

(12) 发明专利申请

(10) 申请公布号 CN 102662940 A

(43) 申请公布日 2012. 09. 12

(21) 申请号 201110424241. 5

(22) 申请日 2011. 12. 13

(30) 优先权数据

2010-278012 2010. 12. 14 JP

(71) 申请人 日本电气株式会社

地址 日本东京都

(72) 发明人 今井祥子

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 吴立明

(51) Int. Cl.

G06F 17/30(2006. 01)

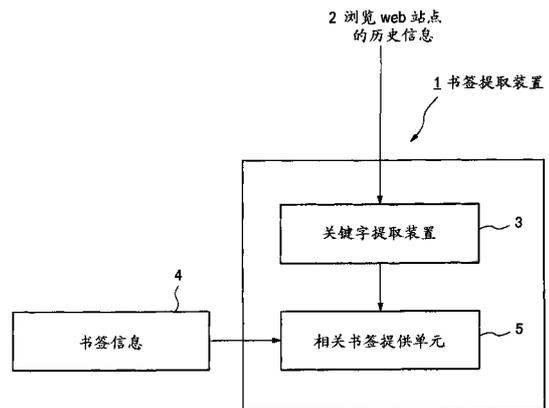
权利要求书 2 页 说明书 20 页 附图 23 页

(54) 发明名称

书签提取装置、方法以及计算机产品

(57) 摘要

本发明涉及一种书签提取装置、方法和计算机程序。公开了书签提取装置，并公开了该装置从提前注册的书签中准确选择与当前正在浏览的 web 站点相关的最适合的书签，并将它们提供给用户。该书签提取装置包括关键字提取单元，该提取单元基于 web 站点至今的浏览历史信息，提取关键字，以及提供单元，该提供单元基于所述关键字，从多个已注册的书签中提供与所述提取单元所提取的所述关键字相关的书签。



1. 一种书签提取装置,包括:
关键字提取装置,用于基于 web 站点至今的浏览历史信息提取关键字;以及
提供单元,用于基于所述关键字,从多个注册的书签中提供与所述关键字提取单元提取的所述关键字相关的书签。
2. 根据权利要求 1 所述的书签提取装置,其中
基于在所述浏览历史信息中包括的、关于当前正在浏览的 web 站点的信息,所述关键字提取单元获取所述当前浏览的 web 站点的内容,并且从所述内容中提取所述关键字。
3. 根据权利要求 2 所述的书签提取装置,其中
所述关键字提取单元包括重要关键字提取单元,用于提取在从所述当前浏览的 web 站点获取的所述内容中包括的重要关键字作为所述关键字,并且其中
所述提供单元基于所述重要关键字,从所述多个书签中提供与所述重要关键字相关的书签。
4. 根据权利要求 3 所述的书签提取装置,其中
所述重要关键字提取单元根据在所述内容中出现的频率确定所述重要关键字。
5. 根据权利要求 1 所述的书签提取装置,其中,
通过分析所述浏览历史信息,当当前正在浏览的 web 站点被确定为已经从搜索站点进行了转变时,所述关键字提取单元提取在到所述当前浏览的 web 站点的所述转变之前已经使用的搜索关键字作为所述关键字,并且其中
所述提供单元基于所述搜索关键字,从所述多个书签中提供与所述搜索关键字相关的书签。
6. 根据权利要求 5 所述的书签提取装置,其中,
当恰好到所述当前浏览的 web 站点的所述转变之前已经被浏览的 web 站点是搜索站点时,所述关键字提取单元提取已经在所述搜索站点中使用的搜索关键字作为所述关键字。
7. 根据权利要求 1 所述的书签提取装置,其中
所述提供单元包括相关性确定单元,用于根据在每个书签的所述属性信息中所述关键字被包括的程度,确定在所述关键字和在关于所述多个书签的信息中包括的书签中的所述每个书签之间的相关性的存在。
8. 根据权利要求 1 所述的书签提取装置,其中
所述提供单元包括相关性确定单元,用于根据与关注的所述书签链接的 web 站点的所述内容中的所述关键字被包括的程度,确定在所述关键字和所述每个书签之间相关性的存在。
9. 根据权利要求 1 所述的书签提取装置,其中
所述提供单元包括相关性确定单元,用于根据在所述每个书签的所述属性信息中的所述关键字被包括的程度以及与关注的所述书签链接的 web 站点的所述内容中的所述关键字被包括的程度的组合,确定在所述关键字和所述每个书签之间相关性的存在。
10. 根据权利要求 7 所述的书签提取装置,其中
所述提供单元以排序的方式,提供已经由所述相关性确定单元确定为具有相关性的多个书签。
11. 根据权利要求 7 所述的书签提取装置,其中,

基于所述相关性确定单元确定的结果,所述提供单元提供给用户与来自所述书签注册的时间的所述书签对应的 web 站点的内容改变的可能性的信息,或者到所述 web 站点的链接失效的信息。

12. 一种书签提取方法,包括:

基于 web 站点至今的浏览历史信息提取关键字;以及

基于所述关键字,从多个提前注册的书签中提供与所述提取的关键字相关的书签。

13. 根据权利要求 12 所述的书签提取方法,其中,

当提取所述关键字时,基于在所述浏览历史信息中包括的、关于当前浏览的 web 站点的信息,获取所述当前浏览的 web 站点的内容,从所述内容中提取所述关键字。

14. 根据权利要求 11 所述的书签提取方法,其中,

当提取所述关键字时,作为分析所述浏览历史信息的结果,在当前浏览的 web 站点被确定为已经从搜索站点进行了转变的情况下,提取在到所述当前浏览的 web 站点的所述转变之前已经使用的搜索关键字作为所述关键字,并且其中

基于所述搜索关键字,从所述多个书签中提供与所述搜索关键字相关的书签。

15. 根据权利要求 14 所述的书签提取方法,其中,

当提取所述关键字时,在恰好到所述当前浏览的 web 站点的所述转变之前已经被浏览的 web 站点是搜索站点的情况下,提取已经在所述搜索站点中使用的搜索关键字作为所述关键字。

16. 一种用于控制书签提取装置的操作的计算机程序,所述计算机程序使得计算机实现:

关键字提取功能,用于基于 web 站点至今的浏览历史信息提取关键字;以及

提供功能,用于基于所述关键字,从多个注册的书签中提供与所述关键字提取功能所提取的所述关键字相关的书签。

17. 根据权利要求 16 所述的计算机程序,其中,

基于在所述浏览历史信息中包括的、关于当前浏览的 web 站点的信息,所述关键字提取功能获取所述当前浏览的 web 站点的内容,并且从所述内容中提取所述关键字。

18. 根据权利要求 16 所述的计算机程序,其中,

作为分析所述浏览历史信息的结果,当当前浏览的 web 站点被确定为已经从搜索站点进行了转变时,所述关键字提取功能提取在到所述当前浏览的 web 站点的所述转变之前已经使用的搜索关键字作为所述关键字,并且其中

所述提供功能基于所述搜索关键字,从所述多个书签中提供与所述搜索关键字相关的书签。

19. 根据权利要求 18 所述的计算机程序,其中,

当恰好到所述当前浏览的 web 站点的所述转变之前浏览的 web 站点是搜索站点的情况下,所述关键字提取功能提取已经在所述搜索站点中使用的搜索关键字作为所述关键字。

20. 一种书签提取装置,包括:

关键字提取装置,用于基于 web 站点至今的浏览历史信息提取关键字;以及

提供装置,用于基于所述关键字,从多个注册的书签中提供与由所述提取装置提取的所述关键字相关的书签。

书签提取装置、方法以及计算机产品

[0001] 本申请基于并且要求 2010 年 12 月 14 日提交的日本专利申请第 2010-278012 号的优先权,通过引用将上述申请的全部公开内容合并于此。

技术领域

[0002] 本发明涉及 web 站点的浏览技术。特别地,本发明涉及在浏览 web 站点时从多个注册的书签为用户提供最合适书签的 web 站点浏览技术。

背景技术

[0003] 使用浏览器(浏览器软件)来浏览国际互联网 web 站点(以下,其可以被称为“web 站点”,或“web 页面”)在人们中广为流行。一般的浏览器具有所谓的书签功能。根据书签功能,通过在浏览器中按照与用户想要的标题相关联的方式存储 web 页面的 URL(统一资源定位符),下一次可以实现用户对所需 web 页面的方便访问。然而,在一般情况下,当浏览器使用了很长的时间段,注册到浏览器的书签的数量趋于增加。在这样的书签功能中,也具有一种已知的功能,其中假设上述情况,用户本身将 web 页面分类(组织)到文件夹,提供用户所期望的文件夹名称。

[0004] 然而,在这种浏览器的实际使用场景中,大量的书签已经被提前注册,更加困难将是用户立即访问对应于所期望 web 页面的书签。甚至当通过分类文件夹中注册的书签时,类似地发生这种趋势。这样的结果是,在这种情况下,即使将大量书签注册到浏览器,出现大部分出现不会被有效地使用的情况。

[0005] 也有已知的具有从多个注册的书签搜索想要的书签的功能的浏览器。然而,为了充分利用这种搜索功能,用户必须记住之前已经注册了所需的书签的事实,而且进一步,必须记住所期望的书签的标题或类似的以将其用作搜索关键字。出于这个原因,用户不容易有效地使用这样的搜索功能,因此它是不方便的。

[0006] 因此,为了改善这种一般浏览器功能和它的书签功能的便利,已经提出了根据用户位置在浏览器屏幕上容易地可访问的(可操作的)位置处指示用户将要访问的书签的技术。如这种技术的一个例子,在专利文件 1(日本专利申请公开号 2007-299170)中,公开了一种技术是,为了促进有效使用大量注册的书签,基于正在浏览的 web 页面的 URL 提取与用户想知道的项有关的书签,并且将它们提供给用户。

[0007] 也就是说,专利文件 1 公开了一种从管理器数据库获取与正在浏览的 web 页面相关性的书签技术,在管理数据库中将 URL 与和 URL 内容相关的书签进行关联。也就是说,在专利文件 1 中,这种管理数据库是由不同于先前的第三人的第三人准备的。作为向管理数据库查询目前正在浏览的 web 页面的 URL 的结果,用户获得相关的 URL 组。然后,在专利文件 1 中,如果落入所获得的 URL 组的书签已经注册了用户使用的浏览器的书签,可以将这些书签的列表提供给用户。在专利文件 2(日本专利申请公开号 2004-070441)中,公开了一种技术,其为了在最合适的时机提供可用于用户的 web 页面的内容的目的,根据时间和用户的位置从注册的书签提供最合适的书签。

[0008] 然而,在专利文件 1 公开的技术中,提前执行将数据注册到管理数据库中。因此,已注册的信息是永远比用户访问时的信息要老的信息。出于这个原因,在专利文件 1 所披露的技术中,它不能立即处理在书签的变化和相关 URL 的内容变化。同样,仅基于 web 页面内容自动提取书签时,可以提取与用户的需要不同的书签,例如当用户想要了解的信息仅由 web 页面的较少部分指示时。

[0009] 此外,在专利文件 1 所公开的技术中,执行按使用分类的 URL 关联。。出于这个原因,根据使用分类的设置,还具有执行不对应于 web 页面内容的分类的可能性。

[0010] 在专利文献 2 所披露的技术中,用户目前正在浏览的 web 页面的内容不被视为提供给用户的书签。

[0011] 因此,本发明的主要目的是提供一种书签提取装置等,其从现在注册的书签中准确地选择与目前正在浏览的 web 页面相关性的最合适的书签,并将他们提供给用户。

发明内容

[0012] 为了实现该目标,根据本发明的书签提取装置具有下面配置。

[0013] 即,作为本发明的一个方面,一种书签提取装置,包括:

[0014] 关键字提取装置,用于基于 web 站点至今的浏览历史信息提取关键字;以及

[0015] 提供单元,用于基于所述关键字,从多个注册的书签中提供与所述关键字提取单元提取的所述关键字相关的书签。

[0016] 在上述情况下,例如基于在所述浏览历史信息中包括的、关于当前正在浏览的 web 站点的信息,所述关键字提取单元获取所述当前浏览的 web 站点的内容,并且从所述内容中提取所述关键字。

[0017] 作为本发明的另一方面,在上述书签提取装置中,当通过分析所述浏览历史信息,确定已经从搜索站点进行了到当前正在浏览的 web 站点的转变时,所述关键字提取单元可以提取在到所述当前浏览的 web 站点的所述转变之前已经使用的搜索关键字作为所述关键字,并且所述提供单元可以基于所述搜索关键字,从所述多个书签中提供与所述搜索关键字相关的书签。

[0018] 在任何上述配置中,所述提供单元可以包括相关性确定单元,用于例如根据所述关键字在包括在关于所述多个书签的信息中的每个所述书签的所述属性信息中所述关键字被包括的程度,确定在所述关键字和所述每个书签的之间的相关性的存在。

[0019] 同时,还使用计算机,通过计算机程序以及通过其中存储了所述计算机程序的计算机可读存储器介质来获得该目标,该计算机程序实现具有前述配置的书签提取装置和对应的方法。

[0020] 根据本发明,实现了提供书签提取装置等,其从提前注册的书签中准确地选择与正在浏览的 web 站点相关的最适合的书签,并且将它们提供给用户。

附图说明

[0021] 结合附图且通过下面的详细说明,本发明的示例性特征和优点将更为明显,其中:

[0022] 图 1 是示出了根据本发明的第一示例性实施方式的书签提取装置的结构框图;

[0023] 图 2 是示出了根据本发明的第二示例性实施方式的包括书签提取系统的通信环境的框图；

[0024] 图 3 是示出了根据本发明的第二示例性实施方式的由书签提取系统 100 处理的书签提取的流程图；

[0025] 图 4 是示出了根据第二实施方式的书签提取处理中图 3 所示的步骤 S102、步骤 S104、步骤 S105 和步骤 S110 中的处理的细节；

[0026] 图 5 是示出了根据第二实施方式的搜索词提取处理和用于关于其是否是搜索站点的单个历史的确定处理的流程图；

[0027] 图 6 是示出了在本发明的第二示例性实施方式中的相关性确定处理的流程图；

[0028] 图 7 是示出了基于概念的在搜索站点信息存储单元 106 中保持的搜索站点信息的示意图；；

[0029] 图 8A 和图 8B 是示例性介绍根据本发明的第二示例性实施方式的实例中用户进行网络浏览时的屏幕转换的示意图；

[0030] 图 9A 和图 9B 是概念地描述图 8A 和图 8B 中示出的屏幕转换情况中书签提取系统的操作的示意图；

[0031] 图 10A 是概念地说明与在浏览器 200 的书签存储单元 206 注册的关于多个书签的书签信息块之间的书签 1 的数据结构的示意图；

[0032] 图 10B 是概念地说明与在浏览器 200 的书签存储单元 206 注册的关于多个书签的书签信息块之间的书签 2 的数据结构的示意图；

[0033] 图 10C 是概念地说明与在浏览器 200 的书签存储单元 206 注册的关于多个书签的书签信息块之间的书签 3 的数据结构的示意图；

[0034] 图 10D 是概念地说明与在浏览器 200 的书签存储单元 206 注册的关于多个书签的书签信息块之间的书签 4 的数据结构的示意图；

[0035] 图 10E 是概念地说明与在浏览器 200 的书签存储单元 206 注册的关于多个书签的书签信息块之间的书签 5 的数据结构的示意图；

[0036] 图 10F 是概念地说明与在浏览器 200 的书签存储单元 206 注册的关于多个书签的书签信息块之间的书签 6 的数据结构的示意图；

[0037] 图 11 是示出了根据本发明的第二示例性实施方式的相关书签的列表屏幕的显示形式的示意图；

[0038] 图 12 是示出了根据本发明的第三示例性实施方式的包括书签提取系统的通信环境的框图；

[0039] 图 13 是示出了根据本发明的第三示例性实施方式的由书签提取系统 100A 进行的书签提取处理的流程图；

[0040] 图 14 是示出了在本发明的第三示例性实施方式中的相关性确定处理中的流程图；

[0041] 图 15 是示出了在本发明的第四示例性实施方式的相关性确定处理中的流程图；

[0042] 图 16 是示出了在本发明的第四示例性实施方式的相关书签的列表屏幕的显示形式的示意图；

[0043] 图 17 是示例性地介绍了可以实现本发明的第一到第四示例性实施方式的计算机

(信息处理装置)的硬件配置的示意图。

具体实施方式

[0044] 下面,将参考附图来详细地介绍用于执行本发明的示例性实施方式。

[0045] <第一示例性实施方式>

[0046] 图1是示出了根据本发明的第一示例性实施方式的书签提取装置的结构框图。图1中示出的书签提取装置1具有关键字提取单元3和相关书签提取单元5。

[0047] 也就是说,关键字提取单元3基于到目前为止的用于web站点(web页面)的浏览历史信息2来提取关键字。浏览历史信息2包括与目前正在浏览的web站点相关的信息。

[0048] 基于关键字,相关书签提供单元5从关于书签(书签信息)4的多个注册信息块选择与关键字提取单元3提取的关键字相关的书签。然后,相关的书签提供单元5提供(或报告)所选的书签给用户。例如,浏览历史信息2和书签信息4可以从图1中未示出的网络浏览器(网络浏览装置)来获得。虽然,作为提供这些选定书签的形式,假设在显示器上的指示(图1中没有示出,但是在之后提到的实例中,可以采用图17中示出的计算机1000的显示器12)),例如,它不仅限于指示。

[0049] 根据以上所述的第一示例性实施方式,它是可以从先前注册的书签精确地选择与目前正在浏览的web站点相关的最适合的书签以提供给用户。也就是说,在根据这个示例性实施方式的书签提取装置1中,关键字提取单元3基于关于web站点的浏览到当前日期的浏览历史信息2来提取关键字,信息2包括关于正在浏览的web站点的信息。然后,通过指示所提取的关键字,书签提取装置1选择与关键字相关的书签并且将所选的书签提供给用户。因此,即使当由于具有大量的注册书签而导致用户不记得每个书签时,或者,甚至当由于已经通过将书签分类到类别来注册书签而导致用户已经忘记将书签注册到哪个类别。书签提取装置1可以将与目前正在浏览的web站点相关的最合适书签提供给用户。

[0050] <第二示例性实施方式>

[0051] 接下来,将通过参考附图来详细地介绍基于上述第一示例性实施方式的第二示例性实施方式。图2是示出了根据本发明的第二示例性实施方式的包括书签提取系统的通信环境的框图。。图2中示出的通信环境包括书签提取系统100,浏览器200,网络300和web站点400。

[0052] 书签提取系统100和web站点400可以通过诸如国际互联网的通信网络(以下,简称为“网络”)进行通信。在这个示例性的实施方式中,它应该是,web站点400是提前注册到浏览器200的书签的链接目的地。在图2中,例如(详细的介绍将在之后通过参照图17来进行),书签提取系统100和浏览器200可以被认为是在图17中示出的单个信息处理装置图(计算机)中执行的软件程序的配置,。然而,通过物理地或虚拟地能够彼此相互通信的不同实体的信息处理装置来单独地实现书签提取系统100和浏览器200。

[0053] (书签提取系统100)

[0054] 书签提取系统100具有浏览历史获取单元102、浏览历史分析单元104、搜索web站点信息存储单元106、浏览的内容获取单元108、浏览的内容分析单元110、书签信息获取单元112、书签属性相关性确定单元114、书签内容获取单元116、书签内容相关性确定单元118和相关的书签显示单元120。书签提取系统100对应于上述第一示例性实施方式中的

书签提取设备 1。

[0055] 也就是说,在书签提取系统 100 中,浏览历史获取单元 102 从浏览器 200 获得 web 站点的浏览历史(其对应于第一示例性实施方式中的 web 站点的浏览历史信息 2)0。搜索 web 站点信息存储单元 106 持有的搜索 web 站点的 URL 并且在诸如图 7 中示出的 web 站点搜索信息 50 的搜索 web 站点中的搜索字提取方法。通过基于从搜索 web 站点信息存储单元 106 获得的浏览历史分析从浏览历史获取单元 102 获得的浏览历史,浏览历史分析单元 104 提取已经在搜索时使用过的关键字。浏览的内容获取单元 108 从浏览器 200 获得在浏览器 200 中浏览的 web 站点的内容。通过分析从浏览器 200 经由浏览的内容获取单元 108 获得的 web 站点的内容,浏览的内容分析单元 110 提取与 web 站点相关的关键字。书签信息获取单元 112 从浏览器 200 获取被注册到浏览器 200 的书签信息。

[0056] 在书签提取系统 100 中,书签属性相关性确定单元 114 确定通过书签信息获取单元 112 获取的书签属性 AT 和由浏览器历史记录分析单元 104 或由浏览的内容分析单元 100 提取的关键字之间的相关性。在这里,“书签的属性 AT”是详细信息(属性信息),其包括字符串和以及类似关于诸如注册为书签的 web 站点的标题和 URL 书签以及文件夹通路,(以下,应当应用系统的含义)。书签内容获取单元 116 获取由书签表示的 web 站点 400(链接目的 URL)的内容,通过网络从 web 站点 400 经由书签信息采集单元 112 获取所述书签。针对通过书签内容获取单元 116 获取的内容,书签内容相关性确定单元 118 确定由浏览历史分析单元 104 或浏览的内容分析单元 110 所提取的关键字的相关性。然后,相关的书签显示单元 120 提供由书签属性相关性确定单元 114 和书签内容相关性确定单元 118 确定的书签,作为具有与通过在指示浏览器 200 的屏幕中指示它们的用户的相关性(此后,它们可以被称为“相关书签”)。

[0057] (浏览器 200)

[0058] 浏览器 200 有显示包括 web 站点 400 的各种 web 站点的内容的功能,使得它们可以被用户阅读。。这个浏览器 200 具有浏览历史存储单元 202、浏览的内容存储单元 204 和书签存储单元 206。

[0059] 也就是说,关于用户已经浏览的 web 站点,浏览历史存储单元 202 持有浏览历史(浏览历史信息),例如 web 站点的 URL。浏览的内容存储单元 204 持有用户正在浏览的 web 站点的内容。书签存储单元 206 持有关于用户已经注册的所需 web 站点的书签的信息(书签信息:它对应于第一示例性实施方式中的书签信息 4)(参考图 10A-10F)。书签存储单元 206 只有,如书签信息,代表已经注册的书签的名称(标题),以及对应于书签的 web 站点的 URL 和属性信息。同时,书签存储单元 206 可以持有额外的信息,如关于注册的另外书签的标签和关键字。

[0060] (操作说明)

[0061] 接下来,将会详细地介绍具有前述配置的这种示例性实施方式的操作。

[0062] 图 3 是示出了根据本发明的第二示例性实施方式的书签提取系统 100 进行的书签提取处理的流程图。根据所述示例性实施方式,书签提取系统 100 响应于用户使用浏览器 200 访问所期望的 web 页面来开始操作。也就是说,根据所述示例性实施方式,响应于检测浏览器 200 对 web 页面的访问来开始书签提取处理。

[0063] 当检测到对 web 页面的访问时,浏览历史获取单元 102 从浏览器 200 的浏览历史

存储单元 202 获取浏览历史（步骤 S 102）。接下来，浏览历史记录分析单元 104 分析在步骤 S 102 获得的浏览历史（步骤 S104）。然后，浏览历史分析单元 104 确定正在被浏览器 200 浏览的 web 页面是否是这样的页面，从指示在搜索 web 站点中指示的搜索结果或者不基于分析结果的页面直接移动到所述页面。。

[0064] 作为步骤 S105 中确定的结果，当目前浏览的 web 页面不是来自搜索 web 站点的移动时，浏览的内容获取单元 108 基于 web 页面的内容执行关键字的提取处理（第一搜索过程）（步骤 S106）。也就是说，在步骤 S106 中，浏览的内容获取单元 108 从浏览器 200 的浏览的内容存储单元 204 获取用户正在浏览的 web 页面的内容。然后，通过分析在步骤 S106 获得的内容，浏览的内容分析单元 110 根据关键字提取 web 页面的内容中的重要字（步骤 S108）。在重要字的提取处理中，例如（其可以通过其它方法来完成，并且此后，应当应用同样的方法），可以基于 web 页面中的字的出现（使用量）的频率来执行特定字是否是重要字。

[0065] 另一方面，当 web 页面是来自搜索 web 站点的移动时，浏览历史记录分析单元 104 执行基于浏览历史的关键词的提取处理（第二搜索过程）（步骤 S110）。

[0066] 同时，由浏览历史分析单元（步骤 S102、步骤 S104 以及步骤 S105）进行浏览历史的分析处理的说明，以及稍后将通过参考图 4 来进行基于浏览历史的关键字提取处理（步骤 S110）。

[0067] 在步骤 S112 和 S114，通过使用由上述的关键字提取处理块中的一个所获得的关键字，提取与关键字相关的书签。也就是说，书签信息获取单元 112 从浏览器 200 书签存储单元 206 获取关于注册到浏览器的书签的信息（书签信息）（步骤 S 112）。接下来，在步骤 S 114，执行关键字和书签信息之间的相关性的确定处理。通过书签属性相关性确定单元 114、书签内容获取单元 116 和书签内容相关性确定单元 118 来实现所述确定处理（后文将通过参考图 5 来详细地说明）。

[0068] 在步骤 S114 完成相关性的确定之后，例如，相关的书签显示单元 120 通过图 11 中示出的显示形式将相关书签的列表提供给用户（步骤 S116）。

[0069] 图 11 是示出了在本发明的第二示例性实施方式中相关书签的列表屏幕的显示形式的示意图。图 11 中示出的实例表明一种状态，其中在目前根据网络 200 显示的网页的左侧框架中指示与 web 页面相关的书签列表。

[0070] 此处，将通过参照图 4 来介绍浏览历史分析单元 104 的操作。图 4 是示出了根据第二示例性实施方式的书签提取处理中包括的图 3 中示出的步骤 S102、步骤 S104、步骤 S105 和步骤 S110 中的详细处理的流程图。

[0071] 首先，浏览历史分析单元 104 获取表示历史的历史信息，直到已经从在浏览器 200 的浏览历史存储单元 202 中存储的浏览历史信息到达（对应于 web 页面的历史）用户正在浏览的 web 页面（步骤 202）。

[0072] 接下来，通过参照在步骤 S202 获取的历史信息，浏览历史分析单元 104 确定当前正在浏览的 Web 页面是否为搜索站点（步骤 S204）。这里，作为用于步骤 S204 中确定是否为搜索站点的方法，例如可以采用现在通用的诸如利用预先存储的搜索站点的 URL 进行图形匹配的处理等的过程。稍后将参照图 5 对这种过程的特定示例进行描述。

[0073] 当确定当前正在浏览的 Web 页面为搜索站点时，浏览历史分析单元 104 提取用户

输入到搜索站点的搜索词（步骤 S210）。另一方面，当确定当前正在浏览的 Web 页面不是搜索站点时，浏览历史分析单元 104 从在步骤 202 获取的浏览信息获取关于恰好在当前正在浏览 Web 页面之前浏览过的 Web 页面的历史信息（步骤 S206）。然后，通过参照在步骤 S206 获取的历史信息，浏览历史分析单元 104 确定恰好在当前 Web 页面之前浏览的 Web 页面是否是搜索站点（步骤 S208）。在步骤 S208 中的确定处理中，可以采用与上述步骤 S204 相同的通用处理过程。

[0074] 图 5 是一个流程图，示出根据第二示例性实施方式的搜索词的提取处理，以及对是搜索站点还是不是以单个历史为目标的确定处理，并且更具体地表明上述图 4 所示一系列处理的一部分（步骤 S204-S210）。

[0075] 浏览历史分析单元 104 获取包括在图 4 所示在步骤 S202 获取的历史信息中的 URL（步骤 S212）。接下来，浏览历史分析单元 104 从搜索站点信息存储单元 106 中获取搜索站点信息片段（步骤 S214）。

[0076] 以恰好在当前正在浏览的 Web 页面之前浏览的 Web 页面为目标，浏览历史分析单元 104 确定 URL 的开头是否等于在步骤 S214 获取的搜索站点信息标识的搜索站点的 URL（步骤 S216）。

[0077] 当在步骤 S216 的确定中已经确定其等于搜索站点，则浏览历史分析单元 104 根据对应于搜索站点的搜索词获取方法获取搜索词（步骤 S220）。另一方面，当确定它不等于搜索站点时，浏览历史分析单元 104 确定还没有验证的搜索站点是否保留在搜索站点信息存储单元 106 中存储的搜索站点信息中（步骤 S218）。当存在在该确定中发现的未验证的搜索站点时，处理返回到步骤 S214。另一方面，当确定不存在未验证的搜索站点时，浏览历史分析单元 104 结束该处理。

[0078] 与此同时，由于关于保持在搜索站点信息存储单元 106 中的搜索站点的信息依赖于搜索站点的实现，所以需要周期性更新。作为这样的更新，可以考虑：通过系统的分发者的周期性信息派送的自动校正；搜索站点信息的分享；以及用户的手动校正。

[0079] 接下来，将参照图 6 描述涉及书签属性相关性确定单元 114、书签内容获取单元 116 和书签内容相关性确定单元 118 的组合的相关性确定处理的细节。

[0080] 图 6 是示出本发明第二示例性实施方式中的相关性确定处理的流程图，并表明上述图 3 所示步骤 S 114 的细节。

[0081] 首先，通过参考在步骤 S 112 通过书签信息获取单元 112 获取的书签信息（图 3），书签属性相关性确定单元 114 以包括在书签的属性 AT 中的字符串为目标执行对关键词的搜索（步骤 S302）。这里，在该步骤中将搜索的关键词是在基于上述 Web 页面的内容的关键词提取处理中（第一搜索处理：图 3 的步骤 S106）或基于浏览历史的关键词提取处理（第二搜索处理：图 3（步骤 S110）、图 4 和图 5）中提取的关键词。

[0082] 作为在步骤 S302 中的搜索结果，当在书签的属性 AT 中存在所有关键字是，书签内容获取单元 116 访问特定的 web 站点 400，该 web 站点是书签的链接目的地（步骤 S308）。由此，书签内容获取单元 116 经由网络 300 获取特定 web 站点 400 的内容。

[0083] 当在步骤 S308 中已经能够获取内容时，鉴于在用户已经注册过书签的时间之后内容的改变，书签内容相关性确定单元 118 在书签内容获取单元 116 已经获取的内容的字符串中搜索关键字（步骤 S309）。

[0084] 作为在步骤 S309 中的搜索结果,当包括至少一个关键字时,书签内容相关性确定单元 118 确定在作为链接目的地的特定 web 站点 400 与当前正在浏览的 web 站点之间存在相关性(步骤 S312)。对比而言,当完全不包括关键字时,书签内容相关性确定单元 118 确定在作为链接目的地的特定 web 站点 400 与当前正在浏览的 web 站点之间不存在相关性(步骤 S310)。同时,在当对应于书签的特定 web 站点 400 是失效链接的情况下,过程前进到步骤 S310。同样在该情况下,书签内容相关性确定单元 118 确定在作为链接目的地的特定 web 站点 400 与当前正在浏览的 web 站点之间不存在相关性。

[0085] 另一方面,作为在步骤 S302 中的搜索结果,当部分或全部关键字不存在书签的属性 AT 中时,书签内容获取单元 116 经由网络 300,通过访问作为书签的链接目的地的 web 站点 400 来获取内容,这是步骤 S308 中的情形(步骤 S304)。继而,通过定向在所获取的内容中包括的字符串,书签内容相关性确定单元 118 搜索还没有被包括在书签的属性 AT 中的关键字(步骤 S306)。同时,为了降低当时的过程中的负载,可以经由网络 300,通过使用搜索站点执行从获取内容到搜索关键字的过程。

[0086] 作为在步骤 S306 中的搜索结果,当确定为在书签的属性 AT 中没有包括的所有关键字包括在步骤 S304 中获取的内容中时,书签内容相关性确定单元 118 确定在作为链接目的地的特定 web 站点 400 和当前正在浏览的 web 站点之间存在相关性(步骤 S312)。

[0087] 另一方面,当没有包括在书签的属性 AT 中的部分或全部关键字不包括在所获取的内容中时,书签内容相关性确定单元 118 确定在所关注的作为链接目的地的特定的 web 站点 400 与当前正在浏览的 web 站点之间不存在相关性。(步骤 S310)。

[0088] (示例)

[0089] 接下来,通过使用具体的示例来描述上述一系列处理步骤的整个画面。首先描述在该说明书中的前提条件。

[0090] 图 8A 和图 8B 是示例性描述了在本发明的第二示例性实施方式的示例中,在的用户浏览 web 的时候的屏幕转换的图。图 9A 和图 9B 是在图 8A 和图 8B 中说明的屏幕转换的情况下的概念性描述书签提取系统的操作的图。

[0091] 也就是说,在图 8A 中说明的模式 1 示出了其中例如通过在其自身的博客中直接输入期望 web 站点的 URL 到达当前正在浏览的 web 站点的情况。另一方面,在图 8B 中说明的模式 2 显示了其中通过在搜索站点中输入关键字获取关于期望关键字的搜索结果,并进一步通过进行到在搜索结果中包括的期望 web 站点的转换,从而当前正在浏览期望的 web 站点的情况。

[0092] 同时,在图 8A,图 8B,图 9A 和图 9B 中,“Internet Explorer”,“Media player”,“Office”,“Windows”是注册商标(下文中,同样适用)。

[0093] 假设在图 7 中概念性地说明的信息被称为从搜索站点信息存储单元 106 中获取的搜索站点信息。同样,假设用户之前已经注册图 10A 到 10F 中概念性地说明的书签(书签 1 到 6)。然而,假设该用户由于注册了大量其它书签,从而导致已经无法发现这些书签 1 到 6。或者,假设由于这些书签已经被注册的事实本身被遗忘,该用户就无法再仅通过自己的操作容易地找到这些书签了。

[0094] 在图 10A 到 10F 中说明的书签信息的示例中,参考文件夹路径,可以发现对大部分书签,用户已经将相关的书签在 Microsoft(注册商标)文件夹中按软件组织起来。然而,

仅有书签 6 (图 10F) 不是基于该组织程序的。例如,一个原因是假设在注册书签 6 时,书签的组织方法已经被改变的情况,或者注册的位置被临时改变的情况。

[0095] 在满足上述前提的一般技术的情况下,为了使用户能够找到书签 6,首先,必不可少的是用户自己记住书签 6 已经被注册的事实。另外,用户除了记住之前有关于书签的组织方法并且按照该组织方法查找书签 6,或者通过搜索调查外别无选择。然而,当使用一般书签搜索功能,并且如果使用的关键字是与在模式 1 (图 8A) 和模式 2 (图 8B) 中例示的页面转换过程中用户想要了解的项相关的、诸如“Office”和“Windows”(Office 和 Windows 是注册商标)之类的词,则无法找到书签 6。

[0096] 与这种一般技术相比,根据该示例实施方式中上文所提到的技术,用户被鼓励有效地使用如以上前提的被埋藏在大量书签中的书签。

[0097] 下文中,针对图 8A 和图 8B 中说明的屏幕转换的书签提取系统 100 的操作将参考图 9A 和图 9B 描述。

[0098] 首先,将描述图 9A 中说明的模式 1 的示例。模式 1 指示“web 页面被浏览”已经被直接从“自己的博客”页面中找到而不需要访问搜索站点的情况。在该模式 1 的示例中,假设用户当前在浏览的 web 页面是具有指示了与信息技术相关的列的 web 页面。假设当用户以前浏览过该 web 页面时,比较 OS (操作系统) 的专题出现在 web 页面上。还假设在当前的浏览中,关于快捷键的专题被发布在 web 页面上。

[0099] 在这种情况下,在书签提取系统 100 中,浏览历史获取单元 102 从浏览历史存储单元 202 中获取浏览历史 (图 3 的步骤 S102)。然后,通过分析获取的浏览历史,浏览历史分析单元 104 确定当前浏览的 web 站点是否是从搜索站点的搜索结果的页面中直接得到转换 (找到) 的页面 (图 3 的步骤 S104)。在模式 1 中,当前浏览的 web 站点的转换不是经由搜索站点的转换。基于这个原因,作为关键字的提取过程,浏览内容获取单元 108 获取从浏览器 200 的浏览内容存储单元 204 获取的浏览内容 (图 3 的步骤 S106)。

[0100] 接下来,通过分析所获取的内容,浏览内容分析单元 110 从“正在被浏览的 web 页面”中提取重要的词作为关键字 (图 3 的步骤 S108)。在图 9A 示出的模式 1 的情况下,假设“Windows”和“快捷键”作为这类关键字被获取。

[0101] 为了从多数已经注册的书签中提取与关键字相关的书签,书签信息获取单元 112 从浏览器 200 的书签存储单元 206 中获取书签信息 (图 3 的步骤 S112)。在这种情况下,图 10A 到 10F 中说明的书签 1-6 被获取。另外,通过联合使用书签属性相关性确定单元 114,书签内容获取单元 116 和书签内容相关性确定单元 118,可以确定这些书签和关键字之间的相关性 (图 3 中的步骤 S114,和图 6)。由此,在该示例中,从书签 1 到 6 中选出书签 1, 3, 5 和 6 作为与关键字相关的书签 (相关书签)。例如,相关书签显示单元 120 通过图 11 示出的显示表格向用户提供相关书签列表 (图 3 的步骤 S116)。

[0102] 这里,关于对模式 1 的情况下的浏览历史的分析处理和相关性确定处理的具体的过程 (处理流程) 将在下文描述。

[0103] 首先,将要描述对浏览历史的分析处理。在模式 1 中 (图 9A),基于被浏览的 web 页面的内容,如上所述,重要的词汇被提取作为关键字。将会增加关于该过程的描述。

[0104] 浏览历史分析单元 104 从由步骤 S102 (图 3) 获取的浏览历史中获取历史作为对正在被浏览的 web 页面的访问信息,该历史体现了直到达到当前浏览的 web 页面的浏览过

程。然后,浏览历史分析单元 104 确认正在被浏览的 web 页面是搜索站点(图 4 的步骤 S204)。在该示例中,正在被浏览的 web 页面不是搜索站点。基于这个原因,浏览历史分析单元 104 从浏览历史中获取 web 页面的历史(图 4 的步骤 S206),该历史恰好在当前正在被浏览的 web 页面之前被浏览。在这种情况下,对“用户自己的博客”的访问信息被如图 9A 所示获取。由于恰好在当前浏览页面之前被浏览的 web 页面依据该访问信息被确定为不是搜索站点(在图 4 的步骤 S208,否),浏览历史分析单元 104 确定用户当前正在浏览的 web 页面不是通过搜索站点找到的页面。

[0105] 接下来,描述相关性确定过程。首先,从书签的属性 AT 中包括的特征字符串中,书签属性相关性确定单元 114 搜索上述在模式 1 中提取的关键词“Windows”和“快捷键”。在模式 1 中,作为包括所有书签的属性 AT 中的关键词的书签,书签 1 和 2 被从书签 1 到 6 中选中。书签 3,4,5 和 6 被从书签 1 到 6 中选中作为书签,其中部分或没有关键字不被包括。

[0106] 关于每个书签 1 和 2,通过经由网路 300 访问链接目标的 web 站点 400,书签内容获取单元 116 获取关注的 web 站点的内容(图 6 的步骤 S308)。书签 2(图 10B)是属于失效链接,其中在步骤 S308 未能获取内容。基于这个原因,书签内容相关性确定单元 118 确定当前正在被浏览的 web 站点和书签 2 没有相关性(图 6 的步骤 S310)。

[0107] 另一方面,关于能够获取内容的书签 1,书签内容相关性确定单元 118 搜索关键字(“Windows”和“快捷键”)以找到获取的内容中包括的特征字符串(图 6 的步骤 S309)。由于书签 1 包括关键字,书签内容相关性确定单元 118 确定当前正在被浏览的 web 站点和书签 1 之间具有相关性(图 6 的步骤 S312)。

[0108] 同样关于书签 3,4,5 和 6,通过经由网路 300 访问链接目标的 web 站点 400,书签内容获取单元获取关注的 web 站点的内容(图 6 的步骤 S 304)。此外,通过定向所获取的内容的特征字符串,书签内容相关性确定单元 118 搜索没有包括在关注的书签的属性 AT 中的关键字(图 6 的步骤 S306)。由此,书签内容相关性确定单元 118 将在其内容中包括所有关键字的书签 3,5 和 6(图 10C,10E 和 10F)为确定为具有相关性的书签(图 6 的步骤 S312)。另一方面,书签内容相关性确定单元 118 将仅包括所有关键字中的部分关键字的书签 4(图 10D)判断为不具有相关性的书签(图 6 的步骤 S310)。

[0109] 在第一次浏览 web 站点期间,用户倾向执行使用关于特定 web 站点的 URL 的使用分类的注册操作。在一般的技术下,当假设重要词汇作为使用分类时,可以认为“OS ;比较”等设置作为使用分类。另一方面,作为该示例(模式 1)的前提,用户当前正在浏览的 web 页面是具有如上所示指示了与信息技术相关的列的 web 页面。由于该 web 页面是列站点,例如出现在其上的内容有很高概率被周期性地更新。对应的,在模式 1 的情况下,与该 web 站点第一次被浏览时的内容不同,它以图 9A 说明的 Windows 的快捷键作为特征。当基于例如某个 web 站点第一次被浏览时的内容的信息执行关联时,例如关键字“OS ;比较”,在一般技术的情况下作为如上所述的使用分类来注册。基于这个原因,依据一般技术,不存在与关键字“Windows”和“快捷键”相关的与当前内容对应的书签。相比之下,在该示例中,利用该次 web 页面被浏览时的信息,书签提取系统 100 将 web 页面和注册的书签进行关联。结果,依据该示例,相关的书签 1,3,5 和 6 可以使用与当前正在浏览的 web 页面的内容对应的关键字“Windows”和“快捷键”来提取。也就是说,可能从预先注册的书签中准确地选择最合适的与当前正在浏览的 web 站点相关的书签,并且将他们提供给用户。

[0110] 接下来,将描述图 9B 中图示的模式 2 中的示例。模式 2 是其中已经经由搜索站点到达“正在浏览的 web 页面”的示例。

[0111] 同样,在模式 2 的情况下,对浏览历史进行分析,正如上述模式 1 那样,以便确定向当前正浏览的 web 页面移动(转移)的方法(图 3 的步骤 S102-S105)。然而,因为模式 2 是经由搜索站点的示例,所以作为关键字提取过程,浏览历史分析单元 104 基于浏览历史提取搜索词(图 3 的步骤 S110),该搜索词是用户已经输入搜索站点中的用于搜索的关键字。这时,在模式 2 中,将“Microsoft”、“Office”、“快捷键”作为关键字获得。

[0112] 接下来,与模式 1 的情况相同(图 3 的步骤 S112),从书签存储单元 206 中获取书签信息,并且在这个情况下,获得图 10A 到 10F 中图示的书签 1 到书签 6。此外,通过以组合的方式使用书签属性相关性确定单元 114,书签内容获取单元 116 和书签内容相关性确定单元 118,对书签和关键字之间的相关性进行确定(图 3 的步骤 S114)。作为结果,从书签 1 到书签 6 中选择书签 3、4 和 6(图 10C、10D 和 10F)作为与关键字有关的相关联书签。

[0113] 在此,关于模式 2 的情况中的浏览历史的分析处理以及相关性的确定处理,下面将描述书签提取系统 100 执行的处理的具体流程。

[0114] 将首先描述浏览历史分析单元 104(图 4)的处理。模式 2 中(图 9B),如上所述将搜索站点中的搜索词作为关键字提取。在模式 2 的情况中,在浏览历史的分析处理中,在步骤 S204 确定正在浏览的 web 页面不是搜索站点之后,浏览历史分析单元 104 从浏览历史中获取与正好在正浏览的 web 页面之前浏览的 web 页面有关的历史信息(步骤 S206)。

[0115] 在图 9B 中示出的示例的情况下,获取与搜索站点的“搜索结果”屏幕有关的访问信息。在模式 2 中,在步骤 S208 中,在步骤 S206 中获取的访问信息与先前存储的搜索站点一致(此处假设它与图 7 中示出的搜索站点信息中的搜索站点 ID 1 相对应)。因此,在步骤 S210 中执行搜索词的提取。

[0116] 例如基于在图 7 中显示的搜索站点信息中包括的“搜索词提取方法”,执行这种搜索词的提取。在这种情况下,浏览历史分析单元 104 从历史信息中获取与“搜索结果”相对应的 URL,并且从 URL 的查询字符串中获取作为搜索词的值 q。在图 7 中显示的“搜索词提取方法”中,当包含作为搜索词的多个词时,通过符号“+”将它们相连。这个符号“+”表示,当为搜索词相连的状态时,也就是说,当存在“关键字 1+关键字 2”的字符串时,对字符串进行划分,并得到“关键字 1”和“关键字 2”。也就是说,在步骤 S210 中,通过使用值 q 对字符串进行划分,获得可以获取的搜索词组(多个搜索词)。在图 9B 中显示的示例的情况下,将“Microsoft”、“Office”、“快捷键”作为搜索词获得。在对书签的内容进行搜索期间,使用所获得的搜索词作为关键字(图 6 的步骤 S306 和步骤 S309)。

[0117] 因此,根据上述示例,可以提取更符合用户想要了解的内容的关键字。与此相反,在采用通用技术的情况下,当已经经由搜索站点到达当前正浏览的 web 站点时(例如图 8B 中显示的模式 2),可以预期从正在浏览的 web 站点的内容中提取出“Windows”和“快捷键”这样重要的(使用频率很高的)词作为关键字。因此,在采用通用技术的情况下,因为可能会提取出与用户想要了解的内容无关的书签,所以不方便。

[0118] 接下来,将以图 9B 中显示的模式 2 作为示例对相关性确定操作进行描述。首先,以书签属性 AT 中包含的字符串为目标,书签属性相关性确定单元 114 搜索关键字“Microsoft”、“Office”、“快捷键”(图 6 的步骤 S302)。在这个情况下,选择出书签 4 为包

含书签属性中的全部关键字的书签。选择出书签 1、2、3、5 为不包含全部关键字中的一部分或全部的书签。

[0119] 对于书签 4 (图 10D), 通过经由网络 300 访问作为链接目的地的特定 web 页面 400, 书签内容获取单元 116 获取该具体 web 站点的内容 (图 6 的步骤 S308)。当已经能够获取内容时, 以该内容中的字符串为目标, 书签内容相关性确定单元 118 搜索关键字 (图 6 的步骤 S309)。在这种情况下, 因为书签 4 包括这些关键字, 所以书签内容相关性确定单元 118 确定书签 4 与正在浏览的 web 站点有相关性 (图 6 的步骤 S312)。

[0120] 同样, 对于书签 1、2、3、5 中的每一个, 通过经由网络 300 访问作为链接目的地的特定 web 站点 400, 书签内容获取单元 116 获取该特定 web 站点的内容 (图 6 的步骤 S304)。此外, 以所获取的内容中的字符串为目标, 书签内容相关性确定单元 118 搜索未包含在所涉及的书签的属性中的关键字 (图 6 的步骤 S306)。在这种情况下, 因为书签 3 (图 10C) 包含该内容中的全部关键字, 书签内容相关性确定单元 118 确定书签 3 和 6 与正浏览的 web 站点有相关性 (图 6 的步骤 S312)。另一方面, 书签 1、2 和 5 (图 10A、B、E) 不包含内容中的关键字的一部分或全部。相应地, 书签内容相关性确定单元 118 确定书签 1、2、5 与正浏览的 web 站点没有相关性 (图 6 的步骤 S310)。

[0121] 类似于之前描述的浏览历史分析单元 104 执行处理的情况, 根据通用技术, 可以预期在模式 2 的情况下, 基于当前正浏览的 web 站点提取出“Windows”和“快捷键”作为关键字, 基于该关键字提取出书签 1、3、5 和 6。在这种情况下, 将省略用户想要了解的基本上包含在正在浏览的 web 站点中的项目, 即, 与“Microsoft”、“Office”、“快捷键”有相关性的书签。与此相反, 根据该示例, 可以将与“Microsoft”、“Office”、“快捷键”相关联的书签 3、4 和 6 作为更符合用户想要了解的内容的书签提供给用户。

[0122] 因此, 在第二示例性实施方式及其上述示例中, 书签提取系统 100 可以使用浏览器 200 从先前注册的书签中准确地选择与用户当前正浏览的 web 站点相关联的最合适的书签, 并且将它们提供给用户。

[0123] 通常而言, 即使已经通过搜索等处理找到 web 站点, 并且之前已经将这些 web 站点作为书签注册了, 仍然有可能因为已经注册了大量的书签而不会再找到需要的书签。可替换地, 通过已经忘记某个已经注册的书签这个事实本身, 用户可能无法再利用它。对于这种情形, 根据第二示例性实施方式以及上述示例, 在上述模式 1 和模式 2 这两种情况下, 与 web 页面相关联的书签可以在该 web 页面被浏览时提供给用户。因此, 用户可以更快速地到达描述了其想要了解的内容的 web 页面。其原因在于, 通过例如书签属性相关性确定单元 114 和书签内容相关性确定单元 118 的功能, 可以准确地提取出与用户当前正浏览的 web 页面相关联的书签。

[0124] 根据第二示例性实施方式及上述提及的示例, 反映用户意图的书签可以被更正确地提取, 其中用户已经参考正在被浏览的网站。特定地声明, 在模式 2 (图 8B), 在利用使用频率或类似的方式对正在被浏览的网站内容进行分析时, “窗口”和“快捷键”似乎根据常规技术作为关键字被获取。但是, 在模式 2 的情况下, 用户已经到达当该网站, 该网站由于执行“office 快捷键”的搜索而被浏览。从而, 在这种情况下, 可以认为用户期望搜索词相关的信息被输入到搜索站。根据第二示例性实施方式及上述例子 (图 9B), 因为书签的提取可以通过利用搜索词来执行, 与用户的期望更一致的书签可被提取。另外, 当用户忘记之前注

册的书签,并在搜索站点执行相同目的搜索时,之前注册的书签可以被提供给用户。这是因为可以通过浏览历史分析单元 104 分析浏览历史来得知用户通过哪种方式参考该网站。

[0125] 同样根据第二示例性实施方式及上述提及的示例,当用户已经访问正在被浏览的 web 页面时,与该 web 页面内容相关的书签可以被提取。从而,相应于自从用户最后一次登录该 web 页面起的内容的变化,与当前状态更一致的书签可以被提取。其理由是,关于与一注册的书签相关的一 web 页面,其相关性被验证,通过被浏览内容分析单元 110 分析被浏览的该 web 页面正被浏览时的内容。

[0126] 还是根据第二示例性实施方式及上述提及的例子,与登录正在被浏览的 web 页面时该书签的内容一致的书签可以被提取。从而,可以进行对不能通过在该书签的属性 AT 中包括的信息找到的相关书签的提取,或者可以进行对相关书签的失效链接的书签的排除。其理由是,在相关性确定过程中,不仅在关键字和属性信息之间的相关性被确定,书签的链接目的地的内容也可以以组合的方式通过书签内容相关性确定单元 118 来确定。

[0127] < 第 3 示例性实施方式 >

[0128] 接下来,基于上述提及的第一和第二示例性实施方式的第三示意性实施方式将被描述。在此后的描述中,将主要描述实施方式的特征部分。在这种情况下,通过与上述示例性实施方式中的相同结构采用相同的参考号码,重复的描述将被省略。

[0129] 图 12 是例示了包括根据本发明的第三示例性实施方式所示的书签提取系统的通信环境的框图。若粗略地归类,如图 12 所示的通信环境包括书签提取系统 100A,浏览器 200,网络 300 以及 web 站点 400。在根据图 2 所示的第二实施方式的书签提取系统 100 之外,该书签提取系统 100A 还进一步包括相关书签高速缓存操作单元 122,相关书签高速缓冲存储器单元 124 以及共同关键字相关性确定单元 126。书签提取系统 100A 对应于上述提及的第一示例性实施方式中的该书签提取装置 1。

[0130] 在该示例性实施方式中,提供相关书签高速缓冲存储器单元 124 和相关书签高速缓存操作单元 122,主要为了减少提取相关书签时的负荷。也即,在该示例性实施方式中,相关书签高速缓冲存储器单元 124 和相关书签高速缓存操作单元 122 实现了各种数据的高速缓存功能。

[0131] 关于到目前为止用户采用浏览器 200 已经浏览的 web 页面(也即,包括正在被浏览的 web 页面),相关书签高速缓冲存储器单元 124 保持数据高速缓存(cache)。该高速缓存与信息对应,在信息中:比如,包括多个以下描述的数据项目;并且它们之间相互关联。

[0132] - 至今为止用户采用浏览器 200 已经浏览的 web 页面(也即,包括正在被浏览的 web 页面)的 URL;

[0133] - 从上述 web 页面中提取的关键字;

[0134] - 上述 web 页面的相关书签的识别信息(ID);

[0135] - 上述 web 页面中的每个的最后登录时间。

[0136] 在此,被提取的关键字是在上述第二示例性实施方式中的第一或第二搜索过程中已经被提取的关键字。相关书签是在有相关性的第二示例性实施方式中由书签属性相关性确定单元 114 和书签内容相关性确定单元 118 确定的书签。

[0137] 假设高速缓存单元保持的信息的生存时间持续到当浏览器 200 的操作被完成(直到浏览器功能的使用结束)或者当系统断电,举例来说。这是为了与 web 页面和书签的内

容的改变相应。共同关键字相关性确定单元 126 确定提取的共同关键字的相关性。

[0138] (操作的描述)

[0139] 接下来,具有此前提到规程的示例性实施方式的操作将被具体地描述。

[0140] 图 13 是示出了根据本发明的第三示例性实施方式的书签提取系统 100A 执行的书签提取过程的流程图。书签提取系统 100A 响应于用户通过使用浏览器 200 登录期望的 web 页面而开始操作,如第二示例性实施方式的情形。

[0141] 首先,通过查阅相关书签高速缓冲存储器单元 124,相关书签高速缓存操作单元 122 获取具有最接近于当前时间的最后登录时间(数据和时间)的若干高速缓存(步骤 S402)。接下来,定向于步骤 S402 获得的高速缓存,共同关键字相关性确定单元 126 提取在上述高速缓存缓存中共同出现的关键字作为与上述高速缓存相关的共同关键字(步骤 S406)。通过查阅相关书签高速缓冲存储器单元 124,相关书签高速缓存操作单元 122 获得与该用户当前正在浏览的 web 页面的 URL 相对应的高速缓存(步骤 S408)。

[0142] 通过步骤 S408 的尝试获得缓存,相关书签高速缓存操作单元 122 确定与正被浏览的 web 页面的对应的高速缓存是否存在(步骤 S410)。在此,高速缓存存在的情形是指这样的情形,其中当前被浏览的 web 页面的 URL 与存储在相关书签高速缓冲存储器单元 124 中的过去曾被浏览的 web 页面的 URL 相同。相关书签高速缓存操作单元 122 在步骤 410 中搜索并检测这样的 URL。

[0143] 当确定存在这样的高速缓存时(在步骤 S410,是),通过从相关书签高速缓冲存储器单元 124 获得的所有相关书签识别信息作为一个关键字,共同关键字相关性确定单元 126 获得关于书签信息获取单元 112 的所有相关书签的属性信息(AT)。共同关键字相关性确定单元 126 经由书签内容获取单元 116 获取上述 web 页面的相关书签的连接目的地的 web 页面的内容。然后,以获得的属性 AT 以及相关书签的内容为目标,共同关键字相关性确定单元 126 测量共同关键字出现的数目,保持所测得的数值(步骤 S428),并且处理前进到步骤 S430。在此,在步骤 S428 从相关书签高速缓冲存储器单元 124 获得的所有书签表示与当前时间相近的若干 web 页面(URLs) 相关联的所有相关书签。

[0144] 另一方面,当确定不存在这样的缓存时(在步骤 S410,否),类似于上述第二示例性实施方式中提及的处理步骤 S 102 到 S 114(图 3)的处理在步骤 S412 到步骤 S426 执行。然而,步骤 S426 中的处理与步骤 S114 部分不同,其将参考图 14 在后续提到。

[0145] 然后,相关书签显示器单元 120 指示在步骤 S426 被确定为有相关性的书签或者在步骤 S428 被作为测量对象的所有书签,以共同关键字的出现次数排序的列表(步骤 S430)。例如,当以图 10A 到 10F 所示的显示形式作为示例描述,书签 1、3 和 5 在书签中按照出现次数的顺序从高到低地显示。

[0146] 在此,在步骤 S426 执行的相关性确定处理将参考图 14 被描述。图 14 是示出了本发明第三示例性实施方式中的相关性确定过程的流程图,并指示了上述提及的图 13 中的步骤 S426 的细节。除去与在第二示例性实施方式中提及的相关性确定过程之外(图 6),该流程图还包括了在步骤 S312 之后的步骤 S501。也即,步骤 S501,以在步骤 S312 确定的有相关性的所有书签作为目标,共同关键字相关性确定单元 126 测量在书签的属性 AT 以及内容中共同关键字的出现次数,并且保持测量值,并且过程前进到步骤 S430。

[0147] 根据上述提及的第三示例性实施方式,有可能从提前注册的书签中选择与当前正

在被浏览的网站最适合的书签,并且将其提供给用户,正如上述提及的第二示例性实施方式所示。此外,根据该示例性实施方式,通过使用相关的书签高速缓存操作单元 122 以及相关的书签高速缓冲存储器单元 124 实现的高速缓存功能,上述的相关信息可以被高速缓存。从而,当分析浏览历史、浏览内容以及书签内容时的负荷可以减少。

[0148] 同样根据上述提及的第三示例性实施方式,通过上述高速缓存功能,有可能通过浏览器导航的对应操作例如“前进”和“返回”。

[0149] < 第四示例性实施方式 >

[0150] 接下来,基于第一和第二示例性实施方式的第四示例性实施方式将被描述。在此后的描述中,将主要描述实施方式的特征部分。在这种情况下,通过与上述示例性实施方式中的相同结构采用相同的参考号码,重复的描述将被省略。

[0151] 在该示例性实施方式中,如下所述,向用户提供另一个显示形式。也就是,在该实施方式中的另一显示,当与用户当前正在浏览的 web 站点相关的书签被提供给用户时,根据其与该 web 站点的相关程度排列的显示形式提供已注册的书签。

[0152] 在下述描述中,作为示例,假设相关性等级有 6 级,从等级 1(相关性:高)到等级 6(相关性:低)。即:

[0153] 等级 1:其中所有关键字在书签属性 AT 和内容这两者中的情形;

[0154] 等级 2:其中书签属性 AT 包括部分关键字,并且内容中包括所有关键字的情形;

[0155] 等级 3:其中书签属性 AT 不包括任意关键字,并且内容中包括所有关键字的情形;

[0156] 等级 4:其中书签属性 AT 包括所有关键字,并且内容中包括部分关键字的情形;

[0157] 等级 5:其中书签属性 AT 包括部分关键字,并且内容中包括所有剩余关键字的情形;以及

[0158] 等级 6:其中书签属性 AT 包括所有关键字,并且内容中不包括关键字的情形。

[0159] 在这个实施方式中,系统配置和整个处理配置几乎与上述第二示例性实施方式相同。然而,如上所述,由书签属性相关性确定单元 114、书签内容获取单元 116 和书签内容相关性确定单元 118 实现的相关性确定处理的配置不同。

[0160] 图 15 是示出了在本发明的第四示例性实施方式中的相关性确定处理的流程图,并且指示了上述图 3 中示出的步骤 S114 的详细内容。

[0161] 首先,与第二示例性实施方式中的步骤 S302(图 6)类似,通过参考由书签信息获取单元 112 在步骤 S112(图 3)获取的书签信息,书签属性相关性确定单元 114 执行以包含在书签属性 AT 中的字符串为目标的关键词查找(步骤 S602)。

[0162] 根据步骤 602 中的确定结果,书签内容获取单元 116 分出执行步骤 S604、步骤 S612 和步骤 S616 中任意步骤的处理的多个支路。也就是说,当书签的属性 AT 中包含关键词的一部分时,执行步骤 S604。当书签的属性 AT 中不包含关键词的一部分时,执行步骤 S612。当书签的属性 AT 中包含所有关键词时,执行步骤 S616。

[0163] 在步骤 S604、步骤 S612 和步骤 S616 中,正如第二示例性实施方式中的步骤 S304 和步骤 S308(图 6)那样,书签内容获取单元 116 访问作为所涉及的书签的链路目的地的特定的网站 400。由此,书签内容获取单元 116 经由网络 300 获取特定 Web 站点 400 的内容。

[0164] 当虽然在步骤 S616 中进行了尝试但仍不能获得内容时,可以确定该特定站点 400

具有失效的链接。因此,书签内容相关性确定单元 118 确定与该网站对应的书签与当前正浏览的网站没有相关性(步骤 S628)。在这种情况下,书签内容相关性确定单元 118 可以通过指示等向用户报告出现了失效的链接。

[0165] 在另一方面,正如第二示例性实施方式中的步骤 S309(图 6)那样,当能够根据步骤 S616 中的处理获得内容时,由于当书签已经由用户注册时内容会随着时间变化,所以书签内容获取单元 116 在所获取的内容的字符串中查找该关键词(步骤 S618)。

[0166] 继而,根据步骤 S618 的查询结果,书签内容获取单元 116 分出执行步骤 S620、步骤 S622 和步骤 S624 的多个支路。也就是说,当在步骤 S616 获取的内容中不包含关键词时,书签内容相关性确定单元 118 给出等级 6(步骤 S620)。当在步骤 S616 获取的内容中包含关键词的一部分时,书签内容相关性确定单元 118 给出等级 4(步骤 S624)。而且,在这三种情况下,如第二示例性实施方式中的步骤 S312(图 6)的情况那样,书签内容相关性确定单元 118 确定作为链接目的地的网站 400 与当前正浏览的网站有相关性(步骤 S630)。

[0167] 正如第二示例性实施方式中的步骤 S306(图 6)那样,当在步骤 S604 和步骤 S612 获取内容后,书签内容相关性确定单元 118 在所获取的内容中包含的字符串中查找未包含在书签的属性 AT 中的关键词(步骤 S306 和步骤 S614)。

[0168] 作为在步骤 S614 中查找的结果,当所获取的内容包含所有关键词时,书签内容相关性确定单元 118 给出等级 3(步骤 S626)。在这种情况下,书签内容相关性确定单元 118 确定作为链接目的地的网站 400 与当前正浏览的网站有相关性(步骤 S630)。与此相反,当所获取的内容不包含任何关键词时,或者当所获取的内容仅包含关键词的一部分时,书签内容相关性确定单元 118 确定作为链接目的地的网站 400 与当前正浏览的网站没有相关性(步骤 S628)。

[0169] 进一步地,作为步骤 S606 中查找的结果,当未包含在书签属性 AT 中的关键词中只有一部分关键词或没有关键词包含在步骤 S604 获取的内容中时,书签内容相关性确定单元 118 确定作为链接目的地的特定网站 400 与当前正浏览的网站没有相关性。

[0170] 与此相反,作为步骤 S606 中查找的结果,当未包含在书签属性 AT 中的所有余下的关键词都包含在步骤 S604 获取的内容中时,书签内容相关性确定单元 118 给出等级 5(步骤 S608)。当在步骤 S604 获取的内容中包含所有关键词时,由书签内容相关性确定单元 118 给出等级 2(步骤 S610)。在这些情况下,书签内容相关性确定单元 118 确定作为链接目的地的特定网站 400 与当前正浏览的网站有相关性(步骤 S630)。

[0171] 图 16 是示例本发明的第四示例性实施方式的相关联书签的列表屏幕的显示格式的示图。显示在图 16 中的该示例标示如下的状态,即,在当前正标示的网页的左侧的框中,由浏览器 200 的功能实体对与网页相关联的书签的列表进行标示。在这种情况下,作为相关联的书签,按照从等级 1 到等级 3 的顺序对书签 1、5、3 进行标示。进一步地,在这个示例中,作为与注册的内容相比可能已经变化的书签,对书签 3 进行标示。同样在这个示例中,作为链路可能失效的书签,对书签 2 进行标示。

[0172] 根据上述第四示例性实施方式,如第一和第二示例性实施方式的情况那样,有可能从先前注册的书签中准确地选择与当前浏览的网站相关的最合适的书签,并将这些书签提供给用户。进一步地,根据这个示例性实施方式,因为在相关性确定处理中执行了等级确定,所以有可能根据相关程度准确地选择相关性更符合正浏览的网页的书签。并且以容易

区分的方式将这些书签提供给用户。

[0173] 同样根据上述第四示例性实施方式,当以容易区分的方式向用户注册了书签时,有可能报告出现失效链路以及内容随时间改变。

[0174] <第一到第四示例性实施方式的通用示例>

[0175] 接下来,将参考图 17 描述第一到第四示例性实施方式的通用示例。将描述使用计算机(信息处理装置)实现根据上述第一到第四示例性实施方式的书签提取系统 100、100A 和浏览器 200 的情况。也就是说,虽然可能由专门的硬件实现根据第一到第四示例性实施方式的装置,但是还可以如本示例的情况那样,由计算机中的计算机程序(软件)作为硬件资源来实现该装置。

[0176] 图 17 是示例性描述能够实现本发明的第一到第四示例性实施方式的计算机(信息处理装置)的硬件配置的示意图。图 17 中显示的计算机 1000 的硬件包括:CPU 11(中央处理单元)、显示器 12 和通信接口(I/F)13、ROM(只读存储器)17、RAM(随机接入存储器)15 和硬盘驱动器(HD)16,并且具有经由总线 17 将这些硬件相连接的结构。在上述示例性实施方式中的每一个中,通信接口 13 是实现经由网络 300 与图 2 和图 12 所示的网站进行通信的通用通信装置。在这种硬件配置中,CPU11 控制至少作为书签提取系统 100(100A)的计算机 1000 的通用操作。

[0177] 在将能够实现描述中提及的流程图(图 3 到图 6 以及图 13 到图 15)的功能的计算机程序提供给图 17 中示出的计算机 1000 之后,通过将计算机程序读入计算机 1000 的 CPU 11 中并执行该计算机程序,实现了以上述第一到第四示例性实施方式为本发明的。提供给计算机 1000 中的计算机程序可以存储在存储设备(存储介质)中,例如临时存储存储器(15)或能够被读写的硬盘设备(16)。

[0178] 此处,在上述每个示例性实施方式的情况下,与每个上述流程图相关联的计算机程序分别表示浏览历史获取单元 102、浏览历史分析单元 104、浏览内容获取单元 108、浏览内容分析单元 110、书签信息获取单元 112、书签属性相关性确定单元 114、书签内容获取单元 116、属性内容相关性确定单元 118、相关联书签显示单元 120、相关联书签高速缓存操作单元 122 以及公共关键词相关性确定单元 126。这种计算机程序例如与存储在硬盘驱动器(HD)16 中的程序 21 相对应。

[0179] 浏览器 200 中的浏览器历史存储单元 202、浏览内容存储单元 204 以及书签存储单元 206 可以例如作为各种存储信息 22 存储在硬盘驱动器 16 中。同样,书签提取系统 100、100A 中的查找站点信息存储单元 106 以及相关联书签高速缓存单元 124 可以例如作为各种存储信息 22 存储在硬盘驱动器 16 中。根据程序 21 的执行对存储信息 22 进行更新。

[0180] 在上述情况下,作为将计算机程序提供给计算机的方法,可以采用诸如经由各种记录介质(例如 CD-ROM)将计算机程序安装在计算机中以及经由通信线路(例如互联网)从外部下载计算机程序的方法之类的、当前通用的步骤。在这种情况下,可以识别出本发明由计算机程序(程序 21)代码或由存储这些代码的存储介质来配置。

[0181] 如上所述,根据上述每个示例性实施方式及其示例,有可能从先前注册的书签中准确地选择与当前正浏览的网站相关联的最合适的书签,并且将这些书签提供给用户。

[0182] 同时,还可以将上述示例性实施方式以及它们的部分或全部修改描述为以下附加内容的形式。然而,通过示例性实施方式示例性地描述的本发明以及上述修改不限于以下

内容。

[0183] (附加 1). 一种书签提取装置,包括:

[0184] 关键字提取装置,用于基于 web 站点至今的浏览历史信息提取关键字;以及

[0185] 提供单元,用于基于所述关键字,从多个注册的书签中提供与所述关键字提取单元提取的所述关键字相关的书签。

[0186] (附加 2). 根据附加 1 所述的书签提取装置,其中

[0187] 基于在所述浏览历史信息中包括的、关于当前正在浏览的 web 站点的信息,所述关键字提取单元获取所述当前浏览的 web 站点的内容,并且从所述内容中提取所述关键字。

[0188] (附加 3). 根据附加 2 所述的书签提取装置,其中

[0189] 所述关键字提取单元包括重要关键字提取单元,用于提取在从所述当前浏览的 web 站点获取的所述内容中包括的重要关键字作为所述关键字,并且其中

[0190] 所述提供单元基于所述重要关键字,从所述多个书签中提供与所述重要关键字相关的书签。

[0191] (附加 4). 根据附加 3 所述的书签提取装置,其中

[0192] 所述重要关键字提取单元根据在所述内容中出现的频率确定所述重要关键字。

[0193] (附加 5). 根据附加 1 所述的书签提取装置,其中,

[0194] 通过分析所述浏览历史信息,当当前正在浏览的 web 站点被确定为已经从搜索站点进行了转变时,所述关键字提取单元提取在到所述当前浏览的 web 站点的所述转变之前已经使用的搜索关键字作为所述关键字,并且其中

[0195] 所述提供单元基于所述搜索关键字,从所述多个书签中提供与所述搜索关键字相关的书签。

[0196] (附加 6). 根据附加 5 所述的书签提取装置,其中,

[0197] 当恰好到所述当前浏览的 web 站点的所述转变之前已经被浏览的 web 站点是搜索站点时,所述关键字提取单元提取已经在所述搜索站点中使用的搜索关键字作为所述关键字。

[0198] (附加 7). 根据附加 1 至 6 之任一所述的书签提取装置,其中

[0199] 所述提供单元包括相关性确定单元,用于根据在每个书签的所述属性信息中所述关键字被包括的程度,确定在所述关键字和在关于所述多个书签的信息中包括的书签中的所述每个书签之间的相关性的存在。

[0200] (附加 8). 根据附加 1 至 6 之任一所述的书签提取装置,其中

[0201] 所述提供单元包括相关性确定单元,用于根据与关注的所述书签链接的 web 站点的所述内容中的所述关键字被包括的程度,确定在所述关键字和所述每个书签之间相关性的存在。

[0202] (附加 9). 根据附加 1 至 6 之任一所述的书签提取装置,其中

[0203] 所述提供单元包括相关性确定单元,用于根据在所述每个书签的所述属性信息中的所述关键字被包括的程度以及与关注的所述书签链接的 web 站点的所述内容中的所述关键字被包括的程度的组合,确定在所述关键字和所述每个书签之间相关性的存在。

[0204] (附加 10). 根据附加 7 至 9 之任一所述的书签提取装置,其中

[0205] 所述提供单源以排序的方式,提供已经由所述相关性确定单元确定为具有相关性的多个书签。

[0206] (附加 11). 根据附加 7 所述的书签提取装置,其中,

[0207] 基于所述相关性确定单元确定的结果,所述提供单元提供给用户与来自所述书签注册的时间的所述书签对应的 web 站点的内容改变的可能性的信息,或者到所述 web 站点的链接失效的信息。

[0208] (附加 12). 一种书签提取方法,包括:

[0209] 基于 web 站点至今的浏览历史信息提取关键字;以及

[0210] 基于所述关键字,从多个提前注册的书签中提供与所述提取的关键字相关的书签。

[0211] (附加 13). 根据附加 12 所述的书签提取方法,其中,

[0212] 当提取所述关键字时,基于在所述浏览历史信息中包括的、关于当前浏览的 web 站点的信息,获取所述当前浏览的 web 站点的内容,从所述内容中提取所述关键字。

[0213] (附加 14). 根据附加 11 所述的书签提取方法,其中,

[0214] 当提取所述关键字时,作为分析所述浏览历史信息的结果,在当前浏览的 web 站点被确定为已经从搜索站点进行了转变的情况下,提取在到所述当前浏览的 web 站点的所述转变之前已经使用的搜索关键字作为所述关键字,并且其中

[0215] 基于所述搜索关键字,从所述多个书签中提供与所述搜索关键字相关的书签。

[0216] (附加 15). 根据附加 14 所述的书签提取方法,其中,

[0217] 当提取所述关键字时,在恰好到所述当前浏览的 web 站点的所述转变之前已经被浏览的 web 站点是搜索站点的情况下,提取已经在所述搜索站点中使用的搜索关键字作为所述关键字。

[0218] (附加 16). 一种用于控制书签提取装置的操作的计算机程序,所述计算机程序使得计算机实现:

[0219] 关键字提取功能,用于基于 web 站点至今的浏览历史信息提取关键字;以及

[0220] 提供功能,用于基于所述关键字,从多个注册的书签中提供与所述关键字提取功能所提取的所述关键字相关的书签。

[0221] (附加 17). 根据附加 16 所述的计算机程序,其中,

[0222] 基于在所述浏览历史信息中包括的、关于当前浏览的 web 站点的信息,所述关键字提取功能获取所述当前浏览的 web 站点的内容,并且从所述内容中提取所述关键字。

[0223] (附加 18). 根据附加 16 所述的计算机程序,其中,

[0224] 作为分析所述浏览历史信息的结果,当当前浏览的 web 站点被确定为已经从搜索站点进行了转变时,所述关键字提取功能提取在到所述当前浏览的 web 站点的所述转变之前已经使用的搜索关键字作为所述关键字,并且其中

[0225] 所述提供功能基于所述搜索关键字,从所述多个书签中提供与所述搜索关键字相关的书签。

[0226] (附加 19). 根据附加 18 所述的计算机程序,其中,

[0227] 当恰好到所述当前浏览的 web 站点的所述转变之前浏览的 web 站点是搜索站点的情况下,所述关键字提取功能提取已经在所述搜索站点中使用的搜索关键字作为所述关键

字。

[0228] 提供上文描述的实施方式以使得本领域技术人员进行和使用本发明。此外,对这些示例性实施方式的各种变型对本领域技术人员来说是明显的,并且在此定义的普通原理和具体示例可以应用到其他实施方式中,而不偏离发明的实质。因而,本发明并不旨在限制于在此描述的示例性实施方式,而是与权利要求及其等同物的限制所定义的最宽范围一致。

[0229] 进而,注意发明人的意图是即使在诉讼期间修改权利要求,也保留所有所要求的发明的等同物。

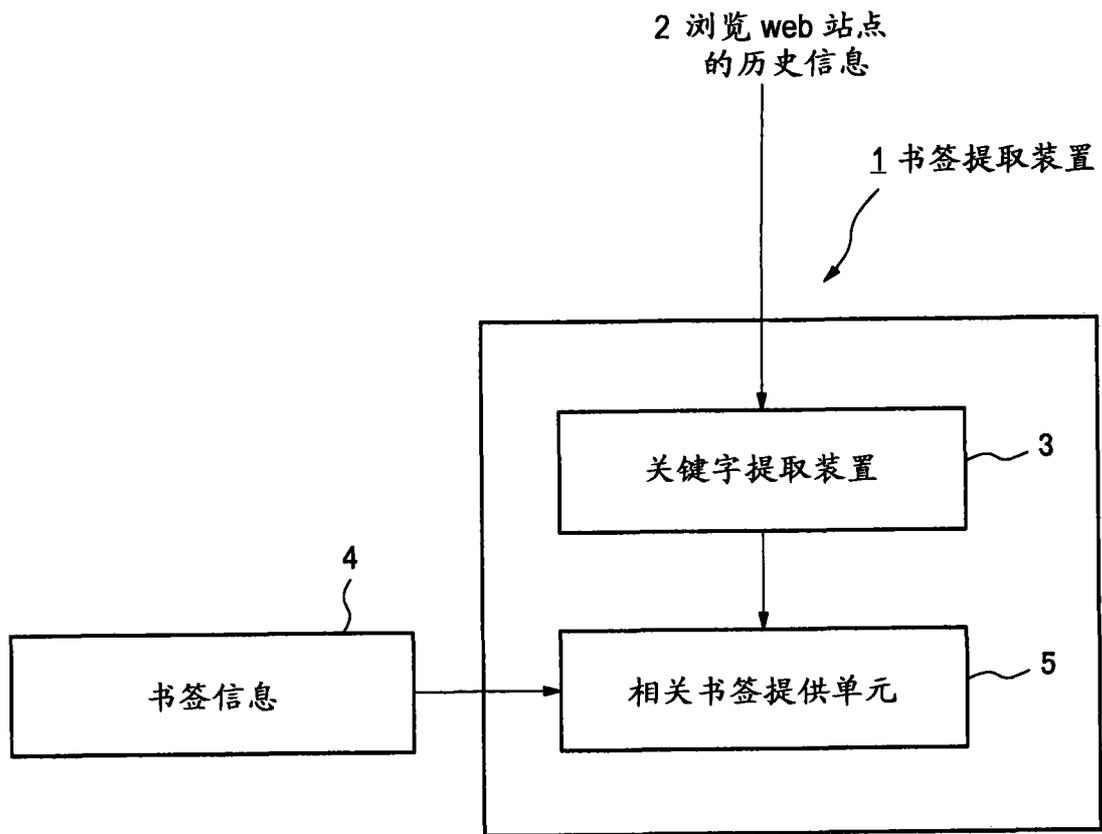


图 1

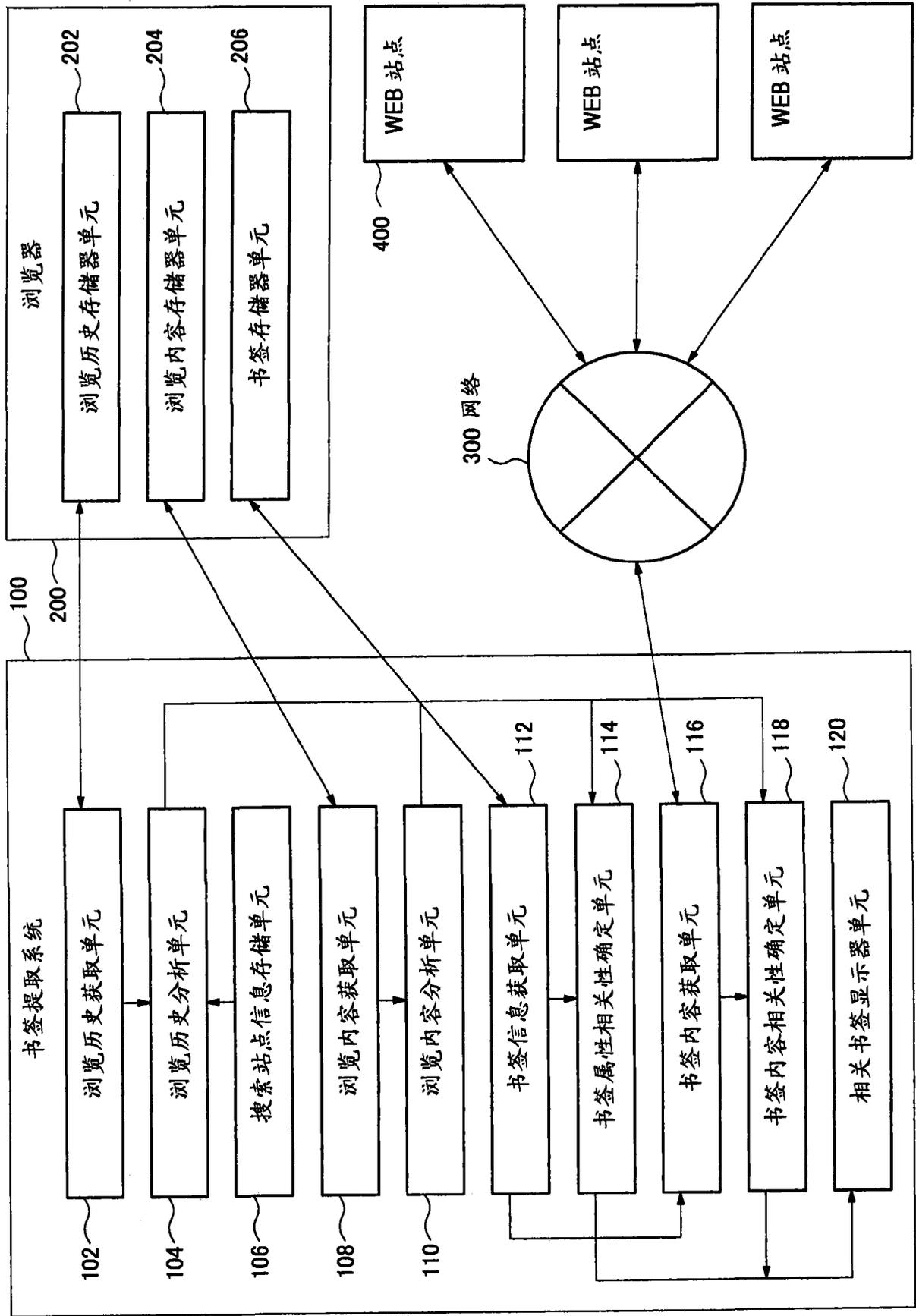


图 2

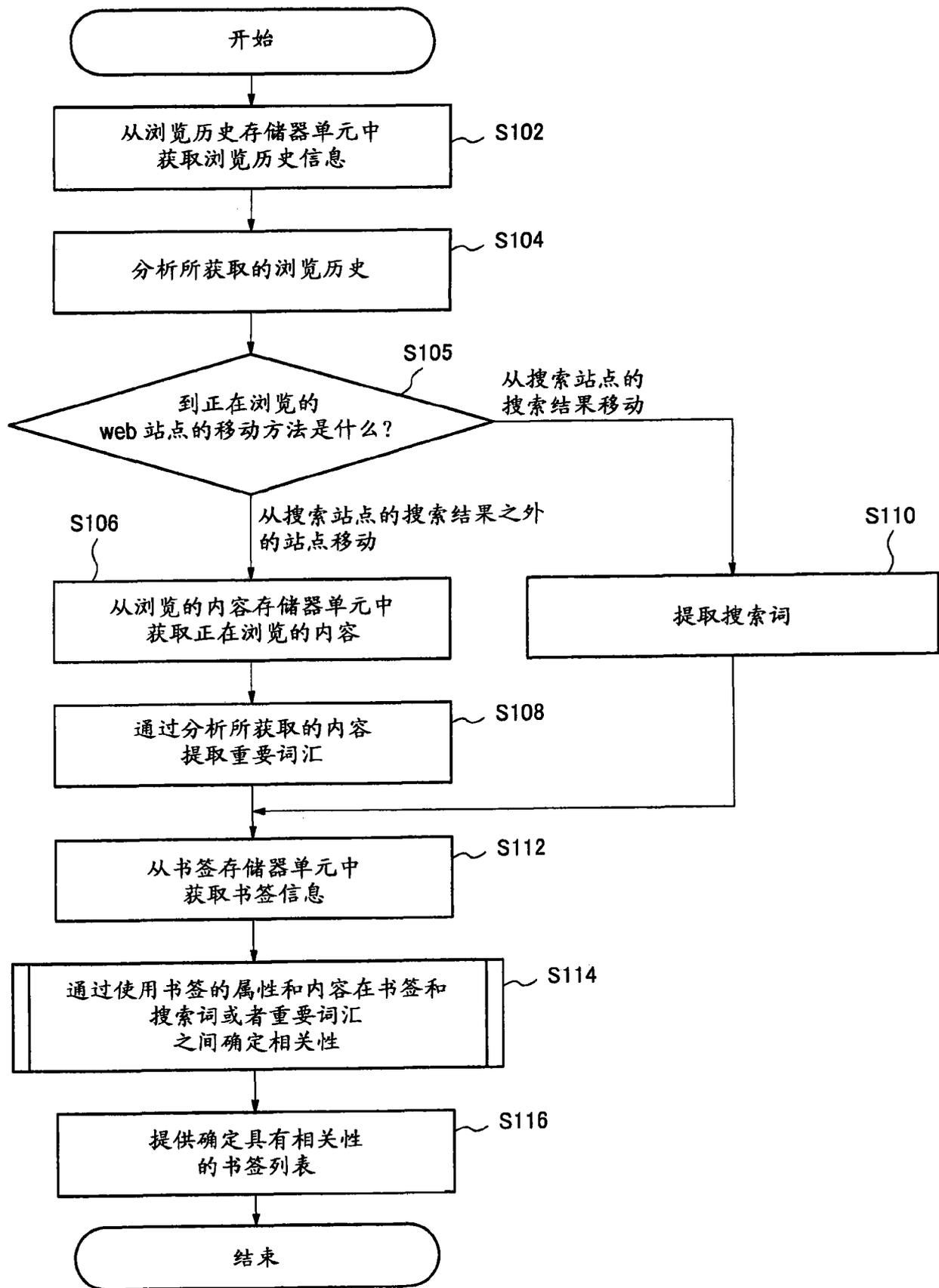


图 3

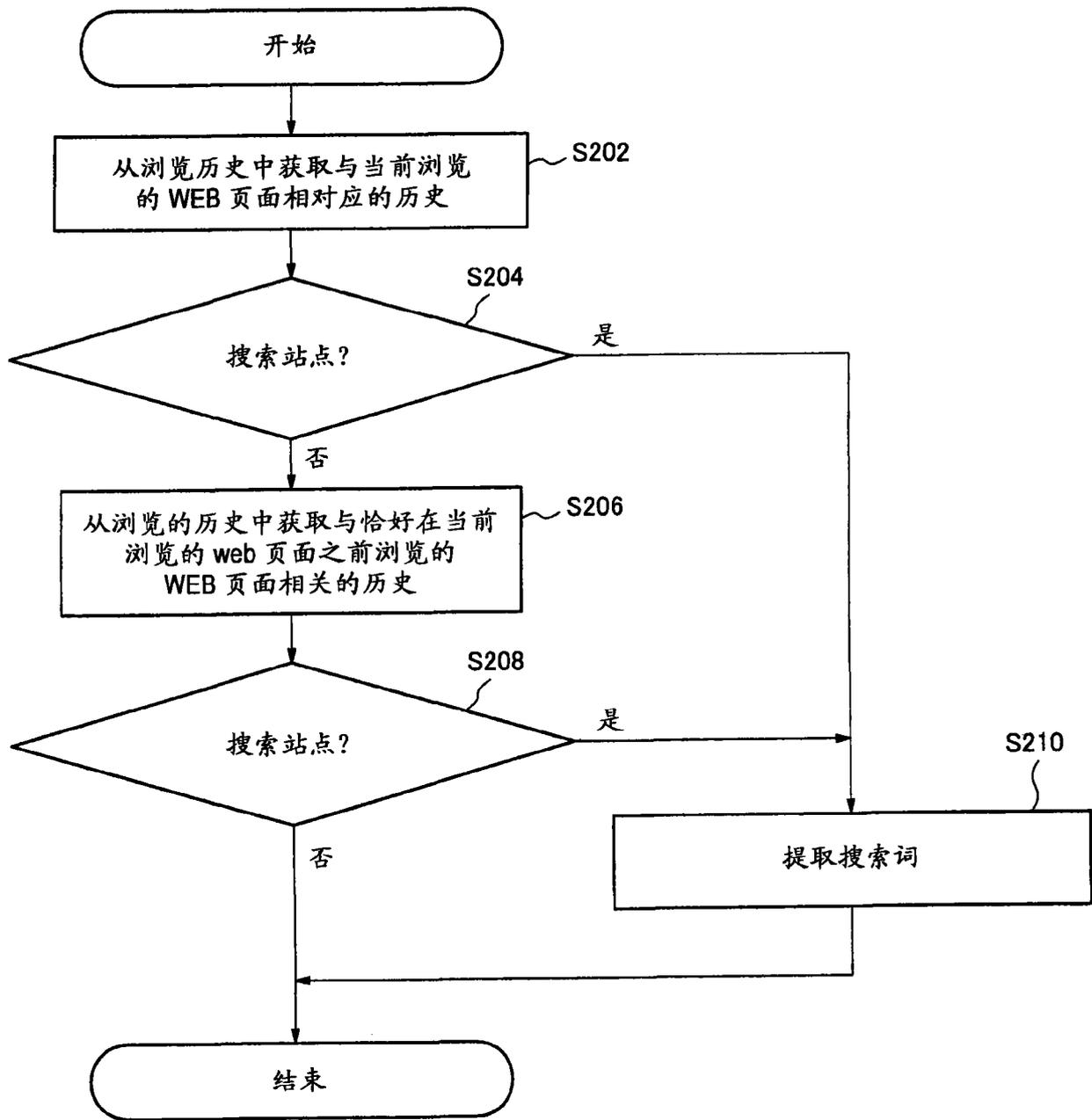


图 4

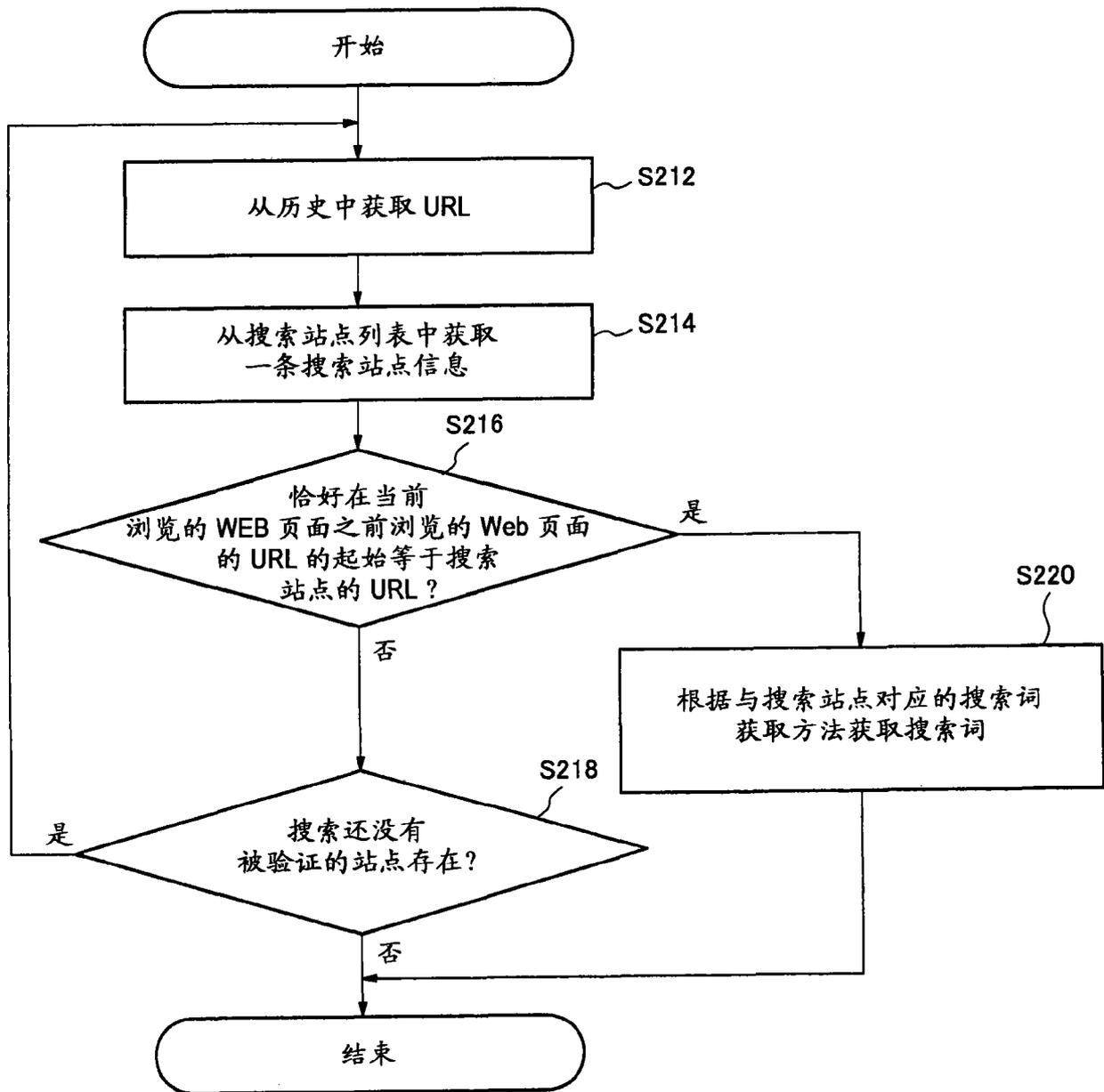


图 5

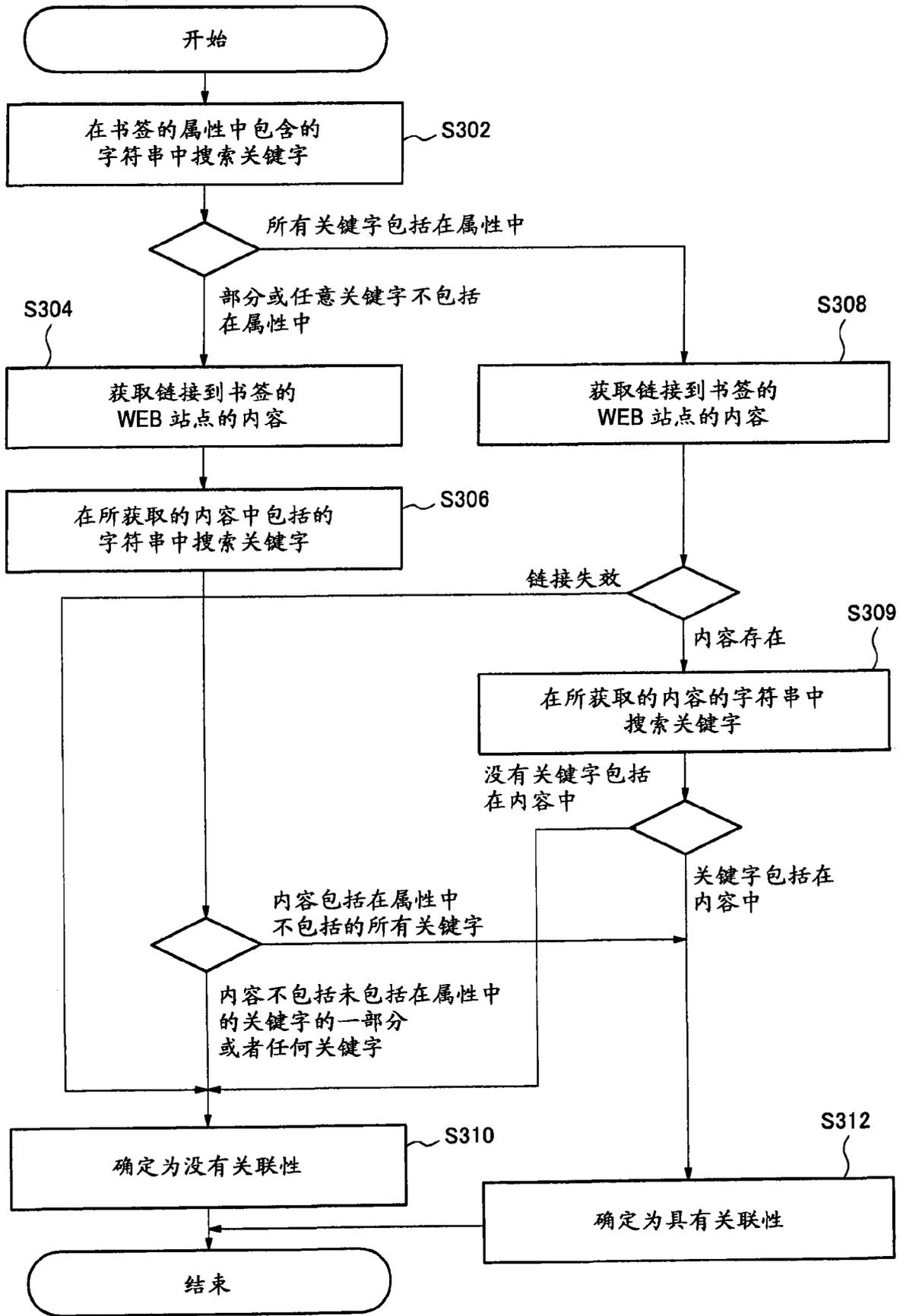


图 6

50 搜索站点信息

ID	URL	搜索词提取方法
1	http://www.google.co.jp/search	query:q:+
2	http://search.yahoo.co.jp/search	query:p:+

图 7

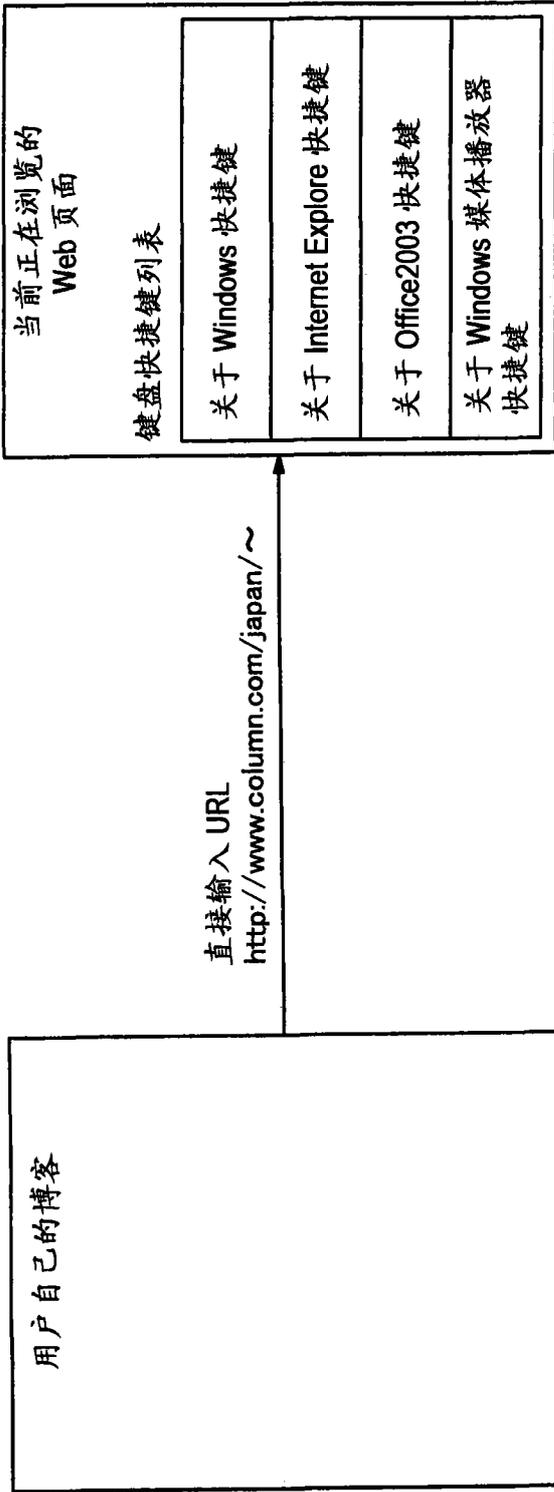


图 8A

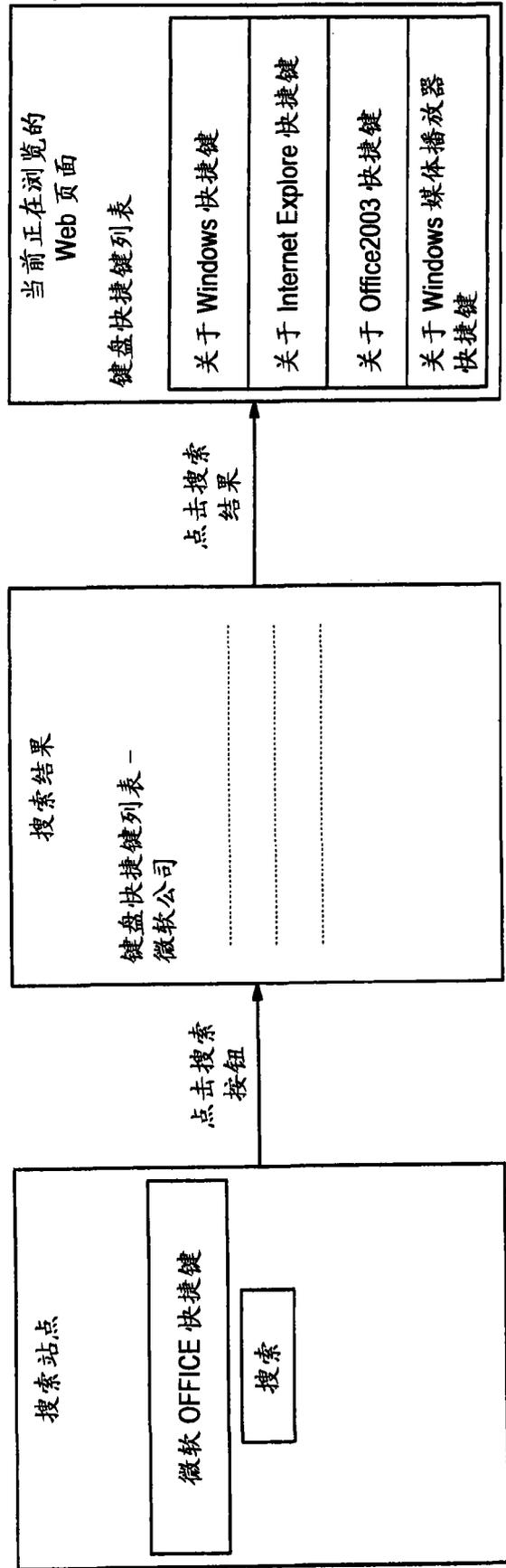


图 8B

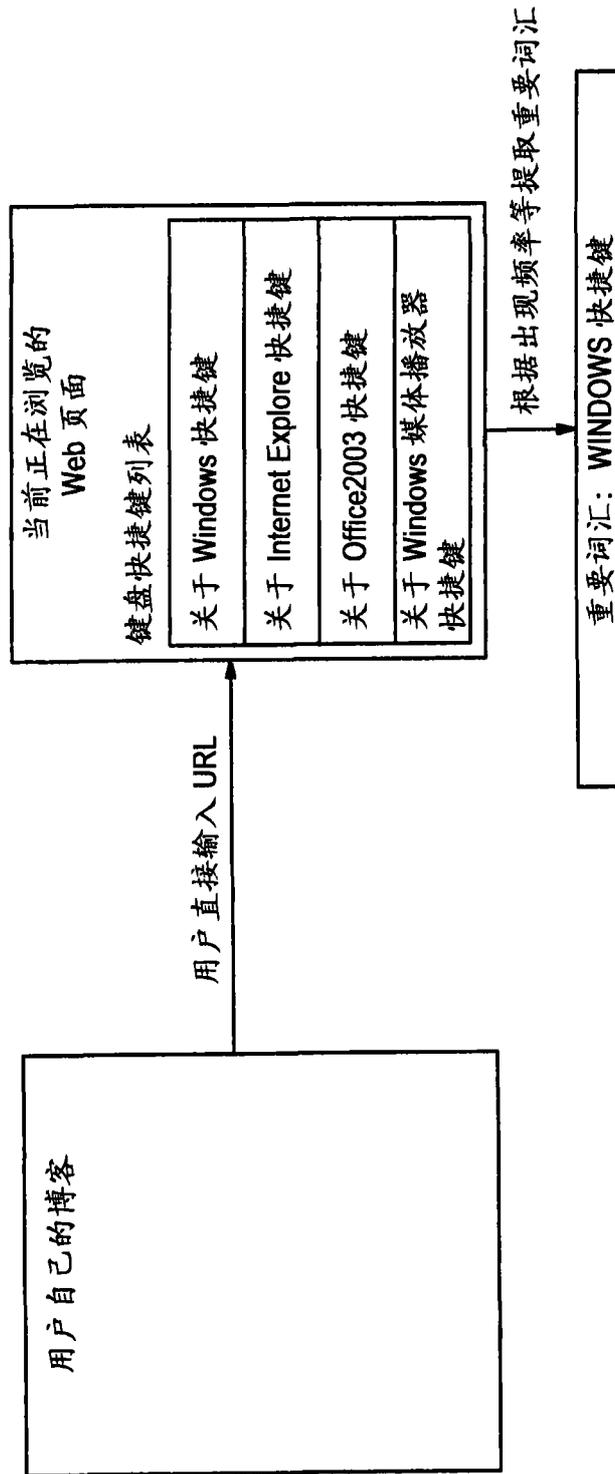


图 9A

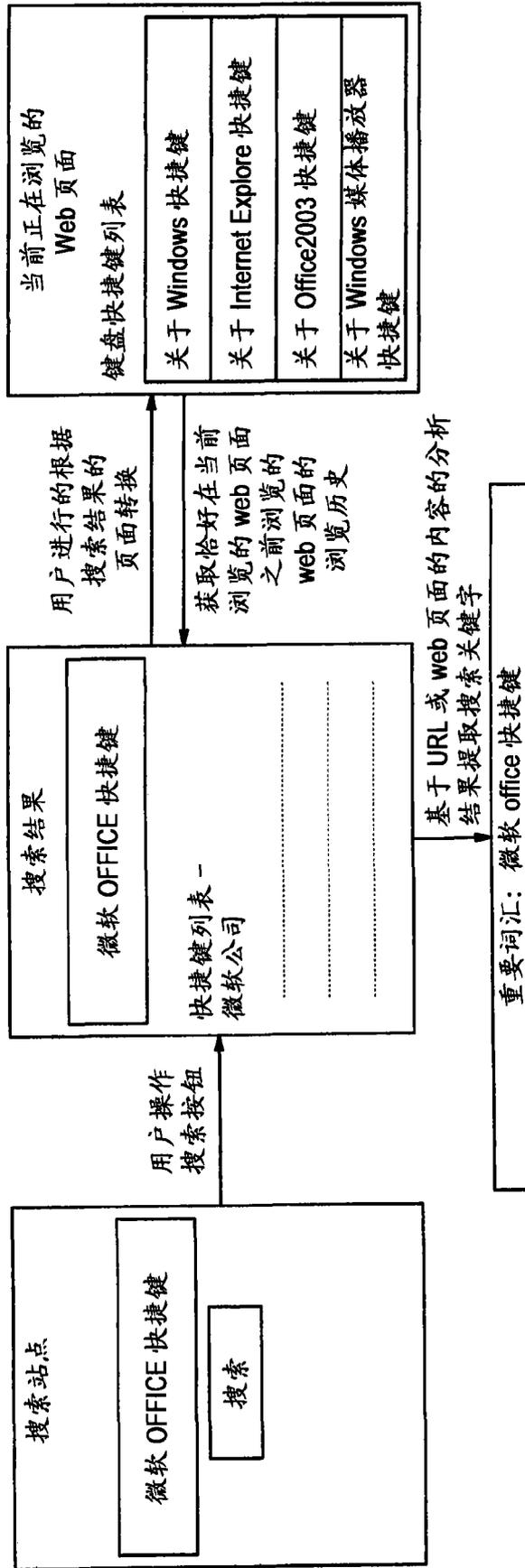


图 9B

书签 1
标题: 快捷键
URL: http://example.blog.co.jp/hoge
文件夹路径: MICROSOFT/WINDOWS
标签: 无
关键字: 无
其他属性: 无
链接目的地内容: 关于 WINDOWS 快捷键 WINDOWS 键 +R: 锁屏 WINDOWS 键 +D: 显示最上层桌面

图 10A

书签 2
标题: 快捷键列表
URL: http://example.blog.co.jp/hoge
文件夹路径: MICROSOFT/WINDOWS
标签: 无
关键字: 无
其他属性: 无
链接目的地内容: 失效的链接

图 10B

书签 3
标题: 键盘快捷方式列表
URL: http://example.blog.co.jp/hoge
文件夹路径: MICROSOFT/IE
标签: 无
关键字: 无
其他属性: 无
链接目的地内容: 关于微软 OFFICE 快捷键 CTRL+S: 保存文件, 覆盖 Windows 快捷键 Windows 键 +R: 锁屏

图 10C

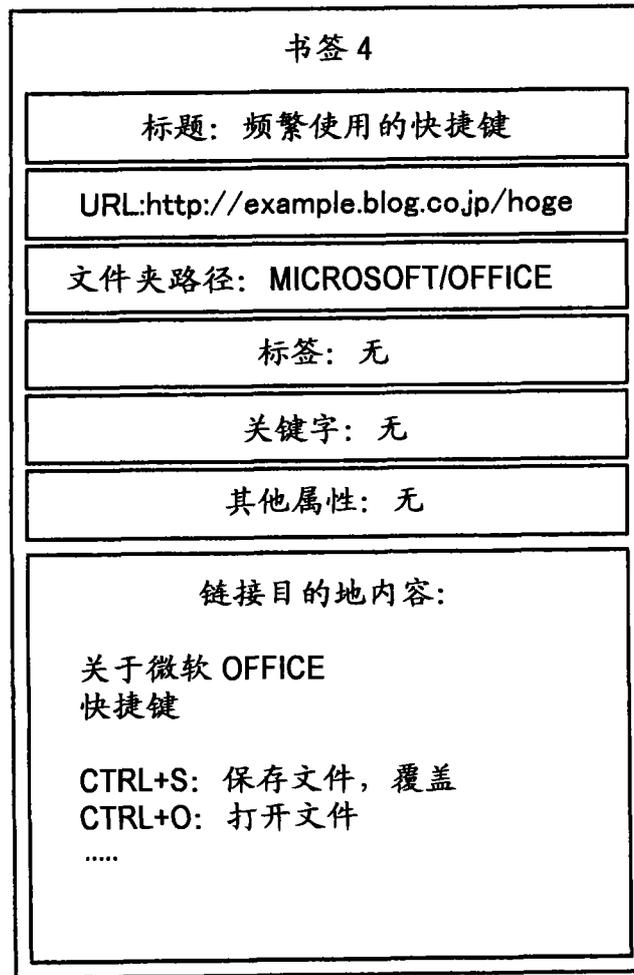


图 10D

书签 5
标题: WINDOWS 手册
URL: http://www.manual.co.jp/windows
文件夹路径: MICROSOFT/WINDOWS
标签: 无
关键字: 无
其他属性: 无
链接目的地内容: 关于 WINDOWS 快捷键 Windows 键 +R: 锁屏 Windows 键 +D: 显示最上层桌面

图 10E

书签 6
标题: 快捷方式
URL: http://example.blog.co.jp/hoge
文件夹路径: WORK/MANUAL/ DOCEMENT/...MICROSOFT
标签: 无
关键字: 无
其他属性: 无
链接目的地内容:
关于微软 OFFICE 快捷键 CTRL+S: 保存文件, 覆盖
关于 WINDOWS 快捷键 Windows 键 +R: 锁屏

图 10F

“WINDOWS 快捷键” 的相关书签	当前正在浏览的 Web 页面				
书签 1	<p>键盘快捷键列表</p> <table border="1"><tbody><tr><td data-bbox="651 452 1378 580">关于 windows 快捷键</td></tr><tr><td data-bbox="651 584 1378 712">关于 internet explore 快捷键</td></tr><tr><td data-bbox="651 716 1378 844">关于 office 2003 快捷键</td></tr><tr><td data-bbox="651 848 1378 976">关于 windows 媒体播放器快捷键</td></tr></tbody></table>	关于 windows 快捷键	关于 internet explore 快捷键	关于 office 2003 快捷键	关于 windows 媒体播放器快捷键
关于 windows 快捷键					
关于 internet explore 快捷键					
关于 office 2003 快捷键					
关于 windows 媒体播放器快捷键					
书签 2					
书签 3					
书签 6					

图 11

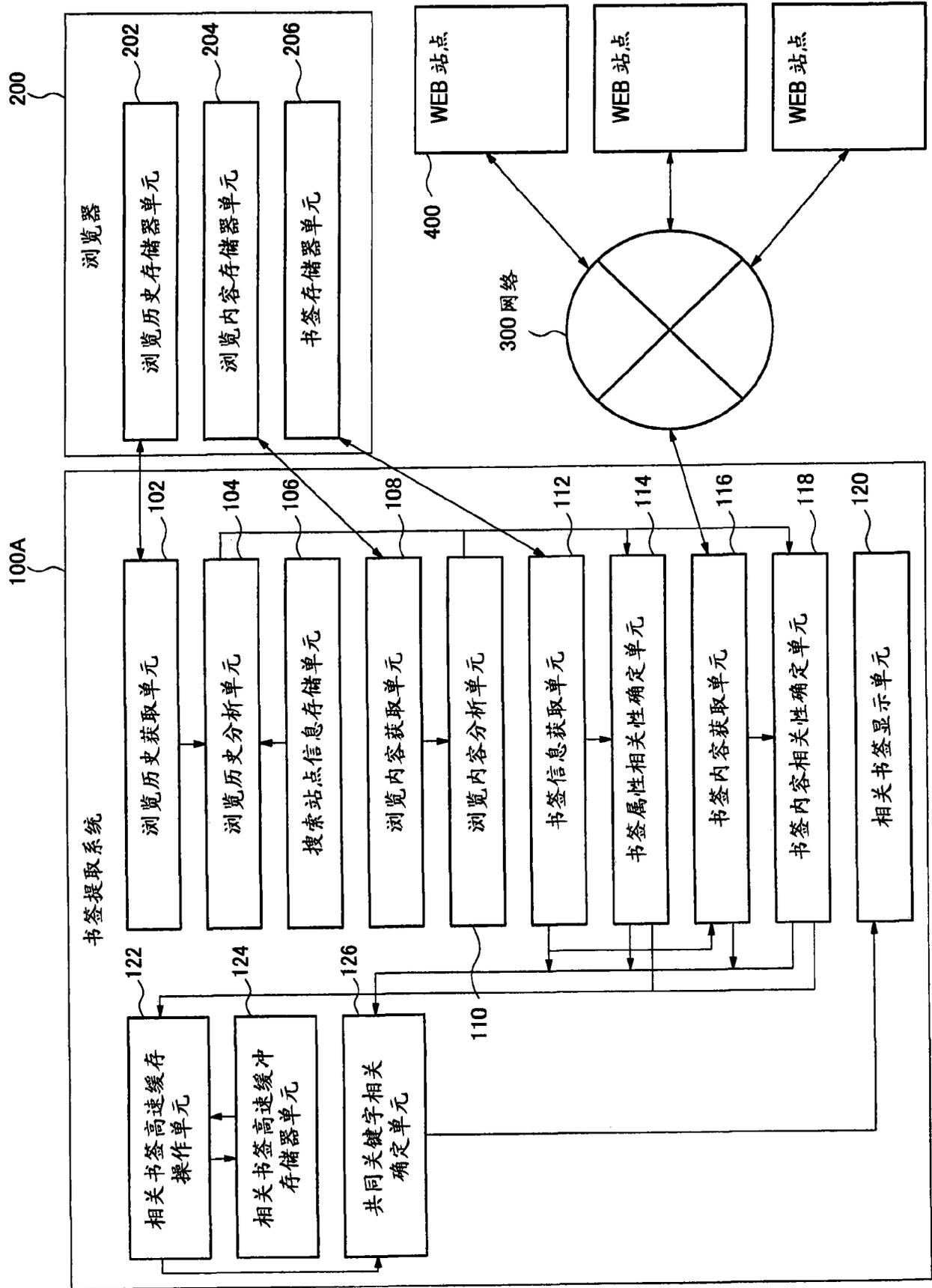


图 12

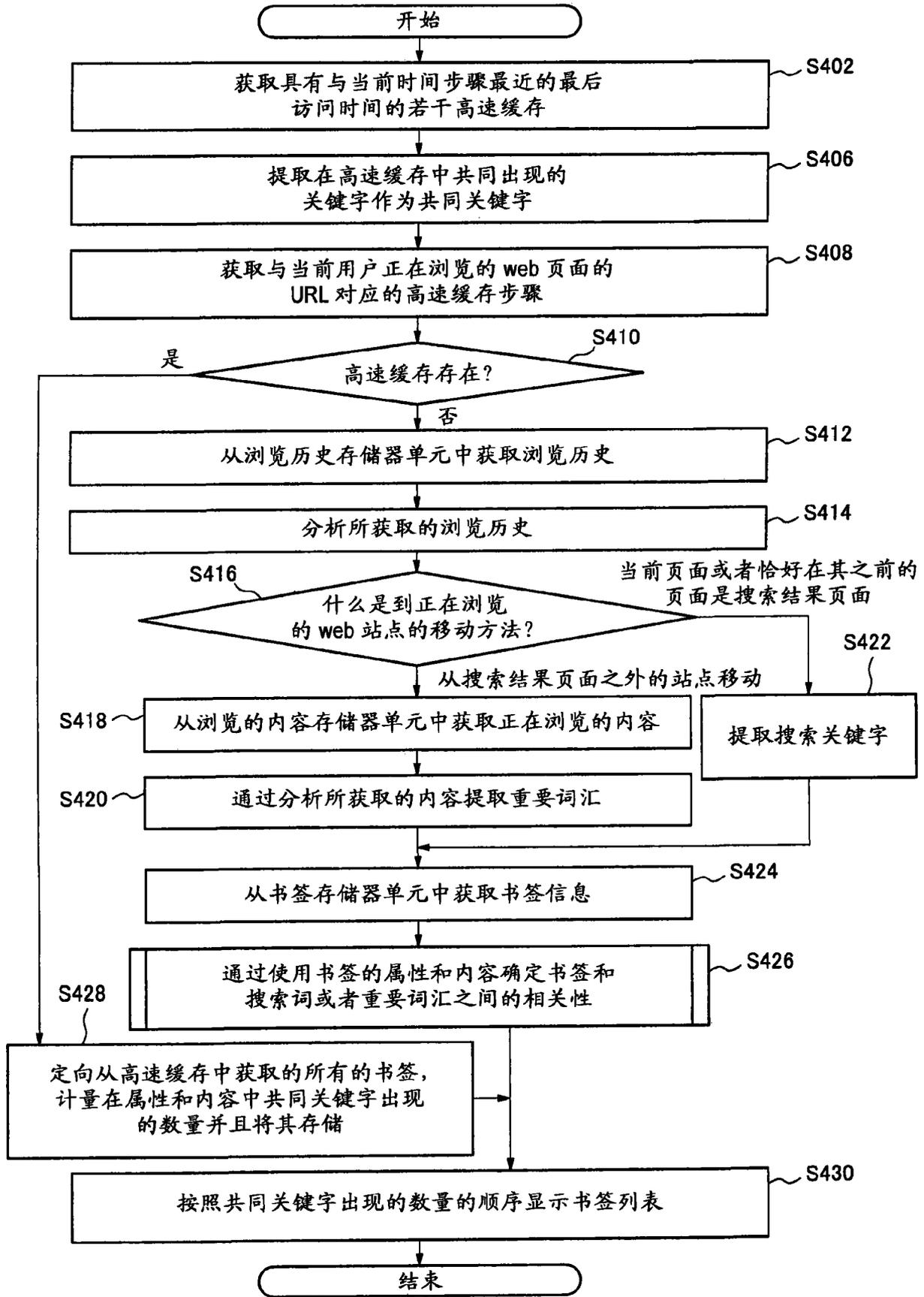


图 13

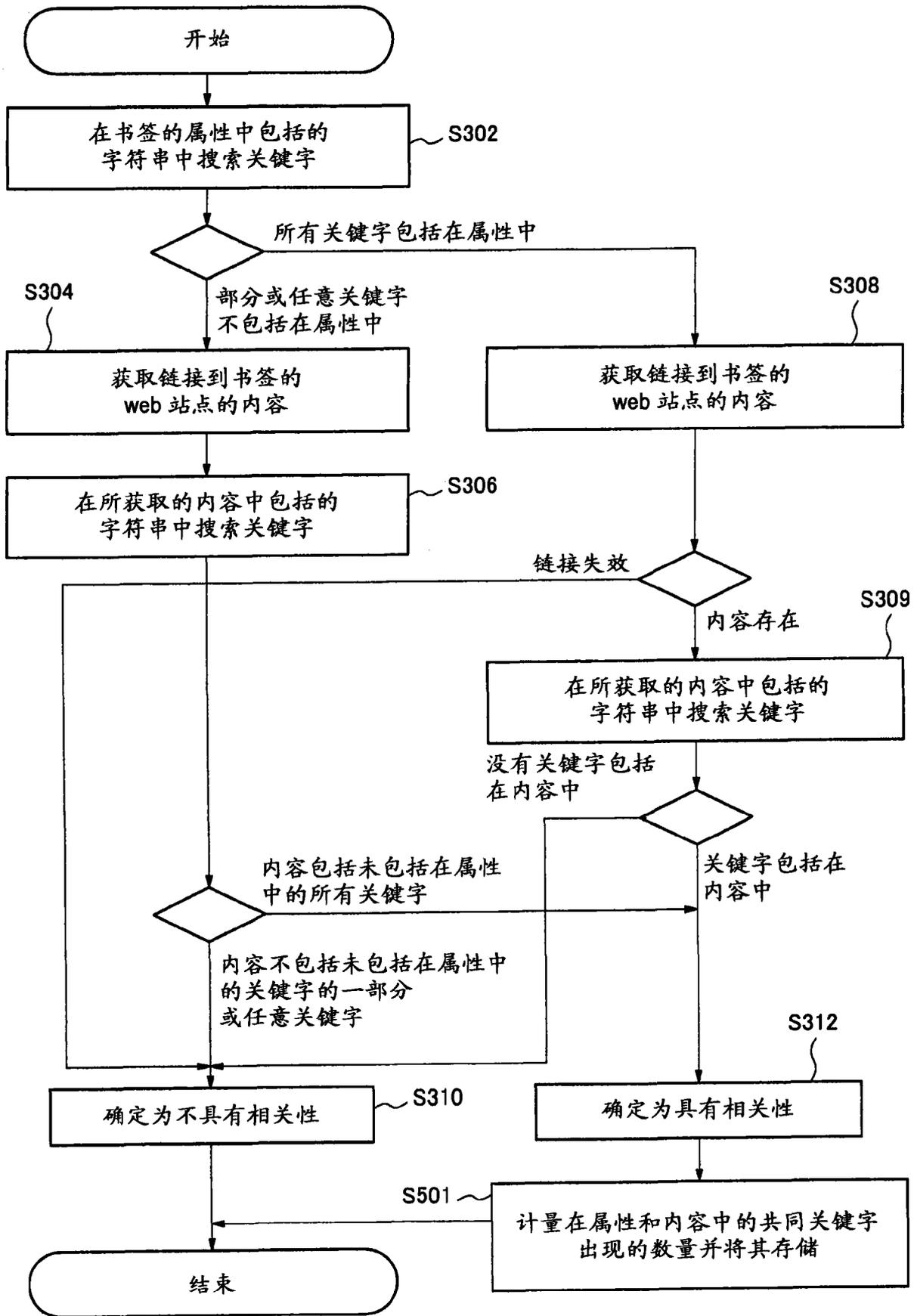


图 14

<p>“WINDOWS 快捷键” 的相关书签</p>	<p>当前正在浏览的 Web 页面</p> <p>键盘快捷键列表</p> <table border="1" data-bbox="651 454 1378 976"> <tr> <td data-bbox="651 454 1378 584">关于 windows 快捷键</td> </tr> <tr> <td data-bbox="651 584 1378 714">关于 internet explore 快捷键</td> </tr> <tr> <td data-bbox="651 714 1378 844">关于 office 2003 快捷键</td> </tr> <tr> <td data-bbox="651 844 1378 976">关于 windows 媒体播放器快捷键</td> </tr> </table>	关于 windows 快捷键	关于 internet explore 快捷键	关于 office 2003 快捷键	关于 windows 媒体播放器快捷键
关于 windows 快捷键					
关于 internet explore 快捷键					
关于 office 2003 快捷键					
关于 windows 媒体播放器快捷键					
<p>(等级 1) 书签 1</p>					
<p>(等级 2) 书签 5</p>					
<p>(等级 3) 书签 3</p>					
<p>在注册之后内容 改变的可能性</p>					
<p>书签 3</p>					
<p>链接失效</p>					
<p>书签 2</p>					

图 16

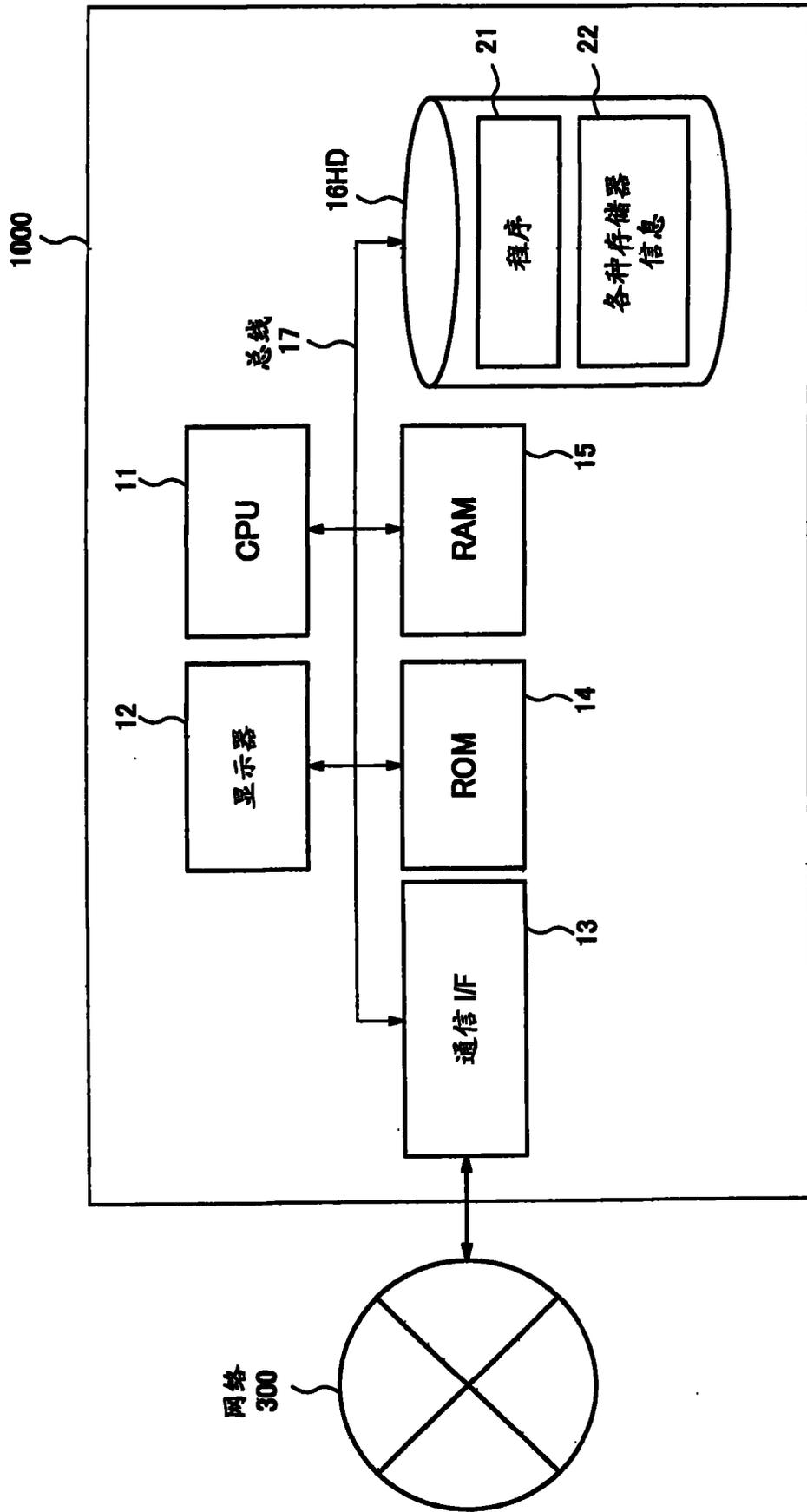


图 17