

(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 1 月 31 日 (31.01.2019)



(10) 国际公布号
WO 2019/019935 A1

(51) 国际专利分类号:
G06F 17/30 (2006.01)

(21) 国际申请号: PCT/CN2018/095836

(22) 国际申请日: 2018 年 7 月 16 日 (16.07.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710620007.7 2017 年 7 月 26 日 (26.07.2017) 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: 孟宾宾(MENG, Binbin); 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)。 陈志博(CHEN, Zhibo); 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 518000 (CN)。

(74) 代理人: 北京康信知识产权代理有限公司 (KANGXIN PARTNERS, P.C.); 中国北京市海淀区知春路甲 48 号盈都大厦 A 座 16 层, Beijing 100098 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,

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

(54) 发明名称: 交互方法、交互终端、存储介质和计算机设备

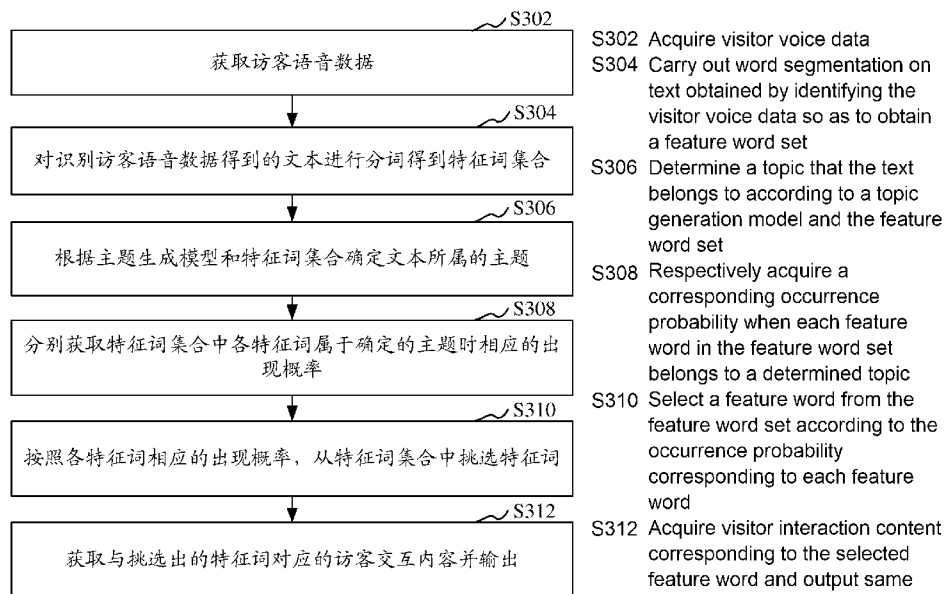


图 3

(57) Abstract: The present invention relates to an interaction method, an interaction terminal, a storage medium, and a computer device. The method comprises: acquiring visitor voice data; carrying out word segmentation on text obtained by identifying the visitor voice data so as to obtain a feature word set; determining a topic that the text belongs to according to a topic generation model and the feature word set; respectively acquiring a corresponding occurrence probability when each feature word in the feature word set belongs to a determined topic; selecting a feature word from the feature word set according to the occurrence probability corresponding to each feature word; and acquiring visitor interaction content corresponding to the selected feature word and outputting same. The solution

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))。

provided in the present application improves the efficiency of interaction with visitors.

(57) 摘要: 本发明涉及一种交互方法、交互终端、存储介质和计算机设备, 所述方法包括: 获取访客语音数据; 对识别所述访客语音数据得到的文本进行分词得到特征词集合; 根据主题生成模型和所述特征词集合确定所述文本所属的主题; 分别获取所述特征词集合中各特征词属于确定的主题时相应的出现概率; 按照各所述特征词相应的出现概率, 从所述特征词集合中挑选特征词; 获取与挑选出的特征词对应的访客交互内容并输出。本申请提供的方案提高了与访客间的交互效率。

交互方法、交互终端、存储介质和计算机设备

本申请要求于 2017 年 07 月 26 日提交中国专利局、申请号为 201710620007.7、发明名称为“交互方法、存储介质和计算机设备”的中国专利
5 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

本公开实施例涉及计算机技术领域，特别是涉及一种交互方法、交互终端、
存储介质和计算机设备。

10

背景技术

随着社会的进步和科技的发展，人们之间的交互越来越频繁，需要进行交互的场景也越来越多。在传统的访客来访场景中，对于来访人员，通常需要由
工作人员通过人工方式获知访客来访意图，再根据访客来访意图与访客进行后
15 续交互。

然而，传统的这种通过人工方式与访客进行交互的交互方式，在交互过程中需要消耗大量的人力物力，而且引入了大量的工作量，导致与访客交互的效率低下。

20 发明内容

基于此，有必要针对传统的交互方式导致的与访客交互的效率低下的问题，
提供一种交互方法、交互终端、存储介质和计算机设备。

一种交互方法，所述方法包括：

获取访客语音数据；

25 对识别所述访客语音数据得到的文本进行分词得到特征词集合；

根据主题生成模型和所述特征词集合确定所述文本所属的主题；

分别获取所述特征词集合中各特征词属于确定的主题时相应的出现概率；

按照各所述特征词相应的出现概率，从所述特征词集合中挑选特征词；

获取与挑选出的特征词对应的访客交互内容并输出。

一种交互终端，

包括：一个或者一个以上处理器、一个或一个以上存储器，所述存储器中
5 存储有至少一段应用程序，所述至少一段应用程序适于处理器进行加载，以执行以下操作：

获取访客语音数据；

对识别所述访客语音数据得到的文本进行分词得到特征词集合；

根据主题生成模型和所述特征词集合确定所述文本所属的主题；

10 分别获取所述特征词集合中各特征词属于确定的主题时相应的出现概率；

按照各所述特征词相应的出现概率，从所述特征词集合中挑选特征词；

获取与挑选出的特征词对应的访客交互内容并输出。

一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读
15 指令，所述计算机可读指令被处理器执行时，使得所述处理器执行上述交互方法的步骤。

一种计算机设备，包括存储器和处理器，所述存储器中储存有计算机可读
指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行上述交
20 互方法的步骤。

上述交互方法、交互终端、存储介质和计算机设备，在获取访客语音数据，
并对识别该访客语音数据得到的文本进行分词得到特征词集合后，即可利用主
题生成模型和分词得到的特征词集合，在访客语音数据稀疏度高的情况下准确
25 地确定访客语音数据所属主题，并自动根据特征词集合中各特征词属于确定的
主题时的出现概率，选取合适的特征词所对应的访客交互内容输出，避免了需
要大量的人工参与而引入的工作量，提高了与访客间的交互效率及输出的访客
交互内容的准确性。

附图说明

图 1 为一个实施例中交互方法的应用环境图；

图 2 为一个实施例中计算机设备的内部结构示意图；

5 图 3 为一个实施例中交互方法的流程示意图；

图 4 为一个实施例中根据主题生成模型和特征词集合确定文本所属的主题的步骤的流程图；

图 5 为一个实施例中主题生成模型中生成文本过程的图模型示意图；

图 6 为一个实施例中地图创建过程的流程示意图；

10 图 7 为一个实施例中在地图中选取趋向访客移动路径的示意图；

图 8 为一个实施例中人脸检测的步骤的流程示意图；

图 9 为一个实施例中对面脸图像进行人脸识别的示意图；

图 10 为另一个实施例中交互方法的流程示意图；

图 11 为一个实施例中交互机器人的架构图；

15 图 12 为一个实施例中交互终端的结构框图；

图 13 为另一个实施例中交互终端的结构框图。

具体实施方式

为了使本实施例的目的、技术方案及优点更加清楚明白，以下结合附图及
20 实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例
仅仅用以解释本发明，并不用于限定本发明。

图1为一个实施例中交互方法的应用环境图。如图1所示，该应用环境包括
交互机器人110和访客120。交互机器人110可通过执行交互方法，与访客120进
行交互。本领域技术人员可以理解，图1中示出的应用环境，仅仅是与本申请方
25 案相关的部分场景，并不构成对本申请方案应用环境的限定，该应用环境可以
应用于室内场景或者室外开阔场景中。

图2为一个实施例中计算机设备的内部结构示意图。如图2所示，该计算机
设备包括通过系统总线连接的处理器、非易失性存储介质、内存储器、摄像头、

声音获取装置、扬声器、显示屏、输入装置和移动装置。其中，计算机设备的非易失性存储介质存储有操作系统，还可存储有计算机可读指令，该计算机可读指令被处理器执行时，可使得处理器实现一种交互方法。该处理器用于提供计算和控制能力，支撑整个计算机设备的运行。该内存储器中也可储存有计算机可读指令，该计算机可读指令被所述处理器执行时，可使得所述处理器执行一种交互方法。计算机设备的显示屏可以是液晶显示屏或者电子墨水显示屏等，输入装置可以是显示屏上覆盖的触摸层，也可以是终端外壳上设置的按键、轨迹球或触控板，也可以是外接的键盘、触控板或鼠标等。该计算机设备是可移动的电子设备，具体可以是如图1中所述的交互机器人110等。该计算机设备还可通过网络与服务器连接，将数据发送至服务器进行处理。比如，将获取的图像帧发送至服务器进行人脸检测等。本领域技术人员可以理解，图2中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的终端的限定，具体的终端可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

如图3所示，在一个实施例中，提供了一种交互方法。本实施例主要以该方法应用于上述图1中的交互机器人来举例说明。参照图3，该交互方法具体包括如下步骤：

S302，获取访客语音数据。

其中，访客语音数据是反映访客访问意图的语音数据。访客语音数据可以是现场采集的语音数据，也可以是事先录制的语音数据，还可以是接收的其他计算机设备传输的语音数据。

在一个实施例中，交互机器人可以直接调用本地的声音获取装置采集环境声音，检测采集的环境声音中是否包括访客语音数据，从而在检测到环境声音中包括访客语音数据时，获取到访客语音数据。

在一个实施例中，交互机器人也可以提供用于进行语音输入的操作入口，检测访客对操作入口的触发操作，根据检测到的触发操作触发相应的语音采集指令，根据该语音采集指令调用本地的声音获取装置采集访客输入的语音数据，从而获取到访客语音数据。其中，操作入口可以是图标、文字链接或者虚拟按

钮等。

在一个实施例中，交互机器人还可以与其他计算机设备建立连接，接收其他计算机设备传输的语音数据，从而获取到访客语音数据。

S304，对识别访客语音数据得到的文本进行分词得到特征词集合。

5 其中，分词是指将一个连续的字符序列切分成多个单独的字符或者字符序列。特征词是指具有语义表达功能的字符或者字符序列。

具体地，交互机器人可对访客语音数据进行特征提取，获得待识别的访客语音特征数据，然后基于声学模型对待识别的访客语音特征数据进行语音分帧处理得到多个音素，根据候选字库中候选字与音素的对应关系，将处理得到的
10 多个音素转化为字符序列，再利用语言模型调整转化得到的字符序列，从而得到符合自然语言模式的文本。

其中，文本是语音数据的字符表示形式。声学模型如 GMM(Gaussian Mixture Model 高斯混合模型)或 DNN(Deep Neural Network 深度神经网络)等。候选字库包括候选字和与候选字对应的音素。语言模型用于按照自然语言模式调整声学
15 模型所识别出的字符序列，比如 N-Gram 模型(CLM, Chinese Language Model 汉语语言模型)等。

进一步地，交互机器人可采用预设的分词方式对识别访客语音数据得到的文本进行分词处理，得到多个字符或者字符序列，从得到的字符序列中筛选出具有实际语义的字符或者字符序列作为特征词，形成特征词集合。特征词集合
20 可以包括一个或多个特征词。其中，预设的分词方式可以是基于字符匹配、基于语义理解或者基于统计的分词方式。

在一个实施例中，交互机器人可采用基于字符匹配的分词方式进行分词处理，将识别访客语音数据得到的文本按照从前到后或者从后到前的顺序逐一切分出单个字符，再将该单个字符与标准词库进行匹配。如果匹配成功，则获取
25 该字符作为一个候选特征词；若匹配失败，则通过增加一个字符继续进行匹配，直至识别访客语音数据得到的文本中包括的字符全部匹配完成。

在一个实施例中，交互机器人也可同时对识别访客语音数据得到的文本进行正向匹配分词和逆向匹配分词。在两种分词方式的分词结果相同时，将分词

得到的多个单独的字符或者字符序列作为候选特征词。在两种分词方式的分词结果不相同，分别计算两种分词方式得到的单独的字符或者字符序列的数量，选取计算的数量少的分词方式得到的单独的字符或者字符序列作为候选特征词。

- 5 进一步地，交互机器人从得到的字符或字符序列中筛选出具有实际语义的字符或者字符序列作为特征词时，具体可从得到的字符或字符序列中过滤掉停用词。其中，停用词是指自然语言中包括的一种功能字符或者字符序列，这类功能字符或者字符序列并无实际语义，包括代表语气的语气字符或字符序列和表示某种逻辑关系连接字符或字符序列等。具体地，语气字符比如“吗”或者
- 10 “呢”等，连接字符比如“的”或“在”等，语气字符序列比如“而已”或者“就是了”等，连接字符序列比如“至于”或“然后”等。

S306，根据主题生成模型和特征词集合确定文本所属的主题。

- 在自然语言处理中，主题生成模型是指一篇文章中的每个词都是按照一定概率选择了某个主题并从选择的主题中选择了某个词这样一个过程得到的。在
- 15 自然语言中主题表示的是若干词所表达的一个概念，在数学上表示为词汇表上词语的条件概率分布。

基于主题生成模型，存在以下公式：

$$p(\text{word}|\text{doc}) = \sum_{\text{topic}} p(\text{word}|\text{topic}) \times p(\text{topic}|\text{doc}) \quad (1)$$

- 其中，*word* 表示特征词，*doc* 表示分词得到特征词集合的文本，*topic* 表示
- 20 主题， $p(\text{word}|\text{doc})$ 表示文本中每个特征词出现的概率， $p(\text{word}|\text{topic})$ 表示各个主题中的特征词出现的概率， $p(\text{topic}|\text{doc})$ 则表示各个文本中的主题出现的概率。

- 本实施例中，上述公式（1）等式左边的 $p(\text{word}|\text{doc})$ 是已知的，可直接通过统计根据文本分词得到的特征词集合中各个特征词的词频而获得，而 $p(\text{word}|\text{topic})$ 和 $p(\text{topic}|\text{doc})$ 是未知的，主题生成模型就是用大量已知的
- 25 $p(\text{word}|\text{doc})$ 经过一系列训练，推理出 $p(\text{word}|\text{topic})$ 和 $p(\text{topic}|\text{doc})$ 。获得 $p(\text{topic}|\text{doc})$ 后，便可以将使得 $p(\text{topic}|\text{doc})$ 最大的主题作为访客语音数据所属的主题。

主题生成模型可采用 pLSA(Probabilistic Latent Semantic Analysis)或者 LDA

(Latent Dirichlet Allocation), pLSA 主题生成模型训练推理主要使用的是 EM(期望最大化) 算法; LDA 主题生成模型训练推理采用的是 Gibbs sampling (吉布斯采样) 方法。

S308, 分别获取特征词集合中各特征词属于确定的主题时相应的出现概率。

5 其中, 特征词属于确定的主题时相应的出现概率, 也就是特征词在确定的主题的条件下出现的频次。特征词属于确定的主题时相应的出现概率反映的是特征词与所属的主题的相关度。特征词与所属的主题的相关度与特征词属于该主题时相应的出现概率正相关, 出现概率越大说明特征词与所属的主题越相关。

具体地, 交互机器人在训练主题生成模型时, 可事先确定预设数量的主题,
10 再为每个主题获取大量与该主题相关的文本作为训练语料, 并对训练语料进行分词处理, 将所有训练语料中分词后获得的词构成词集合, 以进行主题生成模型训练。其中, 作为训练语料的文本可以是互联网上爬取的文本。

进一步地, 交互机器人在对识别访客语音数据得到的文本进行分词得到特征词集合, 并确定文本所属的主题后, 对于特征词集合中的每个特征词, 统计
15 当前特征词在确定的主题对应的训练语料中出现的次数, 进而根据以下公式计算词频:

$$TF = \frac{N_T}{N_{sum}} \quad (2)$$

其中, TF 表示词频, N_T 表示特征词在确定的主题对应的训练语料中出现的次数, N_{sum} 则表示确定的主题对应的训练语料中出现的词的总数量。

20 S310, 按照各特征词相应的出现概率, 从特征词集合中挑选特征词。

具体地, 对于特征词集合中包括的多个特征词, 交互机器人可选取特征词集合包括的全部特征词以进行后续数据处理, 也可从特征词集合选取部分特征词以进行后续数据处理。

在一个实施例中, 交互机器人可选取出现概率超过预设出现概率域值的特征词。
25 交互机器人具体可在确定各特征词相应的出现概率后, 将确定的各出现概率与预设出现概率域值进行比较, 挑选出相应的出现概率超过预设出现概率域值的特征词。交互机器人也可以将各特征词按照相应的出现概率降序排序,

选取排在靠前的预设比例的特征词。

S312, 获取与挑选出的特征词对应的访客交互内容并输出。

其中, 访客交互内容是与访客进行交互的内容。访客交互内容可以包括文本、图片、音频或者视频中的至少一种。访客交互内容可以是根据特征词统一
5 设置的内容, 也可以是与访客身份相关的内容。

特征词与访客交互内容关联, 用于标记访客交互内容。通过特征词可定位到相关联的访客交互内容。一个特征词可以关联一个或者多个访客交互内容。多个特征词也可以关联一个访客交互内容。

在一个实施例中, 交互机器人可事先设置访客交互内容, 并将访客交互内
10 容与特征词关联, 再将设置的访客交互内容存储在数据库或者文件中, 在需要时从数据库或者文件中读取。交互机器人在挑选出特征词后, 可拉取该特征词所关联的访客交互内容。该访客交互内容可以是直接可输出的访客交互内容, 也可以是待补全的访客交互内容模板。

在一个实施例中, 访客交互设备可调用显示屏输出访客交互内容。访客交
15 互内容比如文本、图片或者视频等。访客交互设备还可以获取访客交互内容对应的样式数据, 从而按照该样式数据, 在显示屏中展示访客交互内容。访客交互设备也可以调用扬声器输出访客交互内容。访客交互内容比如音频等。

上述交互方法, 在获取访客语音数据, 并对识别该访客语音数据得到的文本进行分词得到特征词集合后, 即可利用主题生成模型和分词得到的特征词集
20 合, 在访客语音数据稀疏度高的情况下准确地确定访客语音数据所属主题, 并自动根据特征词集合中各特征词属于确定的主题时的出现概率, 选取合适的特征词所对应的访客交互内容输出, 避免了需要大量的人工参与而引入的工作量, 提高了与访客间的交互效率及输出的访客交互内容的准确性。

如图 4 所示, 在一个实施例中, 步骤 S306 具体包括如下步骤:

25 S402, 根据主题生成模型, 对于特征词集合中的每个特征词, 获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数。

其中, 主题生成模型的主要功能是提取可以理解、相对稳定的潜在语义结构。主题模型采用了词袋模型, 将每一篇文本视为一个词频向量, 从而将文本

信息转化为易于建模的数字信息。主题生成模型的主要前提假设是，一系列的主题链接了一系列词和文本集合，主题则可以视为字和词的一种概率分布。本实施例采用 LDA 主题生成模型来确定访客语音数据所属的主题。

在一个实施例中，根据主题生成模型，对于特征词集合中的每个特征词，
5 获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数，包括：根据主题生成模型，对于特征词集合中的每个特征词，按照预设数量的主题与特征词的多项式分布选择当前特征词所属的主题，并按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数。

10 具体地，交互机器人可事先选取预设数量的主题，并获取与选取的主题对应的多篇文本构成训练语料库，初始化预设数量的主题与特征词的多项式分布以及预设数量的主题与文本的多项式分布，根据训练语料库中的训练语料优化这两个多项式分布。

对于训练语料库中的每篇文本，LDA 主题生成模型定义了如下生成过程：

15 给定包括 M 篇文本的训练语料库，训练语料库中的每一篇文本与 K (K 通过反复试验等方法事先给定) 个主题的一个多项式分布相对应，将该多项分布记为 ϕ 。每个主题又与词汇表中的 V 个单词的一个多项式分布相对应，将这个多项分布记为 φ 。 ϕ 和 φ 分别是带有超参数 α 和 β 的 Dirichlet 先验分布。对于一篇文本 m 中的每一个词 w ，从该文本 m 所对应的多项分布 ϕ 中抽取一个主题 z ，然后再从主题 z 所对应的多项分布 φ 中抽取一个词 w ，将这个过程重复 N_m 次，就产生了
20 文本 m ，这里的 N_m 是文本 m 中的词总数。这个生成过程可以用图 5 所示的图模型表示。其中 $m \in [1, M]$ ， $k \in [1, K]$ 。

在一个实施例中，交互机器人可根据主题生成模型，对于特征词集合中的每个特征词，按照预设数量的主题与特征词的多项式分布选择当前特征词所属
25 的主题，并且按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建特征词生成概率函数。其中，预设数量的主题与特征词的多项式分布 $\varphi_k \sim \text{Dirichlet}(\beta)$ ， $\varphi_k \sim \text{Dirichlet}(\beta)$ 表示 φ_k 服从以 β 为超参数的 Dirichlet 分布。文本与主题的多项式分布 $\phi_m \sim \text{Dirichlet}(\alpha)$ 。其中 $\phi_m \sim \text{Dirichlet}(\alpha)$ 表示 ϕ_m 服从以 α 为

超参数的 Dirichlet 分布。

进一步地，交互机器人可根据分布 $\phi_m \sim \text{Dirichlet}(\alpha)$ 生成当前的特征词 n 所属的主题 $z_{m,n} \sim \phi_m$ 。根据分布 $\phi_k \sim \text{Dirichlet}(\beta)$ ，生成当前特征词 $w_{m,n} \sim \phi_{z_{m,n}}$ 。可知，第 m 个文本中第 n 个特征词 t 的生成概率可用如下公式表示：

$$p(w_{m,n} = t | \phi, \phi) = \sum_{k=1}^K p(w_{m,n} = t | \phi_k) p(z_{m,n} = k | \phi_m) \quad (3)$$

其中， m 是文本的标识， n 是根据文本分词得到的特征词集合中的特征词的标识， t 表示文本 m 中的特征词 n 的值， k 表示主题的标识， K 是主题的预设数量； $w_{m,n}$ 表示文本 m 中的特征词 n ， $z_{m,n}$ 表示文本 m 中第 n 个特征词的主题； $p(w_{m,n} = t | \phi, \phi)$ 表示特征词 $w_{m,n}$ 是 t 的概率， $p(z_{m,n} = k | \phi_m)$ 表示在 ϕ_m 的条件下当前特征词所属的主题 $z_{m,n}$ 是 k 的概率， $p(w_{m,n} = t | \phi_k)$ 表示在 ϕ_k 的条件下当前特征词 $w_{m,n}$ 是 t 的概率。 t 比如：天气、晴朗或者游戏等实际的字符或字符串。

在上述实施例中，通过处理自然语言的主题生成模型中训练得到的主题与特征词的关系，及文本与主题的关系，得到访客语音数据包括的特征词属于各主题时相应的特征词生成概率函数，从而可通过概率函数来预测访客语音数据所属的主题，进而合理地概率学的角度上确定访客所要表达的主题。

S404，分别确定特征词集合中各特征词属于各主题时相应的出现概率。

具体地，上述公式 (3) 等式左边的 $p(w_{m,n} = t | \phi, \phi)$ 等于当前特征词 t 在文本 m 中的出现概率，也就是当前特征词 t 在文本 m 中的词频。

S406，对于预设数量的主题中的每个主题，根据属于当前主题的特征词的出现概率和当前主题相应的特征词生成概率函数，获得文本属于当前主题的概率。

具体地， M 个文本是已知的，其中的各个特征词的出现概率也是已知的，可通过吉布斯采样法训练获得 $p(z_{m,n} = k | \phi_m)$ 以及 $p(w_{m,n} = t | \phi_k)$ 的值，再通过 $p(w_{m,n} = t | \phi_k)$ 便可以获得识别访客语音数据得到的文本属于主题 k 的概率。

S408，将对应最大概率的主题确定为文本所属的主题。

具体地，交互机器人可将识别访客语音数据得到的文本属于预设数量的主题中各主题的概率按照降序排序，从而得到概率最大的概率对应的主题，将该

主题确定为识别访客语音数据得到的文本所属的主题。

在其它实施例中，还可以选取概率较大的多个主题作为识别访客语音数据得到的文本所述的主题。

本实施例中，将处理自然语言的主题生成模型运用到对访客语音数据进行
5 访客意图定位，确定访客所要表达的主题后，便可以基于访客表达的主题进行
候选的精确交互，提高了与访客进行交互的准确性。

在一个实施例中，步骤 S306 包括：根据主题生成模型和特征词集合确定文
本所属的主题，以及文本属于确定的主题的概率。步骤 S310 包括：对于特征词
集合中的每个特征词，根据文本属于确定的主题的概率与当前特征词属于确定
10 的主题的出现概率，得到当前特征词对应的主题影响力值；按照各特征词对应
的主题影响力值，从特征词集合中挑选特征词。

其中，主题影响力值是特征词与主题间相关程度的量化数值。在某一主题
条件下，特征词对应的主题影响力值越大，说明特征词与该主题的相关程度越
高。特征词对应的主题影响力值与文本属于确定的主题的概率正相关，且与该
15 特征词属于确定的主题的出现概率正相关。

具体地，交互机器人可直接将文本属于确定的主题的概率与当前特征词属
于确定的主题的出现概率的乘积，作为当前特征词对应的主题影响力值；也可
将文本属于确定的主题的概率与当前特征词属于确定的主题的出现概率进行运
算后得到的数值，作为公众标识相对应的公众影响力值。

20 在一个实施例中，特征词对应的主题影响力值，具体可以由自变量为文本
属于确定的主题的概率，与当前特征词属于确定的主题的出现概率的二元正相
关函数计算出。正相关函数的正相关系数可以是随机设定的数值。

在一个实施例中，交互机器人可选取主题影响力值超过预设主题影响力值
域值的特征词。交互机器人具体可在确定各特征词相应的主题影响力值后，将
25 确定的各主题影响力值与预设主题影响力值域值进行比较，挑选出相应的主题
影响力值超过预设主题影响力值域值的特征词。交互机器人也可以将各特征词
按照相应的主题影响力值降序排序，选取排在靠前的预设比例的特征词。

上述实施例中，在进行特征词挑选时，将文本属于确定的主题的概率与当

前特征词属于确定的主题的出现概率，作为特征词对应的主题影响力值的评估依据，使得对特征词对应的主题影响力值的评估更加科学，从而挑选出更能反映访客意图的特征词，从而提高与访客的交互效率与准确性。

5 在一个实施例中，步骤 S312 包括：从本地查找与挑选出的特征词对应的离线访客交互内容；将挑选出的特征词发送至服务器；接收服务器返回的根据挑选出的特征词实时生成的访客交互内容；将查找到的离线访客交互内容与接收到的访客交互内容取交集或并集后输出。

具体地，交互机器人可针对访客交互场景选取相适配的关键词，并针对这些关键词查找相应的语料，根据这些语料生成个各访客交互场景中各关键词相应的访客交互内容，构建离线访客交互内容库。交互机器人在挑选出代表访客意图的特征词后，可将该特征词与离线访客交互内容库中的各关键词比较，获取与该特征词相匹配的关键词所对应的访客交互内容。交互机器人可直接将获取的离线访客交互内容输出。

交互机器人也可在挑选出的特征词后，将挑选出的特征词发送至服务器。服务器在获取到挑选出的特征词后，在互联网上实时爬取与该特征词相关的语料，从该语料中提取目标内容生成访客交互内容，再反馈给交互机器人，交互机器人可直接将服务器返回的访客交互内容输出。

交互机器人还可在获取离线访客交互内容与接收到服务器返回的访客交互内容后，将查找到的离线访客交互内容与接收到的访客交互内容相融合，再将融合得到的访客交互内容输出。其中，将查找到的离线访客交互内容与接收到的访客交互内容相融合可以是直接将两者取交集，也可以是将两者取并集。

在本实施例中，提供了离线与在线两种方式获取访客交互内容输出。通过离线方式获取的访客交互内容可极大地提高交互机器人的响应效率，通过在线方式获取访客交互内容，使得可输出的访客交互内容灵活且多样。

25 在一个实施例中，该交互方法中，在获取访客语音数据之前还包括以下步骤：获取图像帧；当对图像帧进行人脸检测得到图像帧包括人脸图像时，生成趋向访客移动指令；按照趋向访客移动指令移动。

在一个实施例中，交互机器人可通过摄像头，在摄像头当前的视野下采集

图像帧，获取采集得到的图像帧。其中，摄像头的视野可因交互机器人的姿态和位置的变化而变化。交互机器人具体也可按照固定或动态的帧率采集图像帧，获取采集得到的图像帧。其中，固定或动态的帧率能够使图像帧按照该固定或动态的帧率播放时形成连续的动态画面，以使交互机器人可追踪连续的动态画面中的特定目标。交互机器人还可调用摄像头开启摄像扫描模式，实时扫描当前的视野下的特定目标，并按照一定的帧率实时地生成图像帧，获取生成的图像帧。

其中，摄像头可以是交互机器人内置的摄像头，或者外置的与交互机器人关联的摄像头。摄像头可以是单目摄像头、双目摄像头或者 RGB-D（Red-Green-Blue-Deep）摄像头等。

进一步地，交互机器人可在获取到图像帧后，提取该图像帧中包括的图像数据，并检测该图像数据是否包含人脸特征数据。若交互机器人检测到该图像数据中包含人脸特征数据，则判定该图像帧中包括人脸图像。交互机器人也可在获取到图像帧后，将该图像帧发送至服务器，由服务器完成对图像帧的人脸检测过程，再向交互机器人返回图像帧中是否包括人脸图像的检测结果。其中，检测结果可包括图像帧中存在人脸图像的概率和人脸图像的坐标区域。

更进一步地，交互机器人可在对图像帧进行人脸检测得到图像帧包括人脸图像时，确定人脸图像的目标位置以及交互机器人当前的起始位置，根据起始位置和目标位置生成趋向访客移动指令。趋向访客移动指令用于驱使交互机器人向访客移动。

在一个实施例中，生成趋向访客移动指令的步骤包括：确定人脸图像在地图中相应的目标节点；从地图中挑选与图像帧匹配的起始节点；其中，图像帧的特征与起始节点对应的节点图像的特征相匹配；根据起始节点和目标节点，在地图包括的路径中选取趋向访客移动路径；按照选取的趋向访客移动路径生成趋向访客移动指令。

其中，地图是交互机器人根据从自然空间中采集的图像帧构建的特征分布图。交互机器人可基于 SLAM（Simultaneous Localization And Mapping 定位与地图构建）对自然空间构建相应的地图。交互机器人基于 SLAM 构建的地图具

体可以是三维点图。节点是交互机器人将从自然空间中采集图像帧的位置投影至地图空间中的位置。目标节点是目标在自然空间中的位置投影至地图的节点。比如，目标在自然空间的坐标为 $A(x_1, y_1, z_1)$ ，将 A 投影至地图空间后坐标为 $B(x_2, y_2, z_2)$ ，那么 B 即为目标在地图中的节点。

5 在一个实施例中，地图中可包括若干节点，各节点均存在一一对应的节点图像。地图还可包括从节点图像中提取出的特征点。包括特征点和节点的地图是对自然空间中场景的三维重建。具体地，自然空间中的三维场景中的三维点通过投影矩阵的投影变换，得到交互机器人摄像头摄像平面的二维图像帧中的像素点，二维图像帧中的像素点再经过投影矩阵的投影反变换，得到地图中的
10 三维重建场景中的三维特征点。

交互机器人可在检测到该图像帧中包括人脸图像时，计算该人脸图像在地图中的位置。具体地，交互机器人可确定人脸图像在该图像帧中的坐标位置，根据与交互机器人的摄像头适配的投影矩阵，计算该人脸图像在地图中的位置，在地图中包括的节点中查找与计算得到的位置相应的节点，得到目标节点。

15 进一步地，节点图像是交互机器人在与地图中的节点存在投影关系的自然空间中的位置处采集的图像。图像的特征可以是颜色特征、纹理特征和形状特征中的一种或几种的组合。交互机器人可在构建地图时，对地图中节点对应的节点图像提取特征，将提取的节点图像的特征相对于相应的节点存储在数据库或者缓存中。在一个实施例中，交互机器人可遍历地图中各节点对应的节点图
20 像的特征，判断遍历至的节点图像的特征与图像帧的特征是否匹配。交互机器人可在判定遍历至的节点图像的特征与图像帧的特征匹配时，获取遍历至的节点图像的特征所对应的节点为起始节点。

更进一步地，地图中可包括通过地图中的节点形成的路径。交互机器人可以起始节点作为起点，目标节点为终点，在地图中通过节点形成的路径中选取
25 路径得到趋向访客移动路径，按照选取的趋向访客移动路径生成趋向访客移动指令。在一个实施例中，地图以起始节点作为起点，目标节点为终点的路径可以是一条或者多条。当以起始节点作为起点，目标节点为终点的路径唯一时，交互机器人可直接获取该路径为趋向访客移动路径。当以起始节点作为起点，

目标节点为终点的路径不唯一时，交互机器人可随机选取一条路径作为趋向访客移动路径，也可获取包括的节点数最少的路径作为趋向访客移动路径。

在本实施例中，在检测到获取的图像帧包括人脸图像时，即可自动在地图中确定该人脸图像相应的目标节点，定位目标在地图中的位置，然后以该图像帧的特征与地图中各节点对应的节点图像的特征的匹配关系为依据，即可从地图中挑选与该图像帧匹配的起始节点，定位本机当前在地图中的位置，再根据当前节点和目标节点便可在地图包括的路径中选取趋向访客移动路径，并生成趋向访客移动指令。这样通过图像之间的特征匹配即可完成在地图中的定位，避免了通过传感信号定位引起的环境影响，提高了运动控制的准确性。

上述实施例中，在获取图像帧并检测到获取的图像帧包括人脸图像时，即自动生成趋向访客移动指令并向访客移动，以主动与访客进行交互，提高了与访客的交互效率。

在一个实施例中，该交互方法还包括构建地图的步骤：从按时序采集的图像帧中选取图像帧；判断选取的图像帧的特征是否符合预设的节点图像的特征；当选取的图像帧的特征符合节点图像的特征时，获取选取的图像帧为节点图像；确定获取的节点图像在地图中相应的节点；对应于确定的节点存储获取的节点图像的特征。

图 6 示出了一个实施例中地图创建过程的流程示意图。参考图 6，该地图创建过程包括追踪、建图和闭环检测三个部分。交互机器人在获取构建地图的指令后，可按照固定或动态的帧率采集图像帧。在采集到图像帧后，提取该图像帧的特征点，将提取的特征点与地图中新增的节点对应的节点图像的特征点匹配。当提取的特征点与地图中新增的节点对应的节点图像的特征点匹配失败时，交互机器人可重新获取采集的图像帧进行重定位。

当提取的特征点与地图中新增的节点对应的节点图像的特征点匹配成功时，根据地图中新增的节点预估采集的该图像帧对应与地图中的节点。交互机器人可再追踪地图中与采集的该图像相匹配的特征点，根据相匹配的特征优化该图像帧对应与地图中的节点。在对采集的该图像优化完成后，判断该图像帧的特征点是否符合预设的节点图像的特征点，若否，交互机器人可重新获取采

集的图像帧进行特征点匹配。

若该图像帧的特征点符合预设的节点图像的特征点，交互机器人可获取该图像帧为新增的节点图像。交互机器人可提取该新增的节点图像的特征点，按照预设的统一的格式表示提取的特征点，再按照三角测距算法确定新增的节点图像的特征点在地图中的位置，从而更新局部地图，再进行局部集束调整，去除相似度高于预设相似度阈值的节点图像对应的冗余的节点。

交互机器人在获取该图像帧为新增的节点图像后，可异步进行闭环检测。将新增的节点图像的特征与已有的节点对应的节点图像的特征进行对比，当新增的节点图像的特征与已有的节点对应的节点图像的特征之间的相似度高于预设相似度阈值，交互机器人可判定新增的节点图像在自然空间中的采集位置与已有的节点对应的节点图像在自然空间中的采集位置一致，即存在闭环。交互机器人可再根据新增的节点图像相应的节点，在地图中生成包括位置一致的节点的环形路径，并进行闭环优化和闭环融合。最终得到包括特征点、节点和路径的全局地图

图 7 示出了一个实施例中在地图中选取趋向访客移动路径的示意图。参考图 7，该示意图包括目标节点 701、起始节点 702 以及趋向访客移动路径 703。交互机器人在确定目标节点 701 即目标所在的位置以及起始节点 702 即本机所在位置后，以起始节点 702 为起点，以目标节点 701 为终点，在地图中选取趋向访客移动路径 703。

如图 8 所示，在一个实施例中，该交互方法还包括人脸检测的步骤，人脸检测的步骤具体包括：

S802，将图像帧输入卷积神经网络模型。

其中，卷积神经网络模型是由多层互相连接而形成的复杂网络模型。神经网络模型可包括多层特征转换层，每层特征转换层都有对应的非线性变化算子，每层的非线性变化算子可以是多个，每层特征转换层中一个非线性变化算子对输入的图像进行非线性变化，得到特征图（Feature Map）作为运算结果。

具体地，卷积神经网络模型是以包括人脸图像的图像作为训练数据，进行学习训练得到的用于提取人脸特征的模型。交互机器人在获取到图像帧后，将

图像帧输入卷积神经网络模型，利用卷积神经网络模型对图像帧进行人脸特征提取。其中，人脸特征可以是用于反映出人的性别、人脸的轮廓、发型、眼镜、鼻子、嘴以及各个脸部器官之间的距离等其中的一种或多种特征。

5 在一个实施例中，卷积神经网络模型是以图像作为训练数据，进行学习训练得到的用于提取图像特征的模型。交互机器人在获取到图像帧后，将图像帧输入卷积神经网络模型，利用卷积神经网络模型对图像帧进行图像特征提取。

S804，获取卷积神经网络模型包括的多个网络层输出的特征图。

具体地，交互机器人可获取卷积神经网络模型包括的多个网络层输出的特征图。特征图是由非线性变化算子对输入的图像进行处理得到的响应值构成的。
10 不同的网络层提取的特征不同。交互机器人可利用提取人脸特征的卷积神经网络输出的特征图确定输入的图像相应的人脸特征数据。交互机器人可利用提取图像特征的卷积神经网络输出的特征图确定输入的图像相应的图像特征数据，进而判断该图像特征数据中是否包括人脸特征数据。

举例说明，交互机器人可采用 52 层深度残差网络模型进行图像处理，提取
15 该深度残差网络模型中包括的 4 层全连接层输出的特征图，作为后续输入。

S806，将各特征图依次输入记忆神经网络模型。

其中，记忆神经网络模型是可对序列输入进行综合处理的神经网络模型。记忆神经网络模型是递归神经网络模型。记忆神经网络模型具体可以是 LSTM
(Long Short-Term Memory 长短时记忆神经网络)。具体地，交互机器人可将获
20 取的各特征图依次输入记忆神经网络模型，进行人脸特征检测。

S808，获取记忆神经网络模型输出的图像帧是否包括人脸图像的结果。

具体地，交互机器人可获取记忆神经网络模型根据输入的各特征图综合处理得到的人脸检测结果。人脸检测结果包括存在人脸图像的概率和人脸图像在图像帧中的坐标区域。

25 在一个实施例中，交互机器人还可在提取得到人脸检测结果后，根据人脸检测结果中包括的人脸图像在图像帧中的坐标区域，过滤掉重叠区域超过预设重叠阈值的人脸检测结果，根据过滤后保留的人脸检测结果得到人脸图像在图像帧中的坐标区域。

在一个实施例中，记忆神经网络模型可使用一个矩形窗口，按照预设方向和预设步长在输入的特征图中移动，从而进行窗口扫描，在扫描时提取扫描至的窗口图像中人脸特征数据，根据提取的人脸特征图像，得到扫描至的窗口图像中存在人脸图像的概率。将计算得到的概率排序靠前的窗口图像在图像帧中的坐标区域进行存储，并继续对后续输入的特征图进行处理。

图 9 示出了一个实施例中对面脸图像进行人脸识别的示意图。参考图 9，交互机器人采用的记忆神经网络模型对输入的特征图按照矩形窗口扫描分析，得到与矩形窗口 A 对应的存在人脸图像的概率 P_A ，矩形窗口 B 对应的存在人脸图像的概率 P_B ，矩形窗口 C 对应的存在人脸图像的概率 P_C 。此时， $P_C > P_A > P_B$ ，记忆神经网络模型可将 P_C 对应的矩形窗口 C 进行记录，继续对后续输入的特征图按照矩形窗口扫描分析，并综合多次分析得到矩形窗口以及相应的存在人脸图像的概率，输出交互机器人获取的图像帧中存在人脸图像的概率以及该人脸图像在图像帧中的坐标区域。

在本实施例中，通过卷积神经网络模型的包括的多个网络层充分提取图像特征，再将多层网络层提取的特征输入记忆神经网络模型综合处理，使得人脸检测更准确。

在一个实施例中，该交互方法还包括人脸识别的步骤，人脸识别的步骤具体包括：提取人脸图像的人脸特征数据；根据人脸特征数据查询与人脸图像相匹配的预设人脸图像；根据预设人脸图像得到访客身份识别结果。步骤 S312 具体包括：获取与挑选出的特征词对应、且与所述访客身份识别结果关联的访客交互内容并输出。

其中，访客身份识别结果是用于反映访客身份的数据。访客身份可以是访客的名字、社会地位或者职位信息等。

在一个实施例中，交互机器人上设置有预设人脸图像库，预设人脸图像库中包括若干预设人脸图像。交互机器人可在检测到图像帧中包括人脸图像时，将图像帧中的人脸图像与预设人脸图像库中包括的预设人脸图像比较，检测图像帧中的人脸图像和预设人脸图像之间是否匹配。交互机器人可在图像帧中的人脸图像和预设人脸图像之间匹配时，判定该图像帧包括的人脸图像与预设人

脸图像为相同的人物图像，获取该预设人脸图像对应的访客身份信息作为访客身份识别结果。

其中，预设人脸图像可以是用于反映对应目标的真实人脸图像。可从目标所上传的个人资料、历史发表的图片信息中，由对应目标自定义选取的图像，
5 或由系统自动地分析选取的一张图片，作为相应的预设人脸图像。

在一个实施例中，交互机器人在检测图像帧中的人脸图像和预设人脸图像之间是否匹配，具体可计算图像帧中的人脸图像和预设人脸图像之间的相似度。交互机器人可先提取图像帧中的人脸图像和预设人脸图像各自的特征，从而计算两特征之间的差异，特征之间的差异越大则相似度越低，特征之间的差异越
10 小则相似度越高。其中，交互机器人计算图像帧中的人脸图像和预设人脸图像之间的相似度时，可以采用适于图像处理器的加速算法，提高运算速率。

在一个实施例中，交互机器人可在判定该图像帧中包括人脸图像后从该图像数据中提取人脸特征数据，再将提取的人脸特征数据与预设人脸图像库中各预设人脸图像相对应的人脸特征数据比较，得到访客身份识别结果。

15 在一个实施例中，交互机器人对图像帧进行检测得到的该图像帧包括的人脸图像可以是一个或者多个。交互机器人可确定图像帧中包括的人脸图像占图像帧的占比，提取占比超过预设比例的人脸图像的人脸特征数据；和/或，确定图像帧中包括的人脸图像的清晰度，提取清晰度超过清晰度阈值的人脸图像的人脸特征数据。交互机器人再对提取了人脸特征数据的人脸图像进行识别。

20 交互机器人可再获取与挑选出的特征词对应、且与所述访客身份识别结果关联的访客交互内容并输出。

在一个实施例中，交互机器人可事先设置访客交互内容，并将访客交互内容与访客标识和特征词分别关联，再将设置的访客交互内容存储在数据库或者文件中，在需要时从数据库或者文件中读取。交互机器人在识别得到访客身份
25 识别结果后，可拉取该访客身份识别结果对应的访客标识所关联的、且与挑选出的特征词对应的访客交互内容。该访客交互内容可以是直接可输出的访客交互内容，也可以是待补全的访客交互内容模板。

在本实施例中，根据访客身份信息设置个性化的访客交互内容，使得与访

客的交互内容更丰富，交互内容呈现方式多样化。

在其他实施例中，上述人脸检测以及人脸识别过程均可在服务器上进行。通过服务器进行人脸检测和人脸识别可使得识别结果或者检测结果更准确。

5 在一个实施例中，交互机器人可事先按照特定的排序方式对预设人脸图像库中包括的预设人脸图像进行优先级排序。比如，按照预设人脸图像所对应的真实人物的社会地位或者职位信息等对预设人脸图像进行优先级排序。当交互机器人从获取的图像帧中识别出多个访客身份识别结果时，选取对应优先级最高的预设人脸图像的访客身份识别结果进行响应。

如图 10 所示，在一个具体的实施例中，该交互方法具体包括以下步骤：

10 S1002，获取图像帧。

S1004，将图像帧输入卷积神经网络模型；获取卷积神经网络模型包括的多个网络层输出的特征图；将各特征图依次输入记忆神经网络模型；获取记忆神经网络模型输出的人脸检测结果。

15 S1006，判断人脸检测结果是否表示图像帧包括人脸图像；若是，则跳转到步骤 S1008；若否，则返回步骤 S1002。

S1008，提取人脸图像的人脸特征数据；根据人脸特征数据查询与人脸图像相匹配的预设人脸图像；根据预设人脸图像得到目标身份识别结果。

S1010，确定人脸图像在地图中相应的目标节点；从地图中挑选与图像帧匹配的起始节点；其中，图像帧的特征与起始节点对应的节点图像的特征相匹配。

20 S1012，根据起始节点和目标节点，在地图包括的路径中选取趋向访客移动路径；按照选取的趋向访客移动路径生成趋向访客移动指令，按照趋向访客移动指令移动。

S1014，获取访客语音数据；对识别访客语音数据得到的文本进行分词得到特征词集合。

25 S1016，根据主题生成模型，对于特征词集合中的每个特征词，按照预设数量的主题与特征词的多项式分布选择当前特征词所属的主题，并按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数；分别确定特征

词集合中各特征词属于各主题时相应的出现概率。

S1018, 对于预设数量的主题中的每个主题, 根据属于当前主题的特征词的出现概率和当前主题相应的特征词生成概率函数, 获得文本属于当前主题的概率; 将对应最大概率的主题确定为文本所属的主题。

5 S1020, 分别获取特征词集合中各特征词属于确定的主题时相应的出现概率。

S1022, 对于特征词集合中的每个特征词, 根据文本属于确定的主题的概率与当前特征词属于确定的主题的出现概率, 得到当前特征词对应的主题影响力值; 按照各特征词对应的主题影响力值, 从特征词集合中挑选特征词。

10 S1024, 获取与挑选出的特征词对应、且与访客身份识别结果关联的访客交互内容并输出。

在本实施例中, 在获取访客语音数据, 并对识别该访客语音数据得到的文本进行分词得到特征词集合后, 即可利用主题生成模型和分词得到的特征词集合, 在访客语音数据稀疏度高的情况下准确地确定访客语音数据所属主题, 并
15 自动根据特征词集合中各特征词属于确定的主题时的出现概率, 选取合适的特征词所对应的访客交互内容输出, 避免了需要大量的人工参与而引入的工作量, 提高了与访客间的交互效率及输出的访客交互内容的准确性。

图 11 示出了一个实施例中用于实现上述实施例的交互机器人的架构图。参考图 11, 交互机器人包括用于采集图像帧的摄像头, 采集访客语音数据的声音采集装置, 用于输出访客交互内容的显示屏或扬声器。交互机器人包括图像分
20 析部分和自然语音处理部分。交互机器人在通过摄像头采集到图像帧, 获取该图像帧, 将该图像帧输入神经网络模型(包括卷积神经网络模型和记忆神经网络模型等)进行人脸检测, 在检测得到图像帧中包括人脸图像时再进行人脸识别, 得到访客身份识别结果。交互机器人可在检测得到图像帧中包括人脸图像
25 时异步根据离线场景建模得到的地图选取向访客移动的趋向访客移动路径, 生成趋向访客移动指令并移动。交互机器人再获取访客语音数据, 通过自然语言处理算法(主题生成模型等)结合语料库(在线语料库和离线语料库), 确定访客所要表达的意图, 选取特征词以获取合适的访客交互内容。

如图 12 所示，在一个实施例中，提供了一种交互终端 1200。参照图 12，该一种交互终端 1200，

包括：一个或者一个以上处理器、一个或一个以上存储器，所述存储器中
5 存储有至少一段应用程序，所述至少一段应用程序适于处理器进行加载，以执行以下操作：

获取访客语音数据。

对识别访客语音数据得到的文本进行分词得到特征词集合。

10 根据主题生成模型和特征词集合确定文本所属的主题。

分别获取特征词集合中各特征词属于确定的主题时相应的出现概率。

按照各特征词相应的出现概率，从特征词集合中挑选特征词。

获取与挑选出的特征词对应的访客交互内容并输出。

上述交互终端 1200，在获取访客语音数据，并对识别该访客语音数据得到
15 的文本进行分词得到特征词集合后，即可利用主题生成模型和分词得到的特征词集合，在访客语音数据稀疏度高的情况下准确地确定访客语音数据所属主题，并自动根据特征词集合中各特征词属于确定的主题时的出现概率，选取合适的特征词所对应的访客交互内容输出，避免了需要大量的人工参与而引入的工作量，提高了与访客间的交互效率及输出的访客交互内容的准确性。

20 在一个实施例中，确定模块 1203 还用于根据主题生成模型，对于特征词集合中的每个特征词，获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数；分别确定特征词集合中各特征词属于各主题时相应的出现概率；对于预设数量的主题中的每个主题，根据属于当前主题的特征词的出现概率和当前主题相应的特征词生成概率函数，获得文本属于当前主题的概率；
25 将对应最大概率的主题确定为文本所属的主题。

在本实施例中，将处理自然语言的主题生成模型运用到对访客语音数据进行访客意图定位，确定访客所要表达的主题后，便可以基于访客表达的主题进行候选的精确交互，提高了与访客进行交互的准确性。

在一个实施例中，确定模块 1203 还用于根据主题生成模型，对于特征词集合中的每个特征词，按照预设数量的主题与特征词的多项式分布选择当前特征词所属的主题，并按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数。

在本实施例中，通过处理自然语言的主题生成模型中训练得到的主题与特征词的关系，及文本与主题的关系，得到访客语音数据包括的特征词属于各主题时相应的特征词生成概率函数，从而可通过概率函数来预测访客语音数据所属的主题，进而合理地从概率学的角度上确定访客所要表达的主题。

10 在一个实施例中，挑选模块 1205 还用于对于特征词集合中的每个特征词，根据文本属于确定的主题的概率与当前特征词属于确定的主题的出现概率，得到当前特征词对应的主题影响力值；按照各特征词对应的主题影响力值，从特征词集合中挑选特征词。

15 在本实施例中，在进行特征词挑选时，将文本属于确定的主题的概率与当前特征词属于确定的主题的出现概率，作为特征词对应的主题影响力值的评估依据，使得对特征词对应的主题影响力值的评估更加科学，从而挑选出更能反映访客意图的特征词，从而提高与访客的交互效率与准确性。

20 在一个实施例中，输出模块 1206 还用于从本地查找与挑选出的特征词对应的离线访客交互内容；将挑选出的特征词发送至服务器；接收服务器返回的根据挑选出的特征词实时生成的访客交互内容；将查找到的离线访客交互内容与接收到的访客交互内容取交集或并集后输出。

在本实施例中，提供了离线与在线两种方式获取访客交互内容输出。通过离线方式获取的访客交互内容可极大地提高交互机器人的响应效率，通过在线方式获取访客交互内容，使得可输出的访客交互内容灵活且多样。

25 如图 13 所示，在一个实施例中，交互终端 1200

所述至少一段应用程序适于处理器进行加载，以执行以下操作：

获取图像帧；当对图像帧进行人脸检测得到图像帧包括人脸图像时，生成趋向访客移动指令。

按照趋向访客移动指令移动。

在本实施例中，在获取图像帧并检测到获取的图像帧包括人脸图像时，即自动生成趋向访客移动指令并向访客移动，以主动与访客进行交互，提高了与访客的交互效率。

5 在一个实施例中，所述至少一段应用程序适于处理器进行加载，以执行以下操作：确定人脸图像在地图中相应的目标节点；从地图中挑选与图像帧匹配的起始节点；其中，图像帧的特征与起始节点对应的节点图像的特征相匹配；根据起始节点和目标节点，在地图包括的路径中选取趋向访客移动路径；按照选取的趋向访客移动路径生成趋向访客移动指令。

10 在本实施例中，在检测到获取的图像帧包括人脸图像时，即可自动在地图中确定该人脸图像相应的目标节点，定位目标在地图中的位置，然后以该图像帧的特征与地图中各节点对应的节点图像的特征的匹配关系为依据，即可从地图中挑选与该图像帧匹配的起始节点，定位本机当前在地图中的位置，再根据当前节点和目标节点便可在地图包括的路径中选取趋向访客移动路径，并生成
15 趋向访客移动指令。这样通过图像之间的特征匹配即可完成在地图中的定位，避免了通过传感信号定位引起的环境影响，提高了运动控制的准确性。

在一个实施例中，所述至少一段应用程序适于处理器进行加载，以执行以下操作：

20 将图像帧输入卷积神经网络模型；获取卷积神经网络模型包括的多个网络层输出的特征图；将各特征图依次输入记忆神经网络模型；获取记忆神经网络模型输出的图像帧是否包括人脸图像的结果。

在本实施例中，通过卷积神经网络模型的包括的多个网络层充分提取图像特征，再将多层网络层提取的特征输入记忆神经网络模型综合处理，使得人脸检测更准确。

25 在一个实施例中，所述至少一段应用程序适于处理器进行加载，以执行以下操作：

提取人脸图像的人脸特征数据；根据人脸特征数据查询与人脸图像相匹配的预设人脸图像；根据预设人脸图像得到访客身份识别结果。获取与挑选出的

特征词对应、且与访客身份识别结果关联的访客交互内容并输出。

在本实施例中，根据访客身份信息设置个性化的访客交互内容，使得与访客的交互内容更丰富，交互内容呈现方式多样化。

5 在一个实施例中，提供了一种计算机可读存储介质，计算机可读存储介质上存储有计算机可读指令，计算机可读指令被处理器执行时，使得处理器执行以下步骤：获取访客语音数据；对识别访客语音数据得到的文本进行分词得到特征词集合；根据主题生成模型和特征词集合确定文本所属的主题；分别获取特征词集合中各特征词属于确定的主题时相应的出现概率；按照各特征词相应
10 的出现概率，从特征词集合中挑选特征词；获取与挑选出的特征词对应的访客交互内容并输出。

 在一个实施例中，根据主题生成模型和特征词集合确定文本所属的主题，包括：根据主题生成模型，对于特征词集合中的每个特征词，获得当前特征词
15 分别属于预设数量的主题中各主题时相应的特征词生成概率函数；分别确定特征词集合中各特征词属于各主题时相应的出现概率；对于预设数量的主题中的每个主题，根据属于当前主题的特征词的出现概率和当前主题相应的特征词生成概率函数，获得文本属于当前主题的概率；将对应最大概率的主题确定为文本所属的主题。

 在一个实施例中，根据主题生成模型，对于特征词集合中的每个特征词，
20 获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数，包括：根据主题生成模型，对于特征词集合中的每个特征词，按照预设数量的主题与特征词的多项式分布选择当前特征词所属的主题，并按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建当前特征词
 分别属于预设数量的主题中各主题时相应的特征词生成概率函数。

25 在一个实施例中，按照各特征词相应的出现概率，从特征词集合中挑选特征词，包括：对于特征词集合中的每个特征词，根据文本属于确定的主题的概率与当前特征词属于确定的主题的出现概率，得到当前特征词对应的主题影响力值；按照各特征词对应的主题影响力值，从特征词集合中挑选特征词。

在一个实施例中，获取与挑选出的特征词对应的访客交互内容并输出，包括：从本地查找与挑选出的特征词对应的离线访客交互内容；将挑选出的特征词发送至服务器；接收服务器返回的根据挑选出的特征词实时生成的访客交互内容；将查找到的离线访客交互内容与接收到的访客交互内容取交集或并集后输出。

在一个实施例中，计算机可读指令还使得处理器执行获取访客语音数据之前，执行以下步骤：获取图像帧；当对图像帧进行人脸检测得到图像帧包括人脸图像时，生成趋向访客移动指令；按照趋向访客移动指令移动。

在一个实施例中，生成趋向访客移动指令，包括：确定人脸图像在地图中相应的目标节点；从地图中挑选与图像帧匹配的起始节点；其中，图像帧的特征与起始节点对应的节点图像的特征相匹配；根据起始节点和目标节点，在地图包括的路径中选取趋向访客移动路径；按照选取的趋向访客移动路径生成趋向访客移动指令。

在一个实施例中，计算机可读指令还使得处理器执行获取图像帧之后，执行以下步骤：将图像帧输入卷积神经网络模型；获取卷积神经网络模型包括的多个网络层输出的特征图；将各特征图依次输入记忆神经网络模型；获取记忆神经网络模型输出的图像帧是否包括人脸图像的结果。

在一个实施例中，计算机可读指令还使得处理器执行当对图像帧进行人脸检测得到图像帧包括人脸图像时，生成趋向访客移动指令之后，执行以下步骤：提取人脸图像的人脸特征数据；根据人脸特征数据查询与人脸图像相匹配的预设人脸图像；根据预设人脸图像得到访客身份识别结果。获取与挑选出的特征词对应的访客交互内容并输出，包括：获取与挑选出的特征词对应、且与访客身份识别结果关联的访客交互内容并输出。

上述存储介质，在获取访客语音数据，并对识别该访客语音数据得到的文本进行分词得到特征词集合后，即可利用主题生成模型和分词得到的特征词集合，在访客语音数据稀疏度高的情况下准确地确定访客语音数据所属主题，并自动根据特征词集合中各特征词属于确定的主题时的出现概率，选取合适的特征词所对应的访客交互内容输出，避免了需要大量的人工参与而引入的工作量，

提高了与访客间的交互效率及输出的访客交互内容的准确性。

在一个实施例中，提供了一种交互机器人，包括存储器和处理器，存储器中储存有计算机可读指令，计算机可读指令被处理器执行时，使得处理器执行以下步骤：获取访客语音数据；对识别访客语音数据得到的文本进行分词得到特征词集合；根据主题生成模型和特征词集合确定文本所属的主题；分别获取特征词集合中各特征词属于确定的主题时相应的出现概率；按照各特征词相应的出现概率，从特征词集合中挑选特征词；获取与挑选出的特征词对应的访客交互内容并输出。

10 在一个实施例中，根据主题生成模型和特征词集合确定文本所属的主题，包括：根据主题生成模型，对于特征词集合中的每个特征词，获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数；分别确定特征词集合中各特征词属于各主题时相应的出现概率；对于预设数量的主题中的每个主题，根据属于当前主题的特征词的出现概率和当前主题相应的特征词生成概率函数，获得文本属于当前主题的概率；将对应最大概率的主题确定为文本所属的主题。

在一个实施例中，根据主题生成模型，对于特征词集合中的每个特征词，获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数，包括：根据主题生成模型，对于特征词集合中的每个特征词，按照预设数量的主题与特征词的多项式分布选择当前特征词所属的主题，并按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数。

在一个实施例中，按照各特征词相应的出现概率，从特征词集合中挑选特征词，包括：对于特征词集合中的每个特征词，根据文本属于确定的主题的概率与当前特征词属于确定的主题的出现概率，得到当前特征词对应的主题影响力值；按照各特征词对应的主题影响力值，从特征词集合中挑选特征词。

在一个实施例中，获取与挑选出的特征词对应的访客交互内容并输出，包括：从本地查找与挑选出的特征词对应的离线访客交互内容；将挑选出的特征

词发送至服务器；接收服务器返回的根据挑选出的特征词实时生成的访客交互内容；将查找到的离线访客交互内容与接收到的访客交互内容取交集或并集后输出。

5 在一个实施例中，计算机可读指令还使得处理器执行获取访客语音数据之前，执行以下步骤：获取图像帧；当对图像帧进行人脸检测得到图像帧包括人脸图像时，生成趋向访客移动指令；按照趋向访客移动指令移动。

10 在一个实施例中，生成趋向访客移动指令，包括：确定人脸图像在地图中相应的目标节点；从地图中挑选与图像帧匹配的起始节点；其中，图像帧的特征与起始节点对应的节点图像的特征相匹配；根据起始节点和目标节点，在地图包括的路径中选取趋向访客移动路径；按照选取的趋向访客移动路径生成趋向访客移动指令。

15 在一个实施例中，计算机可读指令还使得处理器执行获取图像帧之后，执行以下步骤：将图像帧输入卷积神经网络模型；获取卷积神经网络模型包括的多个网络层输出的特征图；将各特征图依次输入记忆神经网络模型；获取记忆神经网络模型输出的图像帧是否包括人脸图像的结果。

20 在一个实施例中，计算机可读指令还使得处理器执行当对图像帧进行人脸检测得到图像帧包括人脸图像时，生成趋向访客移动指令之后，执行以下步骤：提取人脸图像的人脸特征数据；根据人脸特征数据查询与人脸图像相匹配的预设人脸图像；根据预设人脸图像得到访客身份识别结果。获取与挑选出的特征词对应的访客交互内容并输出，包括：获取与挑选出的特征词对应、且与访客身份识别结果关联的访客交互内容并输出。

25 上述交互机器人，在获取访客语音数据，并对识别该访客语音数据得到的文本进行分词得到特征词集合后，即可利用主题生成模型和分词得到的特征词集合，在访客语音数据稀疏度高的情况下准确地确定访客语音数据所属主题，并自动根据特征词集合中各特征词属于确定的主题时的出现概率，选取合适的特征词所对应的访客交互内容输出，避免了需要大量的人工参与而引入的工作量，提高了与访客间的交互效率及输出的访客交互内容的准确性。

在一个实施例中，提供了一种计算机设备，包括存储器和处理器，存储器中储存有计算机可读指令，计算机可读指令被处理器执行时，使得处理器执行以下步骤：获取访客语音数据；对识别访客语音数据得到的文本进行分词得到特征词集合；根据主题生成模型和特征词集合确定文本所属的主题；分别获取特征词集合中各特征词属于确定的主题时相应的出现概率；按照各特征词相应的出现概率，从特征词集合中挑选特征词；获取与挑选出的特征词对应的访客交互内容并输出。

在一个实施例中，根据主题生成模型和特征词集合确定文本所属的主题，包括：根据主题生成模型，对于特征词集合中的每个特征词，获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数；分别确定特征词集合中各特征词属于各主题时相应的出现概率；对于预设数量的主题中的每个主题，根据属于当前主题的特征词的出现概率和当前主题相应的特征词生成概率函数，获得文本属于当前主题的概率；将对应最大概率的主题确定为文本所属的主题。

在一个实施例中，根据主题生成模型，对于特征词集合中的每个特征词，获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数，包括：根据主题生成模型，对于特征词集合中的每个特征词，按照预设数量的主题与特征词的多项式分布选择当前特征词所属的主题，并按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数。

在一个实施例中，按照各特征词相应的出现概率，从特征词集合中挑选特征词，包括：对于特征词集合中的每个特征词，根据文本属于确定的主题的概率与当前特征词属于确定的主题的出现概率，得到当前特征词对应的主题影响力值；按照各特征词对应的主题影响力值，从特征词集合中挑选特征词。

在一个实施例中，获取与挑选出的特征词对应的访客交互内容并输出，包括：从本地查找与挑选出的特征词对应的离线访客交互内容；将挑选出的特征词发送至服务器；接收服务器返回的根据挑选出的特征词实时生成的访客交互内容；将查找到的离线访客交互内容与接收到的访客交互内容取交集或并集后

输出。

在一个实施例中，计算机可读指令还使得处理器执行获取访客语音数据之前，执行以下步骤：获取图像帧；当对图像帧进行人脸检测得到图像帧包括人脸图像时，生成趋向访客移动指令；按照趋向访客移动指令移动。

5 在一个实施例中，生成趋向访客移动指令，包括：确定人脸图像在地图中相应的目标节点；从地图中挑选与图像帧匹配的起始节点；其中，图像帧的特征与起始节点对应的节点图像的特征相匹配；根据起始节点和目标节点，在地图包括的路径中选取趋向访客移动路径；按照选取的趋向访客移动路径生成趋向访客移动指令。

10 在一个实施例中，计算机可读指令还使得处理器执行获取图像帧之后，执行以下步骤：将图像帧输入卷积神经网络模型；获取卷积神经网络模型包括的多个网络层输出的特征图；将各特征图依次输入记忆神经网络模型；获取记忆神经网络模型输出的图像帧是否包括人脸图像的结果。

15 在一个实施例中，计算机可读指令还使得处理器执行当对图像帧进行人脸检测得到图像帧包括人脸图像时，生成趋向访客移动指令之后，执行以下步骤：提取人脸图像的人脸特征数据；根据人脸特征数据查询与人脸图像相匹配的预设人脸图像；根据预设人脸图像得到访客身份识别结果。获取与挑选出的特征词对应的访客交互内容并输出，包括：获取与挑选出的特征词对应、且与访客身份识别结果关联的访客交互内容并输出。

20 上述计算机设备，在获取访客语音数据，并对识别该访客语音数据得到的文本进行分词得到特征词集合后，即可利用主题生成模型和分词得到的特征词集合，在访客语音数据稀疏度高的情况下准确地确定访客语音数据所属主题，并自动根据特征词集合中各特征词属于确定的主题时的出现概率，选取合适的特征词所对应的访客交互内容输出，避免了需要大量的人工参与而引入的工作量，提高了与访客间的交互效率及输出的访客交互内容的准确性。

25

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一非易

失性计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory, ROM）等。

5 以上实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

10 以上所述实施例仅表达了本发明的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对本发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干变形和改进，这些都属于本发明的保护范围。因此，本发明专利的保护范围应以所附权利要求为准。

权 利 要 求 书

1、一种交互方法，所述方法包括：

终端获取访客语音数据；

对识别所述访客语音数据得到的文本进行分词得到特征词集合；

根据主题生成模型和所述特征词集合确定所述文本所属的主题；

5 分别获取所述特征词集合中各特征词属于确定的主题时相应的出现概率；

按照各所述特征词相应的出现概率，从所述特征词集合中挑选特征词；

获取与挑选出的特征词对应的访客交互内容并输出。

2、根据权利要求1所述的方法，其中，所述根据主题生成模型和所述特征词集合确定所述文本所属的主题，包括：

10 根据主题生成模型，对于所述特征词集合中的每个特征词，获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数；

分别确定所述特征词集合中各特征词属于各主题时相应的出现概率；

对于预设数量的主题中的每个主题，根据属于当前主题的特征词的出现概率和当前主题相应的特征词生成概率函数，获得所述文本属于当前主题的概率；

15 将对应最大概率的主题确定为所述文本所属的主题。

3、根据权利要求2所述的方法，其中，所述根据主题生成模型，对于所述特征词集合中的每个特征词，获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数，包括：

20 根据主题生成模型，对于所述特征词集合中的每个特征词，按照预设数量的主题与特征词的多项式分布选择当前特征词所属的主题，并按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数。

4、根据权利要求2所述的方法，其中，所述按照各所述特征词相应的出现概率，从所述特征词集合中挑选特征词，包括：

25 对于所述特征词集合中的每个特征词，根据所述文本属于确定的所述主题的概率与当前特征词属于确定的所述主题的出现概率，得到当前特征词对应的主题影响力值；

按照各所述特征词对应的主题影响力值，从所述特征词集合中挑选特征词。

5、根据权利要求1所述的方法，其中，所述获取与挑选出的特征词对应的访客交互内容并输出，包括：

从本地查找与挑选出的特征词对应的离线访客交互内容；

将挑选出的特征词发送至服务器；

5 接收所述服务器返回的根据挑选出的特征词实时生成的访客交互内容；

将查找到的离线访客交互内容与接收到的访客交互内容取交集或并集后输出。

6、根据权利要求1所述的方法，其中，所述获取访客语音数据之前，所述方法还包括：

10 获取图像帧；

当对所述图像帧进行人脸检测得到所述图像帧包括人脸图像时，生成趋向访客移动指令；

按照所述趋向访客移动指令移动。

7、根据权利要求6所述的方法，其中，所述生成趋向访客移动指令，包括：

15 确定所述人脸图像在地图中相应的目标节点；

从所述地图中挑选与所述图像帧匹配的起始节点；其中，所述图像帧的特征与所述起始节点对应的节点图像的特征相匹配；

根据所述起始节点和所述目标节点，在所述地图包括的路径中选取趋向访客移动路径；

20 按照选取的所述趋向访客移动路径生成趋向访客移动指令。

8、根据权利要求6所述的方法，其中，所述获取图像帧之后，所述方法还包括：

将所述图像帧输入卷积神经网络模型；

获取所述卷积神经网络模型包括的多个网络层输出的特征图；

25 将各所述特征图依次输入记忆神经网络模型；

获取所述记忆神经网络模型输出的所述图像帧是否包括人脸图像的结果。

9、根据权利要求6所述的方法，其中，所述当对所述图像帧进行人脸检测得到所述图像帧包括人脸图像时，生成趋向访客移动指令之后，所述方法还包

括:

提取所述人脸图像的人脸特征数据;

根据所述人脸特征数据查询与所述人脸图像相匹配的预设人脸图像;

根据所述预设人脸图像得到访客身份识别结果;

5 所述获取与挑选出的特征词对应的访客交互内容并输出, 包括:

获取与挑选出的特征词对应、且与所述访客身份识别结果关联的访客交互内容并输出。

10 10. 一种交互终端, 包括: 一个或者一个以上处理器、一个或一个以上存储器, 所述存储器中存储有至少一段应用程序, 所述至少一段应用程序适于处理器进行加载, 以执行以下操作:

获取访客语音数据;

对识别所述访客语音数据得到的文本进行分词得到特征词集合;

根据主题生成模型和所述特征词集合确定所述文本所属的主题;

15 分别获取所述特征词集合中各特征词属于确定的主题时相应的出现概率;

按照各所述特征词相应的出现概率, 从所述特征词集合中挑选特征词;

获取与挑选出的特征词对应的访客交互内容并输出。

11、根据权利要求 10 所述的交互终端, 其中,

所述至少一段应用程序适于处理器进行加载, 以执行以下操作:

20 根据主题生成模型, 对于所述特征词集合中的每个特征词, 获得当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数; 分别确定所述特征词集合中各特征词属于各主题时相应的出现概率; 对于预设数量的主题中的每个主题, 根据属于当前主题的特征词的出现概率和当前主题相应的特征词生成概率函数, 获得所述文本属于当前主题的概率; 将对应最大概率的主
25 题确定为所述文本所属的主题。

12、根据权利要求 11 所述的交互终端, 其中,

所述至少一段应用程序适于处理器进行加载, 以执行以下操作:

根据主题生成模型, 对于所述特征词集合中的每个特征词, 按照预设数量

的主题与特征词的多项式分布选择当前特征词所属的主题，并按照文本与主题的多项式分布生成在选择的主题条件下的当前特征词，以构建当前特征词分别属于预设数量的主题中各主题时相应的特征词生成概率函数。

13、根据权利要求 11 所述的交互终端，其中，

5 所述至少一段应用程序适于处理器进行加载，以执行以下操作：

对于所述特征词集合中的每个特征词，根据所述文本属于确定的所述主题的概率与当前特征词属于确定的所述主题的出现概率，得到当前特征词对应的主题影响力值；按照各所述特征词对应的主题影响力值，从所述特征词集合中挑选特征词。

10 14、根据权利要求 10 所述的交互终端，其中，

所述至少一段应用程序适于处理器进行加载，以执行以下操作：

从本地查找与挑选出的特征词对应的离线访客交互内容；将挑选出的特征词发送至服务器；接收所述服务器返回的根据挑选出的特征词实时生成的访客交互内容；将查找到的离线访客交互内容与接收到的访客交互内容取交集或并集后输出。

15 15、根据权利要求 10 所述的交互终端，其中，

所述至少一段应用程序适于处理器进行加载，以执行以下操作：

交互终端图像处理模块，用于获取图像帧；当对所述图像帧进行人脸检测得到所述图像帧包括人脸图像时，生成趋向访客移动指令；

20 运动模块，用于按照所述趋向访客移动指令移动。

16、根据权利要求 15 所述的交互终端，其中，

所述至少一段应用程序适于处理器进行加载，以执行以下操作：

将所述图像帧输入卷积神经网络模型；获取所述卷积神经网络模型包括的多个网络层输出的特征图；将各所述特征图依次输入记忆神经网络模型；获取
25 所述记忆神经网络模型输出的所述图像帧是否包括人脸图像的结果。

17、根据权利要求 15 所述的交互终端，其中，

所述至少一段应用程序适于处理器进行加载，以执行以下操作：

提取所述人脸图像的人脸特征数据；根据所述人脸特征数据查询与所述人

脸图像相匹配的预设人脸图像；根据所述预设人脸图像得到访客身份识别结果；获取与挑选出的特征词对应、且与所述访客身份识别结果关联的访客交互内容并输出。

- 5 18、一种计算机可读存储介质，其中，所述计算机可读存储介质上存储有计算机可读指令，所述计算机可读指令被处理器执行时，使得所述处理器执行如权利要求 1 至 9 中任一项所述的方法的步骤。

- 10 19、一种计算机设备，包括存储器和处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行如权利要求 1 至 9 中任一项所述的方法的步骤。

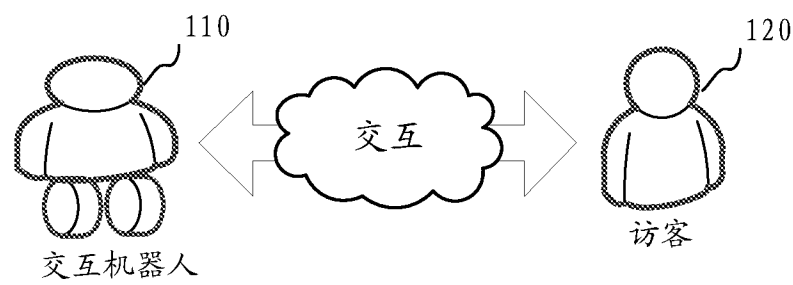


图 1

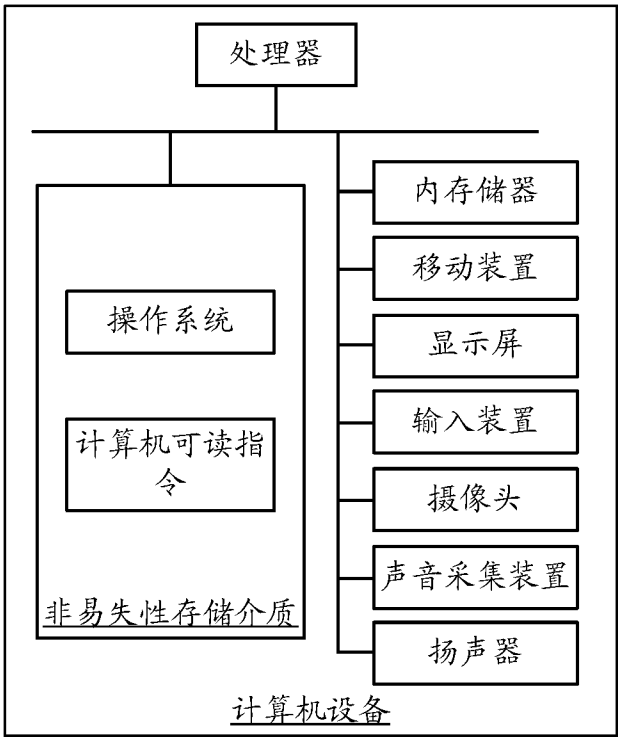


图 2

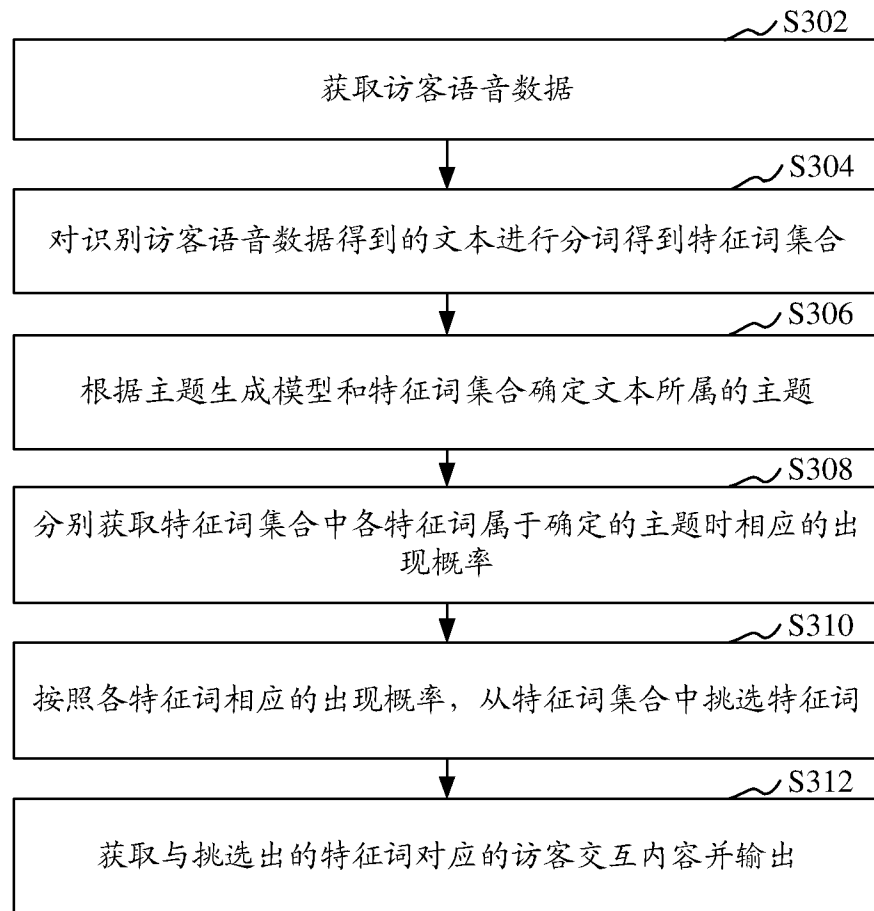


图 3

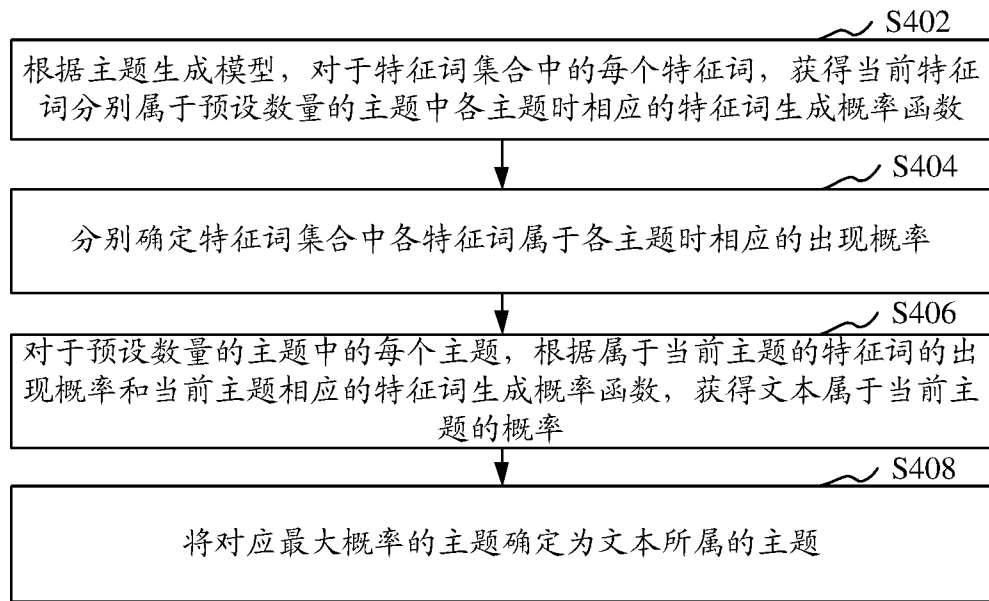


图 4

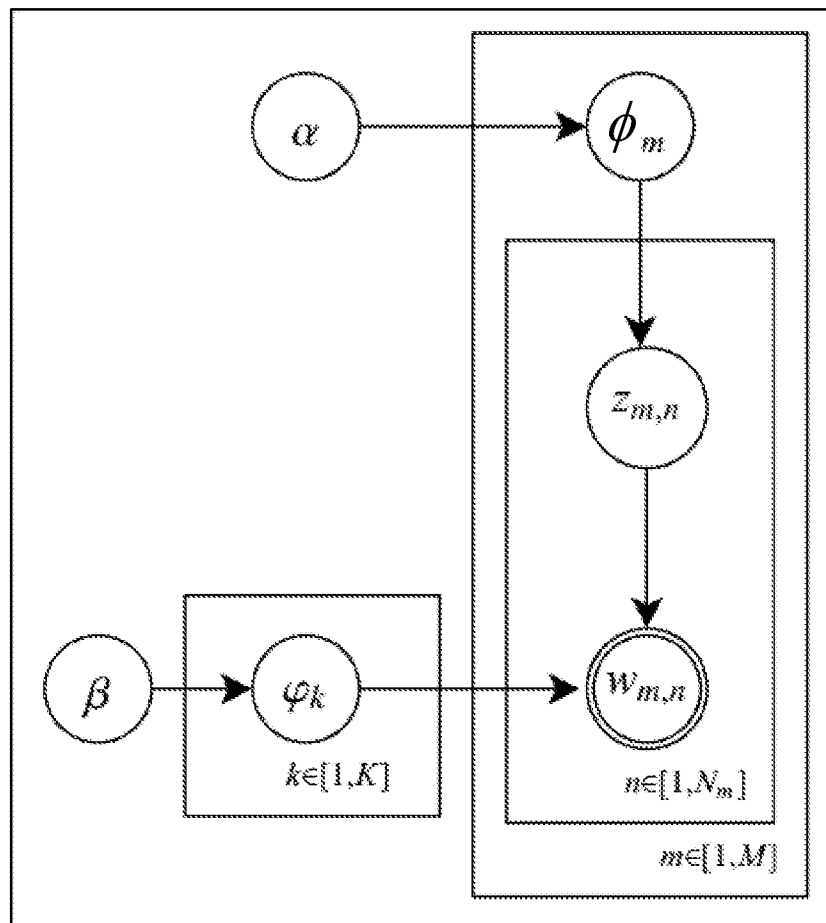


图 5

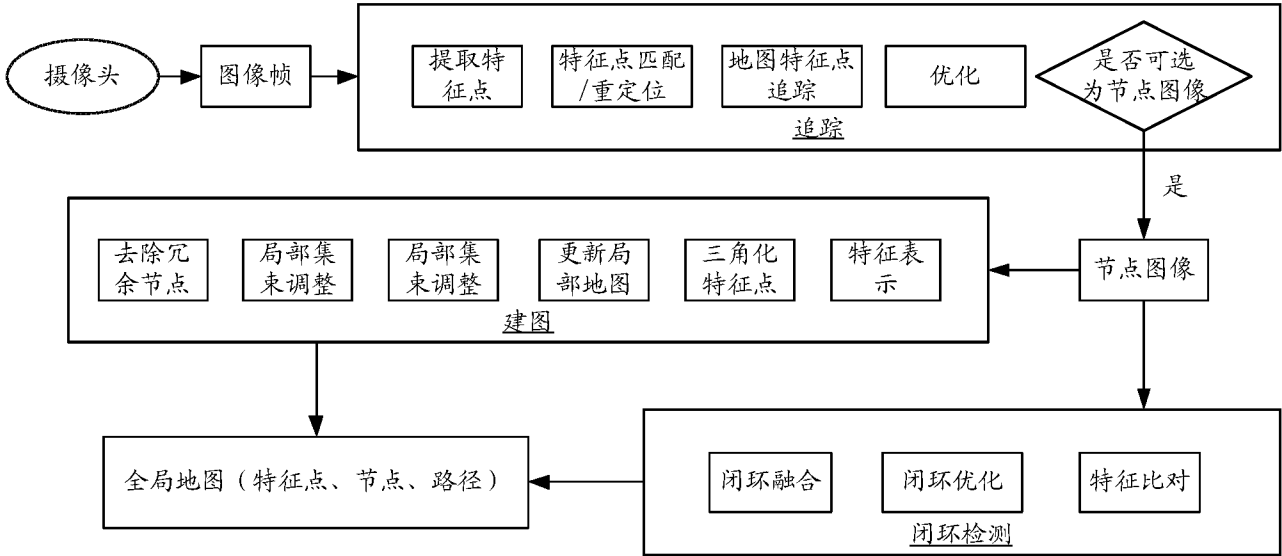


图 6

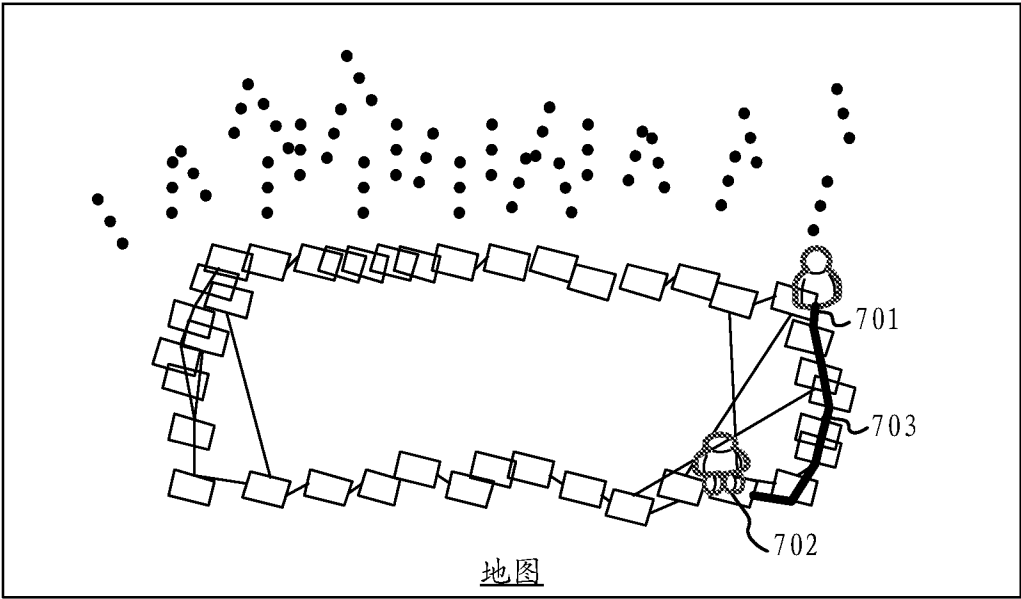


图 7

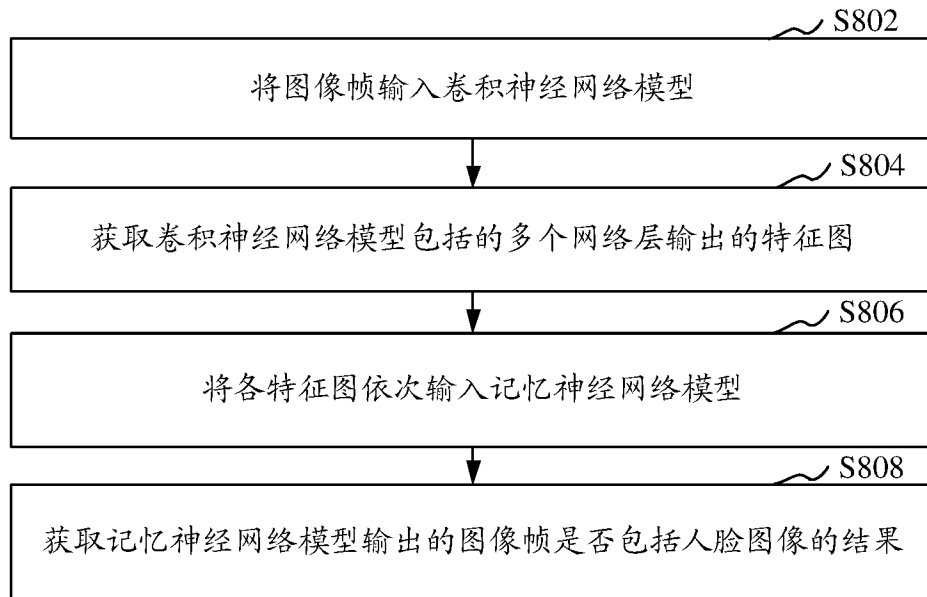


图 8

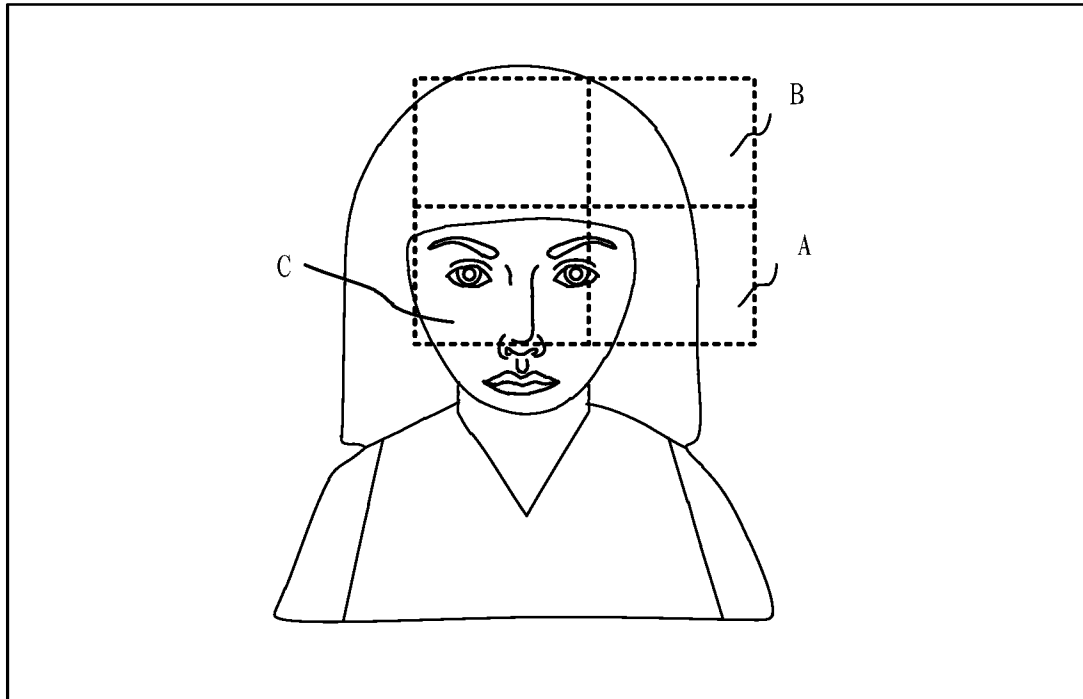


图 9

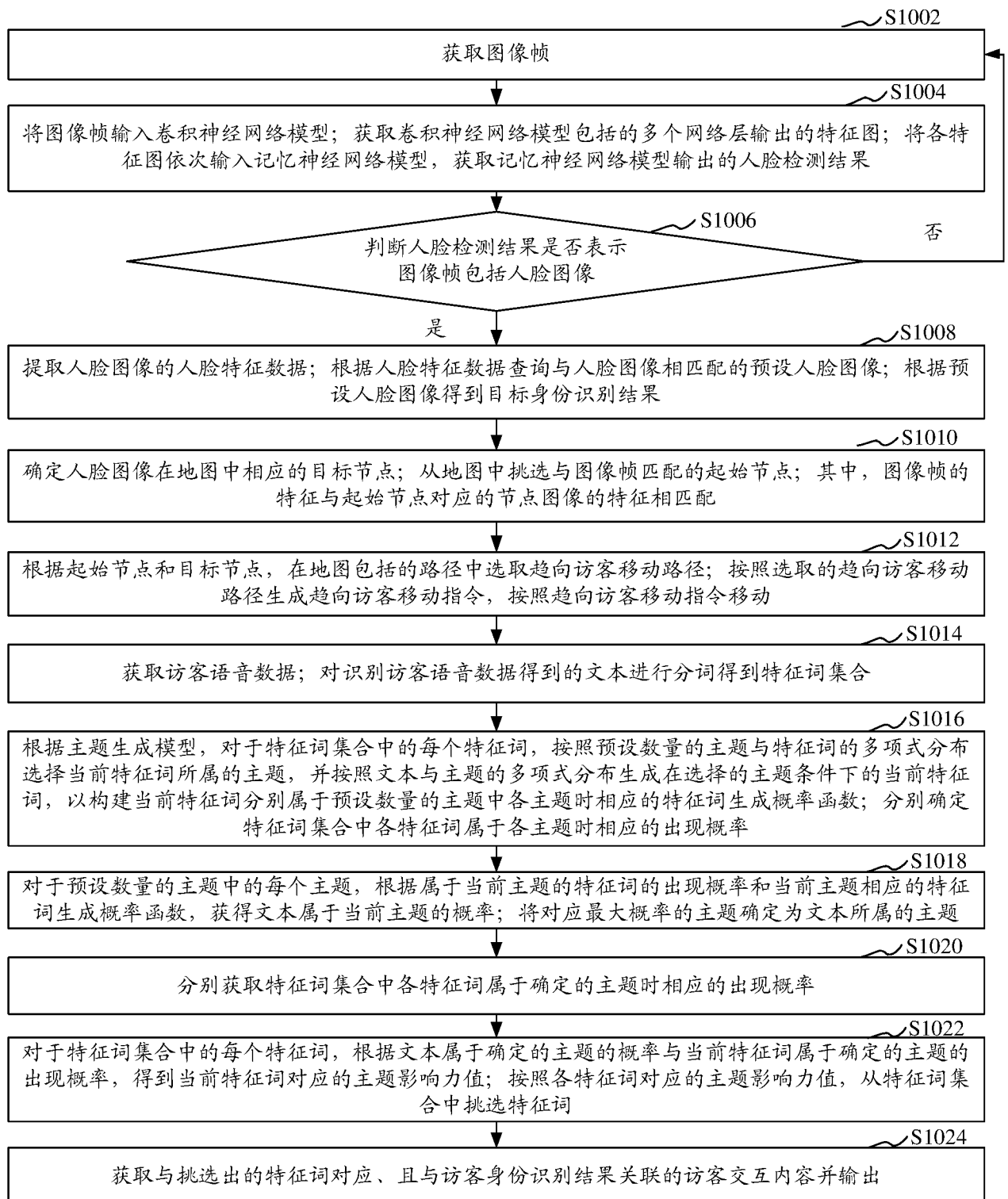


图 10

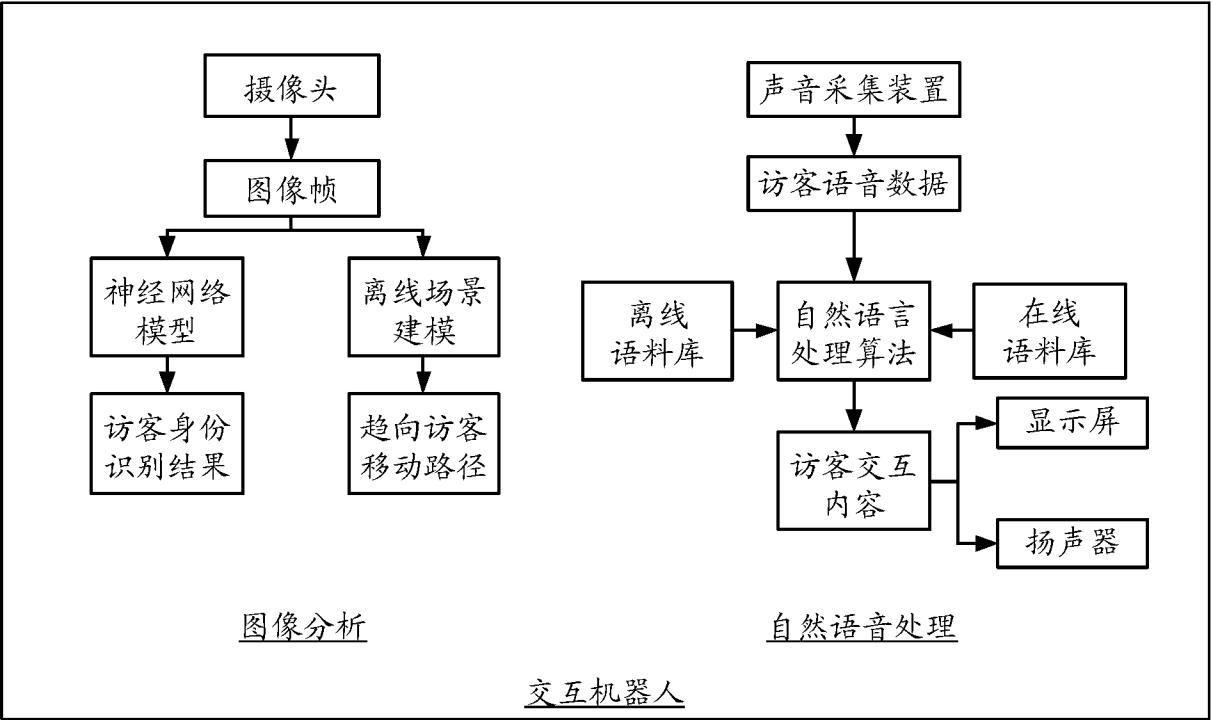


图 11

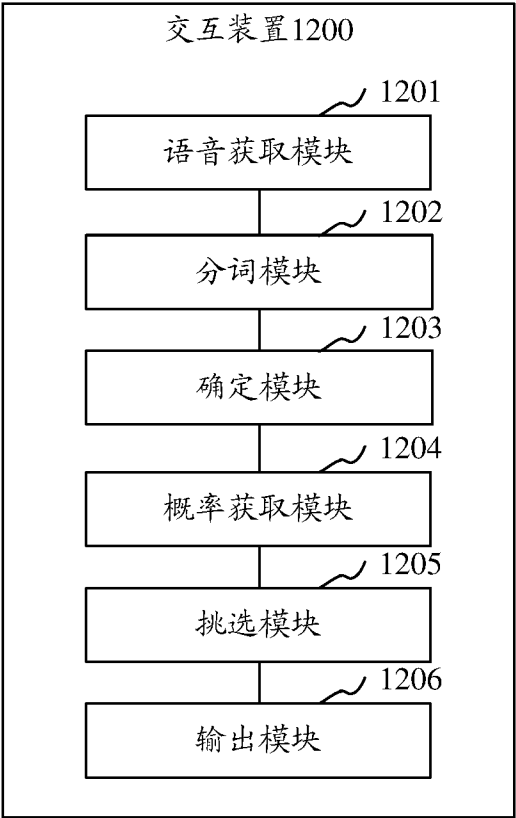


图 12

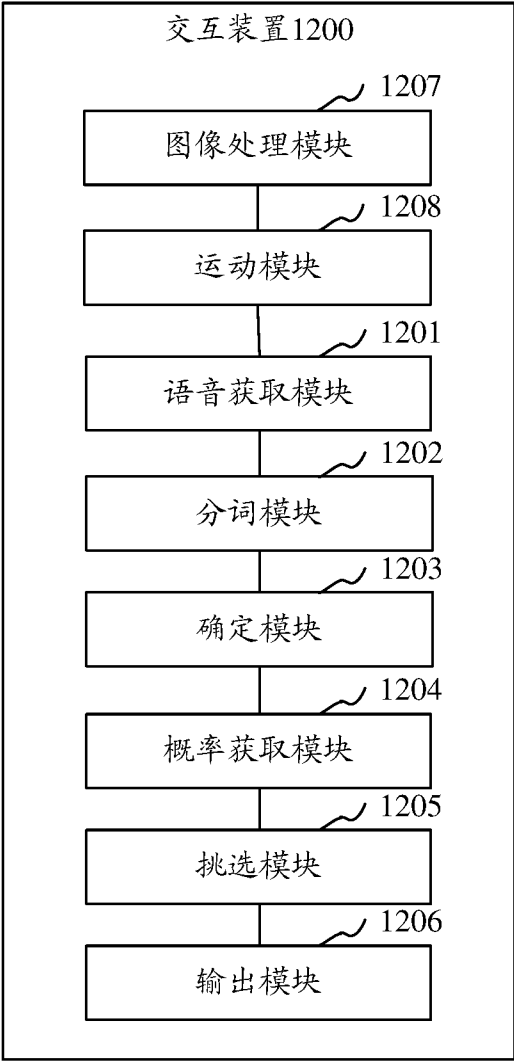


图 13

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/095836

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30(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)

G06F17/-

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, WOTXT, EPTXT: 交互, 语音, 文本, 主题, 特征词, 概率, interact, voice, speech, text, topic, subject, characteristic word, feature word, probability

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107423398 A (TENCENT TECHNOLOGY (SHANGHAI) CO., LTD.) 01 December 2017 (2017-12-01) entire document	1-19
Y	CN 106095834 A (EMOTIBOT TECHNOLOGIES LIMITED) 09 November 2016 (2016-11-09) description, paragraphs 52-71 and 100-101	1-19
Y	CN 105425970 A (SHENZHEN LINGYANG MICRO SERVER ROBOT TECHNOLOGY CO., LTD.) 23 March 2016 (2016-03-23) description, paragraphs 69-70 and 122	1-19
A	US 2002087321 A1 (FUJITSU LTD.) 04 July 2002 (2002-07-04) entire document	1-19



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

09 September 2018

Date of mailing of the international search report

30 September 2018

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/095836

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107423398	A	01 December 2017	None			
CN	106095834	A	09 November 2016	None			
CN	105425970	A	23 March 2016	None			
US	2002087321	A1	04 July 2002	JP	3857047	B2	13 December 2006
				US	7107217	B2	12 September 2006
				JP	2002202882	A	19 July 2002

国际检索报告

国际申请号

PCT/CN2018/095836

A. 主题的分类 G06F 17/30 (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) G06F17/- 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, WOTXT, EPTXT:交互, 语音, 文本, 主题, 特征词, 概率, interact, voice, speech, text, topic, subject, characteristic word, feature word, probability																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 107423398 A (腾讯科技上海有限公司) 2017年 12月 1日 (2017 - 12 - 01) 全文</td> <td>1-19</td> </tr> <tr> <td>Y</td> <td>CN 106095834 A (竹间智能科技上海有限公司) 2016年 11月 9日 (2016 - 11 - 09) 说明书第52-71、100-101段</td> <td>1-19</td> </tr> <tr> <td>Y</td> <td>CN 105425970 A (深圳羚羊微服机器人科技有限公司) 2016年 3月 23日 (2016 - 03 - 23) 说明书第69-70、122段</td> <td>1-19</td> </tr> <tr> <td>A</td> <td>US 2002087321 A1 (FUJITSU LTD) 2002年 7月 4日 (2002 - 07 - 04) 全文</td> <td>1-19</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 107423398 A (腾讯科技上海有限公司) 2017年 12月 1日 (2017 - 12 - 01) 全文	1-19	Y	CN 106095834 A (竹间智能科技上海有限公司) 2016年 11月 9日 (2016 - 11 - 09) 说明书第52-71、100-101段	1-19	Y	CN 105425970 A (深圳羚羊微服机器人科技有限公司) 2016年 3月 23日 (2016 - 03 - 23) 说明书第69-70、122段	1-19	A	US 2002087321 A1 (FUJITSU LTD) 2002年 7月 4日 (2002 - 07 - 04) 全文	1-19
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
PX	CN 107423398 A (腾讯科技上海有限公司) 2017年 12月 1日 (2017 - 12 - 01) 全文	1-19															
Y	CN 106095834 A (竹间智能科技上海有限公司) 2016年 11月 9日 (2016 - 11 - 09) 说明书第52-71、100-101段	1-19															
Y	CN 105425970 A (深圳羚羊微服机器人科技有限公司) 2016年 3月 23日 (2016 - 03 - 23) 说明书第69-70、122段	1-19															
A	US 2002087321 A1 (FUJITSU LTD) 2002年 7月 4日 (2002 - 07 - 04) 全文	1-19															
☐ 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 9月 9日		国际检索报告邮寄日期 2018年 9月 30日															
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (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 王国海 电话号码 86-(20)-28958137															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/095836

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	107423398	A	2017年 12月 1日	无			
CN	106095834	A	2016年 11月 9日	无			
CN	105425970	A	2016年 3月 23日	无			
US	2002087321	A1	2002年 7月 4日	JP	3857047	B2	2006年 12月 13日
				US	7107217	B2	2006年 9月 12日
				JP	2002202882	A	2002年 7月 19日