



[12] 发明专利申请公开说明书

[21] 申请号 200510007838.4

[43] 公开日 2005年8月3日

[11] 公开号 CN 1648902A

[22] 申请日 2005.1.26
 [21] 申请号 200510007838.4
 [30] 优先权
 [32] 2004. 1.26 [33] US [31] 10/765,819
 [71] 申请人 微软公司
 地址 美国华盛顿州
 [72] 发明人 E·L·梅斯 E·B·瓦特森
 O·荷斯特-希勒

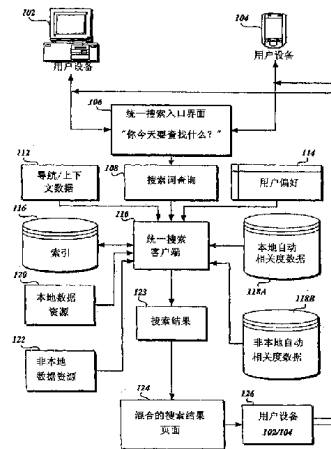
[74] 专利代理机构 上海专利商标事务所有限公司
 代理人 谢喜堂

权利要求书4页 说明书11页 附图8页

[54] 发明名称 统一和混合搜索的系统和方法

[57] 摘要

提供了用于全异资源的统一搜索的系统、方法和计算机可访问媒质。该系统和方法确定要在其中搜索的最相关的全异资源，并混合来自搜索的最相关的结果，以用综合的方式向用户呈现。



I S S N 1 0 0 8 - 4 2 7 4

1. 一种便于从全异的资源中搜索内容的方法，其特征在于，所述方法包括：
响应于一搜索内容的请求，显示一统一搜索入口界面；
- 5 从多个全异资源中确定在其中搜索在所述统一搜索入口界面中输入的搜索词的出现的至少一个相关资源；
归一化所述相关全异资源；以及
混合从在所述归一化相关全异资源中搜索所述搜索词的出现中获得的全异结果。
- 10 2. 如权利要求 1 所述的方法，其特征在于，归一化所述相关全异资源包括：
确定所述搜索词在第一全异资源的一个位置中出现相对于在第二全异资源的另一位置中出现的相对重要性；以及
依照所述相对重要性对所述资源中所述搜索词的出现加权。
3. 如权利要求 2 所述的方法，其特征在于，混合所述全异结果包括按照所述
15 搜索词的出现权值对所述结果分级、按照与所述资源相关联的分类显示所分级的结果、以及显示其中所述搜索词在每一分类的可比较位置中出现的位置。
4. 如权利要求 1 所述的方法，其特征在于，它还包括：
捕捉所述搜索请求的上下文；以及
其中，确定所述至少一个相关资源是基于所述搜索请求上下文。
- 20 5. 如权利要求 1 所述的方法，其特征在于，它还包括：
获得所述多个全异资源的每一个的相关度的自动测量；
其中，确定所述至少一个相关资源是基于所述相关度的自动测量的。
6. 如权利要求 5 所述的方法，其特征在于，所述相关度的自动测量是量化用户与
所述资源的交互的度量。
- 25 7. 如权利要求 6 所述的方法，其特征在于，所量化的用户交互包括用户访问
所述资源的频率、用户访问所述资源的时间长度、以及用户在所述资源上执行的行动的重要性中的至少其中之一。
8. 如权利要求 1 所述的方法，其特征在于，它还包括：
获得指示要在其中搜索的较佳资源的用户偏好；
- 30 其中，确定所述至少一个相关资源是基于所述用户偏好中指示的所述较佳资

源。

9. 如权利要求 1 所述的方法，其特征在于，它还包括：
为所述全异资源构建一索引；以及
在搜索所述搜索词的出现时搜索与所述相关资源相关联的索引。
- 5 10. 一种用于电子内容的统一搜索的系统，其特征在于，所述系统包括：
包含可搜索内容的多个全异资源；
一用于储存为所述资源收集的自动相关度数据的存储库；
一搜索词输入；以及
一处理器，它基于所述自动相关度数据确定所述全异资源中的哪一个与所述
10 搜索词输入最相关、归一化所述相关资源、在所归一化的相关资源中搜索所述搜索
词、以及混合来自所述搜索的结果以向用户显示，其中，所述结果表示从所述归一
化相关资源的至少一个获得的相关但全异的内容。
11. 如权利要求 10 所述的系统，其特征在于，归一化所述相关资源包括从所
述全异资源中确定要在其中搜索所述搜索词的出现可比较数据。
- 15 12. 如权利要求 11 所述的系统，其特征在于，混合来自所述搜索的结果包括
基于所述可比较数据中所述搜索词的出现相对重要性对所述全异结果分级，以及
以统一的方式显示来自每一个所述结果的可比较数据。
13. 如权利要求 10 所述的系统，其特征在于，它还包括一用户偏好输入，并
且其中，确定所述全异资源的哪一个与所述搜索词输入最相关是基于所述用户偏好
20 输入。
14. 如权利要求 10 所述的系统，其特征在于，它还包括一上下文输入，其中，
确定所述全异资源的哪一个与所述搜索词输入最相关是基于所述上下文。
15. 如权利要求 14 所述的系统，其特征在于，所述上下文输入表示正引导到
所述搜索词的输入的导航事件，并且确定所述全异资源的哪一个与所述搜索词输入
25 最相关是基于所述导航事件。
16. 如权利要求 14 所述的系统，其特征在于，所述上下文输入表示所述搜索
词的输入从其起源的应用程序，并且其中，确定所述全异资源的哪一个与所述搜索
词输入最相关是基于所述应用程序。
17. 如权利要求 10 所述的系统，其特征在于，它还包括一与所述全异资源中
30 的内容相关联的索引；并且其中，在所归一化的相关资源中搜索所述搜索词包括
搜索与所述资源相关联的索引。

18. 如权利要求 10 所述的系统，其特征在于，所述自动相关度数据包括一表示用户访问所述资源的频率、用户访问所述资源的时间长度、以及用户在所述资源上执行的行动的重要性中的至少其中之一之测量的度量。

5 19. 如权利要求 18 所述的系统，其特征在于，所述最相关的全异资源是具有最高频率、最大时间长度、以及最重要行动中的至少其中之一的那些资源。

20. 一种具有用于执行电子内容的统一搜索的指令的计算机可访问媒质，其特征在于，所述指令包括：

响应于一搜索内容的请求，显示一统一搜索入口界面；

10 从多个全异资源中确定在其中搜索在所述统一搜索入口界面中输入的搜索词的出现的至少一个相关资源；

归一化所述相关全异资源；以及

混合从在所述归一化的相关全异资源中搜索所述搜索词的出现中获得的全异结果。

15 21. 如权利要求 20 所述的计算机可访问媒质，其特征在于，归一化所述相关全异资源的所述指令包括：

确定所述搜索词出现在第一全异资源的一个位置中相对于所述搜索词出现在第二全异资源的另一位置中的相对重要性；以及

依照所述相对重要性对所述搜索词在所述搜索词的出现加权。

20 22. 如权利要求 20 所述的计算机可访问媒质，其特征在于，混合所述全异结果的指令包括按照所述搜索词的出现权值对所述结果分级、按照与所述资源相关联的分类显示所分级的结果、以及显示在其中所述搜索词在每一分类的可比较位置中出现的位置。

25 23. 如权利要求 20 所述的计算机可访问媒质，其特征在于，所述指令还包括：捕捉所述搜索请求的上下文；以及基于所述搜索请求的上下文确定所述至少一个相关资源。

24. 如权利要求 20 所述的计算机可访问媒质，其特征在于，所述指令还包括：获得所述多个全异资源的每一个的相关度的自动测量；以及基于所述相关度的自动测量确定所述至少一个相关资源。

30 25. 如权利要求 24 所述的计算机可访问媒质，其特征在于，所述相关度的自动测量是表示用户与所述资源的交互的数字，其中，所述用户交互包括用户访问资源的频率、用户访问资源的时间长度、以及用户在资源上执行的行动的重要性中的

至少其中之一，并且当所述自动相关度数字高时，所述资源与所述搜索更相关，当所述自动相关度数字低时，所述资源较不相关。

26. 如权利要求 20 所述的计算机可访问媒质，其特征在于，所述指令还包括：
获取一指示要在其中搜索的较佳资源的用户偏好；以及
- 5 基于所述用户偏好中指示的所述较佳资源确定所述至少一个相关资源。

统一和混合搜索的系统和方法

5 技术领域

本发明一般涉及计算机软件和搜索引擎，尤其涉及便于从不同资源搜索内容的系统和方法。

背景技术

10 数字形式的内容的广泛而迅速的增加提出了一种在各种环境中搜索信息的新机遇。除无所不在的因特网搜索引擎之外，也有其它常见的搜索工具可供用户使用。例如，电子邮件应用程序可向用户提供搜索出现在用户发送或接收的电子邮件的主标题或正文中的信息的工具。在文字处理应用程序中，用户可使用工具来搜索用户所创建或所访问的文档的标题或正文中的单词。其它类型的搜索可包括使用与图像
15 相关联的元数据的图像搜索。

学习如何使用这些不同的搜索工具可能令用户感到迷惑。首先，现有的搜索工具被设计成在假定用户已标识了要观察的特定资源—对许多用户此假定不一定准确—的情况下搜索特定的数据资源。对于现今的用户，知道要到何处寻找尤其困难，因为计算机越来越多地使用因特网或其它类型的网络链接到其它计算机。寻找
20 的可能位置的绝对量，以及一旦到达该位置所要搜索的数据量，令尽可能有效地搜索变得十分重要。确定要到何处寻找，以及挑选出哪些搜索结果是最相关的可以是令人生畏的工作。此外，对要寻找的最佳位置以及最相关结果的确定可能会迅速地改变，由此很难保持搜索既相关又有效。

现有搜索工具的另一问题是每一搜索工具具有不同的界面，用户必须学习这
25 些界面以找出他或她正在寻找的内容。例如，在 Windows XP™操作系统下的搜索（Search）命令调用“搜索助理（search companion）”来引导用户以完全不同于 Microsoft Word®文字处理程序下的搜索（Search）命令的方式进行搜索，后者用具有选项卡的搜索对话框来提示用户。

30 先前有各种努力来为用户简化搜索，但这些努力不足以满足用户的需求。例如，在因特网搜索引擎的环境中，允许用户一次键入搜索词然后同时调用若干不同

的搜索引擎的网站曾经是流行的。但是这类网站没有进一步帮助用户有效地搜索最相关的内容或从各种搜索中挑选出最相关的结果。实际上，当返回了比使用单个搜索引擎时更多的结果时，这类网站可能将用户的任务复杂化。

5 发明内容

为克服上述问题，提供了一种统一搜索的系统、方法和计算机可访问媒质。该系统和方法便于不同资源的统一搜索，并混合来自搜索的最相关结果，以用综合的方式向用户呈现。

依照本发明的一个方面，用户导航到一统一搜索入口界面并输入查询。该统一搜索过程捕捉该查询的导航上的下文，并确定在最相关的资源，在最相关的资源中，不仅基于查询的内容，即搜索词，还基于正引导到该查询的入口的导航的上下文中的一个或多个，即搜索入口点、用户偏好以及随时间推移所收集的潜在资源的相关度的任何自动测量来搜索。

用户偏好可由用户明确地预定义，或基于用户过去的行为来隐含地确定，并可包括一私密或安全偏好，用于指定应当不再搜索哪些资源或哪些资源是可信的，无论其确定的相关度如何。

相关度的自动测量包括使用数据，它反映用户如何在本地和/或非本地环境中与资源交互，并由此反映其与特定搜索查询的潜在相关度。资源包括包含可搜索内容以及内容本身的各种数据库和其它数据存储库。使用数据包括诸如访问资源的频率、访问资源的时间长度、用户在访问之后对资源作出的行动（如，打开、检索、浏览、编辑、发送、复制、保存、创建、购买交易等）的数据。本地使用数据表示本地用户与资源的交互，而非本地使用数据表示许多用户与资源的交互的集成。

依照本发明的另一方面，确定在其中进行搜索的最相关资源可包括确定何时将搜索联合到另外的资源，这些另外的资源一般不会被搜索，因为搜索它们太昂贵或者太耗时，这些另外的资源如可通过虚拟专用网络、企业网络、大数据库，或付费可用的专利数据库访问的远程计算机。在本发明的一个方面，确定是否将搜索联合到其它资源可由用户偏好来控制。

依照本发明的又一方面，一旦确定了最相关资源，该系统和方法使用归一化为搜索词在不同资源上的出现统一了搜索。归一化通过确定出现在例如网页标题中的搜索词相对于出现在电子邮件的主标题中的搜索词的相对重要性来确保统一的搜索将产生有意义的结果。确定出现在一个位置中的搜索词相对于另一位置的搜索

词的相对重要性包括首先确定在何处搜索每一相关资源，然后确定如何权衡搜索词在一个位置中的出现与其在其它位置中的出现。

依照本发明的再一方面，使用随时间推移构建的局部索引来执行搜索，以对
5 被确定为与用户搜索相关的资源进行索引。一旦被索引，对索引的资源任何改变
都将在索引中反映以用于未来的搜索。

依照本发明的还有一个方面，一旦执行了统一搜索，该系统和方法提供了混
合搜索结果并以综合的方式向用户呈现混合的结果。混合搜索结果包括，确定显示
来自哪些资源的哪些结果，以及以什么顺序显示等等。以综合的方式向用户呈现混
10 合结果包括，以易于访问的格式显示不同的结果等等，该易于访问的格式允许用户
快速地标识他或她搜索的特定数据，而无论结果所代表的数据类型如何。

依照本发明的再有一个方面，提供了用于统一搜索的计算机可访问媒质。该
计算机可访问媒质包括数据结构和计算机可执行组件，它们包括一统一搜索处理
器，它便于不同资源的统一搜索，并混合来自搜索的最相关结果，以综合的方式向
15 用户呈现。该数据结构以一般与上述方法一致的方式定义了资源、用户偏好、导航
的上下文数据、自动相关度数据以及索引数据。同样，该计算机可执行组件能够执
行一般与上述方法一致的行动。

附图说明

当结合附图参考以下详细描述时，可以更容易地理解本发明的上述方面和许
20 多附加优点，附图中：

图 1 是依照本发明可在其中执行统一搜索并呈现搜索结果的示例性统一搜索
系统和一个合适的操作环境的描述；

图 2 是依照本发明的一个实施例更详细描述图 1 的统一搜索系统的某些组件
的排列的框图；

25 图 3 是更详细描述用于实现本发明的一个实施例的搜索客户端的某些计算组
件的排列的框图；

图 4A-4B 所示是依照本发明的一个实施例结合图 1 和 2 的统一搜索系统所执
行的用于实现统一搜索并以综合的方式呈现搜索结果的逻辑的流程图；

图 5 是用于实现本发明的一个实施例的示例性统一搜索用户界面的图示；

30 图 6 是依照本发明的一个实施例，在混合了搜索结果并以综合的方式向用户
呈现之后的稍后时刻，图 5 的示例性统一搜索用户界面的图示；以及

图 7 是依照本发明的一个实施例，在选择了显示多媒体搜索结果的选项之后，图 6 的示例性统一搜索用户界面的图示。

具体实施方式

- 5 以下描述旨在提供适用于实现本发明的一个实施例的各个特征的计算系统的简要、通用描述。尽管将在个人计算机和服务器计算机，或可用于在其中辅助任务由通过通信网络链接在一起的远程计算装置执行的分布式计算环境的其它类型的计算装置的一般上下文中描述本计算系统，本领域的技术人员可以理解，本发明也可以用许多其它计算机系统配置来实践，包括多处理器系统、小型机、大型机等等。
- 10 除上述更常规的计算机系统之外，本领域的技术人员将认识到，本发明可在其它计算设备上实践，包括膝上计算机、图形输入板计算机、个人数字助理（PDA）、蜂窝电话和在其上安装了计算机软件或其它数字内容的其它设备。

尽管本发明的各方面可以借助由 web 浏览器结合个人计算机执行的程序或进程，或由搜索引擎结合服务器计算机执行的程序或进程来描述，本领域的技术人员将认识到，这些方面也可以结合其它程序模块来实现。一般而言，程序模块包括例

15 程、子例程、程序、进程、组件、数据结构、函数、接口、对象等等，它们执行特定的任务或实现特定的抽象数据类型。

图 1 是依照本发明的一个实施例，可在其中执行统一搜索，并混合搜索结果的示例性统一搜索系统 100 和一个合适的操作环境的描述。如图所示，操作环境包

20 括驻留在诸如计算机系统 102 或个人数字助理（PDA）104 等用户设备中的统一搜索客户端 110。该统一搜索客户端 110 一般负责如所需要的提供前端搜索服务，并需要时与远程服务器进行通信。由统一搜索客户端 110 提供的前端服务可包括，生成接收用户在统一搜索入口界面 106 中输入的搜索词查询 108，并响应于源自运行在用户设备 102、104 上的应用程序的信息和搜索查询，在混合的搜索结果页面 124

25 中使用超文本传输协议呈现搜索结果 123 的文本和/或图形，以及其它服务。由统一搜索客户端 110 提供的后端搜索服务可包括，在不同的资源上执行统一搜索、混合搜索结果以在用户设备 102、104 上呈现、捕捉自动的相关度数据 118、更新本地搜索索引 116，以及其它服务。

在图 1 所示的环境中，统一搜索客户端 110 响应于用户输入的搜索请求生成

30 一统一搜索入口界面 106，用户可向其中输入搜索词 108 以启动在本地驻留在用户设备上的本地数据资源 120 或非本地数据资源 122 上对相关内容的搜索。在一个实

5 施例中, 统一搜索入口界面 106 可被实现为嵌入在诸如浏览器应用程序、文字处理程序等应用程序中的搜索框, 或诸如 Windows 等图形用户界面中的任务栏。在另一实施例中, 统一搜索入口界面 106 可被排列成单独的应用程序窗口或浏览器应用程序中的页面。如图 2 所示, 本地资源 120 可包括储存在用户设备硬盘驱动器 202 上的数据, 包括电子邮件 204、web 高速缓存 206、照片和图像 208、音乐、以及诸如文字处理文档、电子表格等办公自动文档 212 等等。非本地资源可包括远程地驻留在通过诸如因特网 214、内联网 216、在因特网 214 或内联网 216 上配置的虚拟专用网络以及局域网 218 与统一搜索客户端 110 通信的设备上的资源。

10 在一个实施例中, 统一搜索客户端 110 捕捉搜索请求的导航上下文 112, 即, 用户请求搜索的所在点, 并使用该上下文 112 结合搜索词 108 以及先前储存的用户偏好 114 来确定在其中搜索与搜索词 108 相关的内容的最佳资源 120、122。例如, 当用户从电子邮件应用程序进入统一搜索入口界面 106 时, 统一搜索客户端 110 可确定包含电子邮件文档或作为电子邮件的附件接收的文档的本地资源 120 可能是在其中执行搜索的最相关的资源。

15 在一个实施例中, 当确定哪一资源 120、122 是在其中执行搜索的最相关资源时, 统一搜索客户端 110 也考虑表示相关度的自动测量 118 的任何可用数据。在一个实施例中, 自动相关度数据 118 是量化用户和所述资源之间的交互的度量, 包括用户访问资源的频率、交互 (如打开或激活文件、浏览网页、编辑文档、发送电子邮件、向电子邮件附加文档、移动文档等等) 的时间长度和/或次数和重要性。统一搜索客户端 110 使用数据 118 来测量所考虑的特定资源的潜在相关度。例如, 访问的频率越高, 持续时间越长, 和/或交互的次数和/或重要性越大, 则该资源以及该资源包含的内容将与用户搜索相关的可能性越大。尽管确定相关度的常规方法搜索资源内容 (或表示内容的索引) 来确定搜索词 108 出现的频率, 然而自动相关度数据 118 是使用数据, 它可随时间的推移来捕捉, 以在搜索资源内容之前向统一搜索客户端 110 提供指示 (尽管是间接的) 资源是否可能相关的可计量信息。当与确定相关度的更多常规方法组合时, 统一搜索客户端 110 能够使用自动相关度数据 118 来有利地提高搜索的效率, 因为仅搜索可能包含相关内容的那些资源。

30 在一个实施例中, 自动相关度数据 118 可以是用户专用的或集成的, 以反映资源与特定用户组或与一般用户的相关度。例如, 数据可指示特定数据库与企业的用户的一个部门内的用户的相关度, 或网页与因特网用户的相关度。用户专用自动相关度数据 118A 一般是由本地数据资源 120 的搜索客户端 112 随时间推移捕捉的本地数

据，而集成的相关度数据 118B 的一般是由非本地数据资源 122 的搜索服务器随时间推移捕捉的非本地数据。自动相关度数据 118 在需要时被发送到统一搜索客户端 110 或由其访问，以确定资源 120、122 的潜在相关度。

一旦标识了相关资源，统一搜索客户端 110 使用索引 116 启动搜索以产生搜索结果 123。索引 116 可以是用于便于快速搜索的任何类型的可更新搜索的索引。索引 116 的合适的实现是已知或可购买的，并且尤其是根据本发明的说明，可由本领域的普通技术人员容易地实现。在一个实施例中，索引 116 可使用作为软插件提供的过滤器来索引某些类型的第三方文档，如可移植文档文件（PDF）。

在一个实施例中，统一搜索客户端 110 使用归一化来考虑资源 120、122 的全异特性。归一化包括确定在一个资源的一个位置中相对于在另一资源的另一位置中寻找搜索词（称为“击中”）的相对重要性，这通过首先标识资源 120、122 中搜索哪一位置，然后依照相对重要性对一个位置的结果给予比另一位置的结果更大或更小的权值来完成。位置的相对重要性随着每次搜索而改变，这取决于任意数目的因素，包括导航上下文 112、用户偏好 114 或搜索词查询本身 108 等等。例如，当导航上下文数据 112 指示用户从电子邮件应用程序进入搜索时，搜索词击中电子邮件的主标题的相对重要性高于击中多媒体文件的标题。

在一个实施例中，一旦完成了搜索，统一搜索客户端 110 将搜索结果 120 混合成在用户设备 102、104 上显示的混合搜索结果页面 124。由于搜索结果 123 可在不同的资源 120、122 中找到，结果本身可表示不同类型的数据。例如，如参考图 5-7 所示的示例搜索进一步示出的，搜索词“Madonna”的搜索可获得在新闻和歌迷网站、音乐文件和电子邮件中找到的各种内容。统一搜索客户端 110 确定这些不同结果中的哪些与用户最相关，并相应地显示它们。同时，统一搜索客户端 110 在综合的显示中混合结果，以便于使用图形、定位或便于用户从结果中选择的任何其它显示格式化技术来从其它结果中识别一个结果。例如，在一个实施例中，可将电子邮件可网站组合到分别的位置中，但是电子邮件的主标题可类似于新闻网站的标题行、音乐文件的标题以及歌迷网站的 web 地址来定位和格式化，以允许用户容易地比较各种结果并相应地选择期望的结果。

参考图 1 和 2，用户设备 102、104 通过诸如因特网 214 等一个或多个计算机网络与远程设备或资源通信。用于通过因特网通信的协议和组件对计算机网络通信领域的普通技术人员是众所周知的。用户设备 102、104 之间的通信也可通过本地有线或无线计算机网络连接来实现。图 1 描述的统一搜索客户端 110 也可在作为分

布式计算环境的一部分的用户设备 102、104 中操作，该环境可包括通过通信链路，如使用一个或多个计算机网络或直接连接互联的若干计算机系统。然而，本领域的普通技术人员可以理解，统一搜索客户端 110 同样可在具有比图 1 所示的更少或更多数量的组件的计算机系统中操作。由此，图 1 中的操作环境的描述应当看作示例性的，而非对所附权利要求书的范围的局限。

图 3 是更详细描述主存 (host) 统一搜索客户端 110 并负责图 1 所示的统一搜索系统 110 的操作的用户设备 102、104 的某些示例性计算组件的排列的框图。具体地，示出了用户设备 102、104，包括操作系统 302、处理器 306 和存储器 308，以实现用于该客户端的通用管理和操作的可执行程序指令。用户设备 102、104 还包括与诸如因特网 214 等网络进行通信的网络接口 304，以响应搜索词查询 108 并提供搜索结果 123。对操作系统 302、处理器 306、存储器 308 和网络接口 304 的合适的实现是已知的或可购买的，并且尤其鉴于本发明的说明，可由本领域的普通技术人员容易地实现。

用户设备 102、104 的存储器 308 包括包含统一搜索客户端 110 的计算机可执行程序指令。在某些实施例中，存储器 206 还可包括诸如上述搜索词查询 108 和搜索结果 123 等各种储存的数据，以及如图 1 和 2 中所引用的用户界面 114 和导航/上下文数据 112。

在一个实施例中，统一搜索客户端 110 包括资源相关进程 310、索引进程 312、搜索统一进程 314 和结果混合进程 316。索引进程 312 可以使用已知的或可购买的索引技术来实现。每当执行搜索或出现索引的资源的改变时，就调用索引进程 312。

在一个实施例中，资源相关进程 310 包含那样的逻辑，当执行该逻辑时，它将监控和捕捉关于资源 120、122 的用户活动，并以度量的形式将该活动记录在自动相关度数据 118 中。例如，进程 310 监控硬盘驱动器上的特定文档，并记录用户访问该文档的频率和持续时间。另外，进程 310 监控并捕捉正引导到响应于搜索词查询 108 显示统一搜索入口界面 106 的点的导航上下文数据 112。资源相关进程 310 还包含逻辑，它不仅基于迄今为止捕捉的搜索词查询 108 和自动相关度数据 118，还基于导航/上下文数据 112 和用户偏好 114 来确定要搜索的最佳资源。

在一个实施例中，搜索统一进程 314 包含逻辑，当执行该逻辑时，它确定在每一资源内使用一被称为归一化的进程被确定为与搜索相关的最佳数据。例如，如正文数据相比，电子邮件存储库内最佳的搜索数据可以是主标题数据，而在 Word 文档内最佳搜索数据可以是该文档的标题，或者如果没有可辨别的标题，则为文件

名。归一化提高了搜索结果的相关度，并确保可有意义地比较结果以确定哪一个是最相关的。

5 在一个实施例中，结果混合进程 316 包含逻辑，当执行该逻辑时，它混合搜索结果，当一个结果来自一种类型的资源 120、122，而另一结果来自一个不同类型的资源，比如一个来自电子邮件另一个来自电子表格文档时，考虑结果的全异特性。进程 316 还以便于用户选择期望结果的综合方式向用户呈现混合结果。例如，进程 316 可组合第一网页的每一资源内的最相关结果的表示，使得用户可一眼见到组合结果的所有类别上的所有最佳结果。作为另一示例，进程 316 可交错最相关的结果，并用适当的图标或其它图形指定结果所属的类别。无论用哪一方式，进程 10 316 还混合结果，并通过在所有类别上格式化结果来以综合的方式显示它们，以便于将一个类别或组的击中与另一个进行比较。例如，电子邮件类别中的结果可以与网站类别中的结果垂直对齐，使得电子邮件主标题和网站标题在每一类别内的可比较位置上。

图 4A-4B 所示是依照本发明的一个实施例结合图 1 和 2 的统一搜索客户端 110 15 执行的用于执行统一和混合搜索的逻辑的流程图。在开始椭圆 402 处开始，统一搜索客户端 110 接收起源于运行在主机设备 102、104 上的应用程序的搜索请求 404 的通知。在一个实施例中，请求 404 可以在处理块 406 由客户端 110 截取，其中，客户端 110 向用户启动统一搜索进入图形用户界面 106 的显示，例如，以浏览器程序中显示的网页的形式来显示。统一搜索入口界面提示用户输入他或她的搜索词查询 108。同时，客户端 110 通过对正引导到搜索请求 404 的导航，如，用户在他或她请求搜索的点所使用的应用程序进行排序，来捕捉它将要执行的搜索的上下文。另外，客户端 110 捕捉由用户在统一搜索入口界面中直接输入的，添加到或覆盖先前输入的搜索偏好的任何用户偏好。这类偏好可包括，不被搜索的资源 120、122 20 的指定，无论其与当前搜索词查询 108 的相关度如何；或者相反，要被搜索的资源的指定，即使其确定的与查询的相关度看来是有疑问的；等等。其它偏好可包括私密或安全性偏好，以防止记录或保留搜索私密性。在一个实施例中，用户的搜索偏好根据用户在他或她的用户设备上使用该搜索特征时的过去的行为来隐含地确定。

在处理块 404，客户端 110 从本地源 118A 或非本地源 118B 或两者获取自动 25 相关度数据 118。自动相关度数据 118 是在统一搜索客户端 110 的先前的迭代中捕捉的数据，并反映了用户与各种资源 120、122 和先前的搜索结果 123 的交互的历史。在处理块 412，客户端 110 调用如已参考图 3 所描述的资源相关进程 310，来 30

使用可用的自动相关度数据 118，以确定在其中搜索当前搜索词查询 108 的最佳资源 120、122。

在处理块 414，客户端 110 继续处理，以确定每一资源 120、122 内的最佳数据。客户端 110 调用如已参考图 3 所描述的搜索统一进程 314，来归一化相关资源
5 作为要搜索的最佳数据的确定的一部分，并相应地执行对最佳数据中搜索词 108 的出现的搜索。例如，在已提及的示例中，如果确定在其中搜索词“Madonna”的最佳资源包括电子邮件资源 204 和音乐资源 210，则在归一化这些资源之后，电子邮件资源中搜索“Madonna”的最佳数据可以是主标题，而音乐文件中的最佳数据可以是标题或文件名。

10 参考图 4B，客户端 110 在椭圆 416 处继续，在处理块 418 对照如上确定的最佳资源中的最佳数据执行搜索。在一个实施例中，为最大化搜索效率并最小化对用户设备 102、104 的操作的影响，对照可搜索资源 120、122 的索引 116 来执行搜索。

客户端 110 在处理块 420 生成搜索结果，并前进到处理块 422，通过调用参考图 3 所描述的结果混合进程 314 来混合那些结果。混合结果的进程 314 包括在给定
15 搜索结果网页上可用的有限显示区域中确定显示哪些资源的哪些结果、并格式化结果以用综合的方式向用户呈现，以便于用户选择期望的结果。在处理块 424，客户端向用户显示混合并格式化的结果。混合搜索结果的一个示例在下文参考图 5-7 更详细地描述。

在一个实施例中，一旦向用户显示了混合的搜索结果，则在处理块 426，客户端 110 调用资源相关进程 310 在准备记录本地的自动相关数据库 118A 中监控并捕捉用户与呈现的混合搜索结果的交互，使得客户端 110 能够在进行未来搜索时进一步
20 细化资源相关度的确定，之后，客户端 110 的逻辑在终端椭圆 428 处结束。

图 5-7 示出了浏览器程序 500，它显示了如先前参考图 1 所描述的在网页中实现的统一搜索入口界面 106，其中，用户能够将一个或多个搜索词输入到搜索输入
25 框 502 中。统一搜索入口界面 106 可由统一搜索客户端 110 生成，或可远程地生成并通过因特网传送到用户计算设备 102、104 上的客户端 110。用户在框 502 中输入期望的搜索词，并使用定位设备、键盘或其它输入设备来启动搜索，如通过点击标号 502 标识的箭头按钮来启动。在一个实施例中，可向用户呈现若干搜索偏好 506，用于选择来定制他或她的搜索。例如，可通过点击“My Stuff（我的东西）”
30 文件夹 504 将搜索限于某些本地资源。也可提示用户选择将导致搜索被向外联合到其它在线资源 508 的选项，如多媒体资源（Multimedia）510、因特网（The web）

512、字典 (Dictionary & translation) 514、购物资源 (Shopping) 516、航班资源 (Flights) 518、地图资源 (Maps & directions) 520、百科全书资源 (Encyclopedia) 522、餐馆用餐资源 (Restaurants) 524 以及新闻资源 (News) 526。某些选项可默认地选择，或可基于先前指定的用户搜索偏好 506 来选择。

5 在图 5 所示的示例中，用户输入了搜索词“MADONNA”。向用户显示如依照本发明的一个实施例由客户端 110 实现的统一和混合搜索的结果，如图 6 所示。搜索词“MADONNA”在标号 602 处的文本框中显示。在标号 604 处的下拉框中选择了显示所有结果的选项 (All Results)。如图所示，该显示在页面上列出了五个不同的类别，而在后续的页面上列出更多，因为选择了标号 606 处的其它选项工具条中的选项来显示“MADONNA”的“All Results (所有结果)”。未选中的选项包括标号 618 处的“Grouped by Content Type (按照内容类型分组)”选项、标号 620 处的“Web sites (网站)”选项、标号 622 处的“Multimedia (多媒体)”选项、标号 624 处的“Shopping (购物)”选项、标号 626 处的“News (新闻)”选项、以及标号 628 处的“Learning (学习)”选项。在所示的示例中，备选搜索选项的附加界面选项在标号 630、632、634 和 636 处出现。

在第一页面上显示了用于选中的“All Results (所有结果)”选项的前五个分类标题，即标号 606 处的新闻网站、标号 608 处的电子邮件、标号 610 处的音乐、标号 612 处的歌迷网站、以及标号 614 处的图像，其每一个在该分类标题的下面列出了最佳结果。例如，在标号 606 处的新闻网站分类中，最佳结果是关于重组合 Madonna 独唱的新闻项目 606A，而在标号 608 处的电子邮件分类中，最佳结果是查询 Madonna 视频的电子邮件 608A。类似地，在歌迷站点 612 和图像 614 的分类中，其每一个在紧靠分类标题之下都显示了最佳结果。在每一情况中，尽可能一致地显示来自结果的归一化数据—在这一实例中，在最左边的位置—使得每一分类中的结果都可比较地对齐 (除标号 614 处的 web 图像分类之外，其中，图像本身是水平显示的)。最佳结果的可比较对齐结合按分类标题隔离结果有利地方便了用户对期望结果的选择。

在一个替换的显示中，向用户显示了依照本发明的一个实施例由客户端 110 实现的统一和混合搜索的结果，如图 7 所示。再一次，在标号 702 处的文本框中显示搜索词“MADONNA”，并且搜索被限制在“MADONNA”在因特网上的出现上，如标号 704 的文本框中所指示的。如图所示，该替换的显示列表 342 产生了艺人、音乐、电影、无线电和事件分类，因为选中了标号 722 处的其它选项工具条中

- 的选项，以仅显示“MADONNA”的多媒体结果。未选中的选项包括标号 716 处的“**All Results (所有结果)**”选项、标号 718 处的“**Grouped by Content Type (按内容类型分组)**”选项、标号 720 处的“**Web sites (网站)**”选项、标号 724 处的“**Shopping (购物)**”选项、标号 726 处的“**News (新闻)**”选项、以及标号 728
- 5 处的“**Learning (学习)**”选项。改变显示或改变搜索的附加选项在页面顶端标号 706 处示出。三个分类标题—标号 710 处的艺人网站、标号 712 处的音乐和标号 714 处的电影—类似于图 6 中的分类来显示，其最佳结果在该分类的标题下列出。例如，在标号 710 处的艺人网站分类中，最佳结果是关于 Madonna 的 Madonna 艺人页面，而在标号 712 处的音乐分类中，最佳结果是“**American Life**”的音乐样本—Madonna
- 10 的最新歌曲，并且在标号 714 处的电影分类中，最佳结果是电影“**Die Another Day**”的电影样本。如上所述，所显示的分类 710、712 和 714 的每一个紧靠分类标题的下面显示了最佳结果。在每一情况下，尽可能一致地显示来自结果的归一化数据—在这一实例中，在顶部和最左边的位置中—使得每一分类中的结果都是可比较地对齐的。
- 15 尽管示出并描述了本发明当前的较佳实施例，然而可以理解，可在不脱离本发明的精神和范围的情况下在其中作出各种改变。例如，在本发明的一个实施例中，参考资源相关进程 310、索引进程 312、搜索统一进程 314 和结果混合进程 316 所描述的统一搜索系统 100 的各种组件，尤其是统一搜索客户端 110 的功能，可以进程或程序的不同组合来实现，并可在一个或多个计算设备之间分布。
- 20 尽管示出并描述了本发明的较佳实施例，可以理解，可在不脱离本发明的精神和范围的情况下在其中作出各种改变。

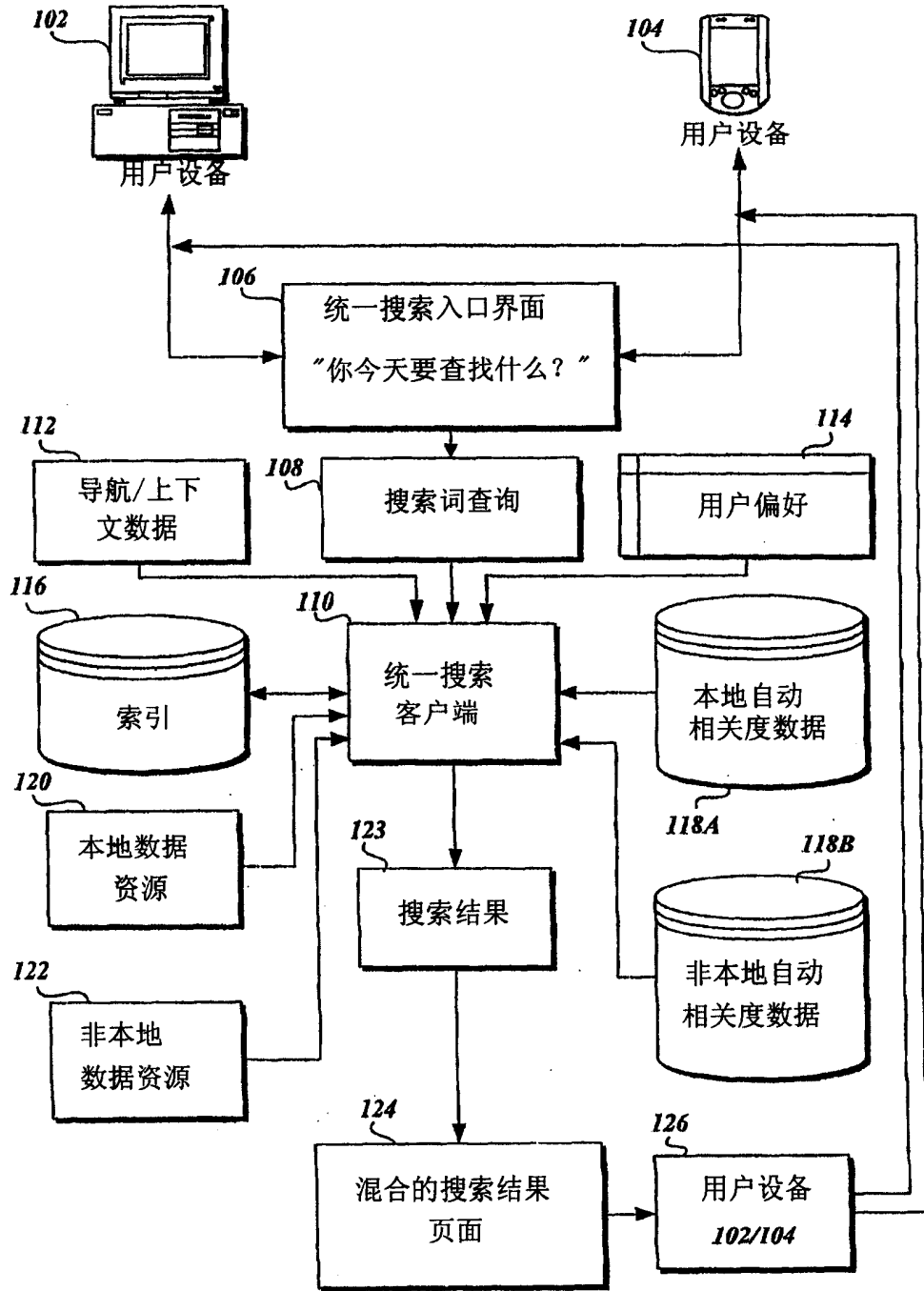


图 1

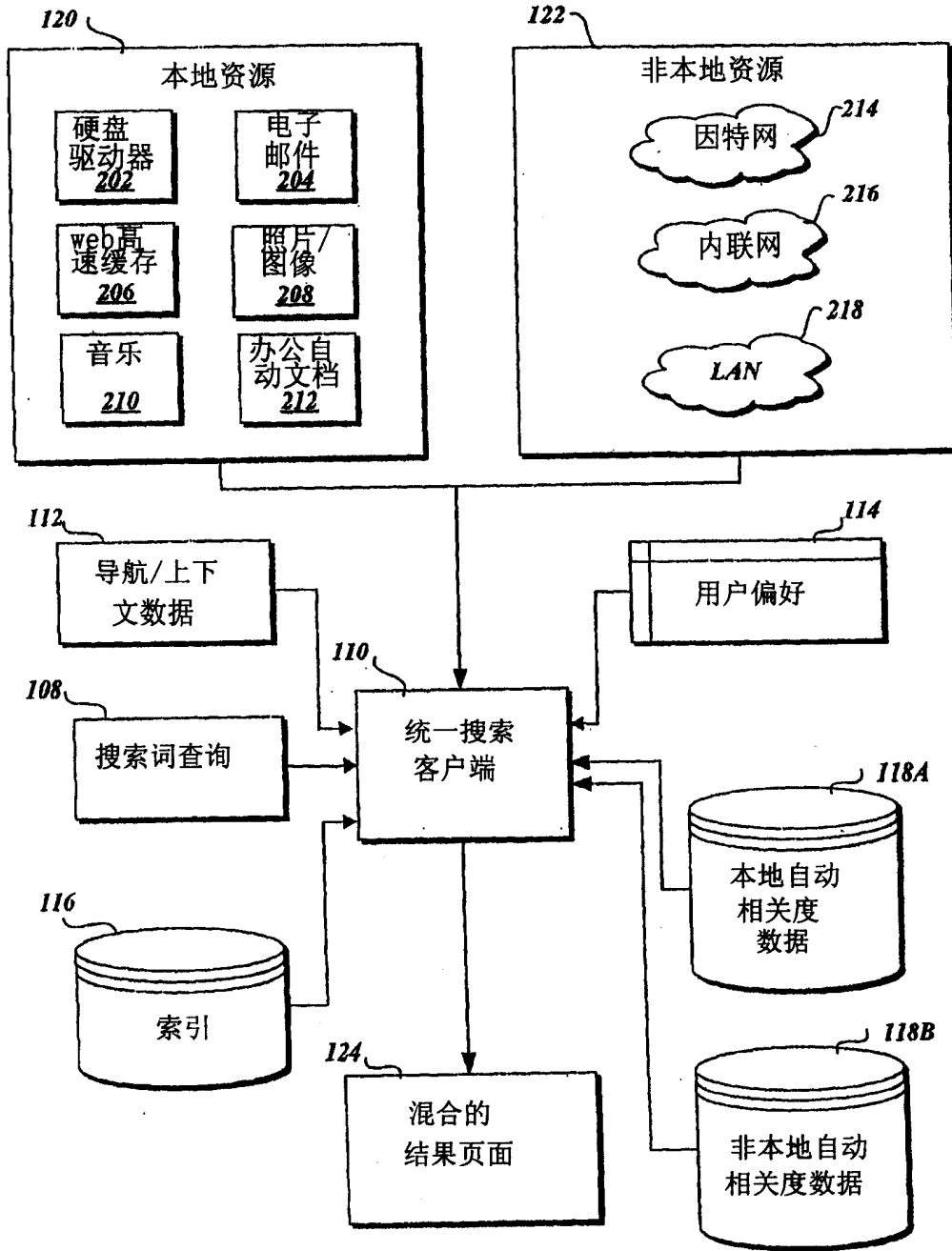


图 2

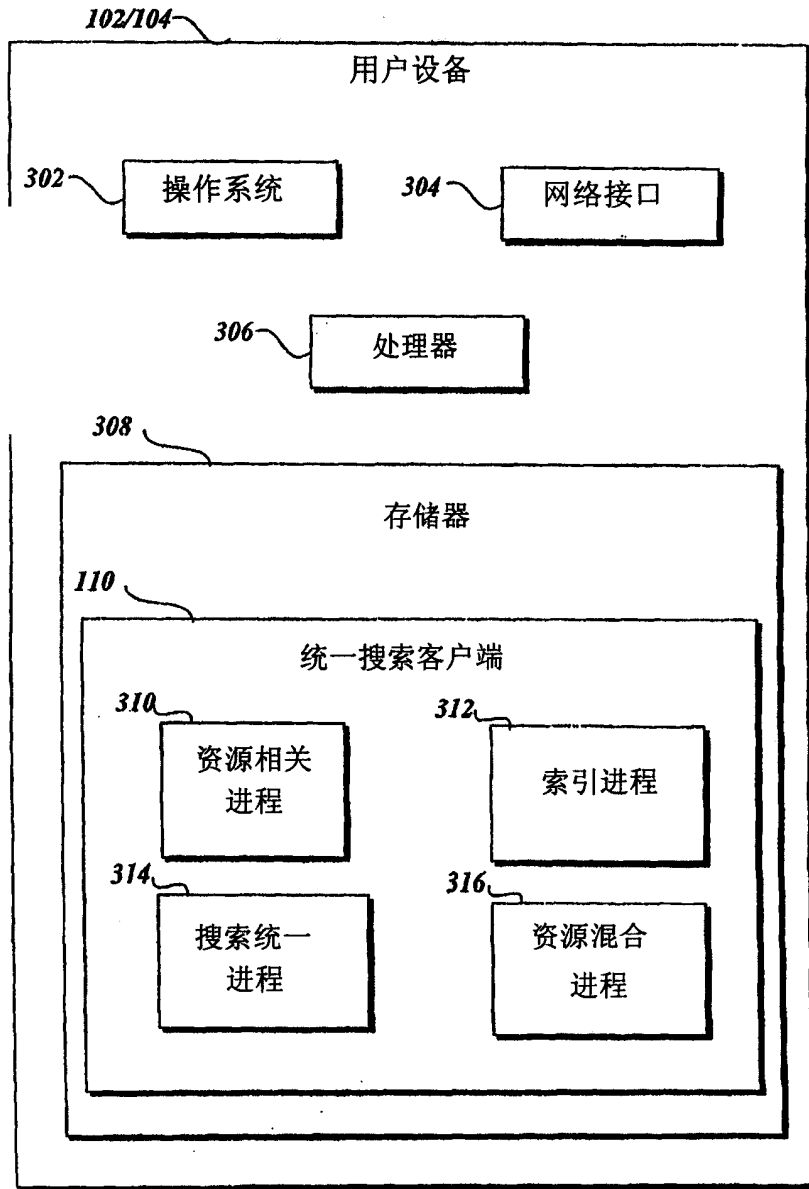


图 3

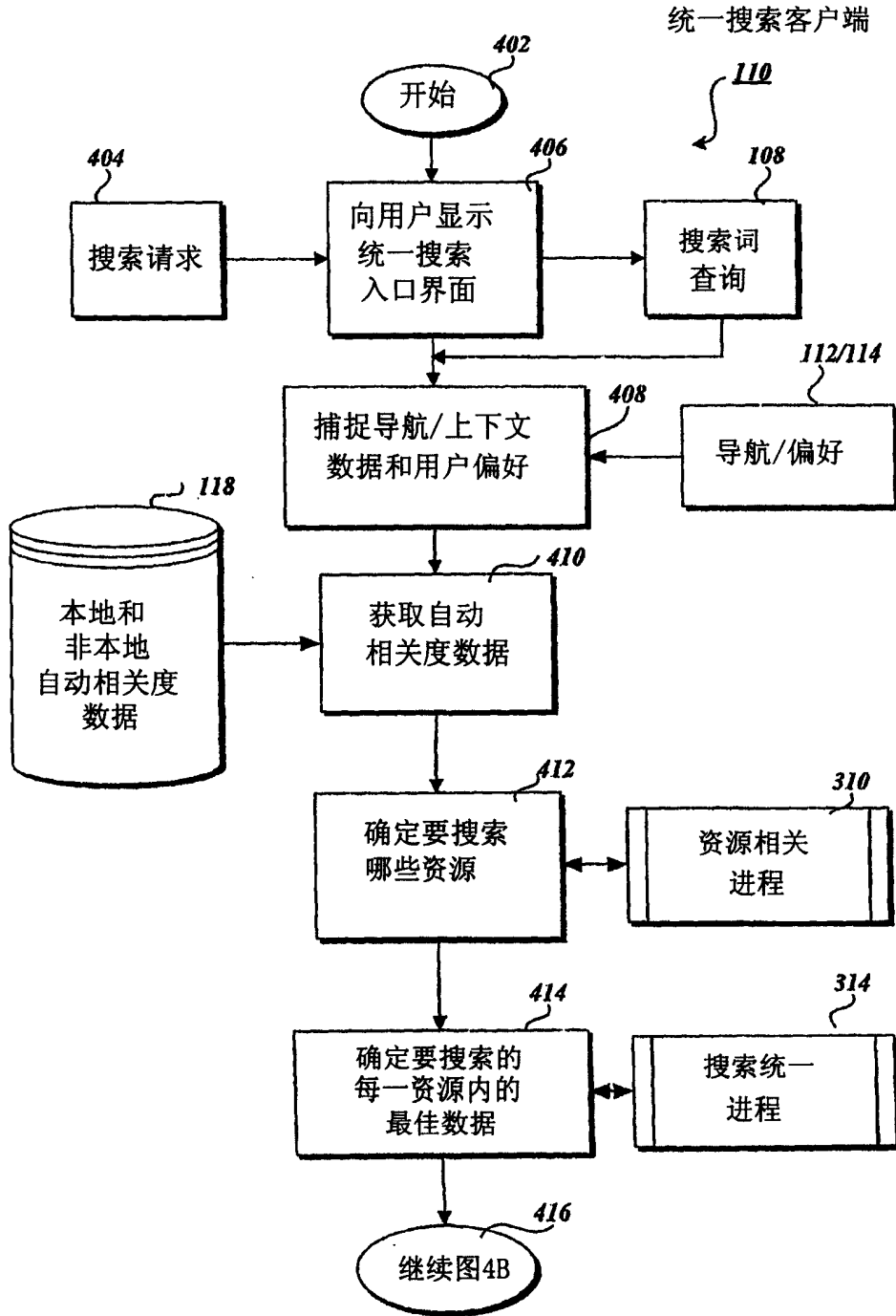


图 4A

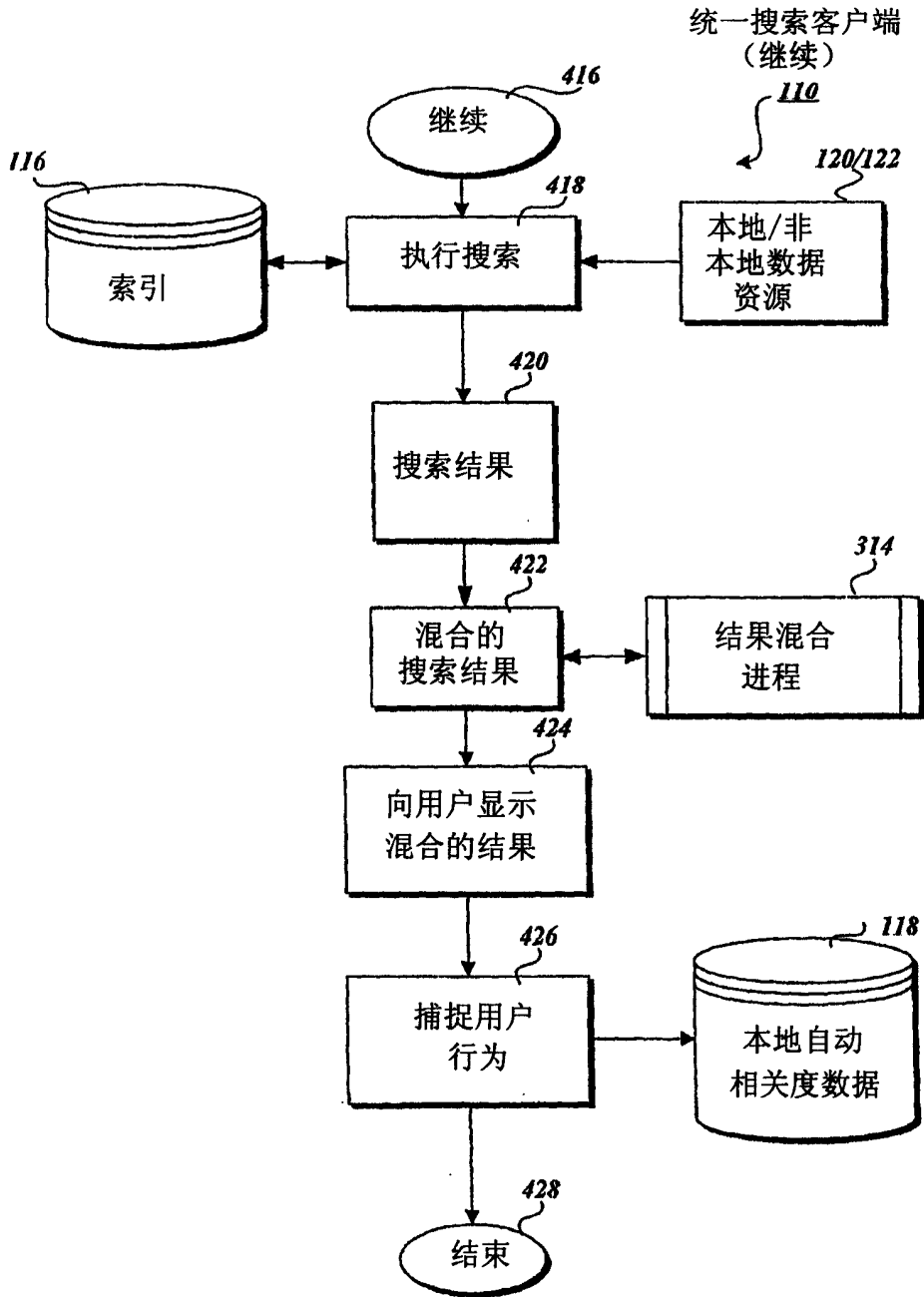


图 4B

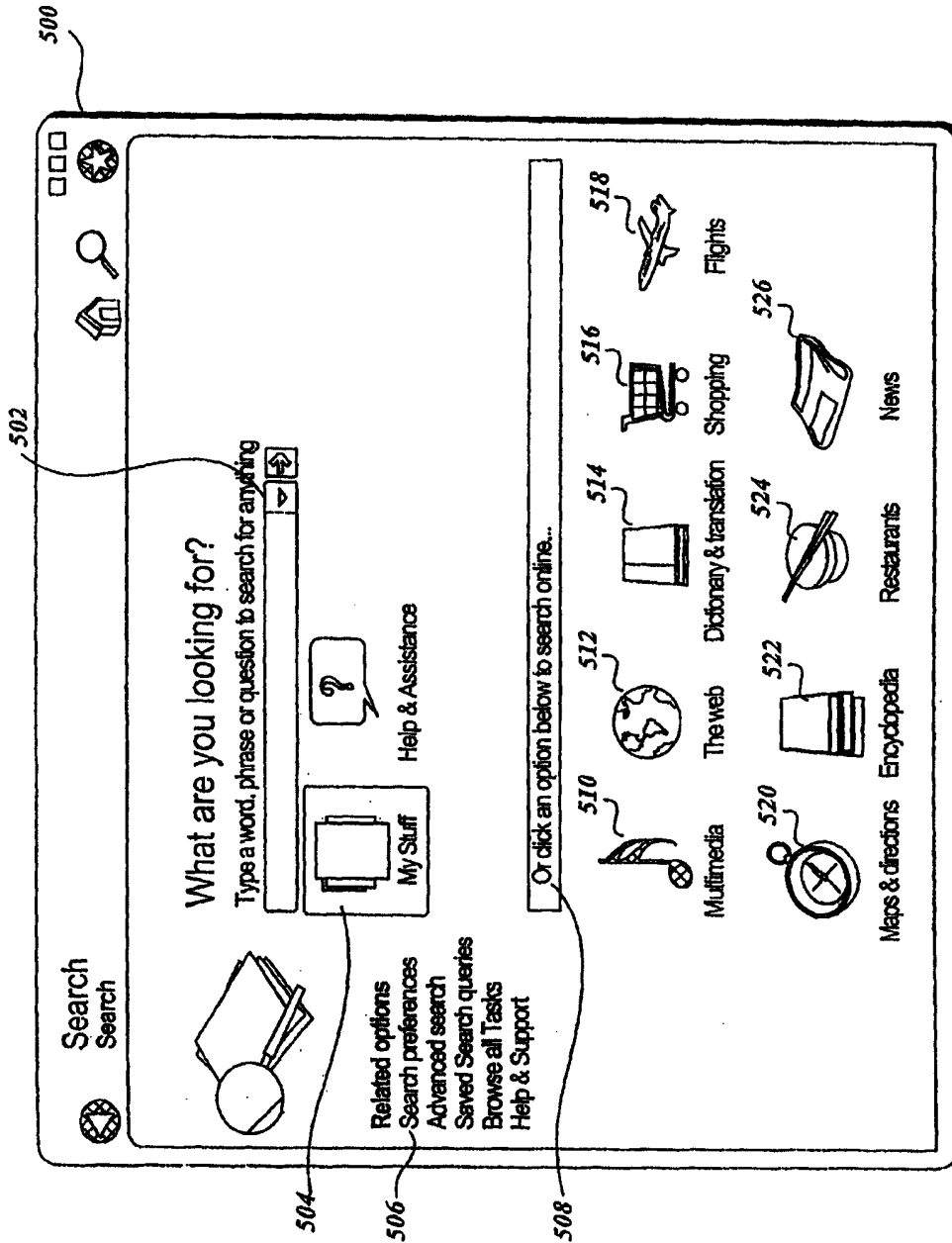


图 5

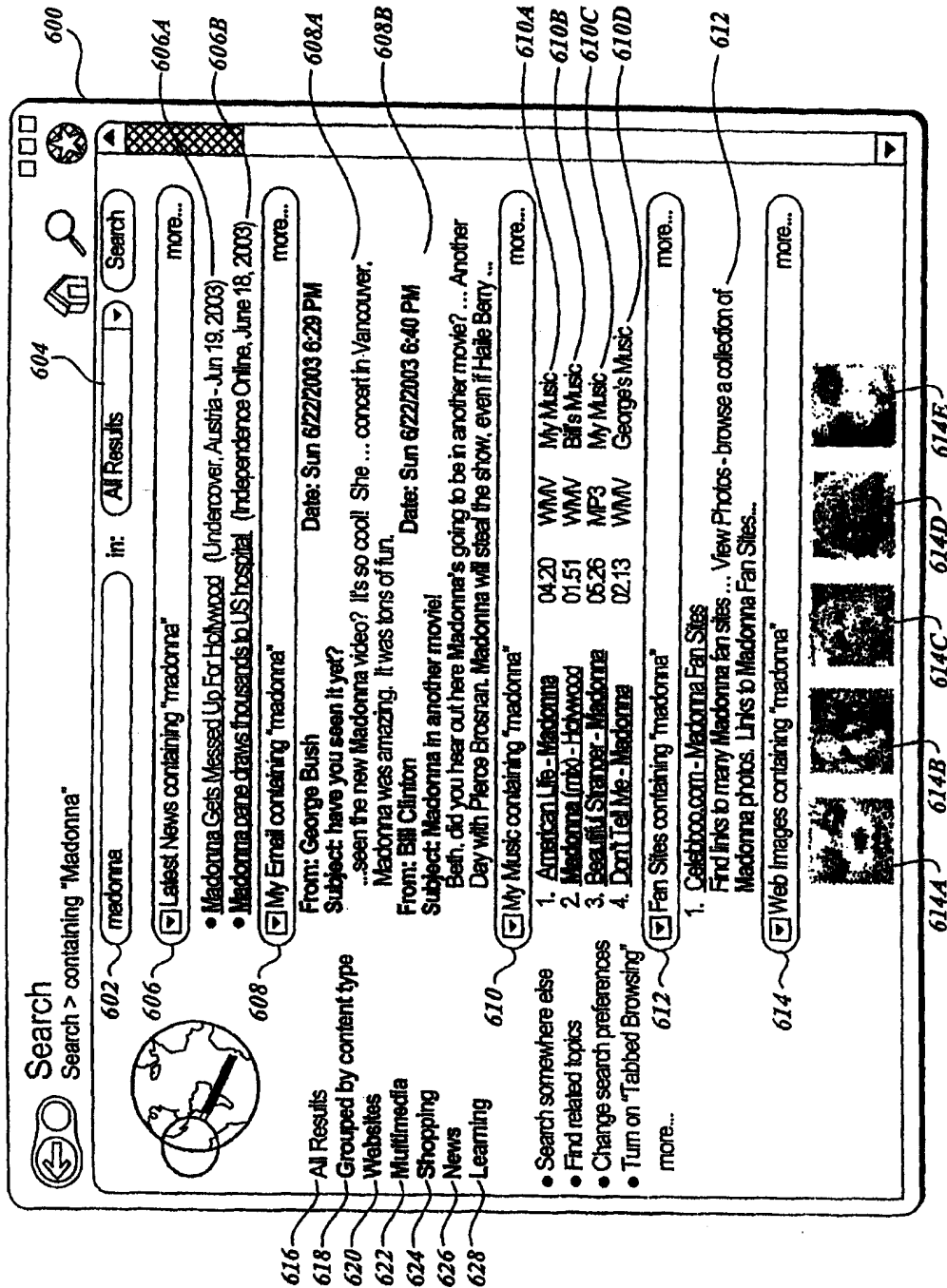
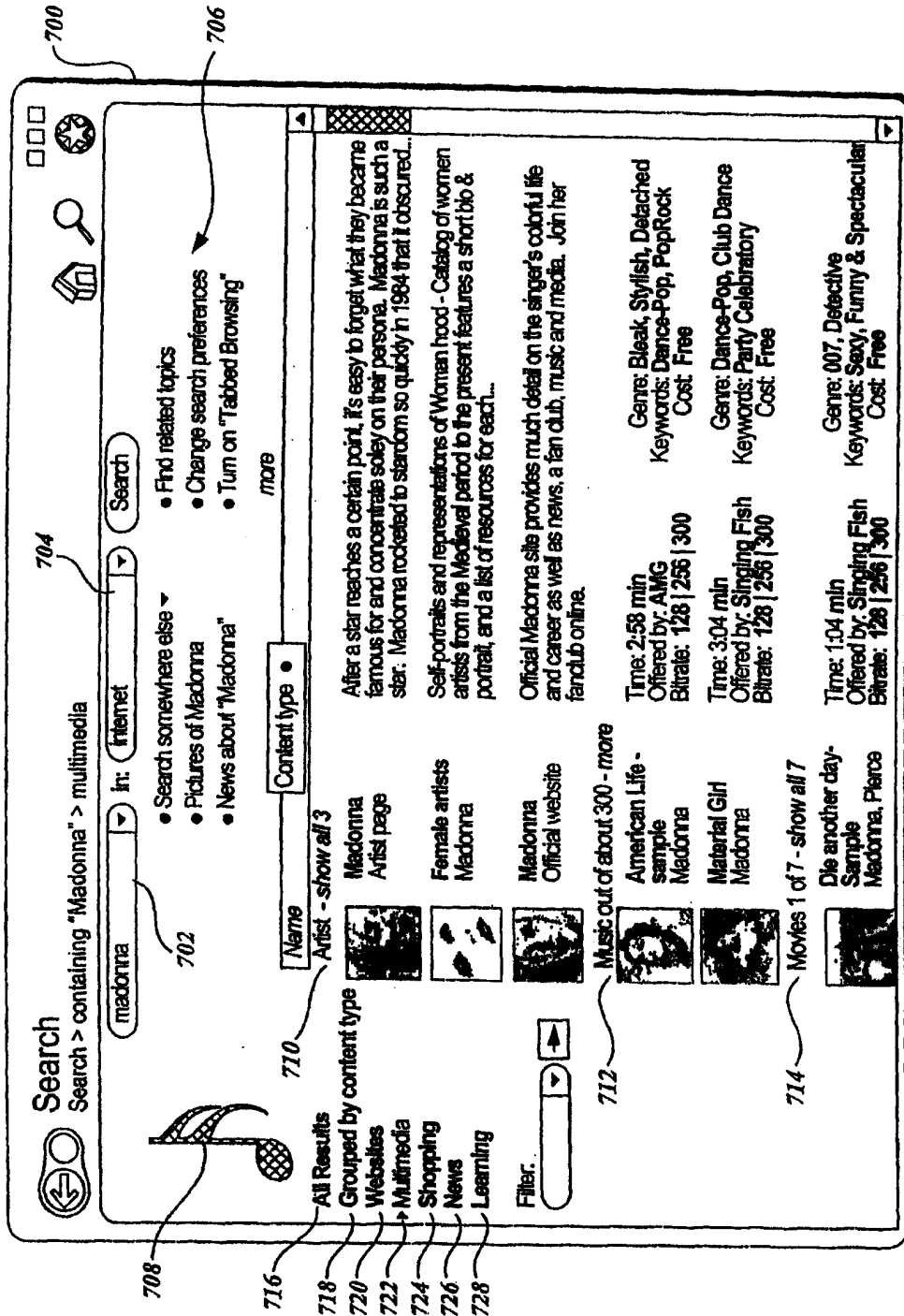


图 6



7