

(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 4 月 18 日 (18.04.2019)



(10) 国际公布号
WO 2019/071607 A1

(51) 国际专利分类号:
G06F 17/30 (2006.01)

(21) 国际申请号: PCT/CN2017/106168

(22) 国际申请日: 2017 年 10 月 13 日 (13.10.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710931504.9 2017年10月9日 (09.10.2017) CN

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 隋志成 (SUI, Zhicheng); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 李艳明 (LI, Yanming); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 北京中博世达专利商标代理有限公司 (BEIJING ZBSD PATENT & TRADEMARK AGENT LTD.); 中国北京市海淀区交大东路 31 号 11 号楼 8 层, Beijing 100044 (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 INFORMATION PROCESSING METHOD AND DEVICE, AND TERMINAL

(54) 发明名称: 一种语音信息处理方法、装置及终端

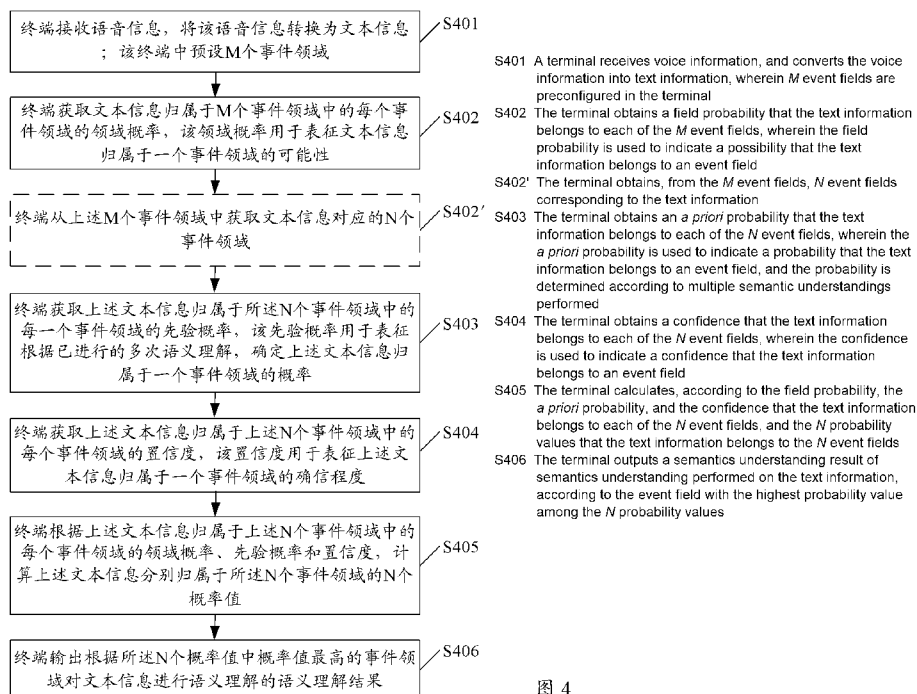


图 4

(57) Abstract: Embodiments of the present application provide a voice information processing method and device and a terminal, relating to the technical field of computers, for improving the efficiency of a terminal executing an event corresponding to a semantics understanding result, and reducing network traffic consumed by semantics understanding. The method comprises: a terminal receiving voice information, and converting the voice information into text information; obtaining a field probability that the text information belongs to each of preset M event fields; obtaining an *a priori* probability that the text information belongs to each of N event fields,

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))。

wherein $N \leq M$; obtaining a confidence that the text information belongs to each of the N event fields; according to the field probability, the *a priori* probability, and the confidence that the text information belongs to each of the N event fields, calculating N probability values that the text information belongs to the N event fields; and outputting a semantics understanding result of semantics understanding performed on the text information, according to the event field with the highest probability value among the N probability values.

(57) 摘要: 本申请实施例提供一种语音信息处理方法、装置及终端, 涉及计算机技术领域, 可以提高终端执行语义理解结果对应的事件的效率, 并节省进行语义理解消耗的网络流量。具体方案包括: 终端接收语音信息, 将该语音信息转换为文本信息; 获取文本信息归属于预设 M 个事件领域中的每个事件领域的领域概率; 获取文本信息归属于 N 个事件领域中的每个事件领域的先验概率, $N \leq M$; 获取文本信息归属于 N 个事件领域中的每个事件领域的置信度; 根据文本信息归属于 N 个事件领域中的每个事件领域的领域概率、先验概率和置信度, 计算文本信息分别归属于 N 个事件领域的 N 个概率值; 输出根据 N 个概率值中概率值最高的事件领域对文本信息进行语义理解的语义理解结果。

一种语音信息处理方法、装置及终端

5 本申请要求于 2017 年 10 月 09 日提交中国专利局、申请号为 201710931504.9、申请名称为“一种语音信息处理方法及装置”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

10 本申请实施例涉及计算机技术领域，尤其涉及一种语音信息处理方法、装置及终端。

背景技术

15 随着电子技术的发展，智能终端的功能越来越多，如终端可以为用户提供语音对话功能，即终端可以接收用户输入的语音信息（如“打开地图应用”），对该语音信息进行语义理解，然后执行语义理解结果对应的事件（开启终端中的地图应用，如百度地图）。

一般而言，由于终端的处理能力有限，终端可以将接收的语音信息发送至云端服务器，由云端服务器对该语音信息进行语义理解，得到语义理解结果；然后，云端服务器可以指示终端执行语义理解结果对应的事件。

20 但是，在上述语音识别的过程中，需要终端与云端服务器进行至少两次数据交互，而终端与云端服务器的数据交互可能会因为网络故障等原因，造成终端不能及时执行语义理解结果对应的事件。并且，由于语音信息的数据量一般较大，因此会消耗大量的网络流量。

发明内容

25 本申请实施例提供一种语音信息处理方法、装置及终端，可以节省由云端服务器进行语义理解时消耗的网络流量。

第一方面，本申请实施例提供一种语音信息处理方法，该语音信息处理方法包括：终端接收语音信息，将该语音信息转换为文本信息；该终端中预设 M 个事件领域；获取上述文本信息归属于 M 个事件领域中的每个事件领域的领域概率，上述文本信息归属于一个事件领域的领域概率用于表征该文本信息归属于该事件领域的可能性；获取上述文本信息归属于 N 个事件领域中的每一个事件领域的先验概率，上述文本信息归属于一个事件领域的先验概率用于表征根据已进行的多次语义理解，确定该文本信息归属于该事件领域的概率，上述 N 个事件领域为上述 M 个事件领域中的 N 个事件领域， N 小于或等于 M ；获取上述文本信息归属于上述 N 个事件领域中的每个事件领域的置信度，上述文本信息归属于一个事件领域的置信度用于表征上述文本信息归属于该事件领域的确信程度；根据上述文本信息归属于上述 N 个事件领域中的每个事件领域的领域概率、先验概率和置信度，计算上述文本信息分别归属于所述 N 个事件领域的 N 个概率值；输出根据上述 N 个概率值中概率值最高的事件领域对文本信息进行语义理解的语义理解结果。可选的，所述输出根据上述 N 个概率值中概率值最高的事件

30

35

领域对文本信息进行语义理解的语义理解结果，可以被替换为：将根据上述 N 个概率值中概率值最高的事件领域对文本信息进行语义理解的语义理解结果作为最终的语义理解结果。

其中，文本信息归属于一个事件领域的先验概率：用于表征历史数据中，文本信息归属于该事件领域的概率；文本信息归属于一个事件领域的领域概率：用于表征该文本信息归属于该事件领域的可能性；文本信息归属于一个事件领域的置信度：用于表征该文本信息归属于该事件领域的确信程度。本申请实施例在选择处理文本信息的事件领域时，不仅参考了对文本信息中包括的词汇进行分析得到的领域概率，还参考了文本信息归属于事件领域的先验概率，以及用于表征该文本信息归属于该事件领域的确信程度的置信度；因此，可以提高选择的事件领域的准确性，进而可以提高语义理解结果的准确性，从而可以提高终端执行事件与用户输入的语音信息指示终端执行的事件的符合度，可以提高用户体验。

在一种可能的设计方法中，当 N 小于 M 时，上述 N 个事件领域是上述预设 M 个事件领域中，领域概率按照由高至低的顺序排列在前 N 位的 N 个事件领域， $N \geq 2$ 。具体的，终端可以按照领域概率由高至低的顺序，从上述 M 个事件领域中选择出领域概率排列在前 N 位的事件领域。

可以理解，终端从 M 个事件领域中选择出 N 个事件领域后，只需要计算上述文本信息归属于上述 N 个事件领域的先验概率和置信度，而不需要计算文本信息归属于 M 个事件领域中所有事件领域的先验概率和置信度，可以减少终端进行语音信息处理时的计算量，提高计算效率。

在另一种可能的设计方法中，在上述终端获取所述文本信息归属于 M 个事件领域中的每个事件领域的领域概率之后，本申请实施例的方法还包括：终端在上述 N 个事件领域，分别对上述文本信息进行语义理解，得到 N 个语义理解结果。

其中，终端在对文本信息进行领域识别后，可以将文本信息传输至识别到的事件领域的对话引擎，由对话引擎对该文本信息进行语义理解，得到语义理解结果。或者，本实施例可以不限定终端进行领域识别和语义理解的顺序，可以同时或基本同时进行领域识别和语义理解，也可以先进行语义理解后进行领域识别。

在另一种可能的设计方法中，上述 M 个事件领域中的每个事件领域对应于一个关键字模型，该关键字模型中包括：对应事件领域的多个关键字。具体的，上述终端获取上述文本信息归属于上述 N 个事件领域中的每个事件领域的置信度，可以包括：终端对所述文本信息进行分词处理，并提取至少一个分词；获取该至少一个分词对应的关键字在上述每个事件领域的关键字模型中的分布信息；根据该分布信息，计算上述文本信息归属于上述 N 个事件领域中的每个事件领域的置信度。

在另一种可能的设计方法中，上述终端获取上述文本信息归属于 M 个事件领域中的每个事件领域的领域概率，包括：终端对上述文本信息进行分词处理，并提取至少一个分词；从上述每个事件领域对应的数据库模型中查找上述至少一个分词对应的特征，上述数据库模型中包括多个特征、每个特征的权重及每个特征对应的分词，上述权重用于指示上述权重对应的特征归属于上述数据库模型中对应的事件领域的概率；其中，每个事件领域对应一个数据库模型；根据从上述每个事件领域对应的数据库模

型中查找到的特征的权重，计算上述文本信息归属于上述每个事件领域的领域概率。

其中，在上述特征数据库中，同一分词在不同事件领域的数据库模型中的特征相同，即在特征数据库中，分词的特征可以唯一标识该分词。但是，同一分词在不同事件领域中的权重不同。

5 在另一种可能的设计方法中，上述 M 个事件领域中的每个事件领域对应于一个关键字模型，上述关键字模型中包括：多个关键字和每个关键字指示文本信息归属于上述关键字模型对应的事件领域的概率。上述终端获取上述文本信息归属于 M 个事件领域中的每个事件领域的领域概率，包括：从上述文本信息中识别至少一个关键字；从上述每个事件领域对应的关键字模型中获取上述至少一个关键字分别指示的概率；根据
10 上述至少一个关键字分别指示的概率，计算上述文本信息归属于上述每个事件领域的领域概率。

其中，由于至少一个关键字中可能包括各个事件领域的关键字模型中的关键字，而每个关键字在不同的事件领域的关键字模型中，可以指示上述文本信息归属于对应事件领域的概率；因此，根据文本信息中包括的各个事件领域的关键字所指示的概率，
15 可以计算得到文本信息归属于各个事件领域的领域概率。

在另一种可能的设计方法中，本申请实施例的方法还可以包括：上述终端输出上述后语义理解结果之后，根据上述语义理解结果，执行语义理解结果对应的操作。

第二方面，本申请实施例提供一种语音信息处理装置，该语音信息处理装置包括：
20 接收单元、转换单元、第一获取单元、第二获取单元、第三获取单元、计算单元和输出单元。其中，上述接收单元，用于接收语音信息。上述转换单元，用于将上述接收单元接收的上述语音信息转换为文本信息；上述终端中预设 M 个事件领域。上述第一获取单元，用于获取上述转换单元转换得到的上述文本信息归属于 M 个事件领域中的每个事件领域的领域概率，上述领域概率用于表征上述文本信息归属于一个事件领域的可能性。上述第二获取单元，用于获取上述转换单元转换得到的上述文本信息归属
25 于上述 N 个事件领域中的每一个事件领域的先验概率，上述先验概率用于表征根据已进行的多次语义理解，确定上述文本信息归属于一个事件领域的概率，上述 N 个事件领域为上述 M 个事件领域中的 N 个事件领域， N 小于或等于 M 。上述第三获取单元，用于获取上述转换单元转换得到的上述文本信息归属于上述 N 个事件领域中的每个事件领域的置信度，上述置信度用于表征上述文本信息归属于一个事件领域的确信程度。
30 上述计算单元，用于根据上述第一获取单元获取的上述文本信息归属于上述 N 个事件领域中的每个事件领域的领域概率、上述第二获取单元获取的先验概率和上述第三获取单元获取的置信度，计算上述文本信息分别归属于上述 N 个事件领域的 N 个概率值。上述输出单元，用于输出根据上述计算单元计算得到的上述 N 个概率值中概率值最高的事件领域对文本信息进行语义理解的语义理解结果。

35 在一种可能的设计方法中，当 N 小于 M 时，上述 N 个事件领域是上述预设 M 个事件领域中，领域概率按照由高至低的顺序排列在前 N 位的 N 个事件领域， $N \geq 2$ 。

在另一种可能的设计方法中，上述语音信息处理装置还包括：语义理解单元。该语义理解单元，用于在上述第一获取单元获取上述文本信息归属于 M 个事件领域中的每个事件领域的领域概率之后，在上述 N 个事件领域，分别对上述文本信息进行语义

理解，得到 N 个语义理解结果。

在另一种可能的设计方法中，上述语音信息处理装置还包括：存储单元。该存储单元，用于保存上述 M 个事件领域中的每个事件领域对应的关键字模型，上述关键字模型中包括：对应事件领域的多个关键字。上述第三获取单元，具体用于：对上述文本信息进行分词处理，并提取至少一个分词；获取上述至少一个分词对应的关键字在上述存储单元保存的上述每个事件领域的关键字模型中的分布信息；根据上述分布信息，计算上述文本信息归属于上述 N 个事件领域中的每个事件领域的置信度。

在另一种可能的设计方法中，上述第一获取单元，具体用于：对上述文本信息进行分词处理，并提取至少一个分词；从上述每个事件领域对应的数据库模型中查找上述至少一个分词对应的特征，上述数据库模型中包括多个特征、每个特征的权重及每个特征对应的分词，上述权重用于指示上述权重对应的特征归属于上述数据库模型中对应的事件领域的概率；其中，每个事件领域对应一个数据库模型；根据从上述每个事件领域对应的数据库模型中查找到的特征的权重，计算上述文本信息归属于上述每个事件领域的领域概率。

在另一种可能的设计方法中，上述语音信息处理装置还包括：存储单元。该存储单元，用于保存上述 M 个事件领域中的每个事件领域对应的关键字模型，上述关键字模型中包括：多个关键字和每个关键字指示文本信息归属于上述关键字模型对应的事件领域的概率。上述第一获取单元，具体用于：从上述文本信息中识别至少一个关键字；从上述每个事件领域对应的关键字模型中获取上述至少一个关键字分别指示的概率；根据上述至少一个关键字分别指示的概率，计算上述文本信息归属于上述每个事件领域的领域概率。

在另一种可能的设计方法中，上述语音信息处理装置还包括：执行单元。该执行单元，用于在上述输出单元输出上述语义理解结果之后，根据上述语义理解结果，执行上述语义理解结果对应的操作。

第三方面，本申请实施例提供一种终端，该终端包括：一个或多个处理器；一个或多个存储器，所述一个或多个存储器中存储有一个或多个计算机程序，所述一个或多个计算机程序包括指令，当所述指令被所述一个或多个处理器执行时，使得所述终端执行如第一方面及其任一种可能的设计方法所述的语音信息处理方法。

第四方面，本申请实施例提供一种电子设备，所述电子设备包括执行如第一方面及其任一种可能的设计方法所述的语音信息处理方法的装置。

第五方面，本申请实施例提供一种包含指令的计算机程序产品，当所述计算机程序产品在电子设备上运行时，使得所述电子设备执行如第一方面及其任一种可能的设计方法所述的语音信息处理方法。

第六方面，本申请实施例提供一种计算机可读存储介质，该计算机可读存储介质包括指令，当所述指令在电子设备上运行时，使得所述电子设备执行如第一方面及其任一种可能的设计方法所述的语音信息处理方法。

可以理解地，上述提供的第二方面所述的装置、第三方面所述的终端、第四方面所述的电子设备，第五方面所述计算机程序产品，以及第六方面所述的计算机存储介质均用于执行上文所提供的对应的方法，因此，其所能达到的有益效果可参考上文所

提供的对应的方法中的有益效果，此处不再赘述。

附图说明

- 图 1 为本申请实施例提供的一种终端的硬件结构示意图；
- 5 图 2 为本申请实施例提供的一种用于进行语音信息处理的架构示意图；
- 图 3 为本申请实施例提供的一种语音信息处理方法流程图一；
- 图 4 为本申请实施例提供的一种语音信息处理方法流程图二；
- 图 5 为本申请实施例提供的一种语音信息处理方法的语义理解结果的历史记录示意图；
- 10 图 6 为本申请实施例提供的一种语音信息处理方法流程图三；
- 图 7 为本申请实施例提供的一种关键字数据库的实例示意图一；
- 图 8 为本申请实施例提供的一种语音信息处理方法的执行过程实例示意图一；
- 图 9 为本申请实施例提供的一种关键字数据库的实例示意图二；
- 图 10 为本申请实施例提供的一种语音信息处理方法流程图四；
- 15 图 11 为本申请实施例提供的一种语音信息处理方法流程图五；
- 图 12 为本申请实施例提供的一种特征数据库的实例示意图一；
- 图 13 为本申请实施例提供的一种特征数据库的实例示意图二；
- 图 14 为本申请实施例提供的一种语音信息处理方法的执行过程实例示意图二；
- 图 15 为本申请实例提供的一种语音信息处理装置的结构组成示意图一；
- 20 图 16 为本申请实例提供的一种语音信息处理装置的结构组成示意图二；
- 图 17 为本申请实例提供的一种终端的结构组成示意图。

具体实施方式

- 以下，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者更多个该特征。在本申请实施例的描述中，除非另有说明，“多个”的含义是两个或两个以上。
- 25

- 本申请实施例提供一种语音信息处理方法及终端，可以应用于终端与用户进行语音对话的过程中。具体应用于终端接收用户输入的语音信息，对该语音信息进行语义理解，并执行语义理解结果对应的事件的过程中。例如用户通过语音控制终端的过程中。
- 30

- 其中，本申请实施例中对语音信息进行语义理解可以包括：将语音信息转换为文本信息，然后分析该文本信息，识别出该文本信息所指示终端所执行的事件。例如，当终端接收到用户输入的语音信息“提醒我 22:00 打开飞行模式”时，终端可以将该语音信息转换成文本信息“提醒我 22:00 打开飞行模式”，然后识别出该文本信息所指示终端执行的事件为“在 22:00 向用户发出“打开飞行模式”的提醒”，而非直接“打开飞行模式”。
- 35

常规方案在上述语音识别的过程中，会出现以下问题：终端与云端服务器的数据交互可能会因为网络故障等原因，造成终端不能及时执行语义理解结果对应的事件；

由于语音信息的数据量一般较大，常规方案会消耗大量的网络流量。为了解决常规方案中的问题，本申请实施例提供的语音信息处理方法，可以由终端执行上述语义理解。

一般而言，终端在执行上述语义理解时，只是简单的对转换得到的文本信息中包括的词汇进行分析，确定出该文本信息所归属的一个事件领域，即判断出该文本信息所归属的事件领域，然后由该文本信息所归属的事件领域的对话引擎，采用该事件领域的语义理解算法对该文本信息进行语义理解，然后执行语义理解结果对应的事件。

但是，存在的问题是，简单的对文本信息中包括的词汇进行分析，确定出的事件领域可能并不准确；由不准确的事件领域的对话引擎，采用该不准确的事件领域的语义理解算法对该文本信息进行语义理解，得到的语义理解结果也不准确。由此，则可能会导致终端执行的语义理解结果对应的事件，与用户输入的语音信息指示终端执行的事件不同，影响用户体验。

本申请实施例中，为了提高终端执行语义理解结果对应的事件的效率，节省由云端服务器进行语义理解时消耗的网络流量的同时，提高终端对语音信息进行语义理解的准确性，可以在将语音信息转换为文本信息后，根据终端进行语义理解的历史数据，获取该文本信息归属于每个事件领域的先验概率；文本信息归属于一个事件领域的先验概率用于表征历史数据中，文本信息归属于该事件领域的概率。然后，分析该文本信息，获取所述文本信息归属于每个事件领域的领域概率；该文本信息归属于一个事件领域的领域概率用于表征该文本信息归属于该事件领域的可能性。随后，终端可以计算文本信息归属于每个事件领域的置信度；文本信息归属于一个事件领域的置信度用于表征该文本信息归属于该事件领域的确信程度。其次，终端可以根据上述文本信息归属于一个事件领域的先验概率、领域概率和置信度，计算该文本信息归属于该事件领域的概率值；进而得到该文本信息归属于每个事件领域的概率值。最后，终端可以将概率值最高的事件领域的对话引擎对该文本信息进行语义理解，得到的语义理解结果作为该文本信息（即上述语音信息）的语义理解结果，终端可以执行该语义理解结果对应的事件。

其中，由于本申请实施例在选择处理文本信息的事件领域时，不仅参考了对文本信息中包括的词汇进行分析得到的领域概率，还参考了文本信息归属于事件领域的先验概率，以及用于表征该文本信息归属于该事件领域的确信程度的置信度；因此，可以提高选择的事件领域的准确性，进而可以提高语义理解结果的准确性，从而提高终端执行事件与用户输入的语音信息指示终端执行的事件的符合度，可以提高用户体验。

其中，本申请实施例中的终端可以是允许用户通过输入语音信息指示终端执行相关操作事件的手机（如图 1 所示的手机 100）、平板电脑、个人计算机（Personal Computer, PC）、个人数字助理（personal digital assistant, PDA）、智能手表、上网本、可穿戴电子设备等，本申请实施例对该设备的具体形式不做特殊限制。

其中，本申请实施例中文本信息所归属的事件领域是指对该文本信息进行语义理解后，语义理解结果所指示终端执行的事件所归属的领域。例如，本申请实施例中的事件领域可以包括音乐领域、设置领域、应用程序（Application, APP）领域等。举例来说，“播放歌曲 a”和“播放下一曲”等文本信息归属于音乐领域，“调低屏幕

亮度”和“打开飞行模式”等文本信息归属于设置领域，“打开微信”和“地图导航至 A 街道 10 号”等文本信息归属于 APP 领域。

如图 1 所示，以手机 100 作为上述终端举例，手机 100 具体可以包括：处理器 101、射频（Radio Frequency，RF）电路 102、存储器 103、触摸屏 104、蓝牙装置 105、一个或多个传感器 106、Wi-Fi 装置 107、定位装置 108、音频电路 109、外设接口 110 以及电源装置 111 等部件。这些部件可通过一根或多根通信总线或信号线（图 1 中未示出）进行通信。本领域技术人员可以理解，图 1 中示出的硬件结构并不构成对手机的限定，手机 100 可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。

下面结合图 1 对手机 100 的各个部件进行具体的介绍：

处理器 101 是手机 100 的控制中心，利用各种接口和线路连接手机 100 的各个部分，通过运行或执行存储在存储器 103 内的应用程序，以及调用存储在存储器 103 内的数据，执行手机 100 的各种功能和处理数据。在一些实施例中，处理器 101 可包括一个或多个处理单元。在本申请实施例一些实施例中，上述处理器 101 还可以包括指纹验证芯片，用于对采集到的指纹进行验证。

射频电路 102 可用于在收发信息或通话过程中，无线信号的接收和发送。特别地，射频电路 102 可以将基站的下行数据接收后，给处理器 101 处理；另外，将涉及上行的数据发送给基站。通常，射频电路包括但不限于天线、至少一个放大器、收发信机、耦合器、低噪声放大器、双工器等。此外，射频电路 102 还可以通过无线通信和其他设备通信。所述无线通信可以使用任一通信标准或协议，包括但不限于全球移动通讯系统、通用分组无线服务、码分多址、宽带码分多址、长期演进、电子邮件、短消息服务等。

存储器 103 用于存储应用程序以及数据，处理器 101 通过运行存储在存储器 103 的应用程序以及数据，执行手机 100 的各种功能以及数据处理。存储器 103 主要包括存储程序区以及存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序（比如声音播放功能、图像播放功能等）；存储数据区可以存储根据使用手机 100 时所创建的数据（比如音频数据、电话本等）。此外，存储器 103 可以包括高速随机存取存储器（RAM），还可以包括非易失存储器，例如磁盘存储器件、闪存器件或其他易失性固态存储器件等。存储器 103 可以存储各种操作系统，例如，iOS®操作系统，Android®操作系统等。上述存储器 103 可以是独立的，通过上述通信总线与处理器 101 相连接；存储器 103 也可以和处理器 101 集成在一起。

触摸屏 104 具体可以包括触控板 104-1 和显示器 104-2。

其中，触控板 104-1 可采集手机 100 的用户在其上或附近的触摸事件（比如用户使用手指、触控笔等任何适合的物体在触控板 104-1 上或在触控板 104-1 附近的操作），并将采集到的触摸信息发送给其他器件（例如处理器 101）。其中，用户在触控板 104-1 附近的触摸事件可以称之为悬浮触控；悬浮触控可以是指，用户无需为了选择、移动或拖动目标（例如图标等）而直接接触触控板，而只需用户位于设备附近以便执行所想要的功能。此外，可以采用电阻式、电容式、红外线

以及表面声波等多种类型来实现触控板 104-1。

显示器（也称为显示屏）104-2 可用于显示由用户输入的信息或提供给用户的信息以及手机 100 的各种菜单。可以采用液晶显示器、有机发光二极管等形式来配置显示器 104-2。触控板 104-1 可以覆盖在显示器 104-2 之上，当触控板 104-1 检测到在其上或附近的触摸事件后，传送给处理器 101 以确定触摸事件的类型，随后处理器 101 可以根据触摸事件的类型在显示器 104-2 上提供相应的视觉输出。虽然在图 1 中，触控板 104-1 与显示屏 104-2 是作为两个独立的部件来实现手机 100 的输入和输出功能，但是在某些实施例中，可以将触控板 104-1 与显示屏 104-2 集成而实现手机 100 的输入和输出功能。可以理解的是，触摸屏 104 是由多层材料堆叠而成，本申请实施例实施例中只展示出了触控板（层）和显示屏（层），其他层在本申请实施例实施例中不予记载。另外，触控板 104-1 可以以全面板的形式配置在手机 100 的正面，显示屏 104-2 也可以以全面板的形式配置在手机 100 的正面，这样在手机的正面就能够实现无边框的结构。

另外，手机 100 还可以具有指纹识别功能。例如，可以在手机 100 的背面（例如后置摄像头的下方）配置指纹识别器 112，或者在手机 100 的正面（例如触摸屏 104 的下方）配置指纹识别器 112。又例如，可以在触摸屏 104 中配置指纹采集器件 112 来实现指纹识别功能，即指纹采集器件 112 可以与触摸屏 104 集成在一起来实现手机 100 的指纹识别功能。在这种情况下，该指纹采集器件 112 配置在触摸屏 104 中，可以是触摸屏 104 的一部分，也可以以其他方式配置在触摸屏 104 中。本申请实施例实施例中的指纹采集器件 112 的主要部件是指纹传感器，该指纹传感器可以采用任何类型的感测技术，包括但不限于光学式、电容式、压电式或超声波传感技术等。

手机 100 还可以包括蓝牙装置 105，用于实现手机 100 与其他短距离的设备（例如手机、智能手表等）之间的数据交换。本申请实施例实施例中的蓝牙装置可以是集成电路或者蓝牙芯片等。

手机 100 还可以包括至少一种传感器 106，比如光传感器、运动传感器以及其他传感器。具体地，光传感器可包括环境光传感器及接近传感器，其中，环境光传感器可根据环境光线的明暗来调节触摸屏 104 的显示器的亮度，接近传感器可在手机 100 移动到耳边时，关闭显示器的电源。作为运动传感器的一种，加速计传感器可检测各个方向上（一般为三轴）加速度的大小，静止时可检测出重力的大小及方向，可用于识别手机姿态的应用（比如横竖屏切换、相关游戏、磁力计姿态校准）、振动识别相关功能（比如计步器、敲击）等；至于手机 100 还可配置的陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器，在此不再赘述。

无线保真（Wireless Fidelity，Wi-Fi）装置 107，用于为手机 100 提供遵循 Wi-Fi 相关标准协议的网络接入，手机 100 可以通过 Wi-Fi 装置 107 接入到 Wi-Fi 接入点，进而帮助用户收发电子邮件、浏览网页和访问流媒体等，它为用户提供了无线的宽带互联网访问。在其他一些实施例中，该 Wi-Fi 装置 107 也可以作为 Wi-Fi 无线接入点，可以为其他设备提供 Wi-Fi 网络接入。

定位装置 108，用于为手机 100 提供地理位置。可以理解的是，该定位装置 108 具体可以是全球定位系统（Global Positioning System，GPS）或北斗卫星导航系统、俄罗斯 GLONASS 等定位系统的接收器。定位装置 108 在接收到上述定位系统发送的地理位置后，将该信息发送给处理器 101 进行处理，或者发送给存储器 103 进行保存。在另外的一些实施例中，该定位装置 108 还可以是辅助全球卫星定位系统（Assisted Global Positioning System，AGPS）的接收器，AGPS 系统通过作为辅助服务器来协助定位装置 108 完成测距和定位服务，在这种情况下，辅助定位服务器通过无线通信网络与设备例如手机 100 的定位装置 108（即 GPS 接收器）通信而提供定位协助。在另外的一些实施例中，该定位装置 108 也可以是基于 Wi-Fi 接入点的定位技术。由于每一个 Wi-Fi 接入点都有一个全球唯一的媒体访问控制（Media Access Control，MAC）地址，设备在开启 Wi-Fi 的情况下即可扫描并收集周围的 Wi-Fi 接入点的广播信号，因此可以获取到 Wi-Fi 接入点广播出来的 MAC 地址；设备将这些能够标示 Wi-Fi 接入点的数据（例如 MAC 地址）通过无线通信网络发送给位置服务器，由位置服务器检索出每一个 Wi-Fi 接入点的地理位置，并结合 Wi-Fi 广播信号的强弱程度，计算出该设备的地理位置并发送到该设备的定位装置 108 中。

音频电路 109、扬声器 113、麦克风 114 可提供用户与手机 100 之间的音频接口。音频电路 109 可将接收到的音频数据转换后的电信号，传输到扬声器 113，由扬声器 113 转换为声音信号输出；另一方面，麦克风 114 将收集的声音信号转换为电信号，由音频电路 109 接收后转换为音频数据，再将音频数据输出至 RF 电路 102 以发送给比如另一手机，或者将音频数据输出至存储器 103 以便进一步处理。

外设接口 110，用于为外部的输入/输出设备（例如键盘、鼠标、外接显示器、外部存储器、用户识别模块卡等）提供各种接口。例如通过通用串行总线（Universal Serial Bus，USB）接口与鼠标连接，通过用户识别模块卡卡槽上的金属触点与电信运营商提供的用户识别模块卡（Subscriber Identification Module，SIM）卡进行连接。外设接口 110 可以被用来将上述外部的输入/输出外围设备耦接到处理器 101 和存储器 103。

在本发明实施例中，手机 100 可通过外设接口 110 与设备组内的其他设备进行通信，例如，通过外设接口 110 可接收其他设备发送的显示数据进行显示等，本发明实施例对此不作任何限制。

手机 100 还可以包括给各个部件供电的电源装置 111（比如电池和电源管理芯片），电池可以通过电源管理芯片与处理器 101 逻辑相连，从而通过电源装置 111 实现管理充电、放电、以及功耗管理等功能。

尽管图 1 未示出，手机 100 还可以包括摄像头（前置摄像头和/或后置摄像头）、闪光灯、微型投影装置、近场通信（Near Field Communication，NFC）装置等，在此不再赘述。

图 2 为本申请实施例提供的一种用于进行语音信息处理的架构示意图，该架构位于终端中。如图 2 所示，该架构包括中控层 201、对话引擎层 202 和算法层 203。

其中，中控层 201 包括：语音服务接口（Voice Service Interface，VSI）2011、

领域识别模块 2012、调度分发模块 2013 和汇总决策 (Decision Summary, DS) 模块 2014。

对话引擎层 202 中包括至少两个对话引擎。例如, 如图 2 所示, 对话引擎层 202 中包括对话引擎 1、对话引擎 2 和对话引擎 3。

5 算法层 203 包括: “模型和算法库” 2031、规则 (Rule) 库 2032、兴趣点 (Points Of Interests, POI) 库 2033 和状态模型 2034。

中控层 201 用于通过 VSI 2011 接收语音信息 (如从第三方应用接收语音信息), 然后将接收到语音信息传输至领域识别模块 2012。

请参考图 2 和图 3, 领域识别模块 2012 用于将接收到的语音信息转换为文本信息, 10 并对该文本信息进行初步领域识别, 识别出该文本信息可能的至少两个事件领域, 然后将识别结果传输至调度分发模块 2013; 其中, 领域识别模块 2012 可以调度算法层 203 中的“模型和算法库” 2031、规则 (Rule) 库 2032、POI 库 2033 和状态模型 2034, 对上述文本信息进行领域识别。

其中, “模型和算法库” 2031 中可以包括多个算法 (也称之为模型), 这多个算 15 法用于支持领域识别模块 2012 和对话引擎层 202 中的对话引擎 (如对话引擎 1) 对文本信息进行分析。举例来说, 如图 2 所示, 算法层 203 中的“模型和算法库” 2031 中包括: 逻辑回归/支持向量机 (Logistic Regression/Support Vector Machine, LR/SVM) 算法、词频-逆向文件频率 (Term Frequency-Inverse Document Frequency, TF-IDF) 算法、N-Gram/WS (Word Segment、分词) 算法、语义角色标注 (Semantic Role Label, 20 SRL) 算法、词性标注 (Part of Speech, POS) 算法、命名实体识别 (Named Entity Recognition, NER) 算法、条件随机场 (Conditional Random Field, CRF) 算法、统计机器翻译 (Statistic Machine Translation, SMT) 算法、深度强化学习网络 (Deep Reinforce learning Network, DRN) 算法、卷积/循环神经网络 (Convolution/Recurrent Neural Net, C/RNN) 算法和长短记忆网络 (Long Short Term Memory, LSTM) 算法等算法。其中, 25 N-Gram 是大词汇连续语音识别中常用的一种语言模型, 对中文而言, 可以称之为汉语语言模型 (Chinese Language Model, CLM)。该汉语语言模型可以利用语音信息的上下文中相邻词间的搭配信息, 可以实现语音信息到汉字 (即文本信息) 的自动转换。

算法层 203 中的 Rule 库 2032 中可以包括归属于各个事件领域的文本信息的语义 30 理解规则。例如, 如图 2 所示, Rule 库 2032 可以包括归属于 APP 领域的文本信息的语义理解规则、归属于设置领域的文本信息的语义理解规则和归属于音乐领域的文本信息的语义理解规则等。其中, Rule 库 2032 中一个事件领域的语义理解规则可以用于指示对归属于该事件领域的文本信息进行语义理解时, 从“模型和算法库” 2031 中要调用的算法。举例来说, Rule 库 2032 中 APP 领域的语义理解规则可以用于指示对归属于 APP 领域的文本信息进行语义理解时, 可以从“模型和算法库” 中调用 LR/SVM 35 算法和 F-IDF 算法。

以导航领域为例, POI 库 2033 可以是包括使用 Rule 库 2032 中的规则的对象名称 (如餐馆名称、学校名称等)、对象地址 (如餐馆地址、学校地址等)、经纬度、类别 (如学校、餐馆、政府机关、商场) 等信息的数据集合。例如, POI 库 2033 中可以包括 Rule 库 2032 中归属于音乐领域的文本信息的语义理解规则中的歌手和歌曲等。

其中，POI 库 2033 中按照不同的地址可以维护多个数据集合；或者，按照不同的类别可以维护多个数据集合。

状态模型 2034 是对话引擎层 202 中的对话引擎管理对话状态的模型，该状态模型 2034 可以是自定义的模型，如确定性模型，概率模型，马尔科夫模型等。状态模型 2034 可以在终端与用户对话过程中，提供不同对话状态之间的转移。例如，概率模型是指用户输入语音信息后，该语音信息对应的文本信息归属于导航领域的概率值大于预设值，则输入在导航领域对文本信息进行语义理解的结果。

其中，调度分发模块 2013 用于将上述文本信息分发给上述识别结果所指示的至少两个事件领域所对应的对话引擎（如对话引擎 1），由对应的对话引擎分别对该文本信息进行自然语言理解（Natural Language Understanding, NLU）、对话管理（Dialogue Management, DM）和自然语言处理（Natural Language Process, NLP）（即对话生成），以得到上述文本信息在对应领域的语音理解结果。

如图 2 所示，对话引擎层 202 中的每个对话引起对应于一个事件领域。例如，对话引擎 1 对应于设置领域，对话引擎 2 对应于 APP 领域，对话引擎 3 对应于音乐领域。以对话引擎 1 为例，各个事件领域的对话引擎中都可以包括：NLU 模块、DM 模块和 NLP 模块，用于对文本信息进行语义理解，得到语音理解结果。其中，各个对话引擎可以调用算法层 203 中的模型和算法库 2031、Rule 库 2032、POI 库 2033 中与该对话引擎对应的模型、算法和规则等，对上述文本信息进行语义理解。

随后，各个对话引擎可以将其得到的语义理解结果传输至 DS 模块 2014，由 DS 模块 2014 执行本申请实施例中的方法步骤，从上述多个对话引擎反馈的语义理解结果中，选择出文本信息的归属概率值最高的事件领域对应的语义理解结果（即如图 3 所示，最优的语义理解结果），然后将选择出的语义理解结果作为上述文本信息的语义理解结果，通过上述 VSI 接口反馈该语义理解结果。

可以理解，图 2 所示的中控层 201 和对话引擎层，以及一部分算法层 203 的功能可以集成在图 1 所示的手机 100 的处理器 101 中实现，图 2 所示的算法层 203 中的算法和规则等信息可以保存在图 1 所示的手机 100 的存储器 103 中。即图 2 所示的用于进行语音信息处理的架构可以位于图 1 所示的手机 100 中。

本申请实施例提供一种语音信息处理方法，如图 4 所示，该语音识别方法包括 S401-S406：

S401、终端接收语音信息，将该语音信息转换为文本信息；该终端中预设 M 个事件领域。

本申请实施例中，终端可以在接收到语音信息后，调用图 2 所示的算法层 203 中用于进行语音文本转换的算法（如 N-Gram/WS 算法），将语音信息转换为文本信息。或者，终端还可以通过调用语音文本转换（voice-to-text, speech-to-text）程序，将语音信息转换为文本信息。本申请实施例中，对该语音信息转换为文本信息的具体方式不作限定。其中，终端接收的语音信息，一般指的是用户发出的语音；即用户发出语音，然后终端接收到该语音后，执行后续的动作。

其中，终端中可以预设 M 个事件领域，如音乐领域、设置领域、应用程序 APP 领域等。其中， $M \geq 2$ 。

S402、终端获取文本信息归属于 M 个事件领域中的每个事件领域的领域概率，该领域概率用于表征文本信息归属于一个事件领域的可能性。

其中，终端可以通过图 2 所示的中控层 201 中的领域识别模块 2012，调用算法层 203 中用于对文本信息进行语义语法分析的算法，获取文本信息归属于 N 个事件领域中的各个事件领域的领域概率。其中，文本信息归属于一个事件领域的领域概率越高，该文本信息归属于该事件领域的可能性则越高。

可选的，终端可以在执行 S402 之后，终端可以从上述 M 个事件领域中获取文本信息对应的 N 个事件领域。具体的，如图 4 所示，在 S402 之后，本申请的方法还可以包括 S402'：

S402'、终端从上述 M 个事件领域中获取文本信息对应的 N 个事件领域。

需要说明的是，终端在执行 S402 之后，如果执行 S402' 从上述 M 个事件领域中获取文本信息对应的 N 个事件领域，那么本申请实施例中的 N 个事件领域则是上述 M 个事件领域中的部分事件领域， $N < M$ 。在这种情况下，上述 N 个事件领域是上述预设 M 个事件领域中，领域概率按照由高至低的顺序排列在前 N 位的 N 个事件领域， $N \geq 2$ 。即终端在执行 S402 后，可以按照领域概率由高至低的顺序，从上述 M 个事件领域中选择出领域概率排列在前 N 位的事件领域。示例性的，假设终端中预先有 4 个事件领域（即 $M=4$ ）：事件领域 1、事件领域 2、事件领域 3 和事件领域 4， $N=3$ 。其中，上述文本信息归属于事件领域 1 的领域概率为 50%，上述文本信息归属于事件领域 2 的领域概率为 25%，上述文本信息归属于事件领域 3 的领域概率为 10%，上述文本信息归属于事件领域 4 的领域概率为 15%。由于 $50\% > 25\% > 15\% > 10\%$ ；因此，终端可以从上述事件领域 1-事件领域 4 中选择出领域概率按照由高至低的顺序排列在前 3 位的 3 个事件领域，即事件领域 1、事件领域 2 和事件领域 4。

如此，终端在执行 S403-S405 时，只需要计算上述文本信息归属于上述 N 个事件领域的先验概率和置信度，而不需要计算文本信息归属于 M 个事件领域中所有事件领域的先验概率和置信度，可以减少终端进行语音信息处理时的计算量，提高计算效率。

当然，在本申请实施例中 S402' 是可选的，终端也可以不执行 S402'，在这种情况下，本申请实施例中的 N 个事件领域即上述 M 个事件领域， $N=M$ 。

本申请中，无论 $N=M$ 或者 $N < M$ ，在 S402 或者 S402' 之后，本申请的方法都可以包括 S403-S406：

S403、终端获取上述文本信息归属于所述 N 个事件领域中的每一个事件领域的先验概率，该先验概率用于表征根据已进行的多次语义理解，确定上述文本信息归属于一个事件领域的概率，该 N 个事件领域为 M 个事件领域中的 N 个事件领域，N 小于或等于 M。

其中，一个事件领域的先验概率用于表征根据已进行的多次语义理解，确定该文本信息归属于该事件领域的概率。终端可以根据以往进行多次语义理解的历史数据，分别获取上述文本信息归属于上述 N 个事件领域中的每一个事件领域的先验概率。

示例性的，本申请这里以终端获取上述文本信息归属于上述 N 个事件领域中的第一事件领域的先验概率为例，对终端获取上述文本信息归属于 N 个事件领域中的各个事件领域的先验概率的方法进行举例说明。其中，第一事件领域可以是上述 N 个事件

领域中的任一事件领域。

在一种可能的实现方式中，终端可以统计该终端进行语义理解的总次数 X ；统计 X 次语义理解中、指示终端执行的事件归属于上述第一事件领域的语义理解结果的个数 y ；计算语义理解结果的个数 y 与语义理解的总次数 X 的比值 y/X ，该 y/X 是第一事件领域的先验概率。其中，终端统计的语义理解的总次数 X 是指终端已进行过的所有语义理解的总次数。此处的“所有语义理解”并不限定语义理解的对象，即包括终端对任意文本信息进行的语义理解。

举例来说，假设终端中预先设置了三个事件领域，如音乐领域、设置领域和 APP 领域。该终端进行语义理解的总次数为 P （即 $X=P$ ），这 P 次语义理解中， a 个语义理解结果指示终端执行的事件归属于音乐领域（简称音乐领域的语义理解）， b 次语义理解结果指示终端执行的事件归属于设置领域（简称设置领域的语义理解）， c 次语义理解结果指示终端执行的事件归属于 APP 领域（简称 APP 领域的语义理解）， $a+b+c=P$ 。那么，终端可以确定待处理的文本信息（如文本信息 K ）归属于音乐领域的先验概率为 a/P ，文本信息 K 归属于设置领域的先验概率为 b/P ，文本信息 K 归属于 APP 领域的先验概率为 c/P 。

一般而言，终端进行的相邻两次语义理解的对象（即文本信息）归属于同一事件领域的可能性较高，或者前一次语义理解的对象归属于一个事件领域，可能会影响后一次语义理解的对象归属于哪一事件领域产生影响。例如，在用户使用手机导航的场景中，手机前一次接收用户输入的语音信息可以是“导航”，随后，该手机接收到用户输入的语音信息是“去 x 街道 100 号”或者其他地点信息的可能性较高。其中，“导航”和“去 x 街道 100 号”都可以归属于上述 APP 领域，用于指示手机调用地图 APP 执行相应事件。

基于上述现象，在一种可能的实现方式中，终端获取文本信息 K 归属于上述 N 个事件领域中的每一个事件领域的先验概率时，如果前一次语义理解的对象归属于事件领域 A 。那么，终端则可以确定文本信息 K 归属于事件领域 A 的先验概率为 a ， $a>0.5$ ，文本信息 K 归属于其他任一事件领域的先验概率均为 $(1-a)/(N-1)$ 。

示例性的，假设 $N=3$ ，上述 N 个事件领域包括音乐领域、设置领域和 APP 领域；前一次语义理解的对象归属于音乐领域。那么终端则可以确定文本信息 K 归属于音乐领域的先验概率为 0.8，文本信息 K 归属于设置领域的先验概率为 $(1-0.8)/(3-1)=0.1$ ，文本信息 K 归属于 APP 领域的先验概率也为 0.1。

基于上述现象，在另一种可能的实现方式中，终端在计算文本信息 K 归属于第一事件领域（如事件领域 P ）的先验概率时，可以参考终端前一次进行语义理解的文本信息所归属的事件领域（记为事件领域 Q ），然后统计该终端已进行过的所有语义理解中，相邻两次的被语义理解的文本信息依次归属于事件领域 Q 和事件领域 P 的概率（即按照时间先后顺序，事件领域 Q 在前，事件领域 P 在后），并将该概率确定为该文本信息 K 归属于第一事件领域（如事件领域 P ）的先验概率。

举例来说，假设终端中预先设置了三个事件领域，如音乐领域、设置领域和 APP 领域，该终端共进行过 Y 次语义理解。并且，如图 5 所示，这 Y 次语义理解的语义理解结果指示终端执行的事件所归属的事件领域依次为：设置领域、APP 领域、设置领

域、音乐领域、音乐领域、设置领域、APP 领域、设置领域... 设置领域。

假设本次语义理解（对文本信息 K 进行语义理解）的前一次语义理解是设置领域的语义理解。如图 5 所示，终端统计上述 Y 次语义理解中，设置领域的语义理解的次数 P。终端获取这 P 次设置领域的语义理解中，每次设置领域相邻的后一次语义理解对应的事件领域；统计出前一次语义理解是设置领域、相邻的后一次语义理解是设置领域的次数为 a，前一次语义理解是设置领域、相邻的后一次语义理解是音乐领域的次数为 b，前一次语义理解是设置领域、相邻的后一次语义理解是 APP 领域的次数为 c， $a+b+c=P$ 。那么，终端可以确定文本信息 K 归属于设置领域的先验概率为 a/P ，文本信息 K 归属于音乐领域的先验概率为 b/P ，文本信息 K 归属于 APP 领域的先验概率为 c/P 。其中，本实施例中出现的前一次、后一次，指的是根据时间的先后顺序，先发生的为前一次，后发生的为后一次。

需要说明的是，本申请实施例中，终端获取文本信息归属于上述 N 个事件领域中的每个事件领域的先验概率的方法包括但不限于上述方法，终端获取文本信息归属于每个事件领域的先验概率的其他方法，本申请实施例这里不再赘述。

S404、终端获取上述文本信息归属于上述 N 个事件领域中的每个事件领域的置信度，该置信度用于表征上述文本信息归属于一个事件领域的确信程度。

本申请实施例中，终端可以针对上述 M 个事件领域中的每一个事件领域，保存一个关键字模型，每一个事件领域的关键字模型中包括该事件领域的多个关键字，该多个关键字是该事件领域中常用的词语和短句等。终端可以对上述文本信息进行分词处理，并提取至少一个分词，然后根据该至少一个分词对应的关键字在上述多个事件领域的关键字模型中的分布情况，计算该文本信息归属于该预设多个事件领域中每个事件领域的置信度。具体的，如图 6 所示，图 4 所示的 S404 可以包括 S404a-S404c：

S404a、终端对上述文本信息进行分词处理，并提取至少一个分词。

其中，终端可以通过图 2 所示的中控层 201 中的领域识别模块 2011，调用算法层 203 对文本信息进行分词处理，并提取至少一个分词。例如，假设上述文本信息为“播放歌手 A 的歌曲 B”，终端可以对该文本信息进行分词处理，并提取出如下分词：“播放”、“歌手 A”、“歌曲 B”。假设终端在播放歌手 A 的歌曲的过程中，接收到文本信息“帮我调低音量”，终端可以对该文本信息进行分词处理，并提取出如下分词：

“帮”、“我”、“调低”和“音量”。

S404b、终端获取上述至少一个分词对应的关键字在上述每个事件领域的关键字模型中的分布信息。

例如，本申请实施例的终端中可以维护一个如图 7 所示的关键字数据库 701，该关键字数据库 701 中可以包括多个事件领域的关键字模型。假设终端中预先设置了两个事件领域，如音乐领域和设置领域。如图 7 所示，关键字数据库 701 中包括音乐领域的关键字模型 702 和设置领域的关键字模型 703。其中，音乐领域的关键字模型 702 中包括音乐领域的多个关键字，如播放、下一曲、播放、歌手、摇滚和歌曲等。设置领域的关键字模型 703 中包括设置领域的多个关键字，如飞行模式、蓝牙、亮度、音量和调低等。

S404c、终端根据所述分布信息，计算所述文本信息归属于所述 N 个事件领域中

的每个事件领域的置信度。

其中，文本信息归属于第一事件领域的置信度用于表征该文本信息归属于该第一事件领域的确信程度。

例如，假设文本信息 1 为“播放歌手 A 的歌曲 B”，终端提取的至少一个分词为“播放”、“歌手 A”和“歌曲 B”。终端可以确定出分词“播放”对应的关键字播放、分词“歌手 A”对应的关键字“歌手”和分词“歌曲 B”对应的关键字“歌曲”，均包含在音乐领域的关键字模型 702 中。即终端可以确定文本信息 1 的所有分词对应的关键字都分布在音乐领域的关键字模型 702 中。在这种情况下，终端可以确定文本信息 1 归属于音乐领域的置信度为 90%，文本信息 1 归属于设置领域的置信度为 10%。

再例如，假设终端中预先设置了三个事件领域，如音乐领域、设置领域和 APP 领域。终端对文本信息 2 进行分词处理得到至少一个分词。当该至少一个分词对应的关键字都分布在设置领域的关键字模型中时，终端可以确定文本信息 2 归属于设置领域的置信度为 80%，文本信息 2 归属于音乐领域的置信度为 10%，文本信息 2 归属于 APP 领域的置信度为 10%。

又例如，假设终端中预先设置了三个事件领域，如音乐领域、设置领域和 APP 领域。终端对文本信息 3 进行分词处理得到 8 个分词。当这 8 个分词中的 5 个分词对应的关键字分布在设置领域的关键字模型中，2 个分词对应的关键字分布在音乐领域的关键字模型中，1 个分词对应的关键字分布在 APP 领域的关键字模型中时，终端可以确定文本信息 3 归属于设置领域的置信度为 $5/8=62.5\%$ ，文本信息 3 归属于音乐领域的置信度为 25%，文本信息 3 归属于 APP 领域的置信度为 12.5%。

需要说明的是，当 $M=N$ 时，本申请实施例对终端执行 S402、S403 和 S404 的先后顺序不作限制。例如，终端可以先执行 S403，再执行 S404，最后执行 S402；或者，终端可以先执行 S404，再执行 S402，最后执行 S404；或者，终端可以基本同时执行 S402、S403 和 S404。

当 $N<M$ 时，在 S402 之后，本申请的方法还可以包括 S402'。在这种情况下，终端可以先执行 S402，再执行 S402'，最后执行 S403 和 S404。本申请实施例对终端执行 S403 和 S404 的先后顺序不作限制。例如，终端可以先执行 S403，再执行 S404；或者，终端可以先执行 S404，再执行 S403；或者，终端可以基本同时执行 S403 和 S404。

S405、终端根据上述文本信息归属于上述 N 个事件领域中的每个事件领域的领域概率、先验概率和置信度，计算上述文本信息分别归属于所述 N 个事件领域的 N 个概率值。

其中，终端可以计算文本信息归属于第一事件领域的先验概率、领域概率和置信度的乘积，将计算得到的乘积确定为该文本信息归属于第一事件领域的概率值。

示例性的，如图 8 所示，假设文本信息 a 归属于音乐领域的先验概率为 40%，文本信息 a 归属于设置领域的先验概率为 30%，文本信息 a 归属于 APP 领域的先验概率为 30%；文本信息 a 归属于音乐领域的领域概率为 40%，文本信息 a 归属于设置领域的领域概率为 20%，文本信息 a 归属于 APP 领域的领域概率为 40%；文本信息 a 归属于音乐领域的置信度为 10%，文本信息 a 归属于设置领域的置信度为 10%，文本信息 a 归属于 APP 领域的置信度为 80%。终端可以计算得到文本信息 a 归属于音乐领域的

概率值为 $40\% \times 40\% \times 10\% = 1.6\%$ ，文本信息 a 归属于设置领域的概率值为 $30\% \times 20\% \times 10\% = 0.6\%$ ，文本信息 a 归属于 APP 领域的概率值为 $30\% \times 40\% \times 80\% = 9.6\%$ 。

S406、终端输出根据所述 N 个概率值中概率值最高的事件领域对文本信息进行语义理解的语义理解结果。

5 可选的，所述 S406 可以被替换为：终端将根据所述 N 个概率值中概率值最高的事件领域对文本信息进行语义理解的语义理解结果作为最终的语义理解结果。

对该文本信息对应的每个事件领域均执行上述步骤 402-405 之后，可以获取该文本信息归属于每个事件领域的概率值，即获取到多个概率值。然后终端获取最高的一个概率值对应的事件领域，即将概率值最高的事件领域识别为该文本信息对应的事件领域。其中，终端在对文本信息进行领域识别后，可以将文本信息传输至识别到的事件领域的对话引擎，由对话引擎对该文本信息进行语义理解，得到语义理解结果。或者，本实施例可以不限定终端进行领域识别和语义理解的顺序，可以同时或基本同时进行领域识别和语义理解，也可以先进行语义理解后进行领域识别。

10 如图 8 所示，音乐领域的对话引擎、设置领域的对话引擎和 APP 领域的对话引擎可以分别对文本信息 a 进行语义理解，得到语义理解结果。在进行领域识别之后，即获知到文本信息 a 归属于 APP 领域的概率值 9.6% 大于文本信息 a 归属于音乐领域的概率值 1.6%，并且文本信息 a 归属于 APP 领域的概率值 9.6% 大于文本信息 a 归属于设置领域的概率值 0.6% 之后；终端可以输出 APP 领域的对话引擎对文本信息 a 进行语义理解得到的语义理解结果。

20 例如，终端在执行 S402 之后，S406 之前，可以在 N 个事件领域，分别对上述文本信息进行语义理解，得到 N 个语义理解结果。具体的，在 S402 之后，S406 之前，本申请实施例的方法还可以包括 S406'：

S406'、终端在 N 个事件领域，分别对上述文本信息进行语义理解，得到 N 个语义理解结果。

25 其中，终端在 N 个事件领域中的每个事件领域，分别对上述文本信息进行语义理解，得到 N 个语义理解结果的方法，可以参考本申请上述实施例中的相关描述，本申请实施例这里不再赘述。

进一步的，终端在输出上述语义理解结果后，还可以根据该语义理解结果，执行所述语义理解结果对应的操作。具体的，在上述 S406 之后，本申请实施例的方法还可以包括 S407：

30 S407、终端输出上述语义理解结果之后，终端根据上述语义理解结果，执行上述语义理解结果对应的操作。

需要说明的是，本实施例中终端将根据所述概率值最高的事件领域对文本信息进行语义理解的语义理解结果作为最终识别的语义理解结果。在确定出最终的语义理解结果之后，所述终端可以向所述终端内部输出该最终的结果，使得所述终端执行该最终结果对应的操作。可以理解的，所述的向所述终端内部输出，可以是终端确定概率值最高的最终结果的过程，也可以是终端向内部的其它部件（硬件或软件）发送最终结果，使得该最终结果对应的操作被所述终端执行。可选的，在确定出最终的语义理解结果之后，所述终端也可以向所述终端的外部输出该最终的语义理解结果，例如所

述终端可以向其它终端发送该最终的结果,使得其它终端获知该最终结果,或者使得其它终端来执行该最终结果对应的动作。可选的,所述终端可以既执行该最终结果对应操作,也把该最终结果向外部输出。

5 本申请实施例提供一种语音信息处理方法,可以在将语音信息转换为文本信息后,根据终端进行语义理解的历史数据,获取该文本信息归属于每个事件领域的先验概率;分析该文本信息,获取所述文本信息归属于每个事件领域的领域概率;并且终端可以计算文本信息归属于每个事件领域的置信度;然后,终端根据上述文本信息归属于一个事件领域的先验概率、领域概率和置信度,计算该文本信息归属于该事件领域的概
10 率值;最后,终端可以将概率值最高的事件领域的对话引擎对该文本信息进行语义理解,得到的语义理解结果作为该文本信息(即上述语音信息)的语义理解结果。

其中,文本信息归属于一个事件领域的先验概率:用于表征历史数据中,文本信息归属于该事件领域的概率;文本信息归属于一个事件领域的领域概率:用于表征该文本信息归属于该事件领域的可能性;文本信息归属于一个事件领域的置信度:用于
15 表征该文本信息归属于该事件领域的确信程度。本申请实施例在选择处理文本信息的事件领域时,不仅参考了对文本信息中包括的词汇进行分析得到的领域概率,还参考了文本信息归属于事件领域的先验概率,以及用于表征该文本信息归属于该事件领域的确信程度的置信度;因此,可以提高选择的事件领域的准确性,进而可以提高语义理解结果的准确性,从而可以提高终端执行事件与用户输入的语音信息指示终端执行的
20 事件的符合度,可以提高用户体验。

可选的,在一种可能的实现方式中,上述关键字模型中不仅可以包括多个关键字,还可以包括每个关键字指示文本信息归属于对应事件领域的概率。例如,如图9所示,关键字数据库901中包括音乐领域的关键字模型902和设置领域的关键字模型903。
25 其中,音乐领域的关键字模型902中还可以包括:关键字“下一曲”指示文本信息归属于音乐领域的概率“概率a”、关键字“播放”指示文本信息归属于音乐领域的概率“概率b”、关键字“歌手”指示文本信息归属于音乐领域的概率“概率c”、关键字“播放”指示文本信息归属于音乐领域的概率“概率d”和关键字“歌曲”指示文本信息归属于音乐领域的概率“概率e”等。设置领域的关键字模型903中还可以包
30 括:关键字“飞行模式”指示文本信息归属于设置领域的概率“概率1”、关键字“蓝牙”指示文本信息归属于设置领域的概率“概率2”、关键字“音量”指示文本信息归属于设置领域的概率“概率3”和关键字“调低”指示文本信息归属于设置领域的概率“概率4”等。

可以想到的是,本申请实施例中的关键字数据库,如关键字数据库701和关键字
35 数据库901,可以保存在终端中。或者,为了减少关键字数据库对终端内存的占用,该关键字数据库也可以保存在云服务器中。终端可以从云服务器保存的关键字数据库中查找对应的关键字以及关键字所指示的概率。

其中,终端可以从上述文本信息中识别至少一个关键字;然后,根据至少一个关键字指示的概率,计算文本信息归属于每个事件领域的领域概率。具体的,上述S402

可以包括 S1001-S1003。例如,如图 10 所示,图 4 中的 S402 可以替换为 S1001-S1003:

S1001、终端从所述文本信息中识别至少一个关键字。

其中,终端可以针对每一个事件领域,识别文本信息中是否包括该事件领域的关键字模型中的关键字。例如,假设终端中预先设置了两个事件领域,如音乐领域和设置领域,文本信息 4 为“播放下一曲时,调低歌曲的音量”。终端可以识别到该文本信息 4 中包括关键字“播放”、“下一曲”“调低”、“歌曲”和“音量”。其中,“播放”、“下一曲”和“歌曲”是音乐领域的关键字模型中的关键字,“调低”和“音量”是设置领域的关键字模型中的关键字。

S1002、终端从上述每个事件领域对应的关键字模型中获取上述至少一个关键字分别指示的概率。

示例性的,如图 9 所示,关键字“播放”指示文本信息归属于音乐领域的概率为概率 b,关键字“下一曲”指示文本信息归属于音乐领域的概率为概率 a,关键字“歌曲”指示文本信息归属于音乐领域的概率为概率 e。如图 9 所示,关键字“调低”指示文本信息归属于设置领域的概率为概率 4,关键字“音量”指示文本信息归属于设置领域的概率为概率 3。

S1003、终端根据上述至少一个关键字分别指示的概率,计算上述文本信息归属于上述每个事件领域的领域概率。

例如,上述文本信息 4 归属于音乐领域的领域概率可以为概率 b、概率 a 与概率 e 之和;上述文本信息 4 归属于设置领域的领域概率可以为概率 4 与概率 3 之和。

可选的,本申请实施例中,终端还可以对上述至少一个关键字指示的概率进行归一化,以计算得到文本信息归属于每个事件领域的领域概率。例如,上述文本信息 4 归属于音乐领域的领域概率可以为(概率 b+概率 a+概率 e)/3;上述文本信息 4 归属于设置领域的领域概率可以为(概率 4+概率 3)/2。

本申请实施例中,终端可以从文本信息中识别至少一个关键字,然后根据至少一个关键字指示的概率,计算文本信息归属于每个事件领域的领域概率。其中,由于至少一个关键字中可能包括各个事件领域的关键字模型中的关键字,而每个关键字在不同的事件领域的关键字模型中,可以指示上述文本信息归属于对应事件领域的概率;因此,根据文本信息中包括的各个事件领域的关键字所指示的概率,可以计算得到文本信息归属于各个事件领域的领域概率。

可选的,在另一种可能的实现方式中,终端可以维护一个特征数据库,该特征数据库中包括上述多个事件领域的数据库模型。每个数据库模型中包括多个特征和每个特征的权重及其对应的分词,该特征的权重用于指示对应特征归属于对应事件领域的概率。终端可以针对任一事件领域,执行以下操作以计算文本信息归属于该事件领域的领域概率:对文本信息进行分词处理提取到至少一个分词,然后从该事件领域的数据库模型中查找至少一个分词对应的特征,再根据查找到的特征的权重,计算该文本信息归属于该事件领域的领域概率。具体的,上述 S402 可以包括 S1101-S1103。例如,如图 11 所示,图 4 中的 S402 可以替换为 S1101-S1103:

S1101、终端对文本信息进行分词处理,并提取至少一个分词。

其中,终端对文本信息进行分词处理,并提取至少一个分词的方法可以参考上述

实施例中 S404a 中的详细描述，本申请实施例这里不再赘述。

S1102、终端从上述每个事件领域对应的数据库模型中查找上述至少一个分词对应的特征，该数据库模型中包括多个特征、每个特征的权重及每个特征对应的分词，该权重用于指示该权重对应的特征归属于上述数据库模型中对应的事件领域的概率；其中，每个事件领域对应一个数据库模型。

其中，终端可以统计上述每个事件领域中出现过的多个分词，然后为每个分词分配一个可以位于标识该分词的特征。例如，终端可以为每个分词分配一个唯一标识该分词的数字，该数字可以是十进制的数字，也可以是二进制的数字，或者该数字还可以是其他格式的数字，本申请实施例对数字的格式不做限制。然后，终端可以根据历史语义理解结果中，上述各个分词归属于各个事件领域的概率值，确定出每个分词对应的特征归属于各个事件领域的概率。

例如，如图 12 所示，终端可以维护特征数据库 1201，该特征数据库 1201 中包括事件领域 1 的数据库模型 1202 和事件领域 2 的数据库模型 1203 等。其中，事件领域 1 的数据库模型 1202 中包括：特征 102、特征 102 对应的分词 a 和特征 102 在事件领域 1 的权重 30%；特征 23、特征 23 对应的分词 b 和特征 23 在事件领域 1 的权重 15%；特征 456、特征 456 对应的分词 c 和特征 456 在事件领域 1 的权重 26%；特征 78、特征 78 对应的分词 d 和特征 78 在事件领域 1 的权重 81%。事件领域 2 的数据库模型 1203 中包括：特征 375、特征 375 对应的分词 e 和特征 375 在事件领域 2 的权重 62%；特征 102、特征 102 对应的分词 a 和特征 102 在事件领域 2 的权重 40%；特征 168、特征 268 对应的分词 f 和特征 168 在事件领域 2 的权重 2%；特征 456、特征 456 对应的分词 c 和特征 456 在事件领域 2 的权重 53%。

需要说明的是，在上述特征数据库中，同一分词在不同事件领域的数据库模型中的特征相同，即在特征数据库中，分词的特征可以唯一标识该分词。但是，同一分词在不同事件领域中的权重不同。

例如，如图 12 所示，在事件领域 1 的数据库模型 1202 和事件领域 2 的数据库模型 1203 中，分词 a 的特征均为 102，分词 c 的特征均为 456。而特征 102 在事件领域 1 的权重 20%，在事件领域 2 的权重为 40%；特征 456 在事件领域 1 的权重 26%，在事件领域 2 的权重为 53%。

示例性的，以上述文本信息为“让蓝光对眼睛的辐射少一点”为例，终端对该文本信息进行分词处理，可以得到一下分词：“让”、“蓝光”、“对”、“眼睛”、“的”、“辐射”、“少”、“一点”。假设分词“让”的特征为 5545、分词“蓝光”的特征为 2313、分词“对”的特征为 2212、分词“眼睛”的特征为 9807、分词“的”的特征为 44、分词“辐射”的特征为 3566、分词“少”的特征为 4324、分词“一点”的特征为 333。终端可以确定文本信息“让蓝光对眼睛的辐射少一点”的特征模型为：5545, 2313, 2212, 9807, 44, 3566, 4324, 333。

S1103、终端根据从所述每个事件领域对应的数据库模型中查找到的特征的权重，计算所述文本信息归属于所述每个事件领域的领域概率。

其中，终端可以针对每个事件领域执行以下操作，计算对应事件领域的领域概率：终端从任一事件领域的数据库模型中，查找上述特征模型中的每个特征在该事件领域

中的权重；终端计算查找到的权重之和。其中，终端计算得到的权重之和为文本信息归属于该事件领域的领域概率。

例如，假设终端中设置了两个事件领域（如音乐领域和设置领域）。上述特征模型“5545，2313，2212，9807，44，3566，4324，333”中，特征5545在音乐领域的权重为21%，特征5545在设置领域的权重为79%；特征2313在音乐领域的权重为12%，特征5545在设置领域的权重为88%；特征2212在音乐领域的权重为69%，特征5545在设置领域的权重为31%；特征9807在音乐领域的权重为56%，特征5545在设置领域的权重为44%；特征44在音乐领域的权重为91%，特征5545在设置领域的权重为9%；特征3566在音乐领域的权重为56%，特征5545在设置领域的权重为44%；特征4324在音乐领域的权重为75%，特征5545在设置领域的权重为25%；特征333在音乐领域的权重为12%，特征5545在设置领域的权重为88%。

那么，终端可以计算得到：文本信息归属于音乐领域的领域概率为 $(21\%+12\%+69\%+56\%+23\%+56\%+75\%+12\%)/8=40.5\%$ ，文本信息归属于设置领域的领域概率为 $(79\%+88\%+31\%+44\%+77\%+44\%+25\%+88\%)/8=59.5\%$ 。

可选的，在另一种可能的实现方式中，终端可以维护一个特征数据库，该特征数据库中包括上述多个事件领域的数据库模型和一个特征关系模型。每个数据库模型中包括多个特征和每个特征的权重，该特征的权重用于指示对应特征归属于对应事件领域的概率。该特征关系模型中包括多个特征以及每个特征对应的分词。

例如，如图13所示，终端可以维护一个特征数据库1301，该特征数据库1201中包括事件领域1的数据库模型1302、事件领域2的数据库模型1303和特征关系模型1304。其中，特征关系模型1304中包括：分词a和分词a对应的特征102；分词b和分词b对应的特征23；分词c和分词c对应的特征456；分词d和分词d对应的特征78；分词e和分词e对应的特征375；……；分词f和分词f对应的特征168等。事件领域1的数据库模型1302中包括：特征102和特征102在事件领域1的权重30%；特征23和特征23在事件领域1的权重15%；特征456和特征456在事件领域1的权重26%；特征78和特征78在事件领域1的权重81%。事件领域2的数据库模型1303中包括：特征375和特征375在事件领域2的权重62%；特征102和特征102在事件领域2的权重40%；特征168和特征168在事件领域2的权重2%；特征456和特征456在事件领域2的权重53%。

其中，终端在执行S1102得到至少一个分词后，可以先从图13所示的特征关系模型1304中，查找该至少一个分词对应的特征，确定出文本信息的特征模型；然后，从事件领域1的数据库模型1302中，查找该特征模型中的特征在上述事件领域1的权重，以计算文本信息归属于事件领域1的领域概率；从事件领域2的数据库模型1303中，查找该特征模型中的特征在上述事件领域2的权重，以计算文本信息归属于事件领域2的领域概率。

需要说明的是，在执行S404a-S404c计算文本信息归属于每个事件领域的确信度时，终端执行了“对文本信息进行分词处理，并提取至少一个分词”；而在执行S1101-S1103计算文本信息归属于每个事件领域的领域概率时，终端还执行了“对文本信息进行分词处理，并提取至少一个分词”。为了避免终端重复执行了以下操作“对

文本信息进行分词处理，并提取至少一个分词”，在终端执行 S404a-S404c 计算文本信息归属于该第一事件领域的确信度，执行 S1101-S1103 计算文本信息归属于第一事件领域的领域概率时，终端可以只执行 S404a，不执行 S1101；或者该终端可以只执行 S1101，不执行 S404a。

5 本申请实施例中提供的方法，不仅可以应用于终端与用户进行的单轮语音对话过程中，还可以应用于终端与用户进行的多轮语音对话的过程中。其中，本申请实施例中所述的单轮语音对话是指用户与终端采用一问一答的模式进行语音对话。但是，在一些场景中，当用户向终端输入一个语音信息（如语音信息 a）后，在该终端还响应该语音信息 a 回复用户时，该用户再次输入了另一语音信息（如语音信息 b）。此时，
10 由于终端几乎同时接收到语音信息 a 和语音信息 b，因此该终端要同时处理该语音信息 a 和语音信息 b。例如，上述语音信息 a 对应的文本信息 a 为“我要去西藏”，上述语音信息 b 对应的文本信息 b 为“今天天气怎么样”。

其中，本申请实施例将终端接收并处理上述语音信息 a 和语音信息 b 的对话过程中，称为多轮语音对话。本申请实施例中，终端可以将上述语音信息 a 和语音信息 b
15 转换为文本信息，并对文本信息 a 和文本信息 b 进行语义理解。

示例性的，由于终端在接收语音信息 b 时，可能还未对文本信息 a 进行语义理解，由此，终端可以认为语音信息 b 的前一个语音信息和语音信息 a 的前一个语言信息是相同的。由于先验概率取决于前一个语音信息对应的事件领域，因此文本信息 a 和文本信息 b 归属于某一事件领域的先验概率相同。例如，如图 14 所示，文本信息 a 和文本信息 b 归属于音乐领域的先验概率 P1 均为 40%，文本信息 a 和文本信息 b 归属于设置领域的先验概率 P1 均为 30%，文本信息 a 和文本信息 b 归属于 APP 领域的先验概率 P1 均为 30%。对于领域概率和置信度，可以分别对文本信息 a 和文本信息 b 进行计算。文本信息 a 归属于音乐领域的领域概率 P2-a 为 40%，文本信息 a 归属于设置领域的领域概率 P2-a 为 20%，文本信息 a 归属于 APP 领域的领域概率 P2-a 为 40%；文本
20 信息 b 归属于音乐领域的领域概率 P2-b 为 20%，文本信息 b 归属于设置领域的领域概率 P2-b 为 10%，文本信息 b 归属于 APP 领域的领域概率 P2-b 为 70%。文本信息 a 归属于音乐领域的置信度 P3-a 为 10%，文本信息 a 归属于设置领域的置信度 P3-a 为 10%，文本信息 a 归属于 APP 领域的置信度 P3-a 为 80%；文本信息 b 归属于音乐领域的置信度 P3-b 为 60%，文本信息 b 归属于设置领域的置信度 P3-b 为 30%，文本信息 b 归
25 属于 APP 领域的置信度 P3-b 为 10%。

其中，如图 14 所示，终端可以同时或基本同时计算得到：文本信息 a 归属于音乐领域的概率值 P-a 为 $40\% \times 40\% \times 10\% = 1.6\%$ ，文本信息 a 归属于设置领域的概率值 P-a 为 $30\% \times 20\% \times 10\% = 0.6\%$ ，文本信息 a 归属于 APP 领域的概率值 P-a 为 $30\% \times 40\% \times 80\% = 9.6\%$ ；文本信息 b 归属于音乐领域的概率值 P-b 为 $40\% \times 20\% \times 60\% = 4.8\%$ ，文本信息 b 归属于设置领域的概率值 P-b 为 $30\% \times 10\% \times 30\% = 0.9\%$ ，文本信息 b 归属于 APP 领域的概率值 P-b 为 $30\% \times 70\% \times 10\% = 2.1\%$ 。
35

如图 14 所示，音乐领域的对话引擎、设置领域的对话引擎和 APP 领域的对话引擎可以分别对文本信息 a 和文本信息 b 进行语义理解，得到语义理解结果。由于文本

信息 a 归属于 APP 领域的概率值 9.6% 大于归属于音乐领域的概率值 1.6%，也大于归属于设置领域的概率值 0.6%；因此，终端可以输出 APP 领域的对话引擎对文本信息 a 进行语义理解得到的语义理解结果。由于文本信息 b 归属于音乐领域的概率值 4.8% 大于归属于 APP 领域的概率值 2.1%，也大于归属于设置领域的概率值 0.9%；因此，终端可以输出音乐领域的对话引擎对文本信息 b 进行语义理解得到的语义理解结果。

本申请实施例提供一种语音信息处理方法，不仅可以应用于终端与用户进行的单轮语音对话过程中，还可以应用于终端与用户进行的多轮语音对话的过程中。本申请实施例提供的方法，无论应用于终端与用户进行的单轮语音对话过程中，还是应用于终端与用户进行的多轮语音对话的过程中，都可以提高选择的事件领域的准确性，进而可以提高语义理解结果的准确性，从而可以提高终端执行事件与用户输入的语音信息指示终端执行的事件的符合度，可以提高用户体验。

可以理解的是，上述终端等为了实现上述功能，其包含了执行各个功能相应的硬件结构和/或软件模块。本领域技术人员应该很容易意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，本发明实施例能够以硬件或硬件和计算机软件的结合形式来实现。某个功能究竟以硬件还是计算机软件驱动硬件的方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明实施例的范围。

本申请实施例可以根据上述方法示例对上述终端等进行功能模块的划分，例如，可以对应各个功能划分各个功能模块，也可以将两个或两个以上的功能集成在一个处理模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。需要说明的是，本发明实施例中对模块的划分是示意性的，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

在采用对应各个功能划分各个功能模块的情况下，图 15 示出了上述实施例中所涉及的终端中的语音信息处理装置的一种可能的结构示意图，该语音信息处理装置 1500 包括：接收单元 1501、转换单元 1502、第一获取单元 1503、第二获取单元 1504、第三获取单元 1505、计算单元 1506 和输出单元 1507。

其中，上述接收单元 1501 用于支持终端执行方法实施例中的 S401 中“接收语音信息”的操作，和/或用于本文所描述的技术的其它过程。

上述转换单元 1502 用于支持终端执行方法实施例中的 S401 中“将语音信息转换为文本信息”的操作，和/或用于本文所描述的技术的其它过程。

上述第一获取单元 1503 用于支持终端执行方法实施例中的 S402、S1001-S1003、S1101-S1103，和/或用于本文所描述的技术的其它过程。

上述第二获取单元 1504 用于支持终端执行方法实施例中的 S403，和/或用于本文所描述的技术的其它过程。

上述第三获取单元 1505 用于支持终端执行方法实施例中的 S404、S404a-S404c，和/或用于本文所描述的技术的其它过程。

上述计算单元 1506 用于支持终端执行方法实施例中的 S405，和/或用于本文所描述的技术的其它过程。

上述输出单元 1507 用于支持终端执行方法实施例中的 S406，和/或用于本文所描

述的技术的其它过程。

进一步的，上述语音信息处理装置 1500 还可以包括：语义理解单元。该语义理解单元用于支持终端执行方法实施例中的 S406'，和/或用于本文所描述的技术的其它过程。

5 进一步的，上述语音信息处理装置 1500 还可以包括：存储单元。该存储单元用于保存方法实施例中所描述的关键字模型和数据库模型等信息。

进一步的，上述语音信息处理装置 1500 还可以包括：执行单元。该执行单元用于支持终端执行方法实施例中的 S407，和/或用于本文所描述的技术的其它过程。

10 其中，上述方法实施例涉及的各步骤的所有相关内容均可以援引到对应功能模块的功能描述，在此不再赘述。

当然，语音信息处理装置 1500 包括但不限于上述所列举的单元模块，例如，如图 16 所示，语音信息处理装置 1500 还可以包括第四获取单元 1508，该第四获取单元用于支持终端执行方法实施例中的 S402'，和/或用于本文所描述的技术的其它过程。

15 并且，上述功能单元的具体所能够实现的功能也包括但不限于上述实例所述的方法步骤对应的功能，语音信息处理装置 1500 的其他单元的详细描述可以参考其所对应方法步骤的详细描述，本申请实施例这里不再赘述。

需要说明的是，上述语义理解单元可以对应于图 2 所示的对话引擎层 202 中的一个或多个对话引擎。上述转换单元 1502 可以对应于图 2 所示的领域识别模块 2012。上述第一获取单元 1503、第二获取单元 1504、第三获取单元 1505、计算单元 1506 的功能可以集成在图 2 所示的 DS 模块 2014 中实现。可以理解，上述语义理解单元、转换单元 1502、第一获取单元 1503、第二获取单元 1504、第三获取单元 1505、计算单元 1506 等的功能都可以集成在一个处理模块中实现，该处理模块可以是处理器或控制器，例如可以是中央处理器（Central Processing Unit, CPU），通用处理器，数字信号处理器（Digital Signal Processor, DSP），专用集成电路（Application-Specific Integrated
20 Circuit, ASIC），现场可编程门阵列（Field Programmable Gate Array, FPGA）或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。其可以实现或执行结合本申请公开内容所描述的各种示例性的逻辑方框，模块和电路。所述处理器也可以是实现计算功能的组合，例如包含一个或多个微处理器组合，DSP 和微处理器的组合等等。上述接收单元 1501 和输出单元 1507 可以对应于图 2 所示的 VSI 2011。该
25 VSI 2011 可以是终端的处理器器的一个接口。上述存储单元可以是用于保存图 2 所示的算法层 203 中的算法和规则等的存储模块。该存储模块可以是存储器。
30

在采用集成的单元的情况下，图 17 示出了上述实施例中所涉及的终端的一种可能的结构示意图。该终端 1700 包括：处理模块 1701 和存储模块 1702。存储模块 1702 用于保存终端的程序代码和数据（如算法和规则等）。处理模块 1701 用于执行存储模块 1702 保存的程序代码执行方法实施例所述的语音信息处理方法。进一步的，该终端 1700 还可以包括通信模块 1703，该通信模块 1703 用于支持终端与其他网络实体的通信。通信模块 1703 可以是收发器、收发电路或通信接口等。存储模块 1702 可以是存储器。
35

当处理模块 1701 为处理器（如图 1 所示的处理器 101），通信模块 1703 为 RF 收

发电路（如图 1 所示的射频电路 102），存储模块 1702 为存储器（如图 1 所示的存储器 103）时，本发明实施例所提供的终端可以为图 1 所示的终端 100。其中，上述通信模块 1703 不仅可以包括射频电路，还可以包括 WiFi 模块和蓝牙模块。射频电路、WiFi 模块和蓝牙模块等通信模块可以统称为通信接口。本申请实施例的终端中可以包括一个或多个处理器和一个或多个存储器，上述一个或多个处理器、一个或多个存储器和通信接口可以通过总线耦合在一起。

本申请实施例该提供一种电子设备，该电子设备包括上述实施例所述的用于执行上述实施例中的语音信息处理方法的语音信息处理装置 1500。

本申请实施例还提供一种计算机存储介质，该计算机存储介质中存储有计算机程序代码，当上述处理器执行该计算机程序代码时，该电子设备执行图 3、图 4、图 6、图 10 和图 11 中任一附图中的相关方法步骤实现上述实施例中的语音信息处理方法。

本申请实施例还提供了一种计算机程序产品，当该计算机程序产品在电子设备上运行时，使得电子设备执行图 3、图 4、图 6、图 10 和图 11 中任一附图中的相关方法步骤实现上述实施例中的语音信息处理方法。

其中，本发明实施例提供的语音信息处理装置 1500、终端 1700、计算机存储介质或者计算机程序产品均用于执行上文所提供的对应的方法，因此，其所能达到的有益效果可参考上文所提供的对应的方法中的有益效果，此处不再赘述。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将装置的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请实施例所提供的几个实施例中，应该理解到，所揭露的系统，装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块或单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请实施例各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请实施例的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件

产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）或处理器执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：快闪存储器、移动硬盘、只读存储器、随机存取存储器、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述，仅为本申请实施例的具体实施方式，但本申请实施例的保护范围并不局限于此，任何在本申请实施例揭露的技术范围内的变化或替换，都应涵盖在本申请实施例的保护范围之内。因此，本申请实施例的保护范围应以所述权利要求的保护范围为准。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将装置的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统，装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块或单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）或处理器执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：快闪存储器、移动硬盘、只读存储器、随机存取存储器、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何在本申请揭露的技术范围内的变化或替换，都应涵盖在本申请的保护范围之内。

内。因此，本申请的保护范围应以所述权利要求的保护范围为准。

权 利 要 求 书

1、一种语音信息处理方法，其特征在于，所述方法包括：

终端接收语音信息，将所述语音信息转换为文本信息；所述终端中预设 M 个事件领域；

5 所述终端获取所述文本信息归属于所述 M 个事件领域中的每个事件领域的领域概率，所述领域概率用于表征所述文本信息归属于一个事件领域的可能性；

所述终端获取所述文本信息归属于 N 个事件领域中的每一个事件领域的先验概率，所述先验概率用于表征根据已进行的多次语义理解，确定所述文本信息归属于一个事件领域的概率，所述 N 个事件领域为所述 M 个事件领域中的 N 个事件领域，N 小于
10 或等于 M；

所述终端获取所述文本信息归属于所述 N 个事件领域中的每个事件领域的置信度，所述置信度用于表征所述文本信息归属于一个事件领域的确信程度；

所述终端根据所述文本信息归属于所述 N 个事件领域中的每个事件领域的领域概率、先验概率和置信度，计算所述文本信息分别归属于所述 N 个事件领域的 N 个概率
15 值；

所述终端输出根据所述 N 个概率值中概率值最高的事件领域对文本信息进行语义理解的语义理解结果。

2、根据权利要求 1 所述的方法，其特征在于，当 N 小于 M 时，所述 N 个事件领域是所述预设 M 个事件领域中，领域概率按照由高至低的顺序排列在前 N 位的 N 个
20 事件领域， $N \geq 2$ 。

3、根据权利要求 1 或 2 所述的方法，其特征在于，在所述终端获取所述文本信息归属于 M 个事件领域中的每个事件领域的领域概率之后，所述方法还包括：

所述终端在所述 N 个事件领域，分别对所述文本信息进行语义理解，得到 N 个语义理解结果。

4、根据权利要求 1-3 中任一项所述的方法，其特征在于，所述 M 个事件领域中的每个事件领域对应于一个关键字模型，所述关键字模型中包括：对应事件领域的多个关键字；
25

所述终端获取所述文本信息归属于所述 N 个事件领域中的每个事件领域的置信度，包括：

30 所述终端对所述文本信息进行分词处理，并提取至少一个分词；

所述终端获取所述至少一个分词对应的关键字在所述每个事件领域的关键字模型中的分布信息；

所述终端根据所述分布信息，计算所述文本信息归属于所述 N 个事件领域中的每个事件领域的置信度。

5、根据权利要求 1-4 中任一项所述的方法，其特征在于，所述终端获取所述文本信息归属于 M 个事件领域中的每个事件领域的领域概率，包括：
35

所述终端对所述文本信息进行分词处理，并提取至少一个分词；

所述终端从所述每个事件领域对应的数据库模型中查找所述至少一个分词对应的特征，所述数据库模型中包括多个特征、每个特征的权重及每个特征对应的分词，所

述权重用于指示所述权重对应的特征归属于所述数据库模型中对应的事件领域的概率；其中，每个事件领域对应一个数据库模型；

所述终端根据从所述每个事件领域对应的数据库模型中查找到的特征的权重，计算所述文本信息归属于所述每个事件领域的领域概率。

- 5 6、根据权利要求 1-4 中任一项所述的方法，其特征在于，所述 M 个事件领域中的每个事件领域对应于一个关键字模型，所述关键字模型中包括：多个关键字和每个关键字指示文本信息归属于所述关键字模型对应的事件领域的概率；

所述终端获取所述文本信息归属于 M 个事件领域中的每个事件领域的领域概率，包括：

- 10 所述终端从所述文本信息中识别至少一个关键字；

所述终端从所述每个事件领域对应的关键字模型中获取所述至少一个关键字分别指示的概率；

所述终端根据所述至少一个关键字分别指示的概率，计算所述文本信息归属于所述每个事件领域的领域概率。

- 15 7、根据权利要求 1-6 中任一项所述的方法，其特征在于，还包括：

所述终端输出所述语义理解结果之后，所述终端根据所述语义理解结果，执行所述语义理解结果对应的操作。

8、一种语音信息处理装置，其特征在于，所述装置包括：

接收单元，用于接收语音信息；

- 20 转换单元，用于将所述接收单元接收的所述语音信息转换为文本信息；所述终端中预设 M 个事件领域；

第一获取单元，用于获取所述转换单元转换得到的所述文本信息归属于所述 M 个事件领域中的每个事件领域的领域概率，所述领域概率用于表征所述文本信息归属于一个事件领域的可能性；

- 25 第二获取单元，用于获取所述转换单元转换得到的所述文本信息归属于 N 个事件领域中的每一个事件领域的先验概率，所述先验概率用于表征根据已进行的多次语义理解，确定所述文本信息归属于一个事件领域的概率，所述 N 个事件领域为所述 M 个事件领域中的 N 个事件领域，N 小于或等于 M；

- 30 第三获取单元，用于获取所述转换单元转换得到的所述文本信息归属于所述 N 个事件领域中的每个事件领域的置信度，所述置信度用于表征所述文本信息归属于一个事件领域的确信程度；

计算单元，用于根据所述第一获取单元获取的所述文本信息归属于所述 N 个事件领域中的每个事件领域的领域概率、所述第二获取单元获取的先验概率和所述第三获取单元获取的置信度，计算所述文本信息分别归属于所述 N 个事件领域的 N 个概率值；

- 35 输出单元，用于输出根据所述计算单元计算得到的所述 N 个概率值中概率值最高的事件领域对文本信息进行语义理解的语义理解结果。

9、根据权利要求 8 所述的装置，其特征在于，当 N 小于 M 时，所述 N 个事件领域是所述预设 M 个事件领域中，领域概率按照由高至低的顺序排列在前 N 位的 N 个事件领域， $N \geq 2$ 。

10、根据权利要求 8 或 9 所述的装置，其特征在于，所述装置还包括：

语义理解单元，用于在所述第一获取单元获取所述文本信息归属于 M 个事件领域中的每个事件领域的领域概率之后，在所述 N 个事件领域，分别对所述文本信息进行语义理解，得到 N 个语义理解结果。

5 11、根据权利要求 8-10 中任一项所述的装置，其特征在于，所述装置还包括：

存储单元，用于保存所述 M 个事件领域中的每个事件领域对应的关键字模型，所述关键字模型中包括：对应事件领域的多个关键字；

所述第三获取单元，具体用于：

对所述文本信息进行分词处理，并提取至少一个分词；

10 获取所述至少一个分词对应的关键字在所述存储单元保存的所述每个事件领域的关键字模型中的分布信息；

根据所述分布信息，计算所述文本信息归属于所述 N 个事件领域中的每个事件领域的置信度。

15 12、根据权利要求 8-11 中任一项所述的装置，其特征在于，所述第一获取单元，具体用于：

对所述文本信息进行分词处理，并提取至少一个分词；

从所述每个事件领域对应的数据库模型中查找所述至少一个分词对应的特征，所述数据库模型中包括多个特征、每个特征的权重及每个特征对应的分词，所述权重用于指示所述权重对应的特征归属于所述数据库模型中对应的事件领域的概率；其中，
20 每个事件领域对应一个数据库模型；

根据从所述每个事件领域对应的数据库模型中查找到的特征的权重，计算所述文本信息归属于所述每个事件领域的领域概率。

13、根据权利要求 8-11 中任一项所述的装置，其特征在于，所述装置还包括：

25 存储单元，用于保存所述 M 个事件领域中的每个事件领域对应的关键字模型，所述关键字模型中包括：多个关键字和每个关键字指示文本信息归属于所述关键字模型对应的事件领域的概率；

所述第一获取单元，具体用于：

从所述文本信息中识别至少一个关键字；

30 从所述每个事件领域对应的关键字模型中获取所述至少一个关键字分别指示的概率；

根据所述至少一个关键字分别指示的概率，计算所述文本信息归属于所述每个事件领域的领域概率。

14、根据权利要求 8-13 中任一项所述的装置，其特征在于，所述装置还包括：

35 执行单元，用于在所述输出单元输出所述语义理解结果之后，根据所述语义理解结果，执行所述语义理解结果对应的操作。

15、一种终端，其特征在于，包括：一个或多个处理器和一个或多个存储器；所述一个或多个存储器中存储有一个或多个计算机程序，所述一个或多个计算机程序包括指令，当所述指令被所述一个或多个处理器执行时，使得所述终端执行如权利要求 1-7 任一所述的语音信息处理方法。

16、一种电子设备，其特征在于，所述电子设备包括执行如权利要求 1-7 中任一项所述的语音信息处理方法的装置。

17、一种包含指令的计算机程序产品，其特征在于，当所述计算机程序产品在电子设备上运行时，使得所述电子设备执行如权利要求 1-7 中任一项所述的语音信息处理
5 方法。

18、一种计算机可读存储介质，包括指令，其特征在于，当所述指令在电子设备上运行时，使得所述电子设备执行如权利要求 1-7 中任一项所述的语音信息处理方法。

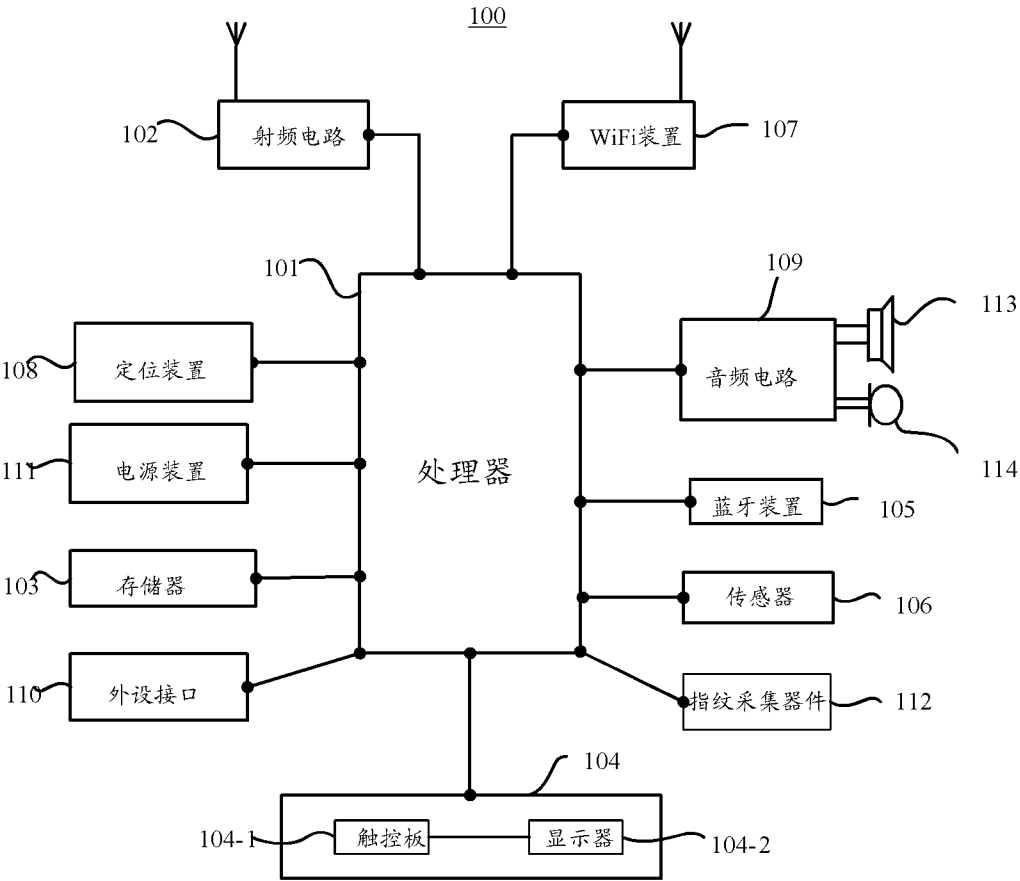


图 1

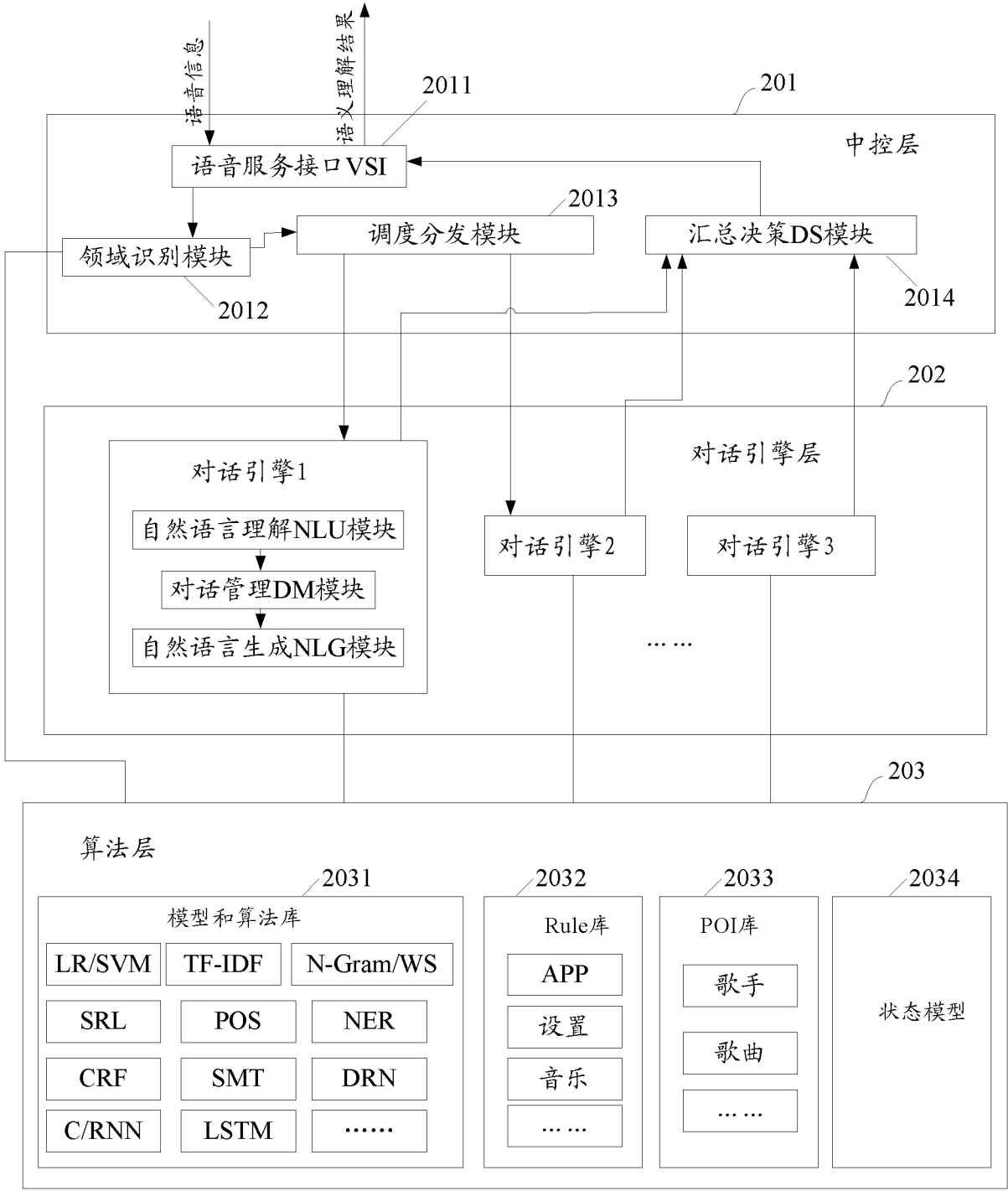


图 2

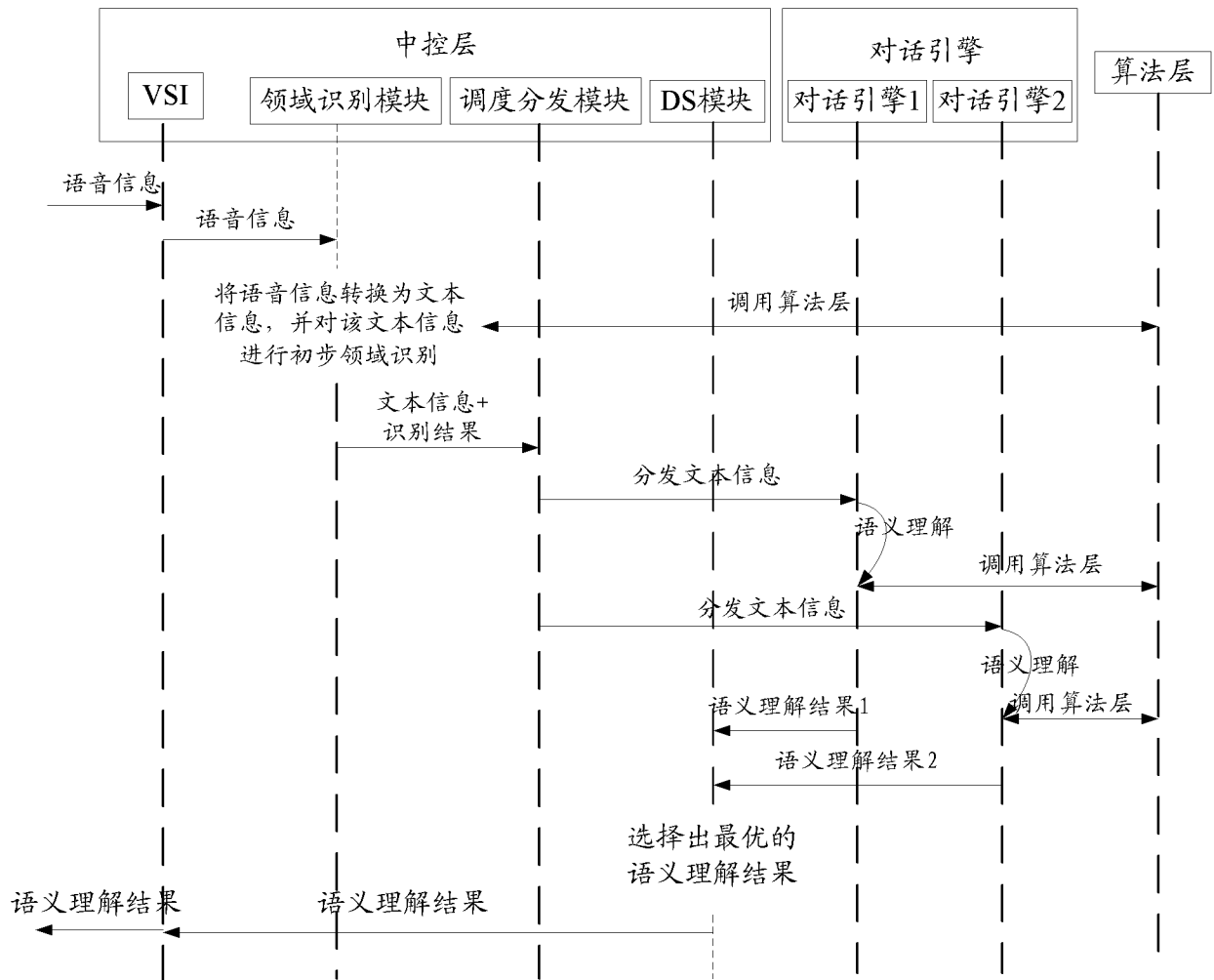


图 3

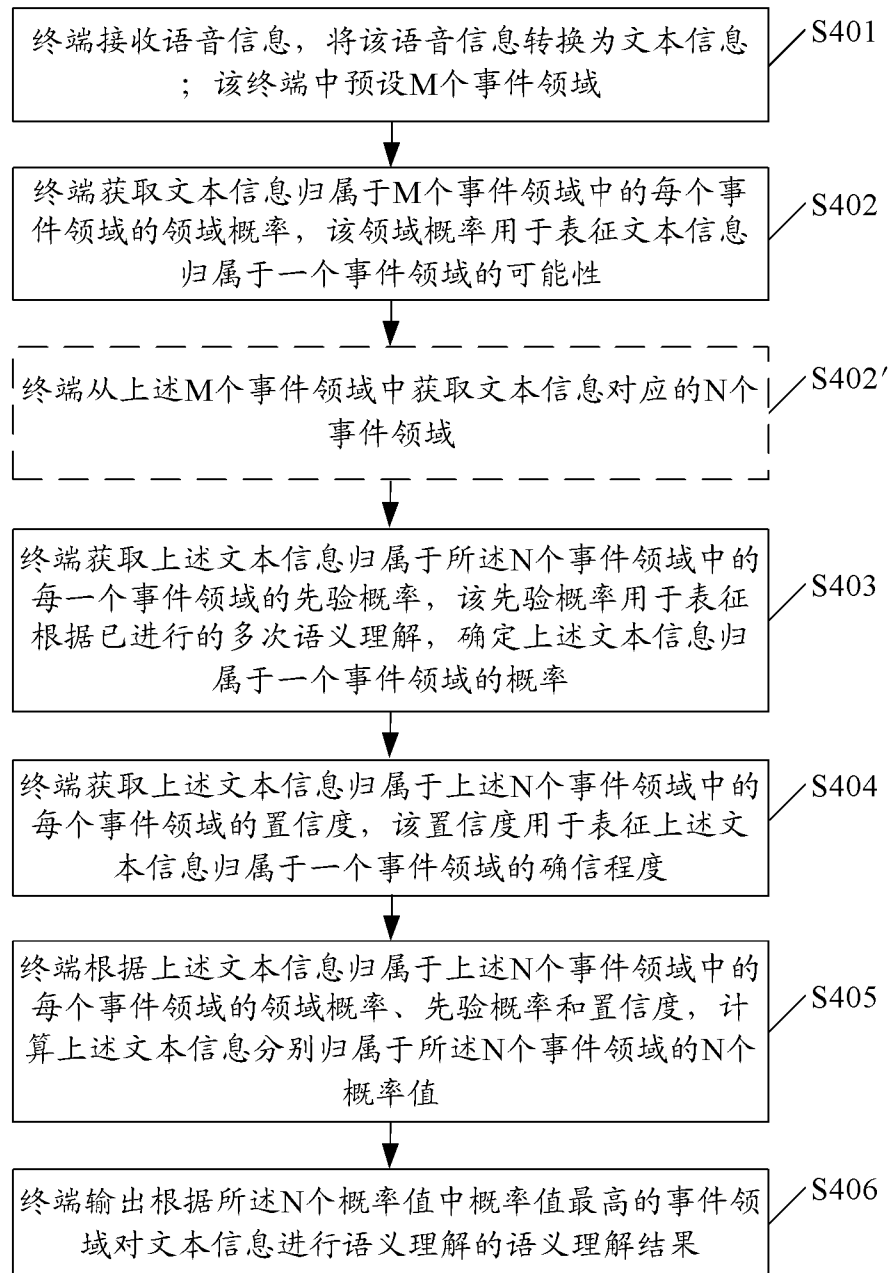


图 4

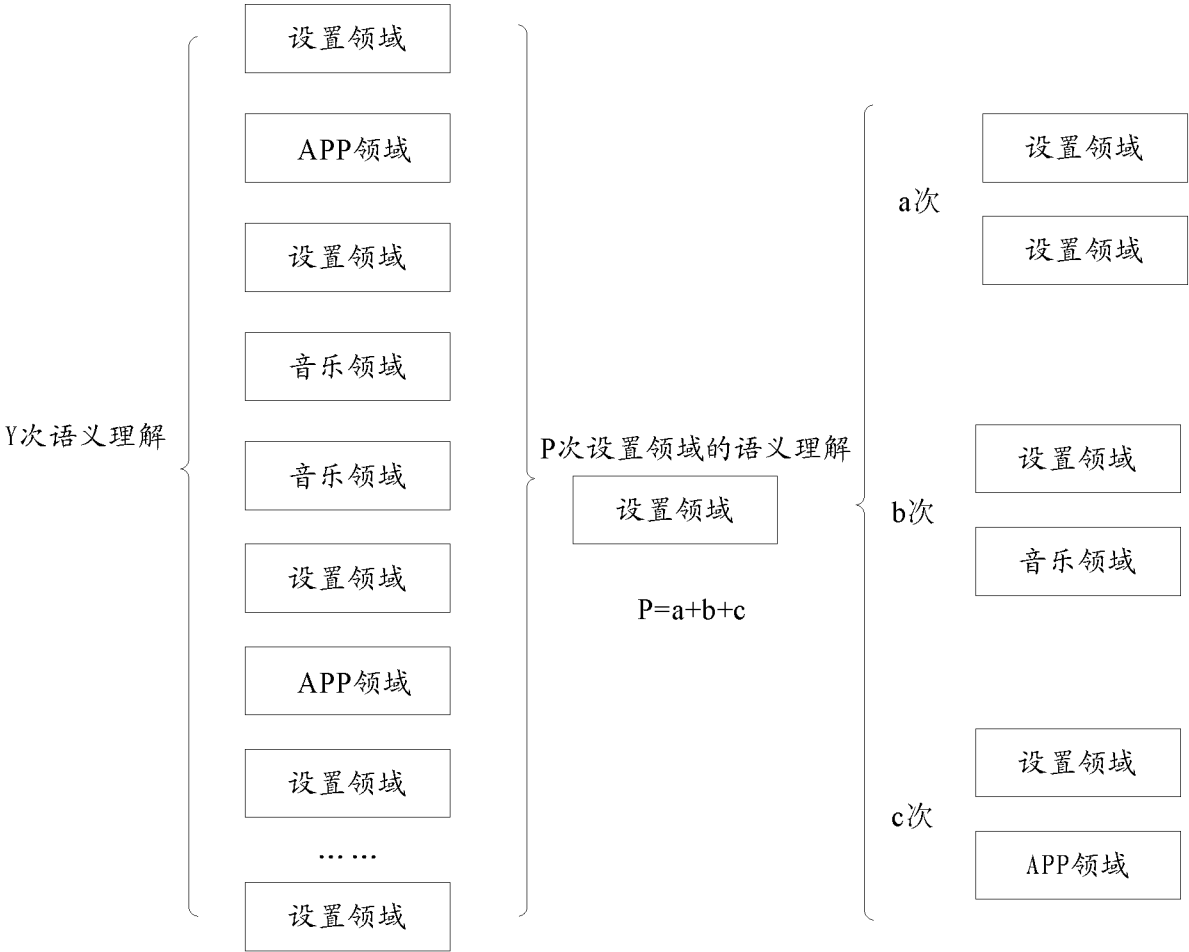


图 5

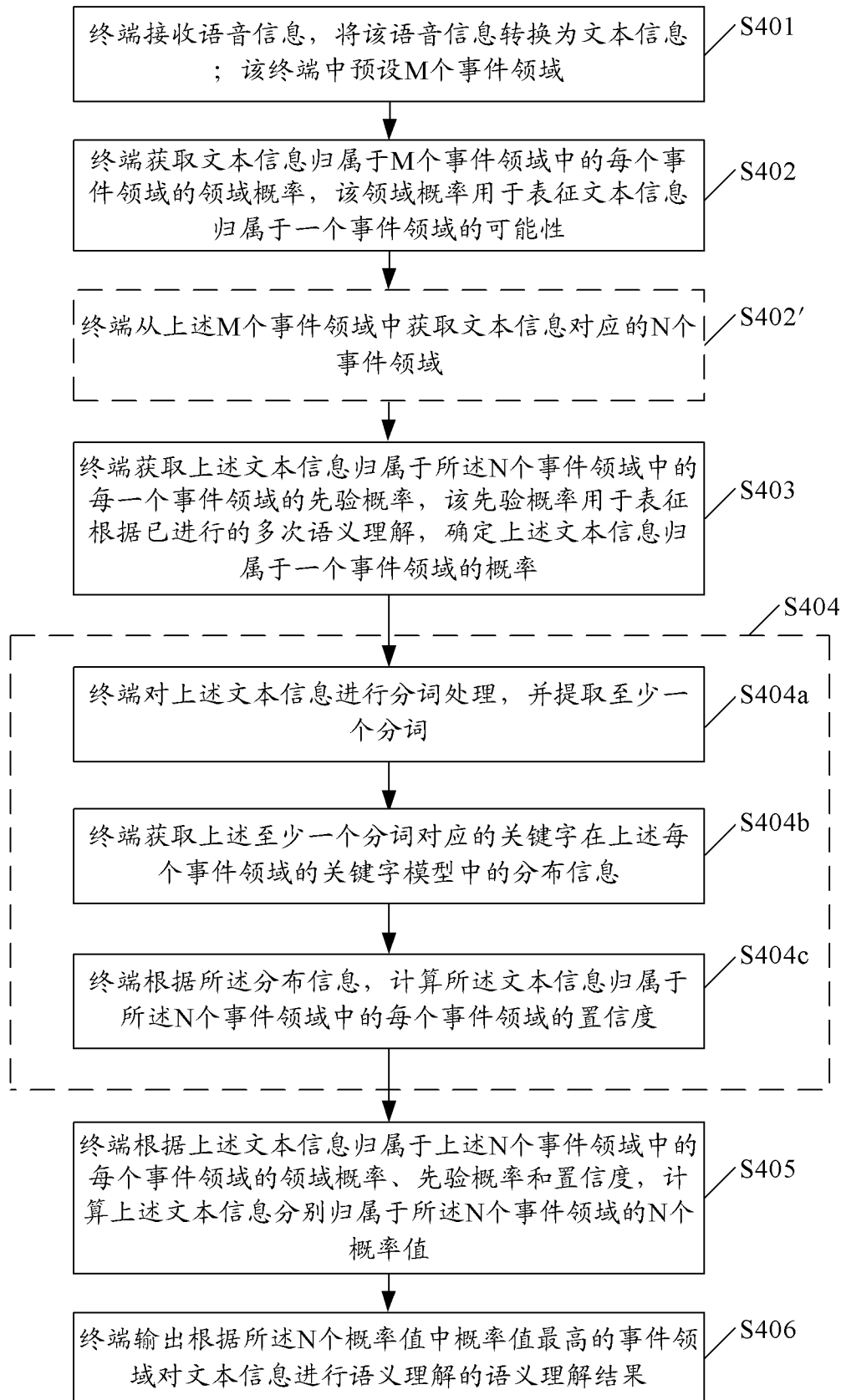


图 6

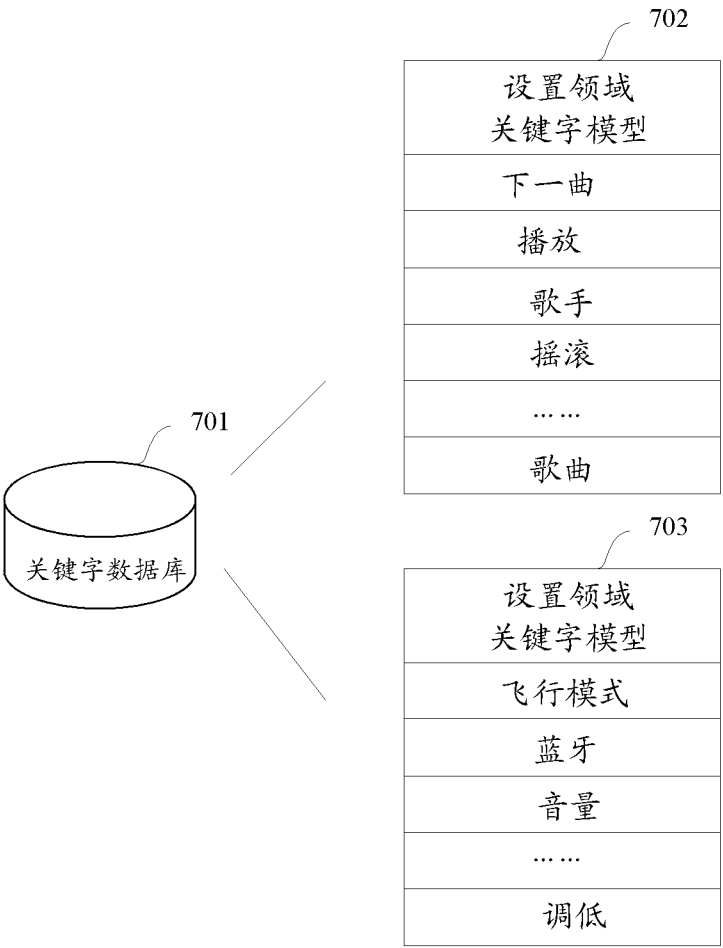


图 7

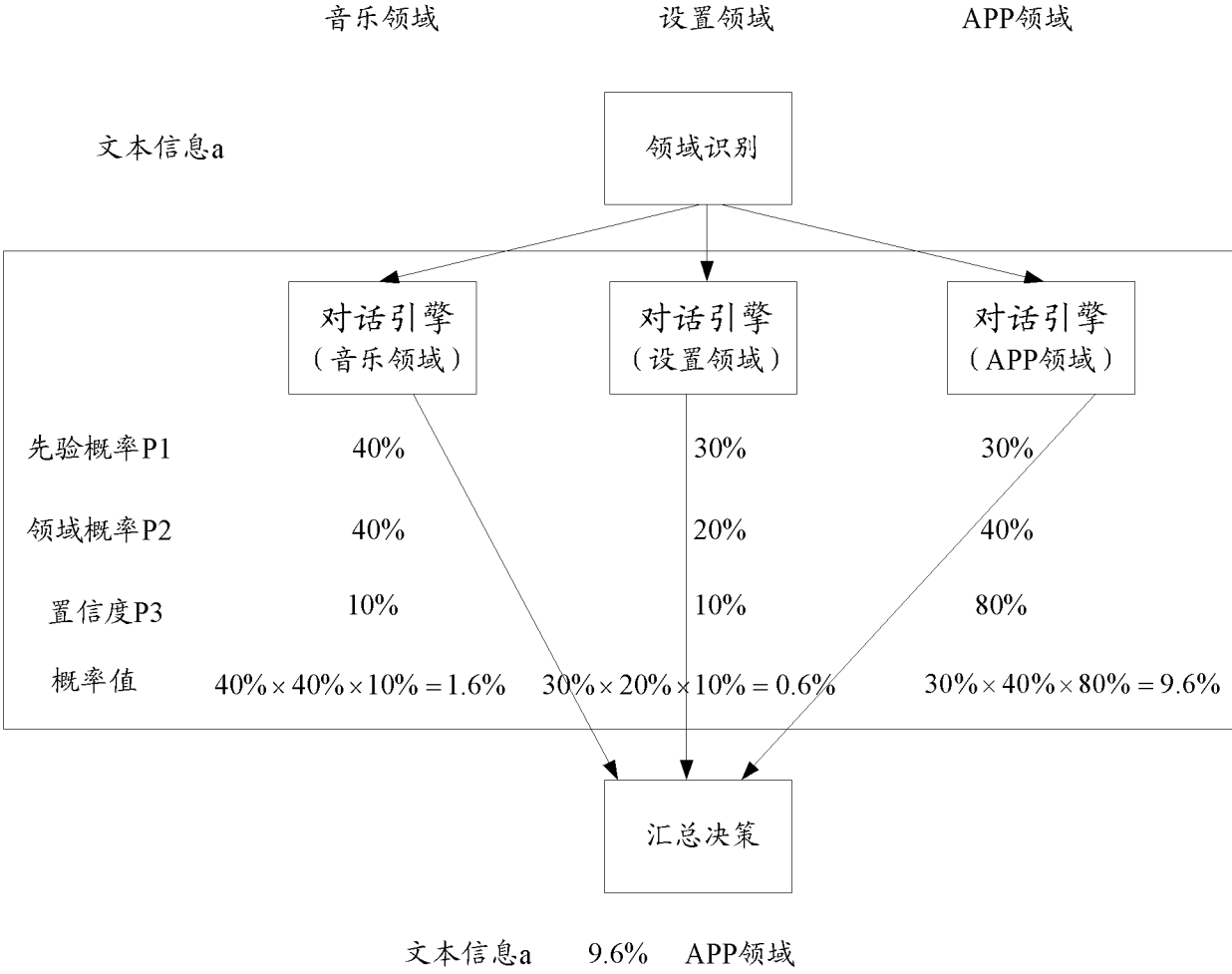


图 8

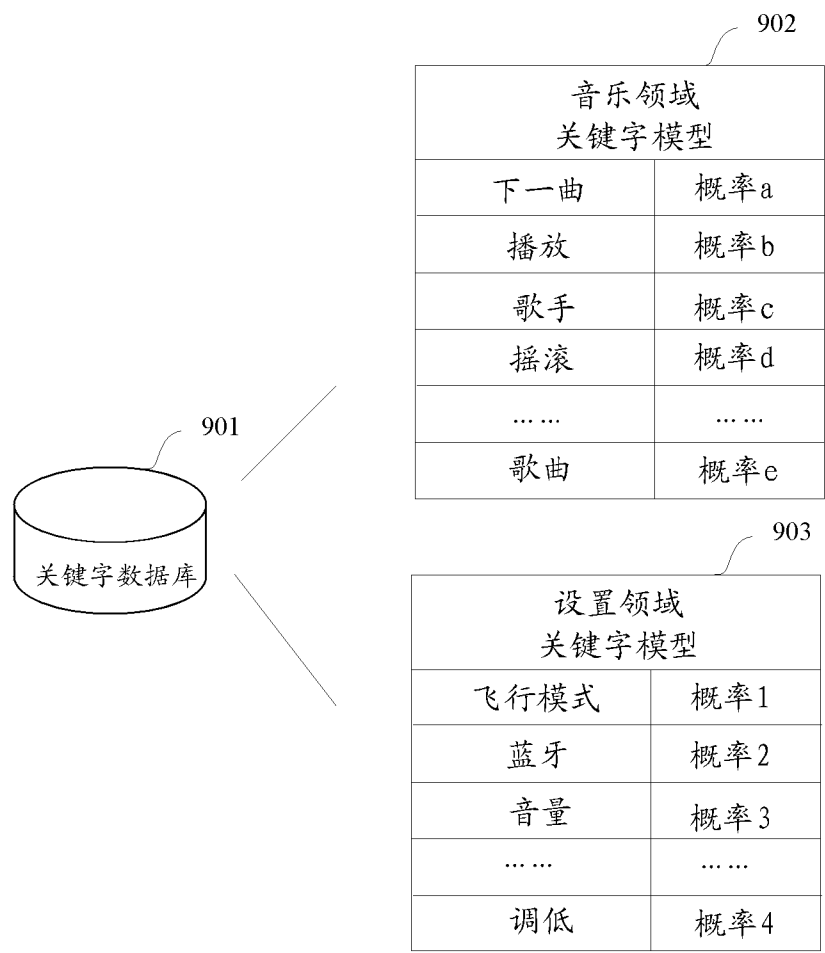


图 9

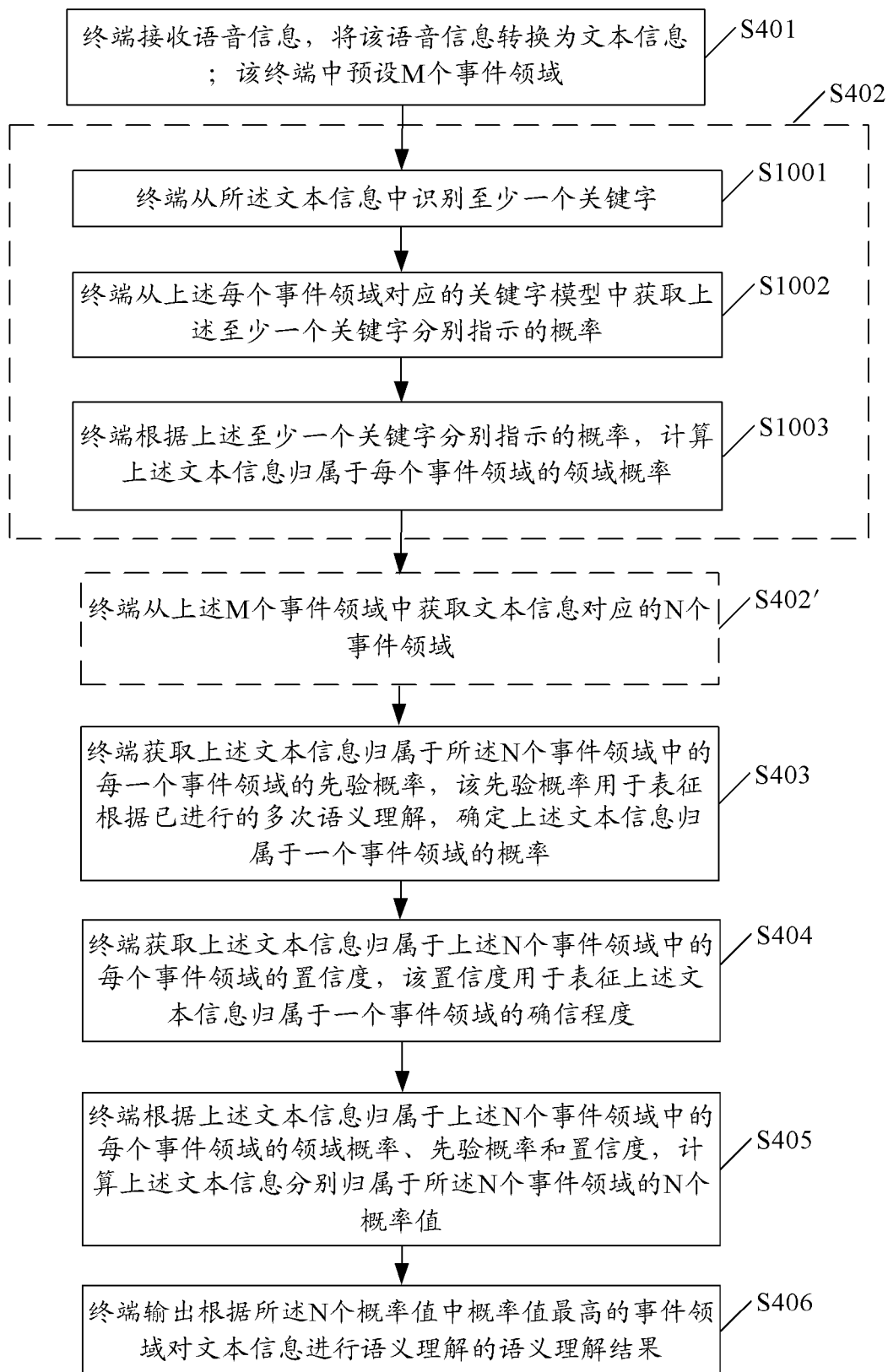


图 10

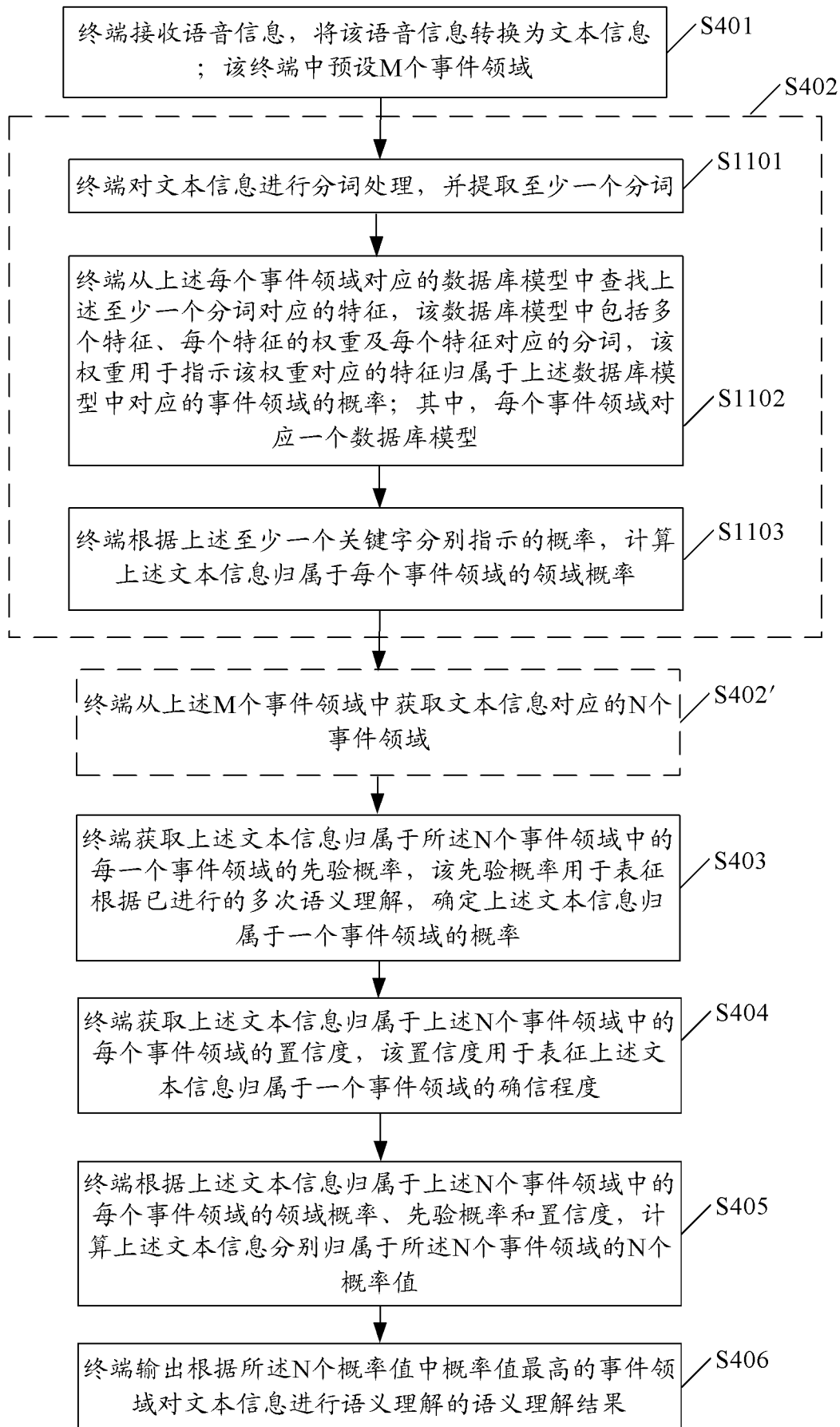


图 11

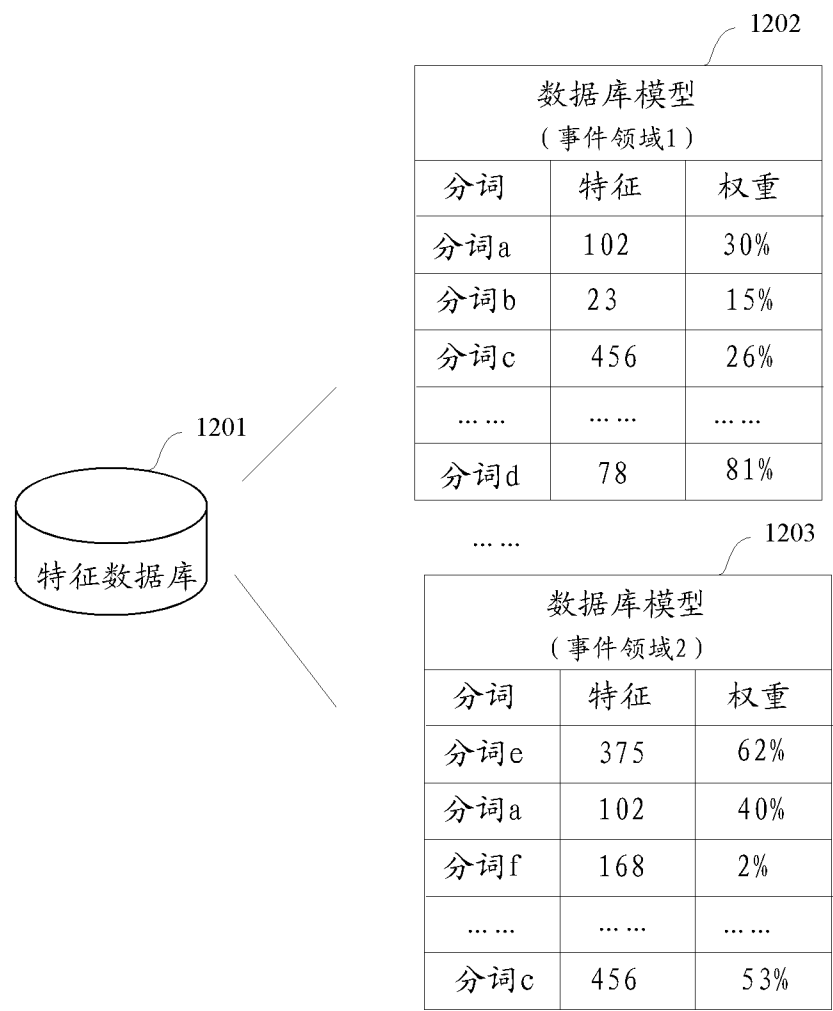


图 12

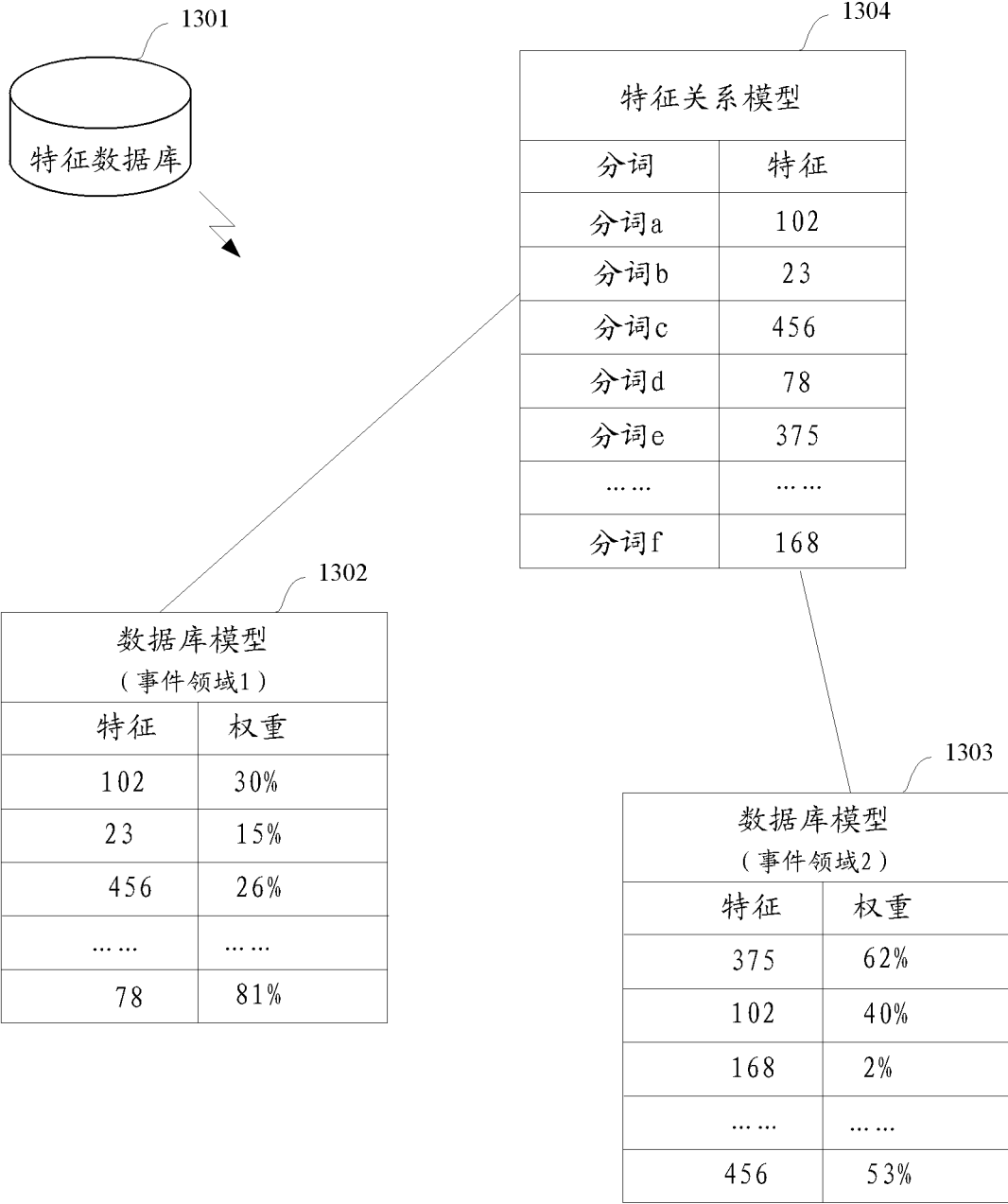


图 13

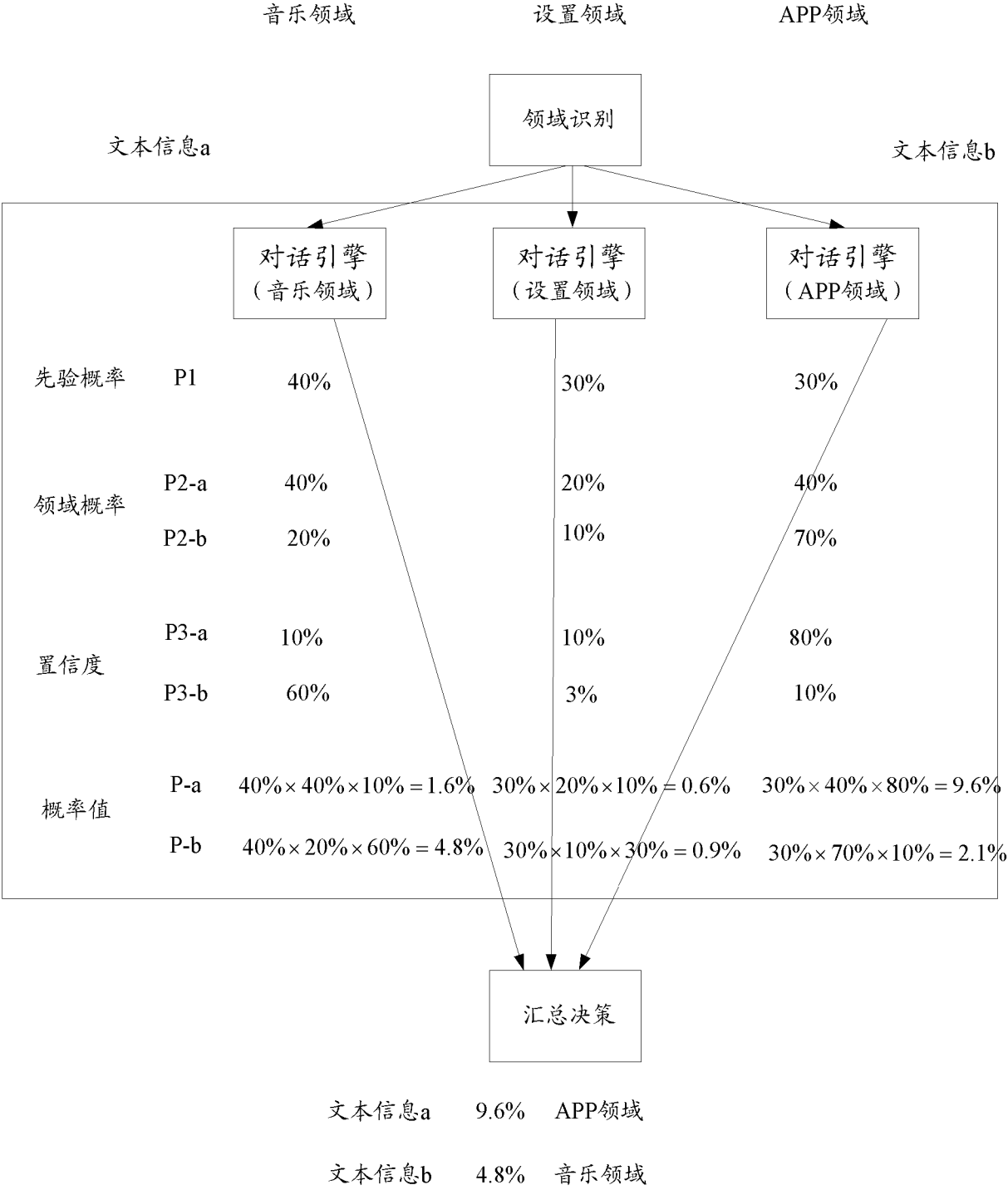


图 14

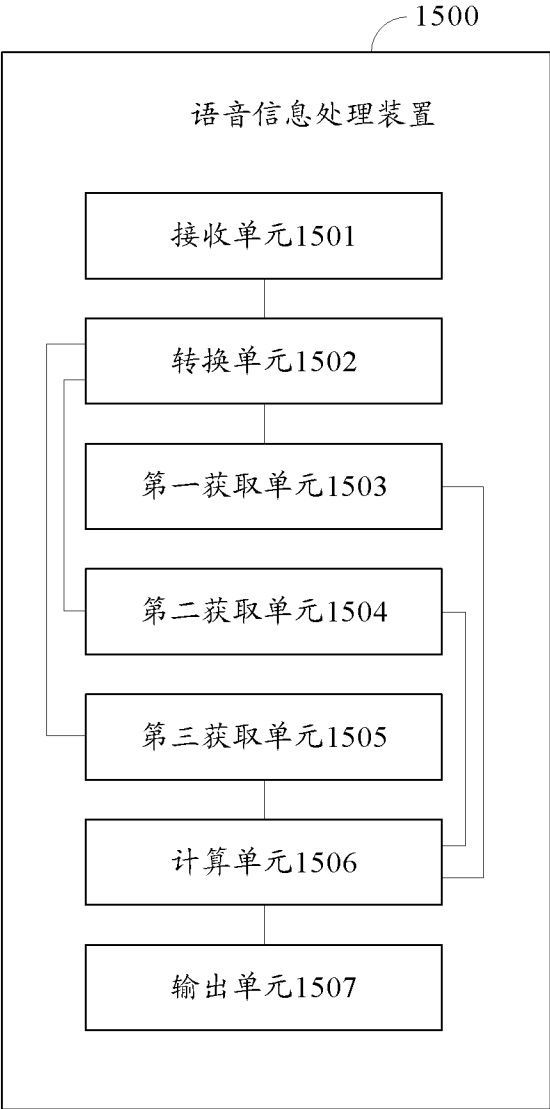


图 15

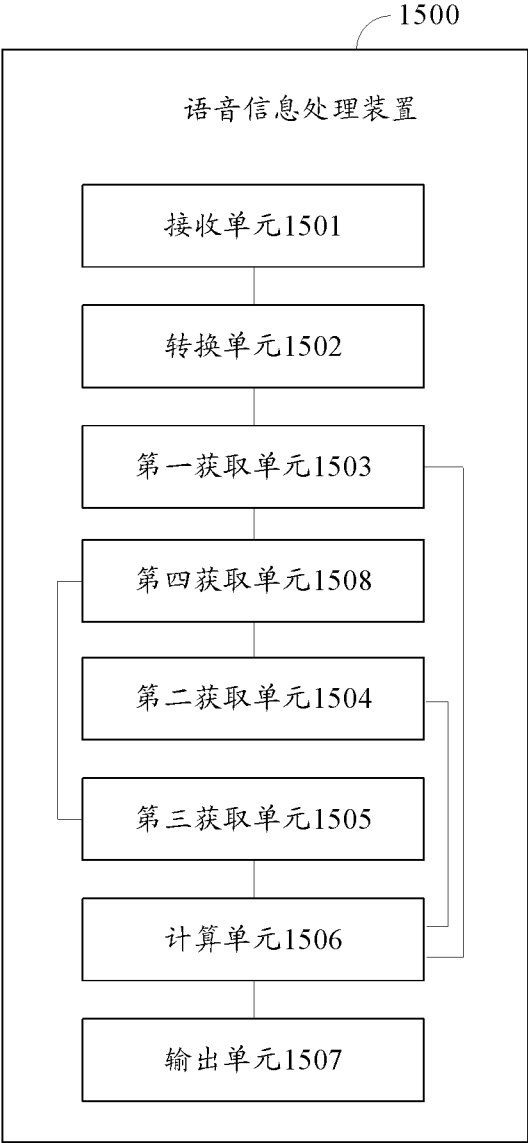


图 16

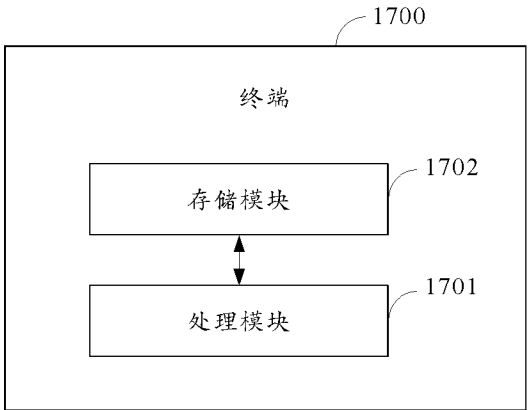


图 17

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/106168

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)

G06F; G10L

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)

USTXT; CPRSABS; CNTXT; CNABS; DWPI: 语音, 文本, 语义, 理解, 分析, 识别, 领域, 事件, 置信度, 先验, 概率, voice, text, semantics, comprehension, analyze, identify, field, event, confidence w level, prior w probability

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 105869629 A (LETV HOLDING (BEIJING) CO., LTD. et al.) 17 August 2016 (17.08.2016), entire document	1-18
A	CN 107092593 A (CENTRAL CHINA NORMAL UNIVERSITY) 25 August 2017 (25.08.2017), entire document	1-18
A	CN 101587493 A (UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA) 25 November 2009 (25.11.2009), entire document	1-18
A	CN 105632487 A (BEIJING QIYI CENTURY TECHNOLOGY CO., LTD.) 01 June 2016 (01.06.2016), entire document	1-18

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 23 March 2018	Date of mailing of the international search report 29 March 2018
Name and mailing address of the ISA 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 CHEN, An'an Telephone No. (86-10) 62411828

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/106168

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105869629 A	17 August 2016	WO 2017166650 A1	05 October 2017
CN 107092593 A	25 August 2017	None	
CN 101587493 A	25 November 2009	CN 101587493 B	04 July 2012
CN 105632487 A	01 June 2016	None	

国际检索报告

国际申请号

PCT/CN2017/106168

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)

G06F; G10L

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))

USTXT;CPRSABS;CNTXT;CNABS;DWPI:语音, 文本, 语义, 理解, 分析, 识别, 领域, 事件, 置信度, 先验, 概率, voice, text, semantics, comprehension, analyze, identify, field, event, confidence w level, prior w probability

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 105869629 A (乐视控股北京有限公司等) 2016年 8月 17日 (2016 - 08 - 17) 全文	1-18
A	CN 107092593 A (华中师范大学) 2017年 8月 25日 (2017 - 08 - 25) 全文	1-18
A	CN 101587493 A (中国科学技术大学) 2009年 11月 25日 (2009 - 11 - 25) 全文	1-18
A	CN 105632487 A (北京奇艺世纪科技有限公司) 2016年 6月 1日 (2016 - 06 - 01) 全文	1-18

☐ 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 3月 23日

国际检索报告邮寄日期

2018年 3月 29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

受权官员

陈安安

电话号码 (86-10)62411828

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/106168

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	105869629	A	2016年 8月 17日	WO	2017166650	A1	2017年 10月 5日
CN	107092593	A	2017年 8月 25日	无			
CN	101587493	A	2009年 11月 25日	CN	101587493	B	2012年 7月 4日
CN	105632487	A	2016年 6月 1日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)