



# (12) 发明专利申请

(10) 申请公布号 CN 102929871 A

(43) 申请公布日 2013. 02. 13

(21) 申请号 201110225313. 3

(22) 申请日 2011. 08. 08

(71) 申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市福田区振兴路赛格科技园 2 栋东 403 室

(72) 发明人 杨巍 张立明 朱靖君

(74) 专利代理机构 广州三环专利代理有限公司

44202

代理人 郝传鑫 熊永强

(51) Int. Cl.

G06F 17/30(2006. 01)

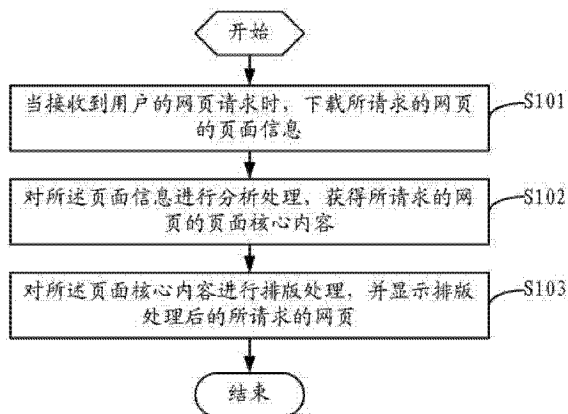
权利要求书 4 页 说明书 13 页 附图 6 页

## (54) 发明名称

一种网页浏览方法、装置及移动终端

## (57) 摘要

本发明实施例公开了一种网页浏览方法,包括:当接收到用户的网页请求时,下载所请求的网页的页面信息;对所述页面信息进行分析处理,获得所请求的网页的页面核心内容;对所述页面核心内容进行排版处理,并显示排版处理后的所请求的网页。本发明实施例还公开了一种网页浏览装置及移动终端。采用本发明,可节省传输带宽,提高网页请求的响应速度,同时提升网页浏览效果和用户的浏览体验。



1. 一种网页浏览方法,其特征在于,包括:  
当接收到用户的网页请求时,下载所请求的网页的页面信息;  
对所述页面信息进行分析处理,获得所请求的网页的页面核心内容;  
对所述页面核心内容进行排版处理,并显示排版处理后的所请求的网页。
2. 如权利要求 1 所述的方法,其特征在于,所述当接收到用户的网页请求时,下载所请求的网页信息,包括:  
当接收到用户的网页请求时,获取所述用户所请求的网页的统一资源定位符 URL;  
从互联网中下载所述 URL 对应的网页的页面信息;  
其中,所述页面信息包括:超文本标记语言 HTML 源网页和相关附件;  
所述 HTML 源网页包括:HTML 源码信息和 HTML 标签;  
所述相关附件包括:层叠样式表 CSS、控件 Frame 及网页脚本语言 JS 中的任一种或多种。
3. 如权利要求 2 所述的方法,其特征在于,所述页面核心内容包括:页面标题和页面核心正文;  
所述对所述页面信息进行分析处理,获得所请求的网页的页面核心内容,包括:  
从所述 HTML 源网页中读取文本标题 H 标签中的文本和标题 Title 标签中的文本;  
判断所述 H 标签中的文本与所述 Title 标签中的文本的相似度是否大于预设的阈值;  
若判断结果为是,将所述 H 标签中的文本确定为所请求的网页的页面标题;  
从所述页面标题在所述 HTML 源网页中所在位置开始,识别所述 HTML 源网页中的 HTML 源码信息;  
当识别到与预设的结尾标记相匹配的 HTML 源码信息时,将所述 HTML 源网页中所述页面标题所在位置至所述匹配的 HTML 源码信息所在位置之间的部分,确定为所请求的网页的页面核心正文;  
其中,所述 H 标签包括:H1 标签、H2 标签、H3 标签、H4 标签、H5 标签和 H6 标签中的任一种;  
所述预设的结尾标记包括:“相关评论”标识、“相关链接”标识、“相关阅读”标识、“相关推荐”标识中的任一种或多种。
4. 如权利要求 3 所述的方法,其特征在于,所述 HTML 源网页还包括:多媒体信息的源地址,所述页面核心内容还包括:页面核心多媒体信息,在确定了所请求的网页的页面核心正文之后,还包括:  
从所述 HTML 源网页中提取位于所述页面核心正文内的多媒体信息的源地址;  
判断提取的所述源地址是否包含广告标识,将未包含广告标识的源地址对应的多媒体信息确定为页面核心多媒体信息;  
其中,所述页面核心多媒体信息为所述页面核心正文相关的多媒体信息,包括:图片、视频、闪存 Flash 中的任一种或多种。
5. 如权利要求 4 所述的方法,其特征在于,在确定了页面核心多媒体信息之后,还包括:  
根据所述页面核心多媒体信息对应的源地址,从互联网中下载所述页面核心多媒体信息;

按照显示界面的要求,对所述页面核心多媒体信息进行缩略处理,获得缩略图并存储,保存所述缩略图的存储地址。

6. 如权利要求 5 所述的方法,其特征在于,若所述页面核心多媒体信息为图片,则所述按照显示界面的要求,对所述页面核心多媒体信息进行缩略处理,获得缩略图,包括:

按照显示界面的长度和宽度,对所述图片进行缩略处理形成缩略图;

若所述页面核心多媒体信息为视频或 Flash,则所述按照显示界面的要求,对所述页面核心多媒体信息进行缩略处理,获得缩略图,包括:

提取所述视频或 Flash 的至少一帧关键帧图像;

按照显示界面的长度和宽度,对所述关键帧图像进行缩略处理形成缩略图;

其中,所述缩略图的长度不大于所述显示界面的长度,所述缩略图的宽度不大于所述显示界面的宽度。

7. 如权利要求 2-6 任一项所述的方法,其特征在于,所述对所述页面核心内容进行排版处理,并显示排版处理后的所请求的网页,包括:

根据所述 HTML 源网页和相关附件,对所述页面核心内容进行排版,形成 HTML 网页;

按照显示界面的要求对所述 HTML 网页进行转换处理,显示转换后的所请求的网页。

8. 如权利要求 7 所述的方法,其特征在于,所述按照显示界面的要求对所述 HTML 网页进行转换处理,显示转换后的所请求的网页,包括:

将所述 HTML 网页中的页面核心多媒体信息对应的源地址替换为相应的缩略图的存储地址;和/或,

按照显示界面的格式要求,将所述 HTML 网页中所述显示界面不支持的 HTML 标签进行替换或删除,形成无线应用协议 WAP 网页;

其中,所述 WAP 网页为 WAP1.0 网页或 WAP2.0 网页。

9. 一种网页浏览装置,其特征在于,包括:

网页下载模块,用于当接收到用户的网页请求时,下载所请求的网页的页面信息;

网页分析模块,用于对所述网页下载模块下载的页面信息进行分析处理,获得所请求的网页的页面核心内容;

排版处理模块,用于对所述网页分析模块获得的页面核心内容进行排版处理;

显示模块,用于显示所述排版处理模块处理后的所述所请求的网页。

10. 如权利要求 9 所述的装置,其特征在于,所述网页下载模块包括:

地址获取单元,用于当接收到用户的网页请求时,获取所述用户所请求的网页的 URL;

下载单元,用于从互联网中下载所述地址获取单元获取的 URL 对应的网页的页面信息;

其中,所述页面信息包括:HTML 源网页和相关附件;

所述 HTML 源网页包括:HTML 源码信息和 HTML 标签;

所述相关附件包括:CSS、Frame 及 JS 中的任一种或多种。

11. 如权利要求 10 所述的装置,其特征在于,所述页面核心内容包括:页面标题和页面核心正文,所述网页分析模块包括:

读取单元,用于从所述 HTML 源网页中读取 H 标签中的文本和 Title 标签中的文本;

第一判断单元,用于判断所述读取单元读取的所述 H 标签中的文本与所述 Title 标签

中的文本的相似度是否大于预设的阈值；

标题确定单元,用于当所述第一判断单元判断结果为是时,将所述 H 标签中的文本确定为所请求的网页的页面标题；

识别单元,用于从所述标题确定单元确定的页面标题在所述 HTML 源网页中所在位置开始,识别所述 HTML 源网页中的 HTML 源码信息；

正文确定单元,用于当所述识别单元识别到与预设的结尾标记相匹配的 HTML 源码信息时,将所述 HTML 源网页中所述页面标题所在位置至所述匹配的 HTML 源码信息所在位置之间的部分,确定为所请求的网页的页面核心正文；

其中,所述 H 标签包括:H1 标签、H2 标签、H3 标签、H4 标签、H5 标签和 H6 标签中的任一种；

所述预设的结尾标记包括:“相关评论”标识、“相关链接”标识、“相关阅读”标识、“相关推荐”标识中的任一种或多种。

12. 如权利要求 11 所述的装置,其特征在于,所述 HTML 源网页还包括:多媒体信息的源地址,所述页面核心内容还包括:页面核心多媒体信息,所述网页分析模块还包括:

提取单元,用于从所述 HTML 源网页中提取位于所述正文确定单元确定的页面核心正文内的多媒体信息的源地址；

第二判断单元,用于判断所述提取单元提取的所述源地址是否包含广告标识；

多媒体确定单元,用于根据所述第二判断单元的判断结果,将未包含广告标识的源地址对应的多媒体信息确定为页面核心多媒体信息；

其中,所述页面核心多媒体信息为所述页面核心正文相关的多媒体信息,包括:图片、视频、Flash 中的任一种或多种。

13. 如权利要求 12 所述的装置,其特征在于,还包括:

多媒体下载模块,用于根据所述页面核心多媒体信息对应的源地址,从互联网中下载所述页面核心多媒体信息；

多媒体处理模块,用于按照显示界面的要求,对所述多媒体下载模块下载的页面核心多媒体信息进行缩略处理,获得缩略图并存储,保存所述缩略图的存储地址。

14. 如权利要求 13 所述的装置,其特征在于,所述多媒体处理模块包括:

第一处理单元,用于当所述页面核心多媒体信息为图片时,按照显示界面的长度和宽度,对所述图片进行缩略处理形成缩略图；

第二处理单元,用于当所述页面核心多媒体信息为视频或 Flash 时,提取所述视频或 Flash 的至少一帧关键帧图像,按照显示界面的长度和宽度,对所述关键帧图像进行缩略处理形成缩略图；

其中,所述缩略图的长度不大于所述显示界面的长度,所述缩略图的宽度不大于所述显示界面的宽度。

15. 如权利要求 10-14 任一项所述的装置,其特征在于,所述排版处理模块包括:

排版单元,用于根据所述 HTML 源网页和相关附件,对所述页面核心内容进行排版,形成 HTML 网页；

转换处理单元,用于按照显示界面的要求对所述排版单元形成的 HTML 网页进行转换处理,形成转换后的所请求的网页。

16. 如权利要求 15 所述的装置,其特征在于,所述转换处理单元按照显示界面的要求对所述排版单元形成的 HTML 网页进行转换处理,包括:

将所述 HTML 网页中的页面核心多媒体信息对应的源地址替换为相应的缩略图的存储地址;和/或,

按照显示界面的格式要求,将所述 HTML 网页中所述显示界面不支持的 HTML 标签进行替换或删除,形成 WAP 网页;

其中,所述 WAP 网页为 WAP1.0 网页或 WAP2.0 网页。

17. 一种移动终端,其特征在于,包括网页浏览装置,所述网页浏览装置包括:

网页下载模块,用于当接收到用户的网页请求时,下载所请求的网页的页面信息;

网页分析模块,用于对所述网页下载模块下载的页面信息进行分析处理,获得所请求的网页的页面核心内容;

排版处理模块,用于对所述网页分析模块获得的页面核心内容进行排版处理;

显示模块,用于显示所述排版处理模块处理后的所述所请求的网页。

## 一种网页浏览方法、装置及移动终端

### 技术领域

[0001] 本发明涉及互联网技术领域,尤其涉及一种网页浏览方法、装置及移动终端。

### 背景技术

[0002] 互联网已成为人们生活中不可缺少的部分。传统的互联网采用 PC (Personal Computer, 个人计算机) 机接入,通过 PC 访问互联网。随着移动终端技术的蓬勃发展,人们已能够通过移动终端无时无刻、随时随地的访问互联网。

[0003] 移动终端访问互联网与传统的 PC 访问互联网有着极大的区别,首先:移动终端的操作平台类型丰富,比如:安卓(Android)平台、赛班(Symbian)平台、苹果(Apple)平台等等,各平台所支持的浏览器设备存在差异,使得各种移动终端无法统一浏览器设备,从而使传统的 HTML (Hypertext Markup Language, 超文本标记语言)网页无法在所有类型的移动终端中进行很好的展示。其次:移动终端的配置通常不及 PC 机,比如:一些非智能手机或普通的智能手机,其配置往往较低,此类移动终端对于一些内容较为丰富的 HTML 网页无法很好的支持。再次,移动终端通常采用无线网络接入,无线网络与传统的有线网络相比,具有带宽低、网速慢的特点,因此对于传统的 HTML 网页(大小通常为几十 K),移动终端的浏览效果非常差。

[0004] 针对上述移动终端无法很好地支持 HTML 网页的问题,目前存在一些解决方案,主要包括:

(1) 将 WAP (Wireless Application Protocol, 无线应用协议) 网页不支持的 HTML 标签进行删除和替换,从而将 HTML 网页转换成标准的 WAP 网页,以便移动终端可以正常访问。由于移动终端的显示界面相比于 PC 机较小,这种方案仅对 HTML 网页的标签进行处理,无法实现网页的压缩和页面核心内容的展现,使得用户无法从移动终端显示的 WAP 网页中直接查看到有用信息。

[0005] (2) 将 HTML 网页中的多媒体信息进行删除,使移动终端能够实现对 HTML 网页的浏览。这种方案使移动终端显示的 HTML 网页缺少丰富的多媒体信息,降低了网页的可读性,降低了用户的阅读兴趣。

[0006] (3) 对 HTML 网页进行剪裁,压缩原始网页,节省传输带宽。这种方案并未对网页的标签、多媒体信息进行处理,使得许多移动终端无法正常显示网页,降低用户的浏览体验。

### 发明内容

[0007] 本发明实施例所要解决的技术问题在于,提供一种网页浏览方法、装置及移动终端,可节省传输带宽,提高网页请求的响应速度,同时提升网页浏览效果和用户的浏览体验。

[0008] 为了解决上述技术问题,本发明实施例提供了一种网页浏览方法,包括:

当接收到用户的网页请求时,下载所请求的网页的页面信息;

对所述页面信息进行分析处理,获得所请求的网页的页面核心内容;

对所述页面核心内容进行排版处理,并显示排版处理后的所请求的网页。

[0009] 相应地,本发明实施例还提供了一种网页浏览装置,包括:

网页下载模块,用于当接收到用户的网页请求时,下载所请求的网页的页面信息;

网页分析模块,用于对所述网页下载模块下载的页面信息进行分析处理,获得所请求的网页的页面核心内容;

排版处理模块,用于对所述网页分析模块获得的页面核心内容进行排版处理;

显示模块,用于显示所述排版处理模块处理后的所述所请求的网页。

[0010] 相应地,本发明实施例还提供了一种移动终端,包括网页浏览装置,所述网页浏览装置包括:

网页下载模块,用于当接收到用户的网页请求时,下载所请求的网页的页面信息;

网页分析模块,用于对所述网页下载模块下载的页面信息进行分析处理,获得所请求的网页的页面核心内容;

排版处理模块,用于对所述网页分析模块获得的页面核心内容进行排版处理;

显示模块,用于显示所述排版处理模块处理后的所述所请求的网页。

[0011] 实施本发明实施例,具有如下有益效果:

本发明实施例对用户所请求的网页进行分析处理,获得所请求的网页的页面核心内容,对该页面核心内容进行排版处理后进行展示;由于仅提取传统的 HTML 网页的页面核心内容,节省了传输带宽,提高了网页请求的响应速度;由于仅对页面核心内容进行排版处理,使得移动终端能够正常的展示排版处理后的网页,同时使用户能够从移动终端显示的网页中直接查看到有用信息(即页面核心内容),提升了网页浏览效果,提升了用户的浏览体验。

## 附图说明

[0012] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0013] 图 1 为本发明的网页浏览装置的第一实施例的结构示意图;

图 2 为本发明的网页浏览装置的第二实施例的结构示意图;

图 3 为本发明的网页浏览装置中的多媒体处理模块的实施例的结构示意图;

图 4 为本发明的网页浏览装置中的网页下载模块的实施例的结构示意图;

图 5 为本发明的网页浏览装置中的网页分析模块的第一实施例的结构示意图;

图 6 为本发明的网页浏览装置中的网页分析模块的第二实施例的结构示意图;

图 7 为本发明的网页浏览装置中的排版处理模块的实施例的结构示意图;

图 8 为本发明的网页浏览方法的第一实施例的流程图;

图 9 为本发明的网页浏览方法的第二实施例的流程图;

图 10 为本发明的网页浏览方法的应用的第一示意图;

图 11 为本发明的网页浏览方法的应用的第二示意图。

## 具体实施方式

[0014] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0015] 本发明实施例公开了一种移动终端,该移动终端包括网页浏览装置,所述网页浏览装置能够为移动终端用户提供较好的网页浏览服务。需要说明的是,所述移动终端可以为:手机、智能手机、PDA(Personal Digital Assistant,个人数字助理)、便携式平板电脑等终端;所述网页浏览装置可以为:浏览器设备。

[0016] 为了更清楚地说明本发明,下面将对本发明的移动终端中的网页浏览装置进行详细介绍。

[0017] 请参见图1,为本发明的网页浏览装置的第一实施例的结构示意图;所述网页浏览装置包括:

网页下载模块101,用于当接收到用户的网页请求时,下载所请求的网页的页面信息。

[0018] 所述网页浏览装置提供人机交互的界面,用户可以在该界面中发出网页请求,比如:所述网页浏览装置为用户提供地址输入栏,用户可在该地址输入栏中输入所请求的网页的URL(Uniform / Universal Resource Locator,统一资源定位符),并在确认之后向所述网页浏览装置发出网页请求,该网页请求中携带用户输入的URL;再如:所述网页浏览装置为用户显示某个网页,当用户在当前网页中点击其他网页的链接时,所述网页浏览装置接收到用户的网页请求,该网页请求中携带用户所请求的网页的链接地址(URL)。当所述网页浏览装置接收到用户的网页请求时,所述网页下载模块101根据该网页请求,从互联网中下载所请求的网页的页面信息,其中,所述页面信息包括:HTML源网页和相关附件;所述HTML源网页可以包括:HTML源码信息、HTML标签和多媒体信息的源地址;所述相关附件包括:CSS(Cascading Style Sheet,层叠样式表)、Frame(一种控件)及JS(JavaScript,一种网页脚本语言)中的任一种或多种。

[0019] 网页分析模块102,用于对所述网页下载模块101下载的页面信息进行分析处理,获得所请求的网页的页面核心内容。

[0020] 所述网页分析模块102对所述网页下载模块101下载的HTML源网页和相关附件进行分析处理,该分析处理的过程可以为:对HTML源网页中的HTML标签进行处理,得到页面标题;和/或对HTML源网页中的页面核心正文进行提取;和/或对HTML源网页中的页面核心多媒体信息进行提取。经所述网页分析模块102分析处理后,可以获得所请求的网页的页面核心内容。可以理解的是,除所述网页分析模块102分析获得的所述页面核心内容之外,所请求的网页中的其他内容即为噪声信息。

[0021] 排版处理模块103,用于对所述网页分析模块102获得的页面核心内容进行排版处理。

[0022] 所述排版处理模块103对所述网页分析模块102获得的页面核心内容进行排版处理,该页面核心内容包括:页面标题、页面核心正文和页面核心多媒体信息。所述排版处理模块103的处理过程可以包括:依据HTML源网页和相关附件,按照所述页面核心内容在原始的所请求的网页的位置和显示方式进行排版;依据所述网页浏览装置所支持的显示界



面的要求,对排版形成的 HTML 网页进行转换处理,得到所述网页浏览装置所支持的网页格式,比如:HTML 网页、WAP1.0 网页格式或 WAP2.0 网页格式。

[0023] 需要说明的是,所述排版处理模块 103 在对页面核心内容进行排版处理的过程中,对于 HTML 网页中除所述页面核心内容之外的其他噪声信息,可以根据实际需要进行丢弃、折叠或隐藏等处理。

[0024] 显示模块 104,用于显示所述排版处理模块 103 处理后的所述所请求的网页。

[0025] 由于所述排版处理模块 103 进行处理后,所述所请求的网页被转换为所述网页浏览装置所支持网页格式,所述显示模块 104 可直接显示所述所请求的网页。可以理解的是,所述显示模块 104 可以在所述网页浏览装置所提供的显示界面中显示所请求的网页,比如:在所述网页浏览装置所在的移动终端的显示界面中显示所请求的网页;所述显示模块 104 还可以在所述网页浏览装置所支持的显示界面中显示所请求的网页,比如:在与所述网页浏览装置相连接的外部显示界面中显示所请求的网页。

[0026] 本发明实施例可节省传输带宽,提高了网页请求的响应速度;同时提升了网页浏览效果和用户的浏览体验。

[0027] 请参见图 2,为本发明的网页浏览装置的第二实施例的结构示意图;与上一实施例相同,所述网页浏览装置包括:网页下载模块 101、网页分析模块 102、排版处理模块 103 和显示模块 104。本实施例中,所述网页浏览装置进一步包括:

多媒体下载模块 105,用于根据所述页面核心多媒体信息对应的源地址,从互联网中下载所述页面核心多媒体信息。

[0028] 如前述,所述网页分析模块 102 可以获得所请求的网页的页面核心多媒体信息,所述多媒体下载模块 105 可以从 HTML 源网页中读取该页面核心多媒体信息对应源地址,并根据该源地址,从互联网中下载所述页面核心多媒体信息。其中,所述页面核心多媒体信息为所述页面核心正文相关的多媒体信息,包括但不限于:图片、视频、Flash 中的任一种或多种。

[0029] 多媒体处理模块 106,用于按照显示界面的要求,对所述多媒体下载模块 105 下载的页面核心多媒体信息进行缩略处理,获得缩略图并存储,保存所述缩略图的存储地址。

[0030] 由于所述网页浏览装置所提供或所支持的显示界面的大小有限,比如:所述网页浏览装置所在的移动终端的显示界面的大小远小于传统 PC 的显示界面大小,所述多媒体下载模块 105 下载的页面核心多媒体信息不适于在所述移动终端的显示界面中进行显示;同时,为了提高多媒体信息的传输速率,提高网页响应速度,所述页面核心多媒体信息需要经过所述多媒体处理模块 106 的缩略处理,该处理过程可以为:若所述页面核心多媒体信息为图片,则按照显示界面的长度和宽度,对所述图片进行缩略处理形成缩略图并存储,保存所述缩略图的存储地址;若所述页面核心多媒体信息为视频或 Flash,则首先提取所述视频或 Flash 的至少一帧关键帧图像;然后按照显示界面的长度和宽度,对所述关键帧图像进行缩略处理形成缩略图并存储,保存该缩略图的存储地址;其中,所述多媒体处理模块 106 获得的缩略图的长度不大于所述显示界面的长度,所述缩略图的宽度不大于所述显示界面的宽度。

[0031] 具体实现中,请一并参见图 3,为本发明的网页浏览装置中的多媒体处理模块的实施例的结构示意图;所述多媒体处理模块 106 包括:

第一处理单元 611,用于当所述页面核心多媒体信息为图片时,按照显示界面的长度和宽度,对所述图片进行缩略处理形成缩略图。

[0032] 第二处理单元 612,用于当所述页面核心多媒体信息为视频或 Flash 时,提取所述视频或 Flash 的至少一帧关键帧图像,按照显示界面的长度和宽度,对所述关键帧图像进行缩略处理形成缩略图。

[0033] 所述第一处理单元 611 和所述第二处理单元 612 获得的缩略图的长度均不大于所述显示界面的长度,宽度均不大于所述显示界面的宽度。可以理解的是,所述第一处理单元 611 和所述第二处理单元 612 的缩略处理过程可以与现有对图片的缩略处理过程相同,在此不赘述。

[0034] 需要说明的是,所述多媒体处理模块 106 获得的缩略图最终提供给所述排版处理模块 103 进行排版处理后进行显示,既能够保证多媒体显示效果,又可以增加网页的可读性,同时可提高多媒体信息的传输速率,提高网页响应速度。

[0035] 本发明实施例可节省传输带宽,提高了网页请求的响应速度;同时保证网页的可读性,提升了网页浏览效果和用户的浏览体验。

[0036] 为了更清楚地说明本发明,下面将对本发明的网页浏览装置中的网页下载模块进行详细介绍。

[0037] 请参见图 4,为本发明的网页浏览装置中的网页下载模块的实施例的结构示意图;所述网页下载模块 101 包括:

地址获取单元 111,用于当接收到用户的网页请求时,获取所述用户所请求的网页的 URL。

[0038] 如前述,用户可以直接输入所请求的网页的 URL,或通过点击所请求的网页的链接地址,向所述网页浏览装置发送网页请求,该网页请求中携带所请求的网页的 URL,当接收到用户的网页请求时,所述地址获取单元 111 从该网页请求中提取所述用户所请求的网页的 URL。

[0039] 下载单元 112,用于从互联网中下载所述地址获取单元 111 获取的 URL 对应的网页的页面信息。所述下载单元 112 根据所述地址获取单元 111 获取的 URL,从互联网中下载用户所请求的网页的页面信息,包括: HTML 源网页和相关附件;所述 HTML 源网页可以包括: HTML 源码信息、HTML 标签和多媒体信息的源地址;所述相关附件包括:CSS、Frame 及 JS 中的任一种或多种。

[0040] 为了更清楚地说明本发明,下面将对本发明的网页浏览装置中的网页分析模块进行详细介绍。

[0041] 请参见图 5,为本发明的网页浏览装置中的网页分析模块的第一实施例的结构示意图;本实施例中,所请求的网页的所述页面核心内容包括:页面标题和页面核心正文,则所述网页分析模块 102 包括:

读取单元 211,用于从所述 HTML 源网页中读取 H 标签(文本标题标签)中的文本和 Title 标签(标题标签)中的文本。

[0042] 其中,所述 H 标签包括:H1 标签、H2 标签、H3 标签、H4 标签、H5 标签和 H6 标签中的任一种。H 标签(H1 标签-H6 标签)是指 HTML 网页中对文本标题所进行的着重强调的一种标签,以标签 <H1>、<H2>、<H3>、<H4>、<H5> 和 <H6> 依次表示重要性的递减,<H1> 标签代

表最重要的文本标题, <H6> 标签代表相对最不重要的文本标题。优选地, 所述读取单元 211 从所述 HTML 源网页中读取 H1 标签中的文本, 并读取 Title 标签中的文本。其中, Title 标签位于 HTML 源网页的 Head 标签内, 被用来定义页面文档的标题, 主要用于告知访问者该篇文章的主题以及提供给搜索引擎索引; Title 标签与 H 标签(H1 标签 -H6 标签)中的文本可以相同, 也可以不同。

[0043] 第一判断单元 212, 用于判断所述读取单元 211 读取的所述 H 标签中的文本与所述 Title 标签中的文本的相似度是否大于预设的阈值。

[0044] 所述预设的阈值可以根据实际需要进行设定, 该阈值的取值范围可以为 :0-100%, 比如 :可设定阈值为 90%, 所述第一判断单元 212 判断 H1 标签中的文本是否与 Title 标签中的文本达到 90% 以上的相似, 此处可以一个例子进行说明, 假设所述读取单元 211 从所述 H1 标签中读取到 11 个字的文本, 从所述 Title 也读取到 11 个字的文本, 所述第一判断单元 212 判断 H1 标签中的文本有 10 个字与所述 Title 中的文本相同, 则判断二者相似度大于 90%。可以理解的是, 所述预设的阈值还可以为其他值, 所述第一判断单元 212 还可以采用其他方式进行相似度的判断, 其他情况下可类似分析, 在此不赘述。

[0045] 标题确定单元 213, 用于当所述第一判断单元 212 判断结果为是时, 将所述 H 标签中的文本确定为所请求的网页的页面标题。

[0046] 如果所述第一判断单元 212 判断 H1 标签中的文本与所述 Title 标签中的文本相似度超过预设的阈值, 所述标题确定单元 213 将 H1 标签中的文本确定为所请求的网页的页面标题。

[0047] 识别单元 214, 用于从所述标题确定单元 213 确定的页面标题在所述 HTML 源网页中所在位置开始, 识别所述 HTML 源网页中的 HTML 源码信息。

[0048] 实际应用中, 页面标题在 HTML 源网页中的位置一般为所请求的该网页的页面核心正文的开头, 当所述标题确定单元 213 确定了页面标题后, 所述识别单元 214 则从所述页面标题在所述 HTML 源网页中的所在位置开始识别所述 HTML 源网页中的 HTML 源码信息。

[0049] 正文确定单元 215, 用于当所述识别单元 214 识别到与预设的结尾标记相匹配的 HTML 源码信息时, 将所述 HTML 源网页中所述页面标题所在位置至所述匹配的 HTML 源码信息所在位置之间的部分, 确定为所请求的网页的页面核心正文。

[0050] 其中, 所述预设的结尾标记包括但不限于 :“相关评论”标识、“相关链接”标识、“相关阅读”标识、“相关推荐”标识中的任一种或多种。一般地, 当网页中出现“相关评论”或“相关链接”的关键字时, 一般表明该网页的页面核心正文的内容已结束, 因此, 出现“相关评论”或“相关链接”的关键字的位置可以确定为该网页的页面核心正文的结尾。当所述识别单元 214 从所述 HTML 源网页中识别到与预设的结尾标记相匹配的 HTML 源码信息时, 所述正文确定单元 215 将所述 HTML 源网页中所述页面标题所在位置(页面核心正文的开头)至所述匹配的 HTML 源码信息所在位置(页面核心正文的结尾)之间的部分, 确定为所请求的网页的页面核心正文。

[0051] 需要说明的是, 本实施例中所述网页分析模块 102 对页面标题和页面核心正文的分析过程仅为举例, 其他情况下, 比如 :所述网页分析模块 102 还可采用朴素贝耶斯分类器、分支定界、阈值分支等方法来确定页面标题和页面核心正文 ;再如 :所述网页分析模块 102 还可先对 HTML 源网页建立 DOM (Document Object Model, 文档对象模型) 树, 基于 DOM

树来实现页面标题和页面核心正文的确定；又如：所述网页分析模块 102 还可对 HTML 源网页进行分块，对重要的网页块进行识别来查找页面标题和页面核心正文，等等，采用上述其他情况均是为了获得页面标题和页面核心正文，在此不赘述。

[0052] 请参见图 6，为本发明的网页浏览装置中的网页分析模块的第二实施例的结构示意图；与上一实施例相同，所述网页分析模块 102 包括：读取单元 211、第一判断单元 212、标题确定单元 213、识别单元 214、正文确定单元 215。本实施例中，所请求的网页的所述页面核心内容包括：页面标题、页面核心正文和页面核心多媒体信息，则所述网页分析模块 102 进一步包括：

提取单元 216，用于从所述 HTML 源网页中提取位于所述正文确定单元 215 确定的页面核心正文内的多媒体信息的源地址。

[0053] 其中，所述页面核心多媒体信息为所述页面核心正文相关的多媒体信息，包括但不限于：图片、视频、Flash 中的任一种或多种。传统的 HTML 网页中的多媒体信息较多，但大部分为该网页的页面核心正文无关的噪声信息，比如：广告图片、广告 Flash 等，这些噪声信息一些位于 HTML 网页中的页面核心正文内，一些则仅仅是环绕在页面核心正文周边，所述提取单元 216 过滤掉环绕在页面核心正文周边的多媒体信息，从所述 HTML 源网页中提取位于所述页面核心正文内的多媒体信息的源地址。

[0054] 第二判断单元 217，用于判断所述提取单元 216 提取的所述源地址是否包含广告标识。

[0055] 由于噪声信息对应的源地址本身会携带广告标识，包括：广告提供商的信息、广告提示信息等。如前述，由于所述页面核心正文内的多媒体信息也有可能是噪声信息，所述第二判断单元 217 进一步判断所述提取单元 216 从所述页面核心正文内提取的多媒体信息的源地址是否包含广告标识，若判断结果为是，则包含广告标识的源地址对应的多媒体信息应当为噪声信息；若判断结果为否，则未包含广告标识的源地址对应的多媒体信息应当为页面核心正文相关的多媒体信息。

[0056] 多媒体确定单元 218，用于根据所述第二判断单元 217 的判断结果，将未包含广告标识的源地址对应的多媒体信息确定为页面核心多媒体信息。

[0057] 为了更清楚地说明本发明，下面将对本发明的网页浏览装置中的排版处理模块进行详细介绍。

[0058] 请参见图 7，为本发明的网页浏览装置中的排版处理模块的实施例的结构示意图；所述排版处理模块 103 包括：

排版单元 311，用于根据所述 HTML 源网页和相关附件，对所述页面核心内容进行排版，形成 HTML 网页。

[0059] 所述排版单元 311 对所述页面核心内容进行排版处理，该页面核心内容包括：页面标题、页面核心正文和页面核心多媒体信息。所述排版单元 311 的处理过程可以包括：依据 HTML 源网页和相关附件，按照所述页面核心内容在原始的所请求的网页的位置和显示方式进行排版，形成 HTML 网页；比如：HTML 源网页中记载了页面标题在 HTML 源网页中的位置，同时 CSS 中记载了页面标题的字体类型和颜色，所述排版单元 311 依据 HTML 源网页的记录，将页面标题排版在 HTML 网页中相应的位置，同时将该页面标题处理为 CSS 记录的相应字体类型和颜色。所述排版单元 311 的排版处理，使得形成的 HTML 网页与原始所请求的

HTML 网页对页面核心内容的展示效果相同,不损失页面的可读性。

[0060] 转换处理单元 312,用于按照显示界面的要求对所述排版单元 311 形成的 HTML 网页进行转换处理,形成转换后的所请求的网页。

[0061] 所述转换处理单元 312 的转换处理过程可以包括:将所述 HTML 网页中的页面核心多媒体信息对应的源地址替换为相应的缩略图的存储地址;和/或,按照显示界面的格式要求,将所述 HTML 网页中所述显示界面不支持的 HTML 标签进行替换或删除,形成 WAP 网页;其中,所述 WAP 网页为 WAP1.0 网页或 WAP2.0 网页。具体实现中,所述转换处理单元 312 基于所述网页浏览装置所支持或所提供的显示界面的要求进行转换处理,比如:若所述网页浏览装置所在的移动终端可以支持 HTML 网页,但无法很好的支持原始 HTML 网页的多媒体信息,所述转换处理单元 312 的转换处理过程可以仅包括:将所述排版后的 HTML 网页中的页面核心多媒体信息对应的源地址替换为相应的缩略图的存储地址,最终形成 HTML 网页提供给移动终端进行显示;再如:若所述网页浏览装置所在的移动终端无法很好的支持 HTML 网页,但其显示界面功能较为强大,可以支持多媒体信息,则所述转换处理单元 312 可以仅仅按照显示界面的格式要求,将所述 HTML 网页中所述显示界面不支持的 HTML 标签进行替换或删除,最终形成 WAP1.0 网页或 WAP2.0 网页提供给移动终端进行显示;又如:若所述网页浏览装置所在的移动终端无法很好的支持 HTML 网页,仅支持 WAP1.0 网页或 WAP2.0 网页,则所述转换处理单元 312 首先将所述 HTML 网页中的页面核心多媒体信息对应的源地址替换为相应的缩略图的存储地址;然后按照显示界面的格式要求,将所述 HTML 网页中所述显示界面不支持的 HTML 标签进行替换或删除,最终形成 WAP1.0 网页或 WAP2.0 网页提供给移动终端进行显示。

[0062] 需要说明的是,本发明实施例的网页浏览装置不仅可以应用于本发明实施例所述的移动终端中,还可应用于 PC 机等传统的固定终端中,当应用于 PC 机等传统终端中时,由于仅对网页的页面核心内容进行排版展示,因此能够提高 PC 机对网页的展示速度,从而提高了网页请求的响应速率,提高了用户的浏览体验。

[0063] 本发明实施例对用户所请求的网页进行分析处理,获得所请求的网页的页面核心内容,对该页面核心内容进行排版处理后进行展示;由于仅提取传统的 HTML 网页的页面核心内容,节省了传输带宽,提高了网页请求的响应速度;由于仅对页面核心内容进行排版处理,使得移动终端能够正常的展示排版处理后的网页,同时使用户能够从移动终端显示的网页中直接查看到有用信息(即页面核心内容),提升了网页浏览效果,提升了用户的浏览体验。

[0064] 为了更清楚地说明本发明,下面将对本发明的网页浏览装置所执行的网页浏览方法进行详细介绍。

[0065] 请参见图 8,为本发明的网页浏览方法的第一实施例的流程图;本实施例中,所述方法的执行主体为上述的网页浏览装置,所述网页浏览装置按以下流程执行所述方法,包括:

S101,当接收到用户的网页请求时,下载所请求的网页的页面信息。

[0066] 所述网页浏览装置提供人机交互的界面,用户可以在该界面中发出网页请求,比如:所述网页浏览装置为用户提供地址输入栏,用户可在该地址输入栏中输入所请求的网页的 URL,并在确认之后向所述网页浏览装置发出网页请求,该网页请求中携带用户输入的

URL ;再如 :所述网页浏览装置为用户显示某个网页,当用户在当前网页中点击其他网页的链接时,所述网页浏览装置接收到用户的网页请求,该网页请求中携带用户所请求的网页的链接地址(URL)。步骤 S101 中,当所述网页浏览装置接收到用户的网页请求时,根据该网页请求,从互联网中下载所请求的网页的页面信息,其中,所述页面信息包括: HTML 源网页和相关附件 ;所述 HTML 源网页可以包括:HTML 源码信息、HTML 标签和多媒体信息的源地址 ;所述相关附件包括:CSS、Frame 及 JS 中的任一种或多种。

[0067] S102,对所述页面信息进行分析处理,获得所请求的网页的页面核心内容。

[0068] 步骤 S102 中,对步骤 S101 下载的 HTML 源网页和相关附件进行分析处理,该分析处理的过程可以为:对 HTML 源网页中的 HTML 标签进行处理,得到页面标题 ;和 / 或对 HTML 源网页中的页面核心正文进行提取 ;和 / 或对 HTML 源网页中的页面核心多媒体信息进行提取。经步骤 S102 分析处理后,可以获得所请求的网页的页面核心内容。可以理解的是,除步骤 S102 分析获得的所述页面核心内容之外,所请求的网页中的其他内容即为噪声信息。

[0069] S103,对所述页面核心内容进行排版处理,并显示排版处理后的所请求的网页。

[0070] 步骤 S103 中,对步骤 S102 获得的页面核心内容进行排版处理,该页面核心内容包括:页面标题、页面核心正文和页面核心多媒体信息。步骤 S103 的处理过程可以包括:依据 HTML 源网页和相关附件,按照所述页面核心内容在原始的所请求的网页的位置和显示方式进行排版 ;依据所述网页浏览装置所支持的显示界面的要求,对排版形成的 HTML 网页进行转换处理,得到所述网页浏览装置所支持的网页格式,比如:HTML 网页、WAP1.0 网页格式或 WAP2.0 网页格式。需要说明的是,步骤 S103 在对页面核心内容进行排版处理的过程中,对于 HTML 网页中除所述页面核心内容之外的其他噪声信息,可以根据实际需要进行丢弃、折叠或隐藏等处理。

[0071] 由于排版处理后,所述所请求的网页被转换为所述网页浏览装置所支持网页格式,可直接显示所述所请求的网页。可以理解的是,步骤 S103 可以在所述网页浏览装置所提供的显示界面中显示所请求的网页,比如:在所述网页浏览装置所在的移动终端的显示界面中显示所请求的网页 ;步骤 S103 还可以在所述网页浏览装置所支持的显示界面中显示所请求的网页,比如:在与所述网页浏览装置相连接的外部显示界面中显示所请求的网页。

[0072] 本发明实施例可节省传输带宽,提高了网页请求的响应速度 ;同时提升了网页浏览效果和用户的浏览体验。

[0073] 请参见图 9,为本发明的网页浏览方法的第二实施例的流程图 ;本实施例中,所述方法的执行主体为上述的网页浏览装置,所述网页浏览装置按以下流程执行所述方法,所述方法包括:

S201,当接收到用户的网页请求时,获取所述用户所请求的网页的 URL。

[0074] 用户可以直接输入所请求的网页的 URL,或通过点击所请求的网页的链接地址,向所述网页浏览装置发送网页请求,该网页请求中携带所请求的网页的 URL,步骤 S201 中,当接收到用户的网页请求时,从该网页请求中提取所述用户所请求的网页的 URL。

[0075] S202,从互联网中下载所述 URL 对应的网页的页面信息。

[0076] 步骤 S102 中,根据步骤 S201 获取的 URL,从互联网中下载用户所请求的网页的页面信息,包括: HTML 源网页和相关附件 ;所述 HTML 源网页包括:HTML 源码信息、HTML 标签

和多媒体信息的源地址 ;所述相关附件包括 :CSS、Frame 及 JS 中的任一种或多种。

[0077] 本实施例中,步骤 S201- 步骤 S202 为上一实施例中的步骤 S101 的具体细化流程。

[0078] S203,从所述 HTML 源网页中读取 H 标签中的文本和 Title 标签中的文本。

[0079] 其中,所述 H 标签包括 :H1 标签、H2 标签、H3 标签、H4 标签、H5 标签和 H6 标签中的任一种。H 标签(H1 标签 -H6 标签)是指 HTML 网页中对文本标题所进行的着重强调的一种标签,以标签 <H1>、<H2>、<H3>、<H4>、<H5> 和 <H6> 依次表示重要性的递减,<H1> 标签代表最重要的文本标题,<H6> 标签代表相对最不重要的文本标题。优选地,步骤 S203 中,从所述 HTML 源网页中读取 H1 标签中的文本,并读取 Title 标签中的文本。其中,Title 标签位于 HTML 源网页的 Head 标签内,被用来定义页面文档的标题,主要用于告知访问者该篇文章的主题以及提供给搜索引擎索引;Title 标签与 H 标签(H1 标签 -H6 标签)中的文本可以相同,也可以不同。

[0080] S204,判断所述 H 标签中的文本与所述 Title 标签中的文本的相似度是否大于预设的阈值 ;若判断结果为是,转入执行步骤 S205 ;否则,结束。

[0081] 所述预设的阈值可以根据实际需要进行设定,该阈值的取值范围可以为 :0-100%,比如 :可设定阈值为 90%,步骤 S204 中,判断 H1 标签中的文本是否与 Title 标签中的文本达到 90% 以上的相似,此处可以一个例子进行说明,假设步骤 S203 从所述 H1 标签中读取到 11 个字的文本,从所述 Title 也读取到 11 个字的文本,步骤 S204 中,判断 H1 标签中的文本有 10 个字与所述 Title 中的文本相同,则判断二者相似度大于 90%。可以理解的是,所述预设的阈值还可以为其他值,步骤 S204 中还可以采用其他方式进行相似度的判断,其他情况下可类似分析,在此不赘述。

[0082] S205,将所述 H 标签中的文本确定为所请求的网页的页面标题。

[0083] 如果步骤 S204 中判断 H1 标签中的文本与所述 Title 标签中的文本相似度超过预设的阈值,步骤 S205 中,则将 H1 标签中的文本确定为所请求的网页的页面标题。

[0084] S206,从所述页面标题在所述 HTML 源网页中所在位置开始,识别所述 HTML 源网页中的 HTML 源码信息。

[0085] 实际应用中,页面标题在 HTML 源网页中的位置一般为所请求的该网页的页面核心正文的开头,当步骤 S205 确定了页面标题后,步骤 S206 中,从所述页面标题在所述 HTML 源网页中的所在位置开始识别所述 HTML 源网页中的 HTML 源码信息。

[0086] S207,当识别到与预设的结尾标记相匹配的 HTML 源码信息时,将所述 HTML 源网页中所述页面标题所在位置至所述匹配的 HTML 源码信息所在位置之间的部分,确定为所请求的网页的页面核心正文。

[0087] 其中,所述预设的结尾标记包括但不限于 :“相关评论”标识、“相关链接”标识、“相关阅读”标识、“相关推荐”标识中的任一种或多种。一般地,当网页中出现“相关评论”或“相关链接”的关键字时,一般表明该网页的页面核心正文的内容已结束,因此,出现“相关评论”或“相关链接”的关键字的位置可以确定为该网页的页面核心正文的结尾。步骤 S207 中,当从所述 HTML 源网页中识别到与预设的结尾标记相匹配的 HTML 源码信息时,将所述 HTML 源网页中所述页面标题所在位置(页面核心正文的开头)至所述匹配的 HTML 源码信息所在位置(页面核心正文的结尾)之间的部分,确定为所请求的网页的页面核心正文。

[0088] S208,从所述 HTML 源网页中提取位于所述页面核心正文内的多媒体信息的源地

址。

[0089] 其中,所述页面核心多媒体信息为所述页面核心正文相关的多媒体信息,包括但不限于:图片、视频、Flash 中的任一种或多种。传统的 HTML 网页中的多媒体信息较多,但大部分为该网页的页面核心正文无关的噪声信息,比如:广告图片、广告 Flash 等,这些噪声信息一些位于 HTML 网页中的页面核心正文内,一些则仅仅是环绕在页面核心正文周边,步骤 S208 中,过滤掉环绕在页面核心正文周边的多媒体信息,从所述 HTML 源网页中提取位于所述页面核心正文内的多媒体信息的源地址。

[0090] S209,判断提取的所述源地址是否包含广告标识;若判断结果为否,转入执行步骤 S210;否则,结束。

[0091] 由于噪声信息对应的源地址本身会携带广告标识,包括:广告提供商的信息、广告提示信息等。如前述,由于所述页面核心正文内的多媒体信息也有可能是噪声信息,步骤 S209 在步骤 S208 的基础上,进一步判断从所述页面核心正文内提取的多媒体信息的源地址是否包含广告标识,若判断结果为是,则包含广告标识的源地址对应的多媒体信息应当为噪声信息;若判断结果为否,则未包含广告标识的源地址对应的多媒体信息应当为页面核心正文相关的多媒体信息。

[0092] S210,将未包含广告标识的源地址对应的多媒体信息确定为页面核心多媒体信息。

[0093] 本实施例中,步骤 S203-步骤 S210 为上一实施例中的步骤 S102 的具体细化流程。

[0094] S211,根据所述页面核心多媒体信息对应的源地址,从互联网中下载所述页面核心多媒体信息。

[0095] 当步骤 S210 确定了请求的网页的页面核心多媒体信息之后,步骤 S211 中,可以从 HTML 源网页中读取该页面核心多媒体信息对应源地址,并根据该源地址,从互联网中下载所述页面核心多媒体信息。其中,所述页面核心多媒体信息为所述页面核心正文相关的多媒体信息,包括但不限于:图片、视频、Flash 中的任一种或多种。

[0096] S212,按照显示界面的要求,对所述页面核心多媒体信息进行缩略处理,获得缩略图并存储,保存所述缩略图的存储地址。

[0097] 由于所述网页浏览装置所提供或所支持的显示界面的大小有限,比如:所述网页浏览装置所在的移动终端的显示界面的大小远小于传统 PC 的显示界面大小,步骤 S211 下载的页面核心多媒体信息不适于在所述移动终端的显示界面中进行显示;同时,为了提高多媒体信息的传输速率,提高网页响应速度,所述页面核心多媒体信息需要经过缩略处理,步骤 S212 中,该缩略处理过程可以为:若所述页面核心多媒体信息为图片,则按照显示界面的长度和宽度,对所述图片进行缩略处理形成缩略图并存储,保存所述缩略图的存储地址;若所述页面核心多媒体信息为视频或 Flash,则首先提取所述视频或 Flash 的至少一帧关键帧图像;然后按照显示界面的长度和宽度,对所述关键帧图像进行缩略处理形成缩略图并存储,保存该缩略图的存储地址;其中,所述缩略图的长度不大于所述显示界面的长度,所述缩略图的宽度不大于所述显示界面的宽度。

[0098] S213,根据所述 HTML 源网页和相关附件,对所述页面核心内容进行排版,形成 HTML 网页。

[0099] 步骤 S213 中,对所述页面核心内容进行排版处理,该页面核心内容包括:页面标



题、页面核心正文和页面核心多媒体信息。步骤 S213 的处理过程可以包括：依据 HTML 源网页和相关附件，按照所述页面核心内容在原始的所请求的网页的位置和显示方式进行排版，形成 HTML 网页；比如：HTML 源网页中记载了页面标题在 HTML 源网页中的位置，同时 CSS 中记载了页面标题的字体类型和颜色，步骤 S213 依据 HTML 源网页的记录，将页面标题排版在 HTML 网页中相应的位置，同时将该页面标题处理为 CSS 记录的相应字体类型和颜色。步骤 S213 中，对所述页面核心内容的排版处理，使得形成的 HTML 网页与原始所请求的 HTML 网页对页面核心内容的展示效果相同，不损失页面的可读性。

[0100] S214，按照显示界面的要求对所述 HTML 网页进行转换处理，显示转换后的所请求的网页。

[0101] 步骤 S214 中，转换处理过程可以包括：将所述 HTML 网页中的页面核心多媒体信息对应的源地址替换为相应的缩略图的存储地址；和 / 或，按照显示界面的格式要求，将所述 HTML 网页中所述显示界面不支持的 HTML 标签进行替换或删除，形成 WAP 网页；其中，所述 WAP 网页为 WAP1.0 网页或 WAP2.0 网页。具体实现中，步骤 S214 基于所述网页浏览装置所支持或所提供的显示界面的要求进行转换处理，比如：若所述网页浏览装置所在的移动终端可以支持 HTML 网页，但无法很好的支持原始 HTML 网页的多媒体信息，所述转换处理过程可以仅包括：将所述排版后的 HTML 网页中的页面核心多媒体信息对应的源地址替换为相应的缩略图的存储地址，最终形成 HTML 网页提供给移动终端进行显示；再如：若所述网页浏览装置所在的移动终端无法很好的支持 HTML 网页，但其显示界面功能较为强大，可以支持多媒体信息，则所述转换处理过程可以仅包括：按照显示界面的格式要求，将所述 HTML 网页中所述显示界面不支持的 HTML 标签进行替换或删除，最终形成 WAP1.0 网页或 WAP2.0 网页提供给移动终端进行显示；又如：若所述网页浏览装置所在的移动终端无法很好的支持 HTML 网页，仅支持 WAP1.0 网页或 WAP2.0 网页，则所述转换处理过程可以为：首先将所述 HTML 网页中的页面核心多媒体信息对应的源地址替换为相应的缩略图的存储地址；然后按照显示界面的格式要求，将所述 HTML 网页中所述显示界面不支持的 HTML 标签进行替换或删除，最终形成 WAP1.0 网页或 WAP2.0 网页提供给移动终端进行显示。

[0102] 为了更清楚地说明本发明，下面将以一个具体示例来进行说明。

[0103] 请参见图 10，为本发明的网页浏览方法的应用的第一示意图；当用户在地址输入栏中输入：<http://news.qq.com/a/20110719/000584.htm>；或在 QQ 新闻主页中点击该网页的链接时，步骤 S201 则可接收用户的网页请求，并获取到所请求的网页的 URL（即：<http://news.qq.com/a/20110719/000584.htm>）；步骤 S202 依据该 URL，从互联网中下载图 10 所示的 HTML 网页对应的页面信息，包括：HTML 源网页和相关附件。步骤 S203 读取 <H1> 标签中的文本为“云南发生伪造矿难事件，6 人谋害智障流浪汉骗取赔偿”，<Title> 标签中的文本为“云南网 云南伪造矿难事件，6 人谋害智障流浪汉骗赔偿”。步骤 S204 判断步骤 S203 读取的二个文本相似度很高，超过预设的阈值如 90%。步骤 S205 则将 <H1> 标签中的文本“云南发生伪造矿难事件，6 人谋害智障流浪汉骗取赔偿”确定为网页的页面标题，同时也是页面核心正文的开始。步骤 S206 从该页面标题所在位置开始识别 HTML 源网页中的 HTML 源码信息，当识别到图 10 所示的“相关阅读”标识时，步骤 S207 确定该位置为页面核心正文的结尾，则从所述 HTML 源网页中所述页面标题所在位置至所述“相关阅读”标识所在位置之间的部分，为所请求的网页的页面核心正文。

[0104] 步骤 S208- 步骤 S212 将页面核心正文内与页面核心正文相关的图片进行下载、缩略处理, 形成缩略图并存储, 保存存储地址。步骤 S213 将页面标题、页面核心正文和页面核心正文内的页面核心多媒体信息(即图片)按原网页的排布和比例进行排版处理, 最后依照移动终端的显示界面要求, 将图片源地址替换为缩略图存储地址, 并将 HTML 标签进行替换或删除处理后, 形成图 11 所示的 WAP1.0 网页, 在移动终端中进行展示。请一并参见图 11, 为本发明的网页浏览方法的应用的第二示意图, 图 11 即为移动终端最终显示的 WAP1.0 网页的示意图, 该 WAP1.0 网页中仅包括了页面核心内容, 其他的周边的噪声信息均被丢弃、折叠或隐藏, 使用户可以直接读取所请求的网页的核心信息, 提高了移动终端的浏览效果。

[0105] 本发明实施例对用户所请求的网页进行分析处理, 获得所请求的网页的页面核心内容, 对该页面核心内容进行排版处理后进行展示; 由于仅提取传统的 HTML 网页的页面核心内容, 节省了传输带宽, 提高了网页请求的响应速度; 由于仅对页面核心内容进行排版处理, 使得移动终端能够正常的展示排版处理后的网页, 同时使用户能够从移动终端显示的网页中直接查看到有用信息(即页面核心内容), 提升了网页浏览效果, 提升了用户的浏览体验。

[0106] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程, 是可以通过计算机程序来指令相关的硬件来完成, 所述的程序可存储于一计算机可读取存储介质中, 该程序在执行时, 可包括如上述各方法的实施例的流程。其中, 所述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory, ROM) 或随机存储记忆体(Random Access Memory, RAM) 等。

[0107] 以上所揭露的仅为本发明较佳实施例而已, 当然不能以此来限定本发明之权利范围, 本领域普通技术人员可以理解实现上述实施例的全部或部分流程, 并依本发明权利要求所作的等同变化, 仍属于发明所涵盖的范围。

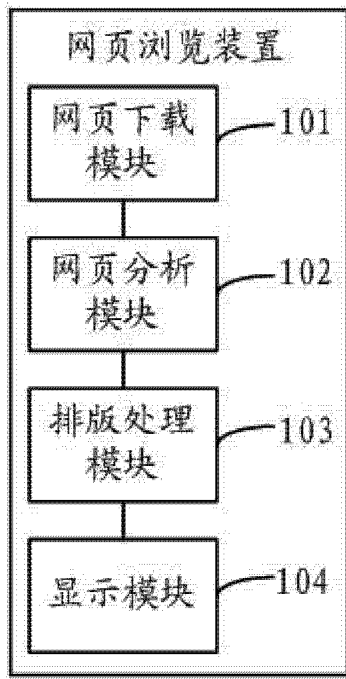


图 1

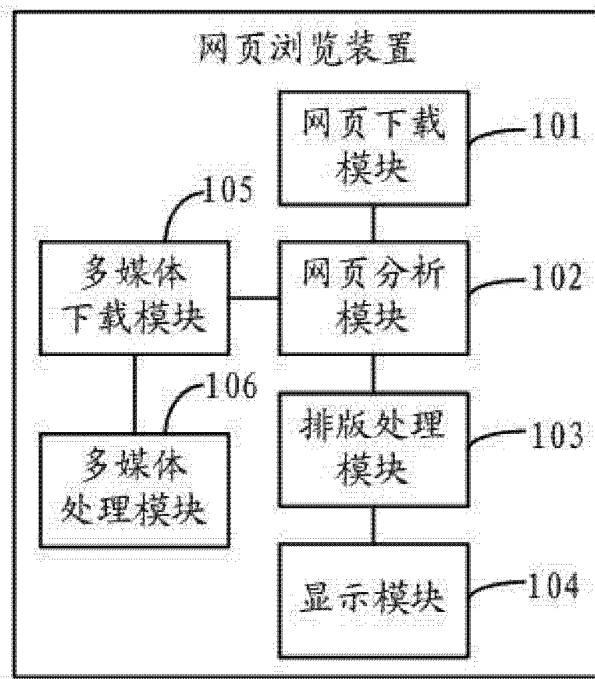


图 2

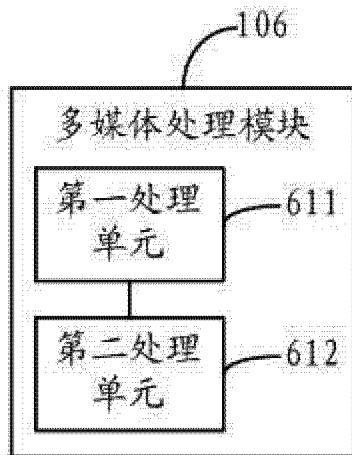


图 3

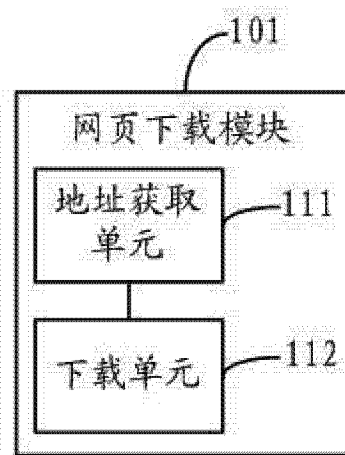


图 4

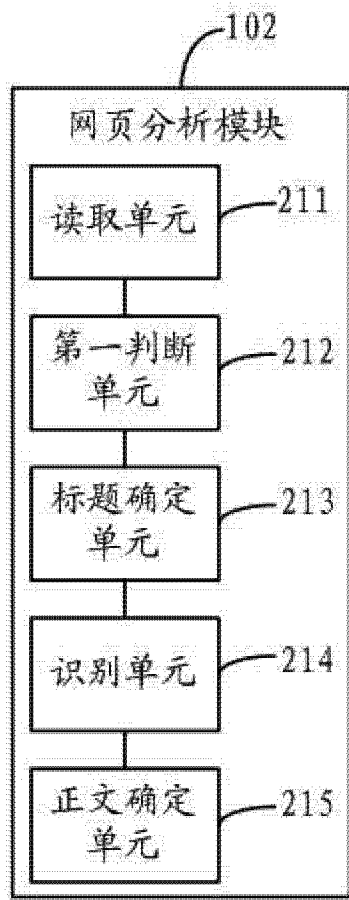


图 5

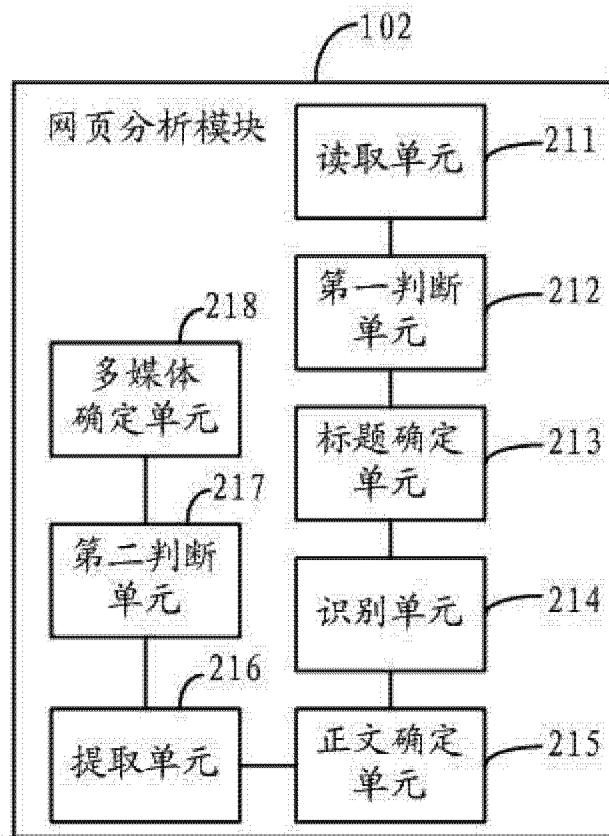


图 6

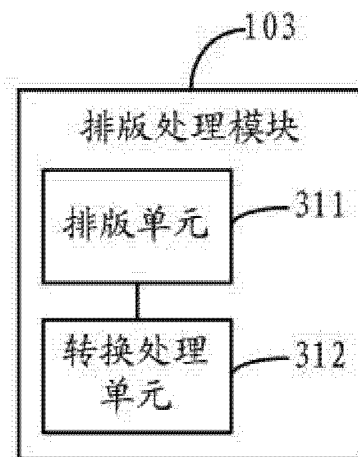


图 7

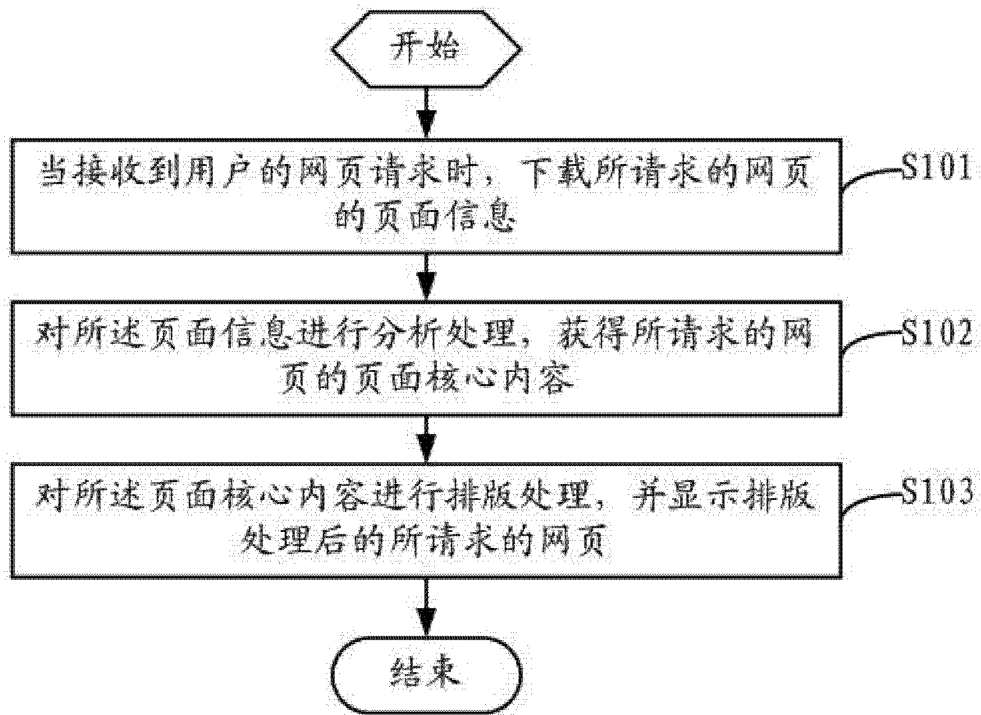


图 8

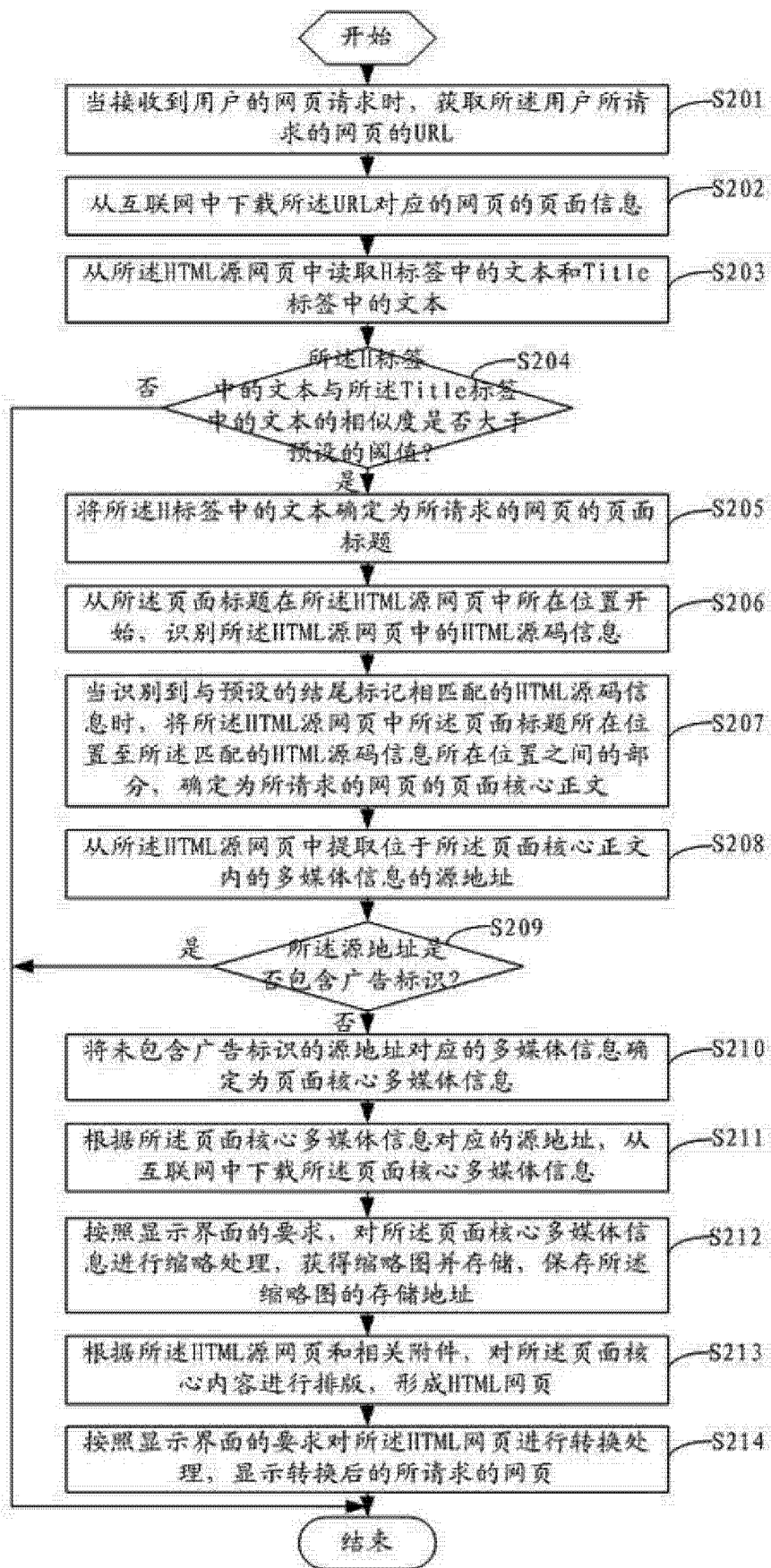


图 9

新闻 视频 图片 评论 娱乐 明星 电影 音乐 汽车 房产 家居 家电 女性 健康 美容 育儿 读书 教育 高考 出国 博客 微博 公益 绿色  
财经 股票 期货 基金 体育 NBA 彩票 大运会 科技 数码 手机 下载 时尚 购物 旅游 小说 游戏 动漫 动画 星座 军事 论坛 儿童 天气

不用跑去银行排队 跨行跨区还免手续费

腾讯新闻 腾讯新闻 社会新闻 社会万象 正文 伪造矿难 搜索

### 云南发生伪造矿难事件 6人谋害智障流浪汉骗赔

2011年07月19日07:16 云南网 中国商报 我要评论(0) 字号: | 打印



流浪汉被打死的地方

昨日，云南昭通市盐津县公安局通报了一起以故意杀人方式骗取煤矿赔偿金的案件，两名矿工在矿井下杀害了一名智障流浪汉后，制造矿难假象，再打电话通知同伙冒充死者的家属和妻子向盐津县滩头乡铜厂沟煤矿（以下简称“铜厂沟煤矿”）索赔。就在双方达成38万元赔偿协议并准备付款时，却被盐津警方识破了。

分享到空间

相关阅读:

- 江西抚州6人杀害智障者 伪造矿难骗取赔偿 2011-07-14
- 云南杨善山矿难谎报被通报 矿主伪造入井记录 2011-04-28

每日推荐

- 李昌奎杀人案一审法官：被告赔偿不积极
- 杭州原副市长许进被执行死刑 总涉案金额3亿元
- 陕西渭南城管与摊贩冲突 200人围观并阻拦执法车
- 男子地窖囚禁两少女藏“性奴”终审被判死刑(图)
- 中国航母年的下水：“瓦良格”号买高次建引发争议
- 中国首架飞船名称网络投票结束 神舟号入选前十名
- 历史 | 1903年苏报案：她清最后一场“文字狱”
- 博客 | “戴笠不靠强奸”揭陈：好少女杀警不判死？
- 深度 | 在中国，谁被成了“精神病”
- 军事 | 印度发展航母道路很智慧值得学习 附博图
- 公益 | 女教师为学生原声呻吟 将病鸡鸡冠症绝期
- 实用 | 头条新闻 公益慈善微博圈 讲堂圈 绿色圈

服务信息

博客·微博

汤镇宗 香港演员

郭文珺 奥运气手枪冠军

- 这届 | 李昌奎案庭审 刘冠群杨松
- 摘译 | 刘冠群刘冠群刘冠群 梅西：秀球秀球
- 陈一冰：福乐2008奥运 吕中：刘冠群

图 10

# 云南发生伪造矿难事件 6人谋害智障流浪汉骗赔

2011年07月19日07:12 云南网 实时报导 我要评论(6)

字号: r | T



流浪汉被打死的现场

昨日，昆明通过盐津县公安局选择了一起以故意杀人方式骗取煤矿赔偿金的案件，两名矿工在矿井不幸害了一名智障流浪汉后，制造矿难假象，再打电话通知同伙冒充死者家属和妻子上同盐津县滩头乡铜厂沟煤矿（以下简称“铜厂沟煤矿”）索赔，就在双方达成88万元赔偿协议并准备付钱时，却被盐津警方识破了。

图 11