

(12) 发明专利

(10) 授权公告号 CN 101336409 B

(45) 授权公告日 2012. 01. 04

(21) 申请号 200680052288. 0

(51) Int. Cl.

(22) 申请日 2006. 12. 08

G06F 7/00(2006. 01)

(30) 优先权数据

(56) 对比文件

60/749, 013 2005. 12. 08 US

US 2005/0149561 A1, 2005. 07. 07, 全文.

US 6701307 B2, 2004. 03. 02, 说明书第 1 栏

(85) PCT 申请进入国家阶段日

2008. 08. 04

第 5 行至第 4 栏第 55 行, 第 11 栏第 30 行至第 59 行、附图 1, 4-5.

(86) PCT 申请的申请数据

PCT/US2006/047010 2006. 12. 08

US 2005/0116966 A1, 2005. 06. 02, 全文.

CN 1275749 A, 2000. 12. 06, 全文.

(87) PCT 申请的公布数据

W02007/067760 EN 2007. 06. 14

审查员 陈佳

(73) 专利权人 德卡尔塔公司

地址 美国加利福尼亚州

(72) 发明人 杰弗里·亨德雷 理查德·波彭

(74) 专利代理机构 北京律盟知识产权代理有限

责任公司 11287

代理人 孟锐

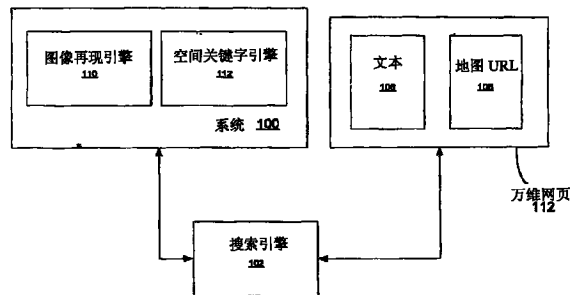
权利要求书 1 页 说明书 7 页 附图 3 页

(54) 发明名称

高精度度因特网局部搜索

(57) 摘要

在因特网上执行高精度度局部搜索。地图图像再现软件供应商将空间关键字嵌入地图中, 然后将所述地图提供给因特网内容的制作者 (例如地图供应商)。例如, 房主可在万维网布告板上贴出消息, 所述消息为其出售房屋做广告且包含显示所述房屋的位置的地图。当搜索引擎的万维网浏览器遇到具有嵌入到图像中的空间关键字的页时, 为所述空间关键字与所述页上的其它内容加索引。由于所述空间关键字识别较小的地理区域, 因此为所述内容与所述空间关键字加索引将允许对搜索查询进行区域限制但仍提供有用结果。因此, 搜索特定区域中“房屋出售”的搜索引擎用户将被引导到满足地理和内容搜索词的万维网页。



1. 一种为万维网页加索引的方法,所述方法包括:
检索万维网页,所述万维网页包含第一内容和指定地图内容的源地址;
在所述万维网页中指定的所述源地址处检索地图内容,所述地图内容包括经嵌入空间关键字;
从所述所检索的地图内容中提取所述经嵌入空间关键字;及
存储所述所提取的空间关键字与所述第一内容之间的关联性。
2. 如权利要求 1 所述的方法,其进一步包括:
接收搜索查询,所述搜索查询包含查询词和位置;
返回所述第一内容和所述所提取的空间关键字。
3. 一种用于为万维网页加索引的设备,其包括:
检索万维网页的装置,所述万维网页包含第一内容和指定地图内容的源地址;
在所述万维网页中指定的所述源地址处检索地图内容的装置,所述地图内容包括经嵌入空间关键字;
从所述所检索的地图内容中提取所述经嵌入空间关键字的装置;及
存储所述所提取的空间关键字与所述第一内容之间的关联性的装置。

高精度度因特网局部搜索

[0001] 相关申请交叉参考案

[0002] 本申请案主张 2005 年 12 月 8 日提出申请的申请案第 60/749,013 号的权利,且所述申请案的全文以引用的方式并入本文中。

技术领域

[0003] 本发明通常涉及搜索因特网。更具体而言,本发明涉及通过将可搜索的空间关键字插入地图图像中来改进在万维网上的搜索。

背景技术

[0004] 尽管因特网上的搜索引擎用于定位来自全世界的内容,但其对于进行局部搜索并不是特别有用。例如,对于搜索查询“CA 旧金山市场街 1 号 3 英里内房屋出售”,搜索引擎不会产生很好的结果。某些搜索引擎提供商已开始维持包含局部商业位置或其它关注点的局部数据库,因此实现例如“94114 附近的星巴克”的搜索。然而,此局部搜索对于搜索引擎来说是昂贵的,因为空间加索引和搜索需要搜索提供商将新技术整合到其索引编制器和搜索算法二者中。为了使用传统搜索引擎技术来执行精确局部搜索,作为索引建立过程的一部分必须从网站剪切地址并将其地理编码为经度和纬度。地理编码容易出错、慢且昂贵。此外,在万维网页中识别地址是困难的,且地址格式是以非标准方式明确地给出的。另一选择是,某些像雅虎 (Yahoo) ! 黄页的服务使用经地理编码的黄页数据库。此数据库是昂贵的且需要周期性更新-因此其不是万维网的空间搜索,而仅仅是黄页(非常小的数据集合)的空间搜索。

[0005] 因此,需要一种万维网搜索技术,其就搜索提供商来说能够用最小的整合成本实现高精度的局部万维网搜索。

发明内容

[0006] 本发明实现能够在因特网上执行的高精度度局部搜索。地图图像再现软件提供商将空间关键字嵌入到地图中,然后又提供给因特网内容的制作者,例如地图提供商、广告商、商业万维网站点、个人万维网站点等。例如,房主可在万维网布告板上贴出为其出售房屋做广告且包含显示所述房屋位置的地图的消息。

[0007] 当搜索引擎的万维网浏览器遇到具有嵌入图像中的空间关键字的页时,为所述空间关键字连同万维网页上的其它内容加索引。由于空间关键字优选地识别小的地理区域,因此为所述内容与所述空间关键字加索引将允许对搜索查询进行区域限制但仍提供有用的结果。因此,在为其房屋做广告的房主的实例中,随后搜索房主区域中“房屋出售”的搜索引擎用户将被引导到房主的万维网页以及满足地理和内容搜索词的其它万维网页。

附图说明

[0008] 图 1 图解说明其中可实施本发明的系统 100 的一个实施例。

[0009] 图 2 是根据本发明的实施例图解说明一种用于与万维网页相关联地为空间关键字加索引的方法的流程图。

[0010] 图 3 是根据本发明的实施例图解说明一种用于检索与空间关键字相关联的 URL 的方法的流程图。

[0011] 图 4 图解说明根据本发明的实施例边界框可如何全部在关注区域内、全部在关注区域外或边界框可如何横切关注区域。

[0012] 图 5 图解说明根据本发明的实施例将空间关键字值整合到图像中。

[0013] 图 6 图解说明根据本发明的实施例如何使用地图产生器来加密或解密空间关键字和边界框。

[0014] 所述图式仅出于图解说明的目的描绘本发明的优选实施例。根据以下的论述,所属领域的技术人员应易于了解,在不违背本文所述本发明原理的前提下,也可采用本文图解说明的结构及方法的替代实施例。

具体实施方式

[0015] 本发明通过将空间关键字嵌入显示为万维网页的一部分的地图中来实现高精度因特网局部搜索。通过为空间关键字连同万维网页的相关内容加索引,搜索引擎因此能够返回与局部信息查询相关的内容。

[0016] 加利福尼亚圣何塞的地开达公司 (deCarta, Inc.) 提供基于空间关键字的空间搜索能力。空间关键字是数值,通常为 32 位数字,其可用作空间参考和用于空间搜索。

[0017] 在文件“丰富的地图引擎软件库:核、路线寻找、属性编辑和数据库叠加 API 参考手册 (Rich Map Engine Software Libraries:Core, Route Finding, Attribute Edit, and Database Overlay API Reference manual),”和“向下钻取服务器参考手册 (Drill Down Server Reference Manual,)”中进一步描述了空间关键字,所述文件分别作为附件 A 和 B 包含在此,且其形成本申请案揭示内容的一部分。

[0018] 许多万维网页 (例如,关于房地产的万维网页) 已包含与所述页内容 (例如,列出的房地产) 相关的地图。使用本发明,供应其中包含空间关键字的所述地图,由此使用现有分布信道将空间关键字附加到代管局部内容的网站。

[0019] 使用本发明的系统,搜索引擎万维网浏览器不再需要执行对地址的地理编码。即使是不具有昂贵地理编码器的最简单的搜索引擎也可为局部内容进行空间加索引且执行高成熟的万维网空间搜索。

[0020] 图 1 图解说明其中可实施本发明的系统 100 的一个实施例。万维网页 112 包含文本 106 以及地图 URL 的参考 108。当图像再现软件 110 接收地图 (在地图 URL 108 指定的 URL 处) 请求时,其返回用空间关键字编码的所请求地图的图像和空间关键字引擎 112 供应的边界框。当搜索引擎 102 浏览万维网页 112 时,其为万维网页 112 的内容与嵌入从系统 100 检索的地图中的空间关键字加索引。

[0021] 图 2 是根据本发明的实施例图解说明一种用于为万维网页和相关联的空间关键字加索引的方法的流程图。

[0022] 如以上所描述,万维网页 112 包含显示万维网页中描述的事物 (例如,房地产列表、餐馆查看等) 的位置的地图。

[0023] 万维网搜索引擎 102 浏览 202 万维网且到达万维网页 112。应注意搜索引擎 102 可出于除已浏览万维网页 112 以外的原因为万维网页 112 加索引——例如,其可已被赋予为所述页加索引的明确指令或出于某些其它原因。

[0024] 一旦搜索引擎 102 已到达页 112,其跟随地图 URL 108 以从通过 URL 识别的源中检索 204 地图图像,其通常列在地图的 HTML “图像源”属性中。如上文所述和下文进一步描述,通过图像再现引擎 110 返回的地图包含通过空间关键字引擎 112 嵌入的空间关键字且在一个实施例中还包含边界框。

[0025] 已从系统 100 检索经再现地图之后,搜索引擎 102 接下来从所包含的二进制数据中提取 206 空间关键字和相关联边界框,且然后与万维网页的 URL 和万维网页 112 的内容相关联地为地图的空间关键字加索引 208。应注意尽管在图 1 中将页的非地图内容识别为文本 106,但所述内容也可以是图形、声音或不失通用性的其它类型的内容。

[0026] 在替代实施例中,空间边界区域以图形收敛方式 (steganographically) 包含在不存在空间关键字的地图中。在此实施例中,加索引搜索引擎 102 产生对应于边界区域的一个或多个空间关键字。

[0027] 参考图 3,执行搜索包含接收 302 搜索查询,从所述搜索查询确定 304 空间关键字和检索 306 被索引到一个或一个以上空间关键字和其它查询词的 URL。例如,在一个实施例中,如果用户访问搜索引擎 102 且输入查询“寻找在公路 101 北 1 英里内的所有出售房屋”,那么搜索引擎 102 确定覆盖公路 101 北的空间关键字并检索被索引到一个或一个以上空间关键字和“出售房屋”的所有 URL。

[0028] 在一个实施例中,地图图像编码空间关键字和边界框。在此实施例中,所述边界框由两点界定。如所述,空间关键字是用于执行加空间索引和搜索的常规技术。空间关键字覆盖可比地图所显示的可见区域大的经界定地理区域。因此,地图图像也可优选地编码边界框。在替代实施例中,使用能够容纳两点以上的边界区域。万维网浏览搜索引擎 102 记录地图的空间关键字和边界框。对比搜索引擎索引执行的初始查询可从空间关键字值范围中提取 URL。可执行子过滤以消除具有落在查询关注区域外的边界框的 URL。

[0029] 图 4 图解说明边界框 404、406、408 可如何全部在关注区域 408 内,例如边界框 404 ;全部在关注区域外,例如边界框 402 ;或边界框可如何横切关注区域,例如边界框 406。

[0030] 我们现在提供如何以假定位图图像的红、绿和蓝 (RGB) 值来以图形收敛的方式编码 32 位空间关键字和空间边界框的实例。尽管我们使用 32 位空间关键字用于此实例,但无疑地可使用其它长度的关键字,如所属领域的技术人员将理解。假设:

[0031] • K 是 32 位空间关键字

[0032] • X_{\min} 是 32 位有符号整数坐标空间中的 X, Y 边界框的最小 X 坐标角

[0033] • X_{\max} 是 32 位有符号整数坐标空间中的边界框的最大 X 坐标

[0034] • Y_{\min} 是 32 位有符号整数坐标空间中的 X, Y 边界框的最小 Y 坐标角

[0035] • Y_{\max} 是 32 位有符号整数坐标空间中边界框的最大 Y 坐标

[0036] 应注意经度和纬度可通过乘以 100,000 且去掉或约掉余数来转换为 32 位有符号整数坐标。

[0037] 使图像中的每一 RGB 值均被表示为 24 位十六进制的数。通常看到表示为 (例如) “0×FFAABB”的 RGB 值,此意指 8 位红像素具有 FF 的十六进制值,绿像素具有 AA 值且

蓝像素具有 BB 值。

[0038] 使用每一 RGB 值来编码 3 位信息。优选地,使用红、绿和蓝像素的最低有效位 (LSB)。可通过用位屏蔽“0×010101”对 RGB 值进行与运算来从给定 RGB 值中提取由 RGB 值编码的 3 个位。所述位屏蔽技术已为业内所熟知。

[0039] 此后,当我们提及图像的字节时,我们指个别 R、G 或 B 值。对于每一 R、G 或 B 值,我们仅使用最低有效位 (LSB) 来以图形收敛的方式编码信息。图 5 图解说明图像的字节。“B0”指示图像的字节零,“B31”指示字节 31 等。为了以图形收敛的方式编码 32 位空间关键字 K,我们优选地使用图像的第一 32 个字节。在每一字节中,我们将使用所述字节的最低有效位来存储 K 的位。如所属领域的技术人员将理解,可代替地使用其它字节。

[0040] 以相同方式对 X_{\min} 、 Y_{\min} 、 X_{\max} 和 Y_{\max} 进行图形收敛编码。

[0041] 在替代实施例中,对额外 (X, Y) 坐标对进行图形收敛编码以容纳不受限于两个坐标的边界区域。

[0042] 在替代实施例中,在经编码坐标之后对额外数据进行编码,如以下进一步描述。

[0043] 对于传输明文形式 (例如可缩放矢量图形 (SVG)) 的图像的向量格式来说,可使用文本文件中的注释来嵌入空间关键字。所述十六进制值可表示经加密或未加密空间关键字:

[0044] <! - 空间关键字 = 0×34567854->

[0045] 以类似方式对边界框进行编码。

[0046] 尽管已相对于位图和 SVG 图像描述了本发明,但是所属领域的技术人员将理解本发明还适用于使用 JPEG、GIF、PNG 和各种其它图像格式的系统。

[0047] 除空间关键字和边界框之外,本发明的替代实施例还允许将任意信息图形收敛地编码在图像中。例如,某人可能需要将指南等级附加到显示餐馆位置的地图。

[0048] 某人可能以图形收敛的方式编码在地图图像中的额外信息的无穷尽的实例包含:

[0049] ○照片

[0050] ○任意的关键字 / 值对

[0051] ○自由文本

[0052] ○URL

[0053] ○邮政编码

[0054] 优选地,使用以下方法中的一者来允许浏览万维网的搜索引擎 102 能够从不含有空间关键字的地图中区分含有空间关键字的地图。

[0055] 可信源

[0056] 在此方法中,图像的 URL 足够确定源字节含有空间关键字。例如,考虑贴在房地产中的雅虎! 地图的 URL,所述 URL 贴出:

[0057] `http://maps.yahoo.com/maps_result?addr=Somerville+Road+at+Melita&csz=santa+r+CA&country=US`

[0058] 在此情形中,“`http://maps.yahoo.com/maps_result?`”足够确定到雅虎! 地图的链接点。如果由雅虎! 产生的每个地图嵌入空间关键字和边界框,那么万维网浏览搜索引擎 102 可从以“`http://maps.yahoo.com/maps_result?`”开始的所有 URL 中提取空间关

键字和边界框。

[0059] 检查和

[0060] 在此实施例中,使用紧随第一 160 个字节之后的字节提供第一若干字节的 32 位检查和,例如经图形收敛编码的 160 个字节或位。

[0061] 万维网浏览器 102 对照第一 160 个字节检查所述检查和。应注意使用此方法,知道检查和算法的任一方有可能在空间上键控任一图像。

[0062] 数字签名

[0063] 使用紧随第一若干字节(例如 160 个字节)之后的字节对图像的数字签名(例如,PGP 签名)进行图形收敛编码。使用私人钥匙通过系统 100 产生签名。读取图像的万维网浏览器 102 必须知道公共钥匙以验证图像和其经图形收敛编码的空间关键字和边界框。

[0064] 此通过可信源的方法的一个优点是在空间上被键控的图像可从其原始源下载并再贴在不同的 URL 上。同样,此通过简单检查和的方法的优点是其允许识别图像的创作者。万维网浏览器 102 维持在空间上经键控的地图的各种提供商的公共钥匙列表且在需要时针对每一提供商运行公共钥匙验证以执行区分。

[0065] 注解

[0066] 许多图像格式支持注解。注解可以是呈如以上针对向量格式所述的注释形式或在作为文件格式本身的一部分的注解机构中。在有注解的地方,注解可用来识别含有经图形收敛编码的空间关键字和边界框的图像。

[0067] 加密

[0068] 可需要加密来防止空间关键字技术的逆向工程设计或允许地图图像的制作者(例如,“maps.xyz.com”)保持独有的读取空间关键字的能力-因此产生用于由 xyz.com 制作的地图的独有空间万维网搜索能力。

[0069] 可采用常规已知的算法(例如 DES 或 RSA)来达成此加密。例如,

[0070] 加密 $(K+X_{\min}+Y_{\min}+X_{\max}+Y_{\max}+Pad)$ ->密码文本显示在空间关键字、边界框和可选填充值的级联上运行以产生经加密值(密码文本)的加密算法。所述填充值可用来消除不安全密码文本的产生。

[0071] 加密机构可采用隐私钥匙来加密空间关键字和边界框。

[0072] 图 6 图解说明根据本发明的实施例可如何使用加密。地图数据供应商 614 向图像再现引擎 110 提供包含经加密的图形收敛空间关键字和边界框的地图数据 616,然后如以上所述,所述图像再现引擎 110 向搜索引擎 102 提供经再现地图。当搜索引擎 102 向系统 100 提供地图和正确的加密钥匙时,优选地返回经解密空间关键字和边界框(未图示)。

[0073] 在一个实施例中,图像的作者可提供其自己的空间关键字和边界框加密。可能有用的一些原因包含:

[0074] 1) 地图图像再现软件提供商已使图像作者不希望与其共享空间关键字的若干方可获得解密软件;

[0075] 2) 图像作者不希望与其共享空间关键字的若干方可以使用公共解密密钥;或

[0076] 3) 代替或除地图图像再现软件提供商所提供的加密之外,图像作者期望提供其自己的加密。

[0077] 根据本发明的实施例再次图解说明查询和响应考虑以下实例:

[0078] 用户输入查询“寻找公路 101 北 1 英里内的所有房屋。”

[0079] a. 搜索引擎对查询执行语义分析。

[0080] b. 搜索引擎识别“房屋”为所需搜索结果

[0081] c. 搜索引擎识别“1 英里内”为空间限定语

[0082] d. 搜索引擎识别“公路 101N”为路

[0083] e. 搜索引擎使用叠加插件来检索覆盖所选定几何形状的公路 101N 周围 1 英里地带（关注区域）的空间关键字范围。（出于实例的目的，认为覆盖关注区域的返回关键字范围为 $[0 \times 12345671, 0 \times 12345674]$ 、 $[0 \times 12345695, 0 \times 12345699]$ 。

[0084] f. 搜索引擎通过在其数据库中搜索来检索所有 URL 和匹配“房屋”标准且具有正确空间关键字的相关联边界框。以下是用于数据库查询的伪码：“SELECT URL AND BOUNDING_BOX WHERE KEYWORD =”

[0085] `home|house|town|home|cabin' AND (SPATIALKEY >= 0 × 12345671 AND`

[0086] `SPATIALKEY <= 0 × 12345674) OR (SPATIALKEY >= 0 × 12345695 AND`

[0087] `SPATIALKEY <= 0 × 12345699) ”`

[0088] g. 搜索引擎简洁表述搜索结果网页，其丢弃从步骤 6 的查询返回的其边界框没有横切步骤 5 中所用的公路 101N 周围的关注区域的 URL。

[0089] 已针对有限数目的实施例特别详细地描述了本发明。所属领域的技术人员将了解本发明可另外在其它实施例中实践。首先，组件的特定命名、术语的大写、属性、数据结构或者任何其它编程或结构方面均不是强制性或具有特殊意义的，且实施本发明或其特征的机制可具有不同的名称、格式或协议。此外，系统可如所述那样经由硬件和软件的组合来实施，或者完全在硬件元件中实施。同样，本文所述的各种系统组件之间的功能性的特定划分仅仅是例示性的，而不是强制性的；由单个系统组件实施的功能可代替地由多个组件来实施，且由多个组件实施的功能可代替地由单个组件来实施。例如，可以许多或一个模块提供地图图像再现软件供应商、地图图像供应商等的特定功能。

[0090] 以上说明的某些部分就算法和关于信息的符号操作表示方面呈现本发明的特征。这些算法描述和表示是所属万维网搜索领域的技术人员用来向所属领域的其它技术人员最有效地传达其工作实质的手段。从功能上和逻辑上进行描述的这些运算应理解为由计算机程序来实施。此外，已经证明有时将操作的这些布置称为模块或代码装置也是方便的，不会丧失通用性。

[0091] 然而，应知道，所有这些术语或类似术语都与适当的物理量相关联且仅是应用于这些物理量的方便标记。除非根据本论述显而易见地另有具体指定，否则应了解，在本说明的通篇中，利用例如“处理”或“计算”或“运算”或“确定”或“显示”或类似词语进行的论述是指计算机系统或类似电子计算装置所进行的动作和过程，所述计算机系统或类似电子计算装置对在计算机系统存储器或寄存器或其它此类信息存储、传输或显示装置中被表示成物理（电子）量的数据进行处理及变换。

[0092] 本发明的某些方面包含本文中以算法的形式描述的过程步骤和指令。应注意，本发明的过程步骤和指令可体现在软件、固件或硬件中，且当体现在软件中时，可进行下载以驻留在由实时网络操作系统使用的不同平台上，并通过所述平台来操作。

[0093] 本发明还涉及一种用于执行本文中的操作的设备。此设备可以针对所需的用途进

行专门构造,或者其可包括通用计算机,所述通用计算机由存储于所述计算机中的计算机程序有选择地启动或重新配置。此计算机程序可存储在计算机可读存储媒体中,例如但不限于任何类型的磁盘,包括软盘、光盘、CD-ROM、磁光盘、只读存储器 (ROM)、随机存取存储器 (RAM)、EPROM、EEPROM、磁卡或光卡、专用集成电路 (ASIC) 或适于存储电子指令且每一者均耦合到计算机系统总线的任何类型的媒体。此外,在本说明书中提及的计算机可包含单个处理器,或可以是采用多处理器设计以用于增加计算能力的构架。

[0094] 本文中所提供的算法及显示并非与任何特定计算机或其它设备内在地相关。各种通用系统还可根据本文中的教导与程序一起使用,或者可证明便于构造用于实施所需方法步骤的更专门的设备。从上文的说明中将显露各种所述系统的所需结构。另外,本发明未参照任何特定编程语言加以阐述。应了解,可使用各种编程语言来实施如本文所述的本发明的教导,且对特定语言的任何提及是为了揭示本发明的可行性和最佳模式而提供的。

[0095] 最后,应注意,本说明书中所使用的语言原则上是出于易读性和指导性目的而选择的,而不是为描述或限制发明的标的物而选择的。因此,本发明的揭示内容旨在图解说明而非限制本发明的范围。

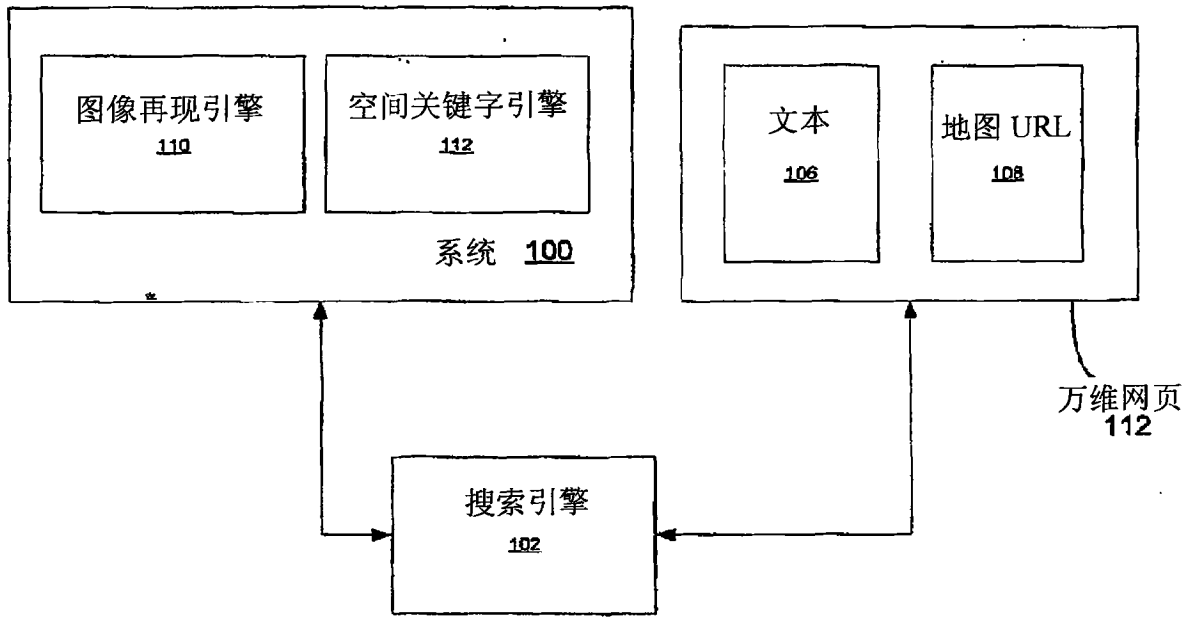


图 1

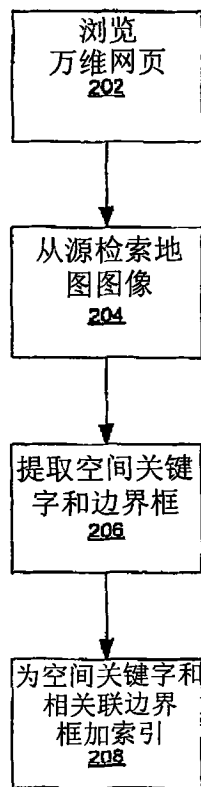


图 2

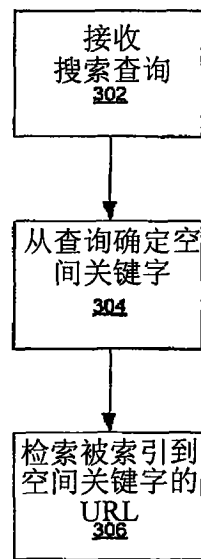


图 3

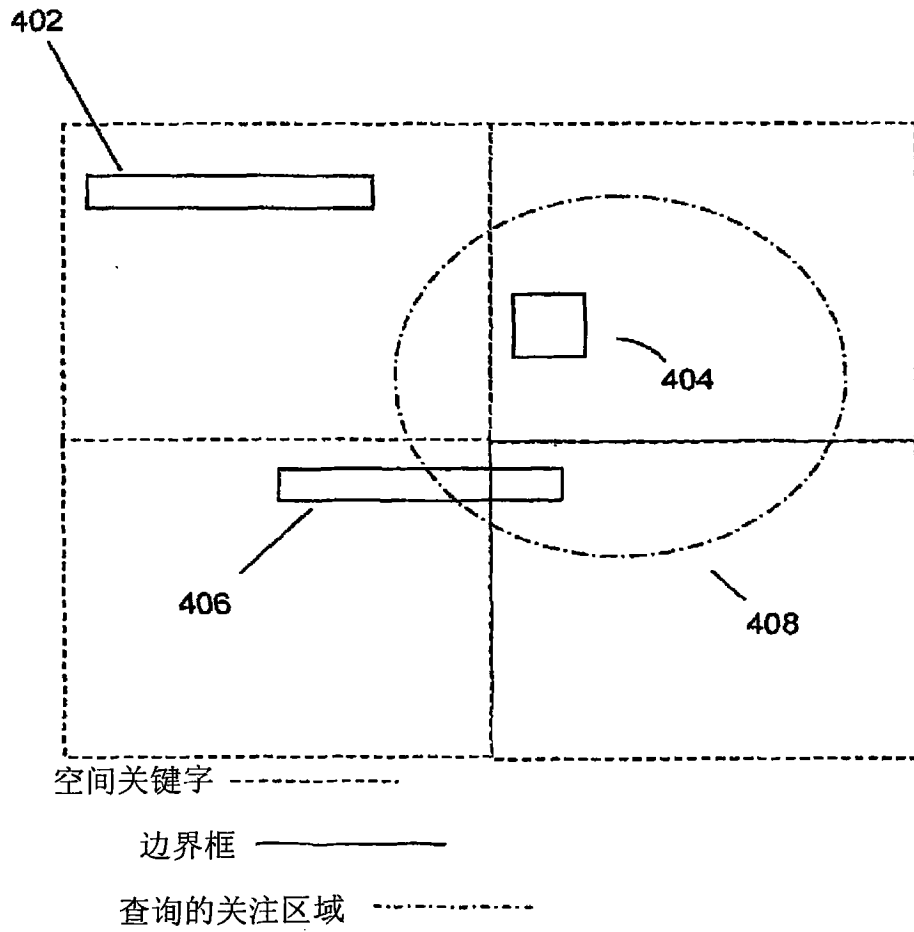


图 4

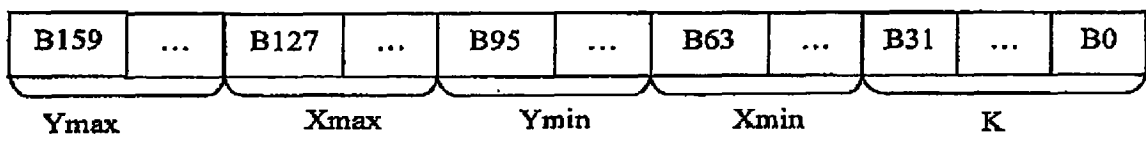


图 5

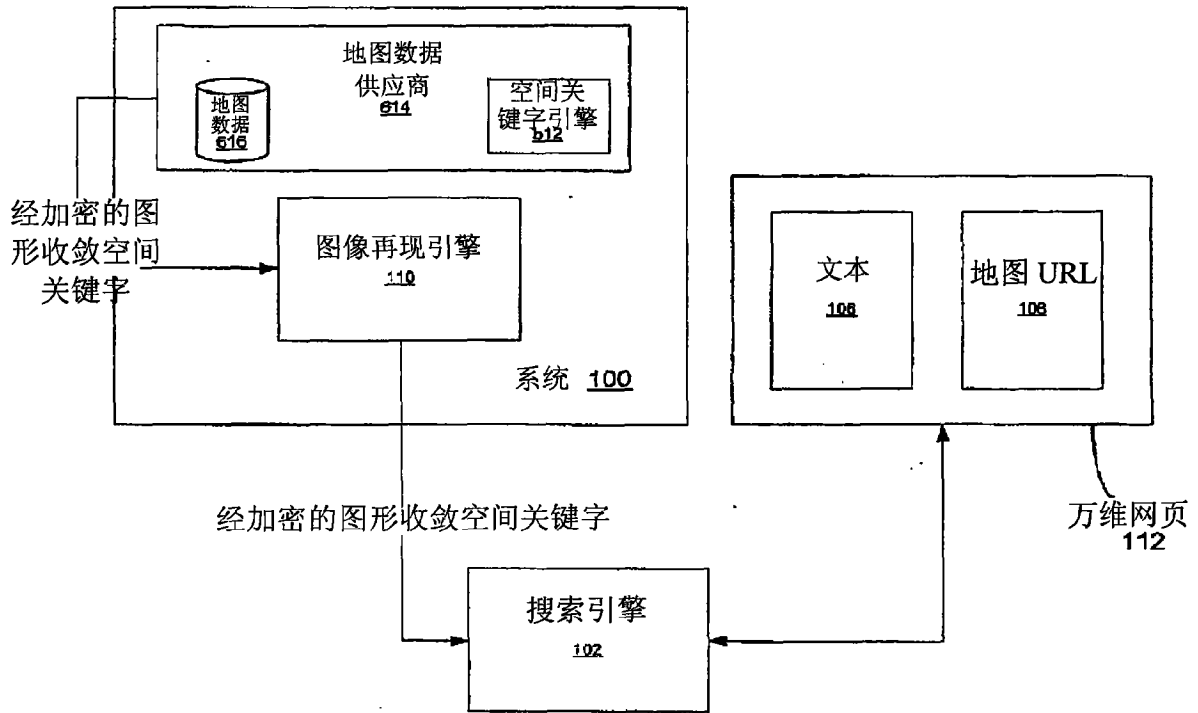


图 6