

[51] Int. Cl.
G06F 17/30 (2006.01)



专利号 ZL 200610066682.1

[11] 授权公告号 CN 100524304C

[74] 专利代理机构 北京怡丰知识产权代理有限公司

代理人 于振强

[32] 2005. 4. 20 [33] JP [31] 122948/2005

地址 日本东京都

[56] 参考文献

CN1588867A 2005.3.2

CN1435782A 2003.8.13

US6055535A 2000.4.25

DE10317408A1 2004.11.25

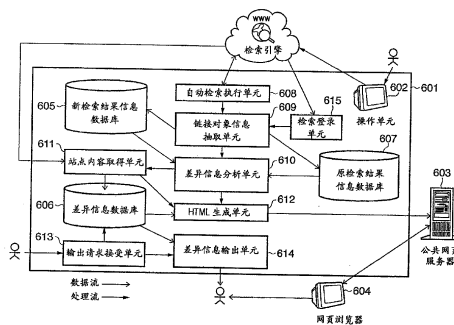
CN1489734A 2004.4.14

审查员 董 刚

权利要求书 3 页 说明书 16 页 附图 13 页

信息处理设备及信息处理方法

本发明提供一种信息处理设备及信息处理方法，能够容易且准确地将不同的检索时期进行了基于同一检索条件的检索后生成的检索结果的差异通知给用户。首先，输入预定检索条件检索网络上的信息，取得第1检索结果。预先登录在进行检索处理时所使用的检索条件。然后，在取得第1检索结果后，用所登录的检索条件对网络上的信息执行再次检索，取得再次检索到的第2检索结果。然后，比较第1检索结果和第2检索结果，检测检索到的信息的差异信息。



1. 一种可经由网络进行通信的信息处理设备，其特征在于，包括：

检索装置，输入检索条件后对上述网络上的信息执行检索处理；

登录装置，登录由上述检索装置执行检索处理时使用的上述检索条件；

第1检索结果取得装置，取得由上述检索装置用上述检索条件检索到的第1检索结果；

再次检索装置，在取得上述第1检索结果后，用由上述登录装置所登录的上述检索条件对上述网络上的信息执行再次检索；

第2检索结果取得装置，取得由上述再次检索装置用上述检索条件再次检索到的第2检索结果；以及

差异检测装置，比较上述第1检索结果和第2检索结果，检测检索到的信息的差异信息。

2. 根据权利要求1所述的信息处理设备，其特征在于，

上述检索装置和再次检索装置，检索上述网络上的网站，

上述差异检测装置，比较基于上述第1检索结果的链接对象的URL群组和基于上述第2检索结果的链接对象的URL群组，检测新检索到的网站和因消失而没有再次检索到的网站，

还包括输出装置，输出上述新检索到的网站和上述因消失而没有再次检索到的网站的URL的一览信息，作为上述差异信息。

3. 根据权利要求2所述的信息处理设备，其特征在于，还包括：

HTML生成装置，将上述新检索到的网站的URL和上述因消失而没有再次检索到的网站的URL的一览信息，作为HTML文件生成；以及

差异站点公开装置，用公共网页服务器公开由上述HTML生成装置所生成的HTML文件。

4. 根据权利要求3所述的信息处理设备，其特征在于，还包括：

显示装置,将由上述公共网页服务器公开的上述 HTML 文件显示在显示画面上。

5. 根据权利要求 1 所述的信息处理设备,其特征在于,还包括:形成装置,根据由上述差异检测装置检测到的上述差异信息,在记录介质上形成反映了上述差异信息的检索结果的图像。

6. 根据权利要求 1 所述的信息处理设备,其特征在于,上述再次检索装置,用上述登录装置所登录的检索条件定期地按固定间隔进行再次检索;

上述第 2 检索结果取得装置,取得由上述再次检索装置定期进行再次检索得到的第 2 检索结果,

其中,所述信息处理设备还包括:指示装置,从用户那里接受关于由上述再次检索装置用上述检索条件进行检索的上述定期的间隔的指示。

7. 根据权利要求 1 所述的信息处理设备,其特征在于,上述再次检索装置,用由上述登录装置所登录的上述检索条件,以接受到差异检索请求这一情况为条件执行再次检索。

8. 根据权利要求 3 所述的信息处理设备,其特征在于,上述 HTML 生成装置,将由上述差异检测装置检测到的网站转换成缩略图像,生成包括该缩略图像的 HTML 文件。

9. 根据权利要求 3 所述的信息处理设备,其特征在于,上述 HTML 生成装置,抽取由上述差异检测装置检测到的网站上的包括上述检索条件的字符串。

10. 一种可经由网络进行通信的信息处理设备的信息处理方法,其特征在于,包括:

检索步骤,输入检索条件后对上述网络上的信息执行检索处理;

登录步骤,登录在上述检索步骤中执行检索处理时使用的上述检索条件;

第 1 检索结果取得步骤,取得在上述检索步骤中用上述检索条件检索到的第 1 检索结果;

再次检索步骤，在取得上述第 1 检索结果后，用在上述登录步骤中所登录的上述检索条件对上述网络上的信息执行再次检索；

第 2 检索结果取得步骤，取得在上述再次检索步骤中用上述检索条件再次检索到的第 2 检索结果；以及

差异检测步骤，比较上述第 1 检索结果和第 2 检索结果，检测检索到的信息的差异信息。

信息处理设备及信息处理方法

技术领域

本发明涉及可安装于经由网络进行通信的多功能设备的信息处理设备及信息处理方法。

背景技术

近年来，以数字复印机为基础，添加了多种功能的多功能设备已被实际应用。该多功能设备具有以下功能：例如，使用数字复印机本来具备的扫描仪、打印机进行传真通信；将从计算机发送来的代码数据展开成位图数据，用打印机进行打印输出；将使用扫描仪读入的图像发送给网络上的其他装置；以及使用打印机打印从网络递送来的图像。

此外，近年来多功能设备的操作板正朝大型化、高精细化方向发展。不仅可以利用扫描仪、打印机对功能进行操作设定，还可以在其操作画面上进行网络上的远程设备的操作、用于监视设备状态的显示、以及基于内部浏览器的网页内容的显示和打印操作。作为与本发明有关的专利文献，有日本特开平 11-134125 号公报。

但是，当通过搭载于上述多功能设备、个人计算机等上的网页浏览器功能，进行使用了搜索引擎的网站检索时，即使按照搜索引擎侧的数据库的内容以相同的关键字进行检索，有时检索结果也不同。

一般地，在 WWW 上总是有新出现的网站、已消失的网站、或者仍然存在但其内容已被更新的网站。并且，搜索引擎侧的数据库的数据，一般是在 WWW 上自动爬行（crawling）而收集到的。即，该数据相当于自动爬行时的快照（snapshot），因此，根据时期的不同数据库的内容往往不同。因此，即使在用户使用某检索条件通过搜索引擎执行了检索时，也可以有根据其检索时期的不同适合该检索条件的

网站不同的情况。

因此，通过分析这种由检索时期引起的检索结果的差异（difference），例如能够了解新出现的站点。但是，当用户要手动分析这种由检索时期引起的检索结果的不同（差异信息）时，检索结果通常是按照与网站的出现时期无关的顺序列举，由此，产生其确认作业所需的负担大这样的问题。此外，需要用户对例如检索时期不同的2个检索结果的URL一个一个地一边观察一边核对。虽然在近年来的多功能设备等信息处理设备中已嵌入了性能较高的屏，但如果用户通过观察确认检索结果，则从屏的大小、可视性、以及操作性方面来看还不够完善，从而产生随着检索结果数量的增加用户的负担进一步加重这样的问题。

发明内容

本发明是考虑到上述问题而完成的，其目的在于，提供一种信息处理设备及信息处理方法，能够容易且准确地将在不同的检索时期进行了基于同一检索条件的检索后生成的检索结果的差异通知给用户。

为了解决上述课题，本发明提供一种可经由网络进行通信的信息处理设备，其特征在于，包括：检索装置，输入检索条件后对上述网络上的信息执行检索处理；登录装置，登录由上述检索装置执行检索处理时使用的上述检索条件；第1检索结果取得装置，取得由上述检索装置用上述检索条件检索到的第1检索结果；再次检索装置，在取得上述第1检索结果后，用由上述登录装置所登录的上述检索条件对上述网络上的信息执行再次检索；第2检索结果取得装置，取得由上述再次检索装置用上述检索条件再次检索到的第2检索结果；以及差异检测装置，比较上述第1检索结果和第2检索结果，检测检索到的信息的差异信息。

优选的是，上述检索装置和再次检索装置，检索上述网络上的网站，上述差异检测装置，比较基于上述第1检索结果的链接对象的URL群组和基于上述第2检索结果的链接对象的URL群组，检测新检索到的网站和因消失而没有再次检索到的网站，还包括输出装置，

输出上述新检索到的网站和上述因消失而没有再次检索到的网站的 URL 的一览信息，作为上述差异信息。

优选的是，还包括：HTML 生成装置，将上述新检索到的网站的 URL 和上述因消失而没有再次检索到的网站的 URL 的一览信息，作为 HTML 文件生成；以及差异站点公开装置，用公共网页服务器公开由上述 HTML 生成装置所生成的 HTML 文件。

优选的是，上述信息处理设备还包括形成装置，根据由上述差异检测装置检测到的上述差异信息，在记录介质上形成反映了上述差异信息的检索结果的图像。

优选的是，上述再次检索装置，用上述登录装置所登录的检索条件定期地按固定间隔进行再次检索；上述第 2 检索结果取得装置，取得由上述再次检索装置定期进行再次检索得到的第 2 检索结果，其中，所述信息处理设备还包括：指示装置，从用户那里接受关于由上述再次检索装置用上述检索条件进行检索的上述定期的间隔的指示。

优选的是，上述再次检索装置，用由上述登录装置所登录的上述检索条件，以接受到差异检索请求这一情况为条件执行再次检索。

优选的是，上述 HTML 生成装置，将由上述差异检测装置检测到的网站转换成缩略图像，生成包括该缩略图像的 HTML 文件。

优选的是，上述 HTML 生成装置，抽取由上述差异检测装置检测到的网站上的包括上述检索条件的字符串。

此外，为了解决上述课题，本发明还提供一种可经由网络进行通信的信息处理设备的信息处理方法，其特征在于，包括：检索步骤，输入检索条件后对上述网络上的信息执行检索处理；登录步骤，登录在上述检索步骤中执行检索处理时使用的上述检索条件；第 1 检索结果取得步骤，取得在上述检索步骤中用上述检索条件检索到的第 1 检索结果；再次检索步骤，在取得上述第 1 检索结果后，用在上述登录步骤中所登录的上述检索条件对上述网络上的信息执行再次检索；第 2 检索结果取得步骤，取得在上述再次检索步骤中用上述检索条件再次检索到的第 2 检索结果；以及差异检测步骤，比较上述第 1 检索结

果和第2检索结果，检测检索到的信息的差异信息。

本发明的其他特征和优点，将通过以下参照附图的说明而得到明确，附图中相同或相似的部分赋予相同的标号。

附图说明

附图包含在说明书中，构成本说明书的一部分，用于说明本发明的实施方式，并与其说明一起用于阐明本发明的原理。

图1是示意地表示本发明的1个实施例中、具有安装了信息处理设备的多功能设备的网络系统的结构的结构图。

图2是表示本发明的1个实施例的复印机1001的要部结构的框图。

图3是本发明的1个实施例的复印机1001的读取单元2070和打印单元2095的硬件结构图。

图4是本发明的1个实施例的复印机1001的操作单元2012的外观图。

图5A是表示本发明的1个实施例的显示在复印机1001的操作单元2012的LCD显示单元2013上的操作画面例的图。

图 5B 是表示在图 5A 所示的操作画面上按下网页浏览器键 504 时启动的浏览器启动画面的一个例子的图。

图 6 是表示本发明的 1 个实施例的安装在复印机 1001 上的信息处理设备的结构的框图。

图 7 是用于说明本发明的 1 个实施例的由信息处理设备进行的检索词登录处理顺序的流程图。

图 8 是用于说明图 7 所示的检索词登录处理中的链接对象信息抽取处理（步骤 S703）的详细情况的流程图。

图 9 是用于说明本发明的 1 个实施例的由信息处理设备进行的自动检索执行处理顺序的流程图。

图 10 是用于说明本发明的 1 个实施例的由信息处理设备进行的待机状态下的用户请求处理顺序的流程图。

图 11 是表示用网页浏览器显示由本实施例的信息处理设备取得、并由公共网页服务器 603 公开的 HTML 文件的例子的图。

图 12 是表示根据用户的请求打印输出由本实施例的信息处理设备取得、并由公共网页服务器 603 公开的 HTML 文件的差异信息的例子的图。

具体实施方式

以下，参照附图详细说明本发明的 1 个实施例的信息处理设备及利用该信息处理设备的信息处理方法。

<网络结构>

图 1 是示意地表示本发明的 1 个实施例中、具有安装了信息处理设备的多功能设备的网络系统的结构的结构图。在本实施例中，作为多功能设备以具有网页浏览器功能的复印机为例进行说明。

如图 1 所示，复印机 1001 在图 1 所示的网络系统中，与具有同复印机 1001 相同的功能的复印机 1002、传真装置 1003、数据库/邮件服务器 1004、以及客户计算机 1005 一起，与由以太网（注册商标）等网络构成的 LAN1006 相连接。此外，复印机 1001 还被包括在公用

线路 (WAN) 1008 中, 可与传真装置 1007 等进行通信。

在本实施例中说明的复印机 1001, 具有复制功能和传真功能, 并且, 还具有读取原稿图像、并将读取到的图像数据发送给收容在 LAN1006 上的各装置的数据发送功能。此外, 复印机 1001 还具有 PDL 功能, 可以按照来自连接在 LAN1006 上的客户计算机 1005 的指示, 接收由该计算机发送来的 PDL 图像并打印。复印机 1001, 可以将由复印机 1001 本身读取到的图像、和由连接在 LAN1006 上的计算机等所发送的 PDL 图像, 保存到该复印机 1001 内部的硬盘中的指定的 BOX 区域, 并可以对保存在 BOX 区域的图像进行打印输出。

复印机 1001, 还可以经由 LAN1006 接收复印机 1002 读取到的图像数据, 将接收到的图像数据保存到复印机 1001 内的硬盘中并进行打印输出。进而, 复印机 1001, 还可以按照客户计算机 1005 的指示, 经由 LAN1006 接收蓄积在数据库服务器 1004 中的图像, 将其保存到该复印机 1001 内并进行打印输出。

传真装置 1003, 可以经由 LAN1006 接收复印机 1001 读取到的数据, 并将接收到的该数据发送给其他装置。此外, 传真装置 1007, 可以经由公用线路 1008, 接收复印机 1001 读取到的数据, 并对接收到的该数据进行打印输出。

另一方面, 数据库/邮件服务器 1004 是具有以下功能的服务器装置, 即, 具有经由 LAN1006 接收复印机 1001 读取到的数据, 将该数据存储到数据库, 并将其作为电子邮件发送的功能。此外, 客户计算机 1005, 可以经由 LAN1006 与数据库/邮件服务器 1004 连接, 从数据库/邮件服务器 1004 取得所需要的数据并显示。此外, 客户计算机 1005, 还可以经由 LAN1006 接收复印机 1001 读取到的数据, 并对接收到的该数据进行加工、编辑。

<复印机的主要结构>

图 2 是表示本发明的 1 个实施例的复印机 1001 的要部结构的框图。如图 2 所示, 复印机 1001 具有作为信息处理装置发挥作用的控制单元 2000。该控制单元 2000 进行以下用于实现复制功能的控制,

即，连接作为图像输入设备的扫描仪 2070 和作为图像输出设备的打印机 2095，由打印机 2095 对由扫描仪 2070 从原稿图像读取到的图像数据进行打印输出。此外，控制单元 2000 还进行用于经由 LAN1006、公用线路（WAN）1008 在与其他装置之间输入输出模式图像（pattern image）、设备信息等的控制。

如图 1 所示，控制单元 2000 具有 CPU2001。CPU2001 利用存储在 ROM2003 中的引导程序（boot program）启动操作系统（OS），通过在该 OS 上执行存储在 HDD（硬盘驱动器）2004 中的应用程序来执行各种处理。作为该 CPU2001 的工作区域，采用 RAM2002。RAM2002 不仅提供 CPU2001 的工作区域，还提供用于暂时存储图像数据的图像存储区域。HDD2004 与上述应用程序一起，存储图像数据。

CPU2001 经由系统总线 2007 与 ROM2003、RAM2002 连接，并且，还与操作单元 I/F（操作单元接口）2006、网络 I/F（网络接口）2010、调制解调器 2050 以及图像总线 I/F（图像总线接口）2005 连接。

操作单元 I/F2006，是与具有触摸板的操作单元 2012 的接口，向操作单元 2012 输出要在操作单元 2012 显示的图像数据。此外，操作单元 I/F2006，向 CPU2001 发送用户在操作单元 2012 中输入的信息。

此外，网络 I/F2010 与 LAN1006 连接，经由该 LAN1006 在与连接于 LAN1006 的各装置之间输入输出信息。调制解调器 2050 与公用线路 1008 连接，经由公用线路 1008 在与其他装置之间输入输出信息。

图像总线 I/F2005，是用于连接与系统总线 2007 高速传送图像数据的图像总线 2008 并转换数据结构的总线桥。图像总线 2008 由 PCI 总线或 IEEE1394 构成。在图像总线 2008 上，设置有光栅图像处理器（RIP）2060、设备 I/F2020、扫描图像处理单元 2080、打印图像处理单元 2090、图像旋转单元 2030、缩略图（thumbnail）生成单元 2035 以及图像压缩单元 2040。

RIP2060 是将 PDL 代码展开成位图图像的处理器。设备 I/F2020 上连接有扫描仪 2070 和打印机 2095，同步或非同步地转换图像数据。

扫描图像处理单元 2080, 对输入图像数据进行修改、加工、编辑处理。打印图像处理单元 2090, 对打印输出图像数据进行打印机的修改、分辨率转换等。图像旋转单元 2030 进行图像数据的旋转。图像压缩单元 2040, 将多值图像数据压缩成 JPEG 数据, 将二值图像数据压缩成 JBIG、MMR、MH 等数据, 并且, 还进行其解压缩处理。

<复印机 1001 的硬件结构>

图 3 是本发明的 1 个实施例的复印机 1001 的读取单元 2070 和打印单元 2095 的硬件结构图。如图 3 所示, 在复印机 1001 中读取单元 2070 和打印单元 2095 被一体地构成。

读取单元 2070 搭载有原稿供纸单元 250。原稿供纸单元 250, 将原稿从最上面开始依次逐页送到台板玻璃 211 上, 每当各原稿的读取动作结束时将该原稿从台板玻璃 211 排出到图中未示出的排出托盘。读取单元 2070, 当原稿被送到台板玻璃 211 上时点亮灯 212, 开始移动移动单元 213。在读取单元 2070 中, 通过移动该移动单元 213, 对台板玻璃 211 上的原稿进行读取扫描。

在读取单元 2070 中, 当正在进行上述读取扫描时, 来自原稿的反射光经过各反射镜 214、215、216 和透镜 217 被导入 CCD 图像传感器 (以下称作 “CCD”) 218, 原稿上的图像在 CCD218 的摄像面上成像。CCD218 将形成在摄像面上的图像转换成电信号, 在对该电信号实施了预定处理后将其输入到图中未示出的控制装置。控制装置相当于图 2 中的控制单元 2000。

打印单元 2095 具有激光驱动器 321。激光驱动器 321 根据从上述控制装置输入的图像数据驱动激光发光单元 322。由此, 从激光发光单元 322 发出对应于图像数据的激光, 该激光一边进行扫描一边被照射到感光鼓 323 上。在感光鼓 323 上由所照射的激光形成静电潜像, 利用由显影器 324 提供的调色剂, 对该静电潜像进行可视化, 成为调色剂像。此外, 与激光的照射时序同步地, 经由传送路径将记录纸从各供纸盒 311、312 供给到感光鼓 323 和转印单元 325 之间, 感光鼓 323 上的调色剂像被转印单元 325 转印到所供给的记录纸上。

转印了调色剂像的记录纸，通过传送带被送到定影辊对（加热辊和加压辊）326。然后，定影辊对 326 对记录纸进行热压，使记录纸上的调色剂像定影在记录纸上。通过该定影辊对 326 后的记录纸，由排纸辊对 327 排出到排纸单元 330。排纸单元 330 由可实施分类、订缀等后处理的片处理装置（sheet processor）构成。此外，当设定了双面记录模式时，在将记录纸传送到排纸辊对 327 后，使排纸辊对 327 向反方向旋转，通过挡板 328 导入再供纸传送路径 329。被导入再供纸传送路径 329 的记录纸，以上述时序被再次供给到感光鼓 323 和转印单元 325 之间，在该记录纸的背面转印调色剂像。

<操作单元 2012 的结构>

图 4 是本发明的 1 个实施例的复印机 1001 的操作单元 2012 的外观图。如图 4 所示，操作单元 2012 中的 LCD 显示单元 2013，在 LCD 上贴有触摸板薄片，显示系统的操作画面，并且，在按下了所显示的键后将其位置信息传送给控制单元 2000 的 CPU2001。

此外，在图 4 中，开始键 2014 在要开始原稿图像的读取动作等时按下。在开始键 2014 的中央部，有绿和红 2 种颜色的 LED2018，通过其颜色表示是否为可使用开始键 2014 的状态。停止键 2015 在要停止运行中的动作时按下。ID 键 2016 在要输入使用者的用户 ID 时使用。此外，复位键 2017 在要将来自操作单元 2012 的设定返回到初始状态时按下。

<操作单元 2012 的操作画面>

图 5A 是表示本发明的 1 个实施例的显示在复印机 1001 的操作单元 2012 的 LCD 显示单元 2013 上的操作画面例的图。如图 5A 所示，在操作画面上部，像复制键 501、发送/FAX 键 502、BOX 键 503、网页浏览器 504 那样，显示有按多个功能标签显示的触摸键。在本实施例中，控制单元 2000 的 ROM2003 中编入了用于显示网页浏览器的程序，可以按照来自操作单元 2012 上的指示显示网页浏览器。

在图 5A 所示的操作画面中，显示着按下了相当于复制键 501 的触摸键时的复制功能的初始画面。在图 5A 中，在显示有“可以复制”

这样的字符的区域 517, 显示复制功能的状况, 在其下方有用于显示倍率、选择的供纸段(selected paper feed stage)、数值(numeric entry)的区域。此外, 在该操作画面上, 作为用于设定复制功能的动作模式的触摸键, 显示着等倍、倍率、用纸选择、分类、双面、中断、字符、用于调整浓度的对应于“变淡”的左箭头键、对应于“变浓”的右箭头键、自动调整浓度的自动键, 无法显示在初始画面上的动作模式的指定, 通过按下应用模式键将设定画面分层显示在画面内。

此外, 在操作画面下部的显示区域 505 显示着复印机 1001 的状态。在该显示区域 505 上, 显示例如卡纸等报警消息、或正在进行 PDL 打印时显示表示正在进行 PDL 打印的状态消息。

进而, 操作画面下部的系统状况/中止键 506, 当按下该触摸键后, 显示表示复印机 1001 的设备信息的画面、和表示打印作业状况的画面, 可以通过这样显示的画面进行作业的中止。

在图 5A 所示的操作画面中, 当用户按下发送/FAX 键 502 后, 复印机 1001, 将由复印机 1001 所读取的图像通过电子邮件或 FTP 等发送给 LAN1006 上的设备。或者, 复印机 1001 显示用于使用公用线路 1008 进行传真发送的设定画面。

此外, 在图 5A 所示的操作画面中, 当用户按下 BOX 键 503 后, 复印机 1001, 将由复印机 1001 所读取的图像保存(登录)到 HDD2004 内的 BOX 区域, 或指定保存在 BOX 区域的图像数据并打印, 或显示用于向 LAN1006 上的设备进行发送的设定画面。

并且, 在图 5A 所示的操作画面中, 当用户按下网页浏览器 504 时, 复印机 1001, 在操作单元 2012 的操作画面上打开图 5B 所示的浏览器启动画面。图 5B 是表示在图 5A 所示的操作画面上按下网页浏览器键 504 时启动的浏览器启动画面的一个例子的图。

在图 5B 所示的画面中, URL 输入单元 507, 在用户要指定 URL 再打开页面时按下该字段。也可以在管理员的权限下进行设定, 以在按下该字段后显示软件盘使用户可以指定 URL。此外, 收藏键 508 可以选择/执行是登录 URL、整理 URL、还是显示 URL。进而, 显示

区域 509 是内容的显示区域。

另外，在图 5B 中，后退键 510 是在要后退到已显示的上一页面时按下的键。而前进键 511 是在要前进到显示的下一页面时按下的键。此外，更新键 512 是要再次加载当前正在显示的页面并显示的键。还有，中止键 513 是在要中止页面的读入时按下的键。还有，主页键 514 是在要移动到设定好的主页时按下的键。

另外，在图 5B 中，打印键 515 是在要打印当前正在显示的页面时按下的键。当按下打印键 515 后显示打印设定对话框。在该打印设定对话框中，显示如何打印边框的设定、关于打印的设定（例如份数、双面、分类等）的画面，可以通过按下该画面内的打印开始按钮开始打印。还有，菜单键 516 是在要显示变更画面的显示倍率、变更字符尺寸、变更字符代码等的画面时按下的键。

<信息处理设备的结构>

图 6 是表示本发明的 1 个实施例的安装在复印机 1001 上的信息处理设备的结构的框图。在图 6 中，信息处理设备 601 安装在图 1 所示的复印机 1001 上。用户经由操作单元 602 访问 WWW，进行网站的检索。并且，关于用户执行的检索，当要日后再次以相同的条件进行检索并取得与上一次检索的检索结果的差异信息时，用检索登录单元 615 将检索中所使用的检索词存储到原检索结果信息数据库 607。此外，检索登录单元 615 取得该检索词的检索结果，再传送给链接对象信息抽取单元 609。链接对象信息抽取单元 609，从作为检索引擎的结果的 HTML 文件中抽取链接对象 URL，并存储到原检索结果信息数据库 607。

此外，图 6 所示的自动检索执行单元 608，定期爬行或被用户请求所触发。该自动检索执行单元 608，自动对在检索登录单元 615 所登录的检索词进行检索，取得其检索结果，传送给链接对象信息抽取单元 609。链接对象信息抽取单元 609，从作为检索引擎的结果的 HTML 文件中，抽取包含在其中的链接对象的 URL，存储到新检索结果信息数据库 605。

此外，在图 6 中，差异信息分析单元 610，分别从新检索结果信息数据库 605 和原检索结果信息数据库 607 取得 URL 的一览，通过逐一比较所取得的 URL，判断在新检索结果信息数据库 605 中存在而在原检索结果信息数据库 607 中不存在的“新站点”，和在新检索结果信息数据库 605 中不存在而在原检索结果信息数据库 607 中存在的“消失站点”，将两者的 URL 传送给站点内容取得单元 611。

站点内容取得单元 611，访问所传送的 URL 的网页，下载页面信息，与 URL 一起存储到差异信息数据库 606。此外，关于消失站点，当 WWW 上的站点本身已完全消失时，将该情况存储到差异信息数据库 606。

进而，在图 6 中，HTML 生成单元 612 从差异信息数据库 606 取得新站点和消失站点的信息，汇总转换成 HTML 文件，并将所生成的 HTML 文件发送给公共网页服务器 603。用户在任何时候都可以经由网页浏览器 604 等阅览所存储的差异信息。

进而，在图 6 中，输出请求接受单元 613，当有用户发出的差异信息的输出请求时，从差异信息数据库 606 取得新站点和消失站点的信息，并传送给差异信息输出单元 614。差异信息输出单元 614 将所接收的差异信息调整格式后输出。

上述各单元，即自动检索执行单元 608、链接对象信息抽取单元 609、差异信息分析单元 610、站点内容取得单元 611、HTML 生成单元 612、输出请求接受单元 613、差异信息输出单元 614、以及检索登录单元 615，可以通过由存储器和 CPU 等构成的处理单元将用于实现这些功能的程序加载到其存储器并执行，来实现其功能。

此外，上述存储器由 RAM 构成，新检索结果信息数据库 605、差异信息数据库 606、以及原检索结果信息数据库 607，由 EEPROM（Electrically Erasable Programmable Read-Only Memory）、硬盘、光磁盘等非易失性的记录装置构成。

<信息处理设备的动作>

接着，参照图 7~图 10 所示的流程图，说明具有上述结构的本实

施例的信息处理设备 601 的动作例。

图 7 是用于说明本发明的 1 个实施例的信息处理设备的检索词登录处理顺序的流程图。

首先,信息处理设备 601,判断是否由用户进行了网站的检索(步骤 S701),当进行了检索时(是)询问并判断是否登录用户进行该检索时使用的检索词、即是否要在日后进行检索时取得差异信息(步骤 S702)。结果,当判断为登录检索词时(是),进行链接对象信息抽取处理(步骤 S703)。

图 8 是用于说明图 7 所示的检索词登录处理中的链接对象信息抽取处理(步骤 S703)的详细情况的流程图。链接对象信息抽取单元 609,首先接受作为检索引擎的输出结果的 HTML 文件的输入(步骤 S801)。接着,链接对象信息抽取单元 609,分析接受了输入的 HTML 文件(步骤 S802),抽取包含在其中的链接对象的 URL(步骤 S803)。

如上所述,当在步骤 S703 抽取了链接对象的信息(链接对象的 URL)后,检索登录单元 615,为了之后进行比较,将通过链接对象信息抽取处理而得到的 URL 的一览,保存到原检索结果信息数据库 607(步骤 S704)。

在信息处理设备 601 中,对所登录的检索词自动执行网站的检索处理,比较用同一检索词在不同的时期进行了检索时的 2 个登录检索词时的检索结果。图 9 是用于说明本发明的 1 个实施例的信息处理设备的自动检索执行处理顺序的流程图。以下的处理,对已登录的所有检索词都要进行。

在此,自动检索执行单元 608,指定登录在原检索结果信息数据库 607 中的检索词中的 1 个(步骤 S901)。接着,自动检索执行单元 608,用在步骤 S901 所指定的检索词检索网站(步骤 S902)。然后,链接对象信息抽取单元 609,对所得到的检索结果进行在前用图 8 说明的链接对象信息抽取处理(步骤 S903),将抽取出的链接对象信息存储到新检索结果信息数据库 605。

接着,差异信息分析单元 610,取得存储在新检索结果信息数据

库 605 和原检索结果信息数据库 607 中的 URL, 例如, 在各自取得了 m 个和 n 个时, 比较 $m \times n$ 次 URL (步骤 S904)。差异信息分析单元 610, 通过进行上述比较, 检测在原检索结果信息数据库 607 中不存在而在新检索结果信息数据库 605 中存在的、新添加的 URL (新站点), 和在原检索结果信息数据库 607 中存在而在新检索结果信息数据库 605 中不存在的 URL (消失站点) (步骤 S905), 并传送给站点内容取得单元 611。

站点内容取得单元 611, 访问从差异信息分析单元 610 传送来的 URL 的网页下载页面信息, 将包括 URL 和该页面信息的差异信息存储到差异信息数据库 606 (步骤 S906)。在进行存储时, 对于被判断为消失站点的站点, 当 WWW 上的站点本身已完全消失而不存在时根本无法访问, 因而, 要将该情况存储到差异信息数据库 606。

接着, HTML 生成单元 612, 从差异信息数据库 606 取得差异信息并将其转换成 HTML 文件 (步骤 S907)。然后, 将所生成的 HTML 文件在公共网页服务器 603 上公开 (步骤 S908)。

然后, 判断是否对已登录的所有检索词都进行了以上说明的自动检索执行处理 (步骤 S909)。结果, 当尚未对所有检索词进行时 (否) 返回到步骤 S901, 对下一检索词进行上述从步骤 S901 到步骤 S908 的处理。另一方面, 当对所有检索词都进行了时 (是) 结束本处理。

信息处理设备 601, 基本上是以一定间隔定期地执行上述自动检索执行处理。此外, 该定期执行的间隔也可以由用户指定。

图 10 是用于说明本发明的 1 个实施例的信息处理设备的待机状态下的用户请求处理顺序的流程图。在待机状态下, 判断是否有由用户发出的差异信息的输出请求 (步骤 S1001)。结果, 当有输出请求时 (是), 输出请求接受单元 613, 接受输出基于哪个检索词的差异信息的指定 (步骤 S1002)。然后, 输出请求接受单元 613, 取得存储在差异信息数据库 606 中的新关于站点和消失站点的差异信息 (步骤 S1003)。然后, 差异信息输出单元 614, 调整格式后输出在步骤 S1003 取得的差异信息 (步骤 S1004)。

此外，当在步骤 S1001 中判断为没有输出请求时（否），判断是否有差异检索请求（步骤 S1005）。结果，当有差异检索请求时（是），信息处理设备 601，进行用图 9 说明的自动检索执行处理（步骤 S1006）。

以上是信息处理设备 601 的动作整体的说明。上述说明的动作流程只是一个例子，因而并不限于上述处理流程。此外，在上述实施例中，将上述信息处理设备 601 作为复印机 1001 所具有的设备进行了说明，但用与互联网连接的通用的个人计算机等信息处理设备，也同样能实现上述处理。

<差异信息显示例>

图 11 是表示用网页浏览器显示由本实施例的信息处理设备取得、并由公共网页服务器 603 公开的 HTML 文件的例子的图。在图 11 中，1101 是显示在信息处理设备的操作画面上的网页浏览器，1102 是检索词。此外，1103 是新站点的列表，标注顺序号显示所有新站点。还有，1104 是消失站点的列表，标注顺序号显示所有消失站点。另外，1105 是站点名，1106 是站点 URL。

图 12 是表示根据用户的请求打印输出由本实施例的信息处理设备取得、并由公共网页服务器 603 公开的 HTML 文件的差异信息的例子的图。在图 12 中，1201 是信息表（sheet）主体，1202 是检索词。此外，1203 是新站点的列表，标注顺序号显示所有新站点。还有，1204 是消失站点的列表，同样标注顺序号显示所有消失站点。另外，1205 是站点名，1206 是站点 URL。

如上所述，在本实施例中，作为差异信息说明了一览显示站点名和 URL 那样的例子，但也可以用缩略图进行显示，所述缩略图是从存储在差异信息数据库 606 的页面信息中抽取各站点的图像，再对其进行转换、缩小而形成的。此外，还可以抽取检索到的网站上的包括检索词的字符串（即，该检索词及其周边的例如 1~2 行字符串）。此外，如上所述，还可以在用同一检索词进行检索却检索不到，关于该消失站点的页面信息在 WWW 上已不存在时，明示站点已经消失。并

且，如果在检索时预先保存了站点的缩略图或页面本身，则即便在再次检索时已消失，也可以通过预先保存的缩略图或页面再次显示或打印。

如上所述，通过分析不同检索时期的检索结果的差异，例如可以了解新出现的站点，可以容易地取得针对某类别（category）的最新信息。此外，通过了解已消失的站点，能够可以访问检索引擎已不适用的站点（即，内容已被变更的站点），能够了解某类别的信息的时间序列流。

此外，作为检索结果的使用方法，除了如上所述地以列表显示之外，例如还可以通过进行仅打印N件（N为任意）新站点那样的指示，从与检索条件一致性高的站点开始打印N件。此外，还可以预先进行打印N件消失站点或打印全部消失站点那样的指示，从而在出现了消失站点时使其自动打印。

此外，关于检索方法，除了用检索词进行检索，还可以采用由读取单元2070读取某图像并检索包括与该图像类似的图像的站点等各种检索方法。

此外，不限于自动进行检索的方法，在用相同的检索词按照用户的手动指示进行了检索时，同样也可以抽取并输出与过去的检索结果的差异。

此外，在以上的实施例中，以在互联网上检索网站为例进行了说明，但例如在独立构筑的网络等各种网络中，在取得各种信息（例如检索存在于网络上的各种装置等）等情况下都可以应用本发明。

根据本发明，能够容易且准确地将在不同的检索时期进行了基于同一检索条件的检索后生成的检索结果的差异通知给用户。

在不脱离本发明的精神和范围的情况下，本发明可以有各种不同的实施例，可以理解为，本发明不限于上述具体实施例，而由权利要求来限定。

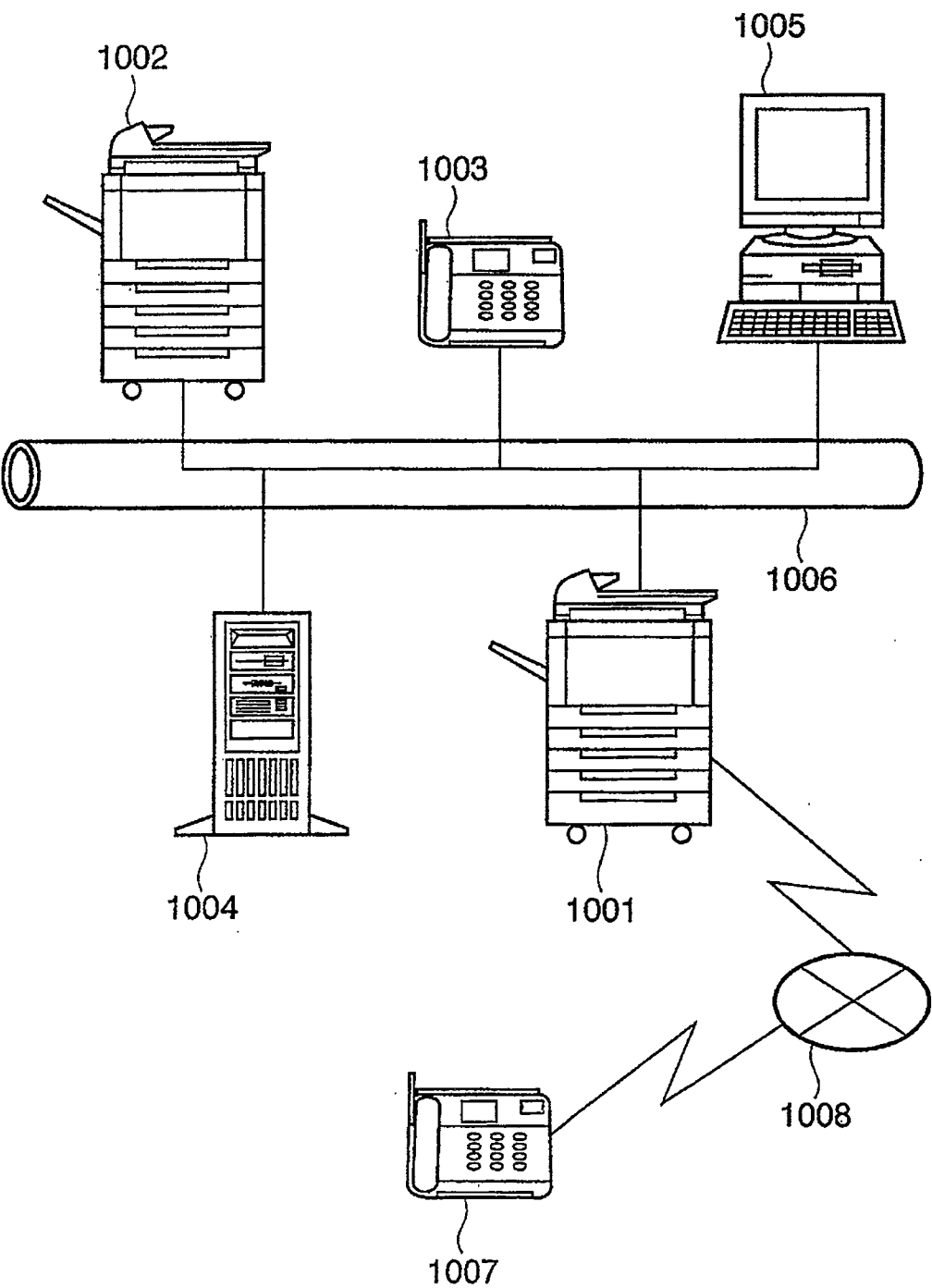


图 1

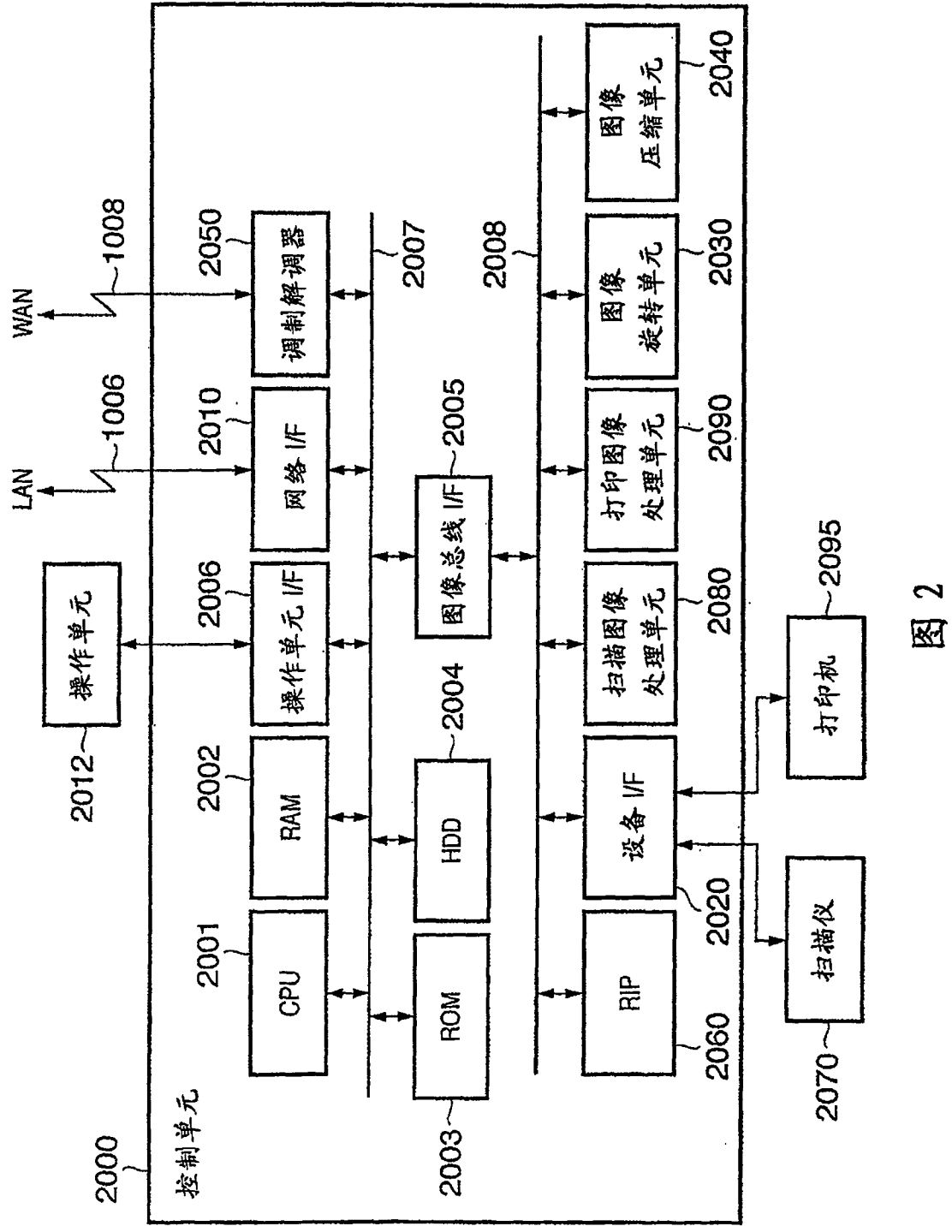


图 2

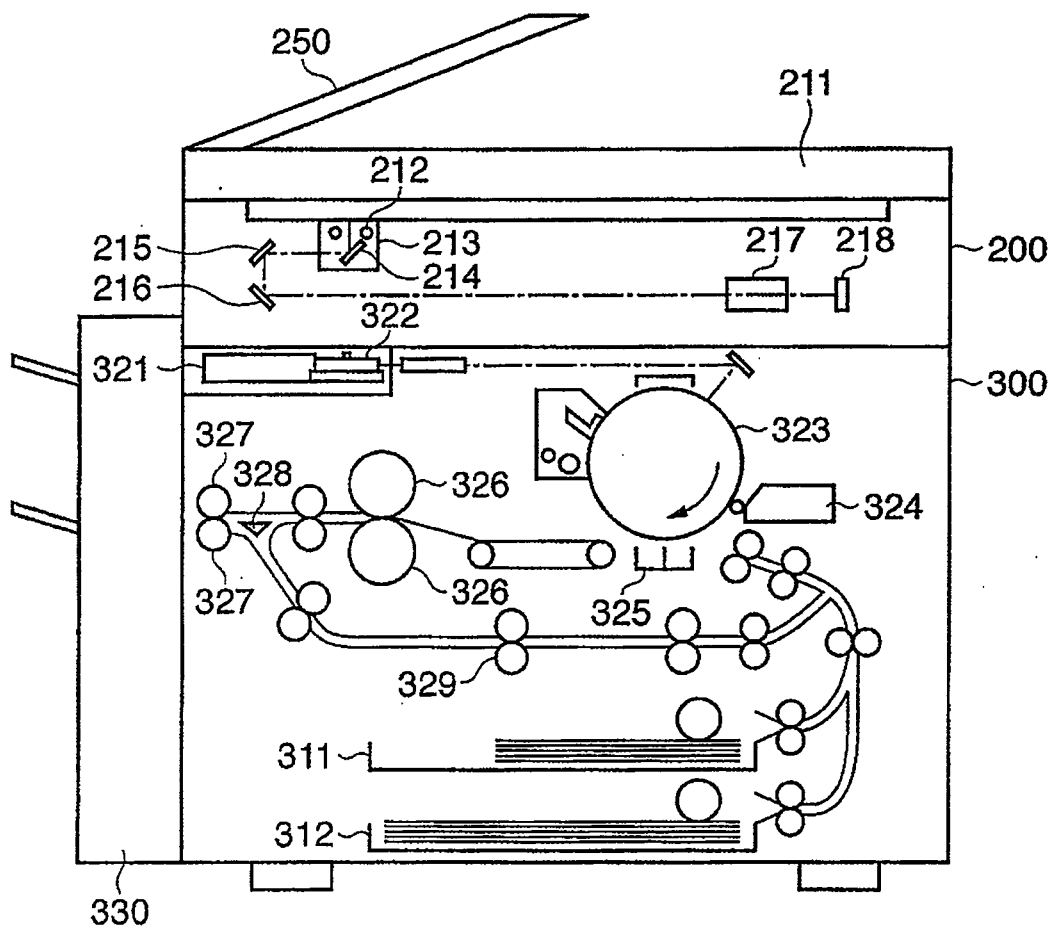


图 3

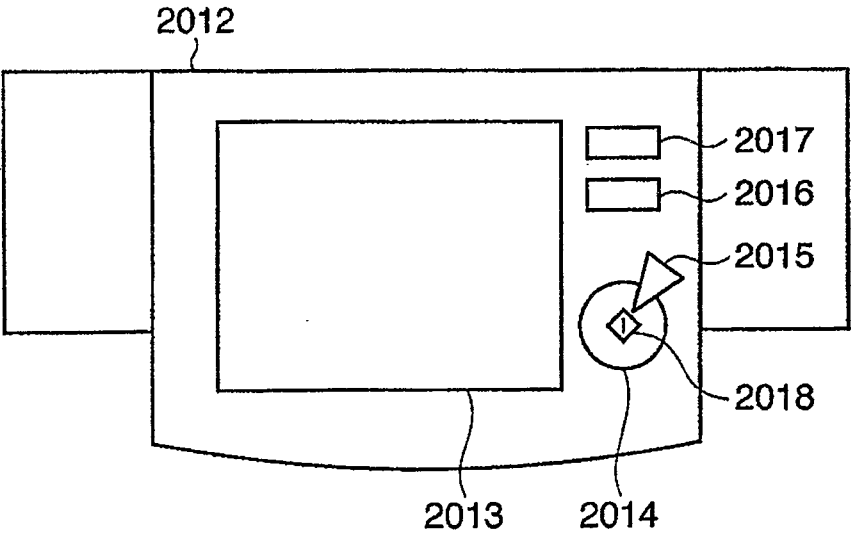


图 4

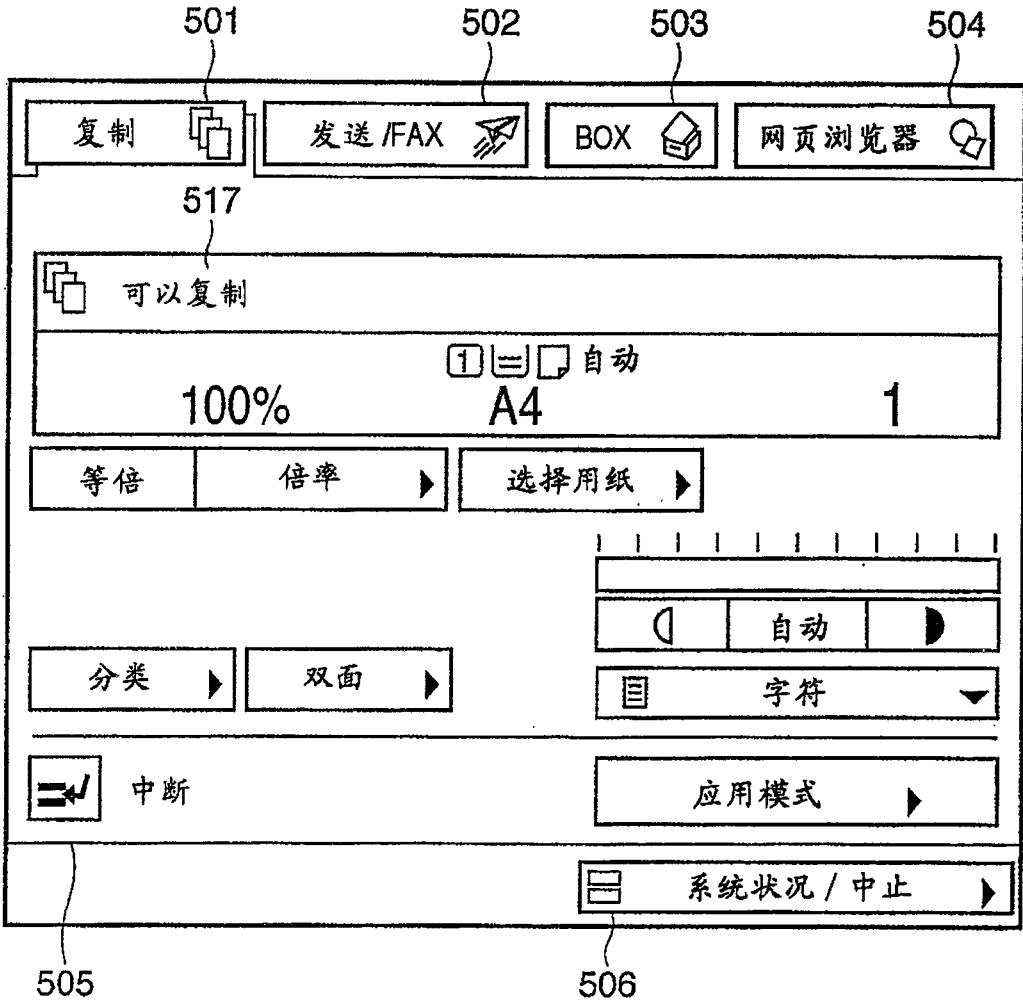


图 5A

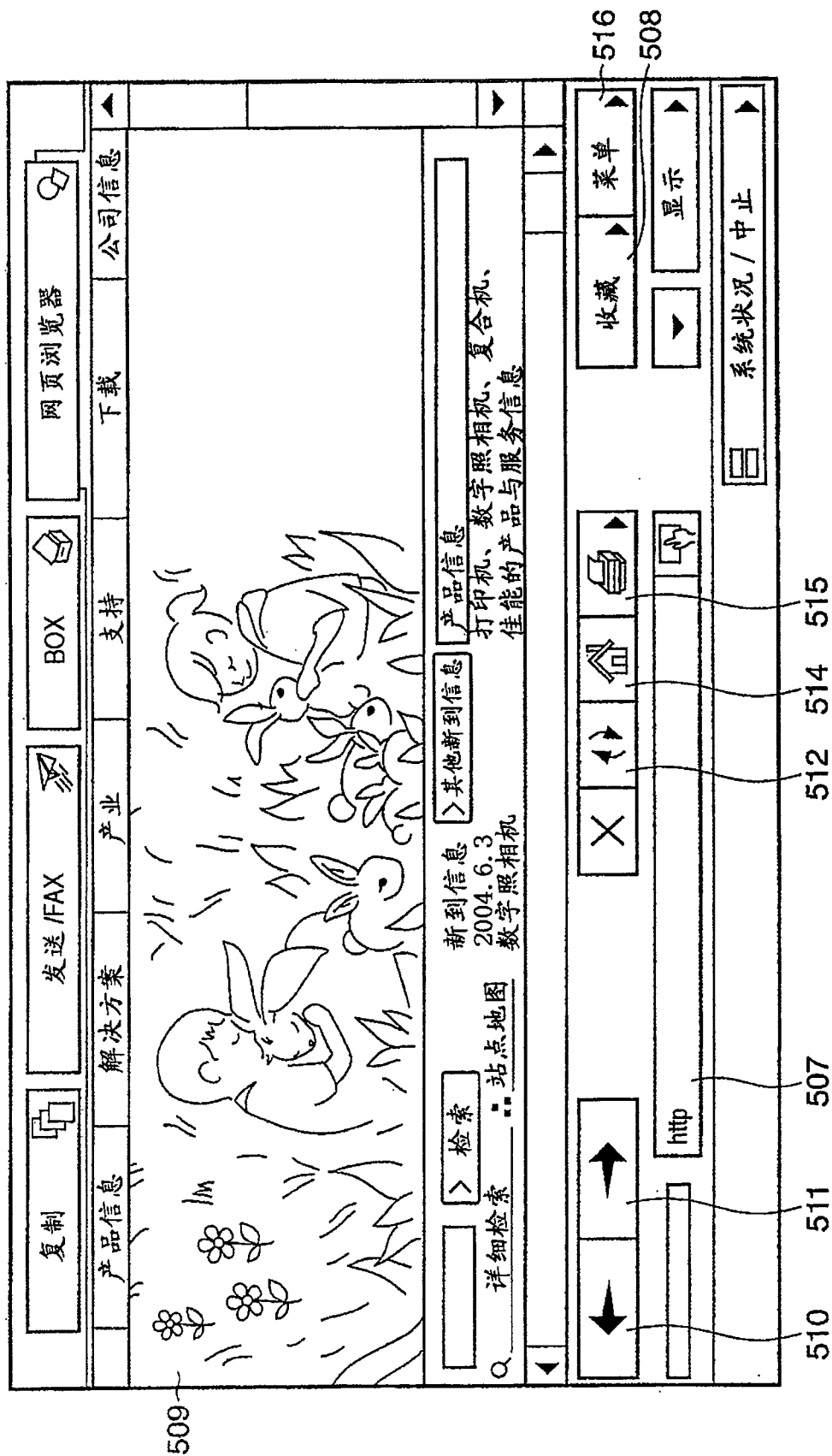


图 5B

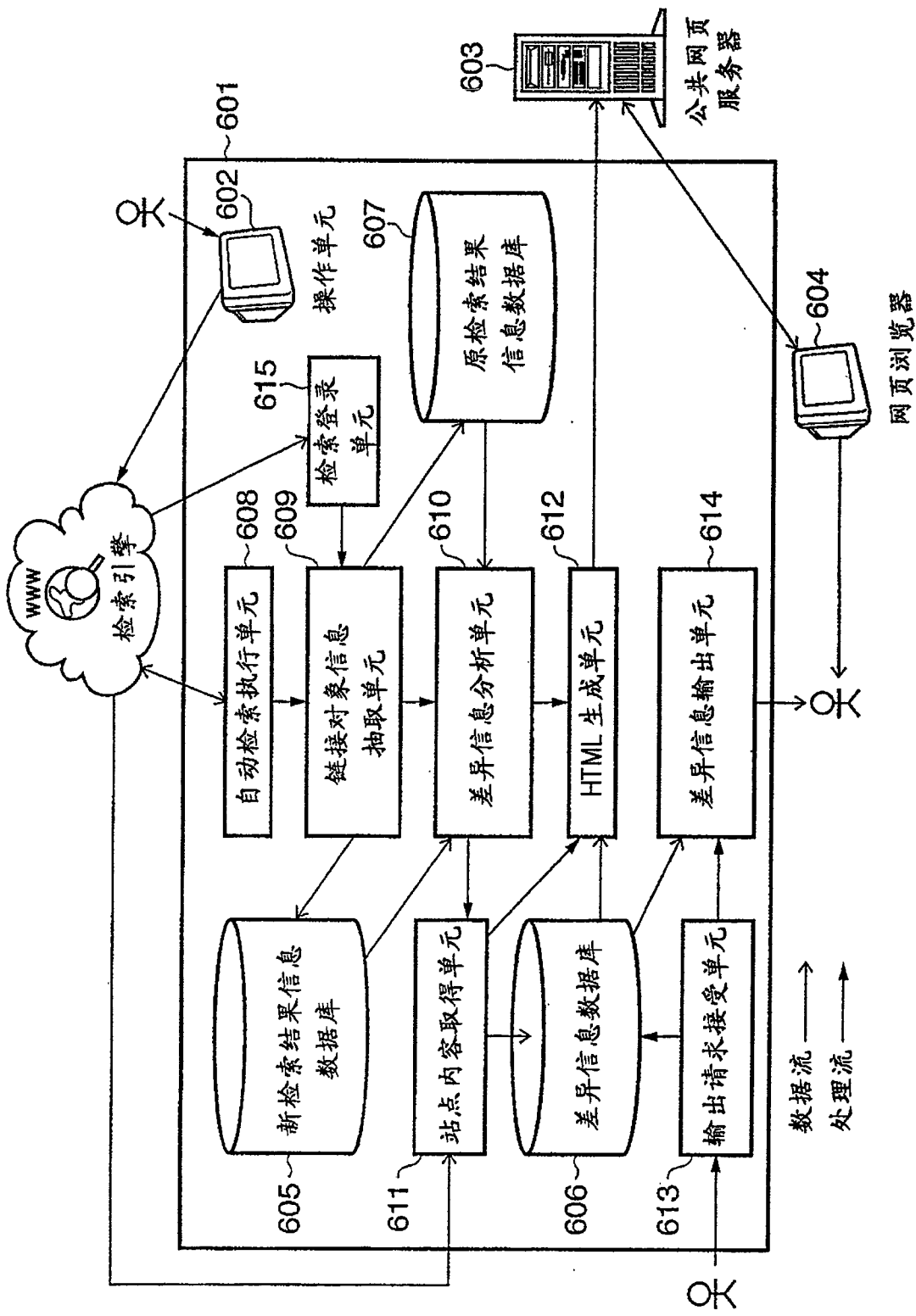


图 6

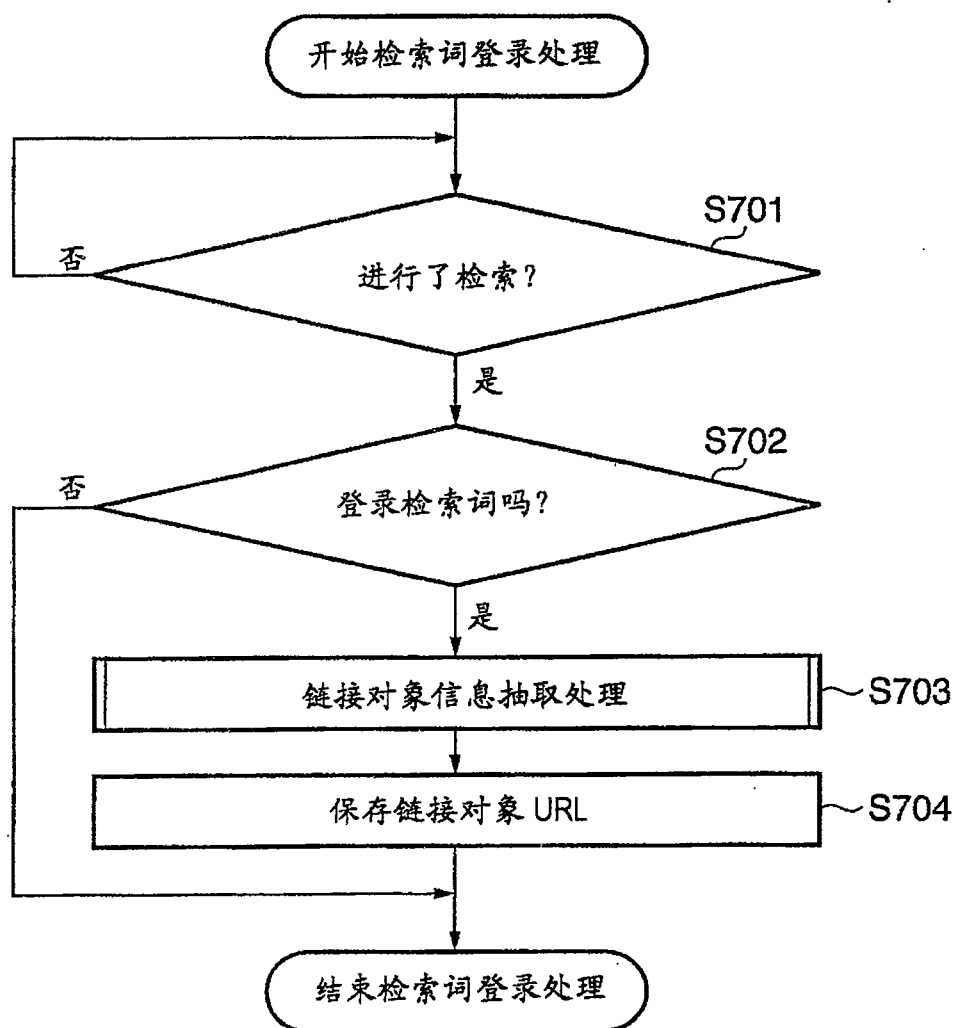


图 7

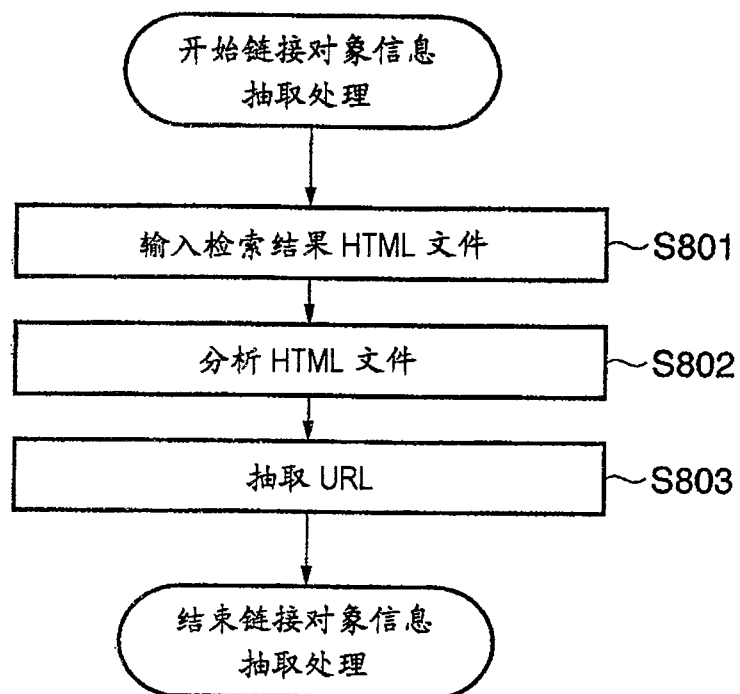


图 8

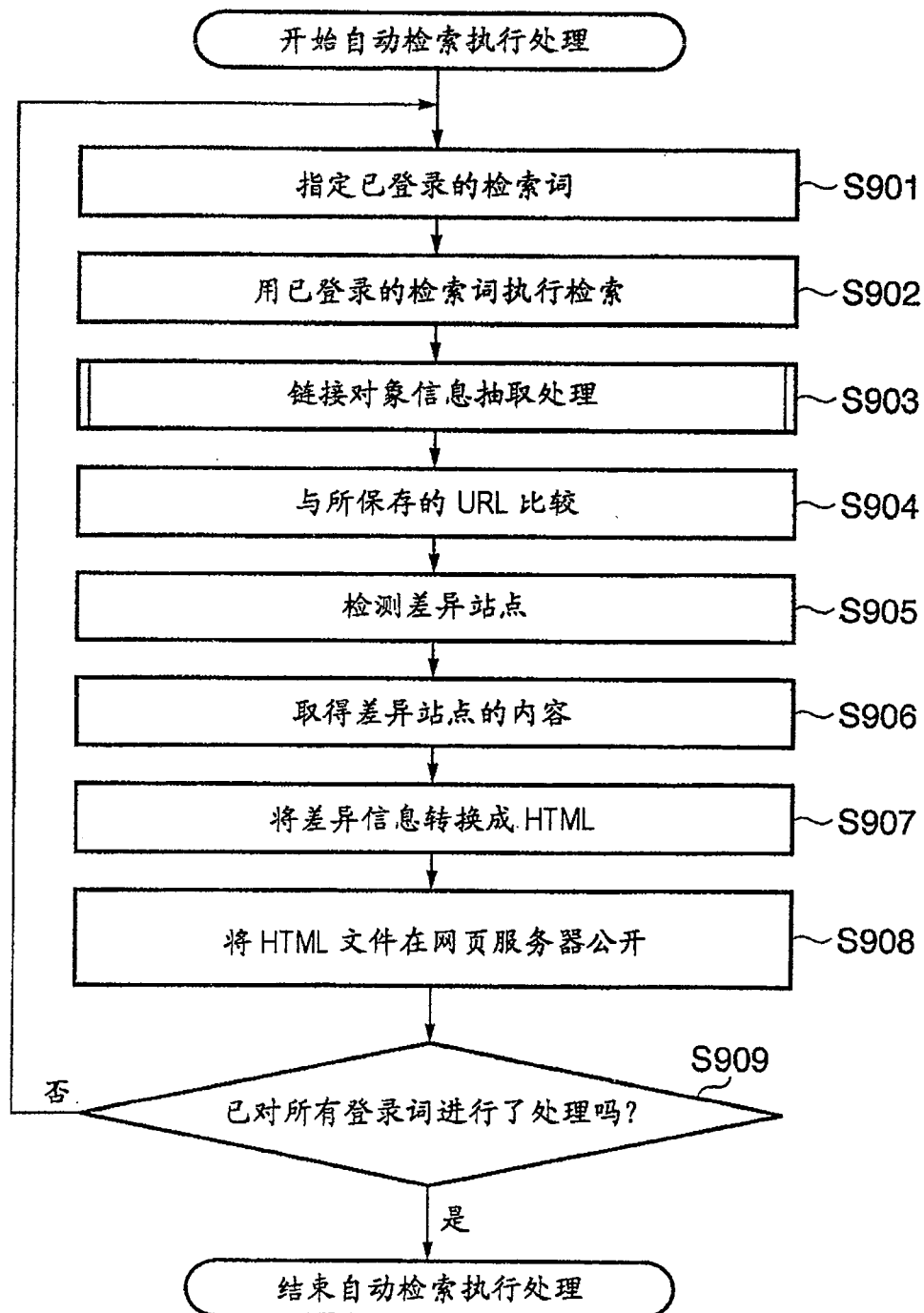


图 9

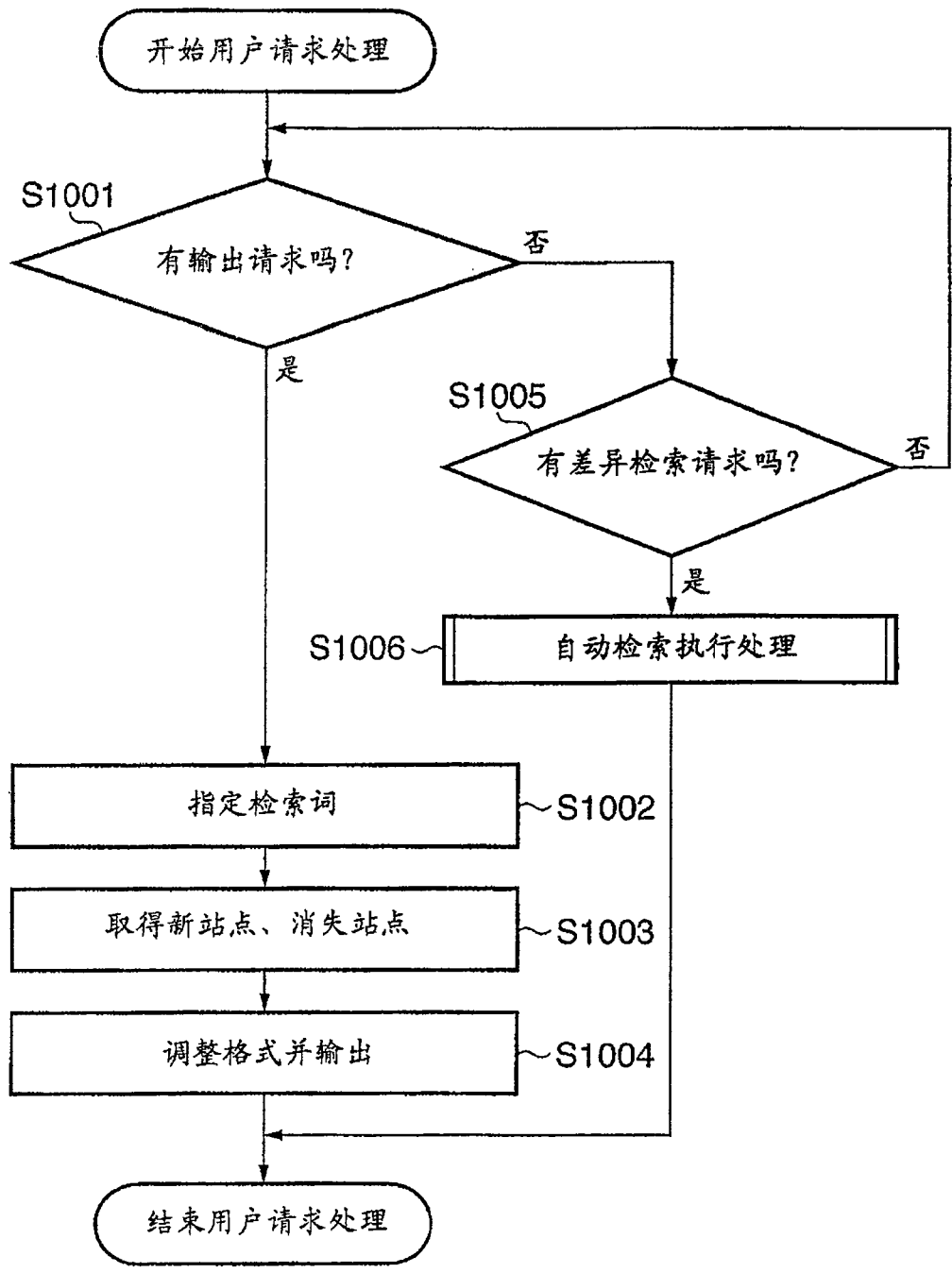


图 10

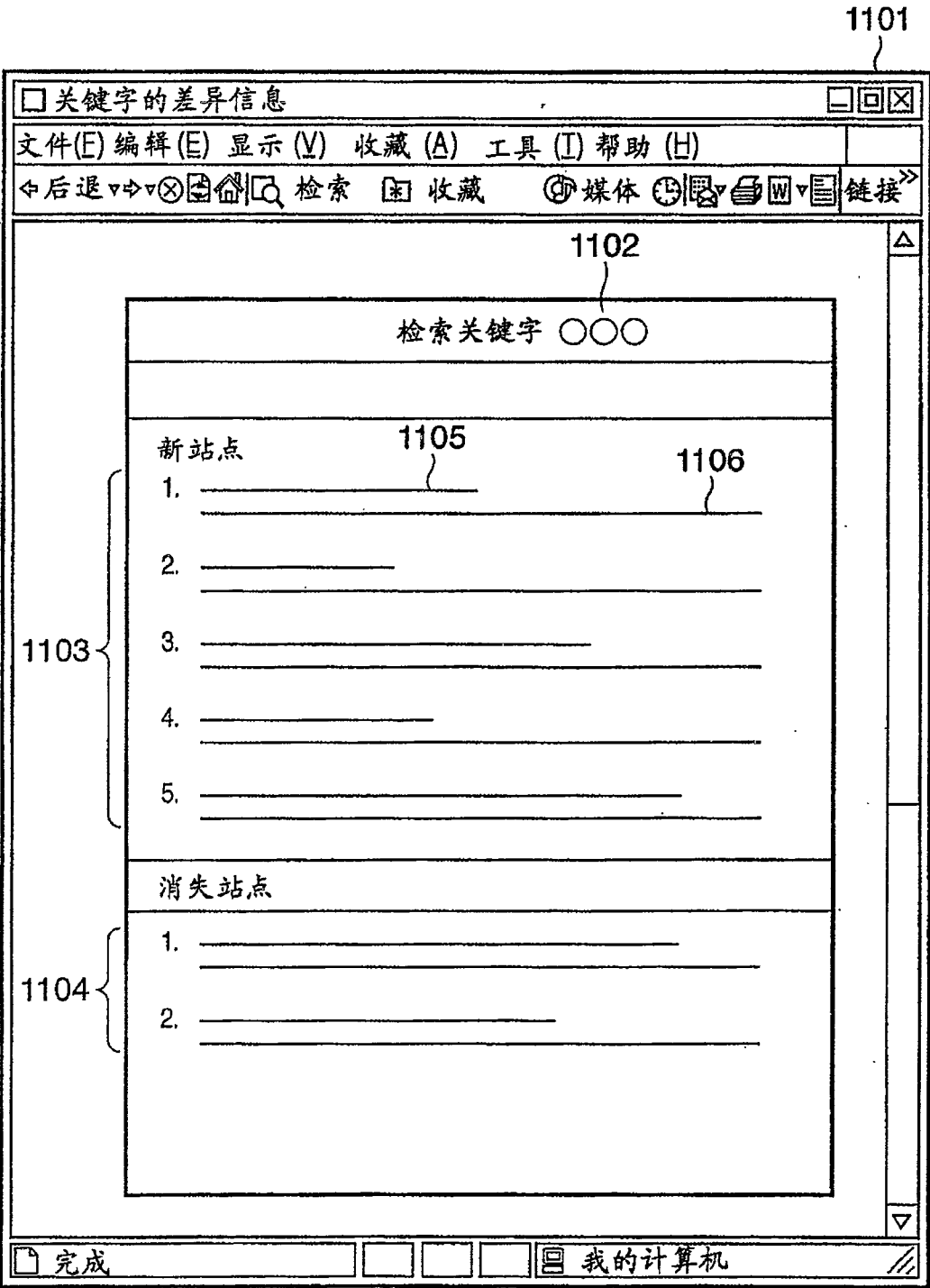


图 11

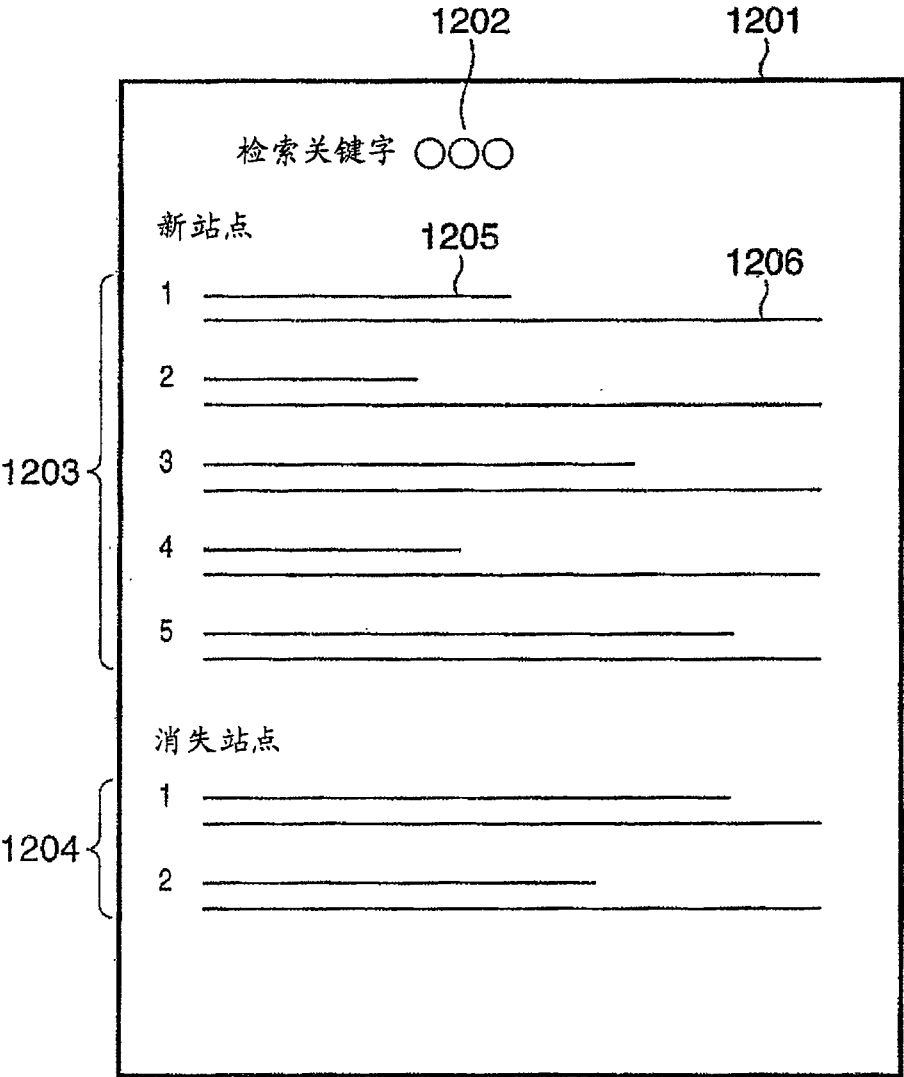


图 12