

(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
2015 年 2 月 5 日 (05.02.2015)



(10) 国际公布号
WO 2015/014122 A1

- (51) 国际专利分类号:
G10L 15/00 (2013.01) *G06F 17/30* (2006.01)
- (21) 国际申请号: PCT/CN2014/072487
- (22) 国际申请日: 2014 年 2 月 25 日 (25.02.2014)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201310330351.4 2013 年 7 月 31 日 (31.07.2013) CN
- (71) 申请人: 深圳创维-RGB 电子有限公司 (SHENZHEN SKYWORTH-RGB ELECTRONIC CO., LTD.) [CN/CN]; 中国广东省深圳市南山区深南大道创维大厦 A 座 13-16 楼, Guangdong 518057 (CN)。
- (72) 发明人: 及
- (71) 申请人 (仅对美国): 付春元 (FU, Chunyuan) [CN/CN]; 中国广东省深圳市南山区深南大道创维大厦 A 座 13-16 楼, Guangdong 518057 (CN)。 雷倍一 (LEI, Beiyi) [CN/CN]; 中国广东省深圳市南山区

深南大道创维大厦 A 座 13-16 楼, Guangdong 518057 (CN)。

- (74) 代理人: 深圳市世纪恒程知识产权代理事务所 (SHENZHEN CENTURY-FOREVER INTELLECTUAL PROPERTY OFFICE); 中国广东省深圳市南山区南山大道 3838 号设计产业园金栋二层 210-212 (原南头城工业村 11 栋), Guangdong 518052 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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,

[见续页]

(54) Title: VOICE INTERACTION METHOD AND SYSTEM AND INTERACTION TERMINAL

(54) 发明名称: 语音交互的方法、系统以及交互终端

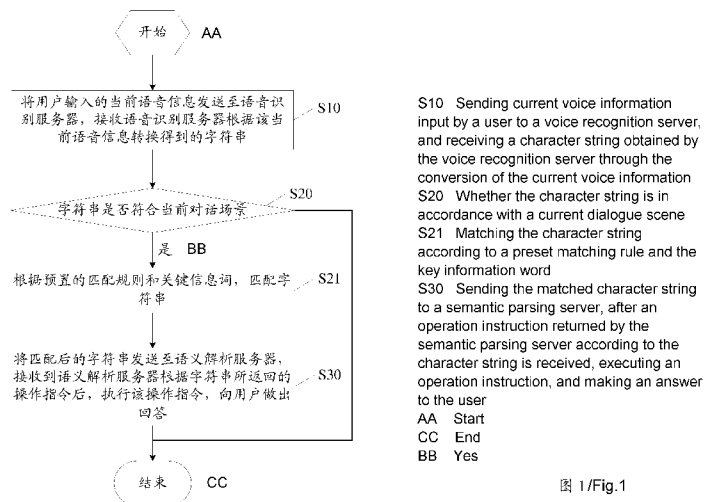


图 1/ Fig.1

(57) Abstract: A voice interaction method, a corresponding system and an interaction terminal. The method comprises: sending current voice information input by a user to a voice recognition server, and receiving a character string obtained by the voice recognition server through the conversion of the current voice information (S10); judging whether the character string is in accordance with a current dialogue scene according to a key information word of the current dialogue scene that is extracted from the context in the dialogue process of a user and stored (S20); matching the character string according to a preset matching rule and the key information word if the character string is in accordance with the current dialogue scene (S21); sending the matched character string to a semantic parsing server, after an operation instruction returned by the semantic parsing server according to the character string is received, executing the operation instruction, and making an answer to the user (S30). By the adoption of the solution disclosed, the sense of experience of the man-machine dialogue is improved, user-friendly service is provided for the user, and user-friendly man-machine interaction is realized.

(57) 摘要:

[见续页]



WO 2015/014122 A1



RW, SD, SL, 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))。

一种语音交互的方法、相应的系统和一种交互终端，所述方法包括：将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当前语音信息转换得到的字符串（S10）；根据从与用户对话过程中的上下文中提取出来并保存的当前对话场景的关键信息词，判断字符串是否符合当前对话场景（S20）；若是，则根据预置的匹配规则和关键信息词，匹配字符串（S21）；将匹配后的字符串发送至语义解析服务器，接收到语义解析服务器根据字符串所返回的操作指令后，执行该操作指令，向用户做出回答（S30）。采用所公开的方案，提升了人机对话的体验感，并且为用户提供了人性化的服务，实现了人性化的人机交互。

说明书

语音交互的方法、系统以及交互终端

5 技术领域

本发明涉及到电子产品技术领域，特别涉及到一种语音交互的方法、系统以及交互终端。

背景技术

10 语音交互是与机器进行语言交流，让机器明白你说的是什么，并给出相应的答案，当输入语音信号时，机器先通过识别和理解过程把语音信号转变为相应的文本或命令，然后归为不同的类别，最后根据不同的类别信息分析出对应的答案。目前，该技术已经进入工业、家电、通信、汽车电子、消费电子产品等各个领域。但是，现有的语音交互主要是停留在一问一答的形式
15 上，用户问什么，智能终端就回答什么，略显呆板；在很多情况下智能终端不能理解用户输入语句的意思，或者不能理解用户输入的缺少关键信息的语句或者省略句的意思；并且，由于智能终端不能根据上下文语义环境分析当前用户输入的语句的意思，因而不能根据当前对话环境给用户必要的提示和引导。这样，就会导致人机对话的体验感较差，使交互欠缺人性化。

20

发明内容

本发明的主要目的为提供一种语音交互的方法、系统以及交互终端，旨在提升人机对话的体验感，并且为用户提供人性化的服务，实现人性化的交互。

25 本发明提供一种语音交互的方法，包括步骤：

将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当前语音信息转换得到的字符串；

根据保存的当前对话场景的关键信息词，判断所述字符串是否符合所述当前对话场景；若是，则根据预置的匹配规则和所述关键信息词，匹配所述
30 字符串；

将匹配后的所述字符串发送至语义解析服务器，接收到语义解析服务器

根据所述字符串所返回的操作指令后，执行该操作指令，向用户做出回答。

优选地，预置的匹配规则包括分词处理、相似度计算、匹配方式和补全信息。

优选地，所述匹配方式包括精确匹配、中心匹配和广泛匹配。

5 优选地，所述将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当前语音信息转换得到的字符串的步骤之前，还包括：

解析用户输入的语音信息，提取该语音信息中的关键信息词，将该关键信息词保存至对应的对话场景中。

10 优选地，所述根据预置的匹配规则匹配所述字符串与所述关键信息词的步骤包括：

计算所述字符串与所述关键信息词的相似度；

根据计算出的相似度以及预置的匹配规则，判断所述字符串与所述关键信息词是否完全匹配；

15 若所述字符串与所述关键信息词部分匹配，则根据所述关键信息词以及该字符串缺失的信息，补全该字符串。

优选地，所述字符串与所述关键信息词的相似度通过以下方法计算：

假设上一次语音信息为 A，当前语音信息为 B，A 包含的关键信息词为 A1、A2、A3、...、An，B 包含的关键信息词为 B1、B2、B3、...、Bm，则 A 和 B 的相似度 S(A, B) 为：

$$20 \quad S(A, B) = \left(\frac{\sum_{i=1}^m a_i}{m} + \frac{\sum_{i=1}^n a_i}{n} \right) / 2 \quad ;$$

其中：

$$\begin{aligned} a_i &= \max(s(A_i, B_1), s(A_i, B_2), \dots, s(A_i, B_m)) \\ b_i &= \max(s(A_1, B_i), s(A_2, B_i), \dots, s(A_n, B_i)) \end{aligned} \quad \circ$$

优选地，在所述判断字符串是否符合所述当前对话场景的步骤之后，还包括：

25 若所述字符串不符合所述当前对话场景，比对该字符串与其他对话场景对应的关键信息词，确定所述字符串所符合的对话场景，并将当前对话场景修改为该对话场景。

本发明进一步提供一种交互终端，包括：

30 发送及接收模块，用于将用户输入的当前语音信息发送至语音识别服务

器，接收语音识别服务器根据该当前语音信息转换得到的字符串；

判断模块，用于根据保存的当前对话场景的关键信息词，判断所述字符串是否符合所述当前对话场景；

匹配模块，用于根据预置的匹配规则和所述关键信息词，匹配所述字符串；

操作模块，用于将匹配后的所述字符串发送至语义解析服务器，接收到语义解析服务器根据所述字符串所返回的操作指令后，执行该操作指令，向用户做出回答。

优选地，预置的匹配规则包括分词处理、相似度计算、匹配方式和补全信息。

优选地，所述匹配方式包括精确匹配、中心匹配和广泛匹配。

优选地，交互终端还包括：

解析及保存模块，用于解析用户输入的语音信息，提取该语音信息中的关键信息词，将该关键信息词保存至对应的对话场景中。

优选地，所述匹配模块包括：

计算单元，用于计算所述字符串与所述关键信息词的相似度；

判断单元，用于根据计算出的相似度以及预置的匹配规则，判断所述字符串与所述关键信息词是否完全匹配；

补全单元，用于根据所述关键信息词以及该字符串缺失的信息，补全该字符串。

优选地，交互终端还包括：

比对模块，用于比对所述字符串与其他对话场景对应的关键信息词；

确定及修改模块，用于确定所述字符串所符合的对话场景，并将当前对话场景修改为该对话场景。

本发明进一步还提供一种语音交互的系统，包括交互终端、语音识别服务器和语义解析服务器，其中：

所述交互终端包括：

发送及接收模块，用于将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当前语音信息转换得到的字符串；

判断模块，用于根据保存的当前对话场景的关键信息词，判断所述字符

串是否符合所述当前对话场景；

匹配模块，用于根据预置的匹配规则匹配所述字符串与所述关键信息词；

操作模块，用于将匹配后的所述字符串发送至语义解析服务器，接收到语义解析服务器根据所述字符串所返回的操作指令后，执行该操作指令，向
5 用户做出回答；

所述语音识别服务器，用于接收所述交互终端发送的用户输入的当前语音信息，将该当前语音信息转换为对应的字符串；

所述语义解析服务器，用于接收所述交互终端发送的匹配后的字符串，解析所述字符串对应的语义，并生成相应的操作指令。

10

本发明通过接收到用户输入的当前语音信息后，将其发送至语音识别服务器，并在接收到语音识别服务器根据该当前语音信息转换得到的字符串后，根据保存的当前对话场景的关键信息词，判断字符串是否符合当前对话场景；若是，根据预置的匹配规则和关键信息词，匹配字符串；然后将匹配后的字符串发送至语义解析服务器，在接收到语义解析服务器根据字符串所返回的操作指令后，执行该操作指令，向用户做出回答，从而提升了人机对话的体验感，并且为用户提供了人性化的服务，实现了人性化的人机交互。
15

附图说明

20 图1为本发明语音交互的方法第一实施例的流程示意图；
图2为本发明语音交互的方法中匹配字符串与关键信息词的流程示意图；
图3为本发明语音交互的方法第二实施例的流程示意图；
图4为本发明语音交互的方法第三实施例的流程示意图；
图5为本发明交互终端第一实施例的结构示意图；
25 图6为本发明交互终端的匹配模块的结构示意图；
图7为本发明交互终端第二实施例的结构示意图；
图8为本发明交互终端第三实施例的结构示意图；
图9为本发明语音交互的系统第一实施例的结构示意图。

30 本发明目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

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。

- 5 本发明提供一种语音交互的方法。在接收到用户输入的当前语音信息后，根据上下文语义的关联性，判断当前语音信息是否符合当前对话场景，并在符合时执行相应的操作指令以及向用户回答。

参照图1，图1为本发明语音交互的方法第一实施例的流程示意图。

本实施例所提供的语音交互的方法，包括：

- 10 步骤 S10，将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当前语音信息转换得到的字符串；

在用户与交互终端进行语音交互的过程中，接收到用户通过语音输入设备所输入的当前语音信息后，将该当前语音信息发送至语音识别服务器，语音识别服务器对该当前语音信息进行识别，并将音频文件转换为对应的以文字信息表示的字符串，将该字符串返回至交互终端。

15

步骤 S20，根据保存的当前对话场景的关键信息词，判断字符串是否符合当前对话场景；若是，则执行步骤 S21；

步骤 S21，根据预置的匹配规则和关键信息词，匹配字符串；

- 接收到语音识别服务器返回的字符串后，根据交互终端中所保存的当前对话场景的关键信息词，判断该字符串是否符合当前对话场景。本实施例中，可以包括多种对话场景，例如：聊天场景、天气场景、地图场景、电影音乐场景、股票场景、菜谱场景、笑话场景、周公解梦场景、日期查询场景、翻译场景和网页浏览场景等可以通过和交互终端对话得到回答的对话场景；在每一种对话场景下进行语音交互时，交互终端都会保存用户对话过程中的关键信息词，并将其与相应的对话场景进行对应。
- 20
- 25

当判断出字符串符合当前对话场景，即字符串所对应的当前语音信息为当前对话场景下的内容时，根据预置的匹配规则以及当前对话场景对应的关键信息词，匹配字符串。本实施例中，预置的匹配规则为预先设定的用于对字符串进行处理的规则，以完善字符串的信息；匹配规则可以包括分词处理、相似度计算、匹配方式和补全信息等步骤。

30

请一并参照图 2，图 2 为本发明语音交互的方法中匹配字符串与关键信息

词的流程图示意图。

在本实施例中，步骤 S21 具体包括：

步骤 S211，计算字符串与关键信息词的相似度；

5 在判断出字符串符合当前对话场景时，根据所保存的当前对话场景下的上一次语音信息，计算当前语音信息和上一次语音信息的相似度，可以根据以下公式计算：

假设上一次语音信息为 A，当前语音信息为 B，A 包含的关键信息词为 A1、A2、A3、...、An，B 包含的关键信息词为 B1、B2、B3、...、Bm，则 A 和 B 的相似度为：

$$10 \quad S(A,B)=\left(\frac{\sum_{i=1}^m a_i}{m}+\frac{\sum_{i=1}^n a_i}{n}\right)/2$$

其中：

$$a_i = \max(s(A_i, B_1), s(A_i, B_2), \dots, s(A_i, B_m))$$

$$b_i = \max(s(A_1, B_i), s(A_2, B_i), \dots, s(A_n, B_i))$$

在相似度计算时，充分考虑了句子的每个词的深层含义。

15 步骤 S212，根据计算出的相似度以及预置的匹配规则，判断字符串与关键信息词是否完全匹配；

本实施例中，预置的匹配规则中的匹配方法包括精确匹配、中心匹配和广泛匹配；其中，精确匹配是指当前语音信息和上一次语音信息的关键信息词完全相同，中心匹配是指当前语音信息中包含上一次语音信息的关键信息词即可，而广泛匹配是指当前语音信息中包含上一次语音信息的关键信息词或包括与该关键信息词相关的词。在计算出当前语音信息和上一次语音信息的相似度后，根据匹配方法，判断字符串与关键信息词是否完全匹配，即是否符合精确匹配的条件。如符合，则说明根据该字符串即可得到完整的语义。

步骤 S213，若字符串与关键信息词部分匹配，则根据关键信息词以及该字符串缺失的信息，补全该字符串。

25 如字符串不符合精确匹配的条件，即字符串与关键信息词部分匹配，则根据关键信息词判断字符串的关键词中所缺失的信息，补全该字符串，使其与当前对话场景下的关键信息词完全匹配。

步骤 S30，将匹配后的字符串发送至语义解析服务器，接收到语义解析服务器根据字符串所返回的操作指令后，执行该操作指令，向用户做出回答。

30 在根据预置的匹配规则和关键信息词匹配字符串，使其与当前对话场景

下的关键信息词完全匹配后，将匹配后的包括完整关键信息词的字符串发送至语义解析服务器，供语义解析服务器解析该字符串的完整语义，语义解析服务器解析出字符串的完整语义后，根据该字符串的语义生成进行相关操作的操作指令，并发送至交互终端。交互终端在接收到语义解析服务器根据字符串所返回的操作指令后，执行该操作指令，并向用户做出相应的回答。

本实施例通过接收到用户输入的当前语音信息后，将其发送至语音识别服务器，并在接收到语音识别服务器根据该当前语音信息转换得到的字符串后，根据保存的当前对话场景的关键信息词，判断字符串是否符合当前对话场景；若是，根据预置的匹配规则和关键信息词，匹配字符串；然后将匹配后的字符串发送至语义解析服务器，在接收到语义解析服务器根据字符串所返回的操作指令后，执行该操作指令，向用户做出回答，从而提升了人机对话的体验感，并且为用户提供了人性化的服务，实现了人性化的人机交互。

参照图3，图3为本发明语音交互的方法第二实施例的流程示意图。

在本发明语音交互的方法第一实施例的基础上，在执行步骤S10之前，该方法还包括：

步骤 S40，解析用户输入的语音信息，提取该语音信息中的关键信息词，将该关键信息词保存至对应的对话场景中。

本实施例中，在相应的对话场景下，接收到用户输入的语音信息，解析该语音信息，并提取出其中的关键信息词，保存至该对话场景下，对每次接收到的语音信息都进行该处理，在之后所接收到当前语音信息时，通过所保存的上一次语音信息的关键信息词，即可以判断出当前语音信息是否符合当前对话场景。例如，在聊天场景中，如用户输入“小李是我的朋友”，交互终端就会通过分词处理得到关键信息“小李/nr”、“我/r”、“朋友/n”，对话场景会记录为“聊天场景”；当用户再输入“谁是我的朋友”时，交互终端就会根据聊天场景解析该语音信息，并采用预置的匹配规则匹配已经记录的关键信息后给出答案“小李”。

在相应的对话场景下，接收到用户输入的语音信息，解析该语音信息，并提取出其中的关键信息词，保存至该对话场景下，在之后所接收到当前语音信息时，通过所保存的上一次语音信息的关键信息词，即可以判断出当前语音信息是否符合当前对话场景，从而为更加智能地识别用户输入的语音信

息提供了基础。

参照图4，图4为本发明语音交互的方法第三实施例的流程示意图。

在本发明语音交互的方法第一实施例和第二实施例的基础上，在执行步骤S20之后，该方法还包括：

步骤 S50，若字符串不符合当前对话场景，比对该字符串与其他对话场景对应的关键信息词，确定字符串所符合的对话场景，并将当前对话场景修改为该对话场景。

在根据交互终端中所保存的当前对话场景的关键信息词，判断出该字符串不符合当前对话场景时，比对该字符串与交互终端中所保存的其他对话场景对应的关键信息词，并根据比对结果确定字符串所符合的对话场景，即当字符串与某一个对话场景的关键信息词相匹配时，则确定该字符串符合该对话场景，同时，将当前对话场景修改为字符串所符合的对话场景。

在判断出字符串不符合当前对话场景时，比对该字符串与交互终端中所保存的其他对话场景对应的关键信息词，确定该字符串符合的对话场景，并将当前对话场景修改为所确定的对话场景，进一步提升了人机对话的体验感，并进一步保证了能够为用户提供人性化的服务。

本发明还提供一种交互终端。

参照图5，图5为本发明交互终端第一实施例的结构示意图。

本实施例所提供的交互终端，包括：

发送及接收模块 101，用于将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当前语音信息转换得到的字符串；

判断模块 102，用于根据保存的当前对话场景的关键信息词，判断字符串是否符合当前对话场景；

匹配模块 103，用于根据预置的匹配规则和关键信息词，匹配字符串；

操作模块 104，用于将匹配后的字符串发送至语义解析服务器，接收到语义解析服务器根据字符串所返回的操作指令后，执行该操作指令，向用户做出回答。

在用户与交互终端进行语音交互的过程中，接收到用户通过语音输入设备所输入的当前语音信息后，发送及接收模块 101 将该当前语音信息发送至

语音识别服务器，语音识别服务器对该当前语音信息进行识别，并将音频文件转换为对应的以文字信息表示的字符串，将该字符串返回至交互终端。

发送及接收模块 101 接收到语音识别服务器返回的字符串后，判断模块 102 根据交互终端中所保存的当前对话场景的关键信息词，判断该字符串是否符合当前对话场景。本实施例中，可以包括多种对话场景，例如：聊天场景、天气场景、地图场景、电影音乐场景、股票场景、菜谱场景、笑话场景、周公解梦场景、日期查询场景、翻译场景和网页浏览场景等可以通过和交互终端对话得到回答的对话场景；在每一种对话场景下进行语音交互时，交互终端都会保存用户对话过程中的关键信息词，并将其与相应的对话场景进行对应。

当判断出字符串符合当前对话场景，即字符串所对应的当前语音信息为当前对话场景下的内容时，匹配模块 103 根据预置的匹配规则以及当前对话场景对应的关键信息词，匹配字符串。本实施例中，预置的匹配规则为预先设定的用于对字符串进行处理的规则，以完善字符串的信息；匹配规则可以包括分词处理、相似度计算、匹配方式和补全信息等步骤。

请一并参照图6，图6为本发明交互终端的匹配模块的结构示意图。

在本实施例中，匹配模块 103 包括：

计算单元 1031，用于计算字符串与关键信息词的相似度；

判断单元 1032，用于根据计算出的相似度以及预置的匹配规则，判断字符串与关键信息词是否完全匹配；

补全单元 1033，用于根据关键信息词以及该字符串缺失的信息，补全该字符串。

在判断出字符串符合当前对话场景时，根据所保存的当前对话场景下的上一次语音信息，通过计算单元 1031 计算当前语音信息和上一次语音信息的相似度，可以根据以下公式计算：

假设上一次语音信息为 A，当前语音信息为 B，A 包含的关键信息词为 A1、A2、A3、...、An，B 包含的关键信息词为 B1、B2、B3、...、Bm，这两个句子的相似度为：

$$S(A,B)=(\frac{\sum_{i=1}^m a_i}{m} + \frac{\sum_{i=1}^n a_i}{n})/2$$

其中：

$$a_i = \max(s(A_i, B_1), s(A_i, B_2), \dots, s(A_i, B_m))$$
$$b_i = \max(s(A_1, B_i), s(A_2, B_i), \dots, s(A_n, B_i))$$

在相似度计算时，充分考虑了句子的每个词的深层含义。

本实施例中，预置的匹配规则中的匹配方法包括精确匹配、中心匹配和广泛匹配；其中，精确匹配是指当前语音信息和上一次语音信息的关键信息词完全相同，中心匹配是指当前语音信息中包含上一次语音信息的关键信息词即可，而广泛匹配是指当前语音信息中包含上一次语音信息的关键信息词或包括与该关键信息词相关的词。在计算出当前语音信息和上一次语音信息的相似度后，判断单元 1032 根据匹配方法，判断字符串与关键信息词是否完全匹配，即是否符合精确匹配的条件。如符合，则说明根据该字符串即可得到完整的语义。

如字符串不符合精确匹配的条件，即字符串与关键信息词部分匹配，则通过补全单元 1033，根据关键信息词判断字符串的关键词中所缺失的信息，补全该字符串，使其与当前对话场景下的关键信息词完全匹配。

在根据预置的匹配规则和关键信息词匹配字符串，使其与当前对话场景下的关键信息词完全匹配后，将匹配后的包括完整关键信息词的字符串发送至语义解析服务器，供语义解析服务器解析该字符串的完整语义，语义解析服务器解析出字符串的完整语义后，根据该字符串的语义生成进行相关操作的操作指令，并发送至交互终端。交互终端在接收到语义解析服务器根据字符串所返回的操作指令后，通过操作模块 104 执行该操作指令，并向用户做出相应的回答。

本实施例通过接收到用户输入的当前语音信息后，将其发送至语音识别服务器，并在接收到语音识别服务器根据该当前语音信息转换得到的字符串后，根据保存的当前对话场景的关键信息词，判断字符串是否符合当前对话场景；若是，根据预置的匹配规则和关键信息词，匹配字符串；然后将匹配后的字符串发送至语义解析服务器，在接收到语义解析服务器根据字符串所返回的操作指令后，执行该操作指令，向用户做出回答，从而提升了人机对话的体验感，并且为用户提供了人性化的服务，实现了人性化的人机交互。

参照图7，图7为本发明交互终端第二实施例的结构示意图。

在本发明交互终端第一实施例的基础上，该交互终端还包括：

解析及保存模块 105, 用于解析用户输入的语音信息, 提取该语音信息中的关键信息词, 将该关键信息词保存至对应的对话场景中。

本实施例中, 在相应的对话场景下, 接收到用户输入的语音信息, 解析及保存模块105解析该语音信息, 并提取出其中的关键信息词, 保存至该对话场景下, 对每次接收到的语音信息都进行该处理, 在之后所接收到当前语音信息时, 通过所保存的上一次语音信息的关键信息词, 即可以判断出当前语音信息是否符合当前对话场景。例如, 在聊天场景中, 如用户输入“小李是我的朋友”, 交互终端就会通过分词处理得到关键信息“小李/nr”、“我/r”、“朋友/n”, 对话场景会记录为“聊天场景”; 当用户再输入“谁是我的朋友”时, 交互终端就会根据聊天场景解析该语音信息, 并采用预置的匹配规则匹配已经记录的关键信息后给出答案“小李”。

在相应的对话场景下, 接收到用户输入的语音信息, 解析该语音信息, 并提取出其中的关键信息词, 保存至该对话场景下, 在之后所接收到当前语音信息时, 通过所保存的上一次语音信息的关键信息词, 即可以判断出当前语音信息是否符合当前对话场景, 从而为更加智能地识别用户输入的语音信息提供了基础。

参照图8, 图8为本发明交互终端第三实施例的结构示意图。

在本发明交互终端第一实施例和第二实施例的基础上, 该交互终端还包括:

比对模块 106, 用于比对字符串与其他对话场景对应的关键信息词;

确定及修改模块 107, 用于确定字符串所符合的对话场景, 并将当前对话场景修改为该对话场景。

在根据交互终端中所保存的当前对话场景的关键信息词, 判断出该字符串不符合当前对话场景时, 通过比对模块106比对该字符串与交互终端中所保存的其他对话场景对应的关键信息词, 确定及修改模块107根据比对结果确定字符串所符合的对话场景, 即当字符串与某一个对话场景的关键信息词相匹配时, 则确定该字符串符合该对话场景, 同时, 确定及修改模块107将当前对话场景修改为字符串所符合的对话场景。

在判断出字符串不符合当前对话场景时, 比对该字符串与交互终端中所保存的其他对话场景对应的关键信息词, 确定该字符串符合的对话场景, 并

将当前对话场景修改为所确定的对话场景，进一步提升了人机对话的体验感，并进一步保证了能够为用户提供人性化的服务。

本发明进一步还提供一种语音交互的系统。

5 参照图9，图9为本发明语音交互的系统第一实施例的结构示意图。

本实施例所提供的语音交互的系统，包括交互终端100、语音识别服务器200和语义解析服务器300，其中：

交互终端 100 包括：

10 发送及接收模块 101，用于将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当前语音信息转换得到的字符串；

判断模块 102，用于根据保存的当前对话场景的关键信息词，判断字符串是否符合当前对话场景；

匹配模块 103，用于根据预置的匹配规则和关键信息词，匹配字符串；

15 操作模块 104，用于将匹配后的字符串发送至语义解析服务器，接收到语义解析服务器根据字符串所返回的操作指令后，执行该操作指令，向用户做出回答；

语音识别服务器 200，用于接收交互终端 100 发送的用户输入的当前语音信息，将该当前语音信息转换为对应的字符串；

20 语义解析服务器 300，用于接收交互终端发送的匹配后的字符串，解析字符串对应的语义，并生成相应的操作指令。

本语音交互的系统包括了上述交互终端全部实施例的全部技术方案，所达到的技术效果也完全相同，在此不做赘述。

25 以上所述仅为本发明的优选实施例，并非因此限制本发明的专利范围，凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本发明的专利保护范围。

权 利 要 求 书

1、一种语音交互的方法，其特征在于，包括步骤：

将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务
5 器根据该当前语音信息转换得到的字符串；

根据保存的当前对话场景的关键信息词，判断所述字符串是否符合所述
当前对话场景；若是，则根据预置的匹配规则和所述关键信息词，匹配所述
字符串；

将匹配后的所述字符串发送至语义解析服务器，接收到语义解析服务器
10 根据所述字符串所返回的操作指令后，执行该操作指令，向用户做出回答。

2、根据权利要求1所述的语音交互的方法，其特征在于，预置的匹配规
则包括分词处理、相似度计算、匹配方式和补全信息。

15 3、根据权利要求2所述的语音交互的方法，其特征在于，所述匹配方式
包括精确匹配、中心匹配和广泛匹配。

4、根据权利要求1所述的语音交互的方法，其特征在于，在所述将用户
输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当
20 前语音信息转换得到的字符串的步骤之前，还包括：

解析用户输入的语音信息，提取该语音信息中的关键信息词，将该关键
信息词保存至对应的对话场景中。

5、根据权利要求4所述的语音交互的方法，其特征在于，所述根据预置
25 的匹配规则匹配所述字符串与所述关键信息词的步骤包括：

计算所述字符串与所述关键信息词的相似度；

根据计算出的相似度以及预置的匹配规则，判断所述字符串与所述关键
信息词是否完全匹配；

若所述字符串与所述关键信息词部分匹配，则根据所述关键信息词以及
30 该字符串缺失的信息，补全该字符串。

6、根据权利要求5所述的语音交互的方法，其特征在于，所述字符串与所述关键信息词的相似度通过以下方法计算：

假设上一次语音信息为A，当前语音信息为B，A包含的关键信息词为A1、A2、A3、...、An，B包含的关键信息词为B1、B2、B3、...、Bm，则A和B的相似度S(A, B)为：

$$S(A,B)=\left(\frac{\sum_{i=1}^m a_i}{m}+\frac{\sum_{i=1}^n a_i}{n}\right)/2 \quad ;$$

其中：

$$\begin{aligned} a_i &= \max(s(A_i, B_1), s(A_i, B_2), \dots, s(A_i, B_m)) \\ b_i &= \max(s(A_1, B_i), s(A_2, B_i), \dots, s(A_n, B_i)) \end{aligned} \quad \circ$$

7、根据权利要求6所述的语音交互的方法，其特征在于，在所述判断字符串是否符合所述当前对话场景的步骤之后，还包括：

若所述字符串不符合所述当前对话场景，比对该字符串与其他对话场景对应的关键信息词，确定所述字符串所符合的对话场景，并将当前对话场景修改为该对话场景。

8、一种交互终端，其特征在于，包括：

发送及接收模块，用于将用户输入的当前语音信息发送至语音识别服务器，接收语音识别服务器根据该当前语音信息转换得到的字符串；

判断模块，用于根据保存的当前对话场景的关键信息词，判断所述字符串是否符合所述当前对话场景；

匹配模块，用于根据预置的匹配规则和所述关键信息词，匹配所述字符串；

操作模块，用于将匹配后的所述字符串发送至语义解析服务器，接收到语义解析服务器根据所述字符串所返回的操作指令后，执行该操作指令，向用户做出回答。

9、根据权利要求8所述的交互终端，其特征在于，预置的匹配规则包括分词处理、相似度计算、匹配方式和补全信息。

10、根据权利要求9所述的交互终端，其特征在于，所述匹配方式包括

精确匹配、中心匹配和广泛匹配。

11、根据权利要求 8 所述的交互终端，其特征在于，还包括：

解析及保存模块，用于解析用户输入的语音信息，提取该语音信息中的
5 关键信息词，将该关键信息词保存至对应的对话场景中。

12、根据权利要求 11 所述的交互终端，其特征在于，所述匹配模块包括：

计算单元，用于计算所述字符串与所述关键信息词的相似度；

判断单元，用于根据计算出的相似度以及预置的匹配规则，判断所述字
10 符串与所述关键信息词是否完全匹配；

补全单元，用于根据所述关键信息词以及该字符串缺失的信息，补全该
字符串。

13、根据权利要求 12 所述的交互终端，其特征在于，还包括：

15 比对模块，用于比对所述字符串与其他对话场景对应的关键信息词；

确定及修改模块，用于确定所述字符串所符合的对话场景，并将当前对
话场景修改为该对话场景。

14、一种语音交互的系统，包括如权利要求 9 至 13 中任一项所述的交互
20 终端、语音识别服务器和语义解析服务器，其特征在于，其中：

所述交互终端包括：

发送及接收模块，用于将用户输入的当前语音信息发送至语音识别服务
器，接收语音识别服务器根据该当前语音信息转换得到的字符串；

判断模块，用于根据保存的当前对话场景的关键信息词，判断所述字符
25 串是否符合所述当前对话场景；

匹配模块，用于根据预置的匹配规则匹配所述字符串与所述关键信息词；

操作模块，用于将匹配后的所述字符串发送至语义解析服务器，接收到
语义解析服务器根据所述字符串所返回的操作指令后，执行该操作指令，向
用户做出回答；

30 所述语音识别服务器，用于接收所述交互终端发送的用户输入的当前语
音信息，将该当前语音信息转换为对应的字符串；

所述语义解析服务器，用于接收所述交互终端发送的匹配后的字符串，解析所述字符串对应的语义，并生成相应的操作指令。

说明书附图

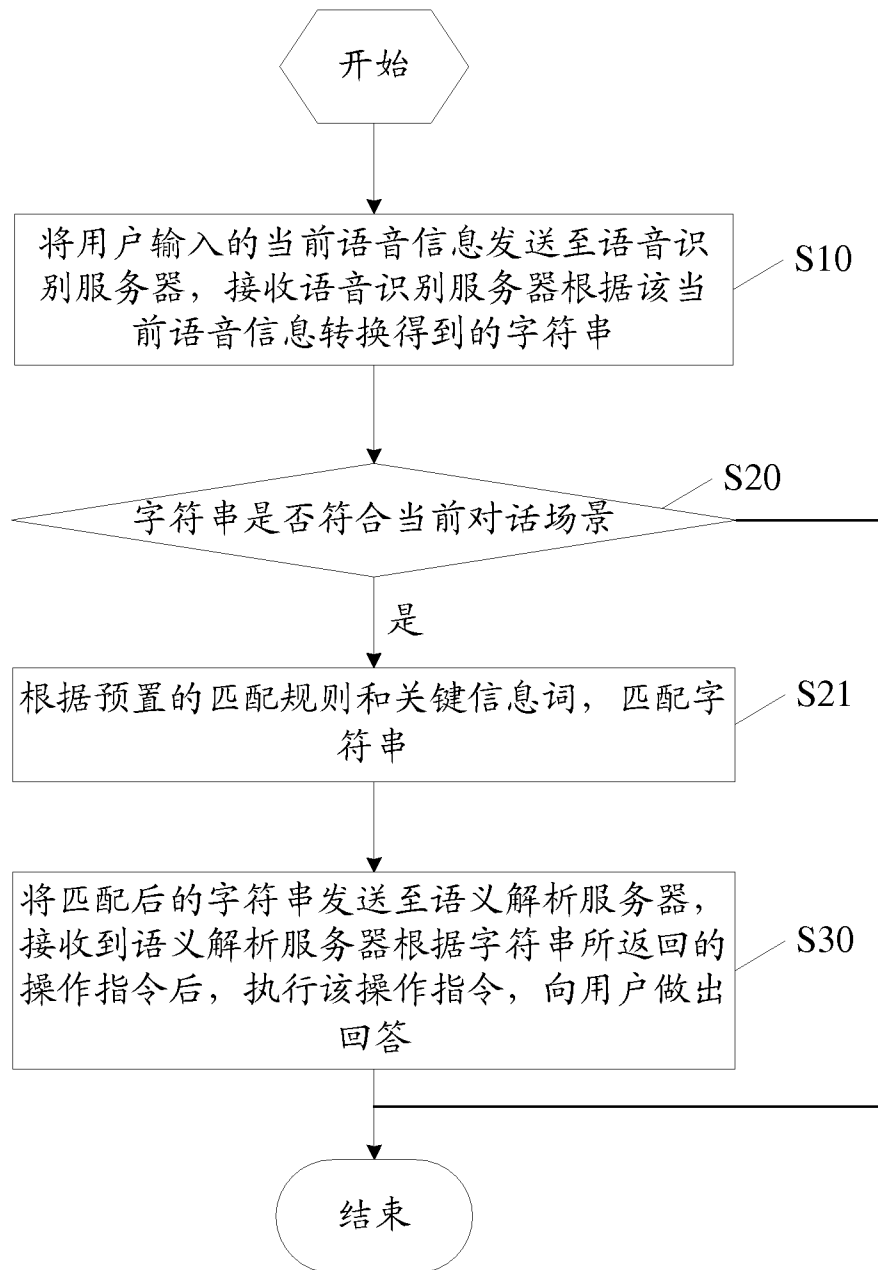


图 1

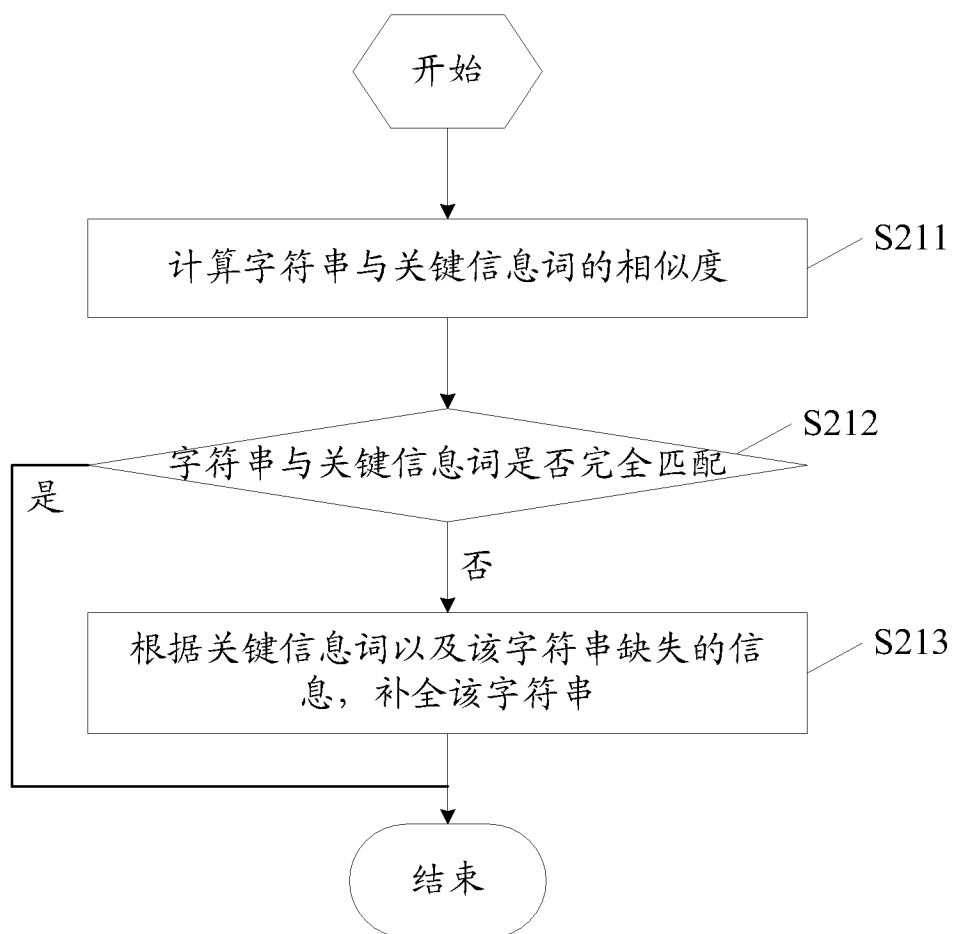


图 2

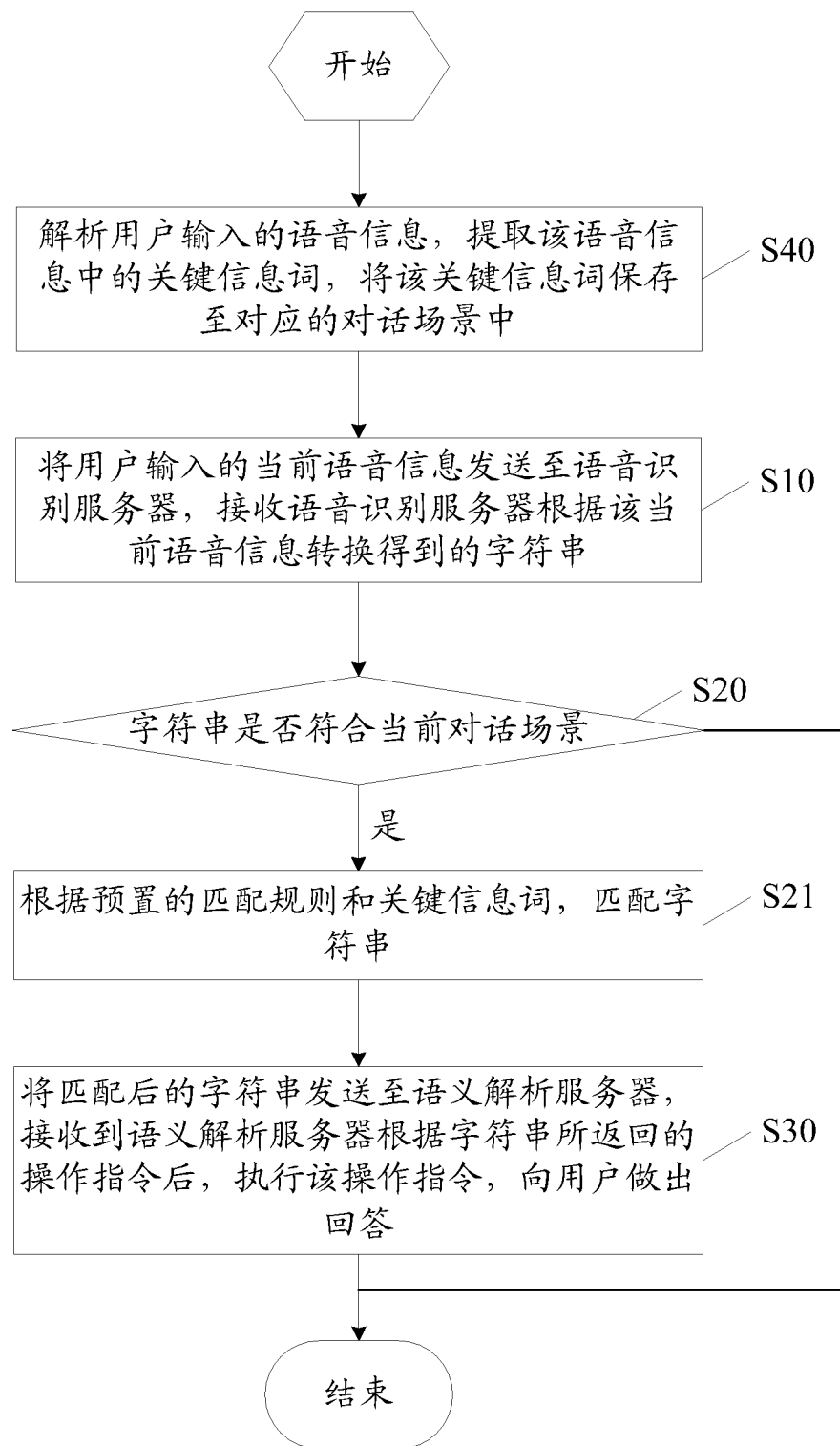


图 3

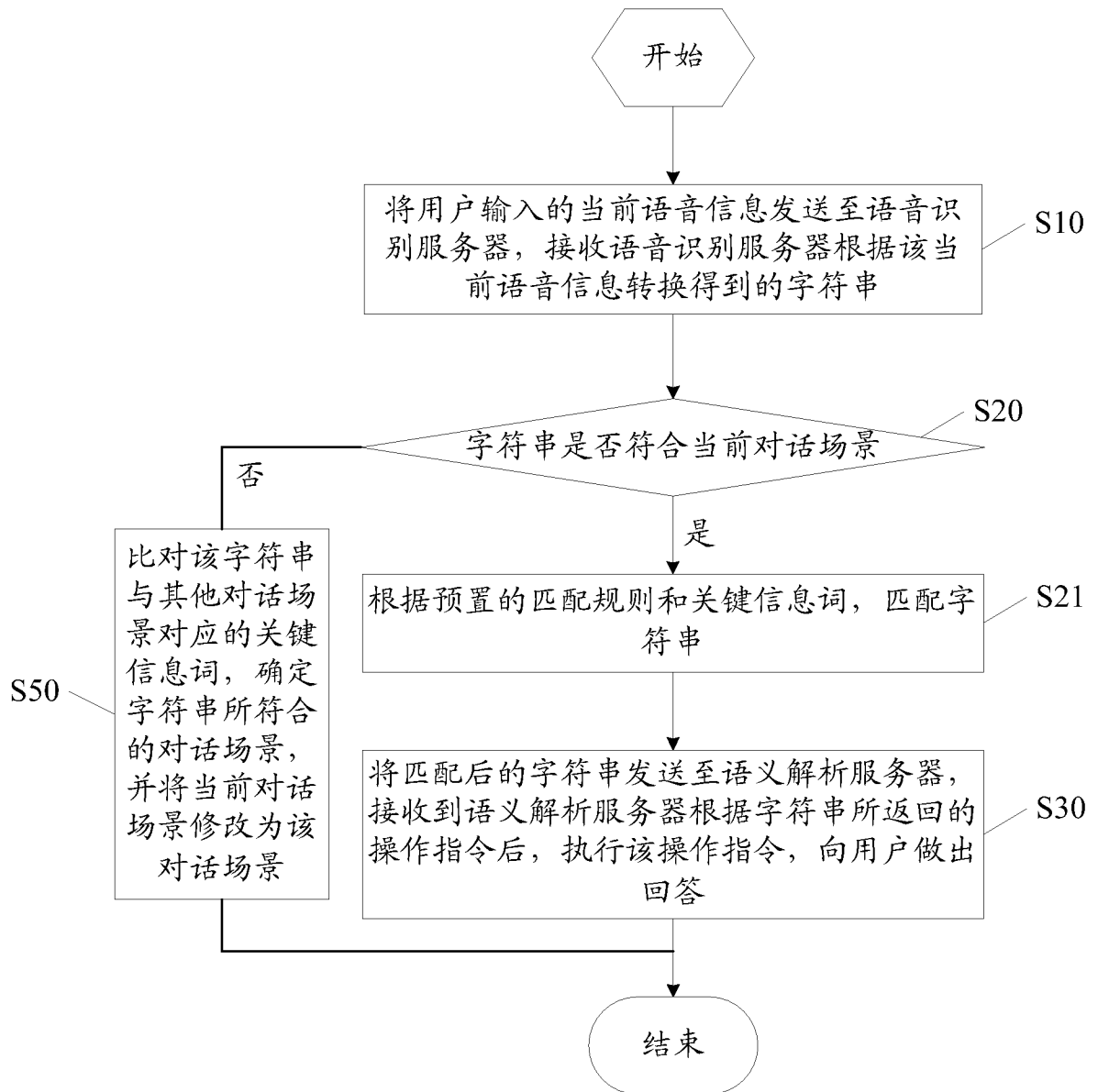


图 4

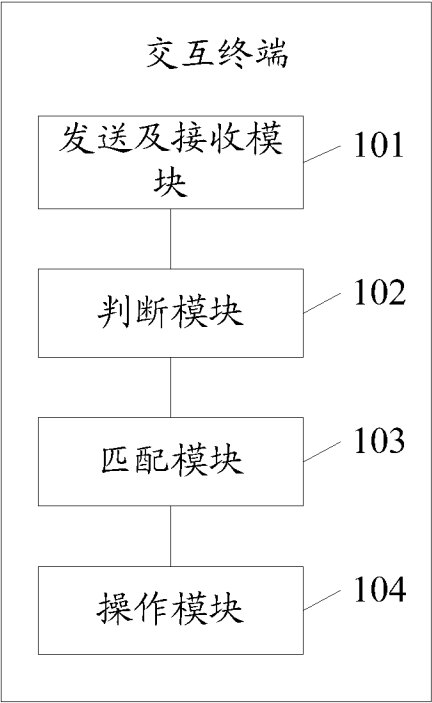


图 5



图 6

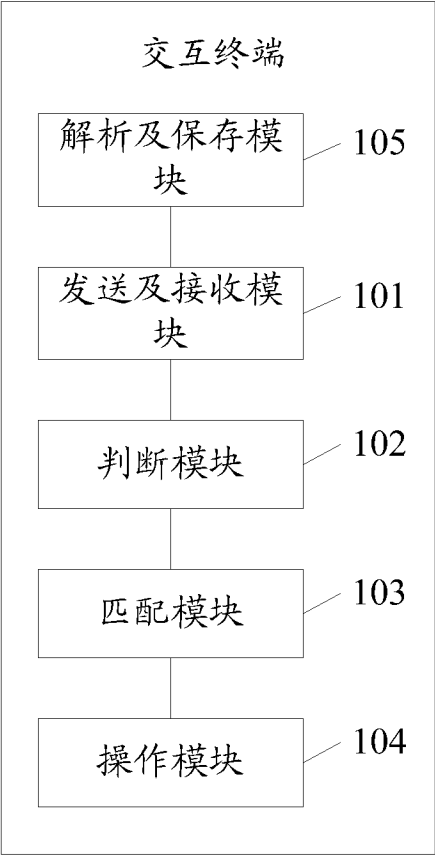


图 7

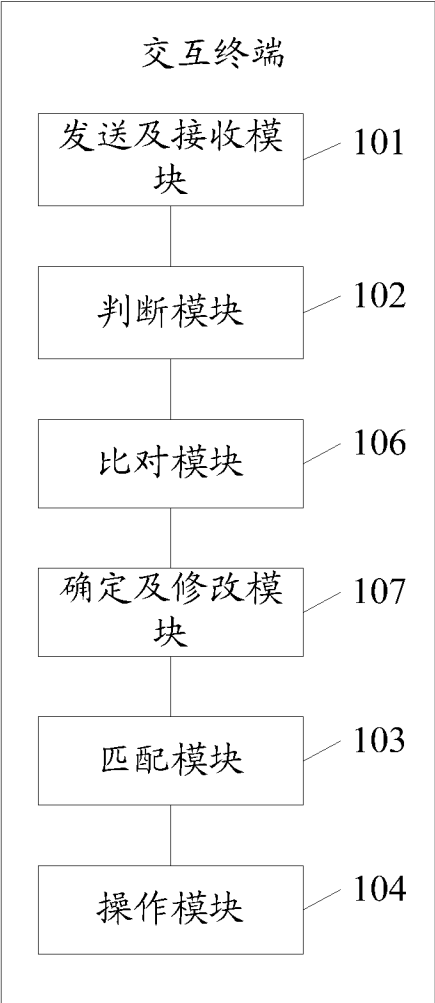


图 8

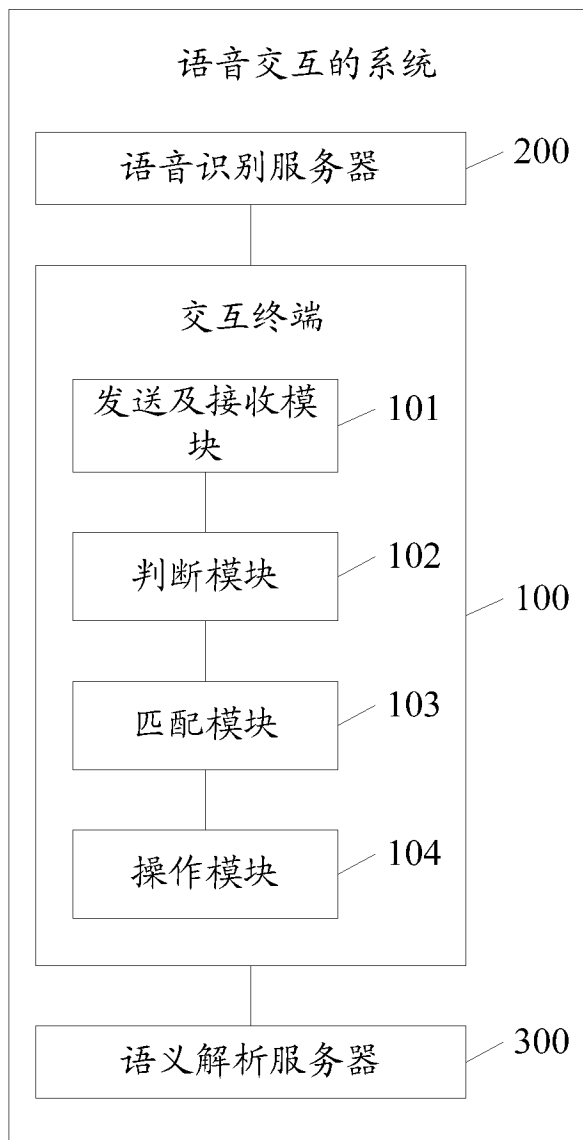


图 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2014/072487

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/00 (2013.01) i;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)

IPC: G10L, G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CPRSABS; CNABS; VEN: SCENE, INSTRUCTION, ORDER, KEY 1w WORD

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103000173 A (YOUSHI TECHNOLOGY CO., LTD.) 27 March 2013 (27.03.2013) description, paragraphs [0081]-[0091], [0094]-[0117], and figures 4-7	1-14
X	CN 1855101 A (SONY COMPUTER ENTERTAINMENT INC.) 01 November 2006 (01.11.2006) description, page 3, line 5 to page 17, line 18, and figures 1-8, 10 and 13	1-14
X	EP 1058453 A1 (MATSUSHITA ELECTRIC IND. CO., LTD. et al.) 06 December 2000 (06.12.2000) description, paragraphs [0048]-[0057], and figure 9	1-14
A	CN 101477798 A (UNIV. BEIJING POSTS & TELECOM) 08 July 2009 (08.07.2009) the whole document	1-14
A	CN 101329667 A (SHENZHEN DAZHENG CHINESE LANGUAGE SOFTWARE CO., LTD.) 24 December 2008 (24.12.2008) the whole document	1-14
A	US 7212972 B2 (DDI CORP.) 01 May 2007 (01.05.2007) the whole document	1-14
PX	CN 103413549 A (SHENZHEN SKYWORTH RGB ELECTRONIC CO., LTD.) 27 November 2013 (27.11.2013) description, paragraphs [0055]-[0139]	1-14

☐ 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 04 May 2014	Date of mailing of the international search report 31 July 2014
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 SUI, Xin Telephone No. (86-10) 62085740

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2014/072487

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103000173 A	27 March 2013	None	
CN 1855101 A	01 November 2006	KR 20060110787 A	25 October 2006
		CN 100449544 C	07 January 2009
		JP 2006301967 A	02 November 2006
		JP 4079275 B2	23 April 2008
		JP 2006301968 A	02 November 2006
		EP 1715475 A1	25 October 2006
		US 2006247932 A1	02 November 2006
EP 1058453 A1	06 December 2000	DE 69915455 D1	15 April 2004
		EP 1058453 B1	10 March 2004
		US 2003200091 A1	23 October 2003
		EP 1357542 A2	29 October 2003
		US 6728673 B2	27 April 2004
		DE 69924765 D1	19 May 2005
		WO 0036833 A1	22 June 2000
		JP 3252282 B2	04 February 2002
		US 6611803 B1	26 August 2003
		EP 1357542 B1	13 April 2005
		JP 2000236494 A	29 August 2000
		DE 69924765 T2	22 September 2005
		JP 3607228 B2	05 January 2005
		JP 2002175304 A	21 June 2002
CN 101477798 A	08 July 2009	CN 101477798 B	05 January 2011
CN 101329667 A	24 December 2008	None	
US 7212972 B2	01 May 2007	US 2001003813 A1	14 June 2001
		JP 2001167557 A	22 June 2001
		JP 2001167109 A	22 June 2001
CN103413549 A	27 November 2013	None	

A. 主题的分类		
G10L 15/00(2013.01)i; 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)		
G10L; G06F		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))		
CPRSABS;CNABS, VEN:场景, 命令, 语义, 指令, SCENE, INSTRUCTION, ORDER, KEY 1w WORD		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 103000173 A (优视科技有限公司) 2013年 3月 27日 (2013 - 03 - 27) 说明书[0081]-[0091], [0094]-[0117]段, 附图4-7	1-14
X	CN 1855101 A (索尼计算机娱乐公司) 2006年 11月 01日 (2006 - 11 - 01) 说明书第3页第5行到第17页第18行, 附图1-8, 10, 13	1-14
X	EP 1058453 A1 (MATSUSHITA ELECTRIC IND. CO., LTD. 等) 2000年 12月 06日 (2000 - 12 - 06) 说明书[0048]-[0057]段, 附图9	1-14
A	CN 101477798 A (北京邮电大学) 2009年 7月 08日 (2009 - 07 - 08) 全文	1-14
A	CN 101329667 A (深圳市大正汉语软件有限公司) 2008年 12月 24日 (2008 - 12 - 24) 全文	1-14
A	US 7212972 B2 (DDI CORP.) 2007年 5月 01日 (2007 - 05 - 01) 全文	1-14
PX	CN 103413549 A (深圳创维—RGB电子有限公司) 2013年 11月 27日 (2013 - 11 - 27) 说明书[0055]段-[0139]段	1-14
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2014年 5月 04日		2014年 7月 31日
ISA/CN的名称和邮寄地址		受权官员
中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国		隋欣
传真号 (86-10)62019451		电话号码 (86-10)62085740

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2014/072487

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 103000173 A	2013年 3月 27日	无	
CN 1855101 A	2006年 11月 01日	KR 20060110787A	2006年 10月 25日
		CN 100449544C	2009年 1月 07日
		JP 2006301967A	2006年 11月 02日
		JP 4079275B2	2008年 4月 23日
		JP 2006301968A	2006年 11月 02日
		EP 1715475A1	2006年 10月 25日
		US 2006247932A1	2006年 11月 02日
EP 1058453 A1	2000年 12月 06日	DE 69915455D1	2004年 4月 15日
		EP 1058453B1	2004年 3月 10日
		US 2003200091A1	2003年 10月 23日
		EP 1357542A2	2003年 10月 29日
		US 6728673B2	2004年 4月 27日
		DE 69924765D1	2005年 5月 19日
		WO 0036833A1	2000年 6月 22日
		JP 3252282B2	2002年 2月 04日
		US 6611803B1	2003年 8月 26日
		EP 1357542B1	2005年 4月 13日
		JP 2000236494A	2000年 8月 29日
		DE 69924765T2	2005年 9月 22日
		JP 3607228B2	2005年 1月 05日
		JP 2002175304A	2002年 6月 21日
CN 101477798 A	2009年 7月 08日	CN 101477798B	2011年 1月 05日
CN 101329667 A	2008年 12月 24日	无	
US 7212972 B2	2007年 5月 01日	US 2001003813A1	2001年 6月 14日
		JP 2001167557A	2001年 6月 22日
		JP 2001167109A	2001年 6月 22日
CN 103413549 A	2013年 11月 27日	无	