(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
2018 年 7 月 12 日 (12.07.2018)

WIPO | PCT

(10) 国际公布号
WO 2018/126935 A1

(51) 国际专利分类号:
*G06F 3/0488* (2013.01)

(21) 国际申请号: PCT/CN2017/118429

(22) 国际申请日: 2017 年 12 月 26 日 (26.12.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710007951.5 2017年1月5日 (05.01.2017) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: 石善文 (SHI, Shanwen); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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,

(54) **Title:** VOICE-BASED INTERACTION METHOD AND APPARATUS, ELECTRONIC DEVICE, AND OPERATING SYSTEM

(54) **发明名称**: 基于语音的交互方法、装置、电子设备及操作系统

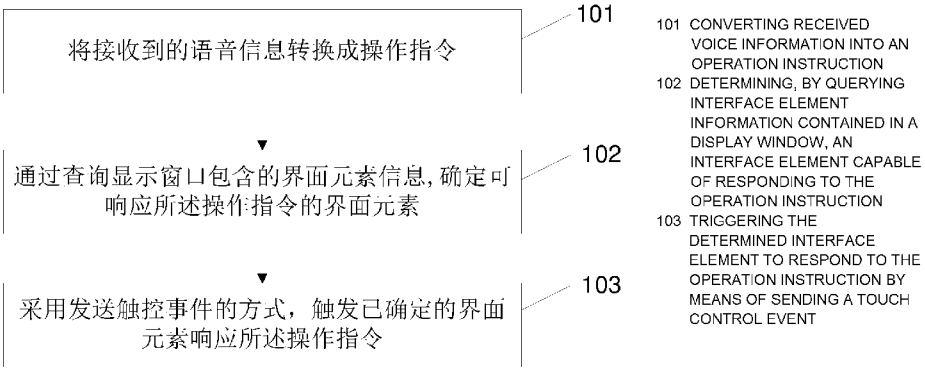



图 1

(57) **Abstract:** The present invention discloses a voice-based interaction method, a voice-based interaction apparatus, an electronic device, and an operating system. The voice-based interaction method comprises: converting received voice information into an operation instruction; determining, by querying interface element information contained in a display window, an interface element capable of responding to the operation instruction; and triggering the determined interface element to respond to the operation instruction by means of sending a touch control event. Since the means of sending a touch control event is used in the method, a voice operation is supported without a developer carrying out, on a program code, any adaptation in terms of the voice. On the one hand, the workload of the developer is reduced, and on the other hand, support for a voice interaction manner by an application window and a system window is realized, and popularization and application of the voice interaction manner in electronic devices is promoted.

(57) **摘要**: 本申请公开了一种基于语音的交互方法, 一种基于语音的交互装置, 一种电子设备, 以及一种操作系统。其中, 所述基于语音的交互方法, 包括: 将接收到的语音信息转换成操作指令; 通过查询显示窗口包含的界面元素信息, 确定可响应所述操作指令的界面元素; 采用发送触控事件的方式, 触发已确定的界面元素响应所述操作指令。上述方法由于采用了发送触控事件的方式, 开发人员无需对程序代码作任何语音方面的适配工作, 就可以支持语音操作, 一方面大大减少了开发人员的工作量, 另一方面, 也实现了应用窗口以及系统窗口对语音交互方式的支持, 促进语音交互方式在电子设备上的推广应用。

WO 2018/126935 A1

[见续页]

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))。

# 基于语音的交互方法、装置、电子设备及操作系统

本申请要求2017年01月05日递交的申请号为201710007951.5、发明名称为“基于语音的交互方法、装置、电子设备及操作系统”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。

## 技术领域

本申请涉及交互方法，具体涉及一种基于语音的交互方法。本申请同时涉及一种基于语音的交互装置，一种电子设备，以及一种操作系统。

## 背景技术

语音是人类惯于使用的交流方式，与其他交流手段相比不仅更加自然，而且具有认知负荷小，资源占用少和交互效率高等优点。随着 HMM 模型和人工神经元网络等技术在语音识别中的成功应用，语音作为一种强大随意的控制入口，目前已经在个人电脑、移动通讯终端等各种电子设备中得到了广泛应用，用户通过录入语音，就可以在电子设备上执行所需的查询、搜索、拨打电话等操作，方便用户的使用。

现有的语音交互方式，通常需要对电子设备中支持语音操作的应用程序进行定制，语音交互过程包括以下处理环节：经过定制的应用程序显示支持语音操作的窗口后，会向操作系统提供的语音服务注册该窗口所支持的语音指令集合；当语音服务接收用户输入的语音信息后，若检测出语音信息与该应用程序已注册的语音指令集中的某语音指令相符，则将语音信息转换成相应的控制指令发送给该应用程序的相应窗口，应用程序通过预先定制的代码进行响应；当应用程序退出支持语音操作的窗口时，则向语音服务注销相应的语音指令集合。

由此可见，现有的语音交互实现方式，若需要应用程序的某窗口支持语音交互，那么需要添加至少以下三方面的操作逻辑：注册语音指令集合、注销语音指令集合、根据接收到的控制指令进行相应的处理。如果应用程序中的两个或者两个以上窗口都有支持语音操作的需求，那么开发人员则需要针对各窗口执行上述定制操作，工作量通常很大，导致了现在的应用程序难以对语音操作提供完善的语音操作支持，阻碍语音交互方式在电子设备上的推广应用。

## 发明内容

本申请实施例提供一种基于语音的交互方法和装置，以解决现有技术需要对支持语音的应用窗口逐一进行定制导致的、开发人员工作量大以及应用程序难以提供完善的语音支持的问题。本申请实施例还提供一种基于语音的交互装置，一种电子设备，以及一种操作系统。

本申请提供一种基于语音的交互方法，包括：

将接收到的语音信息转换成操作指令；

通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；

采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

可选的，所述界面元素包括：可供操作的控件；所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息。

可选的，所述采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令，包括：

根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

可选的，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，包括针对每个已确定的界面元素执行下述操作：

至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

可选的，所述界面元素的布局信息包括：界面元素在显示屏幕中的位置和界面元素的尺寸信息。

可选的，所述手动输入设备包括：触摸屏，或者鼠标。

可选的，在所述将接收到的语音信息转换成操作指令之前，包括：

通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述查询显示窗口包含的界面元素信息，包括：

查询基于上述监控方式记录的显示窗口包含的界面元素信息。

可选的，所述将接收到的语音信息转换成操作指令，包括：

通过语音识别将接收到的语音信息转换成文字形式的指令；

通过语义分析和基于模板的匹配操作，将所述文字形式的指令转换成操作指令。

可选的，所述操作指令的内容包括：被操作界面元素的类型、和操作描述信息。

可选的，所述显示窗口包括：预设类型的显示窗口；所述预设类型的显示窗口包括：应用窗口、和/或系统窗口。

可选的，所述预设类型的显示窗口包括：预设类型的顶层显示窗口。

可选的，所述预设类型的顶层显示窗口包括：顶层应用窗口；所述顶层应用窗口的数量为 1；

所述通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素，包括：通过查询顶层应用窗口包含的界面元素信息，确定所述顶层应用窗口中、可响应所述操作指令的界面元素。

可选的，所述显示窗口的数量大于等于 1；

所述通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素，包括：

遍历各显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

可选的，若确定的可响应所述操作指令的界面元素的数目大于或者等于 2，所述方法还包括：

按照预设方式，从所述确定的界面元素中选择一个界面元素；

所述触发已确定的界面元素响应所述操作指令，包括：触发所选界面元素响应所述操作指令。

可选的，所述方法在具有语音输入装置及显示屏幕的电子设备上实施，所述电子设备包括：移动通讯设备、个人电脑、平板电脑、或应用于互联网汽车的车载电子设备。

相应的，本申请还提供一种基于语音的交互装置，包括：

操作指令生成单元，用于将接收到的语音信息转换成操作指令；

界面元素确定单元，用于通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；

触发响应单元，用于采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

可选的，所述触发响应单元，具体用于根据所述操作指令，向已确定的界面元素所

属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

可选的，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述触发响应单元，包括针对每个已确定的界面元素触发的下列子单元：

触控位置确定子单元，用于至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟触控事件上报子单元，用于模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

触控事件发送子单元，用于由窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

可选的，所述装置还包括：

监控记录单元，用于在将接收到的语音信息转换成操作指令之前，通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述界面元素确定单元，具体用于查询所述监控记录单元记录的显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

此外，本申请还提供一种电子设备，包括：

语音输入装置，用于接收输入的语音信息；

显示器；

处理器；

存储器，用于存储指令；

其中，所述处理器耦合于所述存储器，用于读取所述存储器存储的指令，并执行如下操作：

将接收到的语音信息转换成操作指令；通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

可选的，所述采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令，包括：根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

可选的，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，包括针对每个已确定的界面元素执行下述操作：

至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

可选的，在所述将接收到的语音信息转换成操作指令之前，包括：通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述查询显示窗口包含的界面元素信息，包括：查询基于上述监控方式记录的显示窗口包含的界面元素信息。

此外，本申请还提供一种操作系统，包括：

操作指令生成单元，用于将接收到的语音信息转换成操作指令；

界面元素确定单元，用于通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；

触控事件发送单元，用于发送触控事件，以触发已确定的界面元素响应所述操作指令。

可选的，所述触控事件发送单元，具体用于根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

可选的，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述触控事件发送单元，包括针对每个已确定的界面元素触发的下列子单元：

触控位置确定子单元，用于至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟触控事件上报子单元，用于模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

触控事件发送子单元，用于由窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事

件。

可选的，所述操作系统还包括：

监控记录单元，用于在将接收到的语音信息转换成操作指令之前，通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述界面元素确定单元，具体用于查询所述监控记录单元记录的显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

与现有技术相比，本申请具有以下优点：

本申请提供的基于语音的交互方法，将接收到的语音信息转换成操作指令后，通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素，随后采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

本申请提供的方法，不是针对某一应用程序或者应用程序的某一窗口提供语音支持，而是从系统层面提供语音支持，通过模拟用户触控事件的方式，触发界面元素响应基于语音的操作指令，从而实现基于语音的交互过程。由于采用了发送触控事件的方式，因此触控事件的接收方（例如：窗口）按照常规的触控事件处理逻辑进行响应即可，也即，开发人员无需对程序代码作任何语音方面的适配工作，就可以支持语音操作，一方面大大减少了开发人员的工作量，另一方面，也实现了应用窗口以及系统窗口对语音交互方式的支持，促进语音交互方式在电子设备上的推广应用。

## 附图说明

图 1 是本申请的一种基于语音的交互方法的实施例的流程图；

图 2 是本申请实施例提供的模拟手动输入设备上报触控事件的处理流程图；

图 3 是本申请实施例提供的一个具体实施例子的框架图；

图 4 是本申请的一种基于语音的交互装置的实施例的示意图；

图 5 是本申请的一种电子设备的实施例的示意图；

图 6 是本申请的一种操作系统的实施例的示意图。

## 具体实施方式

在下面的描述中阐述了很多具体细节以便于充分理解本申请。但是，本申请能够以很多不同于在此描述的其它方式来实施，本领域技术人员可以在不违背本申请内涵的情

况下做类似推广，因此，本申请不受下面公开的具体实施的限制。

在本申请中，分别提供了一种基于语音的交互方法、一种基于语音的交互装置、一种电子设备、以及一种操作系统。在下面的实施例中逐一进行详细说明。

请参考图 1，其为本申请的一种基于语音的交互方法的实施例的流程图。为了便于理解，先对本申请提供的技术方案、及实施方式作一个简要说明，再分步骤对具体的实施方式进行描述。

本技术方案，将接收到的语音信息转换成操作指令后，通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素，并采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。由于本技术方案采用了发送触控事件的方式，因此触控事件的接收方按照常规的触控事件处理逻辑进行响应即可，从而开发人员无需针对程序代码作任何语音方面的适配工作，在对语音交互方式提供完善支持的同时，减少了开发人员的工作量，可以促进语音交互方式在电子设备上的推广应用。本技术方案，通常可以集成在操作系统中，从系统层面对应用程序窗口（简称应用窗口）和/或系统程序窗口（简称系统窗口）提供语音支持。

具体实施时，可以根据需求针对预设类型的显示窗口实施本技术方案。例如，所述预设类型的显示窗口可以为：应用窗口。所述预设类型的显示窗口也可以包括：应用窗口和/或系统窗口，所述预设类型的显示窗口的数量可以为一个或者一个以上。以智能移动通讯设备为例，若预设类型包括应用窗口和系统窗口这两种窗口类型，那么在实施本技术方案时，可以遍历当前显示的状态栏、导航栏（状态栏和导航栏都属于系统窗口）、以及应用窗口包含的界面元素信息，并针对其中可响应操作指令的界面元素发送触控事件。

所述显示窗口可以为：顶层显示窗口，例如，在界面上弹出的对话框属于顶层显示窗口，在智能移动通讯设备的界面上显示的状态栏和/或导航栏也属于顶层显示窗口，考虑到用户通常针对顶层显示窗口进行语音交互，本技术方案可以仅针对顶层显示窗口实施，即：所述查询显示窗口包含的界面元素信息的步骤可以为，查询顶层显示窗口包含的界面元素信息。

在具体实施时，上述针对顶层显示窗口的实施方式也可以与针对预设类型显示窗口的实施方式结合，在这种情况下，所述查询显示窗口包含的界面元素信息的步骤可以为，查询预设类型的顶层显示窗口包含的界面元素。预设类型的顶层显示窗口可以包括：顶层应用窗口、和/或顶层系统窗口。

以智能移动通讯设备为例，当预设类型的顶层显示窗口为顶层应用窗口时，由于顶层应用窗口通常为一个，在这种情况下，在实施本技术方案时，仅需要查询顶层应用窗口所包含的界面元素信息，确定所述顶层应用窗口中、可响应所述操作指令的界面元素，并采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

具体实施时，查询显示窗口包含的界面元素信息的步骤，可以基于窗口管理服务（Window Manager Service — WMS）实现。在电子设备采用的操作系统中，通常都提供窗口管理服务，应用程序或者系统程序可以通过与窗口管理服务进行交互，从而将窗口内容显示到屏幕上，以 Android 操作系统为例，窗口管理服务是操作系统的框架层（framework）中的组成部分。因此，在实施本技术方案时，可以向窗口管理服务查询显示窗口包含的界面元素信息；也可以采用监控窗口管理服务从而记录显示窗口包含的界面元素信息的实施方式。其中，后者更具有普遍适用性，并且可以减少与窗口管理服务的耦合，是以下提供的具体实施过程所采用的优选实施方式。

在下面提供的实施例中，以在采用 iOS、Android 或者 YunOS 等智能操作系统的电子设备中、实施本申请提供的基于语音的交互方法（以下简称本方法）为例，对具体实施方式进行说明，其中，显示窗口为顶层应用窗口，且顶层应用窗口的数量为一个。

请参见图 1，其为本申请提供的一种基于语音的交互方法的实施例的示意图。在本实施例中，在执行步骤 101 之前，实施了本方法的语音服务模块可以通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录顶层应用窗口包含的界面元素信息。

具体的，应用程序以及系统程序在显示某一窗口或者退出某一窗口时，通常会导致显示窗口发生变更，而该变更过程通常会由窗口管理服务进行相应的处理，才能向用户呈现变更后的显示效果。因此可以通过监控该过程，记录顶层应用窗口包含的界面元素信息。

在应用程序希望显示某一应用窗口时（即：使该窗口成为顶层应用窗口），通过向窗口管理服务发送请求，触发窗口管理服务调用相应的窗口显示处理函数，在该函数中可以添加监控代码，向语音服务模块发送消息，同时在消息中携带顶层应用窗口所包含的界面元素信息，从而语音服务模块就可以记录顶层应用窗口包含的界面元素信息。

所述界面元素，通常是指可供操作的控件，例如：按钮、列表、文本输入框等。所述界面元素信息，通常包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息。其中，标识界面元素的信息可以是界面元素的布局信息，包括：在显示屏幕中的位置、和尺寸信息等，也可以是界面元素的标识，根据发送触控事件的具

体方式不同，可以记录相应的标识界面元素的信息。

以列表控件为例，其类型可以为对应“列表”这一类型的预设值；对应于该类型的预设属性可以包括：列表内容的项数、和/或、每一项的具体内容等；标识列表控件的信息可以是在屏幕中的显示位置和尺寸信息等，也可以是列表控件的 ID。以按钮控件为例，其类型可以为对应“按钮”这一类型的预设值；对应于该类型的预设属性可以包括：按钮的标题内容；标识按钮控件的信息可以是在屏幕中的显示位置和尺寸信息等，也可以是按钮控件的 ID。

在上述记录顶层应用窗口包含的界面元素信息的基础上，下面对本实施例包括的步骤 101-103 的实施方式进行描述。

步骤 101、将接收到的语音信息转换成操作指令。

用户通过麦克风等语音输入装置输入语音信息，本步骤则将接收到的语音信息转换成操作指令。具体的，针对用户输入的语音信息，可以先通过语音识别将接收到的语音信息转换成文字形式的指令；然后再通过语义分析和基于模板的匹配操作，将所述文字形式的指令转换成操作指令。所述操作指令的内容可以包括：被操作界面元素的类型，和操作描述信息。其中操作描述信息的含义，根据界面元素类型的不同，通常对应于界面元素的不同属性。

例如，用户通过麦克风输入以下语句：“你好 xx，选择第三个。”本步骤则可以针对接收到的语音信息进行语音识别，识别出“你好 xx”为预设的唤醒词，并通过语音识别过程提取出文本形式的指令：“选择第三个”。随后可以通过语义分析将指令信息解析为“选择”和“第三个”，然后通过与预定义的模板进行匹配，将该指令信息翻译为类似“被操作界面元素类型：列表，操作描述信息：3”这样的操作指令。

再如，用户通过麦克风输入以下语句：“我选确定”，本步骤通过上述类似的语音识别、语义分析和基于模板的匹配过程，可以将接收到的语音信息转换为以下操作指令：“被操作界面元素类型：按钮，操作描述信息：确定”。

步骤 102、通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

在本实施例中，本步骤可以通过查询在步骤 101 之前基于监控方式记录的、顶层应用窗口包含的界面元素信息，确定可响应步骤 101 得到的操作指令的界面元素。

具体的，可以查询通过监控方式记录的顶层应用窗口包含的界面元素信息，查看其中是否存在与被操作界面元素类型和操作描述信息相符合的界面元素，若存在，所述相

符合的界面元素即为可响应所述操作指令的界面元素。

例如，操作指令为“被操作界面元素类型：列表，操作描述信息：3”，本步骤则可以查询在已记录的顶层应用窗口包含的界面元素信息中，是否存在类型为列表、并且存在第三条列表内容的界面元素，若存在，那么相应的列表控件即为可响应所述操作指令的界面元素。

再如，操作指令为“被操作界面元素类型：按钮，操作描述信息：确定”，本步骤则可以查询在已记录的顶层应用窗口包含的界面元素信息中，是否存在类型为按钮，并且标题为“确定”的界面元素，若存在，那么相应的按钮控件即为可响应所述操作指令的界面元素。

在应用于智能操作系统的移动通讯设备中、且仅针对数量唯一的顶层应用窗口提供语音支持的场景下，可以采用上面描述的方式确定可响应操作指令的界面元素。在其他实施方式中，如果希望对当前显示的各顶层显示窗口都提供语音支持（例如：状态栏、顶层应用窗口、导航栏即为 3 个顶层显示窗口），那么本步骤可以采用窗口遍历的方式，即：遍历通过监控方式记录的各顶层显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

在具体实施时，如果通过查询操作，没有找到可响应操作指令的界面元素，那么可以通过弹出对话框或者语音播报等方式，向用户提供相应的反馈信息，从而用户可以重新输入语音信息或者采用其他交互方式。

步骤 103、采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

步骤 102 确定的可响应操作指令的界面元素通常为 1 个，本步骤针对所述界面元素，采用发送触控事件的方式，触发所述界面元素响应操作指令。例如：可以向所述界面元素所属窗口或程序，或者是负责控制管理所述界面元素的线程或者进程，发送针对所述界面元素的触控事件，从而触发所述界面元素响应操作指令。

具体到本实施例中，本步骤可以根据操作指令，向所述界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述窗口执行相应的处理操作。

具体的，可以采用模拟手动输入设备上报触控事件的方式向所述界面元素所属的顶层应用窗口发送触控事件，从而触发顶层应用窗口执行相应的处理操作。需要说明的是，如前所述，预先记录的顶层应用窗口包含的每个界面元素信息，包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息，本步骤若采用模拟手动输入设备上报触控事件的实施方式，记录的所述标识界面元素的信息为界面元素的布局信息，

即：界面元素在显示屏幕中的位置和界面元素的尺寸信息等。

针对步骤 102 所确定的可响应操作指令的界面元素，模拟手动输入设备上报触控事件的处理过程，包括如下所示的步骤 103-1 至 103-3，下面结合图 2 进行详细说明。

步骤 103-1、至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置。

例如，所述界面元素为按钮，根据按钮在显示屏幕中的位置和按钮的尺寸信息，可以确定按钮对应于显示屏幕中的显示区域，那么可以选择显示区域中的任意一个点的位置作为触控位置，例如，可以选择显示区域的中心位置作为对应于按钮的触控位置。

再如，所述界面元素为列表，在确定触控位置时，不仅要根据列表的布局信息，还要根据操作指令中的操作描述信息，来共同确定触控位置。例如：操作指令中的操作描述信息为：3，即：被模拟的触控事件应该针对第 3 条列表内容，因此可以根据列表的布局信息，确定其中第 3 条列表内容的显示区域，并从中选择一个位置作为对应于列表的触控位置。

步骤 103-2、模拟手动输入设备向操作系统上报针对所述触控位置的触控事件。

在步骤 103-1 中已经确定了对应于所述界面元素的触控位置，本步骤则模拟手动输入设备向操作系统上报针对所述触控位置的触控事件。

通常情况下，用户通过手指或者鼠标对显示屏幕的触控操作，通常会引发中断，并向文件系统中的相应设备节点写入相应的触控事件，该触控事件通常会上报给操作系统，并分发给窗口管理服务，并由窗口管理服务分发给相应的窗口进行处理。

在本实施例中，并没有用户实施触控操作，而是将已获取的操作指令转化为用户的触控操作。因此本步骤可以基于上述触控操作处理过程的基本原理，通过软件指令向手动输入设备对应的设备节点发送触控事件，以模拟用户实施触控操作的效果。

以基于 linux 内核的 android 操作系统为例，可以利用 Android 操作系统提供的 sendevent 命令向手动输入设备对应的设备节点发送触控事件，例如，可以向设备节点发送以下指令：先发送指定屏幕触控位置的 sendevent 指令，其中携带由步骤 103-1 确定的触控位置对应的坐标值，然后发送对应于触控操作中的按下动作（即通常所述的 down 动作）的 sendevent 命令，最后发送对应于触控操作中的抬起动作（即通常所述的 up 动作）的 sendevent 命令。通过发送上述指令，完整模拟了针对所述触控位置的一次触控操作。

其中，所述手动输入设备包括：触摸屏，或者鼠标。例如，对于个人电脑等没有触

摸屏的电子设备，所述手动输入设备可以为鼠标，即：模拟用户使用鼠标点击的操作；对于移动通讯设备（例如：智能手机）等具有触摸屏的电子设备，所述手动输入设备可以为触摸屏，即：模拟用户用手指触摸屏幕的操作。

步骤 103-3、窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

步骤 103-1 确定了对应于所述界面元素的触控位置，步骤 103-2 模拟了针对所述触控位置的触控事件，本步骤中，在操作系统将接收到触控事件分发给窗口管理服务后，窗口管理服务根据接收到的触控事件对应的触控位置，将接收到的触控事件转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送针对所述界面元素的触控事件，从而触发所述窗口执行相应的处理操作。

例如：步骤 101 获取的操作指令为“被操作界面元素类型：列表，操作描述信息：3”，本步骤中，窗口管理服务接收到针对（x,y）位置的触控事件后将其转换成针对顶层应用窗口中的列表控件第三条列表内容的触控事件，并向顶层应用窗口发送该触控事件，以触发顶层应用窗口执行相应的处理操作，即：触发顶层应用窗口执行针对所述列表控件的触控事件处理函数。例如，顶层应用窗口为乐曲播放窗口，其中包含乐曲列表，顶层应用窗口根据接收到的触控事件执行相应的触控事件处理函数，播放乐曲列表中的第三个曲目，从而实现了基于语音的交互过程。

以上通过步骤 103-1 至 103-3，对模拟底层触控事件的实施方式进行了描述。在具体实施时，在系统架构支持的情况下，本步骤 103 也可以采用直接向顶层应用窗口发送触控消息的实施方式。具体的，在步骤 102 确定可响应操作指令的某一界面元素后，本步骤可以直接向顶层应用窗口发送表征触控事件的触控消息，所述触控消息中包含以下信息：所述界面元素的标识（例如：按钮 ID），或者所述界面元素标识和操作指令中的操作描述信息（例如：列表 ID 和 3，指明相应列表中的第三条列表内容），顶层应用窗口根据接收到的触控消息，执行针对相应界面元素的触控事件处理函数，从而对输入的语音信息做出响应。采用这种实施方式，在步骤 101 之前记录顶层应用窗口包含的界面元素信息时，需要记录界面元素标识，即：界面元素信息中要包含界面元素标识。

如果步骤 102 确定的可响应操作指令的界面元素为 1 个，在这种情况下，可以采用上述两种实施方式（模拟底层触控事件和直接发送触控消息）中的任意一种进行处理。如果步骤 102 确定的界面元素的数目大于或者等于 2，那么本步骤可以针对每个界面元素，采用上述两种方式中的任意一种向其所属窗口发送触控事件，以触发相应窗口执行

相应的处理操作；也可以按照预设方式，从步骤 102 确定的界面元素中选择一个界面元素，然后采用上述两种方式中的任意一种向所选界面元素所属窗口发送触控事件。

其中，按照预设方式选择一个界面元素，可以是根据已确定的各界面元素的布局，选择处于预设方位的一个界面元素，例如：选择其中处于最上方的界面元素，或者选择其中处于最左侧的界面元素。

按照预设方式选择一个界面元素，也可以是采用语音播报方式提醒用户选择一个界面元素，并根据对用户输入语音进行识别的结果，从已确定的各界面元素中选择一个界面元素。

至此，通过上述步骤 101-103，对本实施例提供的基于语音的交互方法的实施方式进行了详细说明。在具体实施时，上述实施例提供的基于语音的交互方法，可以作为运行在操作系统中的一个语音服务模块实施。为了便于维护，也可以采用划分为多个功能模块的方式实施，下面给出一个具体的例子。在其他实施方式中，每个功能模块实现的功能可以不同于本例子，也可以采用不同于本例子的其他模块划分方式。

请参见图 3，在本例子中，语音服务功能由语音管理模块，窗口监控模块，语音引擎模块共同完成。其中，窗口监控模块通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录顶层应用窗口包含的界面元素信息；语音引擎模块通过语音识别将接收到的语音信息转换成文字形式的指令，并发送给语音管理模块，语音管理模块通过语义分析以及匹配等处理过程，将文字形式的指令转换成操作指令，并向窗口监控模块发送携带操作指令的操作请求；窗口监控模块查询已记录的顶层应用窗口包含的界面元素信息，确定可响应所述操作指令的界面元素，并根据所述操作指令，通过窗口管理服务向顶层应用窗口发送针对所述界面元素的触控事件，以触发顶层应用窗口执行相应的处理操作。

需要说明的是，上面提供的实施方式，是在顶层应用窗口只有一个的应用场景下，对顶层应用窗口提供语音支持的实施方式，在其他实施方式中，可以基于相同的原理对本实施例所述的实施方式进行扩展。例如：可以对顶层系统窗口提供语音支持，或者对同时显示的多个顶层显示窗口提供语音支持等。对于前者，可以预先记录顶层系统窗口所包含的界面元素信息，在将接收到的语音信息转换成操作指令后，针对顶层系统窗口进行查询，确定可响应所述操作指令的界面元素并发送相应的触控事件；对于后者，可以预先记录所有顶层显示窗口所包含的界面元素信息，将接收到的语音信息转换成操作指令后，遍历各顶层显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元

素，并发送相应的触控事件。

综上所述，本实施例提供的基于语音的交互方法，在将接收到的语音信息转换成操作指令后，通过发送触控事件的方式，触发界面元素响应操作指令，实现基于语音的交互过程。由于本方法采用了发送触控事件的方式，因此触控事件的接收方按照常规的触控事件处理逻辑进行响应即可，也即，开发人员无需作任何语音方面的代码适配工作，就可以支持语音操作，减少了开发人员的工作量，促进语音交互方式在电子设备上的推广应用。

需要说明的是，本实施例提供的基于语音的交互方法，可以应用于具有语音输入装置及显示屏幕的电子设备上，所述电子设备包括：移动通讯设备（例如：智能手机）、个人电脑、平板电脑、或应用于互联网汽车的车载电子设备。

特别是，由于基于互联网汽车的应用场景更加强调安全性，在互联网汽车的车载电子设备上实施本实施例提供的方法，有益效果更为显著。

例如， 以在安装有 YunOS 智能操作系统的、具用导航功能的车载电子设备上实施所述方法为例，用户在驾车的过程中启用导航功能后，显示屏幕上显示导航目的地列表，用户可以直接说以下语音指示信息“我选第 5 项”，或者“我要去 xx 地”，那么本实施例提供的方法就可以将接收到的语音信息转换成操作指令：“被操作界面元素类型：列表，操作描述信息：5”，然后通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素为：导航目的地列表，并采用发送触控事件的方式，模拟用户触控导航目的地列表中的第 5 项的操作，从而启动用户所需的导航功能。

再如，以安装有 YunOS 智能操作系统的、车载温控设备为例，用户在驾车的过程中，车载温控设备的显示屏幕上弹出对话框，询问用户是否需要开启空调，并显示有标题为“确定”和“取消”的按钮，用户可以直接说以下语音指示信息“我选确定”或者“同意”，那么本实施例提供的方法就可以将接收到的语音信息转换成操作指令“被操作界面元素类型：按钮，操作描述信息：确定”，然后通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素为：标题为“确定”的按钮，并采用发送触控事件的方式，模拟用户触控“确定”按钮的操作，从而实现开启空调的目的。

由此可见，在互联网汽车领域实施本实施例提供的方法，不仅具有上面描述的、可以减少开发人员工作量的有益效果，而且由于用户在无需手动点击屏幕的情况下即可完成与车载电子设备的交互过程，为用户的驾车安全提供有效保证，特别是，由于用户不用严格地按照设定模式输入语音，只要描述出希望操控的界面元素的相关提示信息即可，

因此语音交互过程更加灵活，可以有效提升用户的使用体验。

在上述的实施例中，提供了一种基于语音的交互方法，与之相对应的，本申请还提供一种基于语音的交互装置。请参看图 4，其为本申请的一种基于语音的交互装置的实施例示意图。由于装置实施例基本相似于方法实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。下述描述的装置实施例仅仅是示意性的。

本实施例的一种基于语音的交互装置，包括：操作指令生成单元 401，用于将接收到的语音信息转换成操作指令；界面元素确定单元 402，用于通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；触发响应单元 403，用于采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

可选的，所述界面元素包括：可供操作的控件；所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息。

可选的，所述触发响应单元，具体用于根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

可选的，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述触发响应单元，包括针对每个已确定的界面元素触发的下列子单元：

触控位置确定子单元，用于至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟触控事件上报子单元，用于模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

触控事件发送子单元，用于由窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

可选的，所述界面元素的布局信息包括：界面元素在显示屏幕中的位置和界面元素的尺寸信息。

可选的，所述手动输入设备包括：触摸屏，或者鼠标。

可选的，所述装置还包括：

监控记录单元，用于在将接收到的语音信息转换成操作指令之前，通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述界面元素确定单元，具体用于查询所述监控记录单元记录的显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

可选的，所述操作指令生成单元，包括：

语音识别子单元，用于通过语音识别将接收到的语音信息转换成文字形式的指令；

语音分析匹配子单元，用于通过语义分析和基于模板的匹配操作，将所述文字形式的指令转换成操作指令。

可选的，所述操作指令的内容包括：被操作界面元素的类型、和操作描述信息。

可选的，所述显示窗口包括：预设类型的显示窗口；所述预设类型的显示窗口包括：应用窗口、和/或系统窗口。

可选的，所述预设类型的显示窗口包括：预设类型的顶层显示窗口。

可选的，所述预设类型的顶层显示窗口包括：顶层应用窗口；所述顶层应用窗口的数量为 1；

所述界面元素确定单元，具体用于通过查询顶层应用窗口包含的界面元素信息，确定所述顶层应用窗口中、可响应所述操作指令的界面元素。

可选的，所述显示窗口的数量大于等于 1；

所述界面元素确定单元，具体用于遍历各显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

可选的，所述装置还包括：

界面元素选择单元，用于当所述界面元素确定单元确定的界面元素的数目大于或者等于 2 时，按照预设方式，从所述确定的界面元素中选择一个界面元素；

所述触发响应单元，具体用于采用发送触控事件的方式，触发所选界面元素响应所述操作指令。

可选的，所述装置部署于具有语音输入装置及显示屏幕的电子设备上；所述电子设备包括：移动通讯设备、个人电脑、平板电脑、或应用于互联网汽车的车载电子设备。

此外，本申请还提供了一种电子设备；所述电子设备实施例如下：

请参考图 5，其示出了本申请的一种电子设备的实施例的示意图。

所述电子设备，包括：语音输入装置 501，用于接收输入的语音信息；显示器 502；处理器 503；存储器 504，用于存储指令；其中，所述处理器耦合于所述存储器，用于读取所述存储器存储的指令，并执行如下操作：

将接收到的语音信息转换成操作指令；通过查询显示窗口包含的界面元素信息，确

定可响应所述操作指令的界面元素；采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

可选的，所述采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令，包括：根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

可选的，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，包括针对每个已确定的界面元素执行下述操作：

至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

可选的，在所述将接收到的语音信息转换成操作指令之前，包括：通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述查询显示窗口包含的界面元素信息，包括：查询基于上述监控方式记录的显示窗口包含的界面元素信息。

此外，本申请还提供了一种操作系统，所述操作系统实施例如下：

请参考图 6，其示出了本申请的一种操作系统的实施例的示意图。

所述操作系统包括：操作指令生成单元 601，用于将接收到的语音信息转换成操作指令；界面元素确定单元 602，用于通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；触控事件发送单元 603，用于发送触控事件，以触发已确定的界面元素响应所述操作指令。

可选的，所述触控事件发送单元，具体用于根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

可选的，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述触控事件发送单元，包括针对每个已确定的界面元素触发的下列子单元：

触控位置确定子单元，用于至少根据所述界面元素的布局信息，确定对应于所述界

面元素的触控位置；

模拟触控事件上报子单元，用于模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

触控事件发送子单元，用于由窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

可选的，所述操作系统还包括：

监控记录单元，用于在将接收到的语音信息转换成操作指令之前，通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述界面元素确定单元，具体用于查询所述监控记录单元记录的显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

本申请虽然以较佳实施例公开如上，但其并不是用来限定本申请，任何本领域技术人员在不脱离本申请的精神和范围内，都可以做出可能的变动和修改，因此本申请的保护范围应当以本申请权利要求所界定的范围为准。

在一个典型的配置中，计算设备包括一个或多个处理器 (CPU)、 输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器 (RAM) 和/或非易失性内存等形式，如只读存储器 (ROM) 或闪存(flash RAM)。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、 程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、 其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括非暂存电脑可读媒体 (transitory media)，如调制的数据信号和载波。

本领域技术人员应明白，本申请的实施例可提供为方法、系统或计算机程序产品。

因此，本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

5

# 权 利 要 求 书

1.一种基于语音的交互方法，其特征在于，包括：

将接收到的语音信息转换成操作指令；

通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；

采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

2.根据权利要求 1 所述的方法，其特征在于，所述界面元素包括：可供操作的控件；所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息。

3.根据权利要求 1 所述的方法，其特征在于，所述采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令，包括：

根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

4.根据权利要求 3 所述的方法，其特征在于，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，包括针对每个已确定的界面元素执行下述操作：

至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

5.根据权利要求 4 所述的方法，其特征在于，所述界面元素的布局信息包括：界面元素在显示屏幕中的位置和界面元素的尺寸信息。

6.根据权利要求 4 所述的方法，其特征在于，所述手动输入设备包括：触摸屏，或者鼠标。

7.根据权利要求 1 所述的方法，其特征在于，在所述将接收到的语音信息转换成操作指令之前，包括：

通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述查询显示窗口包含的界面元素信息，包括：

查询基于上述监控方式记录的显示窗口包含的界面元素信息。

8.根据权利要求 1 所述的方法，其特征在于，所述将接收到的语音信息转换成操作指令，包括：

通过语音识别将接收到的语音信息转换成文字形式的指令；

通过语义分析和基于模板的匹配操作，将所述文字形式的指令转换成操作指令。

9.根据权利要求 1 所述的方法，其特征在于，所述操作指令的内容包括：被操作界面元素的类型、和操作描述信息。

10.根据权利要求 1 所述的方法，其特征在于，所述显示窗口包括：预设类型的显示窗口；所述预设类型的显示窗口包括：应用窗口、和/或系统窗口。

11.根据权利要求 10 所述的方法，其特征在于，所述预设类型的显示窗口包括：预设类型的顶层显示窗口。

12.根据权利要求 11 所述的方法，其特征在于，所述预设类型的顶层显示窗口包括：顶层应用窗口；所述顶层应用窗口的数量为 1；

所述通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素，包括：通过查询顶层应用窗口包含的界面元素信息，确定所述顶层应用窗口中、可响应所述操作指令的界面元素。

13.根据权利要求 1 所述的方法，其特征在于，所述显示窗口的数量大于等于 1；

所述通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素，包括：

遍历各显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

14.根据权利要求 1 所述的方法，其特征在于，若确定的可响应所述操作指令的界面元素的数目大于或者等于 2，所述方法还包括：

按照预设方式，从所述确定的界面元素中选择一个界面元素；

所述触发已确定的界面元素响应所述操作指令，包括：触发所选界面元素响应所述操作指令。

15.根据权利要求 1-14 任一项所述的方法，其特征在于，所述方法在具有语音输入装置及显示屏幕的电子设备上实施，所述电子设备包括：移动通讯设备、个人电脑、平板电脑、或应用于互联网汽车的车载电子设备。

16.一种基于语音的交互装置，其特征在于，包括：

操作指令生成单元，用于将接收到的语音信息转换成操作指令；

界面元素确定单元，用于通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；

触发响应单元，用于采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

17.根据权利要求 16 所述的装置，其特征在于，所述触发响应单元，具体用于根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

18.根据权利要求 17 所述的装置，其特征在于，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述触发响应单元，包括针对每个已确定的界面元素触发的下列子单元：

触控位置确定子单元，用于至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟触控事件上报子单元，用于模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

触控事件发送子单元，用于由窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

19.根据权利要求 16 所述的装置，其特征在于，还包括：

监控记录单元，用于在将接收到的语音信息转换成操作指令之前，通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述界面元素确定单元，具体用于查询所述监控记录单元记录的显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

20.一种电子设备，其特征在于，包括：

语音输入装置，用于接收输入的语音信息；

显示器；

处理器；

存储器，用于存储指令；

其中，所述处理器耦合于所述存储器，用于读取所述存储器存储的指令，并执行如下操作：

将接收到的语音信息转换成操作指令；通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令。

21.根据权利要求 20 所述的电子设备，其特征在于，所述采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令，包括：根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

22.根据权利要求 21 所述的电子设备，其特征在于，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，包括针对每个已确定的界面元素执行下述操作：

至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

23.根据权利要求 20 所述的电子设备，其特征在于，在所述将接收到的语音信息转换成操作指令之前，包括：通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述查询显示窗口包含的界面元素信息，包括：查询基于上述监控方式记录的显示窗口包含的界面元素信息。

24.一种操作系统，其特征在于，包括：

操作指令生成单元，用于将接收到的语音信息转换成操作指令；

界面元素确定单元，用于通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素；

触控事件发送单元，用于发送触控事件，以触发已确定的界面元素响应所述操作指令。

25.根据权利要求 24 所述的操作系统，其特征在于，所述触控事件发送单元，具体用于根据所述操作指令，向已确定的界面元素所属窗口发送针对所述界面元素的触控事件，以触发所述界面元素所属窗口执行相应的处理操作。

26.根据权利要求 25 所述的操作系统，其特征在于，所述界面元素信息包括：界面元素的类型、对应于所述类型的预设属性的值、和标识界面元素的信息；所述标识界面元素的信息包括：界面元素的布局信息；

所述触控事件发送单元，包括针对每个已确定的界面元素触发的下列子单元：

触控位置确定子单元，用于至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置；

模拟触控事件上报子单元，用于模拟手动输入设备向操作系统上报针对所述触控位置的触控事件；

触控事件发送子单元，用于由窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件。

27.根据权利要求 24 所述的操作系统，其特征在于，所述操作系统还包括：

监控记录单元，用于在将接收到的语音信息转换成操作指令之前，通过监控窗口管理服务在显示窗口发生变更时的处理过程，记录显示窗口包含的界面元素信息；

所述界面元素确定单元，具体用于查询所述监控记录单元记录的显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素。

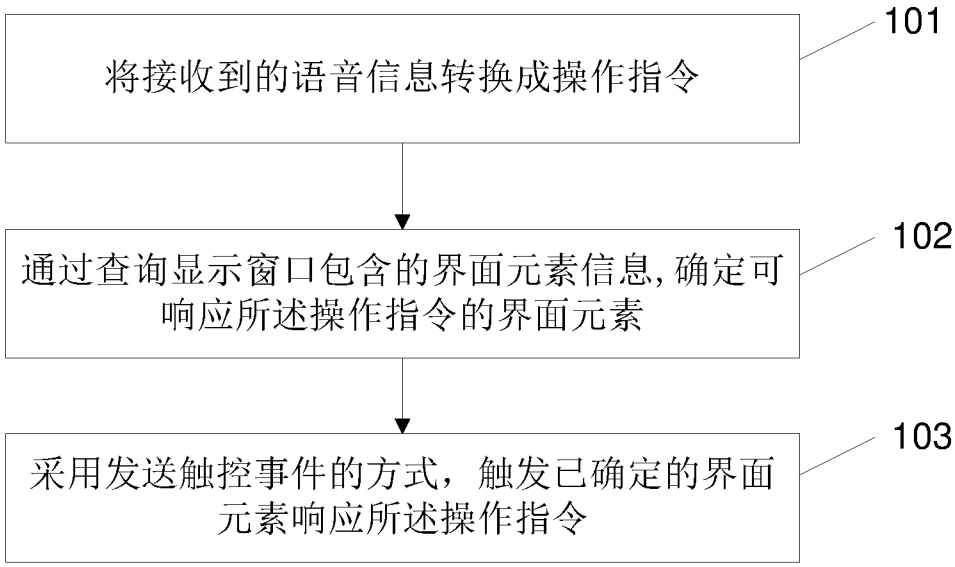

101
将接收到的语音信息转换成操作指令
102
通过查询显示窗口包含的界面元素信息，确定可响应所述操作指令的界面元素
103
采用发送触控事件的方式，触发已确定的界面元素响应所述操作指令

图 1

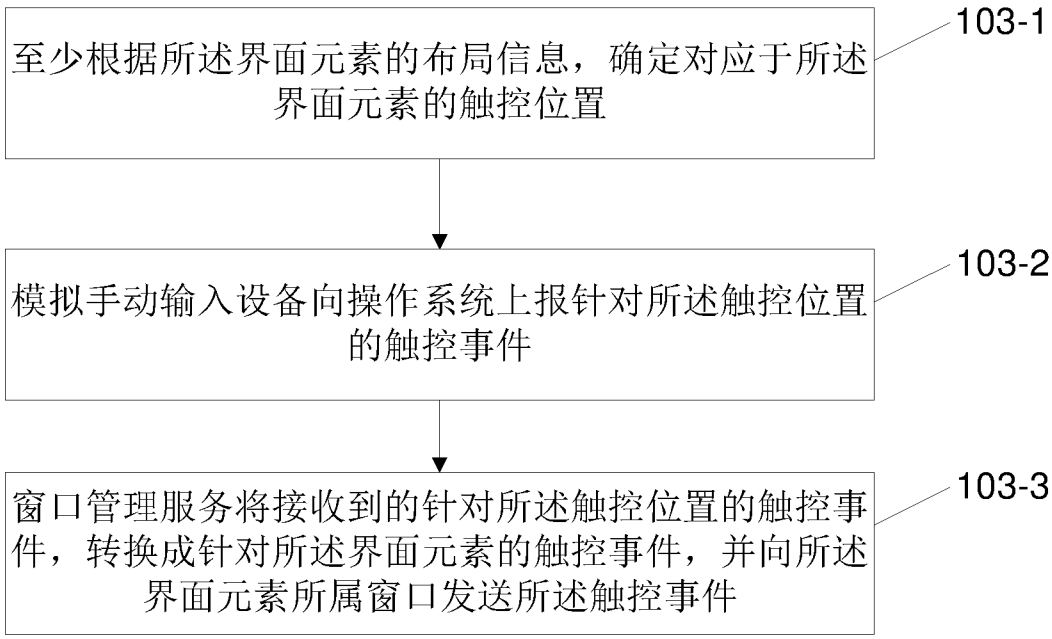

103-1
至少根据所述界面元素的布局信息，确定对应于所述界面元素的触控位置
103-2
模拟手动输入设备向操作系统上报针对所述触控位置的触控事件
103-3
窗口管理服务将接收到的针对所述触控位置的触控事件，转换成针对所述界面元素的触控事件，并向所述界面元素所属窗口发送所述触控事件

图 2

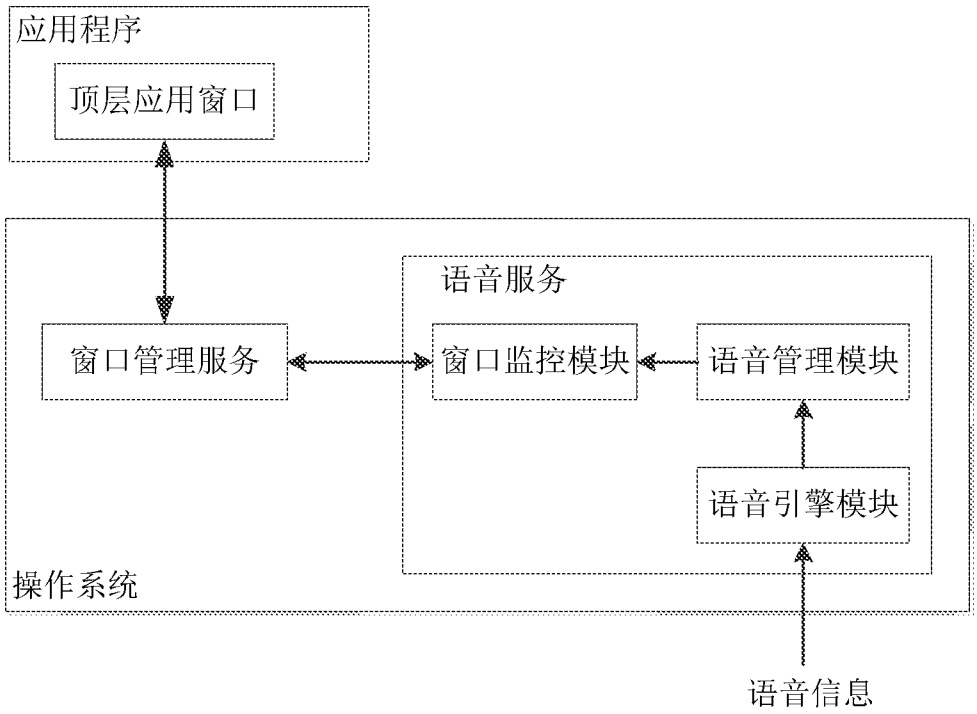

应用程序
顶层应用窗口
语音服务
窗口管理服务
窗口监控模块
语音管理模块
语音引擎模块
操作系统
语音信息

图 3

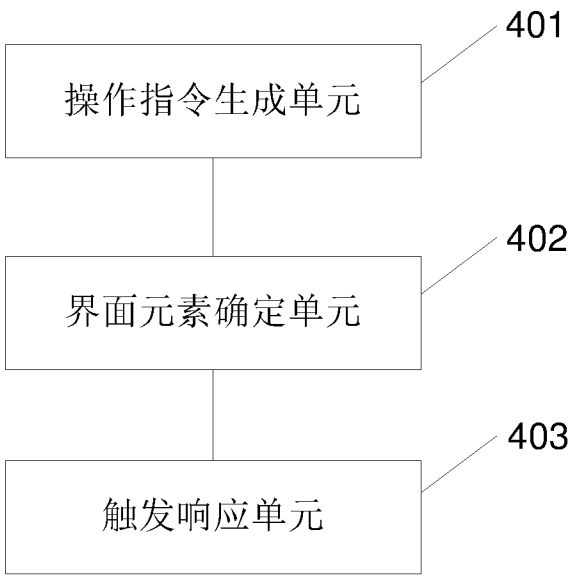

401
操作指令生成单元
402
界面元素确定单元
403
触发响应单元

图 4

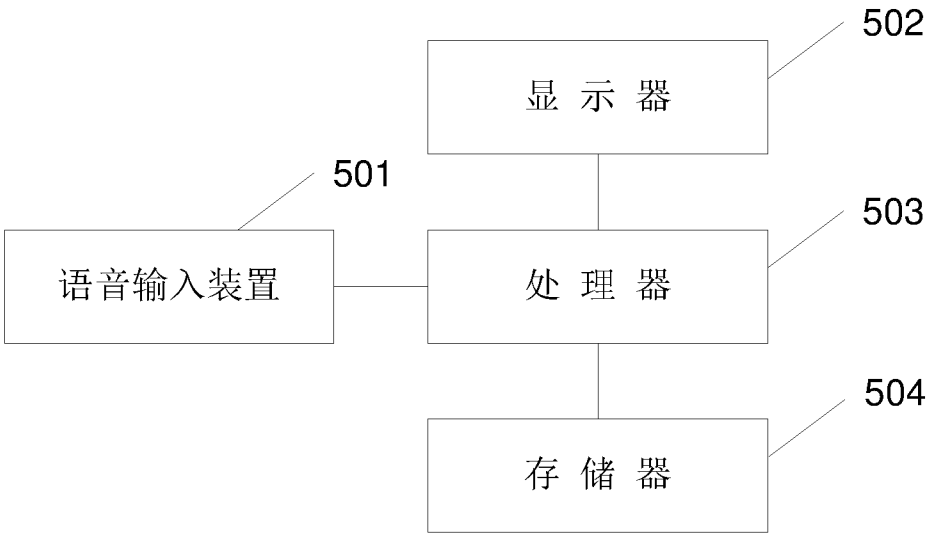

显 示 器
502
501
语音输入装置
处 理 器
503
存 储 器
504

图 5

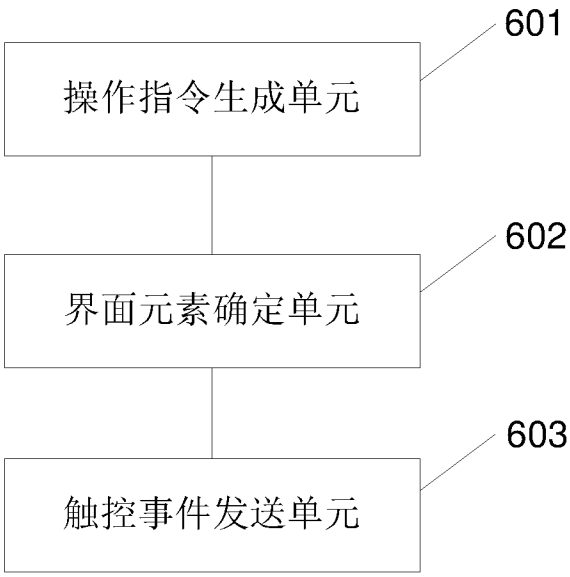

操作指令生成单元
601
界面元素确定单元
602
触控事件发送单元
603

图 6

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/118429

**A. CLASSIFICATION OF SUBJECT MATTER**

G06F 3/0488 (2013.01) 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)

G06F 3/-

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, VEN, USTXT, WOTXT, EPTXT, CNKI: 语音, 交互, 操作指令, 控制指令, 用户界面, 界面元素, 窗口, 应用, 菜单, 项目, 元素, 触控, 触摸, 触发, 事件; voice, sound, interact, operational command, control command, user interface, UI, window, app, menu, item, touch, trigger, event

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | CN 103869931 A (SAMSUNG ELECTRONICS (CHINA) R & D CENTER et al.), 18 June 2014 (18.06.2014), description, paragraphs 0010-0061 and 0066-0154 | 1-27 |
| X | CN 1647023 A (SAP AG), 27 July 2005 (27.07.2005), description, page 4, line 13 to page 14, line 7 | 1-27 |
| A | CN 105653164 A (YULONG COMPUTER TELECOMMUNICATION SCIENTIFIC (SHENZHEN) CO., LTD.), 08 June 2016 (08.06.2016), entire document | 1-27 |
| A | US 2013342457 A1 (COX, C.D. et al.), 26 December 2013 (26.12.2013), entire document | 1-27 |

☐ 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 | Date of mailing of the international search report |
|---|---|
| 07 February 2018 | 21 March 2018 |
| Name and mailing address of the ISA<br>State Intellectual Property Office of the P. R. China<br>No. 6, Xitucheng Road, Jimenqiao<br>Haidian District, Beijing 100088, China<br>Facsimile No. (86-10) 62019451 | Authorized officer<br>XING, Lichao<br>Telephone No. (86-10) 28950382 |

# INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/118429

| Patent Documents referred in the Report | Publication Date | Patent Family | Publication Date |
|---|---|---|---|
| CN 103869931 A | 18 June 2014 | CN 103869931 B | 08 February 2017 |
| CN 1647023 A | 27 July 2005 | WO 03071519 A3 | 11 March 2004 |
| | | CN 1647023 B | 27 June 2012 |
| | | EP 1485790 B1 | 12 December 2007 |
| | | DE 60318021 T2 | 20 November 2008 |
| | | WO 03071519 A8 | 03 March 2005 |
| | | EP 1485790 A4 | 28 February 2007 |
| | | DE 60318021 D1 | 24 January 2008 |
| | | EP 1485790 A2 | 15 December 2004 |
| | | US 6882974 B2 | 19 April 2005 |
| | | US 2003158736 A1 | 21 August 2003 |
| | | AU 2003215226 A1 | 09 September 2003 |
| | | WO 03071519 A2 | 28 August 2003 |
| | | AU 2003215226 A8 | 27 October 2015 |
| | | DE 20321559 U1 | 20 March 2008 |
| CN 105653164 A | 08 June 2016 | None | |
| US 2013342457 A1 | 26 December 2013 | None | |

# 国际检索报告

国际申请号
PCT/CN2017/118429

A. 主题的分类

G06F 3/0488(2013.01)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/-

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,VEN,USTXT,WOTXT,EPTXT,CNKI:语音，交互，操作指令，控制指令，用户界面，界面元素，窗口，应用，菜单，项目，元素，触控，触摸，触发，事件；voice,sound,interact,operational command,control command,user interface,UI,window,app,menu,item,touch,trigger,event

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 103869931 A (三星电子中国研发中心等) 2014年 6月 18日 (2014-06-18)<br>说明书第0010-0061，0066-0154段 | 1-27 |
| X | CN 1647023 A (SAP股份公司) 2005年 7月 27日 (2005-07-27)<br>说明书第4页第13行-第14页第7行 | 1-27 |
| A | CN 105653164 A (宇龙计算机通信科技深圳有限公司) 2016年 6月 8日 (2016-06-08)<br>全文 | 1-27 |
| A | US 2013342457 A1 (COX, CHRISTOPHER DAVID等) 2013年 12月 26日 (2013-12-26)<br>全文 | 1-27 |

☐ 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 | 国际检索报告邮寄日期 |
|---|---|
| 2018年 2月 7日 | 2018年 3月 21日 |
| ISA/CN的名称和邮寄地址<br>中华人民共和国国家知识产权局(ISA/CN)<br>中国北京市海淀区蓟门桥西土城路6号 100088<br>传真号 (86-10)62019451 | 受权官员<br>邢丽超<br>电话号码 (86-20)28950382 |

# 国际检索报告
## 关于同族专利的信息

国际申请号
PCT/CN2017/118429

| 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 | 103869931 | A | 2014年 6月 18日 | CN | 103869931 | B | 2017年 2月 8日 |
| CN | 1647023 | A | 2005年 7月 27日 | WO | 03071519 | A3 | 2004年 3月 11日 |
| | | | | CN | 1647023 | B | 2012年 6月 27日 |
| | | | | EP | 1485790 | B1 | 2007年 12月 12日 |
| | | | | DE | 60318021 | T2 | 2008年 11月 20日 |
| | | | | WO | 03071519 | A8 | 2005年 3月 3日 |
| | | | | EP | 1485790 | A4 | 2007年 2月 28日 |
| | | | | DE | 60318021 | D1 | 2008年 1月 24日 |
| | | | | EP | 1485790 | A2 | 2004年 12月 15日 |
| | | | | US | 6882974 | B2 | 2005年 4月 19日 |
| | | | | US | 2003158736 | A1 | 2003年 8月 21日 |
| | | | | AU | 2003215226 | A1 | 2003年 9月 9日 |
| | | | | WO | 03071519 | A2 | 2003年 8月 28日 |
| | | | | AU | 2003215226 | A8 | 2015年 10月 27日 |
| | | | | DE | 20321559 | U1 | 2008年 3月 20日 |
| CN | 105653164 | A | 2016年 6月 8日 | | 无 | | |
| US | 2013342457 | A1 | 2013年 12月 26日 | | 无 | | |