



[12] 发明专利说明书

[21] ZL 专利号 01111271.9

[45] 授权公告日 2005 年 4 月 6 日

[11] 授权公告号 CN 1196071C

[22] 申请日 2001.3.13 [21] 申请号 01111271.9

[30] 优先权

[32] 2000. 5. 31 [33] US [31] 60/207,969

[32] 2000. 9. 19 [33] KR [31] 54868/2000

[71] 专利权人 三星电子株式会社

地址 韩国京畿道

共同专利权人 加利福尼亚大学董事会

[72] 发明人 申铉科 崔良林 巴里斯·苏门根

班加洛尔·S·曼朱纳思

肖恩·纽塞姆

审查员 丁 文

[74] 专利代理机构 北京市柳沈律师事务所

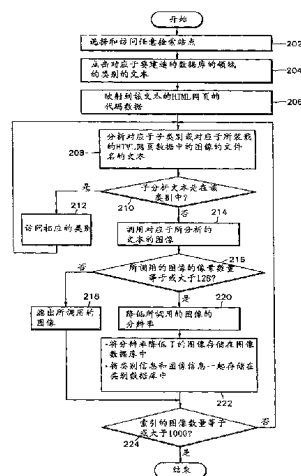
代理人 马 莹

权利要求书 3 页 说明书 14 页 附图 5 页

[54] 发明名称 适用于多媒体内容的数据库构造方法

[57] 摘要

提供了一种构造网络上可用的多媒体资源的数据库的方法，所述方法包括：从网络搜索引擎站点请求分类结点页，网络搜索引擎站点提供构成对资源的树型结构的索引的页；通过跟随到所述索引的子结点页的链接进行搜索，并且对多媒体资源进行结网搜索；请求所有被定位的多媒体资源；分析所请求的多媒体资源的定位符，以取得与其相关的关键词；以及与各自的相关的分类与关键词相关联地，将所请求的多媒体资源添加到所述数据库中。本方法使用关于相应的站点的类别信息，按语义分类多媒体内容并将它们存储在相应的数据库中。在本方法构造的数据库中，可以使用各种检索方法来快速有效地检索所要的多媒体内容。



1. 一种构造网络上可用的多媒体资源的数据库的方法, 所述方法包括:
从网络搜索引擎站点请求分类结点页, 网络搜索引擎站点提供构成对资源
5 源的树型结构的索引的页;
通过跟随到所述索引的子结点页的链接进行搜索, 并且对多媒体资源进行结网搜索;
请求所有被定位的多媒体资源;
分析所请求的多媒体资源的定位符, 以取得与其相关的关键词; 以及
10 与各自的相关的分类与关键词相关联地, 将所请求的多媒体资源添加到所述数据库中。
2. 一种构造网络上可用的多媒体资源的数据库的方法, 所述方法包括:
通过向其发送包含搜索条件的请求, 从网络搜索引擎站点请求搜索结果
15 页;
从所述结果页开始, 进行对多媒体资源的结网搜索;
请求所有被定位的多媒体资源;
分析所请求的多媒体资源的定位符, 以取得与其相关的关键词; 以及
与各自的相关的关键词相关联地, 将所请求的多媒体资源添加到所述数据库
20 中。
3. 如权利要求 1 或 2 所述的方法, 进一步包括:
分析邻近指向多媒体资源的链接的文本, 以取得与其相关的关键词。
4. 如权利要求 1 或 2 所述的方法, 其中通过存储其定位符, 将所请求的多
媒体资源添加到所述数据库中。
5. 如权利要求 4 所述的方法, 其中所请求的多媒体资源是图像数据, 并
25 且通过另外存储其分辨率降低了的版本, 将所请求的多媒体资源添加到所述
数据库中。
6. 如权利要求 1 或 2 所述的方法, 进一步包括:
如果已经向所述数据库添加了 1000 个多媒体资源, 则终止向所述数据库
添加所请求的资源。
7. 如权利要求 1 或 2 所述的方法, 其中如果图像类多媒体资源包含不多
30 于预定数目的像素, 则不将所述图像类多媒体资源添加到所述数据库中。

8. 如权利要求 7 所述的方法, 其中所述预定数目为 128。

9. 如权利要求 1 或 2 所述的方法, 其中对于图像类多媒体资源, 分析这些图像, 以抽取特征信息, 并且根据所述特征信息在所述数据库中对这些图像建立索引。

5 10. 一种构造网络上可用的多媒体资源的数据库的设备, 所述设备包括:
 用于从网络搜索引擎站点请求分类结点页的装置, 该网络搜索引擎站点提供构成对资源的树型结构的索引的页;

 用于通过跟随到所述索引的子结点页的链接, 进行搜索以及对多媒体资源的结网搜索的装置;

10 用于请求所有被定位的多媒体资源的装置;

 用于分析所请求的多媒体资源的定位符, 以取得与其相关的关键词的装置; 以及

 控制装置, 用于与各自的相关的关键词相关联地, 将所请求的多媒体资源添加到所述数据库中。

15 11. 一种构造网络上可用的多媒体资源的数据库的设备, 所述设备包括:
 用于通过向其发送包含搜索条件的请求, 从网络搜索引擎站点请求搜索结果页的装置;

 用于从所述结果页开始, 进行对多媒体资源的结网搜索的装置;

 用于请求所有被定位的多媒体资源的装置;

20 用于分析的装置, 用于分析所请求的多媒体资源的定位符, 以取得与其相关的关键词; 以及

 控制装置, 用于与各自的相关的关键词相关联地, 将所请求的多媒体资源添加到所述数据库中。

25 12. 如权利要求 10 或 11 所述的设备, 其中配置所述用于分析的装置以分析邻近指向多媒体资源的链接的文本, 以取得与其相关的关键词。

 13. 如权利要求 10 或 11 所述的设备, 其中配置所述控制装置以通过存储其定位符, 将所请求的多媒体资源添加到所述数据库中。

30 14. 如权利要求 13 所述的设备, 其中配置所述控制装置以通过另外存储其分辨率降低了的版本, 将所请求的多媒体资源添加到所述数据库中, 其中所请求的多媒体资源是图像数据。

 15. 如权利要求 10 或 11 所述的设备, 配置所述控制装置从而: 如果已

经向所述数据库添加了 1000 个多媒体资源,则终止向所述数据库添加所请求的资源。

16. 如权利要求 10 或 11 所述的设备,其中配置所述控制装置从而:如果图像类多媒体资源包含不多于预定数目的像素,则不将所述图像类多媒体资源添加到所述数据库中。

17. 如权利要求 16 所述的设备,其中所述预定数目为 128。

18. 如权利要求 10 或 11 所述的设备,其中配置所述控制装置从而:对于图像类多媒体资源,分析这些图像,以抽取特征信息,并且根据所述特征信息在所述数据库中对这些图像建立索引。

19. 一种使用根据权利要求 10 来构造的数据库,来检索多媒体资源的方法,所述方法包括:

接收关键词或者分类标识符;

检索在所述数据库中按所述关键词索引的第一组图像;

显示所述被检索的图像的子集;

- 接收指定被显示图像之一的选择输入; 以及

从所述数据库所述第一组中检索并显示第二组图像,该第二组图像按从所选择的图像中抽取的特征信息索引。

20. 一种检索多媒体内容的方法,所述方法包括:

接收来自用户的、相应于用户所要搜索的查询图像的关键词;

- 在预先确定的数据库中检索相应于关键词的图像,并且存储相应于单个图像以及多个图像的关键词;

将所检索的图像显示给用户;

从用户处接收关于被确定为视觉上相似于所述查询图像的所检索的图像的信息; 以及

- 在数据库中的、被确定为视觉上相似于所述查询图像的图像中,检索至少颜色特征、纹理特征和形状之一相似的图像。

21. 如权利要求 20 所述的方法,其中所述多个图像是通过降低原始图像的分辨率所获得的原始图像的微缩图。

适用于多媒体内容的数据库构造方法

5 技术领域

本发明涉及多媒体数据分类，具体地说，涉及适用于多媒体数据（以下称为多媒体内容）的数据库构造方法，在该方法中，多媒体内容按语义分类和存储在一预先确定的数据库中。

10 背景技术

在万维网（World Wide Web, WWW）上，通常要使用大量多媒体内容。但检索方法主要用来检索文本数据，还没有引入用来检索图像、音频数据和具有声音的动态影象数据的快速而有效的方法。

随着近来多媒体数据量的增加，需要适用于多媒体内容的数据库构造方法以及使用已经建立的数据库向用户提供检索服务的方法。

发明内容

为了解决以上问题，本发明的一个目的是提供一种适用于多媒体内容的数据库构造方法，在该方法中，将分散在万维网或其它电信网络上的多媒体内容有效地收集和存储在一个数据库中，以便能够快速检索多媒体内容。

本发明的另一目的是提供一种用于多媒体内容的数据库构造设备，它使用适用于多媒体内容的该数据库构造方法。

本发明的另一目的是提供一种在由适用于多媒体内容的数据库构造方法构造的该数据库中，快速检索多媒体内容的多媒体内容检索方法。

25 本发明的另一目的是提供一种使用适用于多媒体内容的检索方法的多媒体内容检索设备。

为了实现本发明的以上目的，这里提供一种构造网络上可用的多媒体资源的数据库的方法，所述方法包括：从网络搜索引擎站点请求分类结点页，网络搜索引擎站点提供构成对资源的树型结构的索引的页；通过跟随到所述索引的子结点页的链接进行搜索，并且对多媒体资源进行结网搜索；请求所有
30 被定位的多媒体资源；分析所请求的多媒体资源的定位符，以取得与其相关

的关键词;以及与各自的相关的分类与关键词相关联地,将所请求的多媒体资源添加到所述数据库中。

5 为了实现本发明的以上目的,这里提供一种构造网络上可用的多媒体资源的数据库的方法,所述方法包括:通过向其发送包含搜索条件的请求,从网络搜索引擎站点请求搜索结果页;从所述结果页开始,进行对多媒体资源的结网搜索;请求所有被定位的多媒体资源;分析所请求的多媒体资源的定位符,以取得与其相关的关键词;以及与各自的相关的关键词相关联地,将所请求的多媒体资源添加到所述数据库中。

10 为了实现本发明的以上目的,这里提供一种构造网络上可用的多媒体资源的数据库的设备,所述设备包括:用于从网络搜索引擎站点请求分类结点页的装置,该网络搜索引擎站点提供构成对资源的树型结构的索引的页;用于通过跟随到所述索引的子结点页的链接,进行搜索以及对多媒体资源的结网搜索的装置;用于请求所有被定位的多媒体资源的装置;用于分析的装置,用于分析所请求的多媒体资源的定位符,以取得与其相关的关键词;以及控制装置,用于与各自的相关的关键词相关联地,将所请求的多媒体资源添加到所述数据库中。

20 为了实现本发明的以上目的,这里提供一种构造网络上可用的多媒体资源的数据库的设备,所述设备包括:用于通过向其发送包含搜索条件的请求,从网络搜索引擎站点请求搜索结果页的装置;用于从所述结果页开始,进行对多媒体资源的结网搜索的装置;用于请求所有被定位的多媒体资源的装置;用于分析所请求的多媒体资源的定位符,以取得与其相关的关键词的装置;以及控制装置,用于与各自的相关的关键词相关联地,将所请求的多媒体资源添加到所述数据库中。

25 为了实现本发明的以上目的,这里提供一种使用上述方法来构造的数据库,来检索多媒体资源的方法,所述方法包括:接收关键词或者分类标识符;检索在所述数据库中按所述关键词索引的第一组图像;显示所述被检索的图像的子集;接收指定被显示图像之一的选择输入;以及从所述数据库所述第一组中检索并显示第二组图像,该第二组图像按从所选择的图像中抽取的特征信息索引。

30 为了实现本发明的以上目的,这里提供一种检索多媒体内容的方法,所述方法包括:接收来自用户的、相应于用户所要搜索的查询图像的关键词;

在预先确定的数据库中检索相应于关键词的图像，并且存储相应于单个图像以及多个图像的关键词；将所检索的图像显示给用户；从用户处接收关于被确定为视觉上相似于所述查询图像的所检索的图像的信息；以及在数据库中的、被确定为视觉上相似于所述查询图像的图像中，检索至少颜色特征、纹理特征和形状之一相似的图像。

为了实现本发明的以上目的，这里提供一种适用于多媒体内容的数据库构造方法，该方法包括步骤：(a) 经由电信网络存取提供多媒体内容的一任意站点；(b) 通过结网 (spidering) 该站点调用多媒体内容；以及 (c) 根据存储的地址分类该多媒体内容并将其存储到一预先确定的数据库中。

10 此外，该多媒体内容可以为图像数据。

该地址最好为统一资源定位符 (URL)。

该任意站点最好在一检索站点或一入口站点之间选择。

该步骤 (b) 最好还包括子步骤：(b-1) 输入搜索字；(b-2) 相应于与在具有输入的搜索字的检索结果的超文本标示语言 (HTML) 的网页数据中对应于多个子类对应的多个文本的多媒体内容的文件名的多个文本进行分析；以及 (b-3) 将具有相应于分析过的文本的地址的多媒体数据调用。

该方法最好在步骤 (b-3) 之前还包括：(p-b-3-1) 当相应于该子类的文本在载入的 HTML 网页数据中分析时，访问相应的子类。

在步骤 (b-2) 中，最好将代表这些文本的特征的关键词与相应于该子类的文本和相应于该多媒体内容的文件名的文本一起在载入的 HTML 网页数据中进行分析。

该方法最好在步骤 (b-3) 之后还包括：(b-4) 滤出所调用的图像中的噪声图像。

步骤 (b-4) 最好包括子步骤：(b-4-1) 确定一调用图像的像素数是否等于或大于一预先确定的阈值；以及 (b-4-2) 当一调用图像的像素数等于或大于该预先确定的阈值时，索引相应的图像。

该阈值最好是 128。

步骤 (c) 最好还包括子步骤：(c-1) 降低该调用图像的分辨率；以及 (c-2) 根据该分类结构，将分辨率降低了的该图像存储在一预先确定的数据库中。

30 作为替换，在步骤 (c) 中最好使用该 URL 信息将存储该调用多媒体内容数据的网页的 URL 存储在一预先确定的数据库中。

作为替换，在步骤(c)中，最好将至少一个 URL 信息或关键词信息与关于各自图像上的信息一起存储在各自预先确定的数据库中，以便关键词能够链接到单个的图像。

为了实现本发明的另一目的，这里还提供适用于多媒体内容的一数据库构造方法，该方法包括步骤：(a) 使用具有分类的结构的数据存取提供多媒体内容的一任意站点；(b) 通过结网该站点调用多媒体内容；以及(c) 使用该分类的结构，将调用的多媒体内容数据存储到一预先确定的数据库中。

为了实现本发明的另一目的，这里还提供适用于多媒体内容的一数据库构造设备，该设备包括：网站访问器(web visitor)，用于存取提供多媒体内容的一任意站点，并且通过结网该站点调用多媒体内容；以及数据库，使用该站点的数据库的分类结构和存储调用的多媒体内容数据的地址，分类和存储调用的多媒体内容数据。

为了实现本发明的另一目的，这里还提供一种适用于多媒体内容的一检索方法，该方法包括步骤：(a) 接收来自用户的、相应于将要搜索的查询图像的关键词；以及(b) 在存储相应于单个的图像以及多个图像的一预先确定的数据库中检索相应于关键词的图像。

为了实现本发明的另一目的，这里还提供适用于多媒体内容的一检索设备，具有存储多个图像以及相应于单个图像的关键词的数据库；和检索单元，用于接收来自用户的、相应于查询图像的关键词；以及在该数据库中检索相应于该关键词的多媒体内容数据。

附图说明

通过以下参考附图详细描述其优选实施例，将会更容易地理解本发明以上目的和优越性，其中：

图 1 是展示根据本发明的一实施例的、适用于多媒体内容的数据库构造设备的结构的框图。

图 2 是展示使用在图 1 的设备中的、根据本发明的一实施例的、适用于多媒体内容的数据库构造方法的主要步骤的流程图。

图 3 是展示使用在图 1 的设备中的、根据本发明的另一实施例的、适用于多媒体内容的数据库构造方法的主要步骤的流程图。

图 4 是展示根据本发明的一实施例的多媒体内容检索设备的结构的框

图。以及

图 5 展示使用在图 4 的多媒体内容检索设备中的、根据本发明的一实施例的多媒体内容的检索方法的主要步骤的流程图。

5 具体实施方式

以下参考附图详细描述本发明的实施例。本发明不局限于下列实施例，在本发明的精神和范围之内可能有多种变形。根据本发明的实施例是为了更完整地向本领域的普通技术人员解释本发明而提供的。

根据本发明，按语义分类多媒体内容以便有效地实施该检索或浏览。例如如相应于“F-16 战斗机”的多媒体内容可以分类到被称为“海湾战争”的分类中。对于这一点，用到了在检索站点中分类结构的优点。例如，诸如 Yahoo™ 这样的检索站点具有分类的结构。例如如果点击了按“电影”分类的文本，则提供在以诸如“色情”、“动作”或“人文逸事”分类的文本格式中与电影有关的更详细的站点的收集的信息。此外，提供与各种电影相关的详细站点的地址。很好地实现这样的检索站点和入口站点语义分类。因此，本发明在制作适用于多媒体内容的数据库时，使用这样的检索站点和入口站点分类结构。

图 1 是展示根据本发明的一实施例的、适用于多媒体内容的数据库构造设备的框图。图 2 是展示使用在图 1 的设备中的、根据本发明的一实施例的、

适用于多媒体内容的数据库构造方法的主要步骤的流程图。下列解释会经常参考图 2。

对本实施例，取图像作为该多媒体内容的一实例。参考图 1，根据本发明的实施例的、适用于多媒体内容的数据库构造设备 10 连接到万维网 (WWW)12，它具有一网站访问器 100、分析器 102、过滤单元 104、分辨率降低单元 106、图像数据库 1008、类别数据库 110、关键词数据库 112、统一资源定位符(URL)数据库 114 以及控制单元 120。

现在解释该适用于多媒体内容的数据库构造设备的运行。首先，在步骤 202 中用户选择和访问一任意的检索站点，并且在步骤 204 中在该正在访问的主页，即将要构造的数据库的对象上点击用户感兴趣领域所对应的类别的文本。检索站点的内容分类具有一分类结构。在步骤 206 中，响应该用户点击，网站访问器 100 从该文本中装载一映射的超文本标记语言网页数据。接下来在步骤 208 中，分析器 102 分析相应于子类的文本，或者在本实施例中是相应于图像文件名的文本，例如带扩展名“_.JPG”、“_.GIF”或“_.BMF”文本的多媒体内容。接下来在步骤 210 中确定分析过的文本是否包括在一子类中。当确定了分析过的文本包括在一子类中时，在步骤 212 中访问该子类并执行步骤 206。同时，当分析相应于在装载的 HTML 网页数据中的图像的文件名的文本时，在步骤 214 中调用具有相应于分析过的文本的文件名的图像。在这样处理之后，网站访问器 100 分级访问该检索站点的网页并调用图像。这样的操作自动执行并且可以使用被称为网页机器人(web robot)的装置实现这种操作。即可以说该网页机器人站点通过结网所选择的 URL 和它的下级 URL 与所选择的 URL 关联。

此外，在步骤 206 中，分析器 102 最好分析展示该文本以及相应于图像的文件名的文本的特征的关键词。由于关键词一般是名词，所以可以使用已知的方法提取它们。

同时，在调用的图像中用来美化网站的图形等被看成噪声并且在索引时剔除。因此，过滤该调用的图像然后索引。在本实施例中，在步骤 216 中，过滤单元 104 确定一所调用的图像的像素数是否等于或大于 128。当所调用的图像的像素数小于 128 时，在步骤 218 中将所调用的图像确定成一微缩图 (thumb nail)，然后将所调用的图像过滤出并且不索引。当所调用的图像的像素数等于或大于 128 时，在步骤 220 中将所调用的图像确定成非微缩图并且

该分辨率降低单元降低该图像的分辨率。

在步骤 222 中，将降低了分辨率的图像存储在图像数据库 108 中，并且将关于存储在图像数据库 108 中的图像的识别信息以及所访问的网页数据的类别信息都存储在类别数据库 110 中。

5 作为替换，未降低其分辨率的原始数据可以存储在该数据库中，而不用将所调用的图像存储到该数据库中，可以存储具有该图像的网页的 URL，以使得可以链接相应的站点。此外，为了能够让关键词链接到各自图像，可以将相应于各自图像的关键词与存储在图像数据库中的各个图像上的信息一起存储到关键词数据库 112 中。

10 在步骤 224 中，控制单元 120 确定所索引的图像的数量是否等于或大于 1,000。当所索引的图像的数量小于 1,000 时，输出一“低”电平控制信号，而当所索引的图像的数量等于或大于 1,000 时，输出一“高”电平控制信号。响应该输出一“高”电平控制信号，分析器 102 执行步骤 208，而响应该“低”电平控制信号，完成分析。即当所索引的图像的数量等于或
15 大于 1,000 时，完成对一站点的访问。

在根据本发明的实施例的、适用于多媒体内容的数据库构造方法中，使用相应的站点的类别信息，将位于分层地被访问的类别中的多媒体内容，例如降低了图像分辨率的微缩图，或者原始图，按语义分类和存储在相应的数据库中。

20 此外，在根据本发明的实施例的、适用于多媒体内容的数据库构造方法中，使用 URL，并且考虑 WWW 上的站点的目录结构。例如，诸如 GoogleTM 或 AltavistaTM 这样的检索站点提供基于 URL 而不是类别信息的检索。例如，当输入一搜索词“英式足球”时，提供与“英式足球”相关的站点地址作为搜索结果。即使当使用这些检索站点时，也提供与相应的搜索词按语义紧密
25 相关的站点。

在根据本发明的另一实施例的、适用于多媒体内容的数据库构造方法中，能够进行这些检索站点的语义搜索的结构，被用于构造适用于多媒体内容的数据库。图 3 是展示使用在图 1 的设备中的、根据本发明的另一实施例的、适用于多媒体内容的数据库构造方法的主要步骤的流程图。参考图 3，
30 在根据本发明的另一实施例的、适用于多媒体内容的数据库构造方法中，首先在步骤 302 中，网站访问器 100 在选择该站点之后访问一任意检索站点。

接下来，在步骤 304 中用户输入相应于要被构造的数据库的领域的搜索词。该搜索词相应于要包括在该数据库中的多媒体内容的识别符。接下来，在步骤 306 中网站访问器 100 接收关于该输入搜索词的站点的地址，例如，具有 URL 信息的 HTML 网页数据。

5 接下来，在步骤 308 中，分析器 102 分析接收到的 HTML 网页数据中的站点的地址。在步骤 310 中，该网站访问器 100 分层地访问相应于分析过的地址的站点。然后在步骤 312 中该网站访问器 100 从正在访问的检索站点中装载根 HTML 网页数据。在步骤 314 中，分析器 102 分析被装载的 HTML 网页数据中的多媒体内容，在本实施例中例如为相应于图像文件名的文本，
10 诸如带扩展名 “_.JPG”、“_.GIF” 或 “_.BMF” 的文本。作为替换，可以使用用于 HTML 语言中的 ALT 标记符。由于这些图像的名称或 ALT 标记符由网站作者手工输入，所以图像的特征，一般来说，多媒体内容的特征相对好表达。

在步骤 314 中，分析器 102 最好也分析代表所分析的文本的特征的关
15 键词。因为关键词一般为名词，所以可能一已知的方法提取它们。

接下来，在步骤 316 中，网站访问器 100 调用相应于所分析的文本的图像数据。同时，将在所调用的图像数据中用于美化网站的图形看成噪声，并且必须在索引时剔除。因此，过滤单元 104 过滤所调用的图像，以滤出噪声图像。在本实施例中，在步骤 318 中，过滤单元 104 确定所调用的图像的像素数量是否等于或大于 128。当所调用的图像的像素数量小于 128 时，在步骤 320 中将所调用的图像确定成一微缩图，然后将所调用的图像过滤出以便在索引时剔除它。当该所调用的图像的像素数等于或大于 128 时，在步骤 322 中分辨率降低单元 106 确定所调用的图像为非微缩图而是一图像，并且降低该图像的分辨率。在步骤 324 中，将降低了分辨率的图像存储在图像数据库
25 108 中，并且将存储在图像数据库 108 中的各自图像的信息以及所访问的网页数据的 URL 信息都存储在 URL 数据库中。

作为替换，未降低其分辨率的原始数据可以存储在该图像数据库 108 中，并且通过在该数据库中存储存储了该图像的网页的 URL，而不是存储所调用的图像，可以链接相应的站点。最好将相应于各自图像的关键词与存储
30 在图像数据库 108 中的各自图像上的信息一起存储到关键词数据库 112 中。

在步骤 326 中, 控制单元 120 确定所索引的图像的数量是否等于或大于一预先确定的数。当所索引的图像的数量小于 1, 000 时, 网站访问器 100 根据步骤 310 从正在访问的检索站点装载根 HTML 网页数据。当所索引的图像的数量等于或大于 1, 000 时, 完成对一站点的访问。

- 5 同时, 为了有效地检索图像, 可以提取纹理特征和/或颜色以存储在各自的特征数据库(未在图中显示)中。这些特征可以由具有比例和方向系数的 Gabor 滤波器提取。例如当由具有 3 种比例系数和 4 种方向系数的一组 Gabor 滤波器构成的滤波器计算输入的图像的一特征向量, 并且如果平均的一个分布用于该特征向量的分量时, 该特征向量可以表达成下列等式 1:

10
$$f_{texture} = [t_1, t_2, t_2, \dots t_{24}, \dots] \dots (1)$$

使用该特征向量索引图像。在该特征数据库中, 存储该特征向量和相应于该特征向量的图像信息。

- 类似地, 可以提取颜色特征以存储在各自特征数据库中。可以在 CIE LUV 颜色空间中计算出来的颜色分布图表中提取展示颜色图元(primitive)的特征向量。例如, 如果 3 维颜色空间的每一维都按四级进行量化, 则可以将
- 15 其表达成如下列等式 2 的 64 维颜色特征向量:

$$f_{color} = [c_1, c_2, c_2, \dots c_{64}, \dots] \dots (2)$$

在该特征数据库中存储该特征向量和相应于该特征向量的图像信息。

- 在根据本发明的另一实施例的、适用于多媒体内容的数据库构造方法
- 20 中, 图像分辨率降低了的微缩图或原始图像都在使用相应的站点的 URL 信息按语义分类之后, 存储在相应的数据库中, 其中微缩图和原始图像都是从所访问的类别中调用的。存储调用图像的纹理和/或颜色特征到各自的特征数据库中。

- 在根据本发明的、适用于多媒体的数据库构造方法中, 按语义分类和索引 WWW 上的多媒体内容。这样的适用于多媒体内容的数据库构造方法可以用于诸如电视新闻广播这样的多媒体内容, 或者用于使用在线多媒体表达的商品清单。
- 25

- 虽然在以上实施例中示范了构造一个图像数据库, 但本发明可以用于诸如声音片断以及具有声音的电影剪辑这样的各种多媒体内容。即本发明不局
- 30 限于上述实施例, 并且本发明的范围由附随的权利要求书确定。

在由根据上述描述的本发明的、适用于多媒体内容的数据库构造方法构

造的数据库中，使用类别信息或 URL 信息很好地收集了分散在 WWW 上的多媒体内容，并且按语义很好地分类了该多媒体内容。因此，可以使用适用于多媒体的各种检索方法来有效地检索所要的多媒体内容。特别是，当(使用)根据本发明的、用于检索多媒体内容的方法时，可以有效地检索类似于多媒体数据中的查询数据的数据。

图 4 是展示根据本发明的一实施例的多媒体内容检索设备的结构的框图。参考图 4，根据本发明的一实施例的多媒体内容检索设备经由 WWW42 连接到服务器 44，以提供图像检索服务，该服务器经由因特网提供一种服务。

该多媒体内容检索设备具有一关键词检索单元 402、一显示图像选择单元 404、一图像显示单元 406、一图像检索单元 408、一用户接口 410 以及一用来与 WWW42 通信的网络服务器 412。

服务器 44 具有由参考图 2 和 3 所解释的、适用于多媒体内容的数据库构造方法构造的一些数据库，即一图像数据库 440、一类别数据库 442、一 URL 数据库 444 以及一关键词数据库 446。此外，服务器 44 还有一用于与 WWW 通信的网络服务器 448。

图 5 展示使用在图 4 的多媒体内容检索设备中的、根据本发明的一实施例的多媒体内容的检索方法的主要步骤的流程图。图 5 时常被参考。在本实施例中，图像取成多媒体内容的一实例，并假设数据库使用根据参考图 2 解释过的本发明的实施例的、适用于多媒体内容的数据库构造方法构造。

参考图 5，首先，在步骤 502 中接收相应于由一用户所要求的一查询图像的、来自用户的关键词。当用户要求检索带有查询图像、有一定形状的“鞋”时，用户在一计算机中运行存储了用于执行根据本发明所述多媒体内容检索方法的程序码的记录介质，并向位于显示在用户监视器上的运行屏幕上的一检索关键词空间输入关键词“鞋”。

接下来，关键词检索单元 402 经由网络服务器 412，在服务器 44 的关键词数据库 446 中检索与所输入的关键词的相同词。当检索该相同的词时，从图像数据库 440 中调入链接到所检索的词的图像。通过这些处理后，在步骤 504 中，检索相应于所输入的关键词的图像。

同时，由于在数据库中存在很多图像，并且通过仅使用一个关键词在庞大的数据库中获得的所检索的图像可能包括实际上与所要的图像不相似的

那些图像, 所以仅使用一个关键词几乎不可能通过用一次检索检索到所要的图像。因此, 用户最好用裸眼在所按所的图像中核对某些图像, 并选择类似的图像来将所选择的图像反馈给图像检索单元 408, 以便再次执行检索。

对于这一点, 在步骤 504 种, 显示图像选择单元 404 从所检索的图像中
5 间选择预先确定的数量的图像, 并且在步骤 506 中, 图像显示单元 406 为用户显示预先确定数量的所选择的图像。

接下来, 用户通过用裸眼观察所显示的图像, 选择一个或多个与用户要寻找的图像相似的图像, 并且将那些图像确定为查询图像和提供关于它们的信息。在本实施例中, 响应用户的输入, 用户接口 410 选择多个鞋形图像并
10 提供选择信息。通过这样的处理后, 在步骤 508 中, 图像检索单元 408 从该用户处接收关于候选的查询图像的信息, 其中的图像是判决为视觉上相似于所要的图像的图像。

接下来, 在步骤 510 中, 图像检索单元 408 在该图像数据库中, 检索在被确定为视觉上相似于该查询图像的候选的查询图像中至少在颜色特征、纹
15 理特征和形状中的一特征相似的图像。

为了确定是否有两个图像, 即该查询图像和所检索的图像真正相似, 可以通过计算两个图像的特征向量的差异来获得相似性。在本实施例中, 假设图像的特征向量存储在一特征数据库(未在图中显示)中。当 k 为纹理向量的长度时, 两个图像 I 和 j 的纹理特征之间的差异可以由以下等式 3 获得:

$$20 \quad d_{\text{texture}}(i, j) = \sum_{k=1}^{24} |t_k^{(i)} - t_k^{(j)}| \dots\dots(3)$$

此外, 当 k 为颜色向量的长度时, 两个图像 I 和 j 的颜色特征之间的差异可以通过计算两个特征向量的欧氏(Euclidean)距离, 使用下列等式 4 获得:

$$d_{\text{colour}}(i, j) = \left(\sum_{k=1}^{64} (c_k^{(i)} - c_k^{(j)})^2 \right)^{1/2} \dots\dots(4)$$

确定与已给的查询图像的特征向量相比具有最小差异的特征向量的图像为
25 所检索的图像。

当要检索的图像是一原始图像时, 所检索的图像按原样提供给用户。当要检索的图像是微缩图时, 所检索的图像的 URL, 即相应于该微缩图的原始图像的 URL 将在具有相应于该 URL 的站点经由因特网连接上之后, 被用于调用该原始图像, 然后给用户该原始图像。这时, 可以将该 URL 信
30 息与微缩图一其存储在图像数据库 422 中。

在基于内容检索时，用户选择有关的查询图像的一个集合 R 。依据在颜色空间中收集这样的图像集合的紧密程度确定颜色和纹理特征的相对加权值。即当 $|R|$ 为该查询收集的图像数量时，该加权值通过下列等式 5 和 6 获得：

$$\bar{d}_{texture} = \frac{1}{|R|} \sum_{i,j \in R} d_{texture}(i,j) \dots\dots (5)$$

$$\bar{d}_{color} = \frac{1}{|R|} \sum_{i,j \in R} d_{color}(i,j) \dots\dots (6)$$

接下来，当 ε 为用来防止任何一个特征极其显著的预先确定的小值时，该加权值可以经由下列等式 7 和 8 来获得：

$$w_{texture} = \frac{1}{\bar{d}_{texture} + \varepsilon} \dots\dots (7)$$

$$w_{color} = \frac{1}{\bar{d}_{color} + \varepsilon} \dots\dots (8)$$

当 N 是一预先确定的正数时，可以通过计算下列等式 9 获得 N 的最近的邻居：

$$d(\bullet, \bullet) = w_{texture} d_{texture}(\bullet, \bullet) + w_{color} d_{color}(\bullet, \bullet) \dots\dots (9)$$

一般来说，查询由纹理特征向量和颜色特征向量的单一对指定。因此，在本实施例中，当选择多个查询图像时，使用该特征向量和颜色特征向量的平均。即通过下列等式 10 和 11 获得该值：

$$\bar{f}_{texture} = \frac{1}{|R_q|} \sum_{i \in R} f_{texture}^{(i)} \dots\dots (10)$$

$$\bar{f}_{color} = \frac{1}{|R_q|} \sum_{i \in R} f_{color}^{(i)} \dots\dots (11)$$

基于内容的检索可以推广如下。在使用特征向量 $f_{texture}$ 和 f_{color} 的一单一查询图像中，首先，当 i 为 $1, \dots, N/2$ 和 $i \leq j$ 时，假设下列条件 12 和 13 满足：

$$d_{texture}(f_{texture}, s_{texture}^{(i)}) \leq d_{texture}(f_{texture}, s_{texture}^{(j)}) \dots\dots (12)$$

(Here, $x \in S_{texture}$)

$$d_{texture}(f_{texture}, s_{texture}^{(N/2)}) \leq d_{texture}(f_{texture}, x_{texture}^{(j)}) \dots\dots (13)$$

然后可使用下列等式 14：

$$s_{texture} = \{s^{(i)}\} \dots\dots (14)$$

第二，当 i 为 $1, \dots, N/2$ 和 $i \leq j$ 时，假设下列条件 15 和 16 满足：

$$d_{color}(f_{color}, s_{color}^{(i)}) \leq d_{color}(f_{color}, s_{color}^{(j)}) \dots\dots (15)$$

(Here, $x \in S_{color}$)

$$d_{color}(f_{color}, s_{color}^{(N/2)}) \leq d_{color}(f_{color}, x_{color}^{(j)}) \dots\dots (16)$$

然后可使用下列等式 17：

$$s_{color} = \{s^{(i)}\}.....(17)$$

此外，在具有 $f_{texture}$ 和 f_{color} 的多个查询图像中，当 i 为 $1, \dots, N/2$ 和 $i \leq j$ 时，假设下列条件 18 和 19 满足：

$$d(\bar{f}_{texture}, \bar{f}_{color}, (s_{texture}^{(i)}, s_{color}^{(i)})) \leq d(\bar{f}_{texture}, \bar{f}_{color}, (s_{texture}^{(j)}, s_{color}^{(j)})).....(18)$$

(Here, $x \in S_{texture}$)

$$5 \quad d(\bar{f}_{texture}, \bar{f}_{color}, (s_{texture}^{(N)}, s_{color}^{(N)})) \leq d(\bar{f}_{texture}, \bar{f}_{color}, (x_{texture}, x_{color})).....(19)$$

然后可使用下列等式 20：

$$s = \{s^{(i)}\}.....(20)$$

接下来，显示图像选择单元 404 再从所检索的、至少颜色特征、纹理特征和形状之一相似的图像中间选择预先确定的数量的图像，并且在步骤 512 中，图像显示单元 406 为用户显示预先确定数量的所选择的图像。这里，检索范围最好限制在该查询图像类别和邻近类别之内。

当该数据库根据适用于多媒体内容的、根据参考图 4 解释的本发明的第二实施例的数据库构造方法构造时，检索范围最好限制在该查询图像 URL 和邻近 URL 之内。检索的目标图像可以是原始图像也可以是通过降低该原始图像的分辨率获得的微缩图。当该检索目标图像是原始图像时，可以更精确地实现检索，但检索时间会依据数据量和系统性能延长。当该检索目标图像是微缩图时，精确度降低但检索时间缩短。因此，可以根据目的适当地管理数据库。

响应用户的输入，当用户通过用肉眼观察所显示的图像时，用户接口 410 选择一个或多个被确定为与用户所要的图像相似的图像，并且提供被确定为视觉上相似于该查询图像的图像的信息。通过这样的处理后，图像检索单元 408 再次从该用户处接收被确定为视觉上相似于该查询图像的图像的信息。再次接收到的图像看成候选查询图像。接下来，图像检索单元 408 再次在该图像数据库 422 中检索那些图像，其中至少那些图像的颜色特征、纹理特征和形状之一视觉上相似于该查询图像。即在步骤 514 中，确定是否检索到所要的图像，当未检索到所要的图像时，重复执行步骤 508 至 512。这里，检索范围最好限制在该查询图像的类别和邻近类别之内。

该多媒体内容检索方法能够在收集存储多媒体内容的数据库中快速检索所要的图像。

该适用于多媒体内容的数据库构造方法和检索方法可以写成运行于个

人计算机或服务器类计算机的程序。构成该程序的程序代码和代码段可以由该领域的计算机程序员轻而易举地写出。该程序可以存储在计算机可读记录介质中。该记录介质包括磁记录介质、光记录介质和无线电波介质。

- 5 如上所述，根据本发明的、适用于多媒体内容的数据库构造方法，使用相应的站点的类别信息，按语义分类多媒体内容，并将它们存储在相应的数据库中。在由根据本发明的、适用于多媒体内容的数据库构造方法构造的数据库中，很好地收集了分散在 WWW 上的多媒体内容，并且使用类别信息或 URL 信息按语义很好地分类该多媒体内容。因此，可以使用用来检索多媒体内容的各种方法，使得快速有效地检索所要的多媒体内容。

图 1

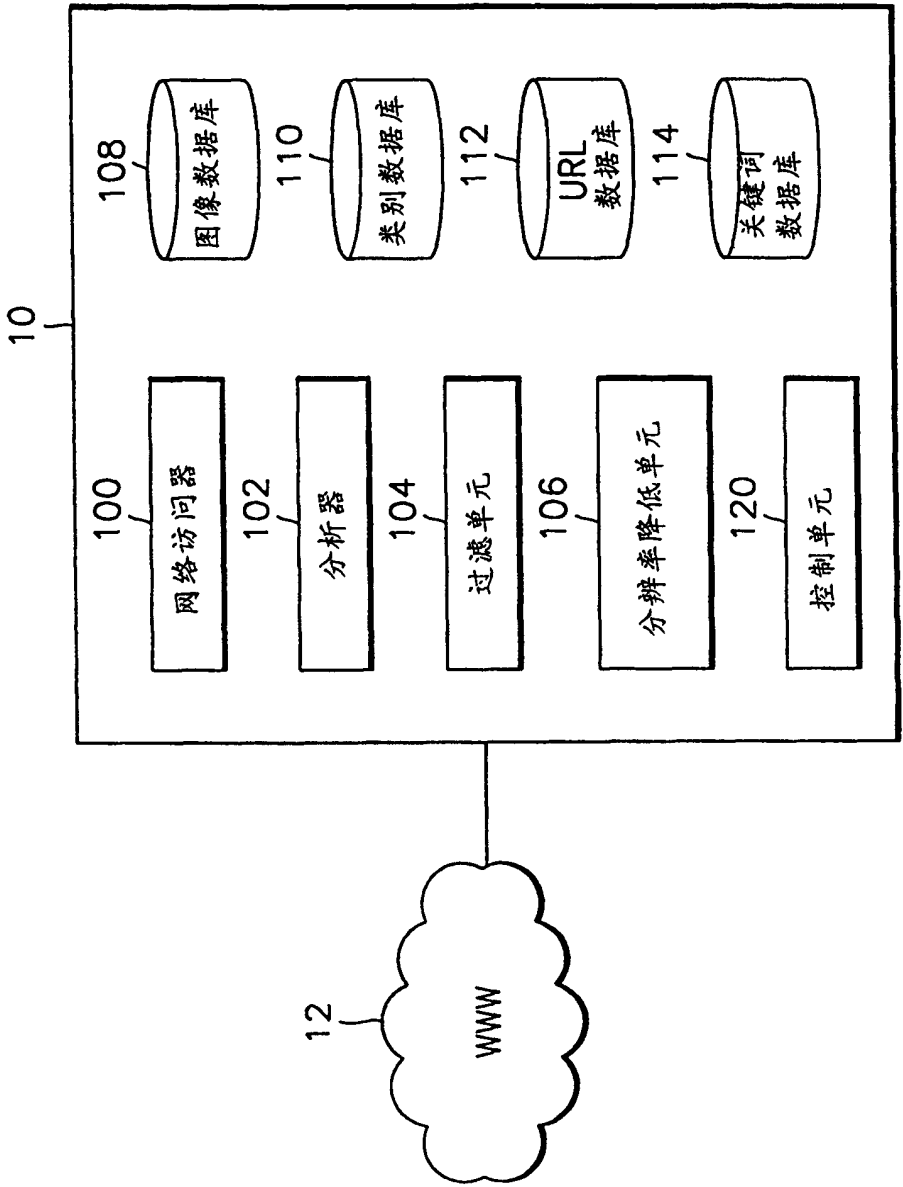


图 2

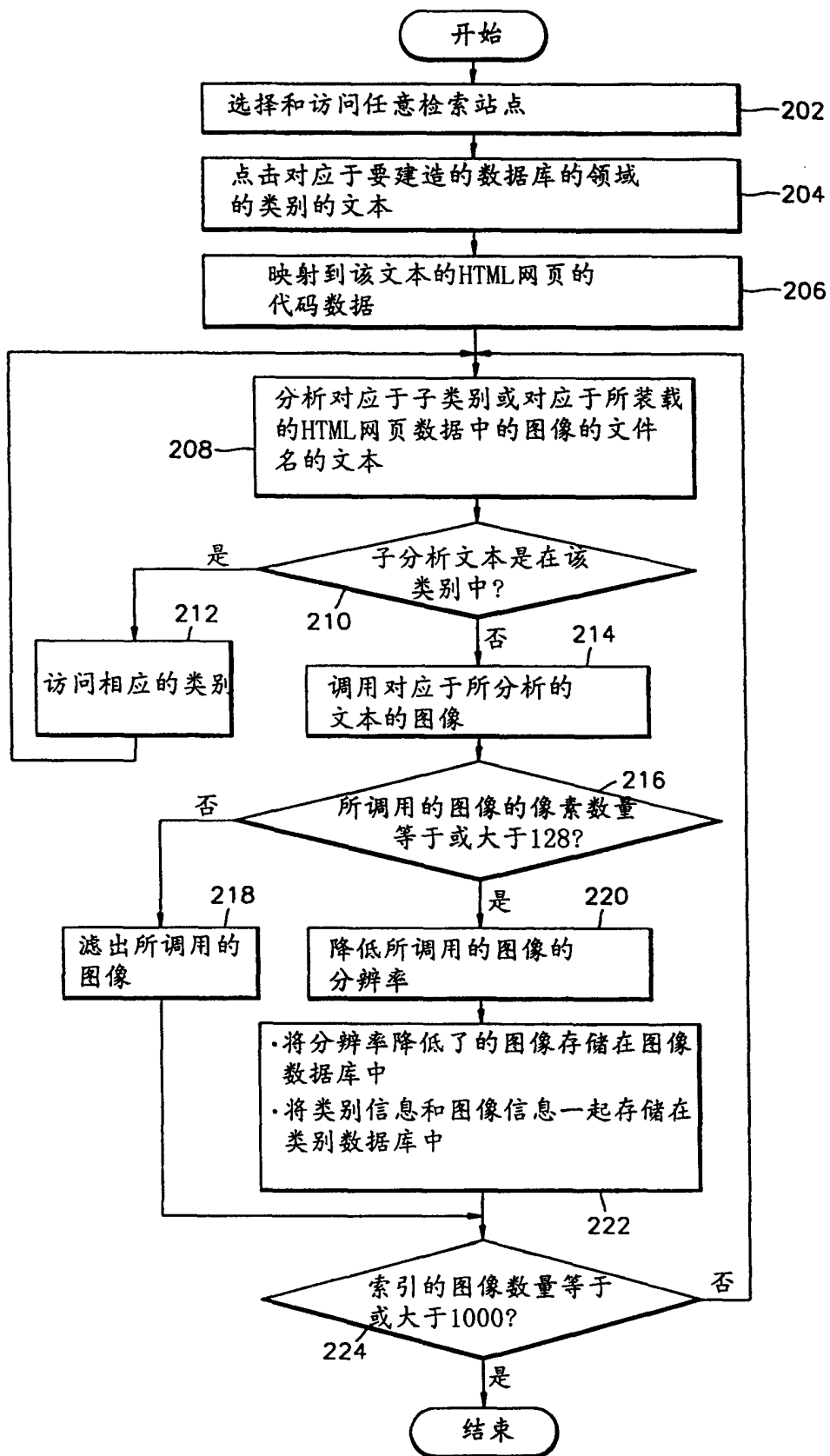


图 3

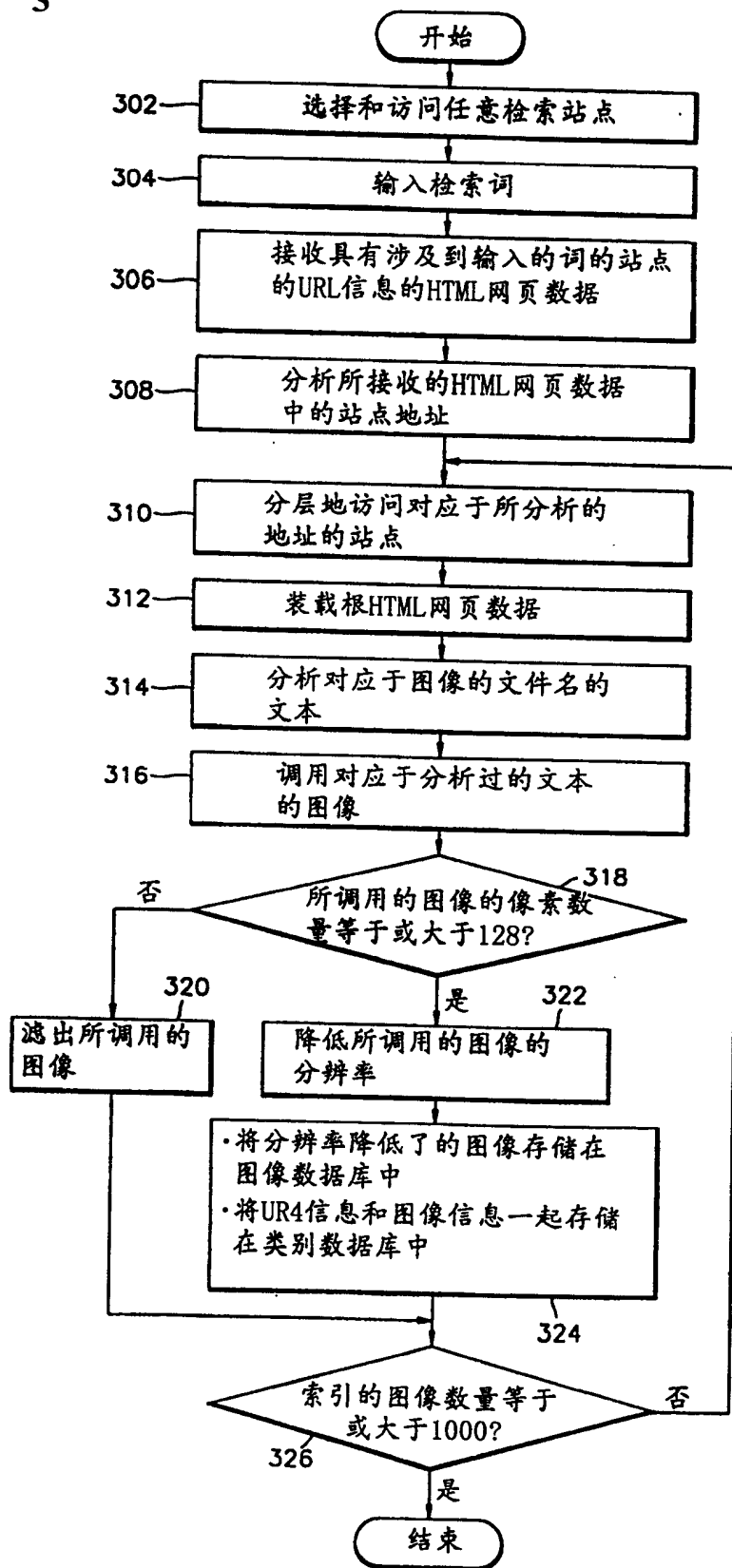


图 4

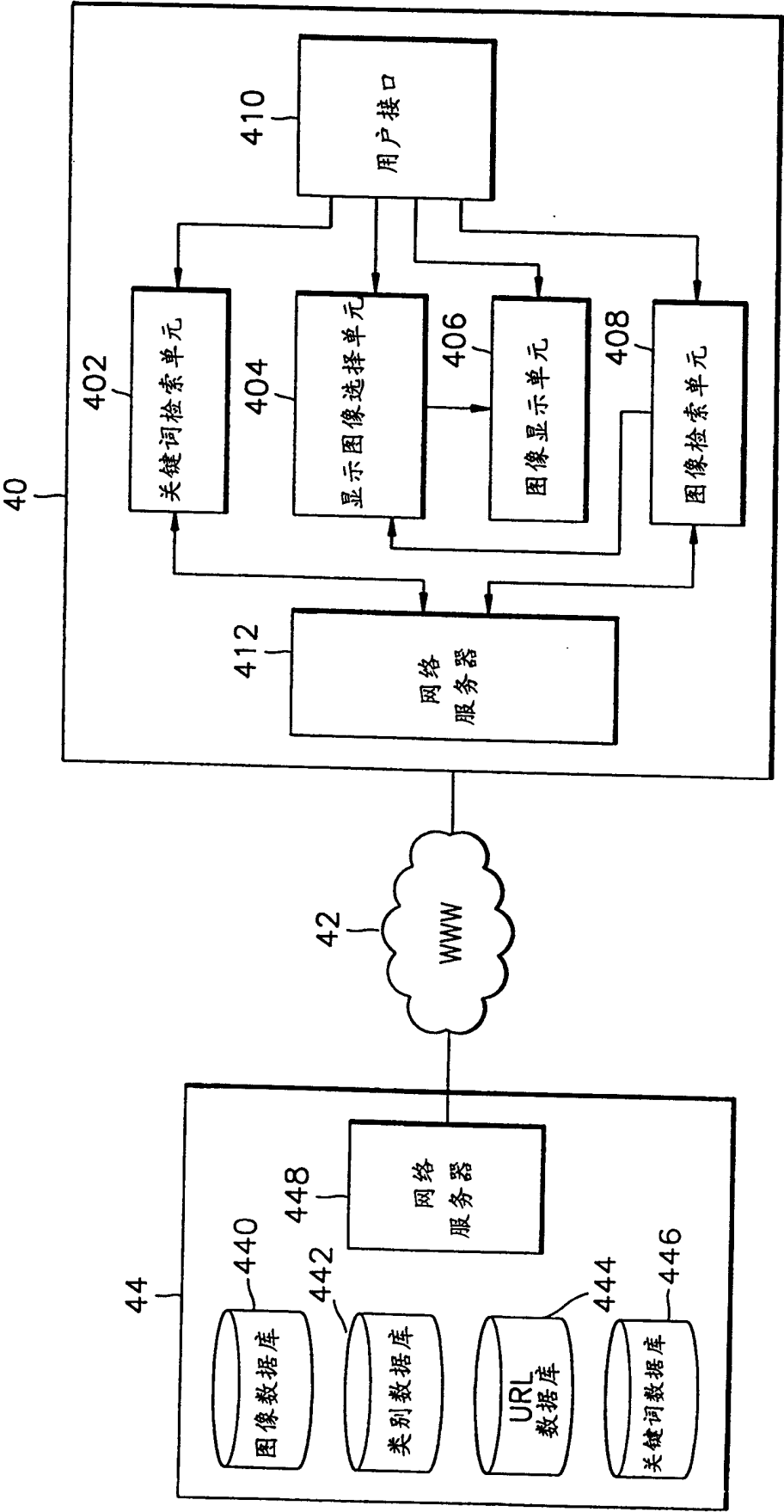


图 5

