



(12) 发明专利申请

(10) 申请公布号 CN 102498464 A

(43) 申请公布日 2012. 06. 13

(21) 申请号 201080041162. X

代理人 黄嵩泉

(22) 申请日 2010. 09. 15

(51) Int. Cl.

(30) 优先权数据

G06F 3/048 (2006. 01)

12/560, 435 2009. 09. 16 US

G06F 9/44 (2006. 01)

(85) PCT申请进入国家阶段日

2012. 03. 15

(86) PCT申请的申请数据

PCT/US2010/049002 2010. 09. 15

(87) PCT申请的公布数据

W02011/034974 EN 2011. 03. 24

(71) 申请人 微软公司

地址 美国华盛顿州

(72) 发明人 G·佩朗塔托斯 K·卡纳瓦特

J·S·瓦纳

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

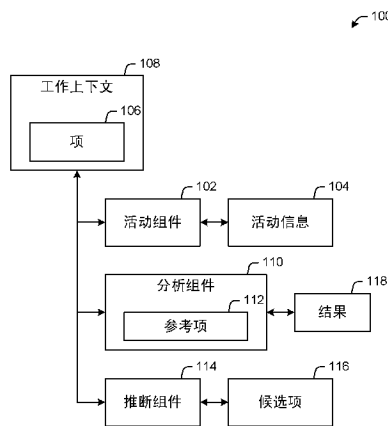
权利要求书 2 页 说明书 10 页 附图 11 页

(54) 发明名称

自动发现任务的上下文相关的项目

(57) 摘要

用于使用户能自动恢复与工作上下文相关联的文档及其他信息以及恢复与特定项目相关联的文档及其他信息人工制品的体系结构。该体系结构允许监测和记录与关于特定工作上下文的信息人工制品的用户交互有关的活动信息。用户可以选择具有与工作上下文有关的工作内容的一部分(例如,文档中的词或其他类型的参考项)的文档。对活动信息和参考项执行词汇分析,以标识词汇相似处。基于词汇相似处,从信息人工制品中推断候选项(例如,相关的文档)列表。向用户呈现与工作上下文有关的候选项,用户可以选择特定的项来重新建立工作上下文。



1. 一种计算机实现的上下文系统 (100), 包括:
监测和记录与和工作上下文相关联的项的用户交互有关的活动信息的活动组件 (102);
对参考项和与所述活动信息相关联的项执行词汇分析的分析组件 (110); 以及
当重新构建所述工作上下文时, 基于所述活动信息和所述词汇分析的结果来推断候选项的推断组件 (114)。
2. 如权利要求 1 所述的系统, 其特征在于, 所述活动组件还包括作为无需用户交互的后台功能来监测用户活动的后台组件。
3. 如权利要求 1 所述的系统, 其特征在于, 所述词汇分析的结果包括在所述参考项和所述候选项中发现的共同的词。
4. 如权利要求 1 所述的系统, 其特征在于, 所述用户交互还包括以下各项操作中的至少一个: 程序或数据之间的切换操作、程序或数据之间的复制粘贴操作、将程序或数据插入到附件中的插入操作、切换频率测量、时间戳操作、设置书签操作、链接操作、将项作为副本保存在另一位置中的保存操作或停留时间。
5. 如权利要求 1 所述的系统, 其特征在于, 还包括用于显式地查询与所述参考项有关的候选项的查询组件。
6. 如权利要求 1 所述的系统, 其特征在于, 还包括用于收集一候选项集合并隐式地查询所述集合以寻找与所述参考项有关的候选项的收集组件。
7. 如权利要求 1 所述的系统, 其特征在于, 还包括用于本地或远程中的至少一项地存储所述活动信息的存储组件。
8. 如权利要求 1 所述的系统, 其特征在于, 还包括:
包括与所述工作上下文有关的至少一个参考项的参考人工制品; 以及
用于呈现与所述工作上下文有关的候选项的呈现组件。
9. 如权利要求 1 所述的系统, 其特征在于, 还包括用于向至少一个预定的活动信息项的相关性分配加权因数的加权组件。
10. 一种计算机实现的上下文方法, 包括:
监测和记录与和工作上下文中的信息人工制品的用户交互有关的活动信息 (900);
对所述活动信息和与所述工作上下文有关的参考项执行词汇分析, 以标识词汇相似处 (902);
基于由所述分析所产生的所述活动信息和所述参考项之间的所述词汇相似处, 推断从所述信息人工制品中选择的候选项 (904); 以及
呈现与所述工作上下文有关的候选项 (906)。
11. 如权利要求 10 所述的方法, 其特征在于, 还包括通过处理所述候选项来重新构建所述工作上下文。
12. 如权利要求 10 所述的方法, 其特征在于, 还包括从选定的候选项中推断相应的相关联的候选项, 以获取更准确的候选项的集合。
13. 如权利要求 10 所述的方法, 其特征在于, 还包括根据所述活动信息和所述参考项之间的词汇相似处的频率, 分层次地排序所述候选项。
14. 如权利要求 10 所述的方法, 其特征在于, 还包括根据所述信息人工制品和包含所

述参考项的参考文档之间的切换的频率,分层次地排序所述候选项。

15. 如权利要求 10 所述的方法,其特征在于,还包括作为后台进程自动地监测所述用户交互。

自动发现任务的上下文相关的项目

背景技术

[0001] 用户在工作时频繁地切换上下文。这可能是由于中断,以及需要撇开对于特定的任务的正在执行的工作,以便切换到另一任务。在返回到原始任务时,用户可能在回想他们处理的所有项(例如,文件、应用程序、位置、人、通信等等)以便恢复原始任务时会遇到问题。

[0002] 如果用户严格地对特定工作上下文中涉及的每件事和每个人作记录,则与上下文切换有关的问题可能被减轻。然而,创建特定工作上下文中所使用的所有项的完整的记录可能就与执行实际工作本身的工作量同样多。另外,不能轻松地记录某些项。例如,可能难以保存或引用电子邮件消息、即时消息谈话、或诸如计算器之类的不产生文件的应用程序。

[0003] 在实践中,用户通常依赖于记忆回想,以重定位和重建工作上下文。然而,这会是耗时的且易于出错的策略。用户也可以依赖于接受用于定位相关网页及其他项的关键字查询的传统的搜索引擎。利用搜索引擎,输入特定短语或参数,以便定位相关项。尽管搜索引擎产生结果,但是,引擎常常产生大量不相关的结果,因此,在回想与特定任务有关的特定组的项时,没有帮助。

[0004] 另外,关键字搜索结果只呈现包含相关项的项列表。即使从关键字搜索中获取相关结果的目标列表,搜索也可能通常只检索文档,而不是应用程序状态。执行搜索会是耗时的,几乎不能保证可以重构以前的工作上下文项的准确的列表。

发明内容

[0005] 下面提供了简明的概述,以便提供对此处所描述的一些新颖实施方式的基本理解。本发明内容不是详尽的概览,并且它不旨在标识关键/重要元素或描绘本发明的范围。其唯一的目的是以简化形式呈现一些概念,作为稍后呈现的更具体实施方式的序言。

[0006] 公开了用于标识与用户在其中工作过的工作上下文有关的、诸如文档及其他类型的信息人工制品的项,并在离开原始上下文之后能检索(例如,手动,自动)项以有效地且高效地重新构建工作上下文的体系结构。

[0007] 可以组合多个信息流,以自动推断给定任务的上下文关系。例如,可以监测和记录诸如在文档之间切换和复制/粘贴操作之类的用户活动。对用户活动执行词汇分析,还使用参考项(reference item)来推断由用户处理的项之间的关系。基于参考项,向用户呈现上下文相关的项,从而产生表示由用户在前面处理的特定用户任务的相关项目的结果,而并非从典型的基于关键字搜索中获取的广泛的、一般化的结果。

[0008] 为了为实现上述及相关目的,本文结合下面的描述和附图来描述某些说明性方面。这些方面指示了可以实施本文所公开的原理的各种方式,所有方面及其等效方面旨在落入所要求保护的的主题的范围内。结合附图阅读下面的具体实施方式,其他优点和新颖特征将变得显而易见。

附图说明

- [0009] 图 1 示出了根据所公开的体系结构的计算机实现的上下文系统。
- [0010] 图 2 示出了包括用于后台监测和分析的额外的实体的上下文系统的替换实施例。
- [0011] 图 3 示出了与上下文系统一起使用的各类型的用户交互。
- [0012] 图 4 示出了包括用于后台收集、查询和存储的额外的实体的上下文系统的替换实施例。
- [0013] 图 5 示出了上下文系统的替换实施例。
- [0014] 图 6 示出了与上下文系统一起使用的各类型的信息人工制品。
- [0015] 图 7 示出了包括用于计分、加权和手动显示的额外的实体的上下文系统的替换实施例。
- [0016] 图 8 示出了根据所公开的体系结构的用于推断上下文关系的实现。
- [0017] 图 9 示出了推断上下文关系的方法。
- [0018] 图 10 示出了推断上下文关系的方法的额外的方面。
- [0019] 图 11 示出了根据所公开的体系结构的可用于提供上下文关系的推断的计算系统的框图。

具体实施方式

[0020] 所公开的体系结构允许当重新构建特定工作上下文时自动恢复与该工作上下文相关联的文档和信息人工制品,以及恢复相关联的文档及其他信息人工制品。该体系结构允许监测和记录与关于特定工作上下文的信息人工制品的用户交互有关的活动信息。用户可以选择具有与工作上下文有关的词 (term) 或其他类型的参考项的文档。可以对参考项和对具有与参考项有关的活动信息执行 (例如,词汇的) 分析以标识相似处。基于从分析导出的相似处,从信息人工制品来推断候选项 (例如,相关的文档) 列表。向用户呈现与工作上下文有关的候选项,然后,用户可以选择,例如,特定的项来重新构建工作上下文。

[0021] 现在将参考附图,全部附图中相同的参考编号用于表示相同的元件。在下面的描述中,为了进行说明,阐述了很多具体细节以便提供对本发明的全面理解。然而,显而易见,可以没有这些具体细节的情况下实施各新颖实施方式。在其他情况下,以框图形式示出了各个公知的结构和设备以便于描述本发明。本发明将涵盖落入所要求保护的的主题的精神和范围内的所有修改、等效方案和替换方案。

[0022] 图 1 示出了根据所公开的体系结构的计算机实现的上下文系统 100。提供了监测和记录与和工作上下文 108 相关联的项 106 的用户交互有关的活动信息 104 的活动组件 102。项 106 可以是文档或其他类型的信息人工制品,如下面所详细描述。分析组件 110 对参考项 112 和与活动信息 104 相关联的项 106 执行 (例如,词汇的) 分析。当重新构建工作上下文 108 时,推断组件 114 基于活动信息 104 和分析的结果 118 来推断候选项 116。

[0023] 如此处所使用的,“上下文”可以是指由用户所查看、读取 (例如,在各应用程序之间或在应用程序和数据之间的读取) 和 / 或创建的内容,其中,内容可以是基于文本的文档和 / 或其他信息人工制品。上下文也可以是指与内容相关联的通信,如电子邮件或即时消息。另外,上下文还可以是指其中内容被读取或使用的一个或多个位置,其可包括网站、网络或本地文件夹、协作站点等等。此外,上下文还可以是指涉及内容的环境,如其中用户正处理内容或与内容进行交互时的设置。环境可以是例如涉及某些参与者的会议中的个人桌

面。上下文也可以是指与通信、位置以及环境等等相关联的其他个人（例如，协作者和 / 或参与者）。

[0024] 如图 1 所示,分析组件 110 对参考项 112 执行词汇分析,参考项 112 可以由用户先前打开的文档或其他类型的信息人工制品。然而,可以理解,可以代替词汇分析与词汇分析相结合来使用其他类型的分析。比较活动信息 104,以确定对项 106 的其他项共同的活动。共同的活动可以是例如,文档之间的切换(导航)。如果对项 106 的子集共同的活动被记录,则对该子集执行词汇分析,以确定项 106 和参考项 112 之间是否有词汇的交叉。作为结果 118,输出这些词汇的交叉。将结果 118 与活动信息 104(例如,文档之间的切换数量)进行比较,以建议候选项 116。

[0025] 如图 1 所示,和依赖于用户输入关键字以产生匹配关键字的结果的搜索引擎相反,上下文系统 100 是基于参考项 112 产生结果的。例如,用户可以通过以当前打开的演示文档开始,返回到计划现场考察(site visit)的工作上下文。演示文档是参考项 112,并作为“查询”而非搜索中的关键字来起作用。上下文系统 100 通过标识用户活动与对该活动的词汇分析的交叉来产生候选项 116。一方面,可以基于用户活动来标识项 106,然后,使用词汇分析来区分项 106,这会产生候选项 116 的非常准确的列表。

[0026] 没有词汇匹配的纯用户活动的实例可以产生不相关的结果,如在对文档进行处理时检查不相关的电子邮件或在不相关的网页上冲浪。没有用户活动的纯词汇匹配的实例可以产生碰巧共享相同字的不相关的项(例如,关于 site visit(现场考察)以及 website(网站)或 drop site(投票站点)的电子邮件中的“site(站点)”)。如此,具有词汇匹配和活动匹配两者的实例表示用户正在试图重建的给定工作上下文的高精度的结果。

[0027] 图 2 示出了包括用于后台监测和分析的额外的实体的上下文系统的替换实施例 200。活动组件 102 可包括作为对用户透明且没有用户交互的操作来监测用户活动的后台组件 202。以此方式,活动信息 104 可以在后台获取,而无需中断用户或需要任何用户反馈。

[0028] 如图 2 所示,词汇分析的结果 118 可包括处理在参考项 112 和候选项 116 中找到的共同的词 204。词汇分析可以检查与信息人工制品相关联的内容和元数据,以基于参考项 112 中的内容的起始点,来将相关的内容聚类到一起。共同的词 204 可包括电子邮件中的共同的发件人和 / 或收件人,主题行中的共同的名词短语,以及用户如何组织内容的指示,诸如,例如,文件夹、类别,以及标记。

[0029] 共同的词 204 也可以具体地包括内容或元数据中发现的共同的名词、名词短语、作者姓名以及参与者。名词可以从电子邮件主题行和文档标题中作为词性提取的,也可以是提取的内容、位置 / 路径、电子邮件发件人 / 收件人等等。在用户介绍关于内容的附加信息的级别的情况下,可以使用任何其他合适的方案。

[0030] 图 3 示出了可以与上下文系统一起使用的各类型的用户交互 300。用户交互 300 可包括程序或数据之间的切换操作 302。切换操作 302 指出共同访问的文档,它可以指出暗示文档之间的关系的模式和活动。切换操作 302 指出在当参考项 112 处于用户的焦点上时的时间附近被用户访问的程序、数据、文档和 / 或其他信息人工制品。切换操作 302 可以被用于确定与被用作参考项 112 的初始文档或电子邮件的可能的关联的起始点。

[0031] 如图 3 所示,用户交互 300 也可以包括程序或数据之间的复制 / 粘贴操作 304。其

他用户交互 300 可包括例如,将程序或数据插入到附件、电子邮件或其他类型的消息中的插入操作 306。切换频率测量 308 可以被用来测量文档之间的切换的实例的数量,该数量可以被用来查明工作上下文 108 的共同的关系。时间戳操作 310 可以被用来确定在参考项 112 附近的合适的时间段内是否创建、编辑或保存了候选项 116。设置书签操作 312 可以被用来确定资源是否被记录以供将来参考。链接操作 314 可以暗示共同的工作上下文 108,因为链接操作提供文档之间的访问。将项作为副本保存在另一位置中的保存操作 316 可以被用来暗示共同的工作上下文 108。在信息人工制品上花费的停留时间 318(例如,总的活动时间)可以被用来暗示工作上下文 108。例如,停留时间 318 可以是花费在处理草案上的总时间量,或花费在审阅草案的时间,并可以作为例如键盘和 / 或鼠标活动来测量。

[0032] 如图 3 所示,提供了这样的示例:切换操作 302 可以被分析组件 110 和推断组件 114 用作活动信息 104,以确定共同访问的文档是否共享共同的上下文。在文档之间的切换可以被记录为指出以下内容:文档在切换操作 302 中是源文档还是目标文档,在切换到目标文档之前源文档是否被关闭,目标是否被用户置于焦点上(前台)持续一预定的最小时间间隔(例如,小于 3 秒),目标是否在后台被打开(例如,在没有处于焦点中的 Web 浏览器选项卡中打开),源文档是否是在目标文档被打开之后从其切换的第一文档,以及目标文档是否是源文档被关闭之前被切换到的最后一个文档。另外,可以记录用户在特定文档之间切换的次数。

[0033] 如前所述的用户交互 300(以及其组合)可以由分析组件 110 以及推断组件 114 评估以在过滤不会指出(以比较高的概率)带有共享的上下文的文档的文档共同访问实例时寻找相关性。其他相关的元数据可以随后被记录,以在过滤结果时寻找相关性。

[0034] 图 4 示出了包括用于后台收集、查询和存储的额外的实体的上下文系统的替换实施例 400。可以提供用于显式地查询与参考项 112 有关的候选项 116 的查询组件 402。查询组件 402 显式地检查参考项 112,并查询系统 400 中的相关项以推断候选项 116。参考项 112 可以是由用户所提供的包括线索作为起始点的选定的文档。

[0035] 可任选地,可以提供用于收集一组候选项 116 并隐式地查询与参考项 112 有关的候选项 116 的集合的收集组件 404。以此方式,收集组件 404 隐式地推断参考项 112,并自动将项收集到一起以向用户呈现,以标识相关项目而无需用户预先提供线索。

[0036] 如图 4 所示,可以提供用于本地和 / 或远程存储活动信息 104 的存储组件 406。存储组件 406 可以是用于本地存储的任何合适的数据存储系统,如计算机内置或外置硬盘驱动器、可写入 CD 或 DVD 或诸如闪存驱动器之类的任何可移动存储器组件。存储组件 406 可以可另选地是用于远程存储的任何合适的数据存储组件,如网络服务器或驱动器、磁带备份或现场以外的存储设施。

[0037] 一方面,上下文系统(例如,系统 100,200,400)不存储用户内容,而是存储与用户相关联的活动信息 104。例如,如果一个文档被打开并且在文档和电子邮件消息之间执行了多个切换操作,该切换操作作为用时间戳来指代文档和电子邮件的标识(ID)号被记录到服务器。ID 号不表示文档内容。可以使用本地桌面搜索索引和对应的应用程序编程接口(API)来解析 ID 号,以执行词汇分析。如此,上下文系统可以使用存储在本地桌面搜索索引中的本地数据来对文档执行词汇分析,并可以从服务器中获取活动信息 104。

[0038] 图 5 示出了上下文系统的替换实施例 500。活动组件 102 监测和记录与和工作上

下文 108 中的信息人工制品 502 的用户交互有关的活动信息 104。如上文所描述的,信息人工制品 502 可以是文档或其他类型的数据,如下面所详细描述。参考人工制品 504 包括与工作上下文有关的参考项 112 中的一个或多个。分析组件 110 对活动信息 104 和参考项 112 执行词汇分析,以标识词汇相似处 506。推断组件 114 基于词汇相似处 506 来推断从信息人工制品 502 中选择的候选项 116。提供了用于呈现与工作上下文 108 有关的候选项 116 的呈现组件 508。

[0039] 图 6 示出了可以与上下文系统一起使用的各类型的信息人工制品 502。信息人工制品 502 可包括一种或多种类型的用户可识别和可使用的数据实体 600。数据实体 600 可包括与任何类型的应用程序相关联的文件 602。数据实体 600 也可以包括诸如携带可以被访问的信息的任何活动进程之类的数据流 604。数据实体 600 还可以包括网页 606,该网页 606 包括可被 web 浏览器或其他合适的读取器消费的数据对象。

[0040] 数据实体 600 另外还可以是电子表格 608、电子邮件消息 610、IM 谈话 612、日历约会 614、便签 616 或包含在如前所述的类型的数据实体 600 中的任何一种中的嵌入式元数据,或任何其他类型的信息人工制品 502。嵌入式元数据 618 可以被包括在,例如,绘图文档中或其他非基于文本的文档中。可以理解,信息人工制品 502 可包括任何类型或格式的文档或其中用户参与到创建或消费的其他数据结构。

[0041] 活动信息 104 可以跨不会产生信息人工制品 502 的应用程序而扩展。如果活动信息 104 可以维持可以用于重新构建工作上下文 108 的应用程序状态,则它可以是有用的。例如,如果用户正在使用电子表格并在电子表格和计算器应用程序之间切换,则活动信息 104 的切换可以对推断工作上下文 108 有用。

[0042] 图 7 示出了包括用于计分、加权和手动显示上下文信息的额外的实体的上下文系统的替换实施例 700。呈现组件 508 可包括用于分层次地排序候选项 116 的分数分配 702。分数分配 702 允许候选项 116 基于诸如被分配用户交互 300 或信息人工制品 502 中的特定一个的较高或较低的值之类的附加信息而被排序。

[0043] 例如,分数分配 702 可以是基于词频和逆文档频率 (TFIDF) 的乘积的词汇分数:

[0044] $TFIDF = \Sigma (\text{词频} * \text{逆文档频率})$

[0045] 其中,对于每一个匹配名词,

[0046] 词频 (归一化) = 文档中的匹配名词的出现次数 / 文档中的名词总数,以及

[0047] 逆文档频率 = $\ln(T/L)$ (即, (文档的总数 / 包含该名词的文档数量) 的自然对数)。

[0048] 在另一示例中,分数分配 702 可以是基于切换频率和逆文档频率 (SFIDF) 的乘积的共同访问分数:

[0049] $SFIDF = \text{切换频率} * \text{逆文档频率}$,

[0050] 其中,

[0051] 切换频率 (归一化) = 与参考文档的切换数量 / 与文档相关联的切换的总数

[0052] 逆文档频率 = $\ln(T/L)$ (即, (文档的总数 / 带有与参考文档的至少一个切换的文档数量) 的自然对数)。

[0053] 另外,词汇分数和共同访问分数可以被合并成单个分数。较大数量的参考点可以指示相关匹配的较大的似然率。两个重叠的交叉与单个交叉相比可以指示较大的相关性,

两个重叠交叉,如文档和电子邮件两者都具有与网页的词汇和活动交叉,提高了网页的排名,因为有两个对工作上下文的参考,而并非一个。

[0054] 如图 7 所示,加权组件 704 可以被用来分配指出至少一个预定的活动信息项的相关性的加权因数。加权组件 704 可以被用来向与特定文档或活动的相关性分配较大的权重或较小的权重。加权可以与各类型的用户交互 300 和信息人工制品 502 相关联,以便分配较大或较小的值。

[0055] 在推断工作上下文 108 时,可以给复制/粘贴操作 304 分配高值,因为内容正在文档之间被复制。基于活动的本质,作为附件向电子邮件中插入文档或将一个链接插入到文档中可以被视作比在文档之间切换更高度相关。另外,还可以基于切换之间的时间来考虑加权,其中,例如,快速的一系列切换可暗示较小的相关性。

[0056] 如图 7 所示,提供了用于在用户界面中手动显示候选项 116 的可操作的菜单元件 706。菜单元件 706 可以是在用户界面中位于文档应用程序中的特定位置处的按钮。可以点击按钮,此后,向用户呈现与文档的工作上下文 108 相关联的候选项 116 的集合。候选项 116 可包括其他文档、相关的电子邮件包含与文档一起打开的相关词的网页等。菜单元件 706 也可以是具有与同文档有关的特定任务或项目相关联的相关项的自动生成的列表的边栏。可另选地,菜单元件也可以是桌面项,其中,一组文档被视为参考项,而并非单个参考项 112。

[0057] 图 8 示出了根据所公开的体系结构的用于推断上下文关系的实现 800。实现 800 是描绘了正在被上下文系统记录的五个项的示例性概念图。参考项 802 表示用户试图恢复的工作上下文的一部分(例如,“site visit agenda(现场考察议程)”文档)。有四个候选项可用。候选项 1804 是表示“plan(计划)”的文档,并具有此文档和参考项之间的某一记录的切换活动。然而,在候选项 1804 和参考项 802 之间没有词汇相似处。

[0058] 候选项 2806 涉及“plan for site visit(现场考察的计划)”并具有记录的切换活动和与参考项的词汇相似处(“site visit(现场考察)”)两者。候选项 3808 包括词“visit this web site(访问此网站)”,但是只有与参考项的词汇相似处(“visit(访问)”,“site(站点)”)。候选项 4810 具有记录的复制/粘贴活动和与参考项的词汇相似处(“agenda(议程)”)。结果是,只有候选项 2806 和候选项 4810 被作为与参考项 802 上下文相关的而被返回。

[0059] 下面是示例性情况以演示此处所公开的上下文体系结构的操作。用户 1,一个营销团队的成员,专注于一公司的平面广告和 web 广告。当接收到 IM 指示用户 1 找出哪些外国杂志适合于做专业齿轮的广告时,用户 1 正在处理一项目。用户 1 过去已经浏览了关于此主题的某些网站,但是,没有正式地调查此事。

[0060] 用户 1 为此任务创建新空间(例如,工作空间或文件夹),将接收到的 IM 添加到该空间中,定位包括与此任务有关的以前的研究的演示文档并将其添加到该空间中。上下文系统自动将与已经在空间中的内容的创建或传播有关的内容添加到空间中,如用户 1 从中复制和粘贴以创建演示文档的源,以及所发送的包括相关内容的电子邮件。上下文系统通过剖析(profiling)由用户 1 查看和处理的信息人工制品,包括与信息人工制品有关的活动,来自动建议与空间中的内容有关的信息人工制品,以帮助重新构建工作上下文。建议可以基于已经在空间中的内容,匹配关键字、作者和协作者以及大约与空间中的内容同时打

开的其他信息人工制品。

[0061] 此处所包括的是一组表示用于执行所公开的体系结构的新颖方面的示例性方法的流程图。尽管出于解释简明的目的,此处例如以流程图形式示出的一个或多个方法被示出并且描述为一系列动作,但是可以理解,各方法不受动作的次序的限制,因为根据本发明,某些动作可以按与此处所示并描述的不同的次序和/或与其他动作同时发生。例如,本领域的技术人员将明白并理解,方法可被替换地表示为一系列相互相关联的状态或事件,诸如以状态图的形式。此外,并非方法中所示出的所有动作都是新颖实现所必需的。

[0062] 图9示出了推断上下文关系的方法。在900,对于与和工作上下文中的信息人工制品的用户交互有关的活动信息执行监测和记录。在902,对活动信息和与工作上下文有关的参考项执行词汇分析,以标识词汇相似处。在904,推断候选项,其中,候选项是基于由分析所产生的活动信息和参考项之间的词汇相似处来从信息人工制品中选出来的。在906,呈现与工作上下文有关的候选项。

[0063] 图10示出了图9的推断上下文关系的方法的额外的方面。在1000,通过处理候选项,来重新构建工作上下文。在1002,从选定的候选项推断相应的相关联的候选项,以获取更准确的候选项的集合。以此方式,特定候选项可以变为其自己的候选项集合的参考项,可以将后一集合与初始集合进行比较,以产生只超出从原始参考项推断的候选项的扩展的候选项集合。

[0064] 如图10所示,在1004,根据活动信息和参考项之间词汇相似处的频率,分层次地排序候选项。在1006,根据信息人工制品和包含参考项的参考文档之间的切换的频率,分层次地排序候选项。在1008,作为后台进程,自动监测用户交互。

[0065] 如此处所公开的,上下文系统和方法通过使活动信息与同该活动信息有关的项的词汇分析相交,来推断由用户处理的项之间的关系。上下文系统和方法通过查询推断的关系并由此产生高精度的结果,基于参考项,向用户呈现上下文相关的项。上下文系统和方法组合两个信息流(即,用户活动和词汇分析),以自动推断与特定任务或工作上下文相关联的上下文关系。上下文系统和方法提供对特定用户活动(包括如上文所提及的文档切换、复制粘贴操作、插入附件或链接,和/或设置书签)的监测,从该监测可以得出推断。以此方式,上下文系统和方法产生表示特定用户任务或工作上下文的相关项的高精度的结果。

[0066] 代替简单地提供一般基于关键字的搜索的结果,上下文系统和方法基于与用户先前看见并处理的信息人工制品的交互,发现与给定工作上下文或任务有关的项。以此方式,上下文系统和方法可以帮助用户比以别的方式更快地返回到给定任务的相关内容。上下文系统和方法可以帮助用户展示可能原本已经被忘记的与特定任务有关的内容。如此,上下文系统和方法可以允许用户看一眼便可直观地感觉到候选项是否具有高价值。上下文系统和方法在与横跨多个工作会话、相隔超过一两天、以及多个应用程序的工作上下文相关联的任务中,以及在用户没有存档或记录看到或完成的作为任务的一部分的一切的情况下特别有用。

[0067] 如在本申请中所使用的,术语“组件”和“系统”旨在表示计算机相关的实体,其可以是硬件、硬件和软件的组合、软件、或者执行中的软件。例如,组件可以是,但不限于,在处理器上运行的进程、处理器、硬盘驱动器、多个存储驱动器(光学,固态和/或磁存储介质)、对象、可执行程序、运行的线程、程序、和/或计算机。作为说明,在服务器上运行的应

用和该服务器两者都可以是组件。一个或多个组件可以驻留在进程和 / 或执行的线程内, 且组件可以位于一个计算机上和 / 或分布在两个或更多的计算机之间。在此可使用词语“示例性”意指用作示例、实例或说明。本文中被描述为“示例性”的任何方面或设计不一定被解释为优选于或更优于其他方面或设计。

[0068] 现在参考图 11, 所示是根据所公开的体系结构的可用于执行上下文关系的推断的计算系统 1100 的框图。为了提供用于其各方面的附加上下文, 图 11 及以下讨论旨在提供对其中可实现各方面的合适的计算系统 1100 的简要概括描述。尽管以上描述是在可在一个或多个计算机上运行的计算机可执行指令的一般上下文中进行的, 但是本领域的技术人员将认识到, 新颖实施例也可结合其它程序模块和 / 或作为硬件和软件的组合来实现。

[0069] 用于实现各方面的计算系统 1100 包括计算机 1102, 其具有处理单元 1104、系统存储器 1106、以及系统总线 1108。处理单元 1104 可以是各种市场上可买到的处理器中的任一种, 诸如单处理器、多处理器、单核单元以及多核单元。此外, 本领域的技术人员可以理解, 各新颖方法可用其它计算机系统配置来实施, 包括小型机、大型计算机、以及个人计算机 (例如, 台式、膝上型等)、手持式计算设备、基于微处理器的或可编程的消费电子产品等, 其每一个都可在操作上耦合到一个或多个相关联的设备。

[0070] 系统存储器 1106 可包括易失性 (VOL) 存储器 1110 (例如, 随机存取存储器 (RAM)) 和非易失性存储器 (NON-VOL) 1112 (如 ROM、EPROM、EEPROM 等)。基本输入 / 输出系统 (BIOS) 可被存储在非易失性存储器 1112 中, 并且包括诸如在启动期间便于在计算机 1102 内的组件之间传递数据和信号的基本例程。易失性存储器 1110 还可包括诸如静态 RAM 等高速 RAM 来用于高速缓存数据。

[0071] 系统总线 1108 提供了用于包括, 但不限于存储器子系统 1106 的系统组件对处理单元 1104 的接口。系统总线 1108 可以是若干种总线结构中的任一种, 这些总线结构还可使用各类可购买到的总线体系结构中的任一种互连到存储器总线 (带有或没有存储器控制器) 以及外围总线 (例如, PCI、PCIe、AGP、LPC 等)。

[0072] 计算机 1102 还包括存储子系统 1114 以及将存储子系统 1116 接口到系统总线 1114 和其他所需计算机组件的存储接口 1108。存储子系统 1114 可包括例如硬盘驱动器 (HDD)、磁软盘驱动器 (FDD) 和 / 或光盘存储驱动器 (例如, CD-ROM 驱动器、DVD 驱动器) 中的一个或多个。存储接口 1116 可包括诸如, 例如 EIDE、ATA、SATA 和 IEEE 1394 等接口技术。

[0073] 一个或多个程序和数据可被存储在存储器子系统 1106、可移动存储器子系统 1118 (例如, 闪存驱动器形状因子技术) 和 / 或存储子系统 1114 (例如, 光、磁、固态) 中, 包括操作系统 1120、一个或多个应用程序 1122、其他程序模块 1124 以及程序数据 1126。

[0074] 一般而言, 程序模块包括执行特定任务或实现特定抽象数据类型的例程、方法、数据结构、其他软件组件等等。操作系统 1120、应用 1122、模块 1124 和 / 或数据 1126 的全部或部分也可被高速缓存在诸如易失性存储器 1110 等存储器中。应该明白, 所公开的体系结构可以用各种市场上可购得的操作系统或操作系统的组合 (例如, 作为虚拟机) 来实施。

[0075] 如前所述的应用程序 1122、程序模块 1124 以及程序数据 1126 可包括图 1 的计算机实现的系统 100、活动组件 102、活动信息 104、项 106、工作上下文 108、分析组件 110、参考项 112、推断组件 114、候选项 116 以及结果 118, 图 2 的系统 200, 该系统 200 包括进一步

的额外的组件,如后台组件 202 和共同的词 204,图 3 的用户交互 300,包括切换操作 302、复制粘贴操作 304、插入操作 306、切换频率测量 308、时间戳操作 310、设置书签操作 312、链接操作 314、保存操作 316,以及停留时间 318,图 4 的系统 400,该系统 400 包括进一步的额外的组件,如查询组件 402、收集组件 404 以及存储组件 406。

[0076] 如前所述的应用程序 1122、程序模块 1124 以及程序数据 1126 还可以包括,例如,系统 500,该系统 500 包括额外的组件,如图 5 的信息人工制品 502、参考人工制品 504、词汇相似处 506 以及呈现组件 508,图 6 的用户可识别和可使用的数据实体 600、文件 602、数据流 604、网页 606、电子表格 608、电子邮件消息 610、即时消息谈话 612、日历约会 614、便签 616 以及嵌入式元数据 618,图 7 的系统 700,该系统 700 包括进一步的额外的组件,如分数分配 702、加权组件 704 以及可操作的菜单元件 706,以及图 8 的实现 800,该实现 800 包括参考项 802、候选项 1 804、候选项 2 806、候选项 3 808 以及候选项 4 810,以及图 9-10 的方法。

[0077] 存储子系统 1114 和存储器子系统 (1106 和 1118) 用作用于数据、数据结构、计算机可执行指令等的易失性和非易失性存储的计算机可读介质。计算机可读介质可以是可以被计算机 1102 访问的任何可用的介质,并包括易失性和非易失性介质、可移动的和不可移动的介质。对于计算机 1102,介质容纳以任何合适的数字格式对数据的存储。本领域的技术人员应当理解,可使用其他类型的计算机可读介质,如 zip 驱动器、磁带、闪存卡、磁带盒等来存储用于执行所公开的体系结构的新颖方法的计算机可执行指令。

[0078] 用户可以使用诸如键盘和鼠标等外部用户输入设备 1128 来与计算机 1102、程序和数据交互。其他外部用户输入设备 1128 可包括话筒、IR(红外)遥控器、操纵杆、游戏手柄、照相机识别系统、指示笔、触摸屏、姿势系统(例如,眼移动、头移动等)和/或类似物。在计算机 1102 是例如便携式计算机的情况下,用户可以使用诸如触摸垫、话筒、键盘等板载用户输入设备 1130 来与计算机 1102、程序和数据交互。这些及其他输入设备常常通过系统总线 1104 并通过输入/输出(I/O)设备接口 1132 连接到处理单元 1108,但是,也可以通过其他接口,如并行端口、IEEE 1394 串行端口、游戏端口、USB 端口、IR 接口等等来进行连接。I/O 设备接口 1132 还便于使用输出外围设备 1134,如打印机、音频设备、照相机设备等,如声卡和/或板载音频处理能力。

[0079] 一个或多个图形接口 1136(通常也称为图形处理单元(GPU))提供计算机 1102 和外部显示器 1138(例如,LCD、等离子)和/或板载显示器 1140(例如,对于便携式计算机)之间的图形和视频信号。图形接口 1136 也可作为计算机系统板的一部分来制造。

[0080] 计算机 1102 可以使用经由有线/无线通信子系统 1142 到一个或多个网络和/或其他计算机的逻辑连接在联网环境(例如,基于 IP 的)中操作。其他计算机可包括工作站、服务器、路由器、个人计算机、基于微处理器的娱乐设备、对等设备或其他常见的网络节点,并且通常包括以上相对于计算机 1102 描述的许多或所有元件。逻辑连接可包括到局域网(LAN)、广域网(WAN)热点等的有线/无线连接。LAN 和 WAN 联网环境常见于办公室和公司,并且方便了诸如内联网等企业范围计算机网络,所有这些都可连接到例如因特网等全球通信网络。

[0081] 当在联网环境中使用时,计算机 1102 经由有线/无线通信子系统 1142(例如,网络接口适配器、板载收发机子系统等)连接到网络来与有线/无线网络、有线/无线打印

机、有线 / 无线输入设备 1144 等通信。计算机 1102 可包括用于通过网络建立通信的调制解调器或其他装置。在联网环境中,相对于计算机 1102 的程序和数据可被存储在远程存储器 / 存储设备中,如与分布式系统相关联。应该理解,所示网络连接是示例性的,并且可以使用在计算机之间建立通信链路的其他手段。

[0082] 计算机 1102 可用于使用诸如 IEEE802. xx 标准家族等无线电技术来与有线 / 无线设备或实体通信,例如在操作上安置在与例如打印机、扫描仪、台式和 / 或便携式计算机、个人数字助理 (PDA)、通信卫星、任何一件与无线可检测标签相关联的设备或位置 (例如,电话亭、报亭、休息室) 以及电话的无线通信 (例如,IEEE802. 11 空中调制技术) 中的无线设备。这至少包括对于热点的 Wi-Fi (或无线保真)、WiMax, 以及 Bluetooth™ 无线技术。由此,通信可以是如对于常规网络那样的预定义结构,或者仅仅是至少两个设备之间的自组织 (ad hoc) 通信。Wi-Fi 网络使用称为 IEEE802. 11x (a、b、g 等) 的无线电技术来提供安全、可靠、快速的无线连接。Wi-Fi 网络可用于将计算机彼此连接、连接到因特网以及连接到有线网络 (使用 IEEE802. 3 相关介质和功能)。

[0083] 所示各方面也可以在其中某些任务由通过通信网络链接的远程处理设备来执行的分布式计算环境中实践。在分布式计算环境中,程序模块可以位于本地和 / 或远程存储器和 / 或存储设备中。

[0084] 上面描述的包括所公开的体系结构的各示例。当然,描述每一个可以想到的组件和 / 或方法的组合是不可能的,但本领域内的普通技术人员应该认识到,许多其他组合和排列都是可能的。因此,该新颖体系结构旨在涵盖所有这些落入所附权利要求书的精神和范围内的更改、修改和变化。此外,就在说明书或者权利要求书中使用的术语“包括”而言,这一术语旨在以与术语“包括”在用作权利要求书中的过渡词时所解释的相似的方式是包含性的。

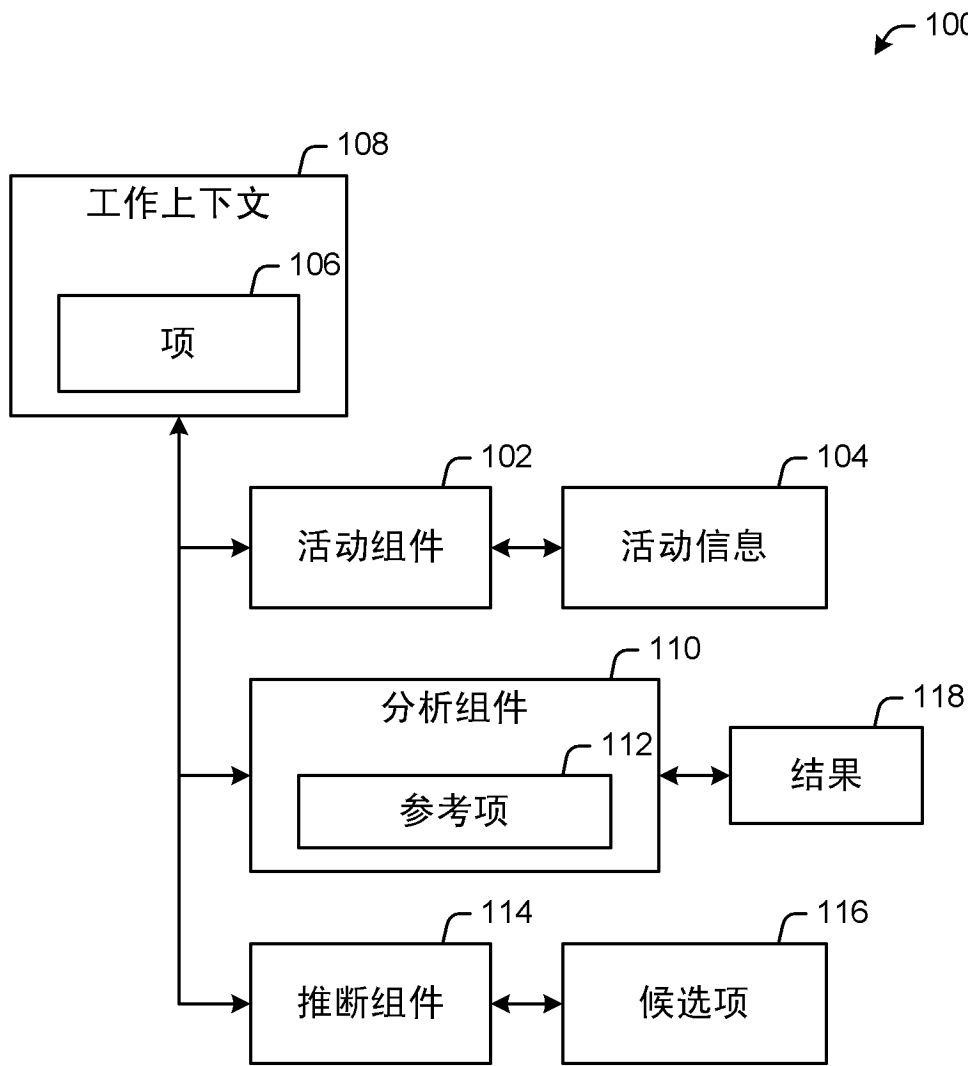


图 1

200

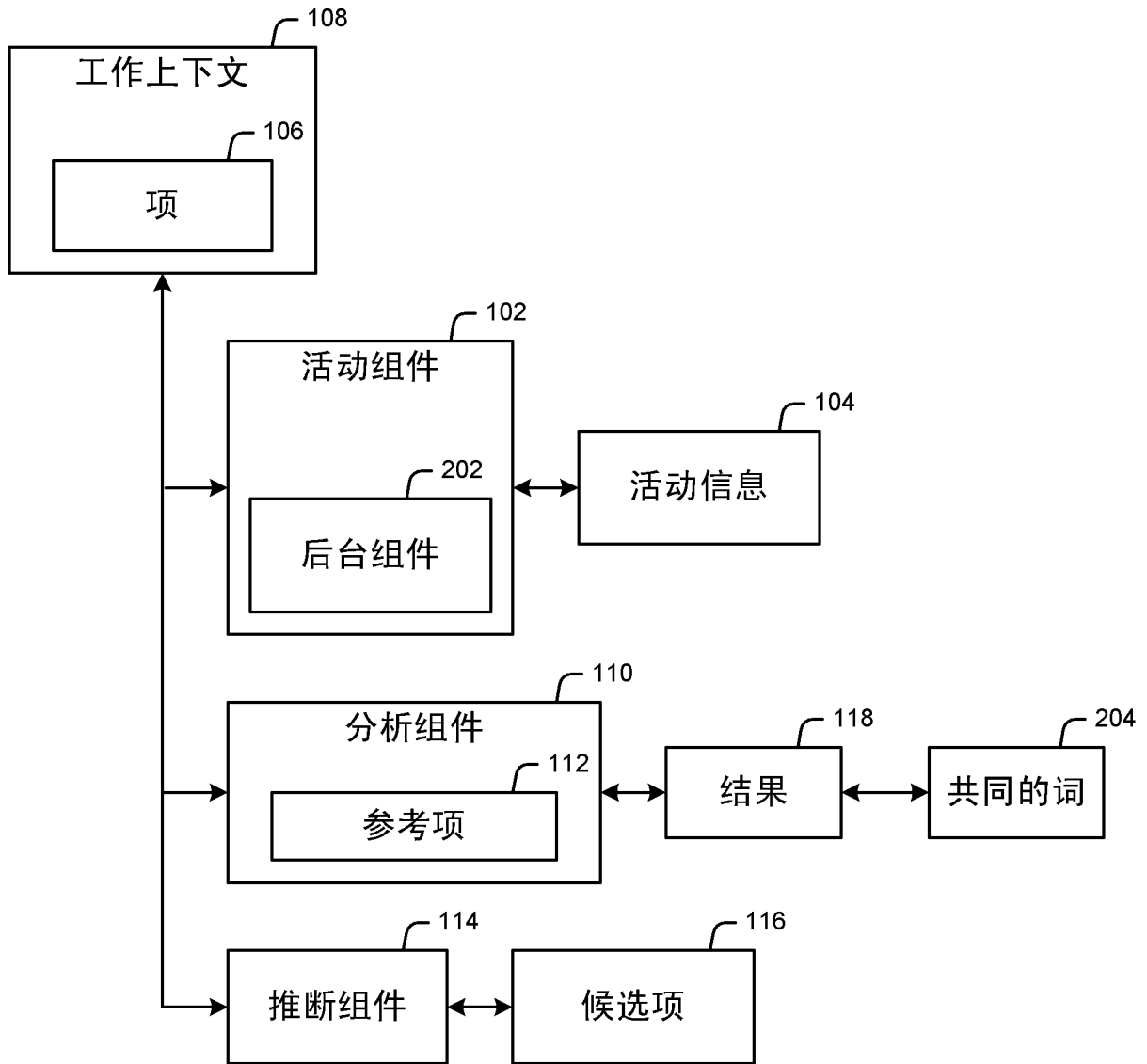


图 2

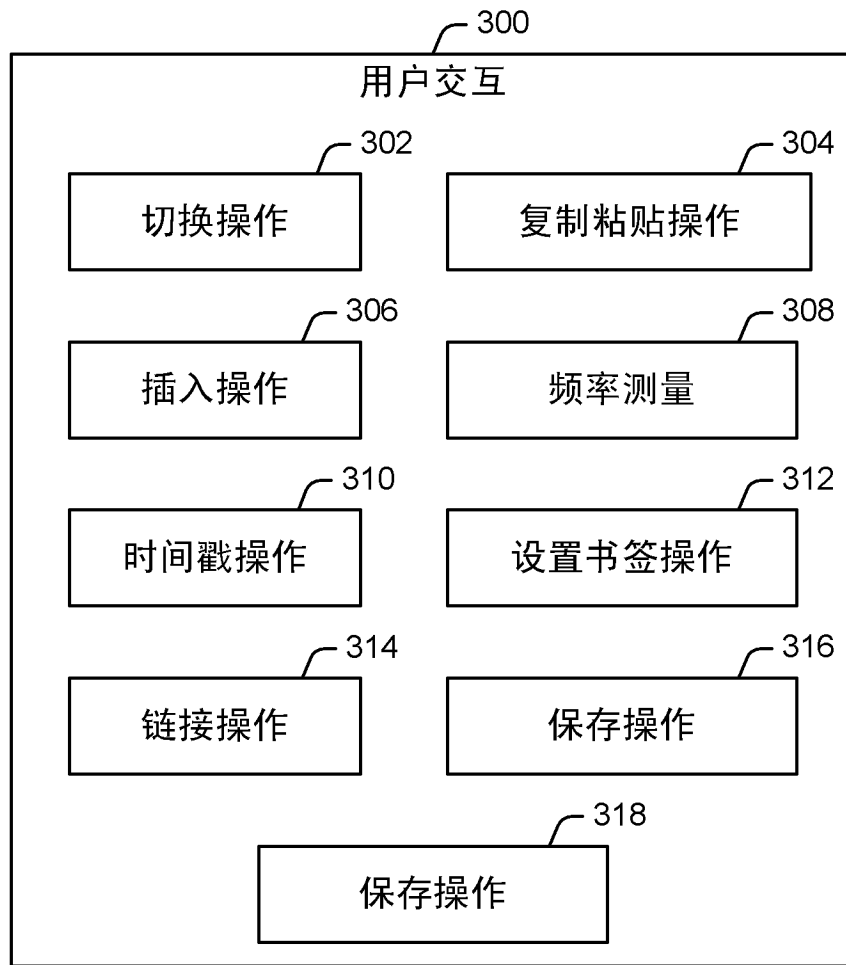


图 3

400

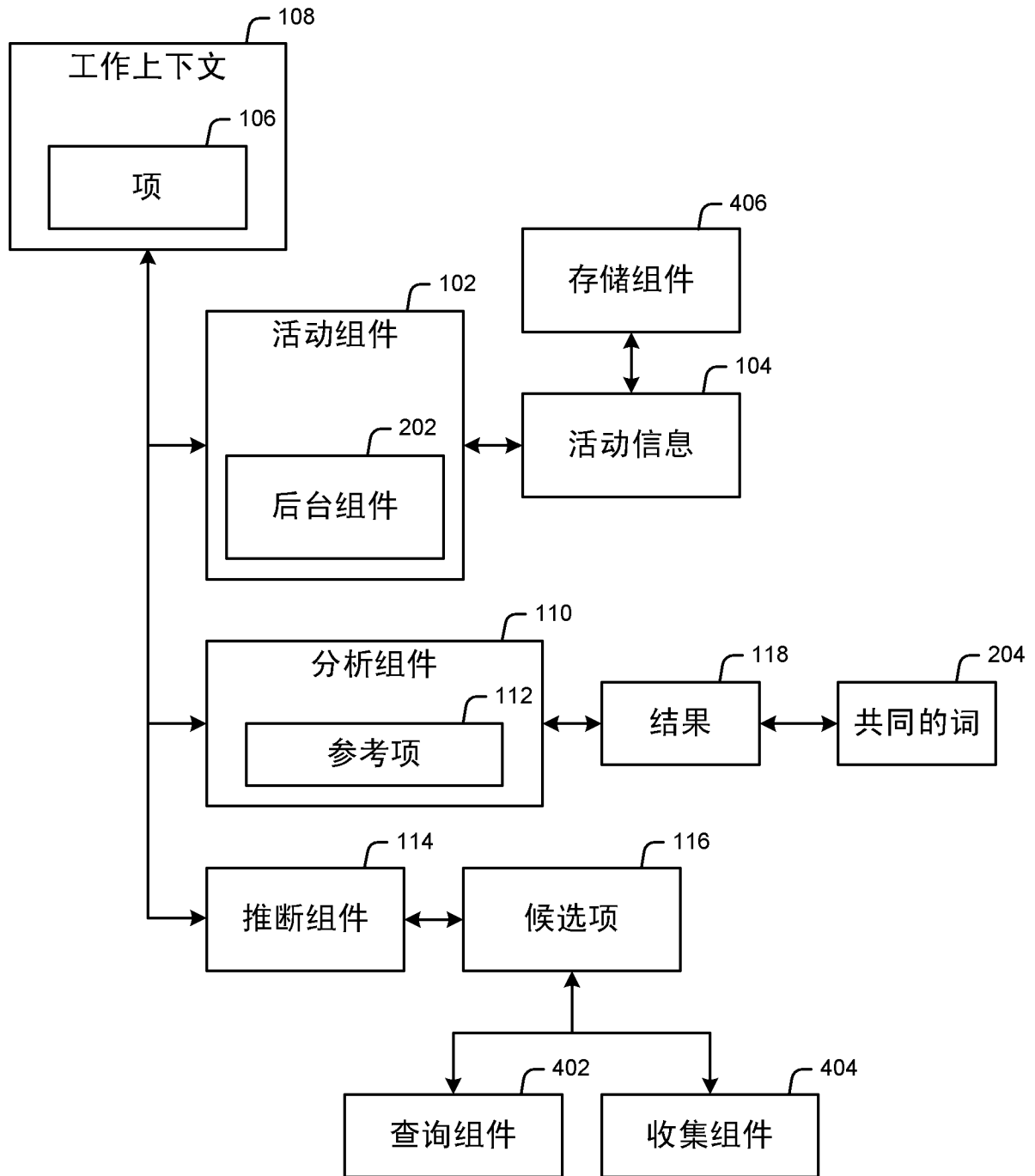


图 4

500

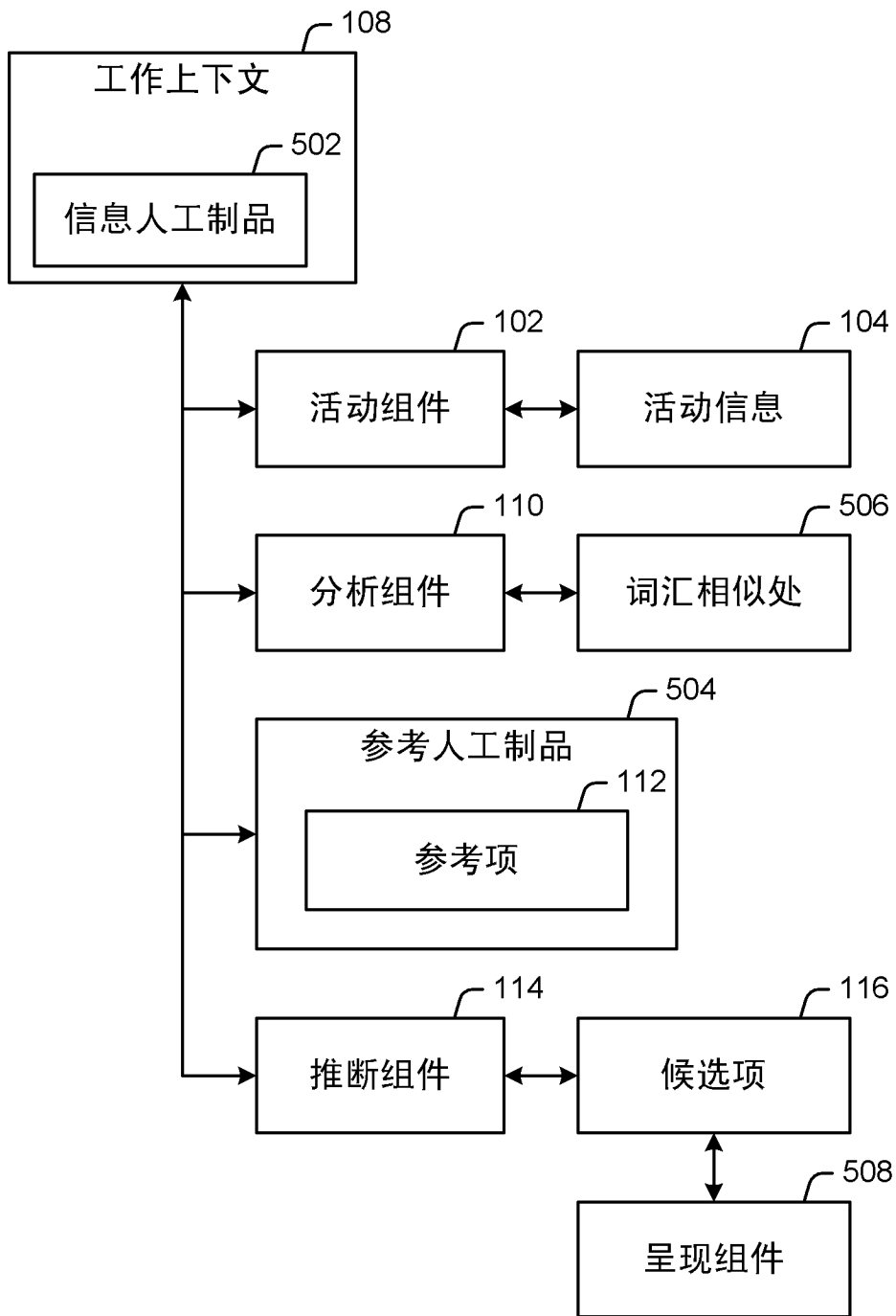


图 5

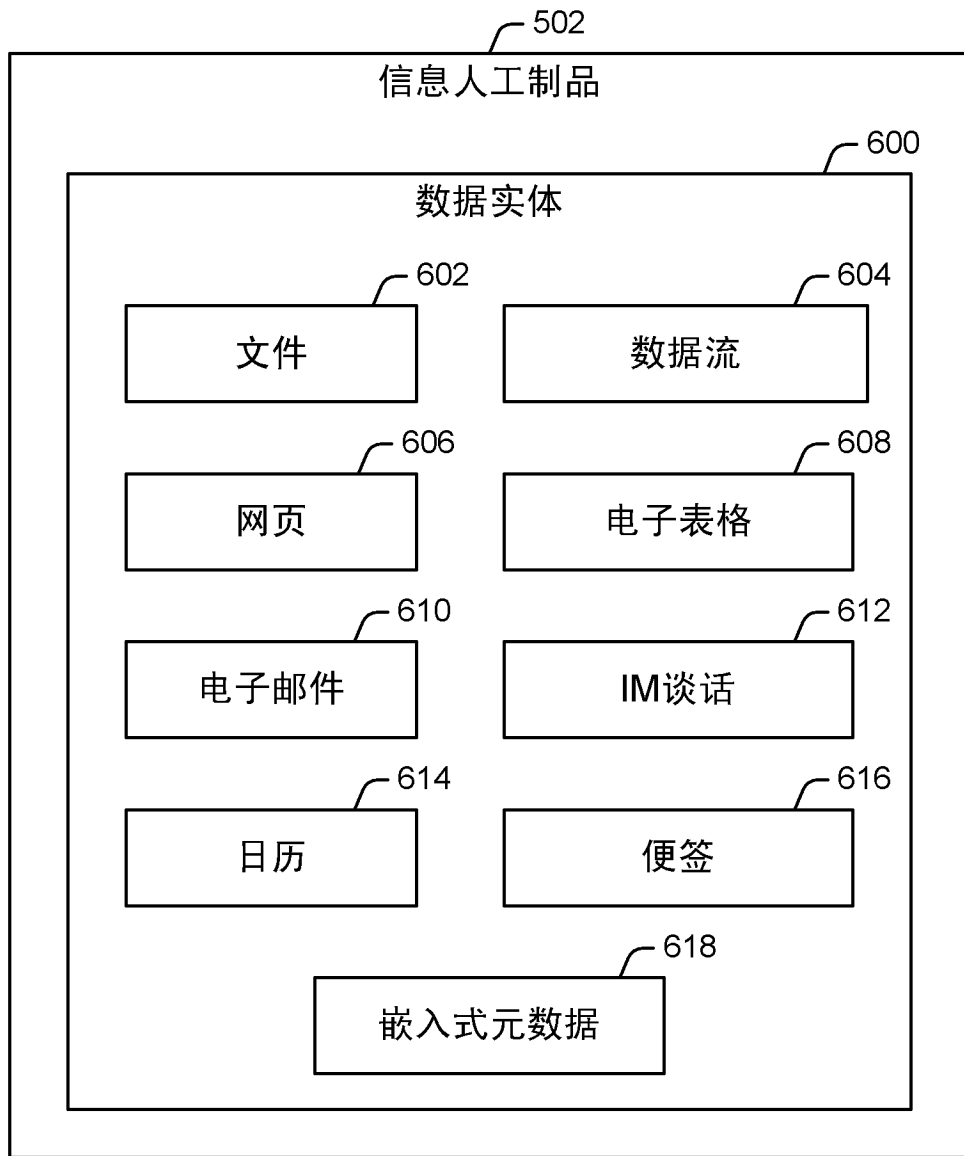


图 6

700

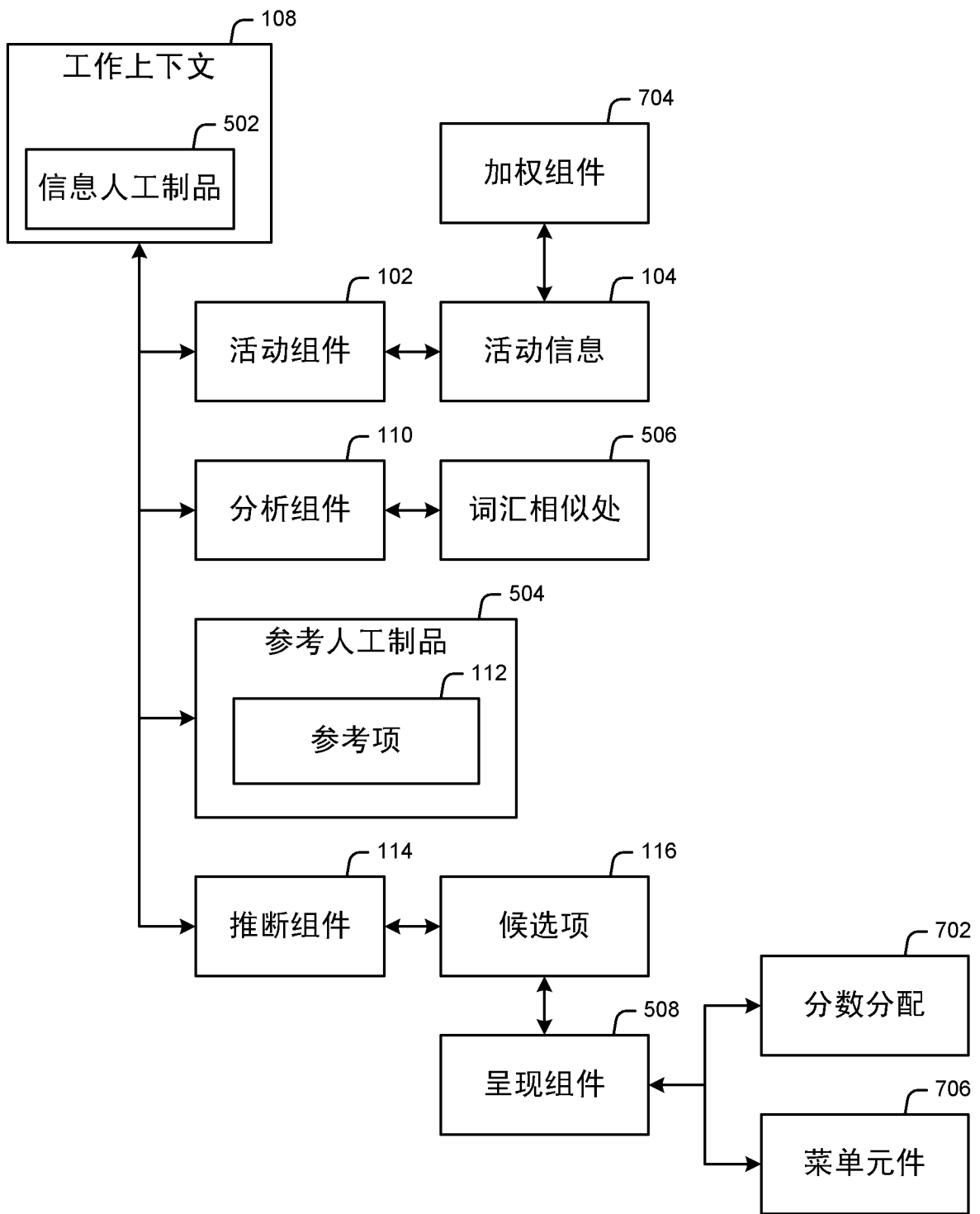


图 7

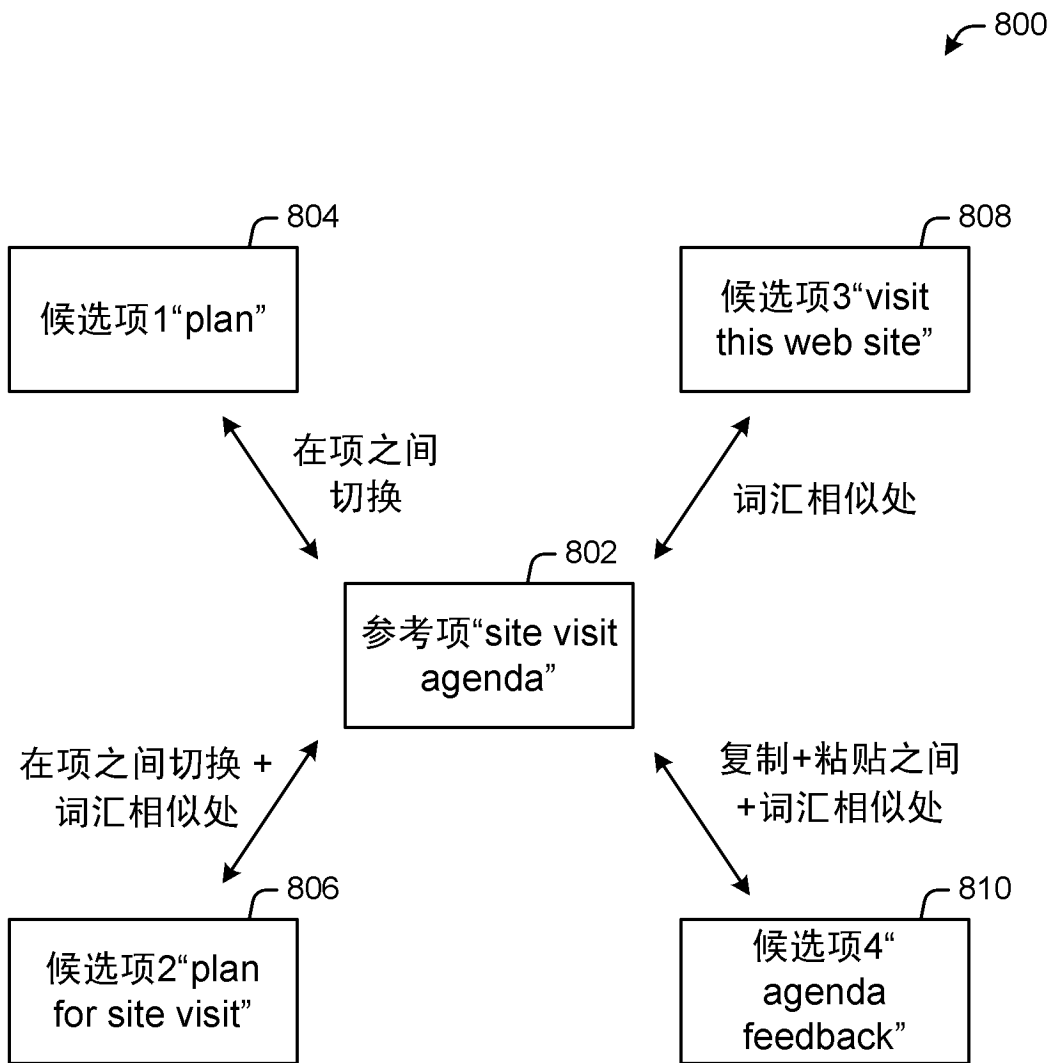


图 8

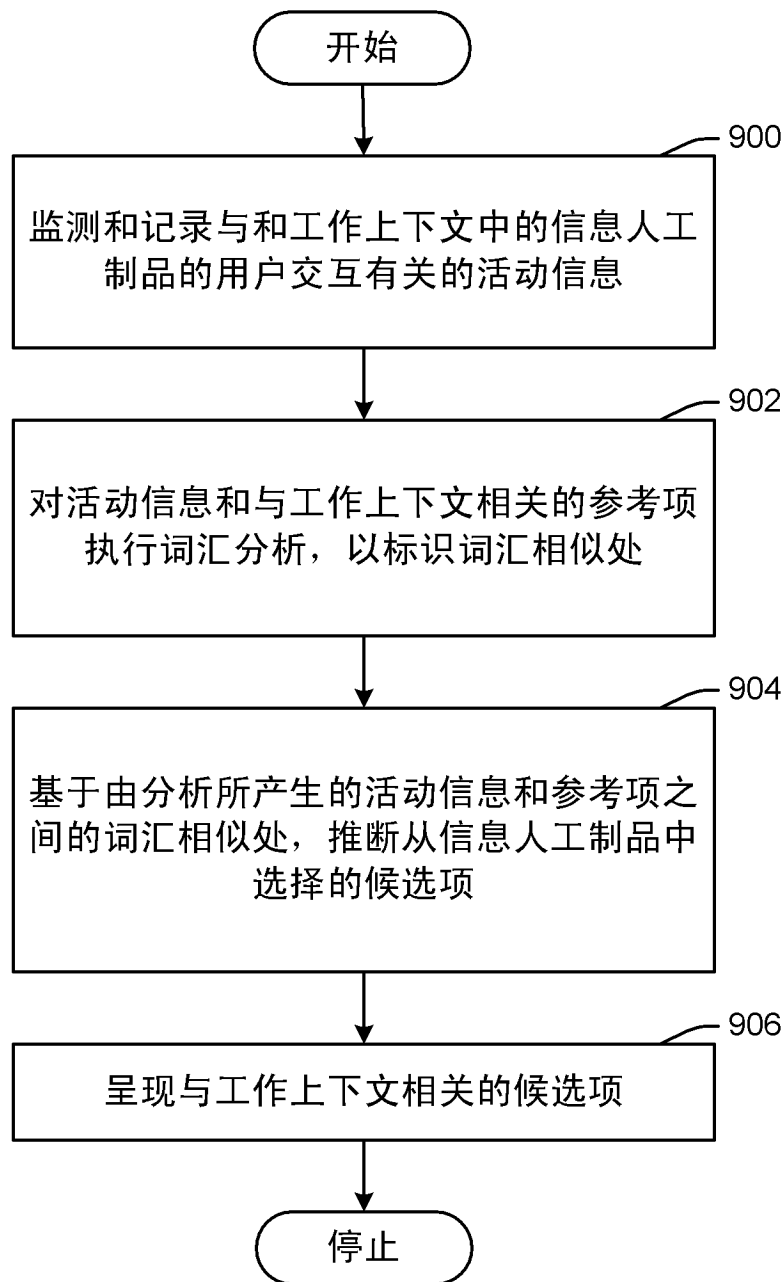


图 9

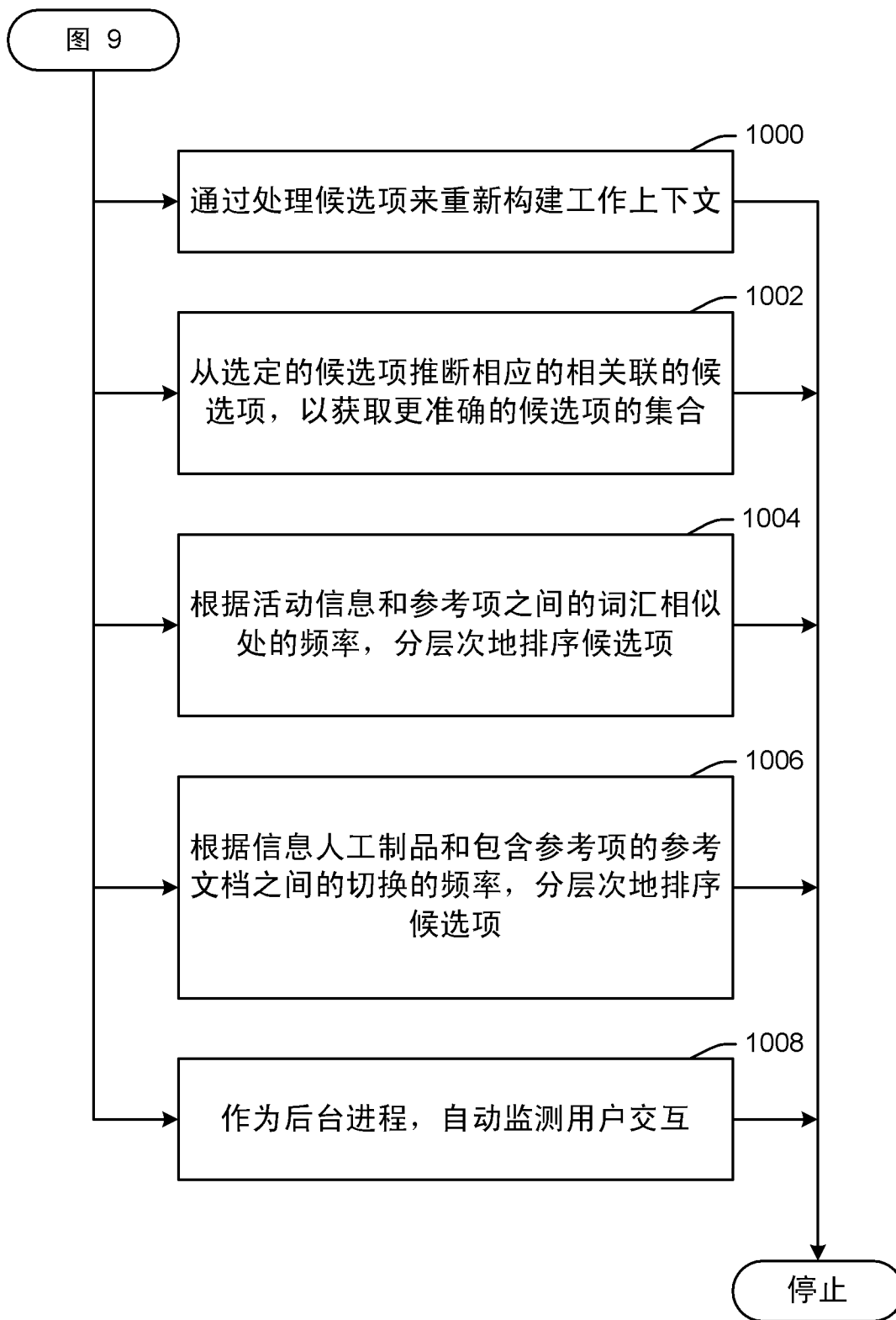


图 10

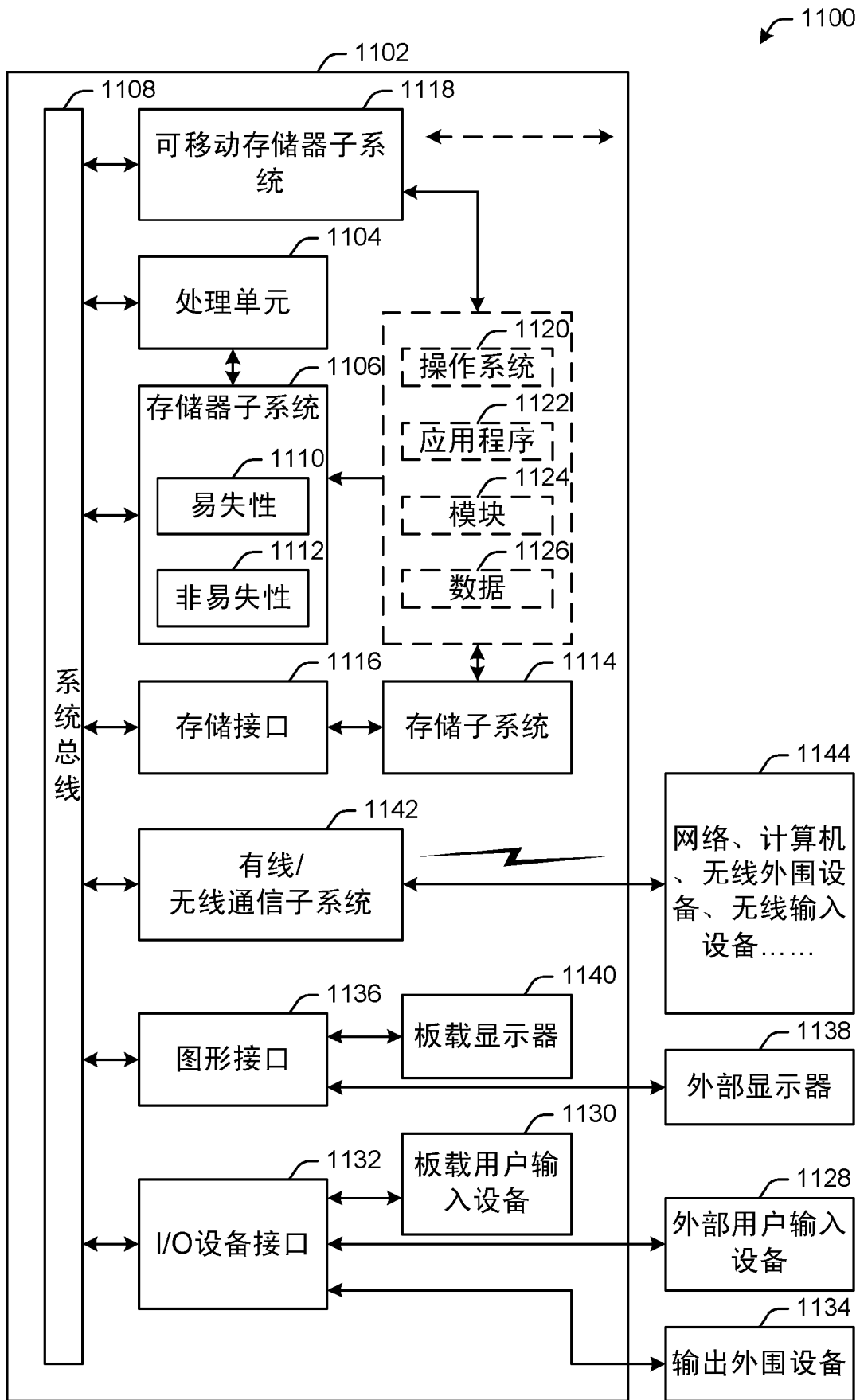


图 11