

一种用于移动式人机协作呼叫机器人的三方通话终端

本公开要求在 2020 年 7 月 13 日提交中国专利局、申请号为 202010669451.X、发明名称为“一种用于移动式人机协作呼叫机器人的三方通话终端”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本公开涉及人工智能领域，更具体地说，它涉及一种用于移动式人机协作呼叫机器人的三方通话终端。

背景技术

随着计算机技术、通信技术、互联网和人工智能技术的快速进步，各种智能家用电器不断走进人们的生活。例如，智能电视、智能冰箱、智能空调、智能音箱、智能手表、智能手环、智能眼镜等等。目前各种品牌的基于语音交互的智能设备已经大量上市，用户可以通过发出语音指令的方式与智能设备进行交互，实现听歌、报时、闲聊、游戏、陪伴、信息查询、设备控制等功能。但目前智能设备主要应用于家庭生活、休闲娱乐或儿童教育等领域，在企业级应用方面很少也很难得到应用。

随着人工智能技术和通信科技的蓬勃发展，电话机器人已广泛应用各行各业，大幅度降低了呼叫中心的人工成本，提高了效率。但目前的语音交互机器人特别是电话人机协作呼叫机器人系统涵盖了以人工智能技术和对话系统为主的人机协作呼叫机器人系统，和以通信网络和 VOIP 技术为主的语音通信系统。两者紧密绑定，系统异常复杂，开发、部署和维护难度都很大，成本很高。而替换其中任何一个组件都非常困难，很不灵活。基于此开发的电话机器人过于复杂和庞大，一旦部署就很难移动。这种电话机器人基于大规模云端服务器，没有一个普通人可以方便认知的实体，无法给人以直观亲切的印象。

要实现移动式的人机协作呼叫机器人，就需要一种支持三方通话并且方便易用的电话三方通话终端。传统的人机协作呼叫机器人一般使用台式机作为通话终端，操作复杂，不方便移动；而新式的移动电话三方通话终端并不支持三方通话，无法实现人机协作呼叫机器人。

发明内容

本公开的目的是提供一种用于移动式人机协作呼叫机器人的三方通话终端，能够与通信系统进行解耦，易于部署，便于切换，提供了可移动性，可以方便地放置在各种场合；便捷的接入个人手机或通话终端。

本公开的上述技术目的是通过以下技术方案得以实现的：
第一方面，本公开提供了一种用于移动式人机协作呼叫机器人的三方通话终端，包括：
第一语音接口，配置为连接至后端处理模块，并在通话对象与所述后端处理模块之间传输通话音频；其中，所述后端处理模块配置为通过预设规则与所述通话对象进行交互；

CODEC1 模块,配置为对所述通话对象与所述后端处理模块之间的通话音频进行编码和/或解码;

- 第二语音接口,配置为连接至人工坐席,并在所述通话对象与所述人工坐席之间传输通话音频;所述第二语音接口还配置为,将所述通话对象与所述后端处理模块之间的通话音频传输至所述人工坐席;

CODEC2 模块,配置为对所述通话对象与所述人工坐席之间的通话音频进行编码和/或解码;

通话控制模块,配置为处理控制信号,以及自动拨打、接听电话、挂断电话;

- 数据处理子模块,配置为处理语音数据以及与所述后端处理模块之间进行数据传递;

联网子模块,配置为与后端处理模块进行网络连接。

作为本公开的一种优选技术方案,所述三方通话终端还包括显示模块,所述显示模块配置为向所述人工坐席和所述通话对象显示所述通话对象与所述后台处理模块的通话记录或通话相关信息。

- 作为本公开的一种优选技术方案,所述三方通话终端还包括按键子模块,所述按键子模块用于输入控制指令。

作为本公开的一种优选技术方案,所述三方通话终端设置于音频设备内部,所述音频设备包括扬声器与麦克风,所述三方通话终端的所述第二语音接口连接至所述音频设备的所述扬声器与所述麦克风。

- 第二方面,本公开还提供了一种通信系统,所述通信系统包括:如第一方面所述的三方通话终端、后端处理模块、人工坐席、以及至少一个通信终端;其中,所述人工坐席通过所述通信终端与所述三方通话终端连接。

作为本公开的一种优选技术方案,所述后端处理模块用于对所述三方通话终端发来的语音数据进行处理并生成应答语音和文字发回给所述三方通话终端。

- 作为本公开的一种优选技术方案,所述后端处理模块包括对话管理子模块、语音识别子模块、意图识别子模块、语音合成子模块;

所述对话管理子模块用于控制对话的流程和逻辑,生成应答文本;

所述语音识别子模块用于识别接收到的通话对象语音并转化为文字;

所述意图识别子模块用于根据识别的语音文本识别出通话对象意图;

- 所述语音合成子模块用于将应答文本合成为语音并发送到所述三方通话终端。

第三方面,本公开还提供了一种通话方法,应用于如第二方面所述的通话系统,所述方法包括:

通过通信终端获取通话对象语音,并通过三方通话终端将所述通话对象语音传输至后端处理模块与所述人工坐席;

- 通过三方通话终端将应答语音传输至所述通信终端,通过所述通信终端将所述应答语音传输至所述通话对象;通过三方通话终端将应答语音和/或应答文字传输至所述人工坐席;

其中,所述应答语音与所述应答文字由所述后端处理模块根据预设规则以及所述通话对象语音生成。

作为本公开的一种优选技术方案，所述通过通信终端获取通话对象语音之前，所述方法还包括：

三方通话终端将根据业务逻辑编写的话术、通话对象的数据同步至后端处理模块；

后端处理模块在接收到话术和通话对象的数据后，开启通信终端与后端处理模块

5 之间的会话；

通过后端处理模块向三方通话终端发送指令，以使三方通话终端进入接听模式；

后端处理模块通过三方通话终端向通话终端发送开场白/欢迎词的语音和文本，以开启后端处理模块与通话对象之间的通话。

作为本公开的一种优选技术方案，所述三方通话终端将根据业务逻辑编写的话术、
10 通话对象的数据同步至后端处理模块之前，还包括：

三方通话终端根据管理员的操作或者事先计划的任务，加载任务清单；

三方通话终端根据任务清单，检索出对应的话术，所述话术代表一套完整的业务流程；

三方通话终端从通话对象数据库中查询通话对象数据。

15 作为本公开的一种优选技术方案，所述通过后端处理模块处理通话对象语音后生成机器人应答语音和文字包括：

后端处理模块根据通话对象语音分析通话对象的意图；

后端处理模块根据通话对象的意图、以及话术内置的策略和规则，生成应答句子文本；

20 后端处理模块根据应答句子文本确定机器人应答语音。

作为本公开的一种优选技术方案，所述后端处理模块根据通话对象语音分析通话对象的意图包括：

后端处理模块将通话对象语音转换成文本；

后端处理模块将文本进行分词，得到分词结果；

25 后端处理模块根据分词结果分析通话对象的意图。

作为本公开的一种优选技术方案，所述方法还包括：

通过三方通话终端显示后端处理模块和通话对象的通话记录或通话相关信息。

作为本公开的一种优选技术方案，所述方法还包括：

30 识别到通信终端与后端处理模块之间的对话结束，通过后端处理模块关闭通信终端与后端处理模块之间的会话；

后端处理模块将通话记录传输至三方通话终端，并保存于三方通话终端的数据库。

作为本公开的一种优选技术方案，所述方法还包括：

人工坐席向三方通话终端发送人工介入指令；

35 三方通话终端响应人工介入指令，切断与后端处理模块的连接，并转换为人工介入模式。

综上所述，本公开通过三方通话终端把通信系统（人工坐席和通信终端）与人机协作呼叫机器人系统（后端处理模块）进行解耦，降低了系统的复杂性，使之易于部署，便于灵活切换，可以大幅度降低电话人机协作呼叫机器人系统的开发、部署和维护成本；为机器人提供了可移动性，三方通话终端可以方便地放置在各种场合，也可以为机器人提供一

个直观、可触摸的实体，使之更具有亲和力；便捷的接入个人手机或者通话终端；提供蓝牙端、音频端口接入方式，应用范围广泛。

附图说明

图 1 是本发明的三方通话终端的模块框图；

5 图 2 是本发明的三方通话系统的结构示意图。

具体实施方式

以下结合附图对本发明作进一步详细说明。

如图 2 所示，本公开提供一种三方通话系统，包括：三方通话终端、后端处理模块、人工坐席、以及至少一个通信终端，如图 2 所示，三方通话系统包括一个通信终端，后端处理模块、人工坐席和通信终端在使用时，可以分别与三方通话系统连接，以通过三方通话系统在后端处理模块、人工坐席和通信终端之间传输语音数据。其中，通信终端是指通话对象所使用的终端设备，人工坐席是指对后端处理模块与通信终端的对话进行监控的终端设备，其用于监控后端处理模块与通信终端之间的语音数据，并可以人工介入，以取代后端处理模块与通信终端直接对话。

15 本公开可以采用如图 1 所示的三方通话终端，三方通话终端包括：第一语音接口、CODEC1 模块、第二语音接口、CODEC2 模块、通话控制模块、数据处理子模块和联网子模块。三方通话终端用于传输后端处理模块、人工坐席和通信终端之间的语音数据。

以下具体介绍三方通话终端、后端处理模块和通信终端的运行过程。

通信终端的运行过程为：

20 M1、连接三方通话终端；

M2、在开启会话后，接收通话对象的通话音频；

M3、将该通话音频通过第二语音接口输入三方通话终端，以通过三方通话终端将通话音频传输至后端处理模块和人工坐席；

M4、通过第二语音接口接收三方通话终端传输的机器人应答语音或者人工坐席语音。

25 三方通话终端中各模块的介绍如下：

第一语音接口，用于传输通话对象和后端处理模块的通话音频；

CODEC1 模块，用于通话对象和后端处理模块之间的通话语音音频编码、解码；

第二语音接口，用于传输人工坐席和通话对象的通话音频；

CODEC2 模块，用于人工坐席和通话对象的通话语音音频编码、解码；

30 第一语音接口和第二语音接口可以为蓝牙端或者音频端口，在通话过程中，人工坐席可以通过第二语音接口接收到通话音频，从而判断是否要进行人工介入，需要介入时，也从第二语音接口将人工坐席语音发给三方通话终端；

其中在音频端口可以设置一个或多个普通麦克风或者 MEMS 麦克风，也可以采用远场麦克风阵列，用于接收通信终端通过语音连接模块发来的通话对象语音，还可以设置扬声器等放音设备，用于播放机器人应答语音；如有必要，在不会导致回响、啸叫和混响的前提下，也可以通过扬声器播放通话对象的语音。

35 联网子模块，通过有线链路、WIFI 或 4G/5G 网络连接后端处理模块，发送/接收语音及其它数据；

数据处理子模块，用于调度控制三方通话终端中的其他模块，用于处理语音数据并发送到后端处理模块，处理来自后端处理模块的语音数据，并发送到通信终端，控制显示屏显示通话的文字记录，接收来自触摸屏的用户指令；

通话控制模块，用于控制通信系统和三方通话终端之间进行批量通话。

- 5 通话控制模块包括：通话对象数据库、系统数据库、任务管理模块、通信控制器子模块；

通话对象数据库，用于存储通话对象相关的数据；

系统数据库，用于存储通话记录及通话过程相关的其它数据；

任务管理模块，用于管理通话任务；

- 10 通信控制器子模块，用于调度其它模块，获取或存入数据；控制通信系统实现批量通话；

此外通话控制模块还包括话术编辑器和话术数据库，话术编辑器用于话术制作人员制作和修改话术，话术数据库用于存储话术制作人员通过话术编辑器制作的话术。

通话控制模块的工作流程如下：

- 15 根据管理员的操作或者事先计划好的任务，通过任务管理模块加载任务清单；

从话术数据库检索出任务需要的、由话术制作人员事先制作好的话术；其中一个话术就代表一套完整的业务流程，包括其对话规则、所有可能的应答句子文本、对通话对象的意向评价的规则等对话和业务相关数据，如果使用录音师录音，则还包括录音音频；

- 20 从通话对象数据库中查询出所需数据，例如电话号码、姓名、性别等，以及其它和业务相关的数据比如欠款金额之类；

通过网络把话术和通话对象数据同步到后端处理模块；

控制通信终端接通（主动拨打或被动接听）通话对象；

等待通话结束，从后端处理模块接收通话记录存储到数据库；

- 25 根据需要，以上流程可以批量执行；也可以并发执行，前提是有多个通信终端和三方通话终端，同时后端处理模块支持并发任务。

此外三方通话终端还包括显示屏和按键子模块，显示屏可以显示人机协作呼叫机器人系统和通话对象的通话记录或其他通话相关信息，也可以使用触摸屏，同时实现按键功能，让用户通过触摸的方式输入控制指令；

三方通话终端还可以设置蓝牙等无线通信系统，用于和语音连接模块进行通信；

- 30 三方通话终端还可以设置音频电路，使得音频的输入输出能够直接以数字形式进行；

三方通话终端还可以设置功率放大模块，用于放大来自语音连接模块的声音信号；

三方通话终端还可以设置语音降噪模块，对收到的音频信号作降噪处理；

三方通话终端还可以设置 AD/DA 转换芯片，用于把接收到的通话对象的语音转成数字信号传输，把接收到的机器人语音转成模拟信号通过扬声器播放；

- 35 三方通话终端还可以设置控制接口，包括按钮、旋钮等，用于外部控制。

三方通话终端的运行过程为：

A1、连接三方通话终端和通信终端；

A2、连接三方通话终端和后端处理模块；

A3、开启会话并接通通话对象；

A4、通话对象语音通过通信终端、三方通话终端传输到后端处理模块；

A5、后端处理模块处理通话对象语音后生成机器人应答语音和文字；

A6、机器人应答语音通过三方通话终端、通信终端传输给通话对象；

A7、机器人应答文字传输到三方通话终端并显示；

- 5 A8、人工坐席通过三方通话终端随时跟踪通话过程，必要时可以转入人工接听，实现人机协作呼叫功能。

后端处理模块用于对三方通话终端发来的语音数据进行意图识别，并根据语音意图生成回复语音发回给三方通话终端。后端处理模块包括对话管理子模块、语音识别子模块、意图识别子模块、语音合成子模块、分词子模块、声音分离子模块、声纹识别、会话管理子模块；后端处理模块部署在云端服务器上，通过有线或无线网络与三方通话终端通信。

- 10 子模块；后端处理模块部署在云端服务器上，通过有线或无线网络与三方通话终端通信。

对话管理子模块用于控制对话的流程和逻辑，生成应答文本；

语音识别子模块用于识别接收到的通话对象语音并转化为文字；

意图识别子模块用于根据识别的语音文本识别出通话对象意图；

语音合成子模块用于将应答文本合成为语音并发送到三方通话终端。

- 15 后端处理模块的对话系统运行过程如下：

S1、通话控制模块通过通信终端接通（主动拨打或被动接听）通话对象的电话后，根据业务逻辑编写的话术、通话对象的数据同步到后端处理模块的会话管理子模块和对话管理子模块；

S2、会话管理子模块开启一个新的会话；

- 20 S3、会话管理子模块向三方通话终端发送指令，使之进入接听模式；

S4、会话管理子模块向三方通话终端发送开场白/欢迎词的语音和文本；

S5、三方通话终端通过语音连接模块和通信系统把语音发送给通话对象，开启机器人和通话对象之间的通话；

- 25 S6、三方通话终端收到通话对象语音，通过网络发送到后端处理模块的语音识别子模块；

S7、语音识别子模块把通话对象语音转成文本，发送到意图识别子模块；

S8、意图识别子模块调用分词子模块先分词，再根据分词结果，结合话术识别出通话对象的意图，发送到对话管理子模块；

S9、可选地，意图识别也可以通过意图识别子模块直接由通话对象的语音得到；

- 30 S10、对话管理子模块根据话术内置的策略和规则，生成应答句子文本，发送到语音合成子模块；

S11、语音合成子模块把文本转换成机器人应答语音；可选地，机器人应答语音也可以事先由录音师录好，根据应答句子检索出来；

- 35 S12、把应答句子文本和语音一起发送到三方通话终端，由三方通话终端播放并通过语音连接模块和通信系统发送给通话对象；

S13、如此循环，直到对话结束；

S14、会话管理子模块关闭会话，把通话记录传输到通话控制模块，保存到系统数据库，供以后查询分析。

本公开把通信系统与人机协作呼叫机器人系统进行解耦，降低了系统的复杂性，使之

易于部署，便于灵活切换，可以大幅度降低电话人机协作呼叫机器人系统的开发、部署和维护成本；为机器人提供了可移动性，三方通话终端可以方便地放置在各种场合，也可以为机器人提供一个直观、可触摸的实体，使之更具有亲和力；便捷的接入个人手机或者通话终端；提供蓝牙端、音频端口接入方式，应用范围广泛。

- 5 同时，本公开中的三方通话终端延续了传统电话客服人机协作呼叫机器人系统功能：提供屏幕显示，可以方便的设置，调取通话记录和切换，使用更便捷；支持外放设备如头戴式耳机，随时进行通话跟踪以便人工坐席介入。

- 具体而言，现有技术中的电话机器人，由于其通信系统、人机协呼系统以及后台通话机器人系统均彼此绑定，故其部署方式仍采用类似于呼叫中心的部署方式，即在使用场合
10 提前部署相应的电话机器人系统，人工坐席或其它使用者需要在部署有该系统的场所工作，才能实现人机协呼的功能。较于上述现有技术，本公开中的三方通话终端的部署方式更为便捷。

- 在一种实现方式中，三方通话终端将现有技术中的通信系统交由使用者自身的通信终端、如手机、固定电话等，进而实现通信系统与人机协呼系统、后台通话机器人系统之间的解绑。以下通过一示例说明本实现方式中三方通话终端的部署方式：
15

在一示例中，某独立销售人员作为三方通话终端的使用者，其直接将三方通话终端通过有线或无线的方式接入至自身的通话终端（如手机）中，以此即完成部署，无需任何其它操作。因此，本实现方式中的三方通话终端在部署效率与便捷性上，较于现有技术得以显著改善。

- 20 实际使用过程中，上述示例中的销售人员一方面可通过音频线连接或蓝牙连接的方式将三方通话终端与自身手机之间进行连接，另一方面，三方通话终端通过连接在三方通话终端之上的耳麦进一步向销售人员提供音频数据（如三方通话终端搭载有扬声器/麦克风也可直接播放）。销售人员通过手机向某意向客户拨打电话（也可直接由三方通话终端自动向某意向用户拨打电话）后，三方通话终端将意向客户输入的语音上传至设置于云端
25 的后端处理模块，后端处理模块中的电话机器人根据规则产生对应的回应后，进一步由三方通话终端将回应语音返回至意向客户，以此实现后端处理模块与意向客户的交互。上述过程中，销售人员可通过耳麦随时监听后端处理模块与意向客户的交互过程，并在需要人工介入时，通过三方通话终端直接与意向客户进行交互。

- 由上述使用过程获知，本实现方式中的三方通话终端在使用过程中，无需提前布置人工坐席，使用者可以通过自身携带的通信终端随时接入三方通话终端，人工坐席可以根据实际需要，选用合适类型的通信终端，而并不限于固定的通信终端。
30

此外，本实现方式的三方通话终端在使用过程中，通信功能是由使用者自身的通信终端发起的，故其不必要将客户电话等敏感信息上传至电话机器人，进而避免了使用过程中由于电话机器人而导致信息泄露的可能。

- 35 在另一种实现方式中，三方通话终端可以直接部署在使用者自身的通信终端（音频设备）中，例如部署在手机、蓝牙耳机等，这些音频设备包括扬声器和麦克，并将三方通话终端中的第二语音接口与音频设备中的扬声器和麦克相连接，这样，该音频设备本身将具有支持三方通话的功能，此时，音频设备的使用者可以通过该音频设备接入人工坐席，并通过麦克向三方通话终端传输语音数据，通过扬声器接收三方通话终端发送的语音数据，

实现三方通话终端的整体可移动性。以下通过一示例说明本实现方式中三方通话终端的部署方式：

在一示例中，某销售人员作为三方通话终端的使用者，其可以直接使用部署有三方通话终端的音频设备（如手机），从而省去将三方通话终端接入自身的通话终端的过程，可以有效提高三方通话终端的部署效率与便捷性。在此基础上，由于音频设备具有较强的可移动性，因此，三方通话终端的使用地点不受限制，本实现方式中的三方通话终端在使用地点的灵活性上，较于现有技术得以显著改善。

实际使用过程中，上述示例中的销售人员可以在手机上开启三方通话功能，具体的，销售人员可以点击手机上的三方通话的应用软件等以启动三方通话终端，并开启三方通话功能，例如，启动三方通话终端后，手机通过联网功能，令三方通话终端连接云端的后端处理模块，并开启将手机与其它通信终端交互的语音数据传输至三方通话终端的功能。该销售人员通过手机向某意向客户拨打电话，可以通过手机接收意向客户发送的语音数据，并将该语音数据传输至三方通话终端，经由三方通话终端上传至后端处理模块。后端处理模块中的电话机器人根据规则产生对应的回应后，将应答数据回传至三方通话终端，三方通话终端通过销售人员的手机将该应答数据发送至意向客户，以此实现后端处理模块与意向客户的交互。上述过程中，销售人员可以通过音频设备的扬声器随时监听后端处理模块与意向客户的交互过程，并在需要人工介入时，通过三方通话终端与意向客户进行交互。

在另一示例中，某销售人员作为三方通话终端的使用者，三方通话终端集成于一蓝牙耳机中，销售人员佩戴该耳机后，即可实现三方通话功能。实际使用过程中，销售人员佩戴集成有三方通话终端的蓝牙耳机，可通过设置于蓝牙耳机之上的触发式按钮或通过手机上安装的应用程序以开启三方通话功能。开启三方通话功能后，该销售人员通过手机向某意向客户拨打电话，手机接收意向客户发送的语音数据后通过蓝牙传输至蓝牙耳机，并将该语音数据传输至三方通话终端，经由三方通话终端上传至后端处理模块。后端处理模块中的电话机器人根据规则产生对应的回应后，将应答数据回传至三方通话终端，三方通话终端通过蓝牙耳机中的麦克将该应答数据发送至意向客户，以此实现后端处理模块与意向客户的交互。上述过程中，销售人员可以通过蓝牙耳机的扬声器随时监听后端处理模块与意向客户的交互过程，并在需要人工介入时，通过三方通话终端与意向客户进行交互。

由上述使用过程可知，本实现方式中的三方通话终端与通信终端之间的集成度更高，使用者可以将三方通话终端随身携带并在任何场景使用，并不限定于固定的使用场景。

本公开通过显示模块提供屏幕显示功能，具体的，三方通话终端通过显示模块显示通话对象的会话内容和机器人的应答内容，其中，在人工坐席的显示屏与三方通话终端的显示屏不同时，通过显示模块推送至人工坐席，以人工坐席的显示屏显示通话对象的会话内容和机器人的应答内容。

具体而言，在一种实现方式中，人工坐席与通话对象同时观看三方通话终端的显示屏，此时，可以通过三方通话终端的显示模块（相当于显示屏）直接显示通话对象的会话内容和机器人的应答内容，以便于通话对象与人工坐席同时掌握通信对象与电话机器人之间的互动过程，便于通话对象与人工坐席之间的交互数据的同步，以便于人工坐席可以及时介入通话，解决电话机器人所无法解决的问题。

在另一种实现方式中，人工坐席可以通过自带的通信终端连接三方通话终端（可以参

考上文三方通话终端部署方式的第一种实现方式)，此时，如果人工坐席所使用的通信终端自带显示屏，则人工坐席可以该自带的显示屏浏览通话对象的会话内容和机器人的应答内容。这样，即使人工坐席使用自带的通信终端，也可以随时掌握通话对象与机器人之间的交互过程。进一步地，后端处理模块可以将通话对象的会话内容和机器人的应答内容进行进一步处理，以使得在三方通话终端上所显示的第一内容和在人工坐席的显示屏上所显示的第二内容不同。例如，人工坐席想要掌握详细的通话对象的会话内容和机器人的应答内容，则后端处理模块将通话对象的会话内容和机器人的应答内容处理为第二内容，并通过显示模块发送给人工坐席，以在人工坐席的自带的显示屏上显示。而通话对象仅想要浏览机器人的应答内容，以便于快速获取所需要的应答信息，则后端处理模块将机器人的应答内容处理为第一内容，并通过显示模块显示在三方通话终端上。这种显示方式具有更强的针对性，以满足通话对象与人工坐席不同的需求。

本公开可以用于实现人机协作呼叫电话机器人，通话系统的运行过程如下：

D1、把三方通话终端装置连接到通信终端设备，连接方式可为蓝牙或 3.5mm 音频接口；

D2、人工坐席接入三方通话终端，接入方式可为蓝牙或 3.5mm 音频接口；

15 D3、三方通话终端连接后端处理模块，连接方式可以为以太网、WIFI、4G 或者 5G；

D4、控制通信终端拨通或者接听通话对象的电话；

D5、会话管理子模块开启会话，向三方通话终端发送指令使它进入等待通话模式，同时把开场白语音/文字、通话对象相关数据等必要的的数据发送给它；

20 D6、三方通话终端检测到电话接通信号，通过通信终端向通话对象发送开场白语音，在显示屏显示相关文字信息；

D7、三方通话终端接收到通话对象的语音，发送到后端处理模块；

D8、后端处理模块调用语音识别子模块把语音转换成文字，再通过意图识别子模块识别出通话对象的意图，经过对话管理模块判断决策后，生成机器人答复文字和语音。可选地，也可以通过意图识别子模块直接把通话对象的语音识别为通话对象的意图；

25 D9、机器人答复文字和语音经由联网子模块发送到三方通话终端；

D10、机器人答复语音经三方通话终端中的数据处理子模块处理后播放出来，同时通过通信终端发送给通话对象；

D12、答复文字经三方通话终端中的数据处理子模块处理后，以通话记录的形式显示在显示屏上；

30 D13、人工坐席可以通过第二语音接口收听机器人语音和通话对象的语音，或者通过观看屏幕文字，了解通话进程；

D14、在人工坐席认为必要的时候，按下三方通话终端上设置的介入按钮开始人工介入；

35 D15、三方通话终端内数据处理子模块器把会话转为人工介入模式（对通信系统本身没有任何影响），切断与后端处理模块的连接；

D16、人工坐席直接对着三方通话终端内的第二语音接口讲话，通过通信终端发送到通话对象，直接和通话对象对话，实现无缝切换。机器人语音和坐席语音已经通过语音合成子模块进行匹配，所以能够较好地模拟坐席的声音；

D17、人工坐席和通话对象的对话也可以通过后端处理模块的语音识别子模块识别成

文字，显示在显示屏；

D18、关闭会话时，保存通话记录等相关数据到数据库。

- 5 以上所述仅是本发明的优选实施方式，本发明的保护范围并不局限于上述实施例，凡属于本发明思路下的技术方案均属于本发明的保护范围。应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明原理前提下的若干改进和润饰，这些改进和润饰也应视为本发明的保护范围。