

(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
2021 年 3 月 4 日 (04.03.2021)



(10) 国际公布号
WO 2021/036622 A1

(51) 国际专利分类号:
G06F 3/041 (2006.01)

(21) 国际申请号: PCT/CN2020/104291

(22) 国际申请日: 2020 年 7 月 24 日 (24.07.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910804635.X 2019 年 8 月 28 日 (28.08.2019) CN

(71) 申请人: 北京市商汤科技开发有限公司(BEIJING SENSETIME TECHNOLOGY DEVELOPMENT CO., LTD.) [CN/CN]; 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。

(72) 发明人: 张子隆(ZHANG, Zilong); 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。刘畅(LIU, Chang); 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。

(74) 代理人: 北京博思佳知识产权代理有限公司(BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街 9 号嘉华大厦 B 座 409, Beijing 100085 (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, IT, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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,

(54) Title: INTERACTION METHOD, APPARATUS, AND DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 交互方法、装置、设备以及存储介质

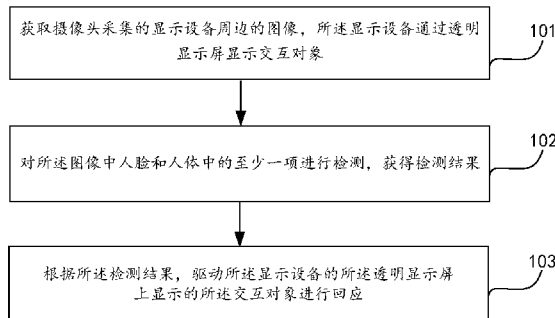


图 1

101 OBTAIN AN IMAGE OF THE PERIPHERY OF A DISPLAY DEVICE COLLECTED BY A CAMERA, THE DISPLAY DEVICE DISPLAYING AN INTERACTION OBJECT BY MEANS OF A TRANSPARENT DISPLAY SCREEN
102 DETECT AT LEAST ONE OF A HUMAN FACE AND A HUMAN BODY IN THE IMAGE TO OBTAIN A DETECTION RESULT
103 DRIVE, ACCORDING TO THE DETECTION RESULT, THE INTERACTION OBJECT DISPLAYED ON THE TRANSPARENT DISPLAY SCREEN OF THE DISPLAY DEVICE TO RESPOND

(57) Abstract: The present disclosure relates to an interaction method, apparatus, and device, and a storage medium. The method comprises: obtaining an image of the periphery of a display device collected by a camera, the display device displaying an interaction object by means of a transparent display screen; detecting at least one of a human face and a human body in the image to obtain a detection result; and driving, according to the detection result, the interaction object displayed on the transparent display screen of the display device to respond.

(57) 摘要: 本公开涉及一种交互方法、装置、设备以及存储介质。所述方法包括: 获取摄像头采集的显示设备周边的图像, 所述显示设备通过透明显示屏显示交互对象; 对所述图像中人脸和人体中的至少一项进行检测, 获得检测结果; 根据所述检测结果, 驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

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))。

交互方法、装置、设备以及存储介质

技术领域

[01]本公开涉及计算机视觉技术领域，具体涉及一种交互方法、装置、设备以及存储介质。

5 背景技术

[02]人机交互的方式大多为：用户基于按键、触摸、语音进行输入，设备通过在显示屏上呈现图像、文本或虚拟人物进行回应。目前虚拟人物多是在语音助理的基础上改进得到的，只是对设备输入的语音进行输出，用户与虚拟人物的交互还停留表面上。

发明内容

10 [03]本公开实施例提供一种交互方案。

[04]第一方面，提供一种交互方法，所述方法包括：获取摄像头采集的显示设备周边的图像，所述显示设备通过透明显示屏显示交互对象；对所述图像中的人脸和人体中的至少一项进行检测，获得检测结果；根据所述检测结果，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

15 [05]在本公开实施例中，通过对显示设备周边的图像进行检测，并根据检测结果驱动显示设备的所述透明显示屏上显示的交互对象进行回应，可以使交互对象的回应更符合实际交互需求，并使用户与所述交互对象之间的交互更加真实、生动，从而提升用户体验。

[06]在一个示例中，所述显示设备通过所述透明显示屏显示所述交互对象的倒影，或者，所述显示设备在底板上显示所述交互对象的倒影。

20 [07]通过在透明显示屏上显示立体画面，并在透明显示屏或底板上形成倒影以实现立体效果，能够使所显示的交互对象更加立体、生动，提升用户的交互感受。

[08]在一个示例中，所述交互对象包括具有立体效果的虚拟人物。

[09]通过利用具有立体效果的虚拟人物与用户进行交互，可以使交互过程更加自然，提升用户的交互感受。

25 [10]在一个示例中，所述检测结果至少包括所述显示设备的当前服务状态；所述当前服

务状态包括等待用户状态、用户离开状态、发现用户状态、服务激活状态、服务中状态中的任一种。

[11]通过结合设备的当前服务状态来驱动所述交互对象进行回应，可以使所述交互对象的回应更符合用户的交互需求。

- 5 [12]在一个示例中，所述对所述图像中的人脸和人体中的至少一项进行检测，获得检测结果，包括：响应于当前时刻未检测到所述人脸和所述人体，且在当前时刻之前的设定时间段内未检测到所述人脸和所述人体，确定所述当前服务状态为所述等待用户状态；或者，响应于当前时刻未检测到所述人脸和所述人体，且在当前时刻之前的设定时间段内检测到所述人脸和所述人体，确定所述当前服务状态为所述用户离开状态；或者，响
- 10 应于当前时刻检测到所述人脸和所述人体中的至少一项，确定所述显示设备的当前服务状态为发现用户状态。

[13]在没有用户与交互对象进行交互的情况下，通过确定显示设备当前处于等待用户状态或用户离开状态，并驱动所述交互对象进行不同的回应，使所述交互对象的展示状态更符合交互需求、更有针对性。

- 15 [14]在一个示例中，所述检测结果还包括用户属性信息和/或用户历史操作信息；所述方法还包括：在确定所述显示设备的所述当前服务状态为所述发现用户状态之后，通过所述图像获得所述用户属性信息，和/或，查找与所述用户的人脸和人体中的至少一项的特征信息相匹配的所述用户历史操作信息。

- 20 [15]通过获取用户历史操作信息，并结合所述用户历史操作信息驱动所述交互对象，可以使所述交互对象更有针对性地对所述用户进行回应。

[16]在一个示例中，所述方法还包括：响应于检测到至少两个用户，获得所述至少两个用户的特征信息；根据所述至少两个用户的特征信息，确定所述至少两个用户中的目标用户；驱动所述显示设备的所述透明显示屏上显示的所述交互对象对所述目标用户进行回应。

- 25 [17]通过根据至少两个用户的特征信息来确定所述至少两个用户中的目标用户，并驱动所述交互对象对所述目标对象进行回应，能够在多用户场景下选择进行交互的目标用户，并实现不同目标用户之间的切换和响应，从而提升用户体验。

[18]在一个示例中，所述方法还包括：获取所述显示设备的环境信息；所述根据所述检

测结果，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应，包括：根据所述检测结果以及所述显示设备的环境信息，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

5 [19]在一个示例中，所述环境信息包括所述显示设备的地理位置、所述显示设备的互联网协议（IP）地址以及所述显示设备所在区域的天气、日期中的至少一项。

[20]通过获取所述显示设备的环境信息，并结合所述环境信息来驱动所述交互对象进行回应，可以使所述交互对象的回应更符合实际交互需求，使用户与交互对象之间的交互更加真实、生动，从而提升用户体验。

10 [21]在一个示例中，根据所述检测结果以及所述环境信息，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应，包括：获得与所述检测结果和所述环境信息相匹配的、预先设定的回应标签；驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出与所述回应标签相应的回应。

15 [22]在一个示例中，所述驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出与所述回应标签相应的回应，包括：将所述回应标签输入至预先训练的神经网络，由所述神经网络输出与所述回应标签对应的驱动内容，所述驱动内容用于驱动所述交互对象输出相应的动作、表情、语言中的一项或多项。

20 [23]通过对不同的检测结果和不同的环境信息的组合配置相应的回应标签，并通过所述回应标签来驱动交互对象输出相应的动作、表情、语言中的一项或多项，可以驱动交互对象根据设备的不同状态、不同的场景，做出不同的回应，以使所述交互对象的回应更加多样化。

25 [24]在一个示例中，所述方法还包括：响应于确定所述当前服务状态为所述发现用户状态，在驱动所述交互对象进行回应之后，追踪所述显示设备周边的图像中所检测到的用户；在追踪所述用户的过程中，响应于检测到所述用户输出的第一触发信息，确定所述显示设备进入所述服务激活状态，并驱动所述交互对象展示与所述第一触发信息匹配的服务。

[25]通过本公开实施例提供的交互方法，用户无需进行按键、触摸或者语音输入，仅站在显示设备的周边，设备中显示的交互对象即可以有针对性地做出欢迎的动作，并按照用户的需求或者兴趣展示服务项目，提升用户的使用感受。

[26]在一个示例中，所述方法还包括：在所述显示设备处于所述服务激活状态时，响应于检测到所述用户输出的第二触发信息，确定所述显示设备进入服务中状态，并驱动所述交互对象展示与所述第二触发信息匹配的服务。

[27]在所述显示设备进入发现用户状态之后，提供两种粒度的识别方式。第一粒度（粗粒度）识别方式为在检测到用户输出的第一触发信息的情况下，使设备进入服务激活状态，并驱动所述交互对象展示与所述第一触发信息匹配的服务；第二粒度（细粒度）识别方式为在检测到用户输出的第二触发信息的情况下，使设备进入服务中状态，并驱动所述交互对象提供相应的服务。通过上述两种粒度的识别方式，能够使用户与交互对象的交互更流畅、更自然。

[28]在一个示例中，所述方法还包括：响应于确定所述当前服务状态为发现用户状态，根据所述用户在所述图像中的位置，获得所述用户相对于所述透明显示屏中展示的所述交互对象的位置信息；根据所述位置信息调整所述交互对象的朝向，使所述交互对象面向所述用户。

[29]通过根据用户的位置来自动调整交互对象的朝向，使所述交互对象始终保持与用户面对面，使交互更加友好，提升了用户的交互体验。

[30]第二方面，提供一种交互装置，所述装置包括：图像获取单元，用于获取摄像头采集的显示设备周边的图像，所述显示设备通过透明显示屏显示交互对象；检测单元，用于对所述图像中的人脸和人体中的至少一项进行检测，获得检测结果；驱动单元，用于根据所述检测结果，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

[31]在一个示例中，所述显示设备还通过所述透明显示屏显示所述交互对象的倒影，或者，所述显示设备还在底板上显示所述交互对象的倒影。

[32]在一个示例中，所述交互对象包括具有立体效果的虚拟人物。

[33]在一个示例中，所述检测结果至少包括所述显示设备的当前服务状态；所述当前服务状态包括等待用户状态、用户离开状态、发现用户状态、服务激活状态、服务中状态中的任一种。

[34]在一个示例中，所述检测单元具体用于：响应于当前时刻未检测到人脸和人体，且在当前时刻之前的设定时间段内未检测到人脸和人体，确定所述当前服务状态为等待用

户状态。

[35]在一个示例中，所述检测单元用于：响应于当前时刻未检测到人脸和人体，且在当前时刻之前的设定时间内段检测到人脸和人体，确定所述当前服务状态为用户离开状态。

5 [36]在一个示例中，所述检测单元具体用于：响应于当前时刻检测到所述人脸和所述人体中的至少一项，确定所述显示设备的当前服务状态为发现用户状态。

[37]在一个示例中，所述检测结果还包括用户属性信息和/或用户历史操作信息；所述装置还包括信息获取单元，所述信息获取单元用于：通过所述图像获得用户属性信息，和/或，查找与所述用户的人脸和人体中的至少一项的特征信息相匹配的用户历史操作信息。

10 [38]在一个示例中，所述装置还包括目标确定单元，所述目标确定单元用于：响应于通过所述检测单元检测到至少两个用户，获得所述至少两个用户的特征信息；根据所述至少两个用户的特征信息，确定所述至少两个用户中的目标用户，其中，所述驱动单元用于驱动所述显示设备的所述透明显示屏上显示的所述交互对象对所述目标用户进行回应。

15 [39]在一个示例中，所述装置还包括用于获取所述显示设备的环境信息的环境信息获取单元，其中，所述驱动单元用于：根据所述检测结果以及所述显示设备的环境信息，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

[40]在一个示例中，所述环境信息至少包括所述显示设备的地理位置、所述显示设备的IP地址，以及所述显示设备所在区域的天气、日期中的一项或多项。

20 [41]在一个示例中，所述驱动单元还用于：获得与所述检测结果和所述环境信息相匹配的、预先设定的回应标签；驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出与所述回应标签相应的回应。

25 [42]在一个示例中，所述驱动单元在用于根据所述回应标签，驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出相应的回应时，具体用于：将所述回应标签输入至预先训练的神经网络，由所述神经网络输出与所述回应标签对应的驱动内容，所述驱动内容用于驱动所述交互对象输出相应的动作、表情、语言中的一项或多项。

[43]在一个示例中，所述装置还包括服务激活单元，所述服务激活单元用于：响应于所述检测单元检测出所述当前服务状态为发现用户状态，在所述驱动单元驱动所述交互对象进行回应之后，追踪在所述显示设备周边的图像中所检测到的用户；在追踪所述用户

的过程中，响应于检测到所述用户输出的第一触发信息，确定所述显示设备进入服务激活状态，并使所述驱动单元驱动所述交互对象展示所提供的服务。

[44] 在一个示例中，所述装置还包括服务单元，所述服务单元用于：在所述显示设备处于所述服务激活状态时，响应于检测到所述用户输出的第二触发信息，确定所述显示设备进入服务中状态，其中，所述驱动单元用于驱动所述交互对象展示与所述第二触发信息匹配的服务。

[45] 在一个示例中，所述装置还包括方向调整单元，所述方向调整单元用于：响应于所述检测单元检测出所述当前服务状态为发现用户状态，根据所述用户在所述图像中的位置，获得所述用户相对于所述透明显示屏中展示的所述交互对象的位置信息；根据所述位置信息调整所述交互对象的朝向，使所述交互对象面向所述用户。

[46] 第三方面，提供一种交互设备，所述设备包括处理器；用于存储可由处理器执行的指令的存储器，当所述指令被执行时，促使所述处理器实现本公开提供的任一实施方式所述的方法。

[47] 第四方面，提供一种计算机可读存储介质，其上存储有计算机程序，当所述计算机程序被处理器执行时，使所述处理器实现本公开提供的任一实施方式所述的方法。

附图说明

[48] 图 1 示出根据本公开至少一个实施例的交互方法的流程图；

[49] 图 2 示出根据本公开至少一个实施例的显示交互对象的示意图；

[50] 图 3 示出根据本公开至少一个实施例的交互装置的结构示意图；

[51] 图 4 示出根据本公开至少一个实施例的交互设备的结构示意图。

具体实施方式

[52] 这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所述的、本公开的一些方面相一致的装置和方法的例子。

[53] 本文中术语“和/或”，仅仅是一种描述关联对象的关联关系，表示可以存在三种关

系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。另外，本文中术语“至少一种”表示多种中的任意一种或多种中的至少两种的任意组合，例如，包括 A、B、C 中的至少一种，可以表示包括从 A、B 和 C 构成的集合中选择的任意一个或多个元素。

5 [54]图 1 示出根据本公开至少一个实施例的交互方法的流程图，如图 1 所示，所述方法包括步骤 101~步骤 103。

[55]在步骤 101 中，获取摄像头采集的显示设备周边的图像，所述显示设备通过透明显示屏显示交互对象。

10 [56]所述显示设备周边，包括所述显示设备的设定范围内任意方向，例如可以包括所述显示设备的前向、侧向、后方、上方中的一个或多个方向。

[57]用于采集图像的摄像头，可以设置在显示设备上，也可以作为外接设备，独立于显示设备之外。并且所述摄像头采集的图像可以在显示设备的透明显示屏上进行显示。所述摄像头的数量可以为多个。

[58]可选的，摄像头所采集的图像可以是视频流中的一帧，也可以是实时获取的图像。

15 [59]在步骤 102 中，对所述图像中的人脸和人体中的至少一项进行检测，获得检测结果。

[60]通过对显示设备周边的图像进行人脸和/或人体检测，获得检测结果，例如所述显示设备周边是否有用户、有几个用户，并可以通过人脸和/或人体识别技术从所述图像中获取关于用户的相关信息，或者通过用户的图像进行查询以获得用户的相关信息；还可以通过图像识别技术识别用户的动作、姿势、手势等等。本领域技术人员应当理解，以上
20 检测结果仅为示例，还可以包括其他检测结果。

[61]在步骤 103 中，根据所述检测结果，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

[62]响应于不同的检测结果，可以驱动所述交互对象进行不同的回应。例如，对于在显示设备周边没有用户的情况，驱动所述交互对象输出欢迎的动作、表情、语音等等。

25 [63]本公开实施例中，通过对显示设备周边的图像进行检测，并根据检测结果驱动显示设备的所述透明显示屏上显示的交互对象进行回应，可以使交互对象的回应更符合用户的交互需求，并使用户与所述交互对象之间的交互更加真实、生动，从而提升用户体验。

[64]在一些实施例中，所述显示设备的透明显示屏显示的交互对象包括具有立体效果的

虚拟人物。

[65]通过利用具有立体效果的虚拟人物与用户进行交互，可以使交互过程更加自然，提升用户的交互感受。

5 [66]本领域技术人员应当理解，交互对象并不限于具有立体效果的虚拟人物，还可以是虚拟动物、虚拟物品、卡通形象等等其他能够实现交互功能的虚拟形象。

[67]在一些实施例中，可以通过以下方法实现透明显示屏所显示的交互对象的立体效果。

[68]人眼看到物体是否为立体的观感，通常由物体本身的外形以及物体的光影效果所决定。该光影效果例如为在物体不同区域的高光和暗光，以及光线照射在物体后在地面的投影（即倒影）。

10 [69]利用以上原理，在一个示例中，在透明显示屏上显示出交互对象的立体视频或图像的画面的同时，还在透明显示屏上显示出该交互对象的倒影，从而使得人眼可以观察到立体效果的交互对象。

[70]在另一个示例中，所述透明显示屏的下方设置有底板，并且所述透明显示与所述底板呈垂直或倾斜状。在透明显示屏显示出交互对象的立体视频或图像的画面的同时，在
15 所述底板上显示出所述交互对象的倒影，从而使得人眼可以观察到立体效果的交互对象。

[71]在一些实施例中，所述显示设备还包括箱体，并且所述箱体的正面设置为透明，例如通过玻璃、塑料等材料实现透明设置。透过箱体的正面能够看到透明显示屏的画面以及透明显示屏或底板上画面的倒影，从而使得人眼可以观察到立体效果的交互对象，如图2所示。

20 [72]在一些实施例中，箱体内还设有一个或多个光源，为透明显示屏提供光线。

[73]在本公开实施例中，通过在透明显示屏上显示交互对象的立体视频或图像的畫面，并在透明显示屏或底板上形成该交互对象的倒影以实现立体效果，能够使所显示的交互对象更加立体、生动，提升用户的交互感受。

[74]在一些实施例中，所述检测结果可以包括所述显示设备的当前服务状态，所述当前
25 服务状态例如包括等待用户状态、发现用户状态、用户离开状态、服务激活状态、服务中状态中的任一种。本领域技术人员应当理解，所述显示设备的当前服务状态还可以包括其他状态，不限于以上所述。

[75]在显示设备周边的图像中未检测到人脸和人体的情况下，表示所述显示设备周边没

有用户，也即该显示设备当前并未处于与用户进行交互的状态。这种状态包含了在当前时刻之前的设定时间段内都没有用户与设备进行交互，也即等待用户状态；还包含了用户在当前时刻之前的设定时间段内与用户进行了交互，显示设备正处于用户离开状态。对于这两种不同的状态，应当驱动所述交互对象进行不同的回应。例如，对于等待用户状态，可以驱动所述交互对象结合当前环境进行欢迎用户的回应；而对于用户离开状态，可以驱动所述交互对象对上一个与其交互的用户进行结束服务的回应。

[76]在一个示例中，可以通过以下方式确定等待用户状态。响应于当前时刻未检测到人脸和人体，且在当前时刻之前的设定时间段内，例如 5 秒钟，未检测到人脸和人体，并且也未追踪到人脸和人体的情况下，确定该显示设备的当前服务状态为等待用户状态。

[77]在一个示例中，可以通过以下方式确定用户离开状态。响应于当前时刻未检测到人脸和人体，且在当前时刻之前的设定时间段内，例如 5 秒钟，检测到了人脸和/或人体，或者追踪到了人脸和/或人体的情况下，确定该显示设备的当前服务状态为用户离开状态。

[78]在显示设备处于等待用户状态或用户离开状态时，可以根据所述显示设备的当前服务状态驱动所述交互对象进行回应。例如，在显示设备处于等待用户状态时，可以驱动显示设备所显示的交互对象做出欢迎的动作或手势，或者做出一些有趣的动作，或者输出欢迎光临的语音。在显示设备处于用户离开状态时，可以驱动所述交互对象做出再见的动作或手势，或者输出再见的语音。

[79]在从显示设备周边的图像中检测到了人脸和/或人体的情况下，表示所述显示设备周边存在用户，则可以将检测到用户这一时刻的当前服务状态确定为发现用户状态。

[80]在检测到显示设备周边存在用户时，可以通过所述图像获得所述用户的用户属性信息。例如，可以通过人脸和/或人体检测的结果确定设备周边存在几个用户；针对每个用户，可以通过人脸和/或人体识别技术，从所述图像中获取关于所述用户的相关信息，例如用户的性别、用户的大致年龄等等，对于不同性别、不同年龄层次的用户，可以驱动交互对象进行不同的回应。

[81]在发现用户状态下，对于所检测到的用户，还可以获取存储在所述显示设备中的用户历史操作信息，和/或，获取存储在云端的用户历史操作信息，以确定该用户是否为老客户，或者是否为 VIP 客户。所述用户历史操作信息还可以包含所述用户的姓名、性别、年龄、服务记录、备注等等。该用户历史操作信息可以包含所述用户自行输入的信息，也可以包括所述显示设备和/或云端记录的信息。通过获取用户历史操作信息，可以驱动

所述交互对象更有针对性地对所述用户进行回应。

[82] 在一个示例中，可以根据所检测到的用户的人脸和/或人体的特征信息查找与所述用户相匹配的用户历史操作信息。

5 [83] 在显示设备处于发现用户状态时，可以根据所述显示设备的当前服务状态、从所述图像获取的用户属性信息、通过查找获取的用户历史操作信息，来驱动所述交互对象进行回应。在初次检测到一个用户的时候，所述用户历史操作信息可以为空，也即根据所述当前服务状态、所述用户属性信息和所述环境信息来驱动所述交互对象。

10 [84] 在显示设备周边的图像中检测到一个用户的情况下，可以首先通过图像对该用户进行人脸和/或人体识别，获得关于所述用户的用户属性信息，比如该用户为女性，年龄在20岁~30岁之间；之后根据该用户的人脸和/或人体特征信息，在显示设备中和/或云端进行搜索，以查找与所述特征信息相匹配的用户历史操作信息，例如该用户的姓名、服务记录等等。之后，在发现用户状态下，驱动所述交互对象对该女性用户作出有针对性的欢迎动作，并向该女性用户展示可以为其提供的服务。根据用户历史操作信息中包括的该用户曾经使用的服务项目，可以调整提供服务的顺序，以使用户能够更快的发现感
15 兴趣的服务项目。

[85] 当在设备周边的图像中检测到至少两个用户的情况下，可以首先获得所述至少两个用户的特征信息，该特征信息可以包括用户姿态信息、用户属性信息中的至少一项，并且所述特征信息与用户历史操作信息对应，其中，所述用户姿态信息可以通过对所述图像中所述用户的动作进行识别而获得。

20 [86] 接下来，根据所获得的所述至少两个用户的特征信息来确定所述至少两个用户中的目标用户。可以结合实际场景综合评估各个用户的特征信息，以确定待进行交互的目标用户。

[87] 在确定了目标用户后，则可以驱动所述显示设备的所述透明显示屏上显示的所述交互对象对所述目标用户进行回应。

25 [88] 在一些实施例，在发现用户状态下，驱动所述交互对象进行回应之后，通过追踪在显示设备周边的图像中所检测到的用户，例如可以追踪所述用户的面部表情，和/或，追踪所述用户的动作，等等，并通过判断所述用户有无主动交互的表情和/或动作来判断是否要使所述显示设备进入服务激活状态。

[89]在一个示例中，在追踪所述用户的过程中，可以设置指定触发信息，例如眨眼、点头、挥手、举手、拍打等常见的人与人之间打招呼的表情和/或动作。为了与下文进行区别，此处不妨将所设置的指定触发信息称为第一触发信息。在检测到所述用户输出的所述第一触发信息的情况下，确定所述显示设备进入服务激活状态，并驱动所述交互对象

5 展示与所述第一触发信息匹配的服务，例如可以利用语言展示，也可以用显示在屏幕上的文字信息来展示。

[90]目前常见的体感交互需要用户先举手一段时间来激活服务，选中服务后需要保持手部位置不动若干秒后才能完成激活。本公开实施例所提供的交互方法，无需用户先举手一段时间激活服务，也无需保持手部位置不同完成选择。通过自动判断用户的指定触发

10 信息，可以自动激活服务，使设备处于服务激活状态，避免了用户举手等待一段时间，提升了用户体验。

[91]在一些实施例中，在服务激活状态下，可以设置指定触发信息，例如特定的手势动作，和/或特定的语音指令等。为了与上文进行区别，此处不妨将所设置的指定触发信息称为第二触发信息。在检测到所述用户输出的所述第二触发信息的情况下，确定所述显

15 示设备进入服务中状态，并驱动所述交互对象展示与所述第二触发信息匹配的服务。

[92]在一个示例中，通过用户输出的第二触发信息来执行相应的服务。例如，可以为用户提供的服务包括：第一服务选项、第二服务选项、第三服务选项等等，可以并且为第一服务选项配置相应的第二触发信息，例如，可以设置语音“一”为与第一服务选项相对应的第二触发信息，设置语音“二”为与第二服务选项相对应的第二触发信息，以此

20 类推。当检测到所述用户输出其中一个语音，则所述显示设备进入与第二触发信息相应的服务选项，并驱动所述交互对象根据该服务选项所设置的内容提供服务。

[93]在本公开实施例中，在所述显示设备进入发现用户状态之后，提供两种粒度的识别方式。第一粒度（粗粒度）识别方式为在检测到用户输出的第一触发信息的情况下，使设备进入服务激活状态，并驱动所述交互对象展示与所述第一触发信息匹配的服务；第

25 二粒度（细粒度）识别方式为在检测到用户输出的第二触发信息的情况下，使设备进入服务中状态，并驱动所述交互对象提供相应的服务。通过上述两种粒度的识别方式，能够使用户与交互对象的交互更流畅、更自然。

[94]通过本公开实施例提供的交互方法，用户无需进行按键、触摸或者语音输入，仅站在显示设备的周边，显示设备中显示的交互对象即可以有针对性地做出欢迎的动作，并

按照用户的需求或者兴趣展示能够提供的服务项目，提升用户的使用感受。

[95] 在一些实施例中，可以获取所述显示设备的环境信息，根据所述检测结果和所述环境信息，来驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

5 [96] 所述显示设备的环境信息可以通过所述显示设备的地理位置和/或所述显示设备的应用场景获取。所述环境信息例如可以是所述显示设备的地理位置、互联网协议 (Internet Protocol, IP) 地址，也可以是所述显示设备所在区域的天气、日期等等。本领域技术人员应当理解，以上环境信息仅为示例，还可以包括其他环境信息。

10 [97] 举例来说，在显示设备处于等待用户状态和用户离开状态时，可以根据所述显示设备的当前服务状态和环境信息驱动所述交互对象进行回应。例如，在所述显示设备处于等待用户状态时，环境信息包括时间、地点、天气情况，可以驱动显示设备所显示的交互对象做出欢迎的动作和手势，或者做出一些有趣的动作，并输出语音“现在是 X 年 X 月 X 日 XX 时刻，天气 XX，欢迎光临 XX 城市的 XX 商场，很高兴为您服务”。在通用的欢迎动作、手势和语音外，还加入了当前时间、地点和天气情况，不但提供了更多资讯，还使交互对象的回应更符合交互需求、更有针对性。

15 [98] 通过对显示设备周边的图像进行用户检测，并根据检测结果和所述显示设备的环境信息，来驱动所述显示设备中显示的交互对象进行回应，使交互对象的回应更符合交互需求，使用户与交互对象之间的交互更加真实、生动，从而提升用户体验。

20 [99] 在一些实施例中，可以根据所述检测结果和所述环境信息，获得相匹配的、预先设定的回应标签；之后根据所述回应标签来驱动所述交互对象做出相应的回应。所述回应标签可以对应于所述交互对象的动作、表情、手势、语言中的一项或多项的驱动文本。对于不同的检测结果和环境信息，可以根据所确定的回应标签获得相应的驱动文本，从而可以驱动所述交互对象输出相应的动作、表情、语言中的一项或多项。

[100] 例如，若当前服务状态为用户等待状态，并且环境信息指示地点为上海，对应的回应标签可以是：动作为欢迎动作，语音为“欢迎来到上海”。

25 [101] 再比如，若当前服务状态为发现用户状态，环境信息指示时间为上午，用户属性信息指示女性，并且用户历史记录指示姓氏为张，对应的回应标签可以是：动作为欢迎动作，语音为“张女士上午好，欢迎光临，很高兴为您提供服务”。

[102] 通过对于不同的检测结果和不同的环境信息的组合配置相应的回应标签，并通

过所述回应标签来驱动交互对象输出相应的动作、表情、语言中的一项或多项，可以驱动交互对象根据设备的不同状态、不同的场景，做出不同的回应，以使所述交互对象的回应更加多样化。

5 [103] 在一些实施例中，可以通过将所述回应标签输入至预先训练的神经网络，输出与所述回应标签对应的驱动文本，以驱动所述交互对象输出相应的动作、表情、语言中的一项或多项。

[104] 其中，所述神经网络可以通过样本回应标签集来进行训练，其中，所述样本回应标签标注了对应的驱动文本。所述神经网络经训练后，对于所输出的回应标签能够输出相应的驱动文本，以驱动所述交互对象输出相应的动作、表情、语言中的一项或多项。
10 相较于直接在显示设备端或云端搜索对应的驱动文本，采用预先训练的神经网络，对于没有预先设置驱动文本的回应标签，也能够生成驱动文本，以驱动所述交互对象进行适当的回应。

[105] 在一些实施例中，针对高频、重要的场景，还可以通过人工配置的方式进行优化。也即，对于出现频次较高的检测结果与环境信息的组合，可以为其对应的回应标签人工配置驱动文本。在该场景出现时，自动调用相应的驱动文本驱动所述交互对象进行
15 回应，以使交互对象的动作、表情更加自然。

[106] 在一个实施例中，响应于所述显示设备处于发现用户状态，根据所述用户在所述图像中的位置，获得所述用户相对于所述透明显示屏中展示的所述交互对象的位置信息；并根据所述位置信息调整所述交互对象的朝向，使所述交互对象面向所述用户。

20 [107] 通过根据用户的位置来自动调整交互对象的身体朝向，使所述交互对象始终保持与用户面对面，使交互更加友好，提升了用户的交互体验。

[108] 在一些实施例中，所述交互对象的图像是通过虚拟摄像头采集的。虚拟摄像头是应用于 3D 软件、用于采集图像的虚拟软件摄像头，交互对象是通过所述虚拟摄像头采集的 3D 图像显示在屏幕上的。因此用户的视角可以理解为 3D 软件中虚拟摄像头的
25 视角，这样就会带来一个问题，就是交互对象无法实现用户之间的眼神交流。

[109] 为了解决以上问题，在本公开至少一个实施例中，在调整交互对象的身体朝向的同时，还使所述交互对象的视线保持对准所述虚拟摄像头。由于交互对象在交互过程中面向用户，并且视线保持对准虚拟摄像头，因此用户会有交互对象正看自己的错觉，可以提升用户与交互对象交互的舒适性。

[110] 图 3 示出根据本公开至少一个实施例的交互装置的结构示意图，如图 3 所示，该装置可以包括：图像获取单元 301、检测单元 302 和驱动单元 303。

[111] 其中，图像获取单元 301，用于获取摄像头采集的显示设备周边的图像，所述显示设备通过透明显示屏显示交互对象；检测单元 302，用于对所述图像中的人脸和人体中的至少一项进行检测，获得检测结果；驱动单元 303，用于根据所述检测结果，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

[112] 在一些实施例中，所述显示设备还通过所述透明显示屏显示所述交互对象的倒影，或者，所述显示设备在底板上显示所述交互对象的倒影。

[113] 在一些实施例中，所述交互对象包括具有立体效果的虚拟人物。

10 [114] 在一些实施例中，所述检测结果至少包括所述显示设备的当前服务状态，所述当前服务状态包括等待用户状态、用户离开状态、发现用户状态、服务激活状态、服务中状态中的任一种。

[115] 在一些实施例中，检测单元 302 具体用于：响应于当前时刻未检测到人脸和人体，且在当前时刻之前的设定时间段内未检测到人脸和人体，确定所述当前服务状态为等待用户状态。

[116] 在一些实施例中，检测单元 302 具体用于：响应于当前时刻未检测到人脸和人体，且在当前时刻之前的设定时间段内检测到人脸和/或人体，确定所述当前服务状态为用户离开状态。

20 [117] 在一些实施例中，检测单元 302 具体用于：响应于检测到所述人脸和所述人体中的至少一项，确定所述显示设备的当前服务状态为发现用户状态。

[118] 在一些实施例中，所述检测结果还包括用户属性信息和/或用户历史操作信息；所述装置还包括信息获取单元，所述信息获取单元用于：通过所述图像获得用户属性信息，和/或，查找与所述用户的人脸和人体中的至少一项的特征信息相匹配的用户历史操作信息。

25 [119] 在一些实施例中，所述装置还包括目标确定单元，所述目标确定单元用于：响应于检测到至少两个用户，获得所述至少两个用户的特征信息；根据所述至少两个用户的特征信息，确定所述至少两个用户中的目标用户。所述驱动单元 303 驱动所述显示设备的所述透明显示屏上显示的所述交互对象对所述目标用户进行回应。

[120] 在一些实施例中，所述装置还包括用于获取环境信息的环境信息获取单元；所述驱动单元 303 具体用于：根据所述检测结果以及所述显示设备的环境信息，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

5 [121] 在一些实施例中，所述环境信息至少包括所述显示设备的地理位置、所述显示设备的 IP 地址，以及所述显示设备所在区域的天气、日期中的一项或多项。

[122] 在一些实施例中，驱动单元 303 具体用于：获得与所述检测结果和所述环境信息相匹配的、预先设定的回应标签；驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出与所述回应标签相应的回应。

10 [123] 在一些实施例中，驱动单元 303 在用于根据所述回应标签，驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出相应的回应时，具体用于：将所述回应标签输入至预先训练的神经网络，由所述神经网络输出与所述回应标签对应的驱动内容，所述驱动内容用于驱动所述交互对象输出相应的动作、表情、语言中的一项或多项。

15 [124] 在一些实施例中，所述装置还包括服务激活单元，所述服务激活单元用于：响应于所述检测单元 302 检测出当前服务状态为发现用户状态，在所述驱动单元 303 驱动所述交互对象进行回应之后，追踪在所述显示设备周边的图像中所检测到的用户；在追踪所述用户的过程中，响应于检测到所述用户输出的第一触发信息，确定所述显示设备进入服务激活状态，并使所述驱动单元 303 驱动所述交互对象展示与所述第一触发信息匹配的服务。

20 [125] 在一些实施例中，所述装置还包括服务单元，所述服务单元用于：在所述显示设备处于所述服务激活状态时，响应于检测到所述用户输出的第二触发信息，确定所述显示设备进入服务中状态，其中，所述驱动单元 303 用于驱动所述交互对象提供与所述第二触发信息匹配的服务。

25 [126] 在一些实施例中，所述装置还包括方向调整单元，所述方向调整单元用于：响应于所述检测单元 302 检测出当前服务状态为发现用户状态，根据所述用户在所述图像中的位置，获得所述用户相对于所述透明显示屏中展示的所述交互对象的位置信息；根据所述位置信息调整所述交互对象的朝向，使所述交互对象面向所述用户。

[127] 本公开至少一个实施例还提供了一种交互设备，如图 4 所示，所述设备包括存储器 401、处理器 402。存储器 401 用于存储可由处理器执行的计算机指令，所述指令被执行时，促使处理器 402 实现本公开任一实施例所述的方法。

[128] 本公开至少一个实施例还提供了一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时，使所述处理器实现本公开任一实施例所述的交互方法。

[129] 本领域技术人员应明白，本公开一个或多个实施例可提供为方法、系统或计算机程序产品。因此，本公开一个或多个实施例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，本公开一个或多个实施例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。

[130] 本公开中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于数据处理设备实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

[131] 上述对本公开特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下，在权利要求书中记载的行为或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外，在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中，多任务处理和并行处理也是可以的或者可能是有利的。

[132] 本公开中的主题及功能操作的实施例可以在以下中实现：数字电子电路、有形体现的计算机软件或固件、包括本公开中公开的结构及其结构性等同物的计算机硬件、或者它们中的一个或多个的组合。本公开中的主题的实施例可以实现为一个或多个计算机程序，即编码在有形非暂时性程序载体上以被数据处理装置执行或控制数据处理装置的操作的计算机程序指令中的一个或多个模块。可替代地或附加地，程序指令可以被编码在人工生成的传播信号上，例如机器生成的电、光或电磁信号，该信号被生成以将信息编码并传输到合适的接收机装置以由数据处理装置执行。计算机存储介质可以是机器可读存储设备、机器可读存储基板、随机或串行存取存储器设备、或它们中的一个或多个的组合。

[133] 本公开中的处理及逻辑流程可以由执行一个或多个计算机程序的一个或多个可编程计算机执行，以通过根据输入数据进行操作并生成输出来执行相应的功能。所述处理及逻辑流程还可以由专用逻辑电路——例如 FPGA（现场可编程门阵列）或 ASIC（专

用集成电路)来执行,并且装置也可以实现为专用逻辑电路。

[134] 适合用于执行计算机程序的计算机包括,例如通用和/或专用微处理器,或任何其他类型的中央处理单元。通常,中央处理单元将从只读存储器和/或随机存取存储器接收指令和数据。计算机的基本组件包括用于实施或执行指令的中央处理单元以及用于存储指令和数据的一个或多个存储器设备。通常,计算机还将包括用于存储数据的一个或多个大容量存储设备,例如磁盘、磁光盘或光盘等,或者计算机将可操作地与此大容量存储设备耦接以从其接收数据或向其传送数据,抑或两种情况兼而有之。然而,计算机不是必须具有这样的设备。此外,计算机可以嵌入在另一设备中,例如移动电话、个人数字助理(PDA)、移动音频或视频播放器、游戏操纵台、全球定位系统(GPS)接收机、或例如通用串行总线(USB)闪存驱动器的便携式存储设备,仅举几例。

[135] 适合于存储计算机程序指令和数据的计算机可读介质包括所有形式的非易失性存储器、媒介和存储器设备,例如包括半导体存储器设备(例如EPROM、EEPROM和闪存设备)、磁盘(例如内部硬盘或可移动盘)、磁光盘以及CD ROM和DVD-ROM盘。处理器和存储器可由专用逻辑电路补充或并入专用逻辑电路中。

[136] 虽然本公开包含许多具体实施细节,但是这些不应被解释为限制本公开的范围或所要求保护的范

15 围,而是主要用于描述本公开的一些实施例的特征。本公开的多个实施例中的某些特征也可以在单个实施例中

20 被组合实施。另一方面,单个实施例中的各种特征也可以在多个实施例中分开实施或以任何合适的子组合来实施。此外,虽然特征可以如上所述在某些组合中起作用并且甚至最初如此要求保护,但是来自所要求保护的组合中的一个或多个特征在某些情况下可以从该组合中去除,并且所要求保护的组合可以指向子组合或子组合的变型。

[137] 类似地,虽然在附图中以特定顺序描绘了操作,但是这不应被理解为要求这些操作以所示的特定顺序执行或顺次执行、或者要求所有例示的操作被执行,以实现期望的结果。在某些情况下,多任务和并行处理可能是有利的。此外,上述实施例中的各种系统模块和组件的分离不应被理解为在所有实施例中均需要这样的分离,并且应当理解,所描述的程序组件和系统通常可以一起集成在单个软件产品中,或者封装成多个软件产品。

[138] 由此,主题的特定实施例已被描述。其他实施例在所附权利要求书的范围以内。在某些情况下,权利要求书中记载的动作可以以不同的顺序执行并且仍实现期望的结果。

此外，附图中描绘的处理并非必需所示的特定顺序或顺次顺序，以实现期望的结果。在某些实现中，多任务和并行处理可能是有利的。

[139] 以上所述仅为本公开的一些实施例而已，并不用以限制本公开。凡在本公开的精神和原则之内所做的任何修改、等同替换、改进等，均应包含在本公开的范围之内。

权利要求书

1、一种交互方法，所述方法包括：

5 获取摄像头采集的显示设备周边的图像，所述显示设备通过透明显示屏显示交互对象；
对所述图像中的人脸和人体中的至少一项进行检测，获得检测结果；
根据所述检测结果，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

10 2、根据权利要求 1 所述的方法，其中，所述显示设备通过所述透明显示屏显示所述交互对象的倒影，或者，所述显示设备在底板上显示所述交互对象的倒影。

3、根据权利要求 1 或 2 所述的方法，其中，所述交互对象包括具有立体效果的虚拟人物。

4、根据权利要求 1 至 3 任一所述的方法，其中，所述检测结果至少包括所述显示设备的当前服务状态；

15 所述当前服务状态包括等待用户状态、用户离开状态、发现用户状态、服务激活状态、服务中状态中的任一种。

5、根据权利要求 4 所述的方法，其中，所述对所述图像中的人脸和人体中的至少一项进行检测，获得检测结果，包括：

20 响应于当前时刻未检测到所述人脸和所述人体，且在当前时刻之前的设定时间段内未检测到所述人脸和所述人体，确定所述当前服务状态为所述等待用户状态；或者，

响应于当前时刻未检测到所述人脸和所述人体，且在当前时刻之前的设定时间段内检测到所述人脸和所述人体，确定所述当前服务状态为所述用户离开状态；或者，

响应于当前时刻检测到所述人脸和所述人体中的至少一项，确定所述显示设备的当前服务状态为发现用户状态。

25 6、根据权利要求 4 所述的方法，其中，所述检测结果还包括用户属性信息和/或用户历史操作信息；

所述方法还包括：在确定所述显示设备的所述当前服务状态为所述发现用户状态之后，通过所述图像获得所述用户属性信息，和/或，查找与所述用户的人脸和人体中的至少一项的特征信息相匹配的所述用户历史操作信息。

30 7、根据权利要求 1 至 6 任一项所述的方法，所述方法还包括：

响应于检测到至少两个用户，获得所述至少两个用户的特征信息；

根据所述至少两个用户的特征信息，确定所述至少两个用户中的目标用户；

驱动所述显示设备的所述透明显示屏上显示的所述交互对象对所述目标用户进行回应。

5 8、根据权利要求 1 至 7 任一所述的方法，所述方法还包括：

获取所述显示设备的环境信息；

所述根据所述检测结果，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应，包括：

10 根据所述检测结果以及所述环境信息，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

9、根据权利要求 8 所述的方法，其中，所述环境信息包括所述显示设备的地理位置、所述显示设备的互联网协议 IP 地址以及所述显示设备所在区域的天气、日期中的至少一项。

15 10、根据权利要求 8 或 9 所述的方法，其中，所述根据所述检测结果以及所述环境信息，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应，包括：

获得与所述检测结果和所述环境信息相匹配的、预先设定的回应标签；

驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出与所述回应标签相应的回应。

20 11、根据权利要求 10 所述的方法，其中，所述驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出与所述回应标签相应的回应，包括：

将所述回应标签输入至预先训练的神经网络，由所述神经网络输出与所述回应标签对应的驱动内容，所述驱动内容用于驱动所述交互对象输出相应的动作、表情、语言中的一项或多项。

12、根据权利要求 4 至 11 任一项所述的方法，所述方法还包括：

25 响应于确定所述当前服务状态为所述发现用户状态，在驱动所述交互对象进行回应之后，追踪所述显示设备周边的图像中所检测到的用户；

在追踪所述用户的过程中，响应于检测到所述用户输出的第一触发信息，确定所述显示设备进入所述服务激活状态，并驱动所述交互对象展示与所述第一触发信息匹配的服务。

30 13、根据权利要求 12 所述的方法，所述方法还包括：

在所述显示设备处于所述服务激活状态时，响应于检测到所述用户输出的第二触发信息，确定所述显示设备进入所述服务中状态，并驱动所述交互对象展示与所述第二触发信息匹配的服务。

14、根据权利要求 4 至 13 任一项所述的方法，所述方法还包括：

5 响应于确定所述当前服务状态为所述发现用户状态，根据所述用户在所述图像中的位置，获得所述用户相对于所述透明显示屏中展示的所述交互对象的位置信息；

根据所述位置信息调整所述交互对象的朝向，使所述交互对象面向所述用户。

15、一种交互装置，所述装置包括：

10 图像获取单元，用于获取摄像头采集的显示设备周边的图像，所述显示设备通过透明显示屏显示交互对象；

检测单元，用于对所述图像中的人脸和人体中的至少一项进行检测，获得检测结果；

驱动单元，用于根据所述检测结果，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

15 16、根据权利要求 15 所述的装置，其中，所述显示设备通过所述透明显示屏显示所述交互对象的倒影，或者，所述显示设备在底板上显示所述交互对象的倒影。

17、根据权利要求 15 或 16 所述的装置，其中，所述交互对象包括具有立体效果的虚拟人物。

18、根据权利要求 15 至 17 任一项所述的装置，其中，所述检测结果至少包括所述显示设备的当前服务状态；

20 所述当前服务状态包括等待用户状态、用户离开状态、发现用户状态、服务激活状态、服务中状态中的任一种。

19、根据权利要求 18 所述的装置，其中，所述检测单元用于：

响应于当前时刻未检测到所述人脸和所述人体，且在当前时刻之前的设定时间段内未检测到所述人脸和所述人体，确定所述当前服务状态为所述等待用户状态；或者，

25 响应于当前时刻未检测到所述人脸和所述人体，且在当前时刻之前的设定时间段内检测到所述人脸和所述人体，确定所述当前服务状态为所述用户离开状态；或者，

响应于当前时刻检测到所述人脸和所述人体中的至少一项，确定所述显示设备的当前服务状态为所述发现用户状态。

30 20、根据权利要求 19 所述的装置，其中，所述检测结果还包括用户属性信息和/或用户历史操作信息；

所述装置还包括信息获取单元，所述信息获取单元用于：

通过所述图像获得所述用户属性信息，和/或，查找与所述用户的人脸和人体中的至少一项的特征信息相匹配的所述用户历史操作信息。

5 21、根据权利要求 15 至 20 任一项所述的装置，所述装置还包括目标确定单元，所述目标确定单元用于：

响应于通过所述检测单元检测到至少两个用户，获得所述至少两个用户的特征信息；根据所述至少两个用户的特征信息，确定所述至少两个用户中的目标用户，

其中，所述驱动单元用于驱动所述显示设备的所述透明显示屏上显示的所述交互对象对所述目标用户进行回应。

10 22、根据权利要求 15 至 21 任一项所述的装置，所述装置还包括用于获取所述显示设备的环境信息的环境信息获取单元，所述驱动单元用于：

根据所述检测结果以及所述环境信息，驱动所述显示设备的所述透明显示屏上显示的所述交互对象进行回应。

15 23、根据权利要求 22 所述的装置，其中，所述环境信息包括所述显示设备的地理位置、所述显示设备的互联网协议 IP 地址以及所述显示设备所在区域的天气、日期中的至少一项。

24、根据权利要求 22 或 23 所述的装置，其中，所述驱动单元用于：

获得与所述检测结果和所述环境信息相匹配的、预先设定的回应标签；

20 驱动所述显示设备的所述透明显示屏上显示的所述交互对象做出与所述回应标签相应的回应。

25、根据权利要求 24 所述的装置，其中，所述驱动单元还用于：

将所述回应标签输入至预先训练的神经网络，由所述神经网络输出与所述回应标签对应的驱动内容，所述驱动内容用于驱动所述交互对象输出相应的动作、表情、语言中的一项或多项。

25 26、根据权利要求 18 至 25 任一项所述的装置，所述装置还包括服务激活单元，所述服务激活单元用于：

响应于所述检测单元检测出所述当前服务状态为所述发现用户状态，在所述驱动单元驱动所述交互对象进行回应之后，追踪所述显示设备周边的图像中所检测到的用户；

30 在追踪所述用户的过程中，响应于检测到所述用户输出的第一触发信息，确定所述显示设备进入所述服务激活状态，其中，所述驱动单元用于驱动所述交互对象展示与

所述第一触发信息匹配的服务。

27、根据权利要求 26 所述的装置，所述装置还包括服务单元，所述服务单元用于：

在所述显示设备处于所述服务激活状态时，响应于检测到所述用户输出的第二触发信息，确定所述显示设备进入所述服务中状态，其中，所述驱动单元用于驱动所述交互

5 对象展示与所述第二触发信息匹配的服务。

28、根据权利要求 18 至 27 任一项所述的装置，所述装置还包括方向调整单元，所述方向调整单元用于：

响应于所述检测单元检测出所述当前服务状态为所述发现用户状态，根据所述用户在所述图像中的位置，获得所述用户相对于所述透明显示屏中展示的所述交互对象的位置信息；

10

根据所述位置信息调整所述交互对象的朝向，使所述交互对象面向所述用户。

29、一种交互设备，所述设备包括：

处理器；以及

用于存储可由所述处理器执行的指令的存储器，

15

其中，所述指令在被执行时，促使所述处理器实现根据权利要求 1 至 14 任一项所述的交互方法。

30、一种计算机可读存储介质，其上存储有计算机程序，其中，所述计算机程序被处理器执行时，使所述处理器实现根据权利要求 1 至 14 任一所述的交互方法。

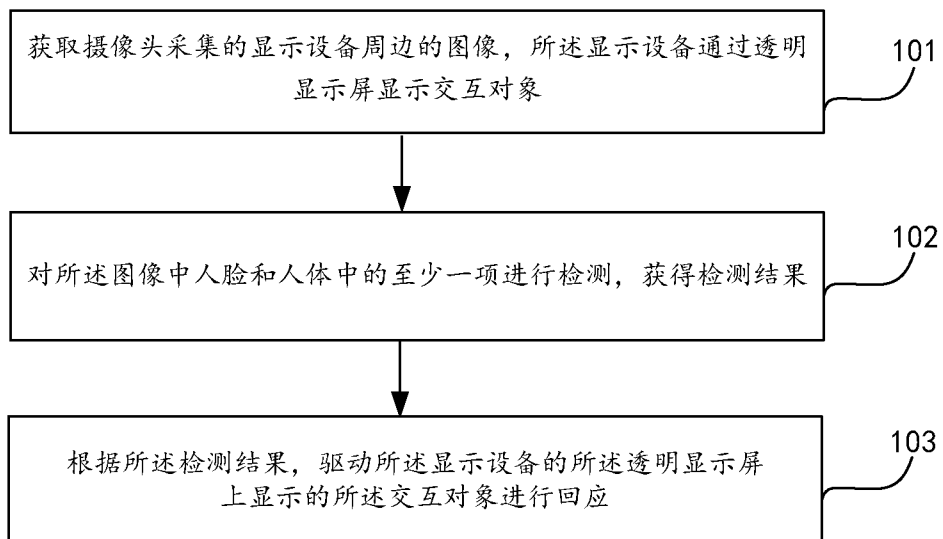


图 1



图 2

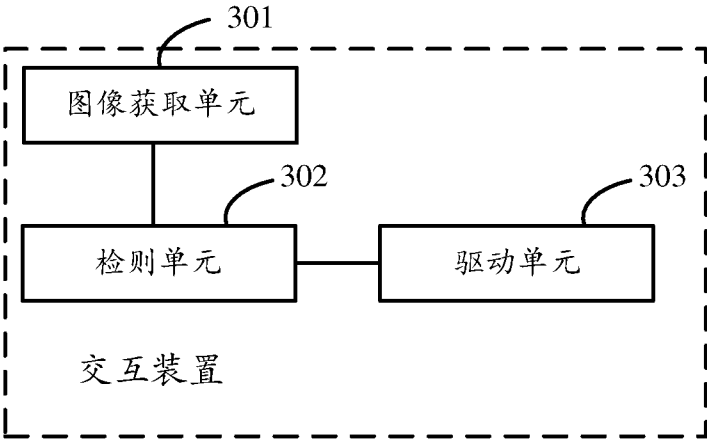


图 3

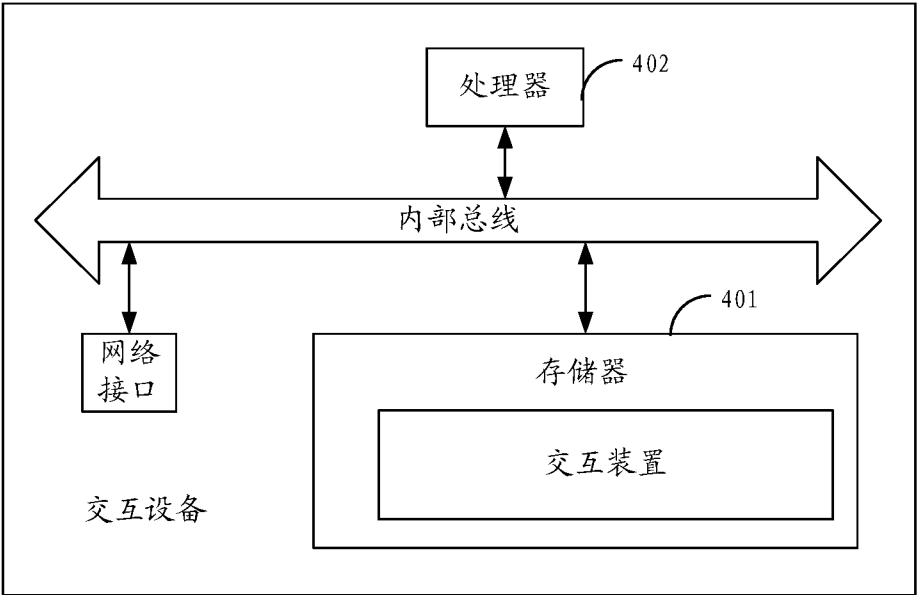


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/104291

A. CLASSIFICATION OF SUBJECT MATTER G06F 3/041(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) G06F 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, USTXT, EPTXT, WOTXT: 相机, 图像, 虚拟, 显示, camera, image, virtual, display		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110716641 A (BEIJING SENSETIME TECHNOLOGY DEVELOPMENT CO., LTD.) 21 January 2020 (2020-01-21) description, paragraphs 4-49	1-30
PX	CN 110716634 A (BEIJING SENSETIME TECHNOLOGY DEVELOPMENT CO., LTD.) 21 January 2020 (2020-01-21) description, paragraphs 4-34	1-30
PX	CN 111052042 A (APPLE INC.) 21 April 2020 (2020-04-21) description, paragraphs 53-90, figure 1A-4	1-30
E	CN 111602391 A (APPLE INC.) 28 August 2020 (2020-08-28) description, paragraphs 21-38, paragraphs 76-113, figures 1-8	1, 8, 15, 22, 29, 30
X	CN 108665744 A (WANG, Hongdong et al.) 16 October 2018 (2018-10-16) description, paragraphs 50-87, figures 1-7	1-30
X	CN 103513753 A (LENOVO (BEIJING) LIMITED) 15 January 2014 (2014-01-15) description, paragraphs 26-135, figures 1-14e	1-30
X	CN 105518582 A (BEIJING MEGVII TECHNOLOGY LIMITED et al.) 20 April 2016 (2016-04-20) description, paragraphs 25-248, figures 1-14	1-30
☒ 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 23 October 2020		Date of mailing of the international search report 29 October 2020
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) 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/CN2020/104291**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2012113140 A1 (MICROSOFT CORPORATION) 10 May 2012 (2012-05-10) description, paragraphs 17-86, figures 1-4	1-30
A	CN 105898346 A (LENOVO (BEIJING) LTD.) 24 August 2016 (2016-08-24) entire document	1-30
A	US 2014300634 A1 (SAMSUNG ELECTRONICS CO., LTD.) 09 October 2014 (2014-10-09) entire document	1-30

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/104291

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110716641	A	21 January 2020	None			
CN	110716634	A	21 January 2020	None			
CN	111052042	A	21 April 2020	KR	20200035103	A	01 April 2020
				WO	2019067902	A1	04 April 2019
				EP	3665550	A1	17 June 2020
CN	111602391	A	28 August 2020	WO	2019144013	A1	25 July 2019
CN	108665744	A	16 October 2018	None			
CN	103513753	A	15 January 2014	CN	103513753	B	27 June 2017
CN	105518582	A	20 April 2016	WO	2017000213	A1	05 January 2017
				US	2018211096	A1	26 July 2018
				CN	105518582	B	02 February 2018
				IN	201727045922	A	20 April 2018
US	2012113140	A1	10 May 2012	US	9529424	B2	27 December 2016
CN	105898346	A	24 August 2016	None			
US	2014300634	A1	09 October 2014	KR	102079097	B1	07 April 2020
				US	9972130	B2	15 May 2018
				KR	20140122126	A	17 October 2014

国际检索报告

国际申请号

PCT/CN2020/104291

A. 主题的分类 G06F 3/041 (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) G06F 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, USTXT, EPTXT, WOTXT: 相机, 图像, 虚拟, 显示, camera, image, virtual, display		
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
PX	CN 110716641 A (北京市商汤科技开发有限公司) 2020年 1月 21日 (2020 - 01 - 21) 说明书第4-49段	1-30
PX	CN 110716634 A (北京市商汤科技开发有限公司) 2020年 1月 21日 (2020 - 01 - 21) 说明书第4-34段	1-30
PX	CN 111052042 A (苹果公司) 2020年 4月 21日 (2020 - 04 - 21) 说明书第53-90段, 图1A-4	1-30
E	CN 111602391 A (苹果公司) 2020年 8月 28日 (2020 - 08 - 28) 说明书第21-38段、第76-113段、图1-8	1, 8, 15, 22, 29, 30
X	CN 108665744 A (王洪冬 等) 2018年 10月 16日 (2018 - 10 - 16) 说明书第50-87段, 图1-7	1-30
X	CN 103513753 A (联想北京有限公司) 2014年 1月 15日 (2014 - 01 - 15) 说明书第26-135段, 图1-14e	1-30
X	CN 105518582 A (北京旷视科技有限公司 等) 2016年 4月 20日 (2016 - 04 - 20) 说明书第25-248段, 图1-14	1-30
X	US 2012113140 A1 (MICROSOFT CORPORATION) 2012年 5月 10日 (2012 - 05 - 10) 说明书第17-86段, 图1-4	1-30
☒ 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 2020年 10月 23日		国际检索报告邮寄日期 2020年 10月 29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		授权官员 胡嫵 电话号码 86-(20)-28950387

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 105898346 A (联想北京有限公司) 2016年 8月 24日 (2016 - 08 - 24) 全文	1-30
A	US 2014300634 A1 (SAMSUNG ELECTRONICS CO LTD) 2014年 10月 9日 (2014 - 10 - 09) 全文	1-30

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/104291

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	110716641	A	2020年 1月 21日	无			
CN	110716634	A	2020年 1月 21日	无			
CN	111052042	A	2020年 4月 21日	KR	20200035103	A	2020年 4月 1日
				WO	2019067902	A1	2019年 4月 4日
				EP	3665550	A1	2020年 6月 17日
CN	111602391	A	2020年 8月 28日	WO	2019144013	A1	2019年 7月 25日
CN	108665744	A	2018年 10月 16日	无			
CN	103513753	A	2014年 1月 15日	CN	103513753	B	2017年 6月 27日
CN	105518582	A	2016年 4月 20日	WO	2017000213	A1	2017年 1月 5日
				US	2018211096	A1	2018年 7月 26日
				CN	105518582	B	2018年 2月 2日
				IN	201727045922	A	2018年 4月 20日
US	2012113140	A1	2012年 5月 10日	US	9529424	B2	2016年 12月 27日
CN	105898346	A	2016年 8月 24日	无			
US	2014300634	A1	2014年 10月 9日	KR	102079097	B1	2020年 4月 7日
				US	9972130	B2	2018年 5月 15日
				KR	20140122126	A	2014年 10月 17日