

[19] 中华人民共和国国家知识产权局

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



[12] 发明专利说明书

专利号 ZL 200710090019.X

[45] 授权公告日 2009 年 9 月 30 日

[11] 授权公告号 CN 100545846C

[22] 申请日 2007.3.23

[21] 申请号 200710090019.X

[30] 优先权

[32] 2006.3.24 [33] JP [31] 2006-083869

[73] 专利权人 佳能株式会社

地址 日本东京都大田区下丸子3-30-2

[72] 发明人 户高伸治

[56] 参考文献

WO2005/038611A2 2005.4.28

US2004/0068486A1 2004.4.8

CN1705895A 2005.12.7

JP2001-101193A 2001.4.13

审查员 唐 嫣

[74] 专利代理机构 北京林达刘知识产权代理事务所

代理人 刘新宇 权鲜枝

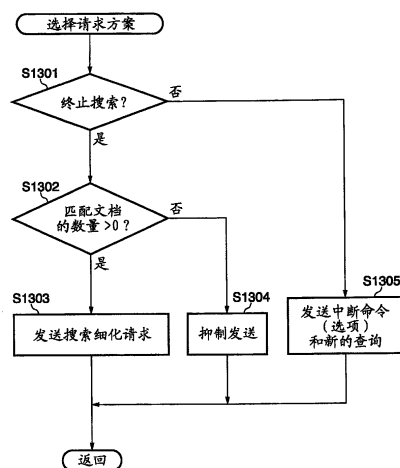
权利要求书4页 说明书24页 附图13页

[54] 发明名称

文档搜索设备和方法

[57] 摘要

本发明提供一种文档搜索设备和方法。文档搜索设备使多个文档管理系统并行执行文档搜索。该文档搜索设备包括：发送单元，其将关于输入的第一搜索条件的第一查询发送给多个文档管理系统中的每个文档管理系统；以及接收单元，其从每个文档管理系统接收对第一查询的搜索结果。特别地，该文档搜索设备包括搜索控制单元，该搜索控制单元在从所有文档管理系统接收到搜索结果之前且在从多个文档管理系统中的至少一个文档管理系统接收到搜索结果之后，使发送单元将用于请求对第一查询进行搜索细化的第二查询发送给每个文档管理系统。



1. 一种使多个文档管理系统并行执行文档搜索的文档搜索设备，该设备包括：

发送单元，其将关于输入的第一搜索条件的第一查询发送给所述多个文档管理系统中的每个文档管理系统；

接收单元，其从所述多个文档管理系统中的每个文档管理系统接收对所述第一查询的搜索结果；

管理单元，其对来自所述多个文档管理系统中的每个文档管理系统的对所述第一查询的搜索结果的应答状态进行管理；

选择单元，其根据由所述管理单元所管理的所述应答状态，针对所述多个文档管理系统中的每个文档管理系统分别选择用于请求对所述第一查询进行搜索细化的方案；以及

搜索控制单元，其在从所有所述文档管理系统接收到对所述第一查询的搜索结果之前且在从所述多个文档管理系统中的至少一个文档管理系统接收到对所述第一查询的搜索结果之后，使所述发送单元基于所述选择单元所选择的方案将用于请求所述搜索细化的第二查询发送给所述文档管理系统中的至少一个文档管理系统。

2. 根据权利要求1所述的文档搜索设备，其特征在于，所述选择单元对应答状态表示搜索终止且命中数量为0的文档管理系统，选择不请求所述搜索细化的方案。

3. 根据权利要求1所述的文档搜索设备，其特征在于，所述搜索控制单元对应答状态表示搜索终止且命中数量为一个或多个的文档管理系统，使所述发送单元发送附有助于识别来自所述第一查询的搜索结果的识别信息和用于搜索细化的第二搜索条件的所述第二查询。

4. 根据权利要求1所述的文档搜索设备，其特征在于，所述搜索控制单元对应答状态表示搜索未完成的文档管理系统，

使所述发送单元发送附有将用于搜索细化的搜索条件添加到所述第一搜索条件的第三搜索条件的所述第二查询。

5. 根据权利要求4所述的文档搜索设备，其特征在于，所述发送单元当将所述第二查询发送给应答状态表示搜索未完成的文档管理系统时，发送用于中断由所述第一查询引起的搜索处理的中断命令。

6. 根据权利要求1所述的文档搜索设备，其特征在于，还包括创建单元，该创建单元当从所述多个文档管理系统中的至少一个文档管理系统接收到搜索结果时，基于该搜索结果创建用于搜索细化的搜索条件。

7. 根据权利要求6所述的文档搜索设备，其特征在于，还包括接受单元，该接受单元从用户或用户设备接受用于创建用于搜索细化的搜索条件的信息，

其中，所述创建单元基于由所述接受单元所接受的信息，创建所述用于搜索细化的搜索条件。

8. 一种使多个文档管理系统并行执行文档搜索的文档搜索设备，该设备包括：

发送单元，其将关于输入的第一搜索条件的第一查询发送给所述多个文档管理系统中的每个文档管理系统；

接收单元，其从所述多个文档管理系统中的至少一个文档管理系统接收对所述第一查询的搜索结果；

搜索控制单元，其使所述发送单元发送用于请求对由所述接收单元接收到的所述搜索结果进行搜索细化的第二查询；

管理单元，其对由所述接收单元从所述多个文档管理系统接收到的搜索结果的应答状态进行管理；以及

选择单元，其根据由所述管理单元所管理的所述应答状态，针对所述文档管理系统中的每个文档管理系统分别选择是否将

所述第二查询发送给所述文档管理系统中的每个文档管理系统。

9. 根据权利要求8所述的文档搜索设备，其特征在于，所述选择单元对应答状态表示搜索终止且命中数量为0的文档管理系统，选择不发送用于搜索细化的所述第二查询。

10. 一种使多个文档管理系统并行执行文档搜索的文档搜索方法，该方法包括以下步骤：

将关于输入的第一搜索条件的第一查询发送给所述多个文档管理系统中的每个文档管理系统；

从所述多个文档管理系统中的每个文档管理系统接收对所述第一查询的搜索结果；

对来自所述多个文档管理系统中的每个文档管理系统的对所述第一查询的搜索结果的应答状态进行管理；

根据所管理的所述应答状态，针对所述文档管理系统中的每个文档管理系统分别选择用于请求对所述第一查询进行搜索细化的方案；以及

在从所有所述文档管理系统接收到对所述第一查询的搜索结果之前且在从所述多个文档管理系统中的至少一个文档管理系统接收到对所述第一查询的搜索结果之后，基于所选择的方案，将用于请求所述搜索细化的第二查询发送给所述文档管理系统中的至少一个文档管理系统。

11. 一种使多个文档管理系统并行执行文档搜索的文档搜索方法，该方法包括以下步骤：

将关于输入的第一搜索条件的第一查询发送给所述多个文档管理系统中的每个文档管理系统；

从所述多个文档管理系统中的至少一个文档管理系统接收对所述第一查询的搜索结果；

发送用于请求对基于所述第一查询的所述搜索结果进行搜索细化的第二查询；

对从所述多个文档管理系统接收到的搜索结果的应答状态进行管理；以及

根据所述应答状态，针对所述文档管理系统中的每个文档管理系统分别选择是否将所述第二查询发送给所述文档管理系统中的每个文档管理系统。

文档搜索设备和方法

技术领域

本发明涉及用于在多个服务器中并行执行搜索的多服务器搜索技术。

背景技术

随着业务的更加复杂和信息的增加，在现代办公环境下必须管理的电子文档（以下简称为“文档”）的量逐渐增加。而且由于信息分类越来越细，使文档的存储位置也越来越分散。在这样的环境下，从在各种不同位置所管理的大量文档中找到想要的文档不是一个简单的任务。

已经提出了使得容易找到所期望文档的文档管理系统。文档管理系统具有用于使用与文档相关的属性信息来搜索文档的文档属性搜索功能、用于搜索文档中所包含的字符串的全文搜索功能、用于搜索文档中所包含的图像的图像搜索功能等。

然而，有时存在许多用作文档管理系统的搜索服务器。在这种情况下，搜索客户机可以通过向多个搜索服务器成批发送查询，来执行所谓的多服务器搜索。

根据日本特开2001-101193号公报，提出了一种用于使所有服务器执行多服务器搜索，以中断搜索并获取中断时刻的搜索结果的方法。

然而，即使在多个服务器上并行执行搜索，也很少能够利用初始搜索条件准确获取所期望的文档。换句话说，可能搜到大量符合搜索条件的文档，从而使得必须对所搜到的文档使用新的搜索条件进行搜索细化（search refinement）。

尽管可以利用在日本特开2001-101193号公报中所记载的

发明使所有服务器在多服务器搜索中间中断搜索，但是却没有努力改进搜索细化。

发明内容

本发明的特征在于当在多个文档管理系统上执行多服务器搜索时提高了搜索细化的效率。注意，通过整个说明书将会理解其它问题。

例如，通过使多个文档管理系统并行执行文档搜索的文档搜索设备来实现本发明。该文档搜索设备包括：发送单元，其将关于输入的第一搜索条件的第一查询发送给所述多个文档管理系统中的每个文档管理系统；接收单元，其从所述多个文档管理系统中的每个文档管理系统接收对所述第一查询的搜索结果；管理单元，其对来自所述多个文档管理系统中的每个文档管理系统的对所述第一查询的搜索结果的应答状态进行管理；选择单元，其根据由所述管理单元所管理的所述应答状态，针对所述多个文档管理系统中的每个文档管理系统分别选择用于请求对所述第一查询进行搜索细化的方案；以及搜索控制单元，其在从所有所述文档管理系统接收到对所述第一查询的搜索结果之前且在从所述多个文档管理系统中的至少一个文档管理系统接收到对所述第一查询的搜索结果之后，使所述发送单元基于所述选择单元所选择的方案将用于请求所述搜索细化的第二查询发送给所述文档管理系统中的至少一个文档管理系统。

而且，例如，通过使多个文档管理系统并行执行文档搜索的文档搜索设备来实现本发明。该文档搜索设备包括：发送单元，其将关于输入的第一搜索条件的第一查询发送给所述多个文档管理系统中的每个文档管理系统；接收单元，其从所述多个文档管理系统中的至少一个文档管理系统接收对所述第一查

询的搜索结果；搜索控制单元，其使所述发送单元发送用于请求对由所述接收单元接收到的所述搜索结果进行搜索细化的第二查询；管理单元，其对由所述接收单元从所述多个文档管理系统接收到的搜索结果的应答状态进行管理；以及选择单元，其根据由所述管理单元所管理的所述应答状态，针对所述文档管理系统中的每个文档管理系统分别选择是否将所述第二查询发送给所述文档管理系统中的每个文档管理系统。

而且，例如，通过一种文档管理系统来实现本发明。该文档管理系统包括：接收单元，其从根据权利要求1所述的文档搜索设备接收查询和用于识别该查询的识别信息；搜索单元，其基于该查询执行文档搜索；发送单元，其将搜索结果发送给所述文档搜索设备；以及存储单元，其相互关联地存储所述搜索结果和所述识别信息，其中，所述搜索单元在从所述文档搜索设备接收到附有所述识别信息的搜索细化请求时，对所存储的搜索结果执行搜索细化。

而且，例如，通过一种文档搜索系统来实现本发明。该文档搜索系统包括上述文档搜索设备和多个上述文档管理系统。

而且，例如，通过使多个文档管理系统并行执行文档搜索的文档搜索方法来实现本发明。该文档搜索方法包括以下步骤：将关于输入的第一搜索条件的第一查询发送给所述多个文档管理系统中的每个文档管理系统；从所述多个文档管理系统中的每个文档管理系统接收对所述第一查询的搜索结果；对来自所述多个文档管理系统中的每个文档管理系统的对所述第一查询的搜索结果的应答状态进行管理；根据所管理的所述应答状态，针对所述文档管理系统中的每个文档管理系统分别选择用于请求对所述第一查询进行搜索细化的方案；以及在从所有所述文档管理系统接收到对所述第一查询的搜索结果之前且在从所述

多个文档管理系统中的至少一个文档管理系统接收到对所述第一查询的搜索结果之后，基于所选择的方案，将用于请求所述搜索细化的第二查询发送给所述文档管理系统中的至少一个文档管理系统。

而且，该文档搜索方法包括以下步骤：将关于输入的第一搜索条件的第一查询发送给所述多个文档管理系统中的每个文档管理系统；从所述多个文档管理系统中的至少一个文档管理系统接收对所述第一查询的搜索结果；发送用于请求对基于所述第一查询的所述搜索结果进行搜索细化的第二查询；对从所述多个文档管理系统接收到的搜索结果的应答状态进行管理；以及根据所述应答状态，针对所述文档管理系统中的每个文档管理系统分别选择是否将所述第二查询发送给所述文档管理系统中的每个文档管理系统。

通过以下（参考附图）对典型实施例的说明，本发明的其它特点将变得显而易见。

附图说明

图1是用于说明根据实施例的文档搜索设备的整体构成的图；

图2是用于说明根据实施例的文档搜索设备的主单元的硬件构成的框图；

图3是示出根据实施例的文档搜索设备的外观的图；

图4是示出根据实施例的文档搜索设备的软件构成的图；

图5是关于根据实施例的文档管理系统的硬件构成的示例性框图；

图6是关于根据实施例的文档管理系统的软件构成的示例性框图；

图7是当使用文档管理系统执行搜索细化时的序列图；

图8是示出由根据实施例的文档搜索设备所存储的查询状况的数据格式的例子的图；

图9是示出根据实施例的搜索结果显示窗口的例子的图；

图10是示出根据实施例的搜索细化条件的输入窗口的例子的图；

图11是根据实施例的搜索方法的例子的流程图；

图12是根据实施例的搜索细化请求的发送方法的例子的流程图；

图13是示出根据实施例请求搜索细化的方案的选择处理的例子的图。

具体实施方式

以下说明本发明的一个实施例。不必说，下述各个实施例有利于理解本发明的上位概念、中位概念和下位概念等各种概念。通过权利要求书的范围确定本发明的技术范围，并且以下各个实施例不以任何方式限制本发明的技术范围。

图1是用于说明根据实施例的文档搜索设备的整体构成的图。文档搜索设备100包括作为图像输入装置的扫描器102、作为图像输出装置的打印机103、作为用户界面装置的操作部分104、以及用于控制这些装置的主单元101。在文档搜索设备100内部或外部设置包含用于保存、管理和搜索文档的功能的文档管理系统105。由该文档搜索设备100和文档管理系统105等构成文档搜索系统。

扫描器102、打印机103和操作部分104通过内部总线与主单元101连接。主单元101与LAN 110和公共交换电话网（public switched telephone network, PSTN）120等通信介质连接。文档管理系统105还与LAN 110连接。因此，主单元101和文档管理系统105通过LAN 110通信数据和命令等。注意，LAN 110可以是例如因特网等广域网。

LAN 110与多个其它文档管理系统130、140和150连接。当请求文档搜索时，文档搜索设备100不仅可以向内部文档管理系统105发送查询，而且还可以向其它文档管理系统130、140和150发送查询。

这里，文档搜索设备100构成为使得可被操作者识别为单个设备。换句话说，为了成为所谓的多功能外围设备，文档搜索设备100具有除文档搜索功能以外的文档复制功能、传真功能、打印机功能和文档存储功能。而且，文档搜索设备100可以是不包含扫描器功能和打印机功能的普通个人计算机。

通常将文档管理系统105称为搜索引擎、搜索服务器和数据库服务器。通常可将文档搜索设备100称为搜索客户机和客户计算机。文档管理系统130、140和150可以是因特网上的搜索引擎和门户网站（portal site）服务器等。在这种情况下，文档搜索设备100是运行web浏览器程序的个人计算机。

图2是用于说明根据实施例的文档搜索设备的主单元的硬件构成的框图。

主单元101是用于进行文档数据、图像信息和装置信息的输入和输出的控制器。主单元101通过控制扫描器102和打印机103实现扫描功能和打印功能。主单元101将文档存储到通过LAN 110连接的文档管理系统105，并通过搜索获取所存储的文档。主单元101通过使用公共交换电话网（public switched telephone network, PSTN）120实现传真功能。主单元101与作为用户界面的操作部分104连接，并控制操作部分104。

CPU 201是用于控制整个主单元101的控制器。RAM 202是用于使CPU 201能够运行的RAM 202。例如，RAM 202是用于保存管理队列的存储设备等，该管理队列用于管理搜索结果的应答状态和存储图像数据。ROM 203是存储引导程序的引导ROM。

HDD 204是硬盘驱动器等大容量存储设备。HDD 204存储各种类型的软件、图像数据及其属性数据、以及其它使用数据等。通过存储在HDD 204中的系统软件来实现根据该实施例的文档搜索功能。

操作部分接口206是将操作部分（UI）104连接到CPU 201的接口。操作部分接口206将所显示的图像数据输出到对操作部分104设置的显示设备。操作部分接口206接收操作者（用户）使用对操作部分104设置的触摸面板和按键输入的信息，并将

该信息发送给CPU 201。搜索结果被显示到操作部分104的显示装置。用户通过操作部分104输入搜索条件。

NIC (network interface card, 网络接口卡) 207是用于与在LAN 110上所连接的文档管理系统105和个人计算机等外部装置进行通信的通信单元。总之, NIC 207用作数据和信息的发送单元或接收单元。

例如, NIC 207支持作为网络通信协议的TCP/IP。CPU 201根据TCP/IP与文档管理系统105进行通信。CPU 201还通过NIC 207与其它文档管理系统130等进行通信。

调制解调器208使用传真功能通过公共交换电话网120将图像信息输入到传真设备或从传真设备输出图像信息。上述装置通过系统总线209相互连接。

图像总线接口205用作用于转换数据结构的总线桥。用于高速传送图像数据的图像总线230和系统总线209通过图像总线接口205连接。

在图像总线230中布置以下装置。光栅图像处理器(raster image processor, RIP) 231将从与LAN 110连接的PC输入的PDL (page description language, 页面描述语言) 代码展开成位图图像。装置接口部分232是用于将扫描器102和打印机103连接到主单元101的接口装置。

扫描器图像处理部分233是用于校正、处理和编辑输入图像数据的装置。打印机图像处理部分234是用于校正、处理和编辑输出图像数据的装置。图像转动部分235与扫描器图像处理部分233链接, 并转动通过扫描器102读取的图像数据, 将该图像数据存储于RAM 202中, 并转动RAM 202中的图像数据。

图像压缩部分236对多值图像数据执行例如JPEG压缩和解压缩处理, 对二值图像数据执行例如JBIG、MMR、MR和

MH压缩和解压缩处理。分辨率转换部分237转换RAM 202中的图像数据的分辨率，并将其存储在RAM 202中。色彩空间转换部分238使用矩阵计算将例如YUV图像转换成Lab图像，并将其存储在RAM 202中。色调转换部分239使用误差扩散处理（error diffusion process）等方法将例如RAM 202中的8位256级图像数据转换成1位2级图像数据，并将其存储在RAM 202中。

图3是示出根据该实施例的文档搜索设备的外观的图。如图1所述，文档搜索设备100是包括文档复制功能、传真功能、打印机功能和文档管理功能的复合机器。

扫描器102通过用光照射原稿纸张上的图像并利用CCD线型传感器扫描该图像，将该图像转换成作为光栅图像数据的电信号。原稿纸张被放置在原稿给送器301的托盘中。当操作者通过操作部分104发出读取命令时，主单元101的CPU 201命令扫描器102进行读取。原稿给送器301根据读取命令每次给送一张原稿纸张。

打印机103是用于将光栅图像数据转换成纸张上的图像的图像形成单元。图像形成系统可以是使用感光鼓和感光带的电子照相系统，或者可以是通过从微喷嘴阵列喷墨而在纸张上直接打印图像的喷墨系统。

打印机103在从CPU 201接收到打印命令时开始图像形成处理。打印机103包括多个供纸盒302，在供纸盒302中，贮存不同大小的纸张和不同方向的纸张。排纸托盘303接收完成了图像形成的纸张。排纸托盘303根据来自CPU 201的命令，执行整理处理和装订处理等。

注意，由于图1中所示的主单元101和文档管理系统105布置在文档搜索设备100内，因而从文档搜索设备100的外观不能

看见它们。

图4是示出根据该实施例的文档搜索设备的软件构成的图。软件模块被存储在HDD 204中。

UI管理器401是用于控制操作部分104的模块。例如，UI管理器401进行将操作者命令通信给下层模块并将来自下层模块的信息显示到操作部分104的控制。

为了从LAN 110接收命令和信息并从主单元101发送命令和信息，网络控制管理器402控制NIC 207。

复制应用程序403从UI管理器401接收复制命令，并将复制作业发送给作业管理器409。发送应用程序404从UI管理器401接收发送命令，并将发送作业发送给作业管理器409。

文档管理应用程序405通过向文档管理系统105发送信息和命令并从文档管理系统105接收信息和命令，实现文档管理功能。文档管理应用程序405执行例如搜索细化的请求、搜索结果的接收、搜索结果的应答状态的管理。文档管理应用程序405与用于提取文档图像中所包含的文本信息的OCR引擎和用于根据文档图像修整图像区域的块选择(block selection, BS)引擎407连接。

公共接口408是用于吸收关于扫描器102和打印机103等的装置控制的装置依赖性的接口。作业管理器409整理从公共接口408所接收的作业数据，并将必要的命令发送到用于控制下层装置的各种控制管理器。例如，在接收到复制作业时，作业管理器409向扫描管理器411发送读取命令，并向打印管理器410发送图像形成命令。注意，打印管理器410是用于控制打印机103的模块。扫描管理器411是用于控制扫描器102的模块。

如果接收到传真发送作业，则作业管理器409将必要的命令发送给扫描管理器411、文件管理器414和传真管理器412。

文件管理器414是用于管理文档文件的模块。传真管理器412是用于控制调制解调器208的模块。

如果以LIPS或PostScript接收到PDL打印作业，则作业管理器409将必要的命令发送给PDL管理器413、文件管理器414和打印管理器410。PDL管理器413控制RIP 231，并将PDL数据转换成光栅图像数据。

图5是关于根据该实施例的文档管理系统的硬件构成的示例性框图。

CPU 501是用于基于计算机程序提供对该文档管理系统中的所有单元的总体的控制的控制单元。RAM 502是用作工作区的易失性存储单元。ROM 503是用于存储固件（firmware）和其它控制程序的非易失性存储单元。硬盘驱动器（hard disk drive, HDD）504是大容量存储单元。通信接口（NIC）505是网络接口卡等通信单元。

图6是关于根据该实施例的文档管理系统的软件构成的示例性框图。文档管理系统105和130~150提供文档的存储、获取和搜索等各种功能。

网络控制模块601控制NIC 505以与主单元101的网络控制管理器402进行通信。例如，网络控制模块601从文档搜索设备100接收查询和用于识别该查询的识别信息（搜索ID），并将搜索结果发送给文档搜索设备100。

搜索控制模块602根据来自主单元101的查询执行搜索，并创建所存储的文档的搜索索引。注意，搜索控制模块602在接收到附有识别信息的搜索细化请求时，可以基于在已提前执行搜索后存储在临时数据区607中的搜索结果，执行搜索细化。在这种情况下，搜索控制模块602将搜索结果和识别信息关联起来，并将其提前存储在临时数据区607中。

图像文档处理模块603根据文档创建缩略图图像，并在需要时执行图像转换。图像文档处理模块603还从文档中提取搜索索引。换句话说，图像文档处理模块603将文档转换成更易于处理的格式，并使用OCR引擎604提取搜索索引和文本信息。图像文档处理模块603使用块选择引擎605以提取文档中所包含的图像区域。

文档管理模块606将与所有文档管理有关的处理分配给各模块。临时数据区607是文档管理系统105等所使用的临时数据的存储位置。

数据库控制模块608对保存在数据库609~612中的数据进行创建、存储和删除等。数据库控制模块608根据来自主单元101的请求从相应数据库读取数据，并根据所读取的数据创建将传送给主单元101的文档。

卷数据库 (volume database) 609是用于存储实际文档的数据库。卷数据库609是概念性的东西，实际上可以是OS中的文件系统。属性数据库610是用于保存关于文档名称、创建日期和注解等属性的信息的数据库。在全文搜索数据库611中，登记根据卷数据库609中所登记的文档的文本数据创建的搜索索引。当主单元101发出全文搜索的命令时，搜索控制模块602搜索全文搜索数据库。

在图像搜索数据库612中，与卷数据库609中所登记的文档相关联地登记从该文档中剪取的特征数据。当主单元101发出图像搜索的命令时，搜索控制模块602根据搜索目标图像的特征搜索图像搜索数据库612，并提取与这些特征相关联的文档。

图7是当使用文档管理系统执行搜索细化时的序列图。这里，多个文档管理系统中的文档管理系统105请求进行搜索。注意，当执行多服务器搜索时，将具有相同搜索条件的查询发

送给每个文档管理系统。

当根据操作者的操作，将具有搜索条件的搜索命令701从操作部分104输入到主单元101时，CPU 201从NIC 207向文档管理系统105发送查询702。当作为软件操作，UI管理器接收具有搜索条件的搜索命令701时，UI管理器将该搜索命令701传送给文档管理应用程序405。文档管理应用程序405创建附有搜索条件的查询702，并将该查询702与文档管理系统105的网络地址一起，传送给网络控制管理器402。网络控制管理器402根据该网络地址将查询702发送给文档管理系统105。

查询702中所包含的参数包括例如用于从该查询识别其它查询的识别信息（以下称为“搜索ID”）和从操作部分104输入的搜索条件。搜索ID必须是唯一的，因此可取的是CPU 201基于网络地址等选择搜索ID。

在通过NIC 505接收到查询702时，文档管理系统105的CPU 501根据查询702执行搜索。关于软件，搜索控制模块602通过网络控制模块601接收查询702，并根据查询702中所包含的搜索条件搜索各种数据库。

注意，CPU 501（搜索控制模块602）根据搜索结果和搜索ID创建临时数据703，并将临时数据703存储在HDD 504中所留出的临时数据区607中。当CPU 501（搜索控制模块602）接收到搜索终止命令713时，或者一旦经过了提前设置的超时间，CPU 501就删除临时数据703。

CPU 501（搜索控制模块602）生成搜索完成通知704，并将该通知发送给请求该搜索的文档搜索设备100。搜索完成通知704的参数有例如文档管理系统105的地址、匹配文档的数量（命中数量）、以及所搜到文档的属性信息等。可以将匹配文档的数量称为搜索结果的数量。

搜索完成通知704不包含来自搜索结果的任何实际数据。因此，为了接收实际数据（实际文档），主单元101的CPU 201（文档管理应用程序405）向文档管理系统105发送搜索结果的获取请求705。此时，CPU 201发送指定搜索结果内的所需范围（例如，命中20~60等）的获取请求705。可以通过操作部分104来指定该范围。不用说，获取请求705包括搜索ID。

文档管理系统105的CPU 501（搜索控制模块602）在接收到获取请求705时，基于附有获取请求705的搜索ID，从临时数据区607读取相应的搜索结果。注意，当指定范围时，CPU 501（搜索控制模块602）仅从临时数据703中提取与该范围相对应的那些搜索结果。

CPU 501（搜索控制模块602）创建结果列表通知706，并将结果列表通知706发送给文档搜索设备100。结果列表通知706包括用于区别该查询的搜索ID和搜索结果列表。

文档搜索设备100的CPU 201（文档管理应用程序405）在接收到结果列表通知706时，创建搜索结果的显示数据707，并将该显示数据707输出到操作部分104。这样，将搜索结果显示到操作部分104的显示设备。

注意，操作部分104进行多任务操作。由于这个原因，即使在具有多个文档管理系统的环境中，操作部分104也按照接收顺序显示来自各文档管理系统的搜索结果。换句话说，甚至在从所有文档管理系统接收到搜索结果之前，就开始显示已经接收到的那些搜索结果。

此时，如果从操作部分104将基于已执行搜索的结果的、附有搜索细化条件的细化命令708输入CPU 201，则CPU 201创建搜索细化请求709，并将该搜索细化请求709发送给文档管理系统105。搜索细化请求709的参数包括例如新发出的搜索

ID、用于指定过去的查询702的原来的搜索ID和用于细化搜索的附加搜索条件。不用说，原来的搜索ID必须与查询702中所包含的ID相匹配。

文档管理系统的CPU 501（搜索控制模块602）在接收到搜索细化请求709时，从搜索细化请求709中提取原来的搜索ID，并从临时数据区607中读取相应的过去搜索的结果。而且，CPU 501（搜索控制模块602）根据从搜索细化请求709中提取的附加搜索条件，基于过去搜索的结果执行搜索细化。CPU 501创建包括这些搜索结果的新的临时数据710，并将该新的临时数据710存储在临时数据区607中。注意，CPU 501可以在此时删除原始临时数据703。

之后，与原始搜索结果一样，发送和接收搜索细化结果的搜索完成通知704、获取请求705和结果列表通知706。将搜索细化结果的显示数据711显示到操作部分104的显示设备。

如果操作者发出再次进行搜索细化的命令，则重复上述一系列搜索细化处理（708～711）。

最后，当将搜索终止命令712从操作部分104输入CPU 201时，CPU 201向文档管理系统发送搜索终止命令713。文档管理系统的CPU 501从临时数据区607中删除临时数据703和710，该临时数据703和710是关于发送搜索终止命令713的文档搜索设备100的一系列搜索结果。

图8是示出由根据该实施例的文档搜索设备所存储的搜索结果的应答状态的数据格式的一个例子的图。CPU 201（文档管理应用程序405）将用于对文档管理系统请求的搜索中的搜索结果的应答状态进行管理的数据存储在RAM 202中。

801是管理队列。使用管理队列801来管理本实施例的搜索结果的应答状态。换句话说，每当生成查询时，CPU 201就将

作为查询目的地的每个文档管理系统的管理数据802、803和804压栈到管理队列801中。

压栈到队列中的管理数据为例如文档管理系统的地址、应答状态、以及匹配文档的数量等。文档管理系统的地址只要是唯一的，可以是任何形式，例如网络地址。应答状态是例如关于搜索状况的数据，例如“查询中”或“搜索完成”等。匹配文档的数量是搜索中所搜到文档的数量（所谓的命中数量）。

在本实施例中，当出现查询时采用用于将管理数据放入管理队列801中的方法，但是本发明不局限于此。例如，当接收到查询时，CPU 201可以将管理数据放入管理队列中。在这种情况下，当广播查询时，仅将关于已返回搜索结果的查询的数据压栈到管理队列中。因此，具有能够简单识别有搜索结果应答的查询的优点。

图9是示出根据该实施例的搜索结果显示窗口的例子的图。将用于显示搜索结果的窗口900显示到操作部分104的显示设备。如上所述，操作部分104采用触摸面板，并且通过操作显示到窗口900的按钮905~909执行与搜索有关的各种功能。

搜索结果的显示区域901是用于显示适用搜索条件的文档列表的区域。操作者可以在显示区域901中选择所期望的文档。

缩略图显示区域902是用于显示当前所选择的文档的缩略图图像的区域。匹配文档数量的显示区域903是用于显示适用搜索条件的文档总数的区域。存储位置的显示区域904是用于显示关于当前所选择的文档的存储位置的信息的区域。

下一按钮905是当超过在显示区域901中可显示的匹配文档的最大数量时所激活的按钮，并且是用于从超过最大数量的文档管理系统获取并显示其它文档的按钮。例如，我们假定匹配文档的数量为800，并且当前在显示区域901中显示编号

200~300。在这种情况下，当操作下一按钮905时，CPU 201为了获取编号300~400的搜索结果，向文档管理系统发送获取请求705。然后，CPU 201将所获取的编号300~400的搜索结果显示到显示区域901。

打开按钮906是用于命令预览显示当前所选择的文档的按钮。文件夹显示按钮907是用于显示存储所选择文档的文件夹的按钮。搜索按钮908是用于命令对当前的搜索结果进行搜索细化的按钮。例如，如果检测到已操作了搜索按钮908，则为了输入在搜索细化中使用的附加搜索条件，CPU 201将例如对话框显示到操作部分104。关闭按钮909是用于命令终止搜索的按钮。例如，如果操作了关闭按钮909，则CPU 201将初始窗口显示到操作部分104。

当将查询发送给多个文档管理系统时，一旦从至少一个文档管理系统返回搜索结果，CPU 201就显示窗口900。之后，每当从其它文档管理系统接收到搜索结果时，CPU 201就实时更新窗口900。

图10是示出根据该实施例的搜索细化条件的输入窗口的例子。当按下窗口900中的搜索按钮908时，显示用于输入搜索条件的输入窗口1000。

显示区域1001是用于示出搜索条件的种类的区域。在本实施例中，提供表示文档名称和文档所有者等文档属性信息的“属性”和允许操作者添加任何属性的“系统索引”的种类。操作者从各种类当中输入他或她希望指定为搜索条件的项。

例如，如果选择“日期”作为种类，则CPU 201显示属于所选择种类的搜索项1002。操作者可以为每个显示项输入搜索条件。在图10中，输入创建日期的条件。

搜索按钮1003是用于命令执行搜索的按钮。关闭按钮1004

是用于终止搜索细化的按钮。当按下关闭按钮1004时，CPU 201关闭输入窗口1000，并将窗口900显示在前面。全部清除按钮1005是用于清除所有输入条件的按钮。

图11是示出根据该实施例的搜索方法的例子的流程图的图。注意，主要通过文档管理应用程序405来执行该搜索方法。

在步骤S1101，CPU 201接收用于使多个文档管理系统105和130~150并行执行文档搜索的搜索条件的输入。根据如上所述操作者在操作部分104中的操作，指定搜索条件的内容。注意，在第一实施例中，基于用户在操作部分104中输入的字符串，创建搜索条件，但是也可以使用其它方法。例如，可以通过扫描器102读取原稿上的图像并将其作为图像数据输入，并且基于该图像数据进行上述图像搜索。可以基于上述OCR引擎604从扫描器102所读取的图像中提取的文本数据，来创建搜索条件。

在步骤S1102，CPU 201创建附加了输入的搜索条件和唯一搜索ID的查询，并将该查询发送给文档管理系统。

在步骤S1103，CPU 201待机直到接收到至少一个搜索结果为止。当接收到至少一个搜索结果时，处理进入步骤S1104。注意，如果在将窗口900显示到操作部分104时未接收到新的搜索结果，则处理可进入步骤S1105。

在步骤S1104，CPU 201获取搜索结果并将其显示到操作部分104的显示设备。如上所述，每当接收到新的搜索结果时，CPU 201就更新窗口900。

在步骤S1105，CPU 201判断是否已通过操作部分104发出了进行搜索细化的命令。如上所述，CPU 201检测是否操作了窗口900所提供的搜索按钮908。如果未发出搜索细化的命令，则处理进入步骤S1108。注意，这里基于来自用户的命令判断

是否进行搜索细化，但是，当S1103中所接收的搜索结果中的命中数量等于或大于预定数量时，可以进行设置以自动进行搜索细化。在这种情况下，需要将创建日期或用于限制搜索目的地（文档管理系统）的信息提前设置为搜索细化的搜索条件。

当发出搜索细化的命令时，在步骤S1106，CPU 201接收通过操作部分104所接收的用于搜索细化的附加条件。

注意，当从至少一个文档管理系统接收到搜索结果时，CPU 201可以基于该搜索结果创建搜索细化的搜索条件。因而使得更容易输入搜索细化的条件。

在步骤S1107，不管是否已从所有文档管理系统接收到搜索结果，CPU 201都发送搜索细化请求。注意，CPU 201可以根据来自各文档管理系统的搜索结果的应答状态，改变或选择用于请求搜索细化的方案。参见以下对用于请求搜索细化的方案的详细说明。

在步骤S1108，CPU 201判断是否通过操作部分104发出了终止搜索的命令。如果未发出终止搜索的命令，则处理返回到步骤S1103。

另一方面，如果发出了终止搜索的命令，则处理进入步骤S1109，并且CPU 201向文档管理系统发送搜索终止命令713。

根据该实施例，可以在不等待从所有多个文档管理系统都接收到搜索结果的情况下，执行搜索细化。因而可以在多服务器搜索中提高搜索细化的处理效率。

通过使用接收到的搜索结果创建搜索细化的条件，使得更容易输入搜索细化的条件。

图12是根据该实施例的搜索细化请求的发送方法的例子的流程图。搜索细化请求的发送方法是上述步骤S1107的子程序。

在步骤S1201, CPU 201将用于管理各文档管理系统的搜索结果的应答状态的管理队列801保存到RAM 202。注意, 当改变应答状态时, 例如, 当返回搜索结果时等, CPU 201更新相应的文档管理系统登记在管理队列801中的管理数据。

在步骤S1202, CPU 201初始化用于区别文档管理系统的索引 (i 是从1到N的自然数; N为文档管理系统的总数)。

在步骤S1203, CPU 201从第*i*个文档管理系统的管理队列801中读取管理数据。在步骤S1204, CPU 201根据该管理数据所表示的应答状态, 选择搜索细化的请求方案, 并发送搜索细化的请求。

在步骤S1205, 判断是否为所有文档管理系统选择了用于请求搜索细化的方案。如果未完成选择和发送, 则处理在步骤S1206中使索引*i*加1后, 返回到步骤S1203。如果完成了选择和发送, 则处理进入主流程图。

根据该实施例, 可以通过提前保存来自各文档管理系统的应答状态, 根据该应答状态来改变用于请求各文档管理系统的搜索细化的方案。根据搜索速度和数据库中的登记文档的数量不同, 预期来自各文档管理系统的搜索结果的应答在搜索时间方面不同。根据应答状态, 通过选择最佳应答方法更进一步地提高了搜索效率。

图13是根据该实施例的用于请求搜索细化的方案的选择处理的例子的流程图。该选择处理是上述步骤S1204的子程序。

在步骤S1301, CPU 201基于管理数据802~804判断所期望的文档管理系统的应答状态是否是“搜索终止”(没有查询; 换句话说, 已返回了搜索结果)。

如果没有查询, 则处理进入步骤S1302, CPU 201判断文档管理系统所提取的文档数量(匹配文档的数量)是否为一个

或多个。

如果匹配文档的数量是一个或多个，则处理进入步骤S1303，将搜索细化请求709发送给文档管理系统。假定如上所述将文档管理系统的地址提前登记在管理数据中。而且，将CPU 201新生成的搜索ID添加到搜索细化请求。

另一方面，如果匹配文档的数量为0，则处理进入步骤S1304，抑制（禁止或跳过）CPU 201向文档管理系统发送搜索细化请求。这是因为请求匹配文档的数量为0的文档管理系统的搜索细化，或换句话说，请求没有命中文档的文档管理系统的搜索细化，将会导致零结果，自然地，可以通过进行以上判断省略无用的查询。

在步骤S1301，如果判断为“搜索未完成”（查询；换句话说，还未返回最初请求的搜索的结果），则处理进入步骤S1305。在步骤S1305，CPU 201在原始搜索条件上利用为搜索细化所添加的条件创建新的搜索条件，并将所创建的搜索条件作为新的查询发送给正被查询的文档管理系统。注意，将CPU 201新生成的搜索ID添加到新的查询。

同样注意，在步骤S1305，CPU 201可以向判断为正在查询的文档管理系统发送中断搜索的命令。优选地，搜索中断命令能够终止价值降低的搜索处理。为了指定被中断的查询，优选将搜索ID附加在搜索中断命令中。

利用该实施例，通过将原始搜索ID和搜索细化条件发送给完成了搜索且存在一个或多个匹配文档的文档管理系统，来进一步提高搜索效率。换句话说，将基于原始搜索的搜索结果存储在临时数据区607中，因此对这些搜索结果执行搜索细化，可以执行比通常更快的搜索处理。如在图7中所述，可以通过作为识别信息的搜索ID来指定原始搜索结果。

而且，可以通过不发送完成了搜索且匹配文档的数量为0的文档管理系统的搜索细化请求，来抑制无用搜索的执行。

此外，对于未完成搜索的文档管理系统，可以通过将搜索细化条件添加到原始搜索条件以创建新的搜索条件，来更有效地执行搜索细化。换句话说，由于实际上可以在不等待原始搜索结果的情况下执行搜索细化，因而提高了搜索效率。

注意，在搜索未终止的文档管理系统中，或换句话说，在应用“搜索未完成”的文档管理系统中，通常对初始查询保持执行搜索处理。由于这个原因，请求搜索细化使得并行执行两个搜索，而不等待原始搜索结果。这增加了文档管理系统的处理负担，因而是不可取的。当然，不用说，完成搜索细化的时间较慢。因此，发出指令以中断第一查询提高了搜索细化的效率，并减轻了文档管理系统的负担。

其它实施例

讨论了略有不同的实施例，并且本发明可应用于由多个装置构成的系统或者可应用于由单个装置组成的设备。例如，其包括扫描仪、打印机、个人计算机、复印机、多功能外围设备和传真装置。

可以通过向系统或设备直接或远程提供实现上述实施例各种功能的软件程序、包括在该系统等中的计算机读取并执行所提供的程序代码，来实现本发明。

因此，为了在计算机上实现本发明的功能和处理，安装在计算机上的程序代码本身实现本发明。换句话说，用于实现这些功能和处理的计算机程序本身就是本发明。

在这种情况下，程序格式无关紧要，可以是目标代码、由解释程序执行的程序、提供给OS的脚本数据等，只要具有该程序的功能即可。

作为用于提供该程序的存储介质，例如有软盘、硬盘、光盘、磁光盘、MO、CD-ROM、CD-R、以及CD-RW等。作为存储介质，还有磁带、非易失性RAM 202卡、ROM、DVD（DVD-ROM、DVD-R）等。

可以使用客户计算机上的浏览器从因特网网站下载该程序。换句话说，可以将本发明的计算机程序或包含自动安装功能的压缩文件从网站下载到硬盘或其它存储介质。可将构成本发明的程序的程序代码分成多个文件，并通过从不同网站下载这些文件来实现该程序代码。换句话说，甚至允许多个操作者下载用于在计算机上实现本发明的功能和处理的程序文件的WWW服务器也是本发明的构成要件。

此外，可以对本发明的程序进行加密，将其存储在CD-ROM或其它存储介质上，并分发给操作者。在这种情况下，可以仅允许满足规定条件的操作者通过因特网从网站下载用于解密的密钥信息，使用该密钥信息对加密程序执行解密，并在计算机上安装该程序。

还可以通过执行由计算机读取的程序来实现上述实施例的功能。而且，可以由运行在计算机上的OS基于该程序的命令，执行全部或部分实际处理。不必说，在这种情况下也可以实现上述实施例的功能。

而且，可以将从存储介质读取的程序写到对插入计算机中的功能扩展板或与计算机连接的功能扩展单元所设置的RAM 202。可以由对该功能扩展板或功能扩展单元所设置的CPU基于该程序的命令，执行全部或部分实际处理。从而实现上述实施例的功能。

利用本发明，可以提高当在多个文档管理系统上执行多服务器搜索时的搜索细化的效率。

尽管参考典型实施例说明了本发明，但应当理解，本发明不局限于所公开的实施例。所附权利要求书的范围符合最宽的解释，以包含所有这类修改和等同结构和功能。

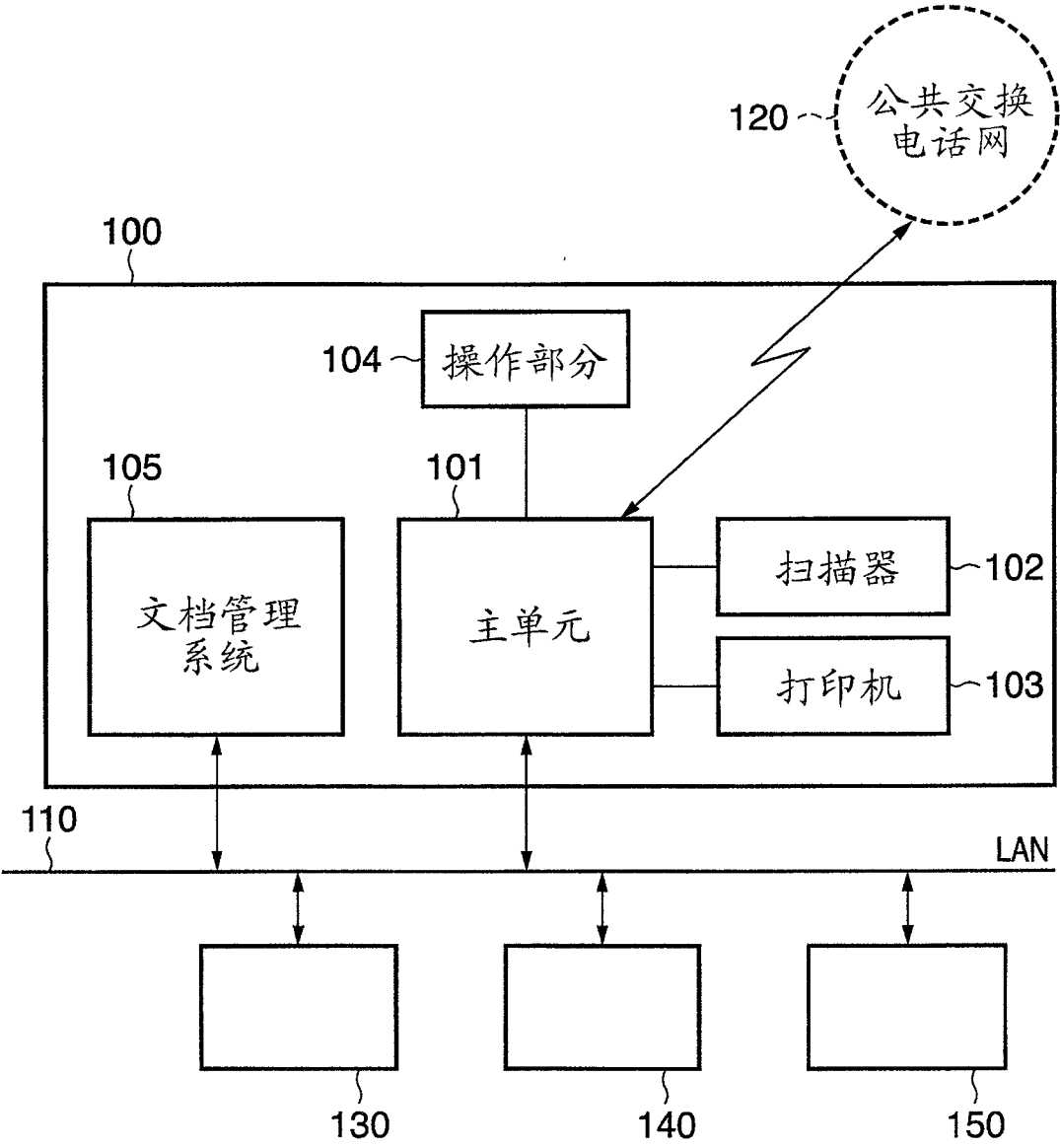


图 1

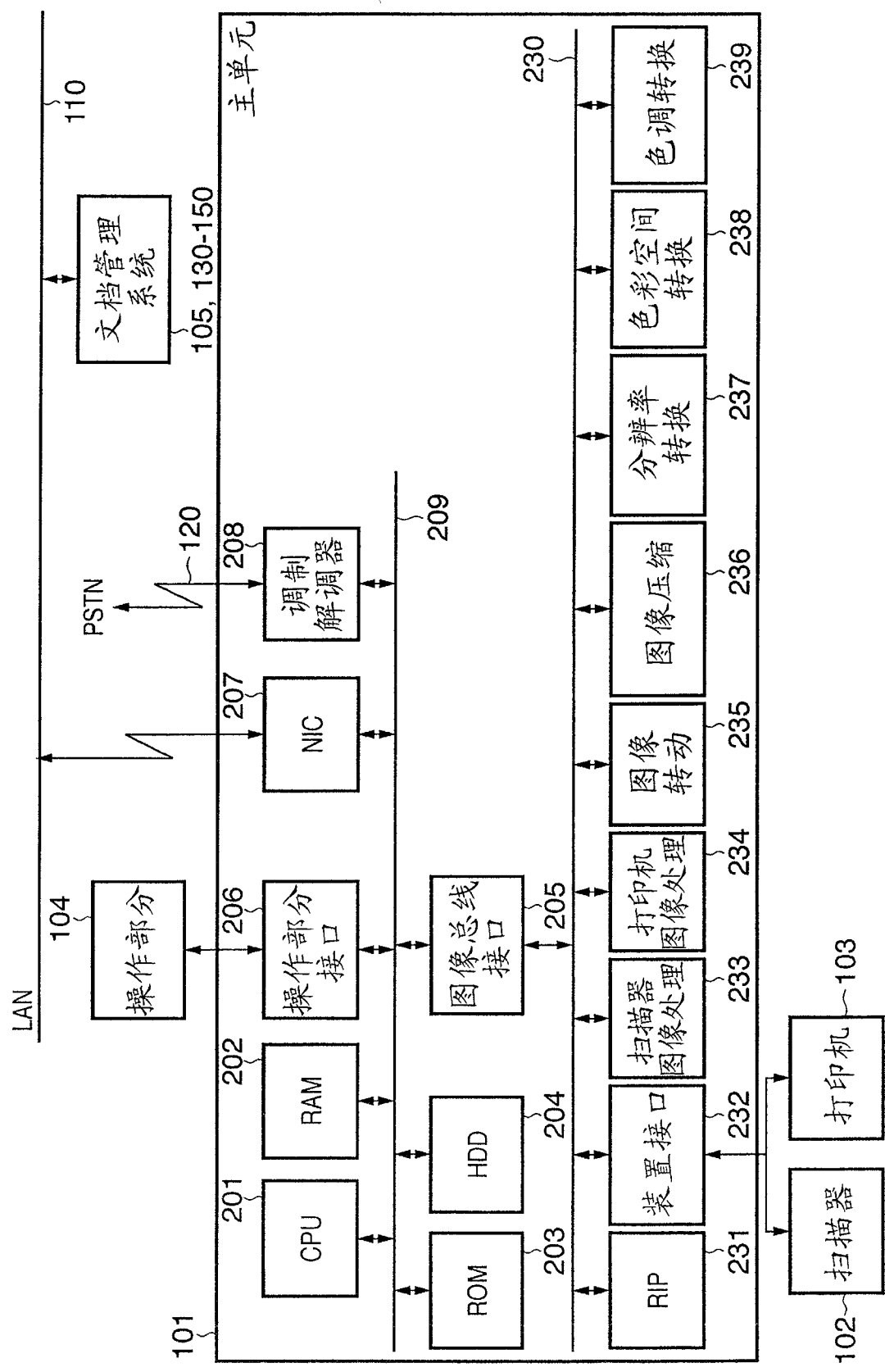


图 2

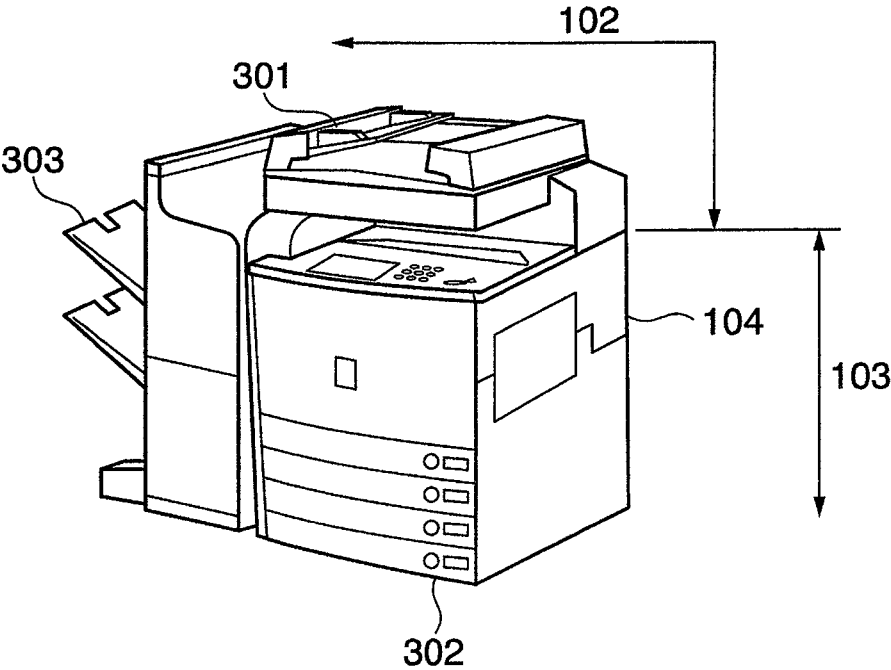


图 3

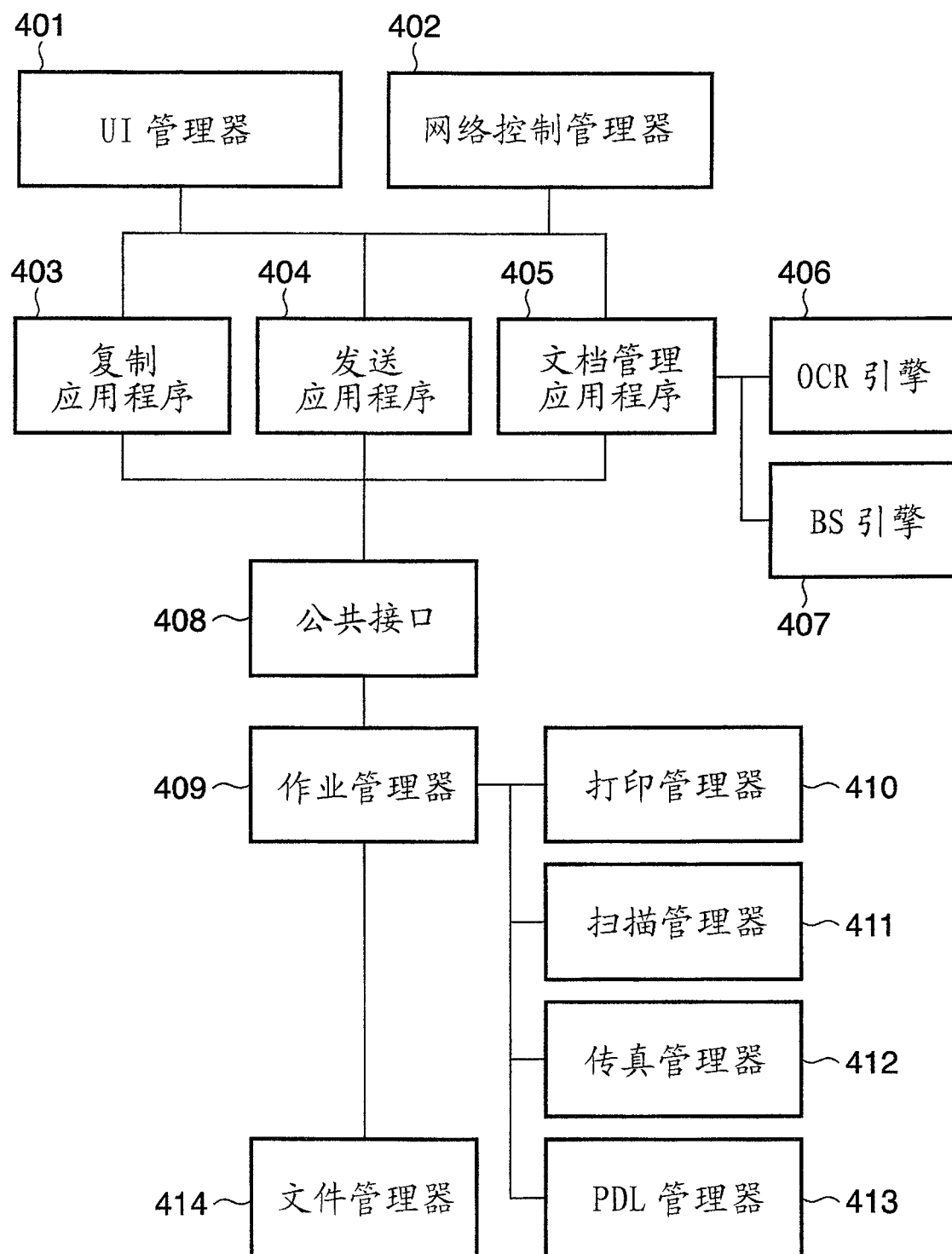


图 4

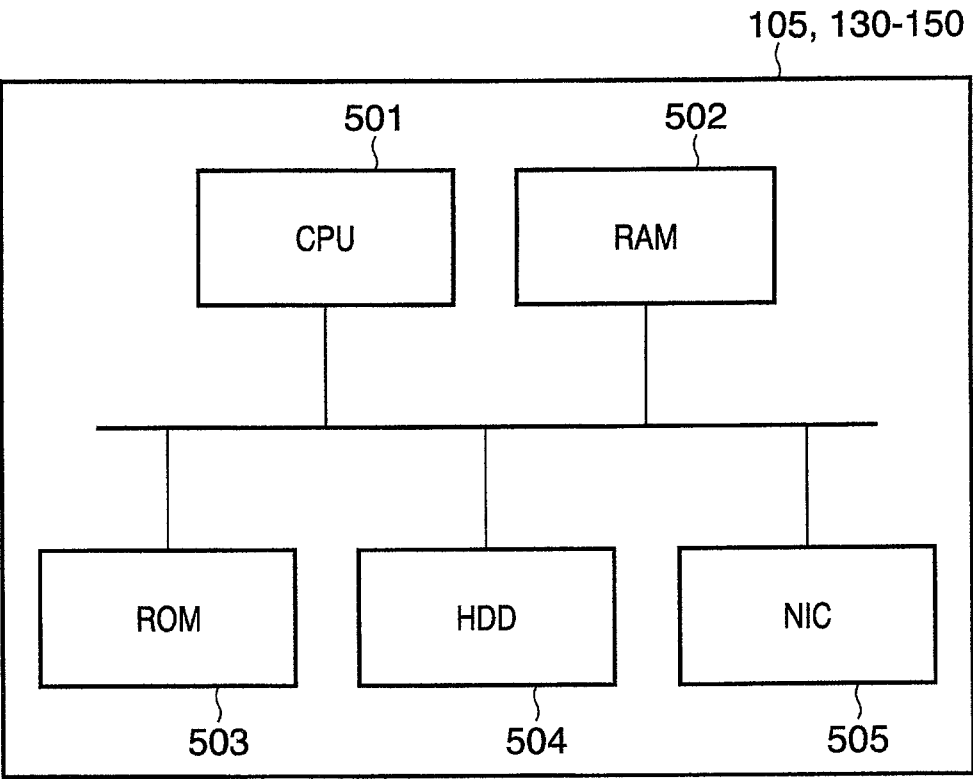


图 5

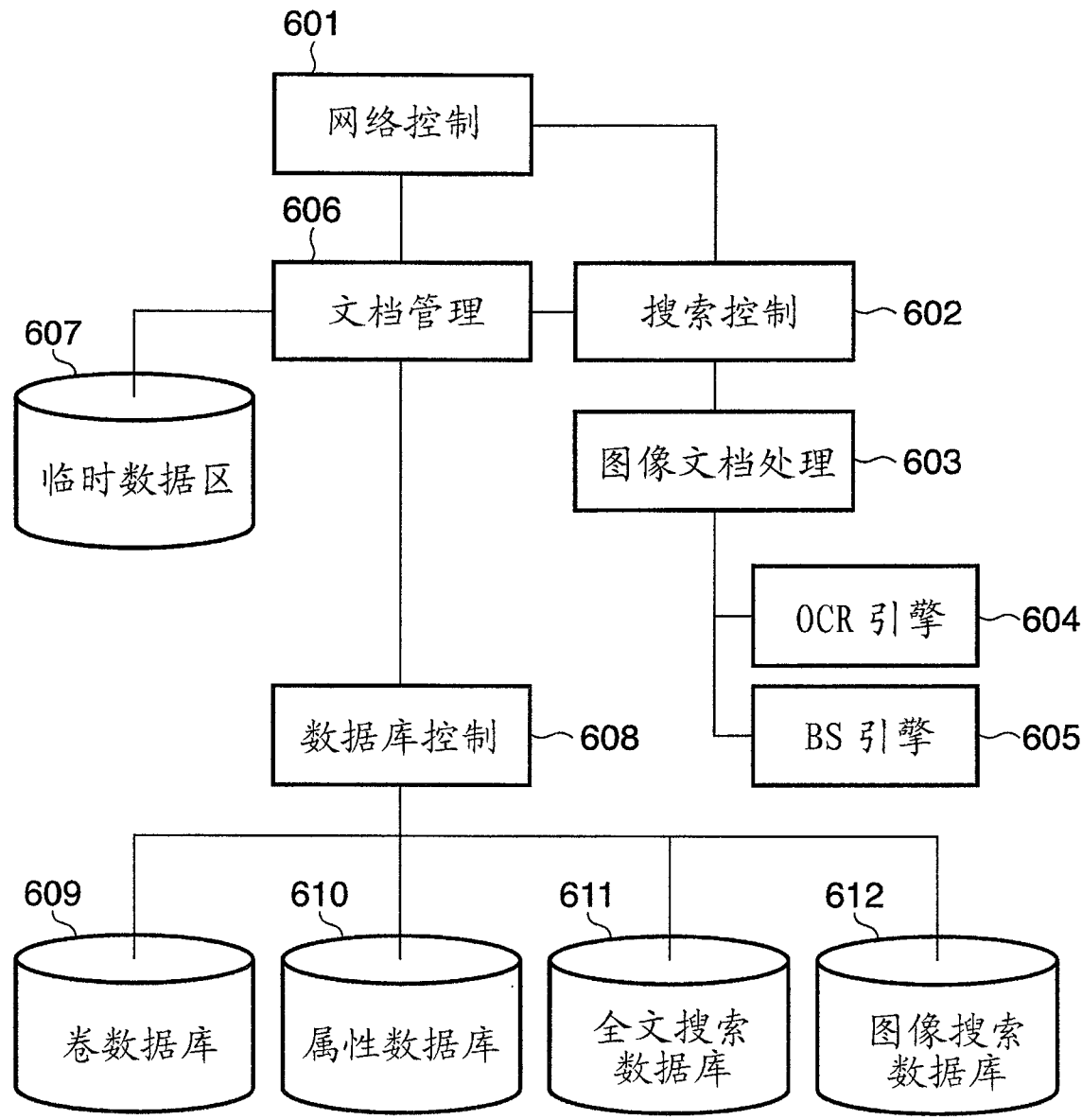


图 6

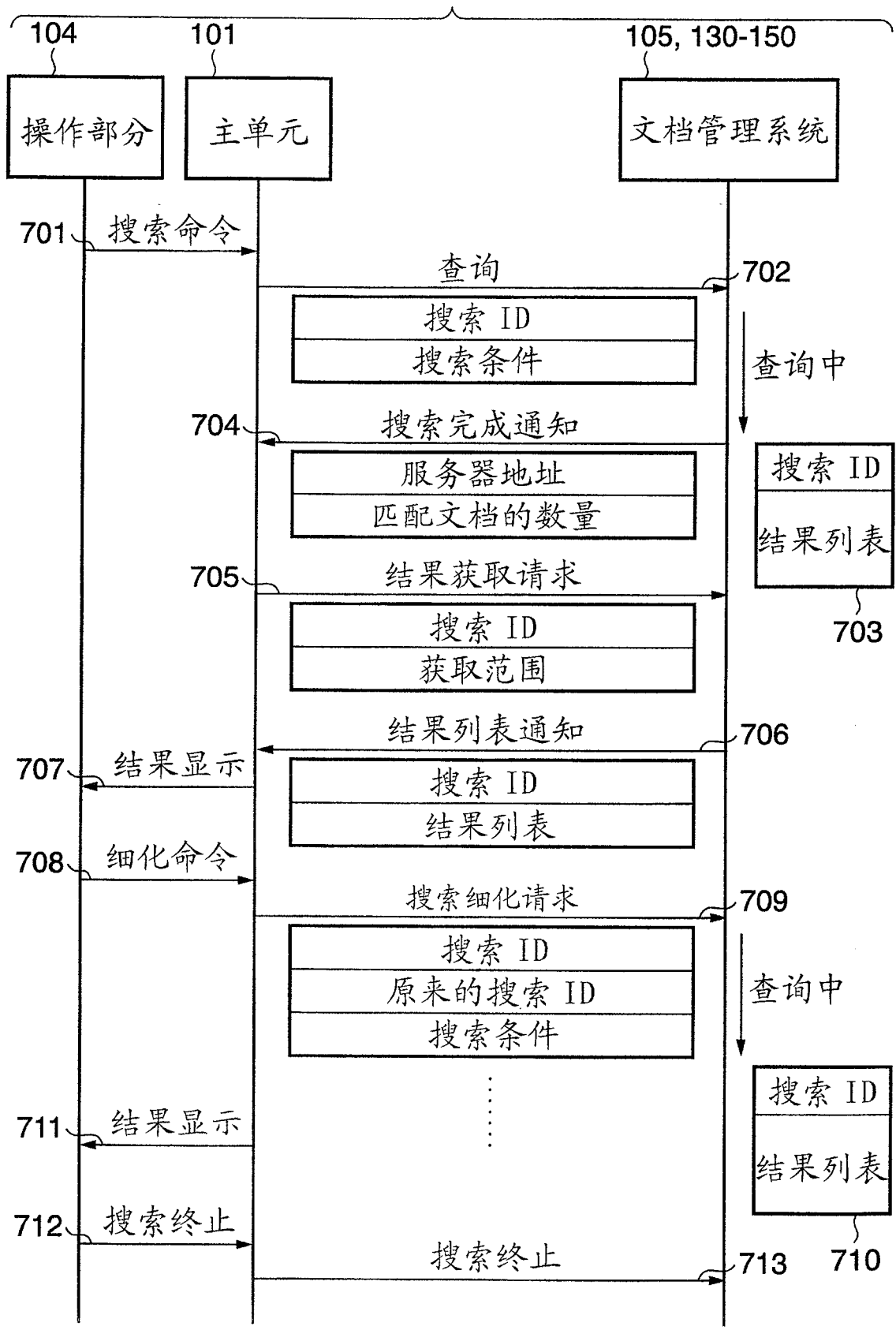


图 7

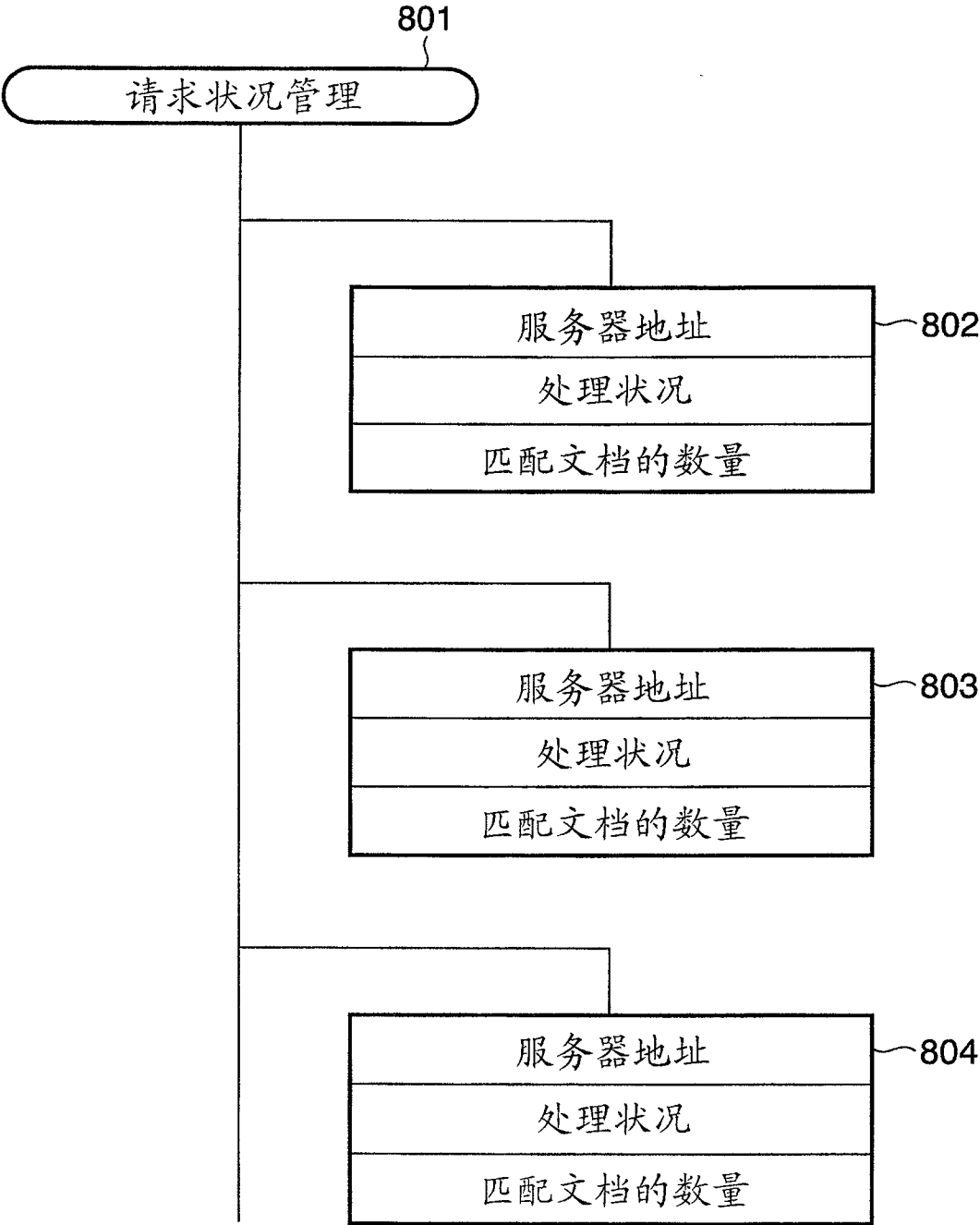


图 8

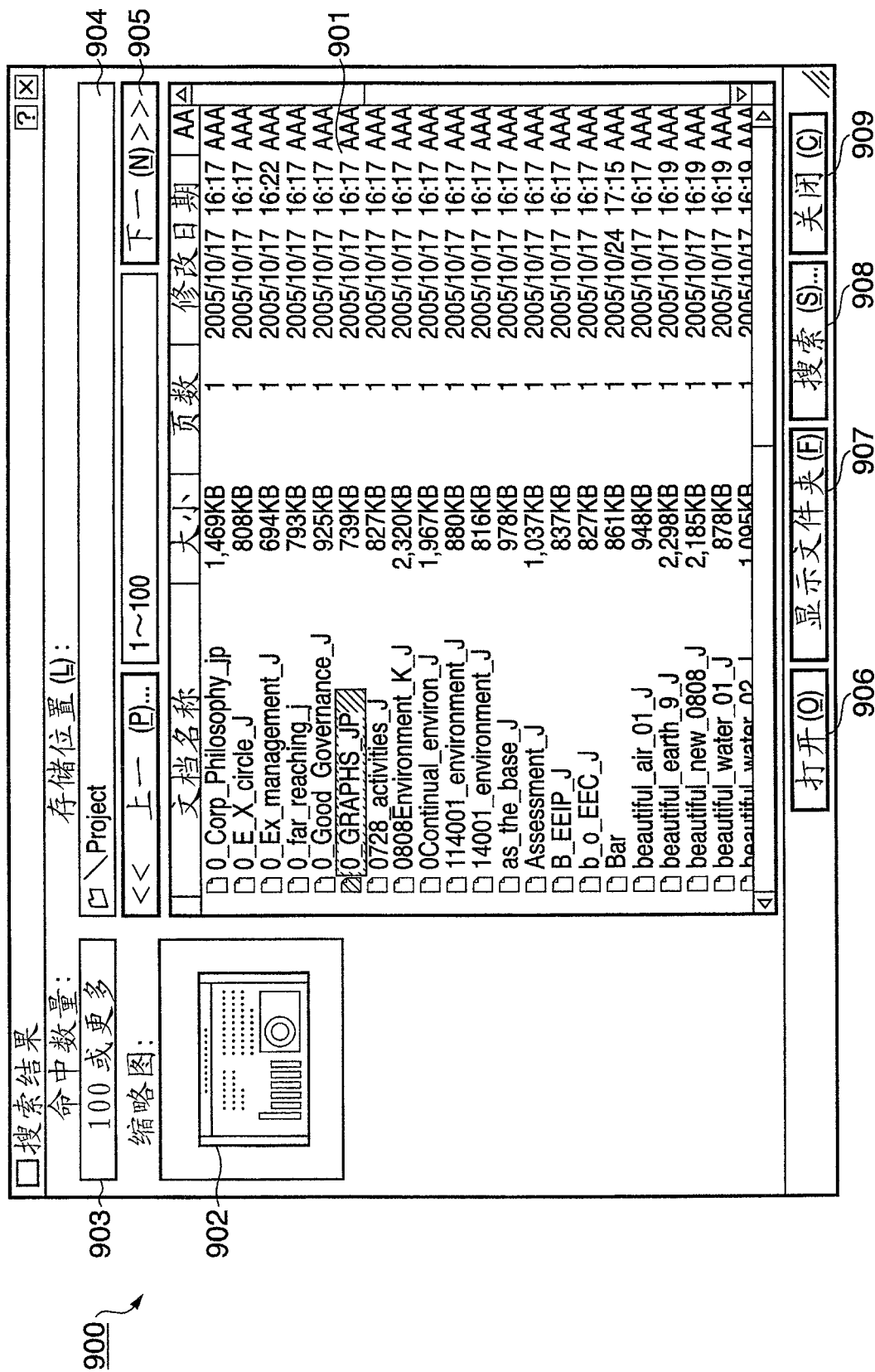


图 9

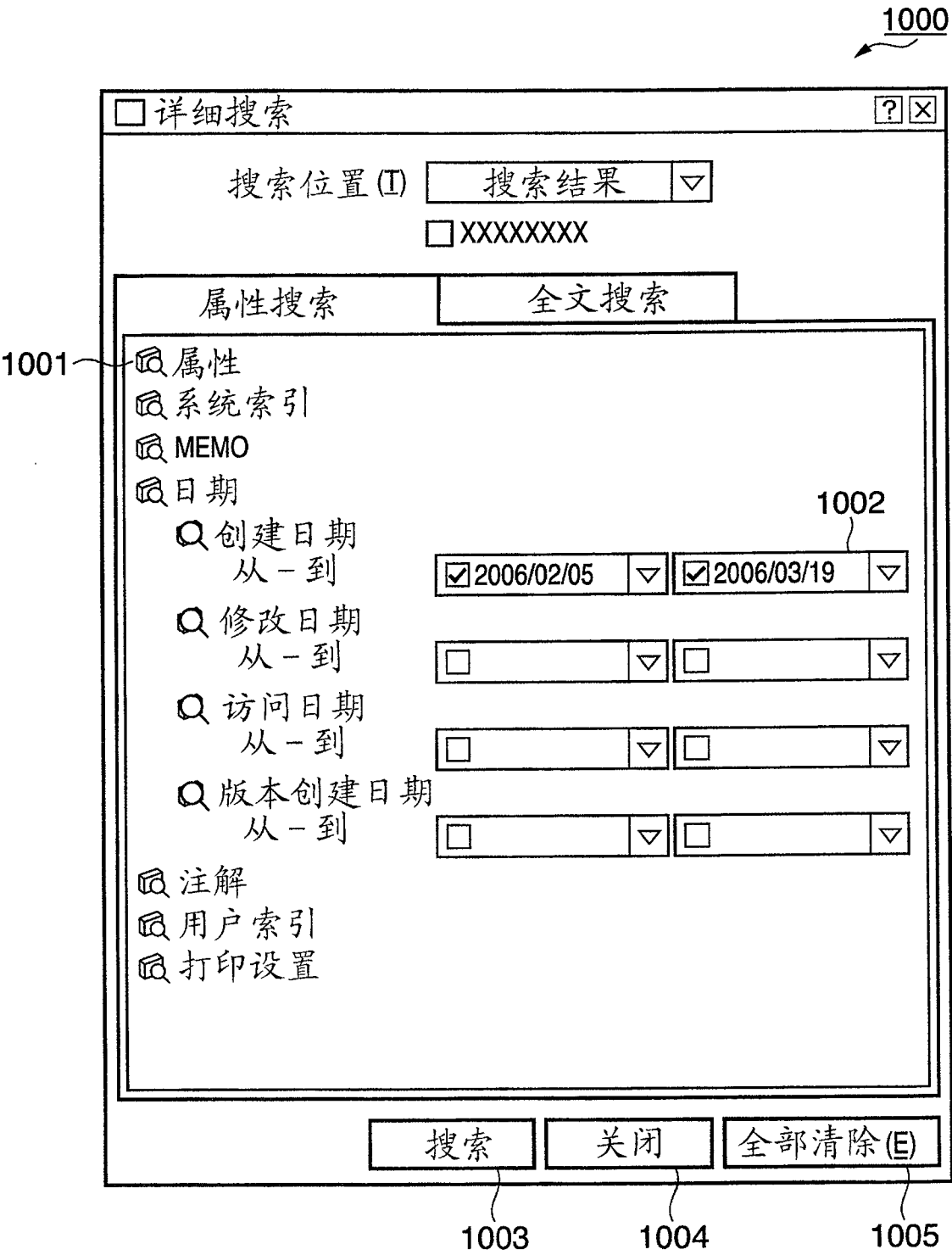


图 10

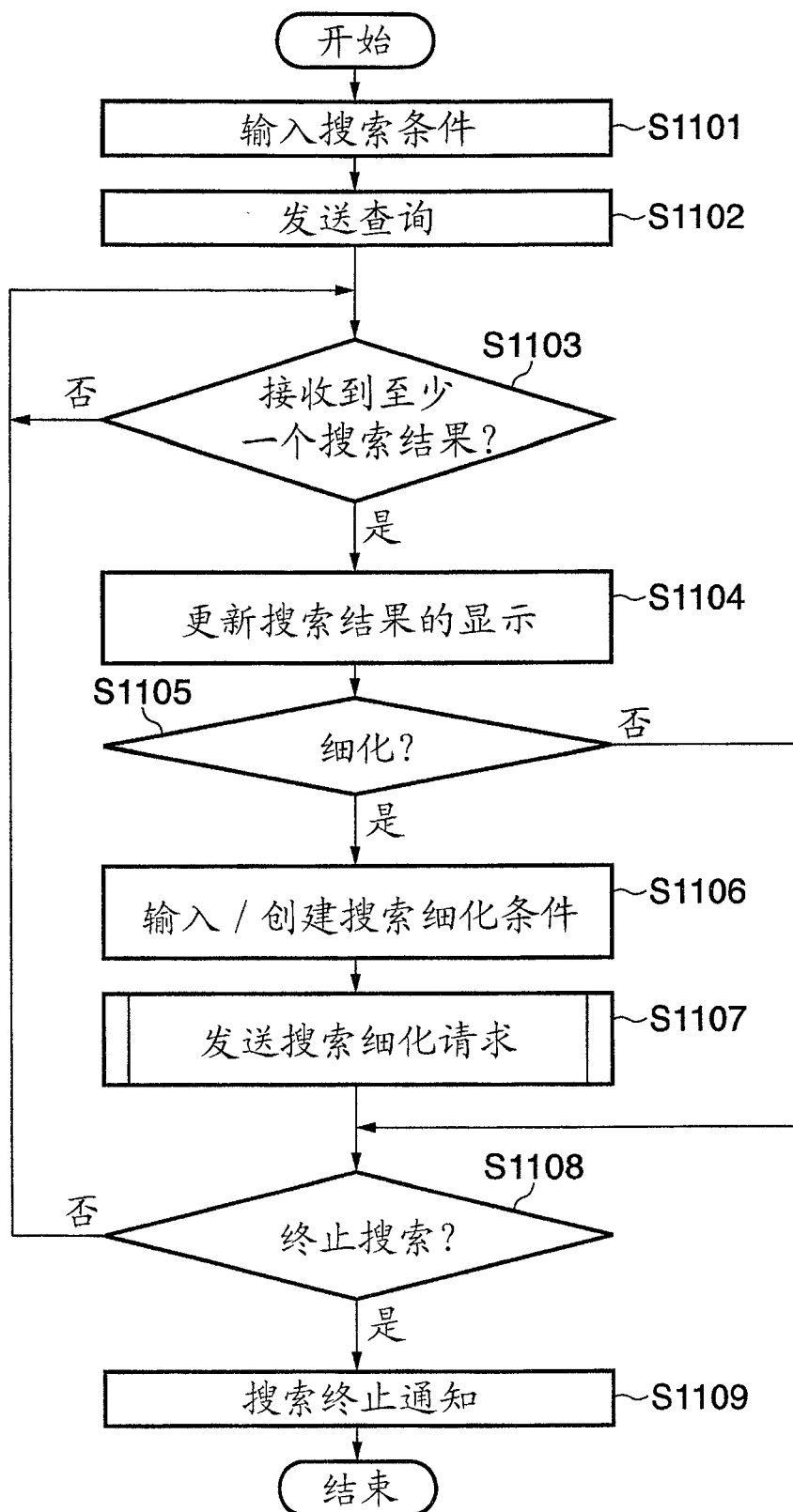


图 11

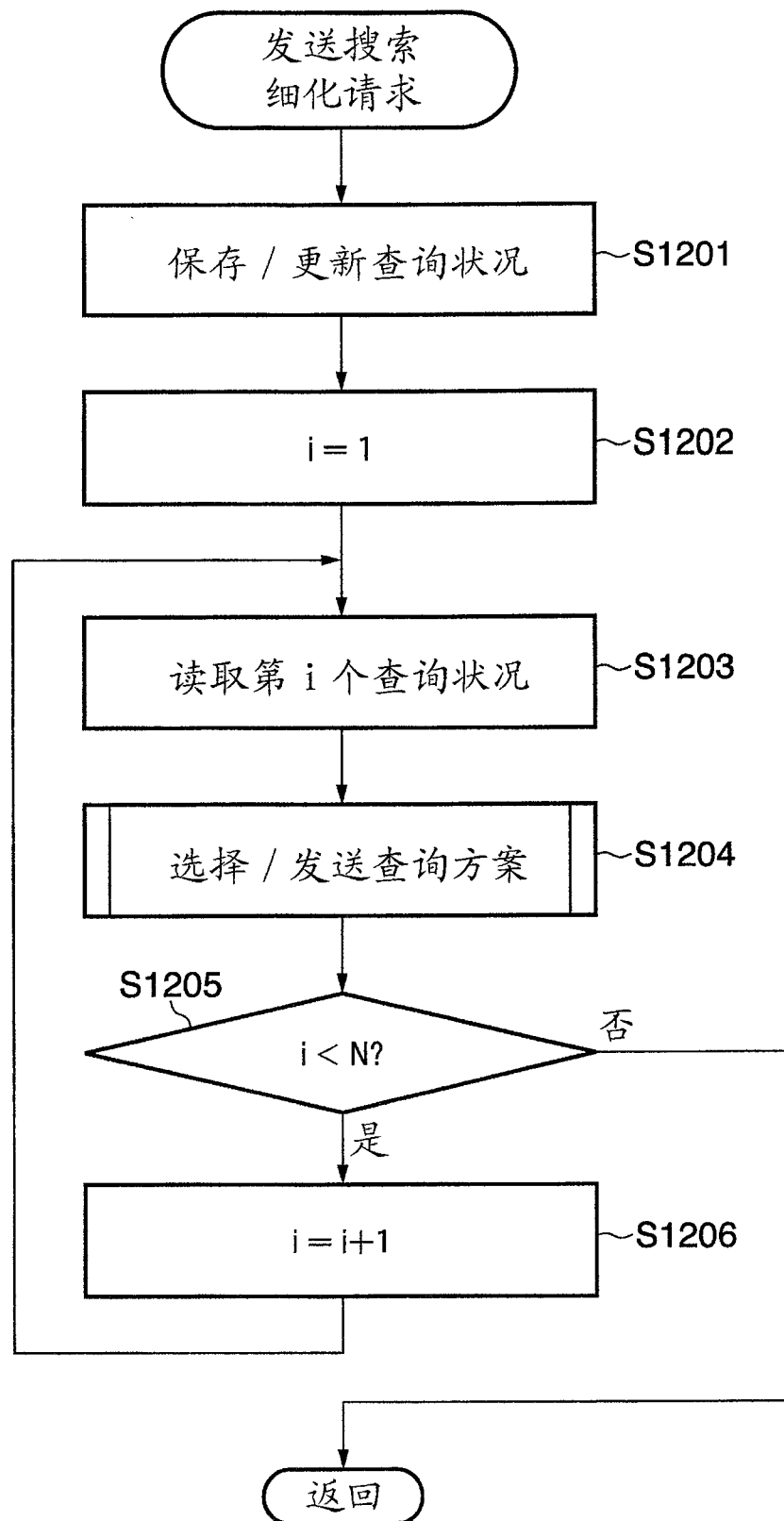


图 12

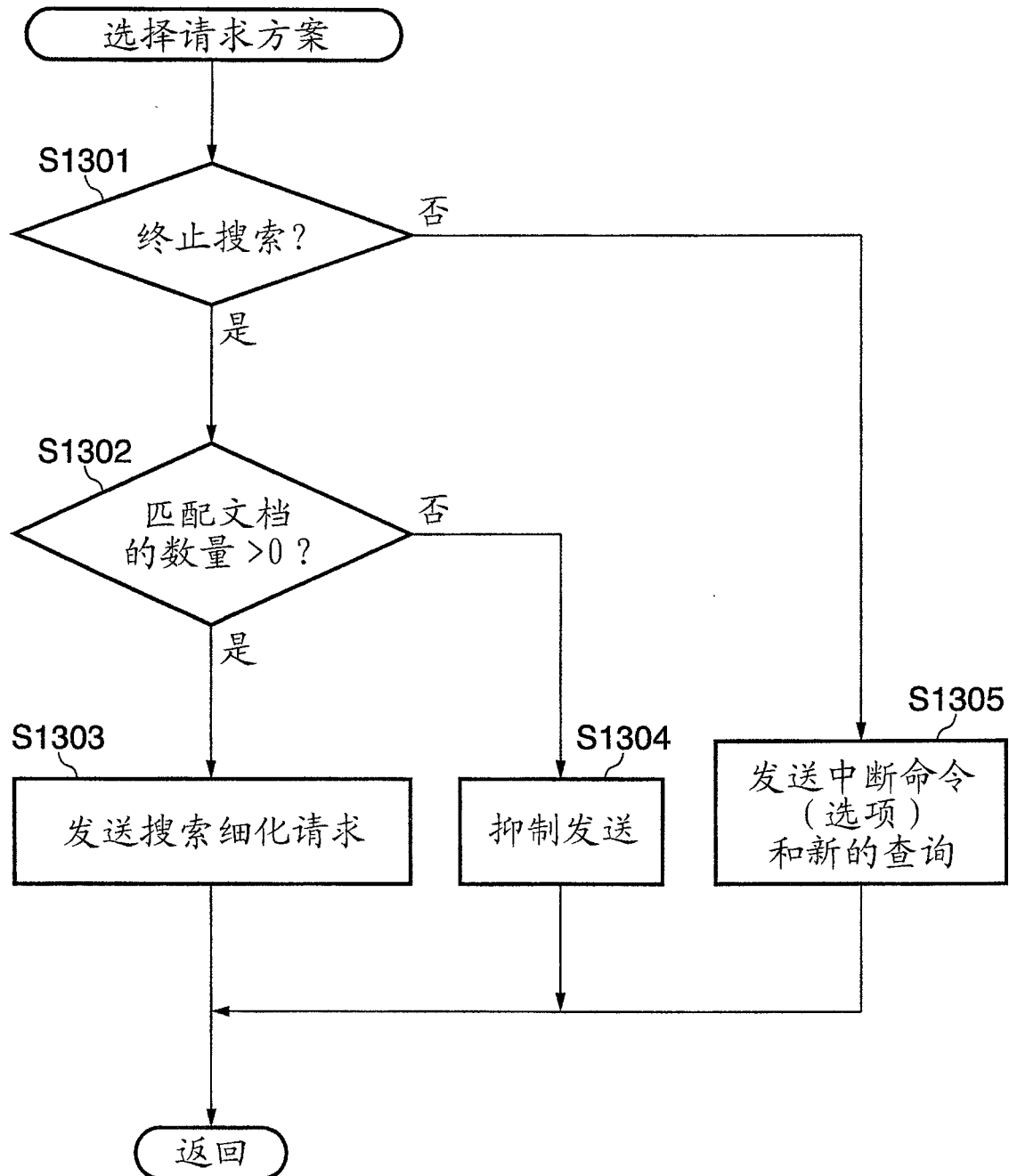


图 13