



(12) 发明专利

(10) 授权公告号 CN 108280081 B

(45) 授权公告日 2023. 02. 24

(21) 申请号 201710009623.9

(22) 申请日 2017.01.06

(65) 同一申请的已公布的文献号

申请公布号 CN 108280081 A

(43) 申请公布日 2018.07.13

(73) 专利权人 百度在线网络技术(北京)有限公司

地址 100085 北京市海淀区上地十街10号  
百度大厦三层

(72) 发明人 姚欣洪 马飞超 黄飞 刘靖雅

(74) 专利代理机构 北京英赛嘉华知识产权代理  
有限责任公司 11204

专利代理师 王达佐 马晓亚

(51) Int. Cl.

G06F 16/9535 (2019.01)

(56) 对比文件

CN 104063454 A, 2014.09.24

CN 104268192 A, 2015.01.07

审查员 唐夏丽

权利要求书3页 说明书13页 附图7页

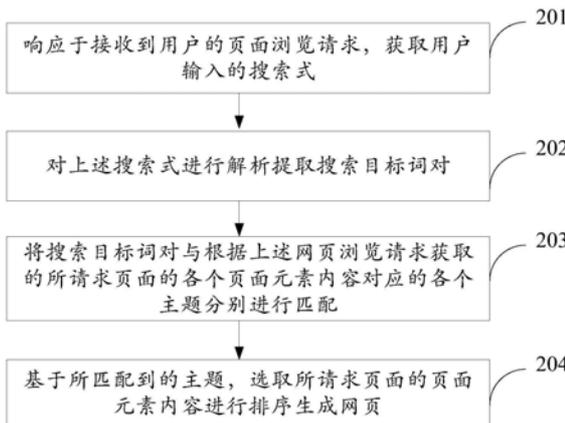
(54) 发明名称

生成网页的方法和装置

(57) 摘要

本申请公开了一种生成网页的方法和装置。该方法的一具体实施方式包括:响应于接收到用户的网页浏览请求,获取用户输入的搜索式;对该搜索式进行解析提取搜索目标词对,其中,搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;将搜索目标词对与根据网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配;基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页。该实施方式可以根据用户的搜索式动态生成新网页,从而提高信息推送的针对性。

200



1. 一种生成网页的方法,其特征在于,所述方法包括:

响应于接收到用户的网页浏览请求,获取所述用户输入的搜索式;其中,所述网页浏览请求包括网址;

通过切词和/或语义分析对所述搜索式进行解析提取搜索目标词对,其中,所述搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;

将所述网页浏览请求获取的所请求页面的各个页面元素内容区分不同的主题,将所述搜索目标词对与根据所述网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配;其中,所述所请求页面为所述网址对应的网页页面;

基于所匹配到的主题以及主题针对所述搜索目标词对的重要度,选取所请求页面的页面元素内容进行排序生成新网页;

其中,所述通过切词和/或语义分析对所述搜索式进行解析提取搜索目标词对,包括:

对所述搜索式进行解析提取至少一个候选搜索目标词对;

根据预先训练的对象重要度系数,从所解析出的至少一个候选搜索目标词对中选取第一候选搜索目标词对,其中,所述第一候选搜索目标词对包含所解析出的至少一个候选搜索目标词对中对象重要度系数最大的搜索对象关键词;所述对象重要度系数用于表示搜索对象关键词对用户所了解的搜索结果的影响程度;

根据预先训练的意图重要度系数,从所述第一候选搜索目标词对中选取第二候选搜索目标词对,其中,所述第二候选搜索目标词对包括所述第一候选搜索目标词对中意图重要度系数最大的搜索对象关键词;所述意图重要度系数用于表示各搜索意图关键词对用户所了解的搜索结果的影响程度;

将所述第二候选搜索目标词对确定为所述搜索目标词对。

2. 根据权利要求1所述的方法,其特征在于,所述对所述搜索式进行解析提取至少一个候选搜索目标词对包括:

将所述搜索式与预设对象关键词集合进行匹配,以确定候选搜索对象关键词;

将所述搜索式与预设意图关键词集合进行匹配,以确定候选搜索意图关键词;

基于对所确定的候选搜索对象关键词和候选搜索意图关键词,组合生成候选搜索目标词对。

3. 根据权利要求1所述的方法,其特征在于,所述根据预先训练的对象重要度系数,从所解析出的至少一个候选搜索目标词对中选取第一候选搜索目标词对包括:

确定各个候选搜索目标词对中的候选搜索对象关键词的所属对象类别;

根据预先训练的搜索对象关键词所属对象类别的对象类别重要度系数,从所解析出的至少一个候选搜索目标词对中选取所述第一候选搜索目标词对,其中,所述第一候选搜索目标词对包含所属对象类别的对象类别重要度系数最大的搜索对象关键词。

4. 根据权利要求1所述的方法,其特征在于,所述根据预先训练的意图重要度系数,从所述第一候选搜索目标词对中选取第二候选搜索目标词对包括:

确定各个候选搜索目标词对中的候选搜索意图关键词的所属意图类别;

根据预先训练的搜索意图关键词所属意图类别的意图类别重要度系数,从所述第一候选搜索目标词对中选取第二候选搜索目标词对,其中,所述第二候选搜索目标词对包含所属类别的意图类别重要度系数最大的搜索意图关键词。

5. 根据权利要求1所述的方法,其特征在于,所述对象重要度系数以及意图重要度系数通过以下方法获取:

获取一定数量的搜索式 and 用户根据搜索式所选择的主题作为样本;

从样本中的每个搜索式中提取至少一个对象关键词和至少一个意图关键词;

基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索对象关键词的对象相似度,训练各个搜索对象关键词针对各个主题的对象重要度系数;

基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索意图关键词的意图相似度,训练各个搜索意图关键词针对各个主题的意图重要度系数。

6. 根据权利要求1-5中任一所述的方法,其特征在于,所述将所述搜索目标词对与根据所述网页浏览请求获取所请求页面的页面内容对应的各个主题分别进行匹配包括:

根据所述网页浏览请求获取所请求页面的各个页面元素内容对应的各个主题;

计算各个主题与所述搜索目标词对的关联度;

基于各个关联度,确定与所述搜索目标词对相匹配的主题。

7. 根据权利要求6所述的方法,其特征在于,所述基于所匹配到的主题以及主题针对所述搜索目标词对的重要度,选取所请求页面的页面元素内容进行排序生成新网页包括:

按照与所述搜索目标词对的关联度由大到小的顺序,对各个所匹配到的主题所对应的页面元素内容进行排序,生成新网页。

8. 一种生成网页的装置,其特征在于,所述装置包括:

获取模块,配置用于响应于接收到用户的网页浏览请求,获取所述用户输入的搜索式;其中,所述网页浏览请求包括网址;

解析模块,配置用于通过切词和/或语义分析对所述搜索式进行解析提取搜索目标词对,其中,所述搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;匹配模块,配置用于对所述网页浏览请求获取的所请求页面的各个页面元素内容区分不同的主题,将所述搜索目标词对与根据所述网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配;其中,所述所请求页面为所述网址对应的网页页面;

生成模块,配置用于基于所匹配到的主题以及主题针对所述搜索目标词对的重要度,选取所请求页面的页面元素内容进行排序生成新网页;

其中,所述解析模块包括:

提取单元,配置用于对所述搜索式进行解析提取至少一个候选搜索目标词对;

第一选择单元,配置用于根据预先训练的对象重要度系数,从所解析出的至少一个候选搜索目标词对中选取第一候选搜索目标词对,其中,所述第一候选搜索目标词对包含所解析出的至少一个候选搜索目标词对中对象重要度系数最大的搜索对象关键词;所述对象重要度系数用于表示搜索对象关键词对用户所了解的搜索结果的影响程度;

第二选择单元,配置用于根据预先训练的意图重要度系数,从所述第一候选搜索目标词对中选取第二候选搜索目标词对,其中,所述第二候选搜索目标词对包括所述第一候选搜索目标词对中意图重要度系数最大的搜索对象关键词;所述意图重要度系数用于表示各搜索意图关键词对用户所了解的搜索结果的影响程度;

确定单元,配置用于将所述第二候选搜索目标词对确定为所述搜索目标词对。

9. 根据权利要求8所述的装置,其特征在于,所述提取单元进一步配置用于:

将所述搜索式与预设对象关键词集合进行匹配,以确定候选搜索对象关键词;  
将所述搜索式与预设意图关键词集合进行匹配,以确定候选搜索意图关键词;  
基于对所确定的候选搜索对象关键词和候选搜索意图关键词,组合生成候选搜索目标词对。

10. 根据权利要求8所述的装置,其特征在于,所述第一选择单元进一步配置用于:  
确定各个候选搜索目标词对中的候选搜索对象关键词的所属对象类别;  
根据预先训练的搜索对象关键词所属对象类别的对象类别重要度系数,从所解析出的至少一个候选搜索目标词对中选取所述第一候选搜索目标词对,其中,所述第一候选搜索目标词对包含所属对象类别的对象类别重要度系数最大的搜索对象关键词。

11. 根据权利要求8所述的装置,其特征在于,所述第二选择单元进一步配置用于:  
确定各个候选搜索目标词对中的候选搜索意图关键词的所属意图类别;  
根据预先训练的搜索意图关键词所属意图类别的意图类别重要度系数,从所述第一候选搜索目标词对中选取第二候选搜索目标词对,其中,所述第二候选搜索目标词对包含所属类别的意图类别重要度系数最大的搜索意图关键词。

12. 根据权利要求8所述的装置,其特征在于,所述装置还包括训练模块,配置用于通过以下方法获取对象重要度系数以及意图重要度系数:

获取一定数量的搜索式 and 用户根据搜索式所选择的主题作为样本;  
从样本中的每个搜索式中提取至少一个对象关键词和至少一个意图关键词;  
基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索对象关键词的对象相似度,训练各个搜索对象关键词针对各个主题的对象重要度系数;  
基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索意图关键词的意图相似度,训练各个搜索意图关键词针对各个主题的意图重要度系数。

13. 根据权利要求8-12中任一所述的装置,其特征在于,所述匹配模块包括:  
获取单元,配置用于根据所述网页浏览请求获取所请求页面的各个页面元素内容对应的各个主题;  
计算单元,配置用于计算各个主题与所述搜索目标词对的关联度;  
匹配单元,配置用于基于各个关联度,确定与所述搜索目标词对相匹配的主题。

14. 根据权利要求13所述的装置,其特征在于,所述生成模块进一步配置用于:  
按照与所述搜索目标词对的关联度由大到小的顺序,对各个所匹配到的主题所对应的页面元素内容进行排序,生成新网页。

15. 一种计算设备,包括:  
一个或多个处理器;  
存储器,用于存储一个或多个程序;  
当所述一个或多个程序被所述一个或多个处理器执行时,使得所述处理器实现如权利要求1到7中任一所述的方法。

## 生成网页的方法和装置

### 技术领域

[0001] 本申请涉及计算机技术领域,具体涉及互联网技术领域,尤其涉及生成网页的方法和装置。

### 背景技术

[0002] 用户通过互联网在终端搜索一些信息时,往往带有明显的搜索意图,例如用户在搜索医疗方面的信息时,输入搜索式“脑膜炎的症状”,意图是想获取“脑膜炎”这种精确疾病的相关信息,而当用户搜索“头痛是怎么回事”时,则主要是想获取引起头痛相关的原因以及可能的病症。目前,用户根据搜索结果点击进入的搜索结果页面往往是固定形式的页面,而当搜索结果页面包含多种信息时,用户在进入搜索结果页面后无法快速获取所需要的信息。因此,现有的搜索结果页面由于网页内容较固定,导致信息推送的针对性不高。

### 发明内容

[0003] 本申请的目的在于提出一种改进的生成网页的方法和装置,来解决以上背景技术部分提到的技术问题。

[0004] 一方面,本申请提供了一种生成网页的方法,该方法包括:响应于接收到用户的网页浏览请求,获取用户输入的搜索式;对该搜索式进行解析提取搜索目标词对,其中,搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;将搜索目标词对与根据网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配;基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页。

[0005] 在一些实施例中,对上述搜索式进行解析提取搜索目标词对包括:对上述搜索式进行解析提取至少一个候选搜索目标词对,其中,每个搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;根据预先训练的对象重要度系数,从所解析出的至少一个候选搜索目标词对中选取出第一候选搜索目标词对,其中,第一候选搜索目标词对包含所解析出的至少一个候选搜索目标词对中对象重要度系数最大的搜索对象关键词;根据预先训练的意图重要度系数,从第一候选搜索目标词对中选取出第二候选搜索目标词对,其中,第二候选搜索目标词对包括第一候选搜索目标词对中意图重要度系数最大的搜索对象关键词;将第二候选搜索目标词对确定为搜索目标词对。

[0006] 在一些实施例中,对上述搜索式进行解析提取至少一个候选搜索目标词对包括:将上述搜索式与预设对象关键词集合进行匹配,以确定候选搜索对象关键词;将上述搜索式与预设意图关键词集合进行匹配,以确定候选搜索意图关键词;基于对所确定的候选搜索对象关键词和候选搜索意图关键词,组合生成候选搜索目标词对。

[0007] 在一些实施例中,根据预先训练的对象重要度系数,从所解析出的至少一个候选搜索目标词对中选取出第一候选搜索目标词对包括:确定各个候选搜索目标词对中的候选搜索对象关键词的所属对象类别;根据预先训练的搜索对象关键词所属对象类别的对象类别重要度系数,从所解析出的至少一个候选搜索目标词对中选取出第一候选搜索目标词

对,其中,第一候选搜索目标词对包含所属对象类别的对象类别重要度系数最大的搜索对象关键词。

[0008] 在一些实施例中,根据预先训练的意图重要度系数,从第一候选搜索目标词对中选取第二候选搜索目标词对包括:确定各个候选搜索目标词对中的候选搜索意图关键词的所属意图类别;根据预先训练的搜索意图关键词所属意图类别的意图类别重要度系数,从第一候选搜索目标词对中选取第二候选搜索目标词对,其中,第二候选搜索目标词对包含所属类别的意图类别重要度系数最大的搜索意图关键词。

[0009] 在一些实施例中,对象重要度系数以及意图重要度系数通过以下方法获取:获取一定数量的搜索式 and 用户根据搜索式所选择的主题作为样本;从样本中的每个搜索式中提取至少一个搜索对象关键词和至少一个搜索意图关键词;基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索对象关键词的对象相似度,训练各个搜索对象关键词针对各个主题的对象重要度系数;基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索意图关键词的意图相似度,训练各个搜索意图关键词针对各个主题的意图重要度系数。

[0010] 在一些实施例中,将搜索目标词对与根据网页浏览请求获取所请求页面的页面内容对应的各个主题分别进行匹配包括:根据网页浏览请求获取所请求页面的各个页面元素内容对应的各个主题;计算各个主题与搜索目标词对的关联度;基于各个关联度,确定与搜索目标词对相匹配的主题。

[0011] 在一些实施例中,基于所匹配到的主题对所请求页面的页面内容进行排序生成网页包括:按照与搜索目标词对的关联度由大到小的顺序,对各个所匹配到的主题所对应的页面元素内容进行排序,生成网页。

[0012] 第二方面,本申请提供了一种生成网页的装置,该装置包括:获取模块,配置用于响应于接收到用户的网页浏览请求,获取用户输入的搜索式;解析模块,配置用于对该搜索式进行解析提取搜索目标词对,其中,搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;匹配模块,配置用于将搜索目标词对与根据网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配;生成模块,配置用于基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页。

[0013] 在一些实施例中,解析模块包括:提取单元,配置用于对上述搜索式进行解析提取至少一个候选搜索目标词对,其中,每个搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;第一选择单元,配置用于根据预先训练的对象重要度系数,从所解析出的至少一个候选搜索目标词对中选取第一候选搜索目标词对,其中,第一候选搜索目标词对包含所解析出的至少一个候选搜索目标词对中对象重要度系数最大的搜索对象关键词;第二选择单元,配置用于根据预先训练的意图重要度系数从第一候选搜索目标词对中选取第二候选搜索目标词对,其中,第二候选搜索目标词对包括第一候选搜索目标词对中意图重要度系数最大的搜索对象关键词;确定单元,配置用于将第二候选搜索目标词对确定为搜索目标词对。

[0014] 在一些实施例中,提取单元进一步配置用于:将上述搜索式与预设对象关键词集合进行匹配,以确定候选搜索对象关键词;将上述搜索式与预设意图关键词集合进行匹配,以确定候选搜索意图关键词;基于对所确定的候选搜索对象关键词和候选搜索意图关键词

词,组合生成候选搜索目标词对。

[0015] 在一些实施例中,第一选择单元进一步配置用于:确定各个候选搜索目标词对中的候选搜索对象关键词的所属对象类别;根据预先训练的搜索对象关键词所属对象类别的对象类别重要度系数,从所解析出的至少一个候选搜索目标词对中选取第一候选搜索目标词对,其中,第一候选搜索目标词对包含所属对象类别的对象类别重要度系数最大的搜索对象关键词。

[0016] 在一些实施例中,第二选择单元进一步配置用于:确定各个候选搜索目标词对中的候选搜索意图关键词的所属意图类别;根据预先训练的搜索意图关键词所属意图类别的意图类别重要度系数,从第一候选搜索目标词对中选取第二候选搜索目标词对,其中,第二候选搜索目标词对包含所属类别的意图类别重要度系数最大的搜索意图关键词。

[0017] 在一些实施例中,上述装置还包括训练模块,配置用于通过以下方法获取对象重要度系数以及意图重要度系数:获取一定数量的搜索式 and 用户根据搜索式所选择的主题作为样本;从样本中的每个搜索式中提取至少一个搜索对象关键词和至少一个搜索意图关键词;基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索对象关键词的对象相似度,训练各个搜索对象关键词针对各个主题的对象重要度系数;基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索意图关键词的意图相似度,训练各个搜索意图关键词针对各个主题的意图重要度系数。

[0018] 在一些实施例中,匹配模块包括:获取单元,配置用于根据网页浏览请求获取所请求页面的各个页面元素内容对应的各个主题;计算单元,配置用于计算各个主题与搜索目标词对的关联度;匹配单元,配置用于基于各个关联度,确定与搜索目标词对相匹配的主题。

[0019] 在一些实施例中,生成模块进一步配置用于:按照与搜索目标词对的关联度由大到小的顺序,对各个所匹配到的主题所对应的页面元素内容进行排序,生成网页。

[0020] 第三方面,本申请还提供了一种计算设备,包括:一个或多个处理器;存储器,用于存储一个或多个程序;当该一个或多个程序被上述一个或多个处理器执行时,使得上述处理器本申请提供的任一生成网页的方法。

[0021] 本申请提供的生成网页的方法和装置,响应于接收到用户的网页浏览请求,获取用户输入的搜索式,接着对该搜索式进行解析提取至少一个搜索目标词对,其中,每个搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词,然后将搜索目标词对与根据网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配,并基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页,由于根据用户的搜索式生成新网页,使网页内容动态生成,从而提高信息推送的针对性。

## 附图说明

[0022] 通过阅读参照以下附图所作的对非限制性实施例的详细描述,本申请的其它特征、目的和优点将会变得更明显:

[0023] 图1示出了可以应用本申请实施例的示例性系统架构;

[0024] 图2是根据本申请的生成网页的方法的一个实施例的流程图;

[0025] 图3a、图3b、图3c是根据本申请的生成网页的方法的一个应用场景的示意图;

- [0026] 图4是根据本申请的生成网页的方法的另一个实施例的流程图；
- [0027] 图5是根据本申请的生成网页的装置的一个实施例的结构示意图；
- [0028] 图6是适于用来实现本申请实施例的电子设备的计算机系统的结构示意图。

### 具体实施方式

[0029] 下面结合附图和实施例对本申请作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释相关发明，而非对该发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与有关发明相关的部分。

[0030] 需要说明的是，在不冲突的情况下，本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

[0031] 图1示出了可以应用本申请实施例的示例性系统架构100。

[0032] 如图1所示，系统架构100可以包括终端设备101、102、网络103和服务器104。网络103用以在终端设备101、102和服务器104之间提供通信链路的介质。网络103可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

[0033] 终端设备101、102可以通过网络103与服务器104交互，以接收或发送消息等。终端设备101、102上可以安装有各种通讯客户端应用，例如浏览器应用、搜索类应用、地图类应用、购物类应用、理财类应用、社交平台应用、邮箱客户端、即时通信工具等等。

[0034] 终端设备101、102可以是支持浏览器应用、搜索类应用等等安装于其上的各种电子设备，包括但不限于智能手机、智能手表、平板电脑、个人数字助理、电子书阅读器、MP3播放器(Moving Picture Experts Group Audio Layer III, 动态影像专家压缩标准音频层面3)、MP4(Moving Picture Experts Group Audio Layer IV, 动态影像专家压缩标准音频层面4)播放器、膝上型便携计算机和台式计算机等等。

[0035] 服务器104可以是提供各种服务的服务器。例如服务器104可以是对终端设备101、102上所显示网页提供支持的后台网页服务器等。服务器可以对接收到的数据进行存储、生成等处理，并将处理结果反馈给终端设备。

[0036] 需要说明的是，本申请实施例所提供的生成网页的方法一般通过服务器104执行，但不排除可以通过终端设备101、102执行的可能。相应地，本申请实施例所提供的生成网页的装置一般设于服务器104中，但不排除可以设于终端设备101、102中的可能。

[0037] 应该理解，图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器。

[0038] 请参考图2，其示出了生成网页的方法的一个实施例的流程200。该生成网页的方法，包括以下步骤：

[0039] 步骤201，响应于接收到用户的页面浏览请求，获取用户输入搜索式。

[0040] 在本实施例中，生成网页的方法运行于其上的电子设备(例如图1所示的服务器104)可以通过有线连接方式或者无线连接方式从用户利用其进行网页浏览的终端接收网页浏览请求。其中，上述无线连接方式可以包括但不限于3G/4G连接、WiFi连接、蓝牙连接、WiMAX连接、Zigbee连接、UWB(ultra wideband)连接、以及其他现在已知或将来开发的无线连接方式。通常，用户利用终端上安装的网页浏览器来搜索和浏览网页，这时，用户可以通过输入搜索式、输入网址或者点击网页浏览器中呈现的网页中的链接来向服务器发起网页

浏览请求以获得网页。在本实施例中,上述网页可以包括html格式、xhtml格式、asp格式、php格式、jsp格式、shtml格式、nsp格式、xml格式的网页或者其它未来将开发的格式的网页(只要这种格式的网页文件可以用浏览器打开并浏览其包含的图片、动画、文字等内容)。当用户通过输入搜索式或者点击由搜索式得到的搜索结果页面中的超链接发起网页浏览请求时,上述电子设备可以获取用户输入的搜索式。这里,搜索式可以是用户进行信息搜索时所输入的搜索关键词、搜索语句、检索式(可以包括词、逻辑运算符等)等等,例如“腹痛”、“××医院肛肠科怎么样”、“腹痛恶心”……。

[0041] 步骤202,对上述搜索式进行解析提取搜索目标词对。

[0042] 在本实施例中,上述运行生成网页的方法的电子设备可以接着对上述搜索式进行识别和分析,从而提取至少一个搜索目标词对。这里,搜索目标词对可以包括一对用于描述用户要通过搜索了解的信息的关键词,一个搜索对象关键词与一个搜索意图关键词。其中,搜索对象关键词可以用于表示所搜索的内容主体,搜索意图关键词可以用于表示用户针对所搜索的内容主体的想了解的信息,如“××医院怎么样”可以包括搜索对象关键词“××医院”,以及搜索意图关键词“怎么样”。

[0043] 上述电子设备可以通过切词方式和/或语义分析方式等对上述搜索式进行分析。例如,对于搜索式“××医院怎么样”,电子设备可以通过切词方法得到搜索对象关键词“××医院”、搜索意图关键词“怎么样”,也可以通过语义分析确定搜索主体“××医院”为医院名称,“怎么样”确定用户针对所搜索的内容主体的想了解的内容为简介、评论等,如此时电子设备可以确定搜索对象关键词为“医院”、搜索意图关键词为“简介”或“评论”等。实践中,电子设备还可以预先对一定数量的自然语言描述信息,通过机器学习方法(如最大期望EM算法之类的分类算法等)确定各个搜索对象关键词及搜索意图关键词对应的词汇集合,例如搜索意图关键词“简介”对应的词汇集合可以包括词汇“怎么样”、“如何”、“好不好”等等,当搜索式出现搜索意图关键词“简介”对应的词汇集合中的任一词汇时,可以确定搜索意图关键词为“简介”。在一些实现中,上述搜索式可能只包含一个关键词,例如搜索式为“肚子疼”,此时,电子设备例如可以根据该搜索式确定搜索对象关键词

[0044] 步骤203,将搜索目标词对与根据上述网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配。

[0045] 在本实施例中,上述电子设备可以将页面上的各个页面元素内容根据主题进行区分和存储,响应于接收到用户的网页浏览请求,上述电子设备还可以根据上述网页浏览请求获取所请求的页面中各个页面元素内容所对应的各个主题,并在步骤202提取搜索目标词对之后,将各个主题分别与上述搜索目标词对进行匹配,以确定与搜索目标词对最相关的一个或较相关的多个主题。

[0046] 可以理解,网页浏览请求可以包括用户期望浏览的网页的地址,即网址。实践中,网址一般由统一资源定位符(Uniform Resource Locator,URL)来表示。电子设备首先可以根据上述网址获取所对应的网页页面。本实施例中,所请求页面的页面元素内容可以按照主题进行区分和存储,例如描述一种药的页面,其可以包括该药的药品说明、生产企业、所针对的疾病的病理描述等等多个主题。电子设备可以预先通过统计分析方式和/或语义分析方式等对所请求的页面中各个页面元素内容进行处理,以区分不同主题。具体地,例如电子设备可以根据页面所属领域预先确定好主题,再计算页面上各个页面元素内容与预先确

定的主题的相关性,确定相关页面元素内容属于相应主题;再例如电子设备可以对上述网页的各个页面元素内容中存在的各个词语的出现频率进行统计和排序,选取出现频率排序靠前的一个词语作为相应页面元素内容的关键词,将根据jaccard相似度之类的相似度算法计算的关键词相似度高于预设相似度阈值的两个或两个以上页面元素内容归入同一主题。显然,一个主题可以对应页面内容的部分或全部页面元素,相应地,电子设备获取的网页页面的页面内容可以对应一个或多个主题。

[0047] 电子设备可以将步骤202中提取的搜索目标词对与所获取的一个或多个主题进行匹配。匹配时,电子设备可以将主题与搜索目标词对进行匹配,也可以将主题对应的页面元素内容与搜索目标词对进行匹配,本申请对此不做限定。实践中,以将主题与搜索目标词对进行匹配为例,电子设备可以预先存储有主题与搜索目标词对的对应关系,电子设备可以根据该对应关系进行匹配。例如主题“治疗方法”可以对应搜索目标词对“精确疾病问诊”等等,当用户的搜索式为“胃溃疡怎么办”时,电子设备可以提取关键词“胃溃疡”和“怎么办”,“胃溃疡”可以对应搜索对象关键词“精确疾病”,“怎么办”可以对应搜索意图关键词“问诊”,进一步地,“精确疾病问诊”与主题“治疗方法”相匹配。在一些实现中,一个主题可以与多个搜索目标词对具有对应关系,一个搜索目标词对也可以与多个主题具有对应关系,则匹配时,一个搜索目标词对可以匹配到多个主题,例如搜索目标词对“精确疾病问诊”可以对应主题“治疗方法”和“治疗药物”等等。此时,这多个主题还可以具有不同的重要度,例如对于搜索目标词对“精确疾病问诊”,主题“治疗方法”的重要度大于“治疗药物”。可选地,该重要度的量值可以通过重要度系数来表示。

[0048] 步骤204,基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页。

[0049] 在本实施例中,生成网页的方法运行于电子设备接着可以选取所匹配到的主题对应的页面元素内容,根据匹配结果对这些页面元素内容进行排序,生成网页以推送给用户。例如,当只匹配到一个主题时,电子设备可以将这一个主题所对应的页面元素内容生成网页推送给用户,当匹配到多个主题时,电子设备可以根据这多个主题针对上述搜索目标词对的重要度由高到低的顺序,对相应的页面元素内容进行排序生成网页。

[0050] 在本实施例的一些可选实现方式中,电子设备可以在根据网页浏览请求获取所请求页面的各个页面元素内容对应的各个主题后,先计算各个主题与搜索目标词对的关联度。电子设备接着可以基于各个关联度,确定与搜索目标词对相匹配的主题。例如按照预设的关联度阈值或者个数阈值,选取出与搜索目标词对的关联度大于预设的关联度阈值的主题,或者与搜索目标词对的关联度最高的该个数阈值数目的主题作为与搜索目标词对相匹配的主题。其中,关联度可以用于表示主题所对应的页面元素内容与搜索目标词对的关联程度,例如可以是上述主题针对搜索目标词对的重要度的量值。上述电子设备可以通过诸如jaccard相似度、余弦相似度之类的相似度方法计算搜索目标词对与各个主题的关联度,也可以通过自然语言的语义分析方法计算搜索目标词对与各个主题的关联度,本申请对此不做限定。以jaccard相似度为例,电子设备可以先提取某个主题对应的页面元素内容的关键词生成主题关键词集合,并根据搜索目标词对中每个词的同义词和/或近义词(例如搜索对象关键词根据同义词,搜索意图关键词根据同义词和近义词等)进行扩展生成搜索目标关键词集合,计算这两个关键词集合的相似度作为搜索目标词对与该主题的关联度,如:两

个关键词集合的jaccard相似度=两个关键词集合的共有词数量/两个关键词集合的词数量之和。

[0051] 作为一个应用场景,本实施例的生成网页的方法例如可以应用于为移动终端运行的浏览器应用提供支持的后台服务器。一般的,用户可以通过终端上运行的浏览器应用进行搜索或点击超链接等向服务器发起页面请求。请参考图3a-图3c,以下结合附图3a-图3c对本实施例的生成网页的方法的该应用场景进行详细说明。

[0052] 如图3a所示,当用户通过终端300上运行的浏览器应用进行搜索时,在搜索引擎301中输入搜索式“拉肚子怎么办”,搜索引擎按照预设的搜索规则(如广告主的竞价排序规则、内容关联度排序规则等等各种可能的搜索规则,这不属于本申请的发明内容,在此不再赘述),展示搜索结果页面302,并在搜索结果页面302上展示各页面3021、3022……等的简介及超链接。假设用户点击了页面3021的超链接,则终端向后台服务器发送页面3021的网页浏览请求。此时,后台服务器响应于接收到该用户对页面3021的网页浏览请求,可以获取该用户输入的搜索式“拉肚子怎么办”,接着,后台服务器可以对该搜索式进行解析以提取搜索目标词对,如所提取的搜索目标词对包括搜索对象关键词“拉肚子”和搜索意图关键词“问诊”。

[0053] 同时,响应于接收到该用户对页面3021的网页浏览请求,上述后台服务器还可以获取页面3021的页面元素内容对应的各个主题。如图3b所示,页面3021的页面元素内容可以预先通过后台服务器的处理并按照主题30211“原因”、30212“药物”、30213“治疗方法”……进行存储。

[0054] 之后,后台服务器可以将上述搜索目标词对“拉肚子”和“问诊”与图3b所示的各个主题30211、30212……分别进行匹配。例如匹配到主题30213“治疗方法”和主题30212“药物”,则后台服务器可以生成如图3c所示的页面303发送至终端300进行显示。如图3c所示,页面303只包括主题30213“治疗方法”和主题30212“药物”对应的页面元素内容。可选地,页面303可以为广告主的落地页,在页面303中药物30212部分的药物介绍可以是超链接形式,用户可以通过点击超链接进入相应购买页面。

[0055] 在本实施例中,由于根据用户的搜索式提取搜索目标词对,进一步匹配相关主题动态生成新网页,从而提高信息推送的针对性。

[0056] 请参考图4,其示出了生成网页的方法的另一个实施例的流程400,流程400包括以下步骤:

[0057] 步骤401,响应于接收到用户的页面浏览请求,获取用户输入的搜索式。

[0058] 在本实施例中,在本实施例中,生成网页的方法运行于其上的电子设备(例如图1所示的服务器104)可以通过有线连接方式或者无线连接方式从用户利用其进行网页浏览的终端接收网页浏览请求。其中,当用户通过输入搜索式或者点击由搜索式得到的搜索结果页面中的超链接发起网页浏览请求时,上述电子设备可以获取用户输入的搜索式。这里,搜索式可以是用户进行信息搜索时所输入的搜索关键词、搜索语句、检索式(可以包括词、逻辑运算符等)等等。

[0059] 步骤402,对上述搜索式进行解析提取至少一个候选搜索目标词对。

[0060] 在本实施例中,基于步骤401中获取的搜索式,上述电子设备可以利用各种分析手段对该搜索式的内容进行分析,从而提取出一个或多个候选搜索目标词对。例如,电子设备

可以通过统计方法对包括词和逻辑运算符的检索式进行分析,或者通过切词方法对搜索语句进行分析等等。其中,一个候选搜索目标词对可以包括一个候选搜索对象关键词和一个候选搜索意图关键词。

[0061] 在本实施例的一些可选实现方式中,电子设备可以将上述搜索式与预设对象关键词集合中的各个词进行匹配,以确定搜索对象关键词。例如,当在预设对象关键词集合中匹配到至少一个词时,根据所匹配到的词确定候选搜索对象关键词,如将匹配到的词确定为搜索对象关键词。同样地,电子设备可以将上述搜索式与预设意图关键词集合中的各个词进行匹配,以确定候选搜索意图关键词。例如,当在预设意图关键词集合中匹配到至少一个词时,根据所匹配到的词确定候选搜索意图关键词。然后,上述电子设备可以基于对所确定的候选搜索对象关键词和候选搜索意图关键词进行组合,生成候选搜索目标词对。作为示例,假设搜索式为“××医院对孩子拉肚子的治疗效果怎么样”,根据自然语言模型选择最优切词序列得到“××医院”、“对”、“孩子”、“拉肚子”、“的”、“治疗”、“效果”、“怎么样”(在一些实现中还可以过滤掉常用词和/或虚词“对”、“的”)。假设在预设对象关键词集合中匹配到“××医院”、“拉肚子”,电子设备可以将它们作为候选搜索对象关键词;假设在预设意图关键词集合中匹配到“治疗”、“怎么样”,电子设备可以将它们作为候选搜索意图关键词;进一步地,组合得到的候选搜索目标词对可以包括“××医院治疗”、“××医院怎么样”、“拉肚子治疗”、“拉肚子怎么样”。

[0062] 在一些实现中,搜索对象可以按照类别进行分类,相应地,预设对象关键词集合可以有多个,每个对象关键词集合对应一个搜索对象的对象类别,如对象类别“泛症状”(表示可能为多种疾病引起的症状)可以对应对象关键词集合{拉肚子头痛发烧……},对象类别“精确疾病”可以对应对象关键词集合{胃溃疡咽炎鼻炎……},此时,在一个对象关键词集合中匹配到相应词后,电子设备也可以将该词对应的对象类别作为候选搜索对象关键词。

[0063] 同理,在另一些实现中,预设意图关键词集合也可以对应一个搜索意图的意图类别,如意图类别“问诊”可以对应意图关键词集合{怎么办治疗吃什么药……}。其中,预设搜索对象关键词集合和搜索意图关键词集合可以根据大量的自言语言描述信息通过统计或机器学习方法分类获得,例如本申请的生成网页的方法应用于医疗领域的网页生成中时,电子设备可以预先抓取一定数量的医疗问答数据和/或医疗广告页面的转化数据训练分类模型。

[0064] 步骤403,根据预先训练的对象重要度系数,从所解析出的至少一个候选搜索目标词对中选取出第一候选搜索目标词对。

[0065] 在本实施例中,上述电子设备可以对各搜索对象关键词预先训练有对象重要度系数,在该步骤中,电子设备可以根据预先训练的对象重要度系数,对步骤402所提取的至少一个候选搜索目标词对中的各候选搜索对象关键词的对象重要度系数进行对比,并选取出对象重要度系数最大的候选搜索对象关键词,将包含该候选搜索对象关键词的候选搜索目标词对作为第一候选搜索目标词对。

[0066] 其中,上述对象重要度系数可以用于表示各搜索对象关键词对用户所了解的搜索结果的影响程度。例如前述示例中,对于搜索式“××医院对孩子拉肚子的治疗效果怎么样”,候选搜索目标词对可以包括“××医院治疗”、“××医院怎么样”、“拉肚子治疗”、“拉肚子怎么样”。其中包括了两个候选搜索对象关键词“××医院”和“拉肚子”,通常,用户更

倾向于了解拉肚子这种症状,而非××医院,因此,候选搜索对象关键词“拉肚子”应该比“××医院”具有更大的预设对象重要度系数。电子设备选择的第一候选搜索目标词对可以包括“拉肚子治疗”、“拉肚子怎么样”。

[0067] 实践中,电子设备可以通过机器学习方法训练对象重要度系数。例如,当本申请的生成网页的方法应用于搜索引擎的广告信息网页生成时,可以将多个搜索式和相应的转化数据作为样本,从每个搜索式中提取搜索对象关键词,计算完成转化的页面内容与各搜索对象关键词的对象相似度,通过机器学习方法训练各个搜索对象关键词对转化结果的影响程度量值作为对象重要度系数。其中,转化数据例如可以包括但不限于用户点击浏览的商品信息、用户购买的服务等等,当上述广告信息是医疗方面的广告信息时,转化数据还可以包括用户选择的医院科室、用户购买的药品信息等等。

[0068] 在本实施例的一些可选实现方式中,搜索对象关键词可以对应不同的对象类别,如医疗领域,搜索对象关键词“拉肚子”可以对应对象类别“泛症状”,搜索对象关键词“咽炎”可以对应对象类别“精确疾病”,上述电子设备可以根据对象类别对搜索式中包含的搜索对象关键词按照对象类别进行统计,通过诸如机器学习之类的方法训练每个对象类别的对象类别重要度系数,当搜索式中提取的搜索对象关键词对应不同对象类别时,选择对象类别重要度系数最大的搜索对象关键词对应的候选搜索目标词对作为第一候选搜索目标词对。例如泛症状的对象类别重要度系数大于医院的对象类别重要度系数,则在上述示例中选择包含泛症状“拉肚子”的候选搜索目标词对作为第一候选搜索目标词对,而不选择包含医院“××医院”的候选搜索目标词对。

[0069] 步骤404,根据预先训练的意图重要度系数从第一候选搜索目标词对中选取第二候选搜索目标词对。

[0070] 在本实施例中,上述电子设备还可以对各搜索意图关键词预先训练有意图重要度系数,在该步骤中,电子设备可以根据预先训练的意图重要度系数,对步骤403所选取的第一候选搜索目标词对中的各候选搜索意图关键词的意图重要度系数进行对比,并选取意图重要度系数最大的候选搜索意图关键词,将包含该候选搜索意图关键词的第一候选搜索目标词对作为第二候选搜索目标词对。

[0071] 其中,上述意图重要度系数可以用于表示各搜索意图关键词对用户所了解的搜索结果的影响程度。例如前述示例中,对于搜索式“××医院对孩子拉肚子的治疗效果怎么样”,候选搜索目标词对可以包括“××医院治疗”、“××医院怎么样”、“拉肚子治疗”、“拉肚子怎么样”。其中包括了两个候选搜索意图关键词“治疗”和“怎么样”,通常,用户更倾向于了解针对拉肚子这种症状的治疗方法或效果,而非怎么样,因此,候选搜索意图关键词“治疗”应该比“怎么样”具有更大的预设意图重要度系数。从而,电子设备选择的第二候选搜索目标词对可以包括“拉肚子治疗”。

[0072] 实践中,电子设备可以通过诸如机器学习之类的方法训练意图重要度系数。例如,当本申请的生成网页的方法应用于搜索引擎的广告信息网页生成时,可以将多个搜索式和相应的转化数据作为样本,从每个搜索式中提取搜索意图关键词,计算完成转化的页面内容与各搜索意图关键词的意图相似度,通过机器学习方法训练各个搜索意图关键词对转化结果的影响程度量值作为意图重要度系数。在一些实现中,搜索意图关键词可以对应不同的意图类别,上述电子设备可以根据意图类别对搜索式中包含的搜索意图关键词按照类别

进行统计,通过机器学习方法训练每个意图类别的意图类别重要度系数,当搜索式中提取的搜索意图关键词对应不同意图类别时,选择意图类别重要度系数最大的搜索意图关键词对应的第一候选搜索目标词对作为第二候选搜索目标词对。

[0073] 步骤405,将第二候选搜索目标词对确定为搜索目标词对。

[0074] 在本实施例中,电子设备可以进一步将通过步骤404选取的第二候选搜索目标词对确定为搜索目标词对。例如,在上述的示例中,将“拉肚子治疗”作为搜索目标词对。

[0075] 步骤406,将搜索目标词对与根据上述网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配。

[0076] 在本实施例中,上述电子设备可以将页面上的各个页面元素内容根据主题进行区分和存储,响应于接收到用户的网页浏览请求,上述电子设备还可以根据上述网页浏览请求获取所请求的页面中各个页面元素内容所对应的各个主题,并在提取搜索目标词对之后,将各个主题分别与上述搜索目标词对进行匹配,以确定与搜索目标词对最相关的一个或较相关的多个主题。这里,所请求页面的页面元素内容可以按照主题进行区分和存储。电子设备可以将提取到的搜索目标词对与所获取的一个或多个主题进行匹配。匹配时,电子设备可以将主题与搜索目标词对进行匹配,也可以将主题对应的页面元素内容与搜索目标词对进行匹配,本申请对此不做限定。

[0077] 步骤407,基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页。

[0078] 在本实施例中,生成网页的方法运行于电子设备接着可以获取所匹配到的主题对应的页面元素内容,根据匹配结果对这些页面元素内容进行排序,生成网页以推送给用户。当匹配到多个主题时,电子设备可以根据这多个主题针对上述搜索目标词对的重要度由高到低的顺序,对相应的页面元素内容进行排列生成网页。

[0079] 在本实施例中,上述实现流程中的步骤401、步骤406和步骤407分别与前述实施例中的步骤201、步骤203和步骤204基本相同,在此不再赘述。

[0080] 从图4中可以看出,与图2对应的实施例不同的是,本实施例中生成网页的方法的流程400通过在搜索式包括多个搜索对象关键词或者搜索意图关键词时根据重要度系数确定搜索目标词对的步骤402-405代替了步骤202,通过对搜索目标词对的选择,进一步提高了对用户搜索意图把握的准确性。

[0081] 进一步参考图5,作为对上述各图所示方法的实现,本申请提供了一种生成网页的装置的一个实施例,该装置实施例与图2所示的方法实施例相对应,该装置具体可以应用于电子设备中。

[0082] 如图5所示,本实施例的生成网页的装置500包括:获取模块501、解析模块502、匹配模块503及生成模块504。其中,获取模块501可以配置用于响应于接收到用户的网页浏览请求,获取用户输入的搜索式;解析模块502可以配置用于对该搜索式进行解析提取搜索目标词对,其中,搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;匹配模块503可以配置用于将搜索目标词对与根据网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配;生成模块504可以配置用于基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页。

[0083] 在本实施例的一些可选实现方式中,解析模块502可以包括:提取单元(未示出),

可以配置用于对上述搜索式进行解析提取至少一个候选搜索目标词对,其中,每个搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;第一选择单元(未示出),可以配置用于根据预先训练的对象重要度系数,从所解析出的至少一个候选搜索目标词对中选取第一候选搜索目标词对,其中,第一候选搜索目标词对包含对象重要度系数最大的搜索对象关键词;第二选择单元(未示出),可以配置用于根据预先训练的意图重要度系数从第一候选搜索目标词对中选取第二候选搜索目标词对,其中,第二候选搜索目标词对包括意图重要度系数最大的搜索对象关键词;确定单元(未示出),可以配置用于将第二候选搜索目标词对确定为搜索目标词对。

[0084] 在本实施例的一些可选实现方式中,提取单元进一步可以配置用于:将上述搜索式与预设对象关键词集合进行匹配,以确定候选搜索对象关键词;将上述搜索式与预设意图关键词集合进行匹配,以确定候选搜索意图关键词;基于对所确定的候选搜索对象关键词和候选搜索意图关键词,组合生成候选搜索目标词对。

[0085] 在本实施例的一些可选实现方式中,第一选择单元进一步可以配置用于:确定各个候选搜索目标词对中的候选搜索对象关键词的所属对象类别;根据预先训练的搜索对象关键词所属对象类别的对象类别重要度系数,从所解析出的至少一个候选搜索目标词对中选取第一候选搜索目标词对,其中,第一候选搜索目标词对包含所属对象类别的对象类别重要度系数最大的搜索对象关键词。

[0086] 在本实施例的一些可选实现方式中,第二选择单元进一步可以配置用于:确定各个候选搜索目标词对中的候选搜索意图关键词的所属意图类别;根据预先训练的搜索意图关键词所属意图类别的意图类别重要度系数,从第一候选搜索目标词对中选取第二候选搜索目标词对,其中,第二候选搜索目标词对包含所属类别的意图类别重要度系数最大的搜索意图关键词。

[0087] 在本实施例的一些可选实现方式中,装置500还包括训练模块(未示出),可以配置用于通过以下方法获取对象重要度系数以及意图重要度系数:获取一定数量的搜索式和用户根据搜索式所选择的主题作为样本;从样本中的每个搜索式中提取至少一个搜索对象关键词和至少一个搜索意图关键词;基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索对象关键词的对象相似度,训练各个搜索对象关键词针对各个主题的对象重要度系数;基于用户根据搜索式所选择的主题与该搜索式所对应的各个搜索意图关键词的意图相似度,训练各个搜索意图关键词针对各个主题的意图重要度系数。

[0088] 在本实施例的一些可选实现方式中,匹配模块503可以包括:获取单元(未示出),可以配置用于根据网页浏览请求获取所请求页面的各个页面元素内容对应的各个主题;计算单元(未示出),可以配置用于计算各个主题与搜索目标词对的关联度;匹配单元(未示出),可以配置用于基于各个关联度,确定与搜索目标词对相匹配的主题。

[0089] 在本实施例的一些可选实现方式中,生成模块504可以进一步配置用于:按照与搜索目标词对的关联度由大到小的顺序,对各个所匹配到的主题所对应的页面元素内容进行排序,生成网页。

[0090] 值得说明的是,生成网页的装置500中记载的诸模块与参考图2描述的方法中的各个步骤相对应。由此,上文针对方法描述的操作和特征同样适用于生成网页的装置500及其所包含的模块或单元,在此不再赘述。

[0091] 本领域技术人员可以理解,上述生成网页的装置500还包括一些其他公知结构,例如处理器、存储器等,为了不必要地模糊本公开的实施例,这些公知的结构在图5中未示出。

[0092] 下面参考图6,其示出了适于用来实现本申请实施例的电子设备的计算机系统600的结构示意图。

[0093] 如图6所示,计算机系统600包括中央处理单元(CPU)601,其可以根据存储在只读存储器(ROM)602中的程序或者从存储部分608加载到随机访问存储器(RAM)603中的程序而执行各种适当的动作和处理。在RAM 603中,还存储有系统600操作所需的各种程序和数据。CPU 601、ROM 602以及RAM 603通过总线604彼此相连。输入/输出(I/O)接口605也连接至总线604。

[0094] 以下部件连接至I/O接口605:包括键盘、鼠标等的输入部分606;包括诸如阴极射线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分607;包括硬盘等的存储部分608;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分609。通信部分609经由诸如因特网的网络执行通信处理。驱动器610也根据需要连接至I/O接口605。可拆卸介质611,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器610上,以便于从其上读出的计算机程序根据需要被安装入存储部分608。

[0095] 特别地,根据本申请的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本申请的实施例包括一种计算机程序产品,其包括有形地包含在机器可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分609从网络上被下载和安装,和/或从可拆卸介质611被安装。

[0096] 本申请实施例中所涉及到的单元可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的模块也可以设置在处理器中,例如,可以描述为:一种处理器包括获取模块、解析模块、匹配模块及生成模块。其中这些模块的名称在某种情况下并不构成对该模块本身的限定,例如,生成模块还可以被描述为“配置用于基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页的模块”。

[0097] 作为另一方面,本申请还提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例中的装置中所包含的计算机可读存储介质;也可以是单独存在,未装配入终端中的计算机可读存储介质。上述计算机可读存储介质存储有一个或者一个以上程序,当该程序被一个或者一个以上的处理器执行时,使得上述设备:响应于接收到用户的网页浏览请求,获取用户输入的搜索式;对该搜索式进行解析提取搜索目标词对,其中,搜索目标词对包括一个搜索对象关键词与一个搜索意图关键词;将搜索目标词对与根据网页浏览请求获取的所请求页面的各个页面元素内容对应的各个主题分别进行匹配;基于所匹配到的主题,选取所请求页面的页面元素内容进行排序生成网页。

[0098] 作为再一方面,本申请还提供了一种计算设备,包括:一个或多个处理器;存储器,用于存储一个或多个程序(该存储器例如可以包括上述的计算机可读存储介质);其中,当上述一个或多个程序被上述一个或多个处理器执行时,使得上述处理器实现本申请提供的生成网页的方法。

[0099] 以上描述仅为本申请的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解,本申请中所涉及的发明范围,并不限于上述技术特征的特定组合而成的技术

方案,同时也应涵盖在不脱离上述发明构思的情况下,由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本申请中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。

100

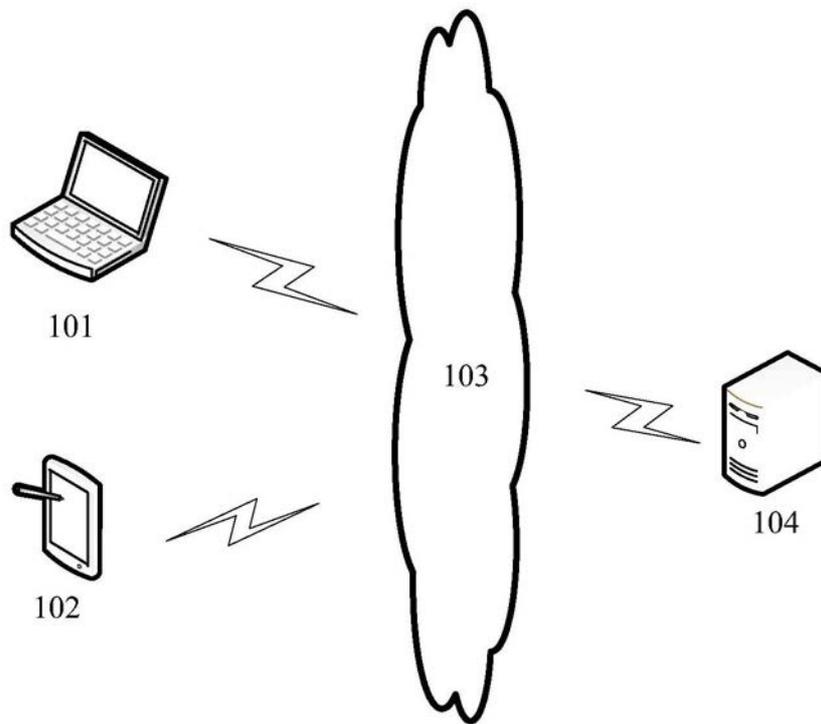


图1

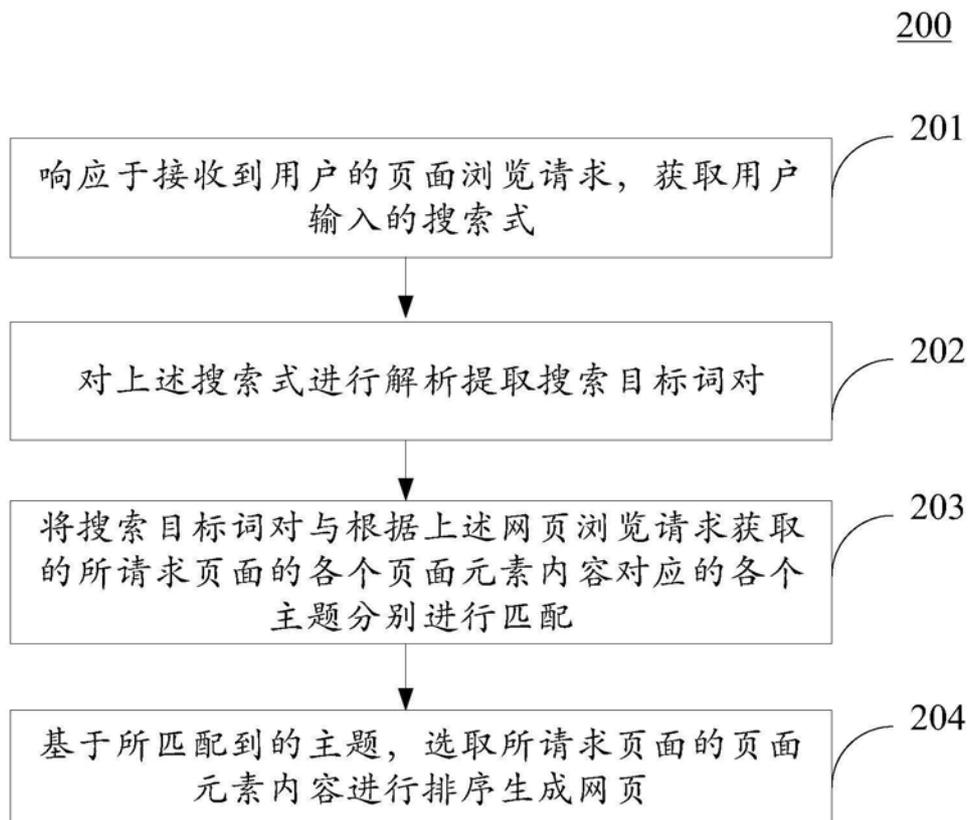


图2

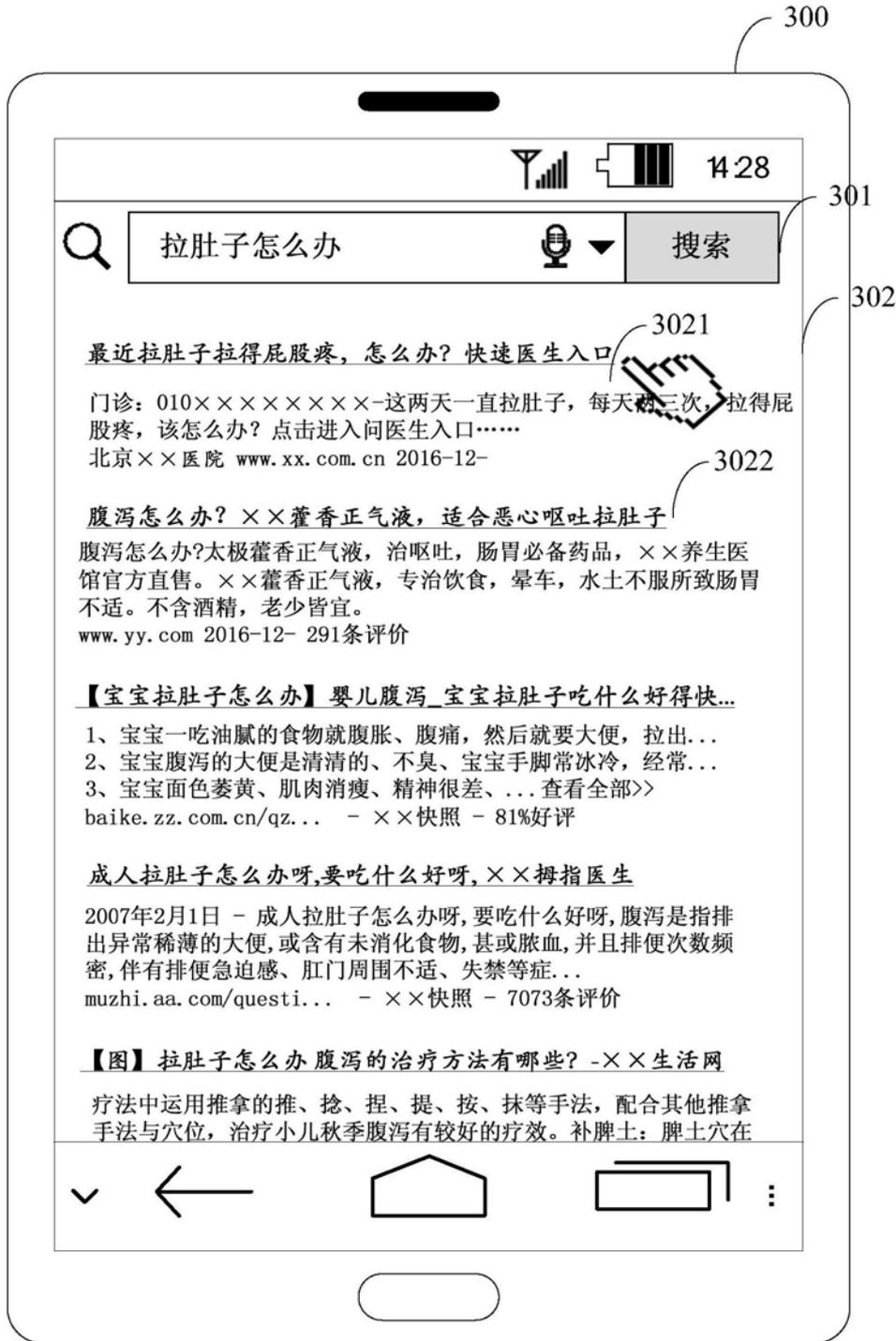


图3a

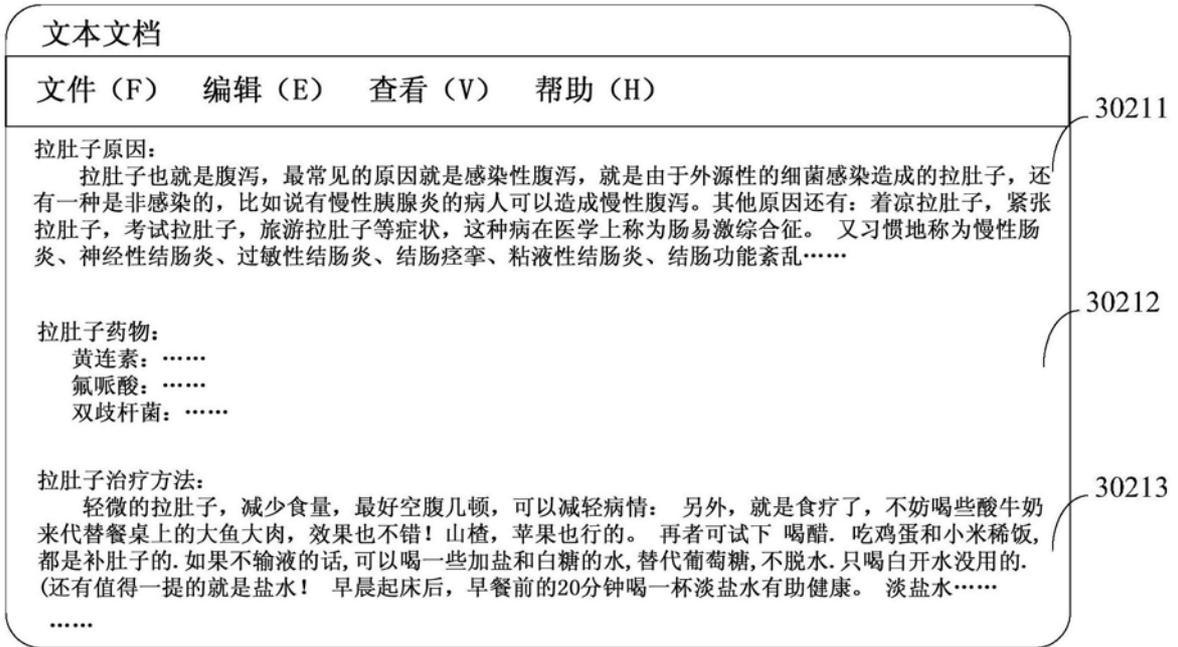


图3b

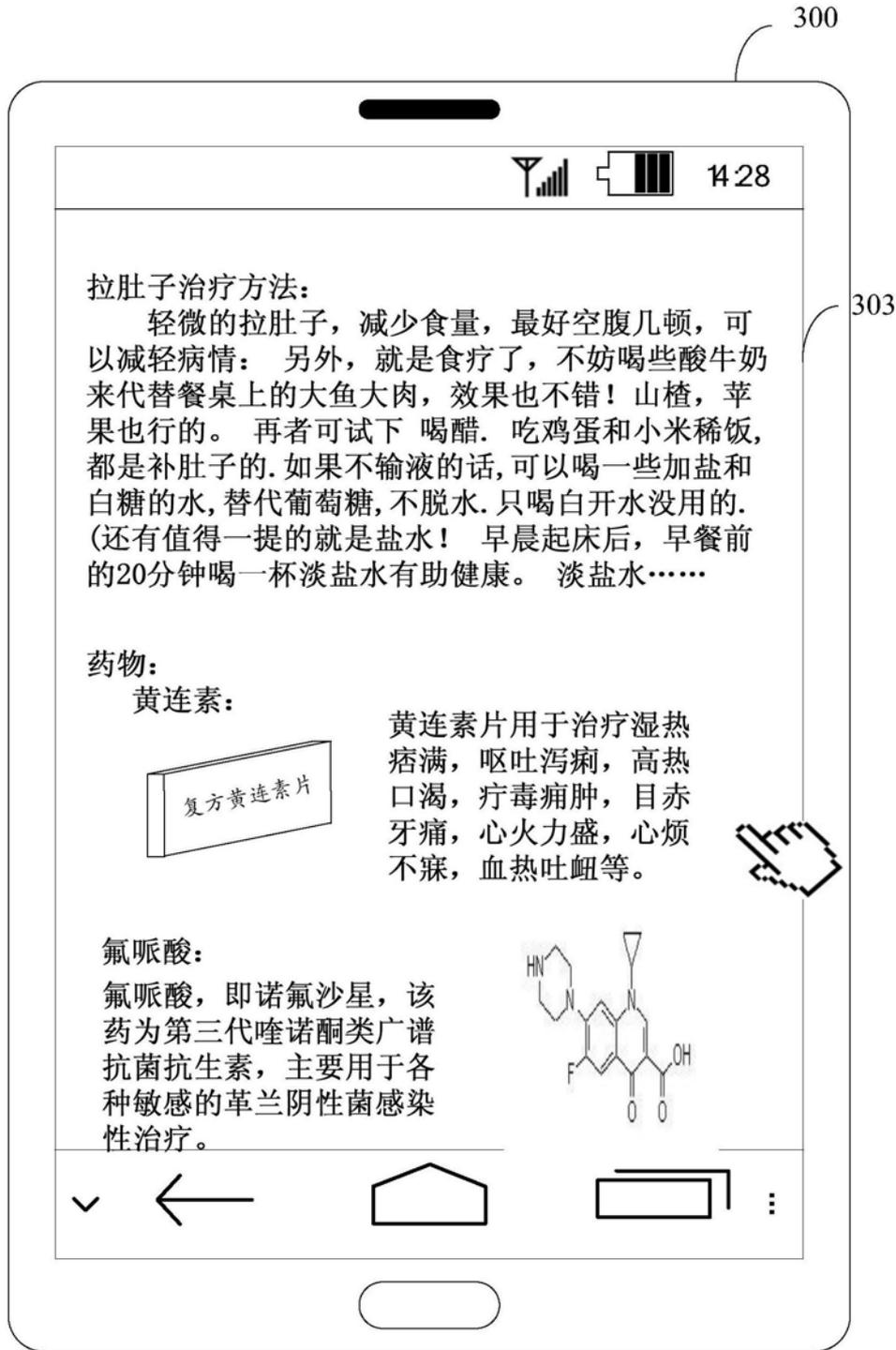


图3c

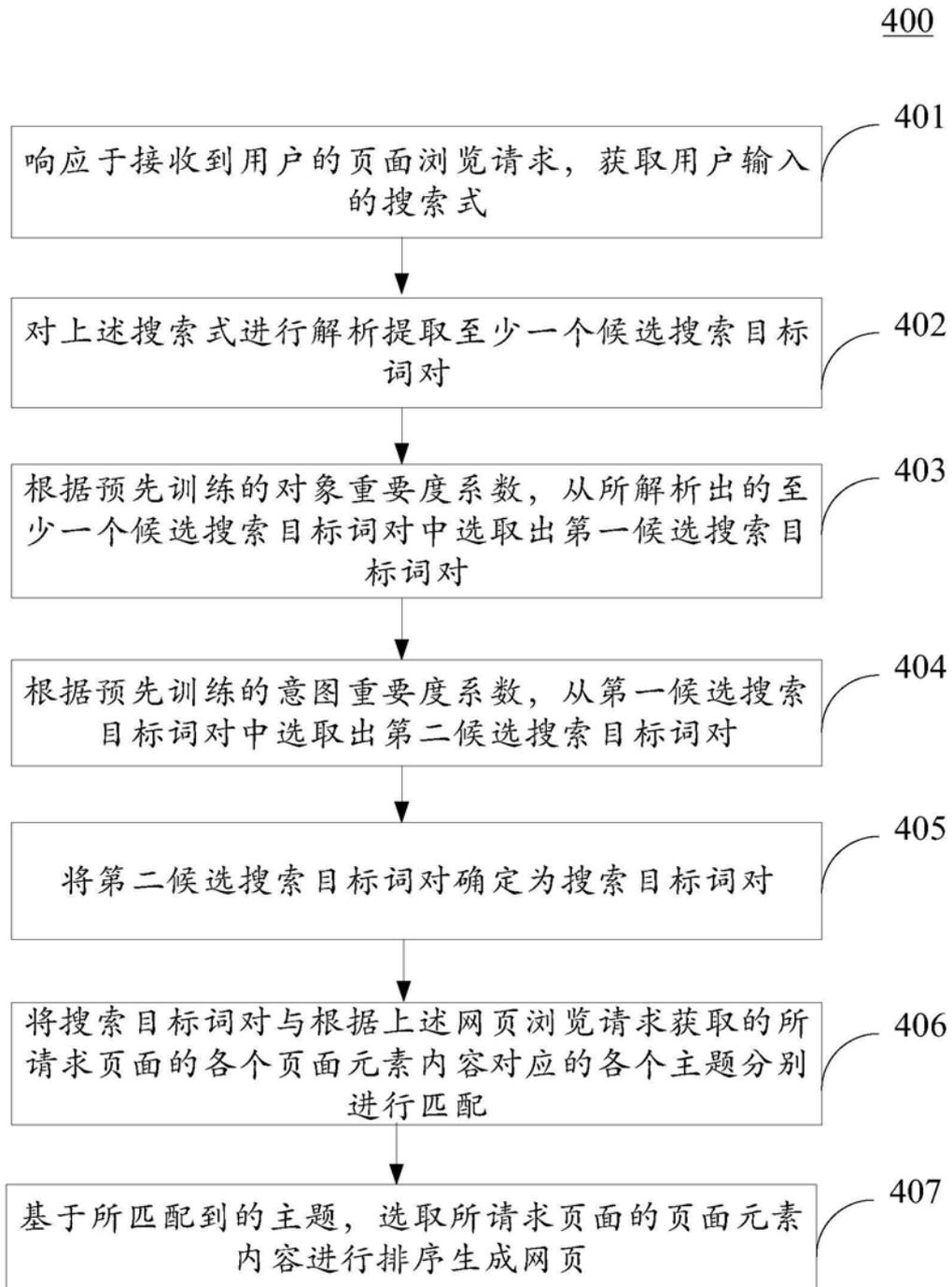


图4

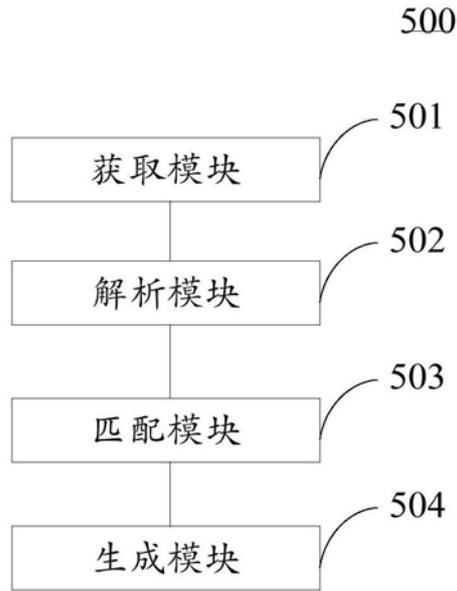


图5

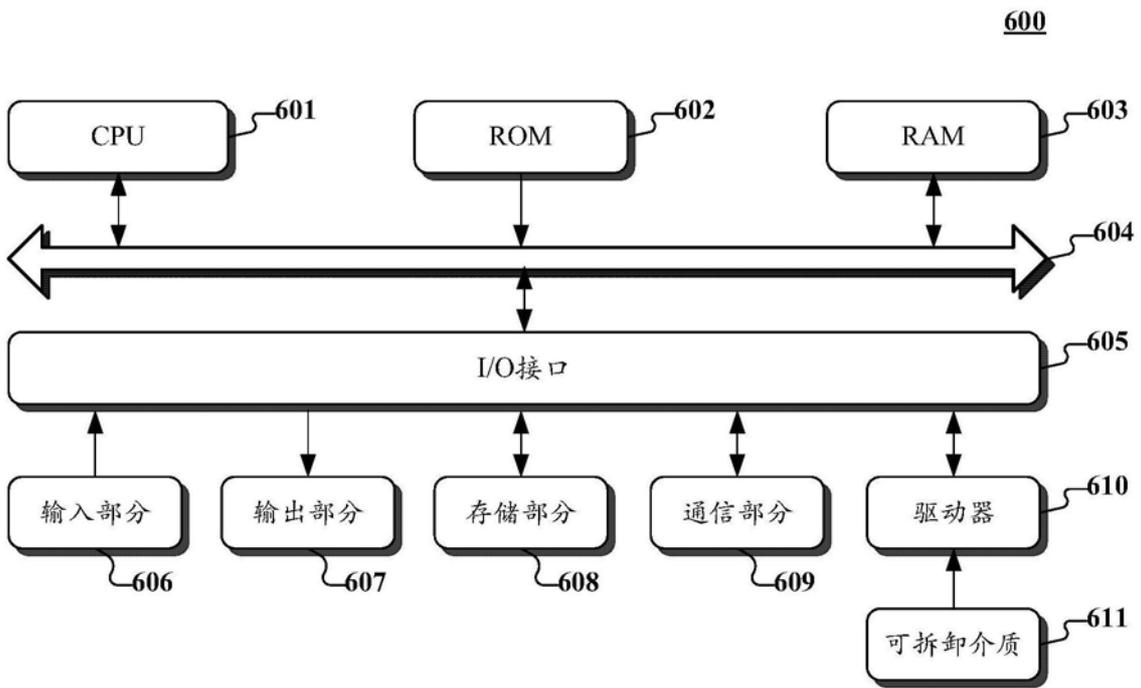


图6