



(12)发明专利

(10)授权公告号 CN 104268166 B

(45)授权公告日 2017.04.19

(21)申请号 201410456861.0

G06F 3/023(2006.01)

(22)申请日 2014.09.09

审查员 徐波

(65)同一申请的已公布的文献号

申请公布号 CN 104268166 A

(43)申请公布日 2015.01.07

(73)专利权人 北京搜狗科技发展有限公司

地址 100084 北京市海淀区中关村东路1号
院9号楼搜狐网络大厦9层01房间

(72)发明人 崔欣 任尚昆 苏其龙 张扬

(74)专利代理机构 北京润泽恒知识产权代理有
限公司 11319

代理人 苏培华

(51)Int.Cl.

G06F 17/30(2006.01)

G06F 17/27(2006.01)

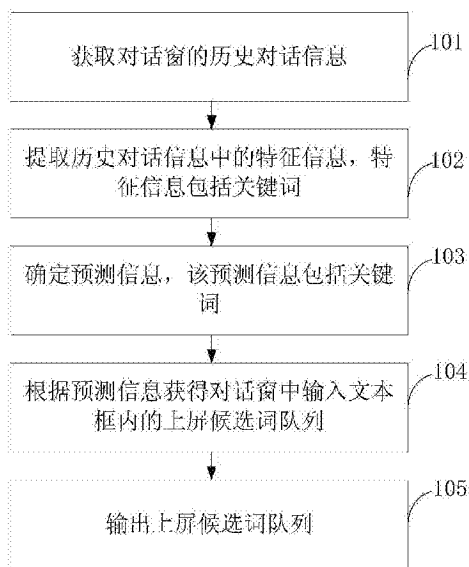
权利要求书3页 说明书11页 附图4页

(54)发明名称

一种输入方法、装置和电子设备

(57)摘要

本发明实施例提供了一种输入方法、装置及电子设备。该输入方法包括：获取对话框的历史对话信息；提取所述历史对话信息中的特征信息，所述特征信息包括关键词；确定预测信息，所述预测信息包括所述关键词；根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列；输出所述上屏候选词队列。本发明实施例通过充分挖掘对话框的历史对话信息，综合从中提取的特征信息来确定预测信息，并依据预测信息查找相应的联想候选词库获得了最终的上屏候选队列。该方法无论输入文本框内容是否为空仍然可以获得上屏候选队列，而且所获得的上屏候选队列显然比单单依据用户上一次上屏词条获得的上屏候选队列更加合理，更加准确。



1. 一种输入方法,其特征在于,包括:
获取对话框的历史对话信息;
提取所述历史对话信息中的特征信息,所述特征信息包括关键词;
确定预测信息,所述预测信息包括所述关键词;
根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列;
输出所述上屏候选词队列;
其中,所述确定预测信息包括:
判断所述对话框的输入文本框内是否存在文本信息;
若否,确定所述预测信息为所述关键词。
2. 根据权利要求1所述的方法,其特征在于,所述获取对话框的历史对话信息,包括:
获取所述对话框预设历史时长内的历史对话信息;或者,
获取所述对话框内预设信息条数的历史对话信息。
3. 根据权利要求1所述的方法,其特征在于,所述确定预测信息还包括:
在所述对话框的输入文本框内存在文本信息时,判断所述文本信息与所述关键词之间是否存在话题转移;
若否,确定所述预测信息为所述关键词和所述文本信息。
4. 根据权利要求3所述的方法,其特征在于,所述判断所述文本信息与所述关键词之间是否存在话题转移,包括:
判断所述文本信息命中的情景特征标签与所述关键词命中的情景特征标签是否存在重叠,若是,则不存在话题转移,若否,则存在话题转移。
5. 根据权利要求3所述的方法,其特征在于,所述根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列,包括:
确定所述预测信息命中的情景特征标签;
对所述情景特征标签进行排序;
按照所述情景特征标签的排序,根据所述预测信息在所述情景特征标签的联想候选词库中查找上屏候选词;
根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列。
6. 根据权利要求5所述的方法,其特征在于,所述对所述情景特征标签进行排序,包括:
根据命中所述情景特征标签的所述预测信息的个数,以及所述预测信息在所述对话框中出现的时间对所述情景特征标签进行排序。
7. 根据权利要求5所述的方法,其特征在于,所述特征信息包括标识词,在所述根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列之前,还包括:
对所述标识词按照情景特征标签进行分类存储;
按照所述情景特征标签的排序,查找具有相应情景特征标签的标识词,将查找到的标识词作为所述输入文本框内的上屏候选词。
8. 根据权利要求5所述的方法,其特征在于,在所述根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列之前,还包括:
根据所述预测信息在语言模型的联想候选词库中查找上屏候选词;其中,当所述预测信息为所述关键词时,以所述关键词为左元查询,所述语言模型为二元语言模型;当所述预

测信息为所述关键词及所述文本信息时,以所述关键词为左元,以所述文本信息为中元查询,所述语言模型为三元语言模型,或者,以所述文本信息为左元查询,所述语言模型为二元语言模型。

9. 一种输入装置,其特征在于,包括:

信息获取单元,用于获取对话框的历史对话信息;

信息提取单元,用于提取所述历史对话信息中的特征信息,所述特征信息包括关键词;

信息确定单元,用于确定预测信息,所述预测信息包括所述关键词;

预测单元,用于根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列;

输出单元,用于输出所述上屏候选词队列;

其中,所述信息确定单元包括:

第一判断子单元,用于判断所述对话框的输入文本框内是否存在文本信息;

第一确定子单元,用于当所述第一判断子单元判断不存在文本信息时,确定所述预测信息为所述关键词。

10. 根据权利要求9所述的装置,其特征在于,

所述信息获取单元,具体用于获取所述对话框预设历史时长内的历史对话信息;或者,获取所述对话框内预设信息条数的历史对话信息。

11. 根据权利要求9所述的装置,其特征在于,所述信息确定单元还包括:

第二判断子单元,用于当所述第一判断子单元判断存在文本信息时,判断所述文本信息与所述关键词之间是否存在话题转移;

第二确定子单元,用于当所述第二判断子单元判断不存在话题转移时,确定所述预测信息为所述关键词和所述文本信息。

12. 根据权利要求11所述的装置,其特征在于,

所述第二判断子单元,具体用于判断所述文本信息命中的情景特征标签与所述关键词命中的情景特征标签是否存在重叠,若是,则不存在话题转移,若否,则存在话题转移。

13. 根据权利要求11所述的装置,其特征在于,所述预测单元包括:

标签确定子单元,用于确定所述预测信息命中的情景特征标签;

排序子单元,用于对所述情景特征标签进行排序;

第一查找子单元,用于按照所述情景特征标签的排序,根据所述预测信息在所述情景特征标签的联想候选词库中查找上屏候选词;

队列获得子单元,用于根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列。

14. 根据权利要求13所述的装置,其特征在于,

所述排序子单元,具体用于根据命中所述情景特征标签的所述预测信息的个数,以及所述预测信息在所述对话框中出现的时间对所述情景特征标签进行排序。

15. 根据权利要求13所述的装置,其特征在于,所述特征信息包括标识词,所述预测单元还包括:

存储子单元,用于在所述队列获得子单元获得所述输入文本框内的上屏候选词队列之前,对所述标识词按照情景特征标签进行分类存储;

第二查找子单元,用于按照所述情景特征标签的排序,查找具有相应情景特征标签的标识词,将查找到的标识词作为所述输入文本框内的上屏候选词。

16. 根据权利要求13所述的装置,其特征在于,所述预测单元还包括:

第三查找子单元,用于在所述队列获得子单元获得所述输入文本框内的上屏候选词队列之前,根据所述预测信息在语言模型的联想候选词库中查找上屏候选词;其中,当所述预测信息为所述关键词时,以所述关键词为左元查询,所述语言模型为二元语言模型;当所述预测信息为所述关键词及所述文本信息时,以所述关键词为左元,以所述文本信息为中元查询,所述语言模型为三元语言模型,或者,以所述文本信息为左元查询,所述语言模型为二元语言模型。

17. 一种电子设备,其特征在于,包括存储器和处理器,所述存储器用于存储计算机指令或代码,所述处理器和所述存储器耦合,用于执行所述存储器中的计算机指令或代码,实现以下方法:

获取对话框的历史对话信息;

提取所述历史对话信息中的特征信息,所述特征信息包括关键词;

确定预测信息,所述预测信息包括所述关键词;

根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列;

输出所述上屏候选词队列;

其中,所述确定预测信息包括:

判断所述对话框的输入文本框内是否存在文本信息;

若否,确定所述预测信息为所述关键词。

一种输入方法、装置和电子设备

技术领域

[0001] 本发明涉及通信技术领域,尤其涉及一种输入方法、装置及电子设备。

背景技术

[0002] 拼音输入是一种最简单的汉字输入方法,它的发展非常快,从第一代的以字输入为主,即用户一次只能输入一个汉字,发展到第二代以词输入为主并具有智能调频功能,这个时候主要依赖的是输入法的词典,发展到第三代,用户可以进行语句的输入,输入法词典中没有的句子也可以进行输入,组词功能对输入的体验有着很大的影响。

[0003] 输入法联想功能是拼音输入法主动输入的一种扩展,它的出现减少了用户主动输入的次数、按键的次数,并增加了输入法的智能性。该输入法的实现过程是首先获取用户在输入文本框中上一次上屏的词条,并根据该词条查询系统二元库等预建词库来获取上屏候选词队列,然后输出该上屏候选词队列。

[0004] 然而,该输入法中的上屏候选词队列仅依赖于上一次上屏的词条,会导致预测出的上屏候选词队列可能是片面的或不合理的,甚至当输入文本框中不存在上一次上屏的词条时,就无法给出上屏候选词队列。

发明内容

[0005] 本发明实施例所要解决的技术问题是提供一种输入方法,能够给出更合理的上屏候选词队列。

[0006] 相应的,本发明实施例还提供了一种输入装置及电子设备,用以保证上述方法的实现及应用。

[0007] 为了解决上述问题,本发明公开了一种输入方法,包括:

[0008] 获取对话窗的历史对话信息;

[0009] 提取所述历史对话信息中的特征信息,所述特征信息包括关键词;

[0010] 确定预测信息,所述预测信息包括所述关键词;

[0011] 根据所述预测信息获得所述对话窗中输入文本框内的上屏候选词队列;

[0012] 输出所述上屏候选词队列。

[0013] 进一步,所述获取对话窗的历史对话信息,包括:

[0014] 获取所述对话窗预设历史时长内的历史对话信息;或者,

[0015] 获取所述对话窗内预设信息条数的历史对话信息。

[0016] 进一步,所述确定预测信息包括:

[0017] 判断所述对话窗的输入文本框内是否存在文本信息;

[0018] 若否,确定所述预测信息为所述关键词;若是,判断所述文本信息与所述关键词之间是否存在话题转移;

[0019] 若否,确定所述预测信息为所述关键词和所述文本信息。

[0020] 进一步,所述判断所述文本信息与所述关键词之间是否存在话题转移,包括:

[0021] 判断所述文本信息命中的情景特征标签与所述关键词命中的情景特征标签是否存在重叠,若是,则不存在话题转移,若否,则存在话题转移。

[0022] 进一步,所述根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列,包括:

[0023] 确定所述预测信息命中的情景特征标签;

[0024] 对所述情景特征标签进行排序;

[0025] 按照所述情景特征标签的排序,根据所述预测信息在所述情景特征标签的联想候选词库中查找上屏候选词;

[0026] 根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列。

[0027] 进一步,所述对所述情景特征标签进行排序,包括:

[0028] 根据命中所述情景特征标签的所述预测信息的个数,以及所述预测信息在所述对话框中出现的时间对所述情景特征标签进行排序。

[0029] 进一步,所述特征信息包括标识词,在所述根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列之前,还包括:

[0030] 对所述标识词按照情景特征标签进行分类存储;

[0031] 按照所述情景特征标签的排序,查找具有相应情景特征标签的标识词,将查找到的标识词作为所述输入文本框内的上屏候选词。

[0032] 进一步,在所述根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列之前,还包括:

[0033] 根据所述预测信息在语言模型的联想候选词库中查找上屏候选词;其中,当所述预测信息为所述关键词时,以所述关键词为左元查询,所述语言模型为二元语言模型;当所述预测信息为所述关键词及所述文本信息时,以所述关键词为左元,以所述文本信息为中元查询,所述语言模型为三元语言模型,或者,以所述文本信息为左元查询,所述语言模型为二元语言模型。

[0034] 本发明还公开了一种输入装置,包括:

[0035] 信息获取单元,用于获取对话框的历史对话信息;

[0036] 信息提取单元,用于提取所述历史对话信息中的特征信息,所述特征信息包括关键词;

[0037] 信息确定单元,用于确定预测信息,所述预测信息包括所述关键词;

[0038] 预测单元,用于根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列;

[0039] 输出单元,用于输出所述上屏候选词队列。

[0040] 进一步,所述信息获取单元,具体用于获取所述对话框预设历史时长内的历史对话信息;或者,获取所述对话框内预设信息条数的历史对话信息。

[0041] 进一步,所述信息确定单元包括:

[0042] 第一判断子单元,用于判断所述对话框的输入文本框内是否存在文本信息;

[0043] 第一确定子单元,用于当所述第一判断子单元判断不存在文本信息时,确定所述预测信息为所述关键词;

[0044] 第二判断子单元,用于当所述第一判断子单元判断存在文本信息时,判断所述文

本信息与所述关键词之间是否存在话题转移；

[0045] 第二确定子单元,用于当所述第二判断子单元判断不存在话题转移时,确定所述预测信息为所述关键词和所述文本信息。

[0046] 进一步,所述第二判断子单元,具体用于判断所述文本信息命中的情景特征标签与所述关键词命中的情景特征标签是否存在重叠,若是,则不存在话题转移,若否,则存在话题转移。

[0047] 进一步,所述预测单元包括:

[0048] 标签确定子单元,用于确定所述预测信息命中的情景特征标签;

[0049] 排序子单元,用于对所述情景特征标签进行排序;

[0050] 第一查找子单元,用于按照所述情景特征标签的排序,根据所述预测信息在所述情景特征标签的联想候选词库中查找上屏候选词;

[0051] 队列获得子单元,用于根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列。

[0052] 进一步,所述排序子单元,具体用于根据命中所述情景特征标签的所述预测信息的个数,以及所述预测信息在所述对话框中出现的时间对所述情景特征标签进行排序。

[0053] 进一步,所述特征信息包括标识词,所述预测单元还包括:

[0054] 存储子单元,用于在所述队列获得子单元获得所述输入文本框内的上屏候选词队列之前,对所述标识词按照情景特征标签进行分类存储;

[0055] 第二查找子单元,用于按照所述情景特征标签的排序,查找具有相应情景特征标签的标识词,将查找到的标识词作为所述输入文本框内的上屏候选词。

[0056] 进一步,所述预测单元还包括:

[0057] 第三查找子单元,用于在所述队列获得子单元获得所述输入文本框内的上屏候选词队列之前,根据所述预测信息在语言模型的联想候选词库中查找上屏候选词;其中,当所述预测信息为所述关键词时,以所述关键词为左元查询,所述语言模型为二元语言模型;当所述预测信息为所述关键词及所述文本信息时,以所述关键词为左元,以所述文本信息为中元查询,所述语言模型为三元语言模型,或者,以所述文本信息为左元查询,所述语言模型为二元语言模型。

[0058] 本发明还公开了一种电子设备,包括存储器和处理器,所述存储器用于存储计算机指令或代码,所述处理器和所述存储器耦合,用于执行所述存储器中的计算机指令或代码,实现以下方法:

[0059] 获取对话框的历史对话信息;

[0060] 提取所述历史对话信息中的特征信息,所述特征信息包括关键词;

[0061] 确定预测信息,所述预测信息包括所述关键词;

[0062] 根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列;

[0063] 输出所述上屏候选词队列。

[0064] 与现有技术相比,本发明实施例包括以下优点:

[0065] 本发明实施例通过充分挖掘对话框的历史对话信息,综合从中提取的特征信息来确定预测信息,并依据预测信息查找相应的联想候选词库获得了最终的上屏候选队列。该方法无论输入文本框内容是否为空仍然可以获得上屏候选队列,而且所获得的上屏候选队

列显然比单单依据用户上一次上屏词条获得的上屏候选队列更加合理,更加准确。

附图说明

- [0066] 图1是本发明实施例一种输入方法的流程图;
- [0067] 图2是本发明实施例中一种确定预测信息的方法流程图;
- [0068] 图3是本发明实施例中输入文本框为空时的输入方法流程图;
- [0069] 图4是本发明实施例中输入文本框不为空时的输入方法流程图;
- [0070] 图5是本发明实施例一种输入装置的结构示意图;
- [0071] 图6是本发明实施例中一种信息确定单元的结构示意图;
- [0072] 图7是本发明实施例中一种预测单元的结构示意图。

具体实施方式

[0073] 为使本发明的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0074] 参照图1,为本发明实施例一种输入方法的流程图。

[0075] 该方法可以包括:

[0076] 步骤101,获取对话框的历史对话信息。

[0077] 本实施例中,对话框可以是指与某一联系人之间的短信息对话窗口,或者是与某一联系人之间的网络聊天对话窗口,历史对话信息是指对话框中记录的已经发生的对话信息。

[0078] 本步骤可以以要获取上屏候选队列时的时间作为基准,向前获取对话框中的历史对话信息。具体可以获取该对话框内预设历史时长内的历史对话信息,例如假设当前时间为7:00,则选择6:00至7:00之间的历史对话信息;或者,可以获取该对话框内预设信息条数的历史对话信息,例如当前输入文本框的前三条等。其中,不同的app对话环境可以设置不同的时间阈值或者信息条数阈值。

[0079] 步骤102,提取历史对话信息中的特征信息,特征信息包括关键词。

[0080] 该特征信息可以是关键词,也可以是关键词及关键词出现在历史对话信息中的时间,还可以包括从历史对话信息中提取的标识词。

[0081] 其中,可以预先设置关键词元词表,该关键词元词表是一个集合,其中包含可以作为关键词的词条。可以约定凡是出现在该关键词元词表中的词条都可以作为关键词,不在这个关键词元词表中的词条都不作为关键词。

[0082] 每条历史对话信息都会有时间记录,关键词在历史对话信息中出现的时间也就是该历史对话信息建立的时间,也即关键词在对话框中出现的时间。

[0083] 标识词可以是人名(如张三),地名(如五道口),或商业服务机构名称(如酒店名称)等便于识别的词条,还可以是验证码,例如大众点评发送的验证码信息。标识词还可以是应用名称或可以触发应用名称等,当提取出该类标识词时,可以同时触发到相应的应用去获取信息,例如去通讯录查找某人的联系电话等。在提取出该标识词后对该标识词进行分类存储。

[0084] 步骤103,确定预测信息,该预测信息包括关键词。

[0085] 该预测信息用于获取输入文本框内的上屏候选队列,在不同的条件下,预测信息不同,具体可以按照如图2所示的方法进行确定:

[0086] 步骤201,判断对话框的输入文本框内是否存在文本信息。

[0087] 若不存在文本信息,则执行步骤202,若存在文本信息,则进一步执行步骤203。

[0088] 步骤202,确定预测信息为关键词。

[0089] 步骤203,判断文本信息与关键词之间是否存在话题转移。

[0090] 具体的判断方法可以是判断文本信息命中的情景特征标签与关键词命中的情景特征标签是否存在重叠,若是,则不存在话题转移,若否,则存在话题转移。

[0091] 若存在话题转移,则执行步骤204,若不存在话题转移,则执行步骤205。

[0092] 步骤204,确定预测信息为文本信息。

[0093] 步骤205,确定预测信息为关键词和文本信息。

[0094] 在确定预测信息后即可进一步执行步骤104。

[0095] 步骤104,根据预测信息获得对话框中输入文本框内的上屏候选词队列。

[0096] 该获得上屏候选词队列的过程可以是根据预测信息的情景特征标签查询联想候选词库,也可以是根据预测信息查询相应的二元和/或三元语言模型的联想候选词库,还可以是在查询预先存储的标识词;也可以上述三类查询均执行,然后将获得的上屏候选词队列进行合并,确定最终的上屏候选词队列。具体请参见后续实施例的描述。

[0097] 步骤105,输出上屏候选词队列。

[0098] 本发明实施例通过充分挖掘对话框的历史对话信息,综合从中提取的特征信息来确定预测信息,并依据预测信息查询相应的联想候选词库获得了最终的上屏候选队列。该方法无论输入文本框内容是否为空仍然可以获得上屏候选队列,而且所获得的上屏候选队列显然比单单依据用户上一次上屏词条获得的上屏候选队列更加合理,更加准确。

[0099] 以下针对输入文本框内是否存在文本信息分别对该输入方法进行描述。

[0100] 参见图3,为本发明实施例另一种输入方法的流程图。

[0101] 本实施例以对话框的当前输入文本框为空为例进行说明,也即当前输入文本框内不存在文本信息。该方法可以包括:

[0102] 步骤301,获取对话框的历史对话信息。

[0103] 该步骤与前述步骤101类似。

[0104] 步骤302,提取历史对话信息中的特征信息,该特征信息包括关键词、关键词在历史对话信息中出现的时间及标识词。

[0105] 本实施例中,在提取出标识词后,还可以进一步根据标识词中的应用名称或可以触发应用的名词获取标识词的相关信息,例如根据标识词“地点”获取GPS地点,根据“现在”获取当前时间,根据“张三的电话”获取通讯录中张三的电话号码等。

[0106] 在提取出关键词及关键词在历史对话信息中出现的时间后,执行步骤303。在提取出标识词以及根据标识词获取的相关信息后,可以进一步对标识词及其相关信息执行步骤303'。

[0107] 步骤303,确定关键词命中的情景特征标签。

[0108] 其中,每个情景标签对应一个情景话题,同一关键词可能命中不同的情景标签。例如,关键词“吃饭”可能命中“聚会”情景,也可能命中“家庭活动”情景等。

[0109] 步骤303',对标识词以及根据标识词获取的相关信息按照情景特征标签进行分类存储。

[0110] 步骤304,对关键词命中的情景特征标签进行排序。

[0111] 具体的,可以根据命中情景特征标签的关键词的个数,以及关键词在历史对话信息中出现的时间对情景特征标签进行排序。具体的,可以采用综合打分的方式进行排序,例如,按照关键词在历史对话信息中出现的时间设置权重,出现时间越早权重越低,然后累计各情景特征标签中关键词的综合权重,权重越高的情景特征标签综合打分越高,排序越靠前。

[0112] 步骤305,按照情景特征标签的排序,获得输入文本框内的上屏候选词队列。

[0113] 按照情景特征标签的排序,依次查找具有相应情景特征标签的标识词及其相关信息,依次获得的标识词及其相关信息形成第一上屏候选词队列。

[0114] 按照情景特征标签的排序,根据关键词查找情景特征标签的联想候选词库,获得第二上屏候选词队列。其中,每个情景特征标签对应有联想候选词库。

[0115] 以关键词为左元,上屏候选词为右元,查询二元语言模型的联想候选词库,获得第三上屏候选词队列。

[0116] 对上述获得的三个上屏候选词队列分别设置不同的权重,其中,第一上屏候选词队列的权重最高,第三上屏候选词队列的权重最低。然后,对三个上屏候选词队列中的上屏候选词按照权重线性叠加的方式进行合并,最后按照权重由高到低对合并后的上屏候选词队列进行排序,获得最终的上屏候选词队列。

[0117] 本实施例通过充分挖掘对话窗的历史对话信息,综合从中提取的关键词、标识词等信息,并依据标识词、情景特征标签的联想候选词库及二元语言模型的联想候选词库获得了最终的上屏候选队列。该方法在输入文本框内容为空的情况下仍然可以获得上屏候选队列,而且所获得的上屏候选队列显然比单单依据用户上一次上屏词条获得的上屏候选队列更加合理,更加准确。

[0118] 参见图4,为本发明实施例另一种输入方法的流程图。

[0119] 本实施例以对话窗的当前输入文本框不为空为例进行说明,也即当前输入文本框内存在文本信息。该方法可以包括:

[0120] 步骤401,获取对话窗的历史对话信息。

[0121] 步骤402,提取历史对话信息中的特征信息,该特征信息包括关键词、关键词在历史对话信息中出现的时间及标识词。

[0122] 步骤403,确定关键词命中的情景特征标签。

[0123] 步骤403',对标识词以及根据标识词获取的相关信息按照情景特征标签进行分类存储。

[0124] 步骤401~步骤(403) 403'与前述实施例中的步骤301~步骤(303) 303'完全相同,此处不再赘述。

[0125] 步骤404,确定输入文本框中的文本信息命中的情景特征标签。

[0126] 具体的,可以通过查找预先设置的关键词元词表来提取文本信息中的关键词,然后确定提取的关键词命中的情景特征标签。

[0127] 步骤405,判断文本信息命中的情景特征标签与关键词命中的情景特征标签是否

有重叠。

[0128] 如果不存在重叠,则说明存在话题转移的现象,此时只针对输入文本框中的文本信息采用现有的预测方法获得上屏候选词队列即可。

[0129] 如果存在重叠,则说明话题没有转移,此时执行步骤406。

[0130] 步骤406,对文本信息命中的情景标签和关键词命中的情景特征标签进行排序。

[0131] 具体的,可以提取文本信息中的关键词,然后与历史对话信息中的关键词一起采用前述步骤304中的方法进行排序,最终,文本信息和关键词均命中的相同的情景标签排在最靠前。

[0132] 步骤407,按照情景特征标签的排序,获得输入文本框内的上屏候选词队列。

[0133] 按照情景特征标签的排序,依次查找具有相应情景特征标签的标识词及其相关信息,依次获得的标识词及其相关信息形成第一上屏候选词队列。

[0134] 按照情景特征标签的排序,根据输入对话框的文本信息中的关键词查找情景特征标签的联想候选词库,获得第二上屏候选词队列。其中,每个情景特征标签对应有联想候选词库。

[0135] 按照情景特征标签的排序,根据历史对话信息中的关键词查找情景特征标签的联想候选词库,获得第三上屏候选词队列。其中,每个情景特征标签对应有联想候选词库。

[0136] 以历史对话信息中的关键词为左元,以输入文本框中文本信息的关键词为中元,上屏候选词为右元,查询三元语言模型的联想候选词库,获得第四上屏候选词队列。

[0137] 以输入文本框中文本信息的关键词为左元,上屏候选词为右元,查询二元语言模型的联想候选词库,获得第五上屏候选词队列。

[0138] 以历史对话信息中的关键词为左元查询二元语言模型的联想候选词库,获得第六上屏候选词队列。

[0139] 对上述获得的六个上屏候选词队列分别设置不同的权重,第一、第二、第三、第四、第五、第六上屏候选词队列的权重依次由大到小设置。然后,对六个上屏候选词队列中的上屏候选词按照权重线性叠加的方式进行合并,最后按照权重由高到低对合并后的上屏候选词队列进行排序,获得最终的上屏候选词队列。

[0140] 本实施例通过充分挖掘对话窗的历史对话信息,综合从中提取的关键词、标识词等信息,并结合输入对话框中的文本信息,依据标识词、情景特征标签的联想候选词库及二元、三元语言模型的联想候选词库获得了最终的上屏候选队列。该方法在输入文本框中存在文本信息的情况下所获得的上屏候选队列显然比单单依据用户上一次上屏词条获得的上屏候选队列更加合理,更加准确。

[0141] 下面通过具体实例进行说明。

[0142] 例如,历史对话信息为“在哪儿吃饭?”,如果当前输入文本框为空,则提取该历史对话信息中的关键词“哪儿”、“吃饭”,该关键词命中的情景特征标签为地点和咨询标签,动作、活动标签、餐饮标签,如果历史对话信息中包含商家的标识词“五道口日昌”,则此时根据情景特征标签可以查询到该标识词“五道口日昌”,进而可以将“五道口日昌”作为上屏候选词提供给用户。如果没有查询到相关的标识词,则可以从GPS获取到用户当前的位置信息为“五道口”,进而搜索“五道口”附近的商家,其中“五道口日昌”距离用户当前位置最近,那么“五道口日昌”这个候选则会提供给用户。

[0143] 同样是历史对话信息为“在哪儿吃饭?”如果当前输入文本框的文本信息为“五道口”,则提取该历史对话信息中的关键词“哪儿”、“吃饭”,“哪”带有地点和咨询情景特征标签,“吃饭”带有动作、活动情景特征标签、餐饮情景特征标签,通过排序可以获得情景特征标签队列。当输入文本框内存在了“五道口”信息后,“五道口”被挖掘成为关键词,确定为地点情景特征标签,该地点情景特征标签与历史对话信息中的地点和咨询情景特征标签重叠,可以确定话题没有转移,通过查询所有以“五道口”为左元,右元带有“动作”“活动”“餐饮”情景特征标签的联想候选词库,即可获得上屏候选词队列,例如上屏候选词“日昌”。

[0144] 再例如,历史对话信息“在哪个会议室?”,提取关键词“哪个”、“会议室”,确定了“查询”情景特征标签和“地点”情景特征标签,“会议室”的标识词触发了对Outlook等日历软件查询动作,Outlook返回了一个即将发生的会议事件,并获取该标识词的相关信息“清华同方D座20层阿空加瓜”,在按照情景特征标签队列查询标识词及其相关信息时,即可获取该信息“清华同方D座20层阿空加瓜”作为上屏候选词加入上屏候选词队列。

[0145] 同样,标识词还可以触发对其他应用的查询,例如对通讯录的查询,获取某一的电话号码进行存储。

[0146] 需要说明的是,对于方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本发明实施例并不受所描述的动作顺序的限制,因为依据本发明实施例,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作并不一定是本发明实施例所必须的。

[0147] 参照图5,为本发明实施例一种输入装置的结构框图。

[0148] 该装置可以包括:

[0149] 信息获取单元501,用于获取对话框的历史对话信息。

[0150] 信息提取单元502,用于提取所述历史对话信息中的特征信息,所述特征信息包括关键词。

[0151] 信息确定单元503,用于确定预测信息,所述预测信息包括所述关键词。

[0152] 预测单元504,用于根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列。

[0153] 输出单元505,用于输出所述上屏候选词队列。

[0154] 本发明实施例通过上述单元充分挖掘对话框的历史对话信息,综合从中提取的特征信息来确定预测信息,并依据预测信息查找相应的联想候选词库获得了最终的上屏候选队列。该装置无论输入文本框内容是否为空仍然可以获得上屏候选队列,而且所获得的上屏候选队列显然比单单依据用户上一次上屏词条获得的上屏候选队列更加合理,更加准确。

[0155] 在本发明的另一实施例中,信息获取单元501,具体用于获取所述对话框预设历史时长内的历史对话信息;或者,获取所述对话框内预设信息条数的历史对话信息。

[0156] 在本发明的另一实施例中,如图6所示,信息确定单元503可以进一步包括:

[0157] 第一判断子单元601,用于判断所述对话框的输入文本框内是否存在文本信息。

[0158] 第一确定子单元602,用于当所述第一判断子单元601判断不存在文本信息时,确定所述预测信息为所述关键词。

[0159] 第二判断子单元603,用于当所述第一判断子单元601判断存在文本信息时,判断所述文本信息与所述关键词之间是否存在话题转移。

[0160] 第二确定子单元604,用于当所述第二判断子单元603判断不存在话题转移时,确定所述预测信息为所述关键词和所述文本信息。

[0161] 其中,第二判断子单元603,具体用于判断所述文本信息命中的情景特征标签与所述关键词命中的情景特征标签是否存在重叠,若是,则不存在话题转移,若否,则存在话题转移。

[0162] 在本发明的另一实施例中,如图7所示,预测单元504可以进一步包括:

[0163] 标签确定子单元701,用于确定所述预测信息命中的情景特征标签。

[0164] 排序子单元702,用于对所述情景特征标签进行排序。

[0165] 第一查找子单元703,用于按照所述情景特征标签的排序,根据所述预测信息在所述情景特征标签的联想候选词库中查找上屏候选词。

[0166] 队列获得子单元704,用于根据查找到的上屏候选词获得所述输入文本框内的上屏候选词队列。

[0167] 其中,排序子单元702,具体用于根据命中所述情景特征标签的所述预测信息的个数,以及所述预测信息在所述对话框中出现的时间对所述情景特征标签进行排序。

[0168] 另外,信息提取单元502提取的特征信息还可以包括标识词,该预测单元504还可以包括:

[0169] 存储子单元705,用于在所述队列获得子单元704获得所述输入文本框内的上屏候选词队列之前,对所述标识词按照情景特征标签进行分类存储;

[0170] 第二查找子单元706,用于按照所述情景特征标签的排序,查找具有相应情景特征标签的标识词,将查找到的标识词作为所述输入文本框内的上屏候选词。

[0171] 第三查找子单元707,用于在所述队列获得子单元704获得所述输入文本框内的上屏候选词队列之前,根据所述预测信息在语言模型的联想候选词库中查找上屏候选词;其中,当所述预测信息为所述关键词时,以所述关键词为左元查询,所述语言模型为二元语言模型;当所述预测信息为所述关键词及所述文本信息时,以所述关键词为左元,以所述文本信息为中元查询,所述语言模型为三元语言模型,或者,以所述文本信息为左元查询,所述语言模型为二元语言模型。

[0172] 在通过上述第一、第二、第三查找子单元查找到各自的上屏候选词后,可以按照一定的权重对获取到的上屏候选词排序获得上屏候选词队列。

[0173] 本实施例通过上述单元充分挖掘对话框的历史对话信息,综合从中提取的关键词、标识词等信息,并依据标识词、情景特征标签的联想候选词库及二元、三元语言模型的联想候选词库获得了最终的上屏候选队列。该装置无论在输入文本框内容是否为空的情况下仍然可以获得上屏候选队列,而且所获得的上屏候选队列显然比单单依据用户上一次上屏词条获得的上屏候选队列更加合理,更加准确。

[0174] 本发明实施例还提供了一种电子设备,包括存储器和处理器,所述存储器用于存储计算机指令或代码,所述处理器和所述存储器耦合,用于执行所述存储器中的计算机指令或代码,实现以下方法:

[0175] 获取对话框的历史对话信息;

[0176] 提取所述历史对话信息中的特征信息,所述特征信息包括关键词;

[0177] 确定预测信息,所述预测信息包括所述关键词;

[0178] 根据所述预测信息获得所述对话框中输入文本框内的上屏候选词队列;

[0179] 输出所述上屏候选词队列。

[0180] 对于装置实施例而言,由于其与方法实施例基本相似,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0181] 本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。

[0182] 本领域内的技术人员应明白,本发明实施例的实施例可提供为方法、装置、或计算机程序产品。因此,本发明实施例可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0183] 本发明实施例是参照根据本发明实施例的方法、终端设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理终端设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理终端设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0184] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理终端设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0185] 这些计算机程序指令也可装载到计算机或其他可编程数据处理终端设备上,使得在计算机或其他可编程终端设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程终端设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0186] 尽管已描述了本发明实施例的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例做出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本发明实施例范围的所有变更和修改。

[0187] 最后,还需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者终端设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者终端设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者终端设备中还存在另外的相同要素。

[0188] 以上对本发明所提供的一种输入方法、输入装置和一种电子设备,进行了详细介

绍,本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

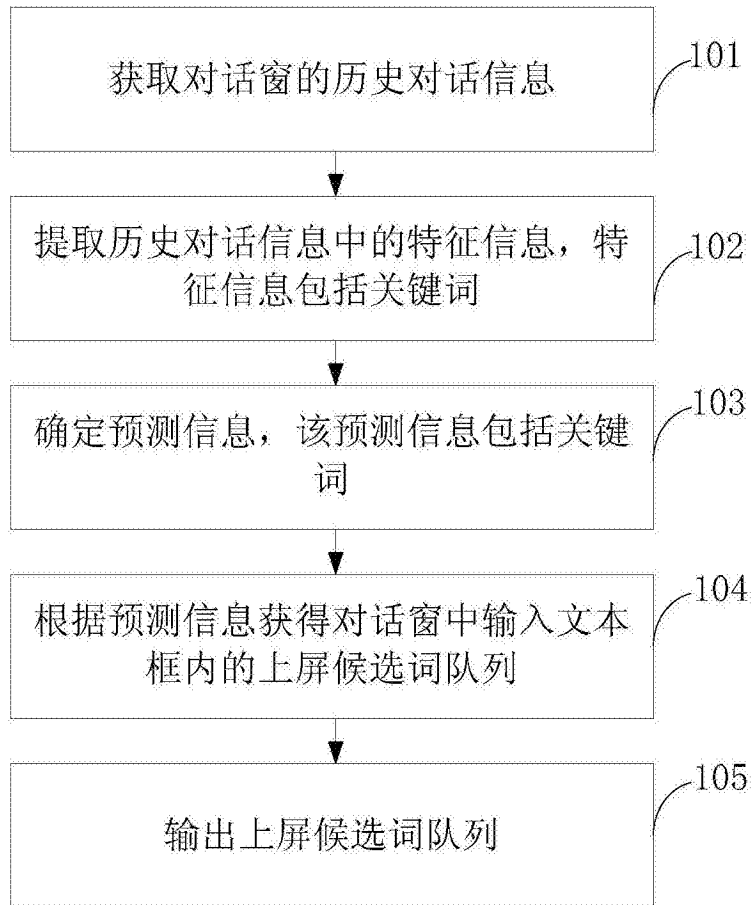


图1

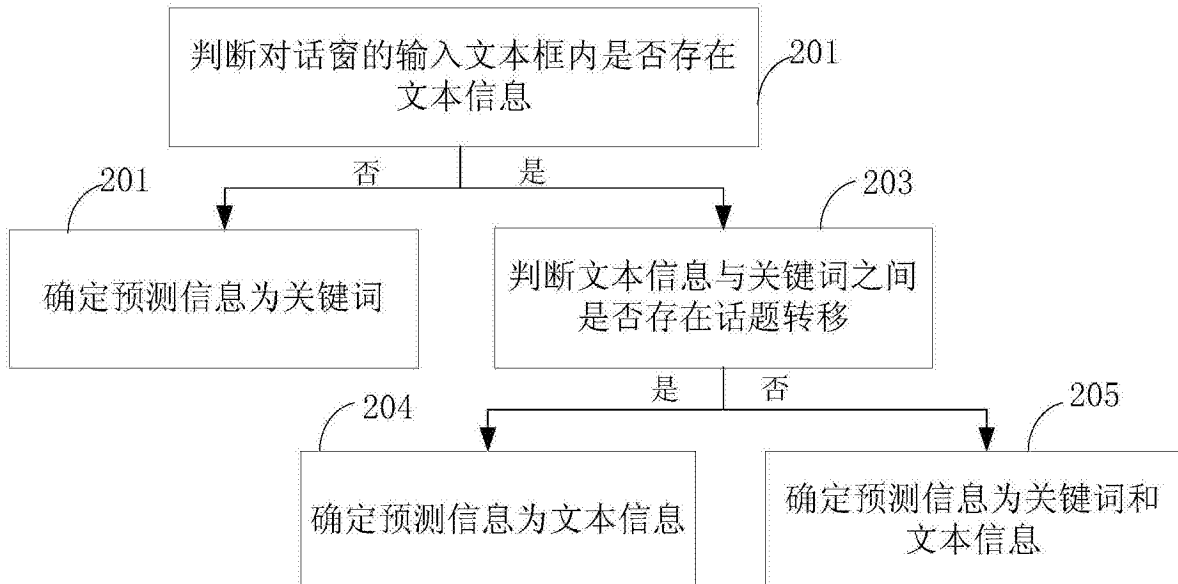


图2

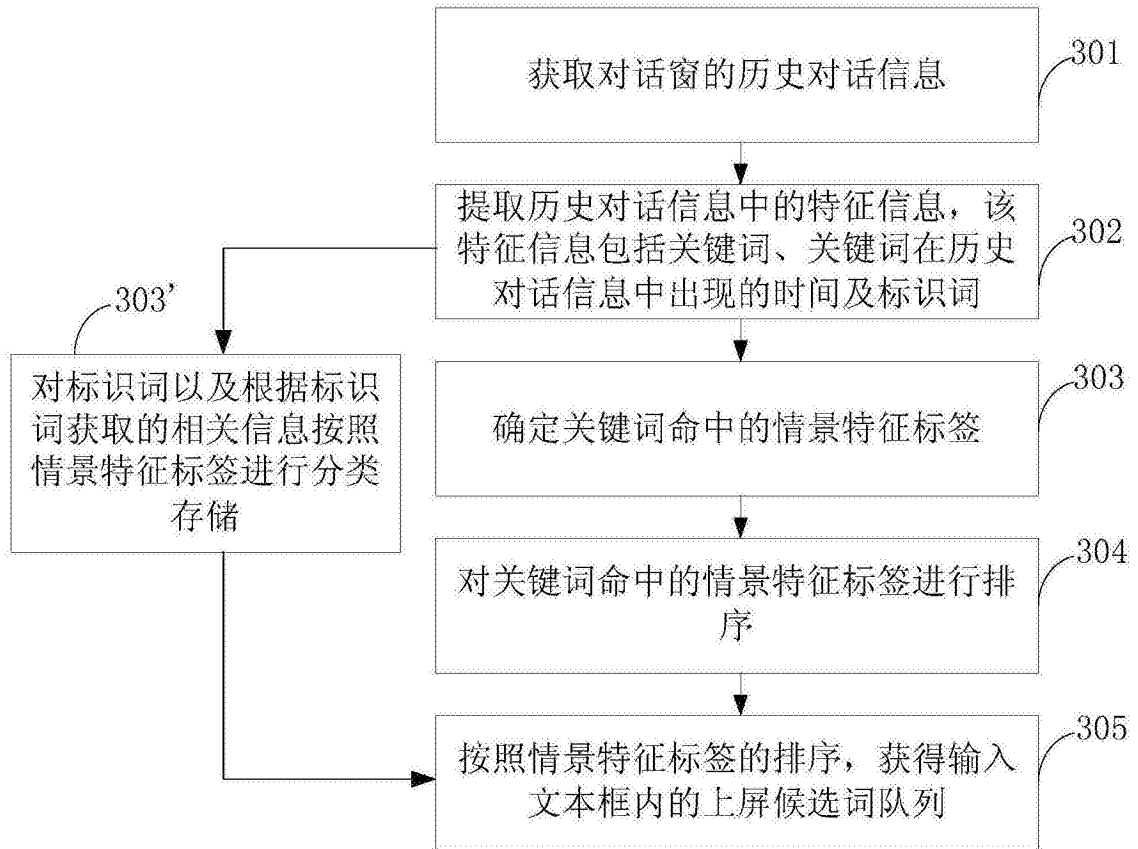


图3

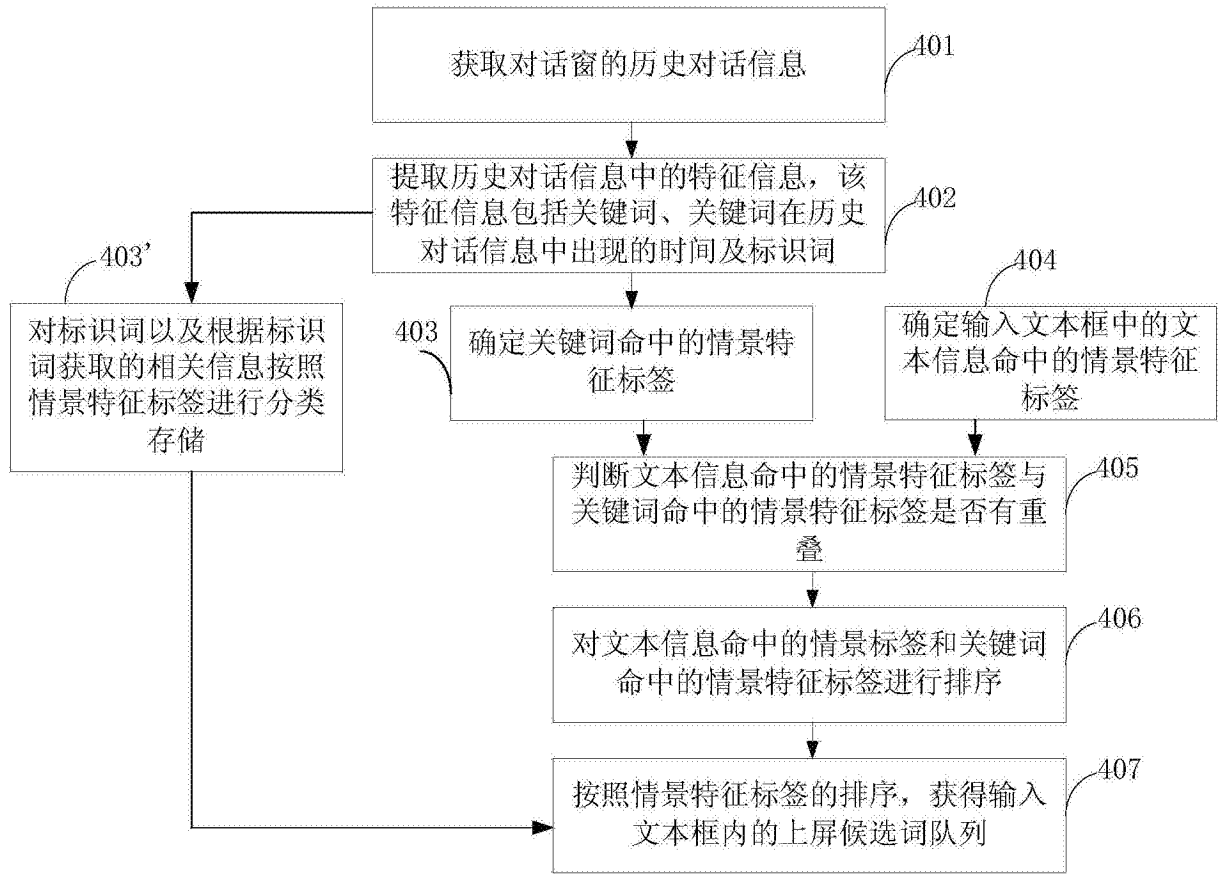


图4

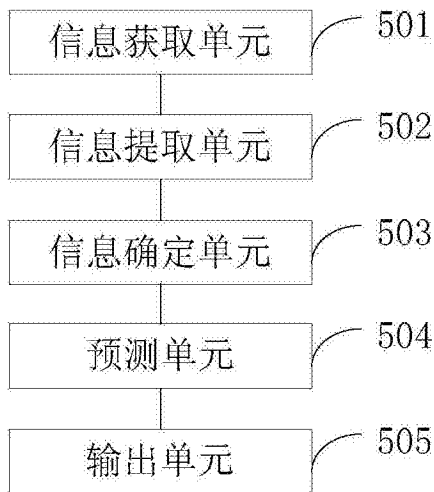


图5

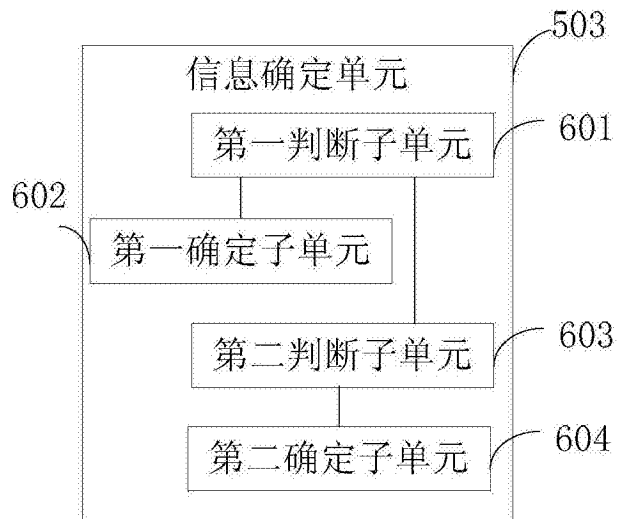


图6

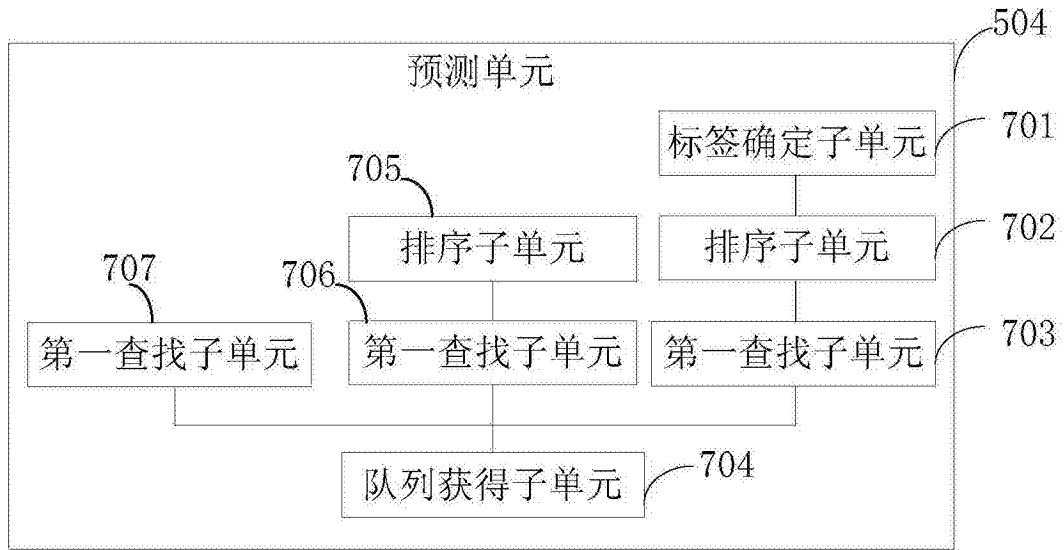


图7