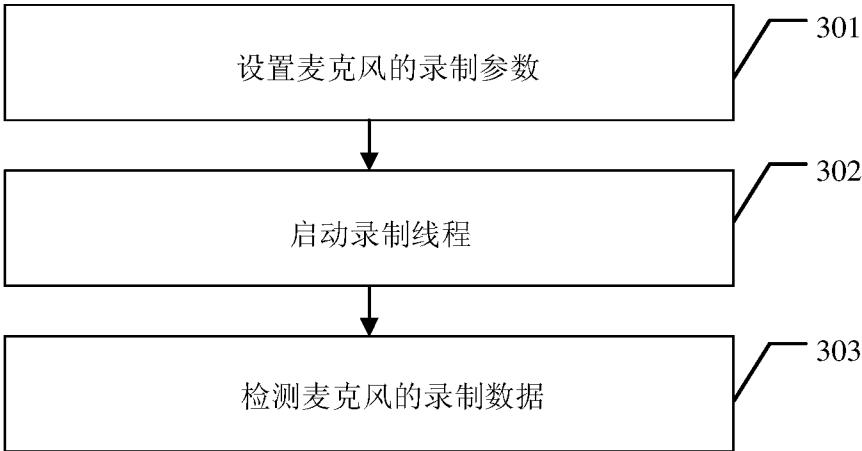


图 3

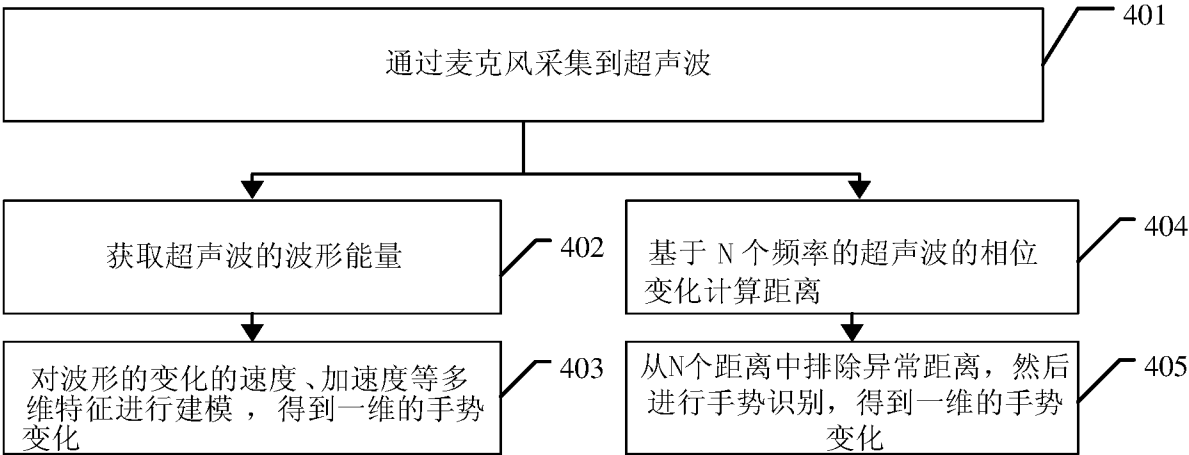


图 4

4/8

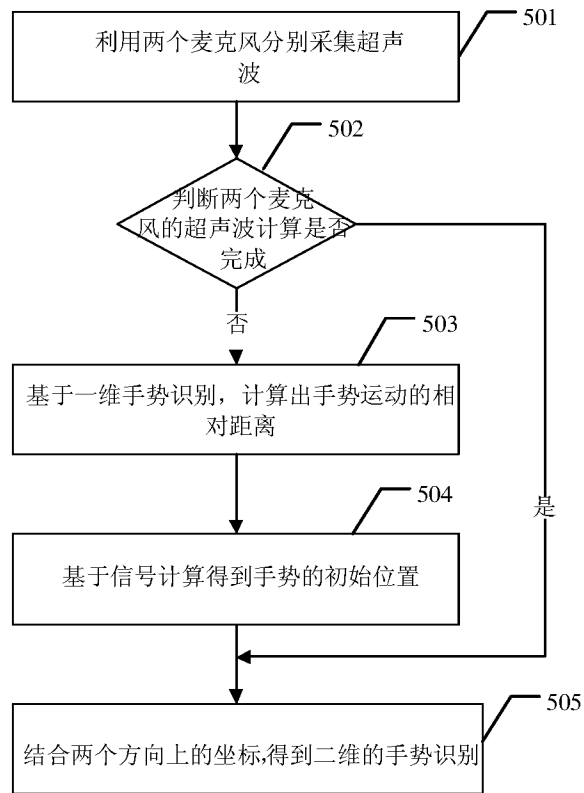


图 5

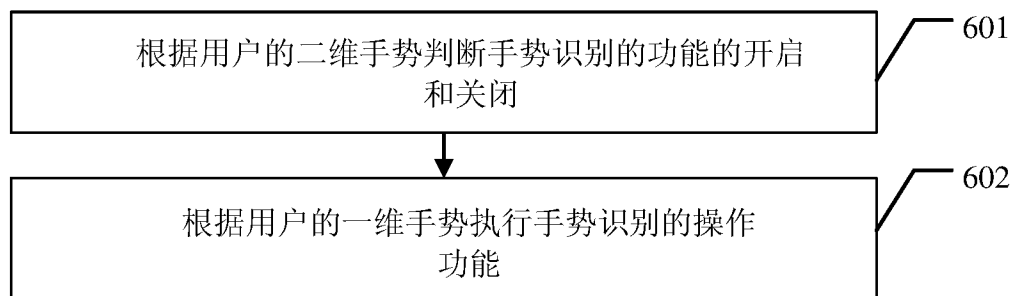


图 6

5/8

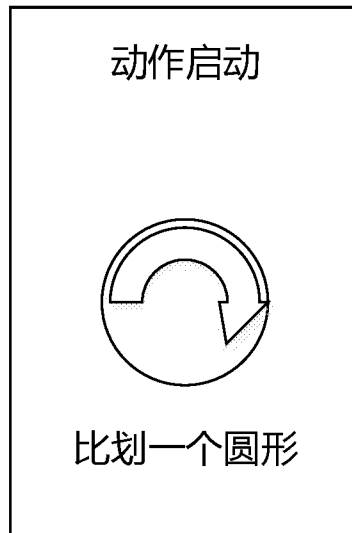


图 7-a

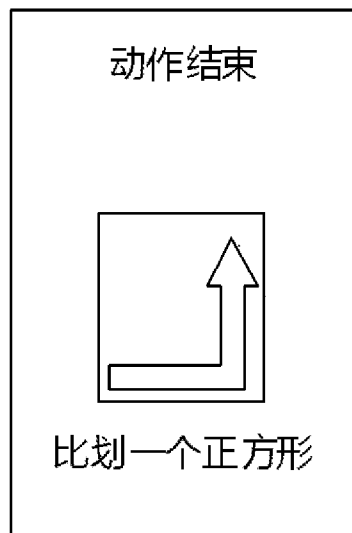


图 7-b

6/8

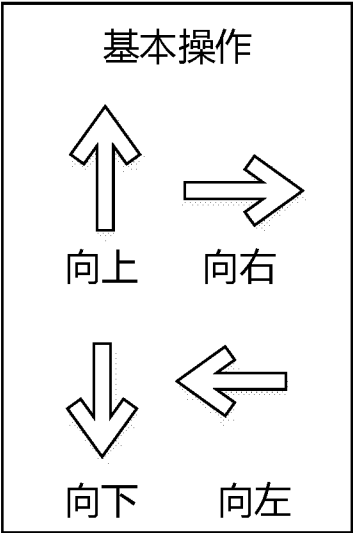


图 8

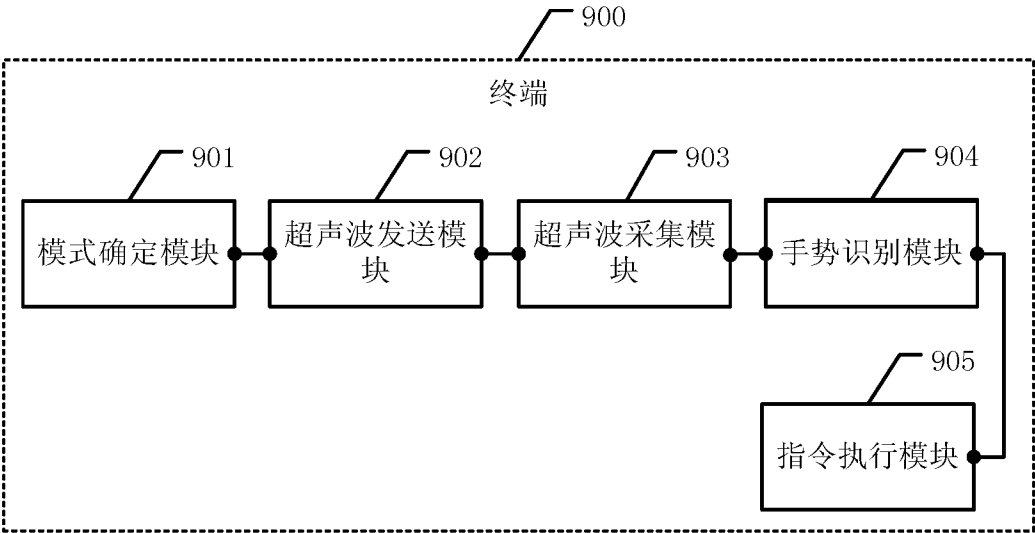


图 9-a

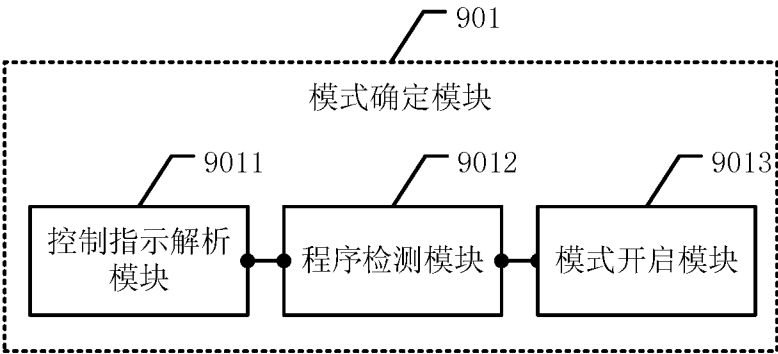


图 9-b

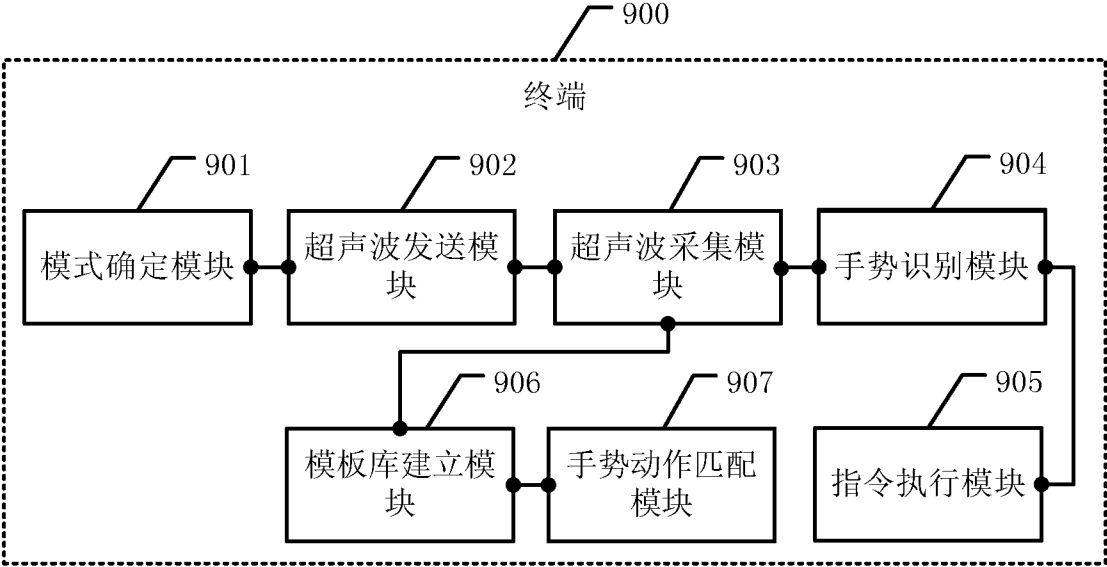


图 9-c

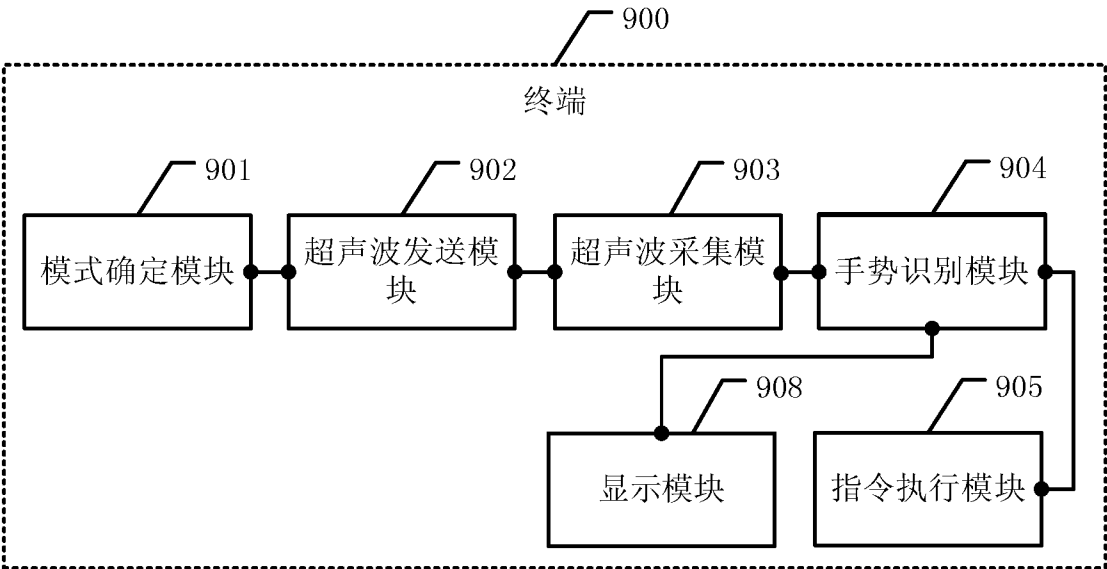


图 9-d

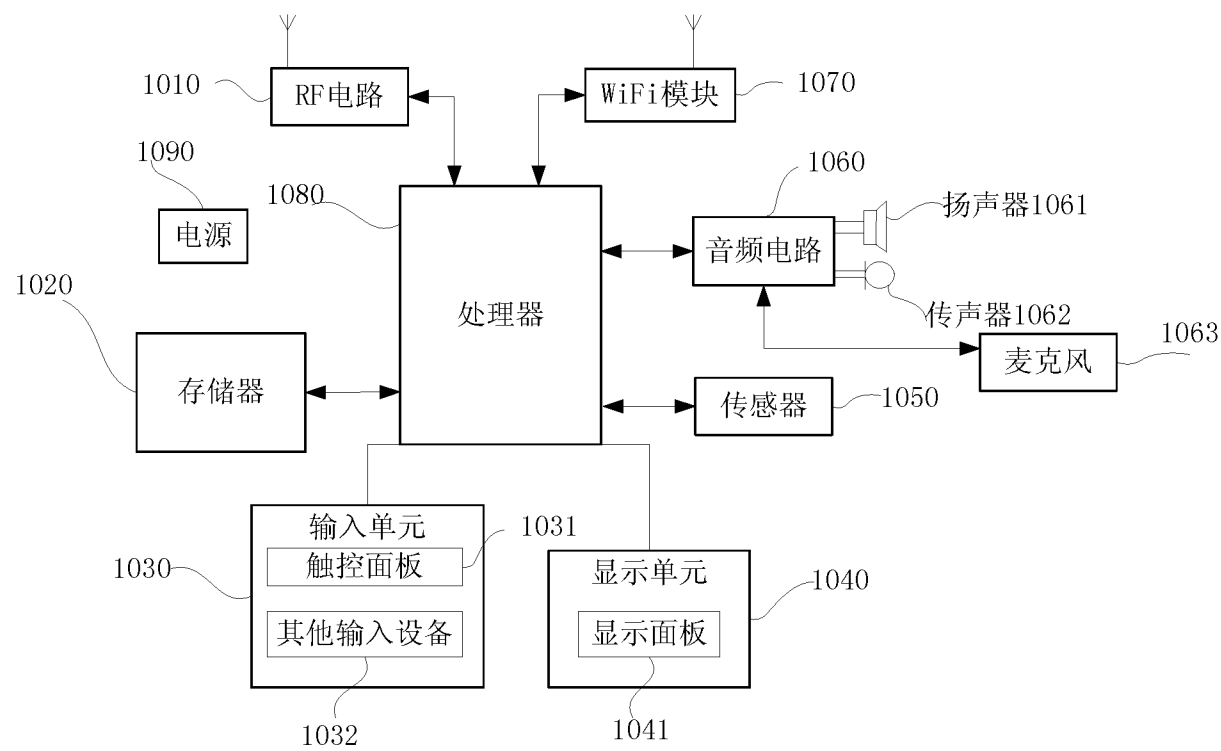


图 10

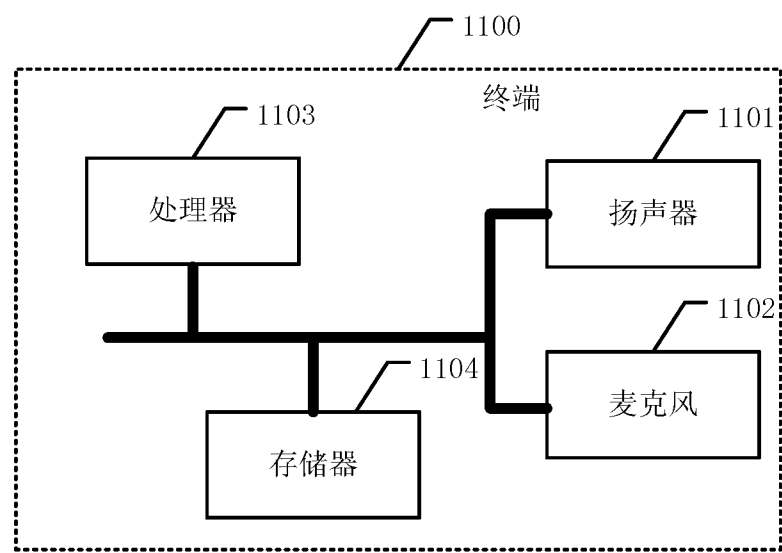


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/117864

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/01(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)

G06F3/,G01S

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, DWPI, SIPOABS: 手势, 识别, 超声, 扬声器, 麦克风, gesture, recogni+, identif+, ultrasound, hyperacoustic, ultrasonic, loudspeaker, speaker, microphone

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 105718064 A (NANJING UNIVERSITY) 29 June 2016 (2016-06-29) description, paragraphs [0030]-[0043]	1-10
X	CN 105807923 A (INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 27 July 2016 (2016-07-27) description, paragraphs [0039]-[0103]	1-10
A	CN 103577107 A (GUANGDONG OPPO MOBILE TELECOMMUNICATIONS CORP., LTD.) 12 February 2014 (2014-02-12) entire document	1-10

☐ 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

28 December 2018

Date of mailing of the international search report

16 January 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (ISA/CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/117864

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	105718064	A	29 June 2016	None			
CN	105807923	A	27 July 2016	WO	2017152531	A1	14 September 2017
CN	103577107	A	12 February 2014	None			

国际检索报告

国际申请号

PCT/CN2018/117864

A. 主题的分类

G06F 3/01 (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)

G06F3/, G01S

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, DWPI, SIPOABS: 手势, 识别, 超声, 扬声器, 麦克风, gesture, recogni+, identif+, ultrasound, hyperacoustic, ultrasonic, loudspeaker, speaker, microphone

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	CN 105718064 A (南京大学) 2016年 6月 29日 (2016 - 06 - 29) 说明书第[0030]-[0043]段	1-10
X	CN 105807923 A (中国科学院计算技术研究所) 2016年 7月 27日 (2016 - 07 - 27) 说明书第[0039]-[0103]段	1-10
A	CN 103577107 A (广东欧珀移动通信有限公司) 2014年 2月 12日 (2014 - 02 - 12) 全文	1-10

☐ 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

2018年 12月 28日

国际检索报告邮寄日期

2019年 1月 16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

传真号 (86-10) 62019451

受权官员

田丽娜

电话号码 86-(010)-62412072

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/117864

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)
CN	105718064	A	2016年 6月 29日	无			
CN	105807923	A	2016年 7月 27日	WO	2017152531	A1	2017年 9月 14日
CN	103577107	A	2014年 2月 12日	无			

表 PCT/ISA/210 (同族专利附件) (2015年1月)