

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2016 年 7 月 7 日 (07.07.2016)



(10) 国际公布号
WO 2016/107126 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2015/083395
- (22) 国际申请日: 2015 年 7 月 6 日 (06.07.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410843496.9 2014 年 12 月 30 日 (30.12.2014) CN
- (71) 申请人: 百度在线网络技术(北京)有限公司
(BAIDU ONLINE NETWORK TECHNOLOGY
(BEIJING) CO., LTD.) [CN/CN]; 中国北京市海淀区
上地十街 10 号百度大厦三层, Beijing 100085 (CN)。
- (72) 发明人: 文石磊 (WEN, Shilei); 中国北京市海淀区
上地十街 10 号百度大厦三层, Beijing 100085 (CN)。
李旭斌 (LI, Xubin); 中国北京市海淀区上地十街 10
号百度大厦三层, Beijing 100085 (CN)。
陈世佳 (CHEN, Shijia); 中国北京市海淀区上地十街 10 号
百度大厦三层, Beijing 100085 (CN)。
秦首科 (QIN, Shouke); 中国北京市海淀区上地十街 10 号百度大

厦三层, Beijing 100085 (CN)。
张泽明 (ZHANG, Zeming); 中国北京市海淀区上地十街 10 号百度大厦三层, Beijing 100085 (CN)。
韩友 (HAN, You); 中国北京市海淀区上地十街 10 号百度大厦三层, Beijing 100085 (CN)。
江焱 (JIANG, Yan); 中国北京市海淀区上地十街 10 号百度大厦三层, Beijing 100085 (CN)。
陈志扬 (CHEN, Zhiyang); 中国北京市海淀区上地十街 10 号百度大厦三层, Beijing 100085 (CN)。

(74) 代理人: 北京清亦华知识产权代理事务所(普通合伙)
(TSINGYIHUA INTELLECTUAL PROPERTY LLC); 中国北京市海淀区清华园清华大学照
澜院商业楼 301 室, Beijing 100084 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST,

[见续页]

(54) Title: IMAGE SEARCH METHOD AND DEVICE

(54) 发明名称: 图片搜索方法和装置

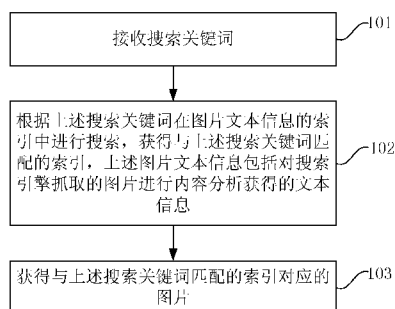


图 1/ Fig. 1

- 101 Receive a search keyword
- 102 Search in indexes of image text information according to the search keyword, and obtain the index matching with the search keyword, wherein the image text information comprises text information that is obtained by performing content analysis on images captured by a search engine
- 103 Obtain the image corresponding to the index matching with the search keyword

(57) Abstract: An image search method and device. The image search method comprises: receiving a search keyword (101); searching in indexes of image text information according to the search keyword, and obtaining the index matching with the search keyword, wherein the image text information comprises text information that is obtained by performing content analysis on images captured by a search engine (102); and obtaining the image corresponding to the index matching with the search keyword (103). By performing content analysis on images and adding description of image text information, more related images can be returned when a user searches for an image, and therefore the user demands can be better satisfied and the user experience is improved.

(57) 摘要: 一种图片搜索方法和装置, 该图片搜索方法包括: 接收搜索关键词(101); 根据所述搜索关键词在图片文本信息的索引中进行搜索, 获得与所述搜索关键词匹配的索引; 所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息(102); 获得与所述搜索关键词匹配的索引对应的图片(103)。通过对图片进行内容分析, 增加图片文本信息的描述, 从而可以在用户搜索时返回更多更相关的图片, 进而可以更好地满足用户的需求, 提高用户体验度。



WO 2016/107126 A1



SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, ZA, ZM, ZW。

HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO,
PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ,
CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE,
SN, TD, TG)。

- (84) **指定国** (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚
(AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT,
BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR,

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

图片搜索方法和装置

相关申请的交叉引用

5 本申请要求百度在线网络技术（北京）有限公司于 2014 年 12 月 30 日提交的、发明名称为“图片搜索方法和装置”的、中国专利申请号“201410843496.9”的优先权。

技术领域

本发明涉及互联网技术领域，尤其涉及一种图片搜索方法和装置。

10

背景技术

图片相关文本信息的获取，是指通过获取图片相关的文本段，作为对图片的描述，为搜索图片和/或理解图片提供更精准的技术服务。

常见的图片相关文本信息的获取方法，是通过分析图片所在网页的文本信息，抽取出
15 若干文本段，作为图片的文本描述。其中，图片所在网页的标题信息、图片的链接文本、
图片的属性文本和图片的环绕文本都是非常有效的描述图片的文本信息，上述信息对于用户搜索图片和/或理解图片是必要的。在传统的图片搜索引擎中，网络设备通常基于用户的搜索关键词，在图片的相关文本信息的描述集合中，利用与搜索关键词匹配等技术来返回给用户一个或者多个搜索结果图片。

20 因此，用户在利用网络设备上网检索图片时，只能基于图片所在网页的周边文本信息来进行检索，于是，对于网页文本上缺乏文本描述信息的图片，用户就无法检索到，导致用户需求不能被很好地满足，用户体验较差。

发明内容

25 本发明的目的旨在至少在一定程度上解决相关技术中的技术问题之一。

为此，本发明的第一个目的在于提出一种图片搜索方法。该方法通过对图片进行内容分析，增加图片文本信息的描述，从而可以在用户搜索时返回更多更相关的图片，可以更好地满足用户的需求，提高用户体验度。

本发明的第二个目的在于提出一种图片搜索装置。

30 为了实现上述目的，本发明第一方面实施例的图片搜索方法，包括：接收搜索关键词；根据所述搜索关键词在图片文本信息的索引中进行搜索，获得与所述搜索关键词匹配的索引；所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；获得与

所述搜索关键词匹配的索引对应的图片。

本发明实施例的图片搜索方法，接收搜索关键词之后，根据上述搜索关键词在图片文本信息的索引中进行搜索，获得与上述搜索关键词匹配的索引；进而获得与上述搜索关键词匹配的索引对应的图片，其中上述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；该方法可以通过对图片进行内容分析，增加图片文本信息的描述，从而可以在用户搜索时返回更多更相关的图片，进而可以更好地满足用户的需求，提高用户体验度。

为了实现上述目的，本发明第二方面实施例的图片搜索装置，包括：接收模块，用于接收搜索关键词；搜索模块，用于根据所述接收模块接收的搜索关键词在图片文本信息的索引中进行搜索，获得与所述搜索关键词匹配的索引；以及获得与所述搜索关键词匹配的索引对应的图片；所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；发送模块，用于将所述搜索模块获得的图片发送给所述终端设备。

本发明实施例的图片搜索装置，接收模块接收用户通过终端设备输入的搜索关键词之后，搜索模块根据上述搜索关键词在图片文本信息的索引中进行搜索，获得与上述搜索关键词匹配的索引，进而获得与上述搜索关键词匹配的索引对应的图片，其中上述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；该装置可以通过对图片进行内容分析，增加图片文本信息的描述，从而可以在用户搜索时返回更多更相关的图片，进而可以更好地满足用户的需求，提高用户体验度。

为了实现上述目的，本发明第三方面实施例提供了一种电子设备，包括：一个或者多个处理器；存储器；一个或者多个程序，所述一个或者多个程序存储在所述存储器中，当被所述一个或者多个处理器执行时进行如下操作：接收搜索关键词；根据所述搜索关键词在图片文本信息的索引中进行搜索，获得与所述搜索关键词匹配的索引；所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；获得与所述搜索关键词匹配的索引对应的图片。

为了实现上述目的，本发明第四方面实施例提供了一种非易失性计算机存储介质，所述计算机存储介质存储有一个或者多个模块，当所述一个或者多个模块被执行时进行如下操作：接收搜索关键词；根据所述搜索关键词在图片文本信息的索引中进行搜索，获得与所述搜索关键词匹配的索引；所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；获得与所述搜索关键词匹配的索引对应的图片。

本发明附加的方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本发明的实践了解到。

附图说明

本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

图 1 为本发明图片搜索方法一个实施例的流程图；

5 图 2 (a) ~图 2 (b) 为本发明对图片进行内容分析一个实施例的示意图；

图 3 (a) ~图 3 (b) 为本发明对图片进行内容分析另一个实施例的示意图；

图 4 为本发明图片搜索装置一个实施例的结构示意图；

图 5 为本发明图片搜索装置另一个实施例的结构示意图。

10 具体实施方式

下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，仅用于解释本发明，而不能理解为对本发明的限制。相反，本发明的实施例包括落入所附加权利要求书的精神和内涵范围内的所有变化、修改和等同物。

15 图 1 为本发明图片搜索方法一个实施例的流程图，如图 1 所示，该图片搜索方法可以包括：

步骤 101，接收搜索关键词。

步骤 102，根据上述搜索关键词在图片文本信息的索引中进行搜索，获得与上述搜索关键词匹配的索引，上述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文
20 本信息。

步骤 103，获得与上述搜索关键词匹配的索引对应的图片。

本实施例的一种实现方式中，进一步地，接收搜索关键词之前，搜索引擎还可以对该搜索引擎抓取的图片进行内容分析，获得上述图片的文本信息。

25 本实施例的另一种实现方式中，进一步地，接收搜索关键词之前，搜索引擎还可以对该搜索引擎抓取的图片进行内容分析和对上述图片所在网页的文本进行分析，获得上述图片的文本信息。

具体地，对搜索引擎抓取的图片进行内容分析可以为：对搜索引擎抓取的图片中的物体信息、风景信息、人物信息和文字信息中的至少一种进行分析。也就是从图片内容的角度进行分析，包括但不限于分析图片中的物体信息、风景信息、人物信息和文字信息等信
30 息中的至少一种。

图 2 (a) ~图 2 (b) 为本发明对图片进行内容分析一个实施例的示意图，图 2 (a) 为搜索引擎抓取的图片，图 2 (b) 为对图 2 (a) 所示图片进行内容分析获得的图 2 (a) 所

示图片的文本信息。图 2 (b) 中, “类别” 这一列中的文本为对图 2 (a) 所示图片进行内容分析获得的文本信息, “概率” 这一列中的概率为图 2 (a) 所示图片属于每条文本信息的概率。

图 3 (a) ~图 3 (b) 为本发明对图片进行内容分析另一个实施例的示意图, 图 3 (a) 为搜索引擎抓取的图片, 图 3 (b) 为对图 3 (a) 所示图片进行文字识别获得的图 3 (a) 所示图片的文本信息。图 3 (a) 所示图片中包含文字信息, 对图 3 (a) 所示图片中的文字信息进行识别, 获得图 3 (b) 所示的文本信息, 从图 3 (a) 和图 3 (b) 中可以看出, 由于图片清晰度等因素的影响, 文字识别的结果与原图中的文字信息可能会有一定的误差, 但这并不影响本发明的实现, 所获得的文本信息仍是非常有效的描述图片的文本信息。

具体地, 对上述图片所在网页的文本进行分析可以为: 对上述图片所在网页的结构化字段中的文本进行分析, 上述图片所在网页的结构化字段包括上述网页的网页标题字段、锚文本字段和/或图片属性字段。

上述两种实现方式中, 获得上述图片的文本信息之后, 接收搜索关键词之前, 搜索引擎还可以对上述图片的文本信息进行整理, 建立上述图片的文本信息的索引。

具体地, 搜索引擎对上述图片的文本信息进行整理可以为: 搜索引擎对上述图片的文本信息进行切词等处理。

上述实施例中, 接收搜索关键词之后, 根据上述搜索关键词在图片文本信息的索引中进行搜索, 获得与上述搜索关键词匹配的索引; 进而获得与上述搜索关键词匹配的索引对应的图片, 其中上述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息; 该方法可以通过对图片进行内容分析, 增加图片文本信息的描述, 从而可以在用户搜索时返回更多更相关的图片, 进而可以更好地满足用户的需求, 提高用户体验度。

图 4 为本发明图片搜索装置一个实施例的结构示意图, 本实施例中的图片搜索装置可以作为搜索引擎, 或者搜索引擎的一部分实现本发明图 1 所示实施例的流程, 如图 4 所示, 该图片搜索装置可以包括: 接收模块 41 和搜索模块 42;

其中, 接收模块 41, 用于接收搜索关键词;

搜索模块 42, 用于根据接收模块 41 接收的搜索关键词在图片文本信息的索引中进行搜索, 获得与上述搜索关键词匹配的索引; 以及获得与上述搜索关键词匹配的索引对应的图片; 其中, 上述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息。

上述图片搜索装置中, 接收模块 41 接收搜索关键词之后, 搜索模块 42 根据上述搜索关键词在图片文本信息的索引中进行搜索, 获得与上述搜索关键词匹配的索引, 进而获得与上述搜索关键词匹配的索引对应的图片, 其中上述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息; 该装置可以通过对图片进行内容分析, 增加图片文本

信息的描述，从而可以在用户搜索时返回更多更相关的图片，进而可以更好地满足用户的需求，提高用户体验度。

图 5 为本发明图片搜索装置另一个实施例的结构示意图，与图 4 所示的图片搜索装置相比，不同之处在于，本实施例的一种实现方式中，图 5 所示的图片搜索装置还可以包括：

5 获得模块 43；

获得模块 43，用于在接收模块 41 接收搜索关键词之前，对搜索引擎抓取的图片进行内容分析，获得上述图片的文本信息。

本实施例的另一种实现方式中，图 5 所示的图片搜索装置还可以包括：获得模块 43；

10 获得模块 43，用于在接收模块 41 接收搜索关键词之前，对搜索引擎抓取的图片进行内容分析和对上述图片所在网页的文本进行分析，获得上述图片的文本信息。

具体地，获得模块 43 用于对搜索引擎抓取的图片进行内容分析可以为：获得模块 43，具体用于对搜索引擎抓取的图片中的物体信息、风景信息、人物信息和文字信息中的至少一种进行分析。也就是从图片内容的角度进行分析，包括但不限于分析图片中的物体信息、风景信息、人物信息和文字信息等信息中的至少一种。

15 参见图 2 (a) ~图 2 (b)，图 2 (a) 为搜索引擎抓取的图片，图 2 (b) 为对图 2 (a) 所示图片进行内容分析获得的图 2 (a) 所示图片的文本信息。图 2 (b) 中，“类别”这一列中的文本为对图 2 (a) 所示图片进行内容分析获得的文本信息，“概率”这一列中的概率为图 2 (a) 所示图片属于每条文本信息的概率。

继续参见图 3 (a) ~图 3 (b)，图 3 (a) 为搜索引擎抓取的图片，图 3 (b) 为对图 3 (a) 所示图片进行文字识别获得的图 3 (a) 所示图片的文本信息。图 3 (a) 所示图片中包含文字信息，对图 3 (a) 所示图片中的文字信息进行识别，获得图 3 (b) 所示的文本信息，从图 3 (a) 和图 3 (b) 中可以看出，由于图片清晰度等因素的影响，文字识别的结果与原图中的文字信息可能会有一定的误差，但这并不影响本发明的实现，所获得的文本信息仍是非常有效的描述图片的文本信息。

25 具体地，获得模块 43 用于对上述图片所在网页的文本进行分析可以为：获得模块 43，具体用于对上述图片所在网页的结构化字段中的文本进行分析，上述图片所在网页的结构化字段包括上述网页的网页标题字段、锚文本字段和/或图片属性字段。

本实施例中，图片搜索装置还可以包括：建立模块 44，用于在获得模块 43 获得上述图片的文本信息之后，接收模块 41 接收搜索关键词之前，对上述图片的文本信息进行整理，30 建立上述图片的文本信息的索引。

上述图片搜索装置可以在用户搜索时返回更多更相关的图片，进而可以更好地满足用户的需求，提高用户体验度。

本发明提供的图片搜索方法和装置，一方面，通过分析图片内容可增加图片相关文本信息的描述，从而在搜索图片时可以获取更多的图片搜索结果；另一方面，在某些类别的图片搜索结果中，譬如具有商业价值的图片上，基于图片光学字符识别的图片内容识别，可以带来更为精确的图片相关文本，能更精确的满足用户需求。因此本发明提供的图片搜索方法和装置，不管是在图片搜索还是在具有商业价值的广告配图上，能显著地提高搜索或者配图的性能。

本发明实施例还提供了一种电子设备，包括：一个或者多个处理器；存储器；一个或者多个程序，所述一个或者多个程序存储在所述存储器中，当被所述一个或者多个处理器执行时进行如下操作：接收搜索关键词；根据所述搜索关键词在图片文本信息的索引中进行搜索，获得与所述搜索关键词匹配的索引；所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；获得与所述搜索关键词匹配的索引对应的图片。

本发明实施例还提供了一种非易失性计算机存储介质，所述计算机存储介质存储有一个或者多个模块，当所述一个或者多个模块被执行时进行如下操作：接收搜索关键词；根据所述搜索关键词在图片文本信息的索引中进行搜索，获得与所述搜索关键词匹配的索引；所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；获得与所述搜索关键词匹配的索引对应的图片。

需要说明的是，在本发明的描述中，除非另有说明，“多个”的含义是两个或两个以上。

流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为，表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分，并且本发明的优选实施方式的范围包括另外的实现，其中可以不按所示出或讨论的顺序，包括根据所涉及的功能按基本同时的方式或按相反的顺序，来执行功能，这应被本发明的实施例所属技术领域的技术人员所理解。

应当理解，本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中，多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。例如，如果用硬件来实现，和在另一实施方式中一样，可用本领域公知的下列技术中的任一项或他们的组合来实现：具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路，具有合适的组合逻辑门电路的专用集成电路，可编程门阵列（Programmable Gate Array；以下简称：PGA），现场可编程门阵列（Field Programmable Gate Array；以下简称：FPGA）等。

本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通过程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，

该程序在执行时，包括方法实施例的步骤之一或其组合。

此外，本发明各个实施例中的各功能模块可以集成在一个处理模块中，也可以是各个模块单独物理存在，也可以两个或两个以上模块集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用时，也可以存储在一个计算机可读取存储介质中。

上述提到的存储介质可以是只读存储器，磁盘或光盘等。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中，对上述术语的示意性表述不一定指的是相同的实施例或示例。而且，描述的具体特征、结构、材料或者特点可以在任何一个或多个实施例或示例中以合适的方式结合。

尽管上面已经示出和描述了本发明的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本发明的限制，本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

权利要求书

1、一种图片搜索方法，其特征在于，包括：

接收搜索关键词；

5 根据所述搜索关键词在图片文本信息的索引中进行搜索，获得与所述搜索关键词匹配的索引；所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息；获得与所述搜索关键词匹配的索引对应的图片。

2、根据权利要求1所述的方法，其特征在于，所述接收搜索关键词之前，还包括：对搜索引擎抓取的图片进行内容分析，获得所述图片的文本信息。

10 3、根据权利要求1所述的方法，其特征在于，所述接收搜索关键词之前，还包括：对搜索引擎抓取的图片进行内容分析和对所述图片所在网页的文本进行分析，获得所述图片的文本信息。

4、根据权利要求2或3所述的方法，其特征在于，所述获得所述图片的文本信息之后，所述接收搜索关键词之前，还包括：

15 对所述图片的文本信息进行整理，建立所述图片的文本信息的索引。

5、根据权利要求2或3所述的方法，其特征在于，所述对搜索引擎抓取的图片进行内容分析包括：

对搜索引擎抓取的图片中的物体信息、风景信息、人物信息和文字信息中的至少一种进行分析。

20 6、根据权利要求3所述的方法，其特征在于，所述对所述图片所在网页的文本进行分析包括：

对所述图片所在网页的结构化字段中的文本进行分析，所述图片所在网页的结构化字段包括所述网页的网页标题字段、锚文本字段和/或图片属性字段。

7、一种图片搜索装置，其特征在于，包括：

25 接收模块，用于接收搜索关键词；

搜索模块，用于根据所述接收模块接收的搜索关键词在图片文本信息的索引中进行搜索，获得与所述搜索关键词匹配的索引；以及获得与所述搜索关键词匹配的索引对应的图片；所述图片文本信息包括对搜索引擎抓取的图片进行内容分析获得的文本信息。

8、根据权利要求7所述的装置，其特征在于，还包括：获得模块；

30 所述获得模块，用于在所述接收模块接收搜索关键词之前，对搜索引擎抓取的图片进行内容分析，获得所述图片的文本信息。

9、根据权利要求7所述的装置，其特征在于，还包括：获得模块；

所述获得模块，用于在所述接收模块接收搜索关键词之前，对搜索引擎抓取的图片进行内容分析和对所述图片所在网页的文本进行分析，获得所述图片的文本信息。

10、根据权利要求 8 或 9 所述的装置，其特征在于，还包括：

5 建立模块，用于在所述获得模块获得所述图片的文本信息之后，所述接收模块接收搜索关键词之前，对所述图片的文本信息进行整理，建立所述图片的文本信息的索引。

11、根据权利要求 8 或 9 所述的装置，其特征在于，所述获得模块用于对搜索引擎抓取的图片进行内容分析包括：

所述获得模块，具体用于对搜索引擎抓取的图片中的物体信息、风景信息、人物信息和文字信息中的至少一种进行分析。

10 12、根据权利要求 9 所述的装置，其特征在于，所述获得模块用于对所述图片所在网页的文本进行分析包括：

所述获得模块，具体用于对所述图片所在网页的结构化字段中的文本进行分析，所述图片所在网页的结构化字段包括所述网页的网页标题字段、锚文本字段和/或图片属性字段。

15 13、一种电子设备，其特征在于，包括：

一个或者多个处理器；

存储器；

一个或者多个程序，所述一个或者多个程序存储在所述存储器中，当被所述一个或者多个处理器执行时：

20 执行如权利要求 1-6 任一项所述的方法。

14、一种非易失性计算机存储介质，其特征在于，所述计算机存储介质存储有一个或者多个模块，当所述一个或者多个模块被执行时：

执行如权利要求 1-6 任一项所述的方法。

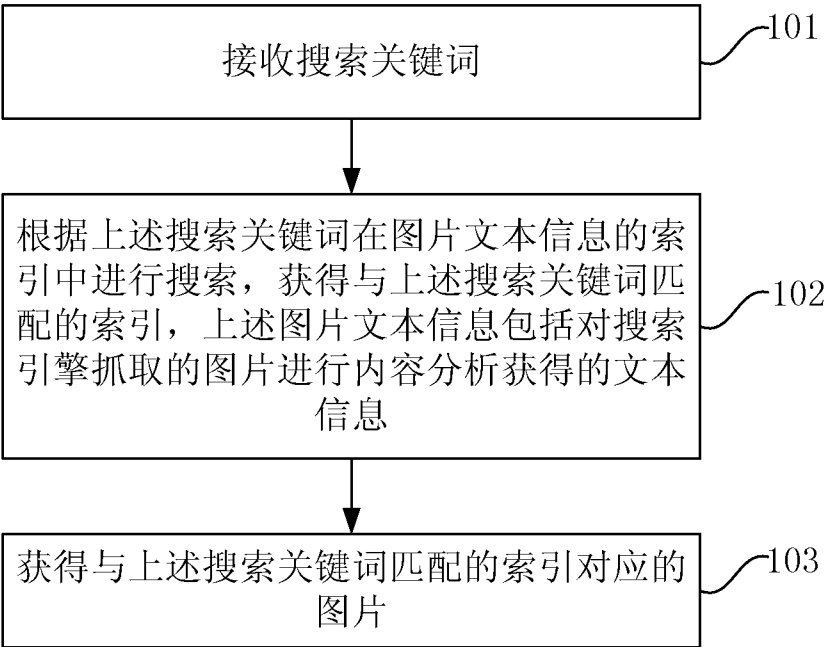


图 1

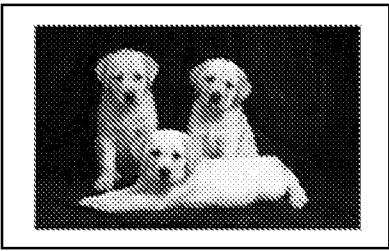


图 2 (a)

类别	概率
可爱小狗	0.06899890
拉布拉多	0.05316645
拉布拉多猎犬	0.04411934
拉布拉多猎狗	0.04006365
拉布拉多犬	0.03541176

图 2 (b)

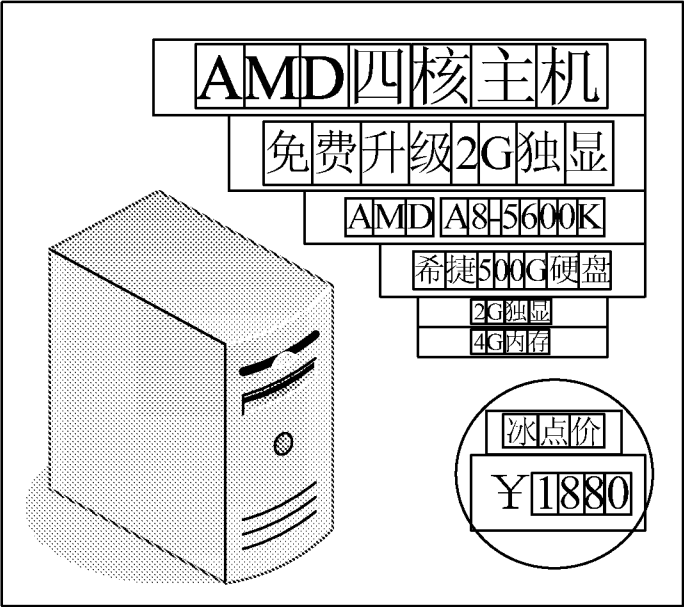


图 3 (a)

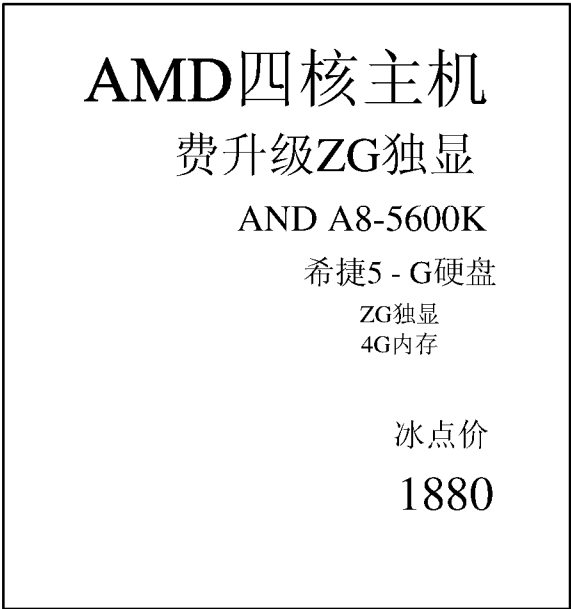


图 3 (b)

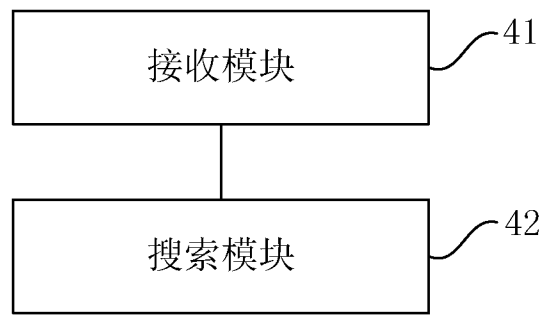


图 4

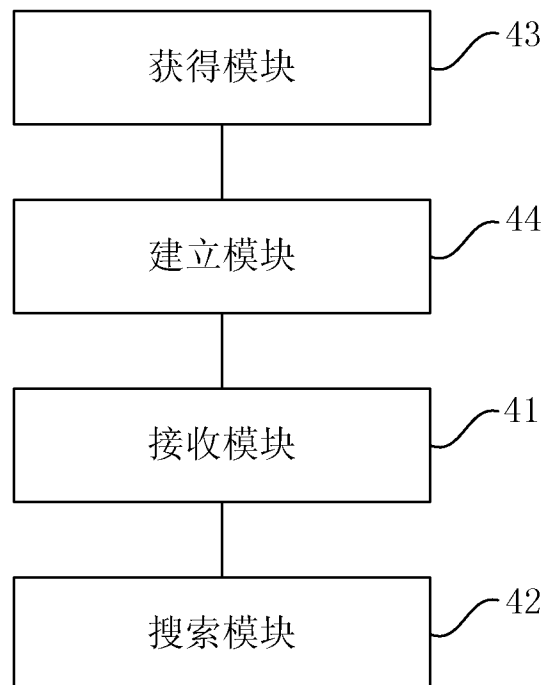


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/083395

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, EPODOC, WPI, IEEE: key word, text, BAIDU, analyze, image?, search+, content, key, word?, extract+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103761326 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 30 April 2014 (30.04.2014), description, paragraphs [0032]-[0044] and [0064]-[0085], and figures 2-3 and 8	1-14
X	CN 103793498 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 14 May 2014 (14.05.2014), description, paragraphs [0051]-[0074], [0100]-[0120] and [0141]-[0145], and figures 2a-6 and 11-13	1-14
PX	CN 104504109 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 08 April 2015 (08.04.2015), description, paragraphs [0020]-[0046] and [0050]-[0053], and figures 1-5	1-14
PX	CN 104462590 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 25 March 2015 (25.03.2015), description, paragraphs [0019]-[0053] and [0063]-[0066], and figures 1a-4	1-14
X	CN 103559220 A (BEIJING QIHOO TECHNOLOGY CO., LTD. et al.), 05 February 2014 (05.02.2014), description, paragraphs [0140]-[0115], [0118]-[0126] and [0147], and figures 4-6	1-14
X	CN 103136228 A (ALIBABA GROUP HOLDING LIMITED), 05 June 2013 (05.06.2013), description, paragraphs [0050]-[0060], [0076], [0092]-[0104] and [0187]-[0190], and figures 3 and 7	1-2, 4-5, 7-8, 10-11, 13-14
A	EP 1103900 A2 (MEDISON CO., LTD.), 30 May 2001 (30.05.2001), the whole document	1-14

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 06 September 2015 (06.09.2015)	Date of mailing of the international search report 08 October 2015 (08.10.2015)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer LI, Duo Telephone No.: (86-10) 62414054

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2015/083395

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103761326 A	30 April 2014	None	
CN 103793498 A	14 May 2014	None	
CN 104504109 A	08 April 2015	None	
CN 104462590 A	25 March 2015	None	
CN 103559220 A	05 February 2014	None	
CN 103136228 A	05 June 2013	WO 2013078307 A1	30 May 2013
		US 2013138636 A1	30 May 2013
		TW 201322021 A	01 June 2013
		HK 1181481 A0	08 November 2013
		EP 2783310 A1	01 October 2014
		JP 2014533868 A	15 December 2014
EP 1103900 A2	30 May 2001	KR 20010048272 A	15 June 2001
		JP 2001167129 A	22 June 2001
		EP 1103900 A3	16 January 2002
		US 6488627 B1	03 December 2002
		KR 100367932 B1	14 January 2003

国际检索报告

国际申请号

PCT/CN2015/083395

A. 主题的分类

G06F 17/30(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT, CNKI, EPODOC, WPI, IEEE: 图, 索, 关键词, 内容, 文本, 百度, 分析, image?, search+, content, key, word?, extract+

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 103761326 A (百度在线网络技术北京有限公司) 2014年 4月 30日 (2014 - 04 - 30) 说明书第[0032]-[0044], [0064]-[0085]段, 附图2-3, 8	1-14
X	CN 103793498 A (百度在线网络技术北京有限公司) 2014年 5月 14日 (2014 - 05 - 14) 说明书第[0051]-[0074], [0100]-[0120], [0141]-[0145]段, 附图2a-6, 11-13	1-14
PX	CN 104504109 A (百度在线网络技术北京有限公司) 2015年 4月 8日 (2015 - 04 - 08) 说明书第[0020]-[0046], [0050]-[0053]段, 附图1-5	1-14
PX	CN 104462590 A (百度在线网络技术北京有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0019]-[0053], [0063]-[0066]段, 附图1a-4	1-14
X	CN 103559220 A (北京奇虎科技有限公司 等) 2014年 2月 5日 (2014 - 02 - 05) 说明书第[0140]-[0115], [0118]-[0126], [0147]段, 附图4-6	1-14
X	CN 103136228 A (阿里巴巴集团控股有限公司) 2013年 6月 5日 (2013 - 06 - 05) 说明书第[0050]-[0060], [0076], [0092]-[0104], [0187]-[0190]段, 附图3, 7	1-2, 4-5, 7-8, 10-11, 13-14
A	EP 1103900 A2 (MEDISON CO., LTD.) 2001年 5月 30日 (2001 - 05 - 30) 全文	1-14

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2015年 9月 6日

国际检索报告邮寄日期

2015年 10月 8日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
北京市海淀区蓟门桥西土城路6号
100088 中国

传真号 (86-10)62019451

授权官员

李铎

电话号码 (86-10)62414054

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/083395

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	103761326	A	2014年 4月 30日	无			
CN	103793498	A	2014年 5月 14日	无			
CN	104504109	A	2015年 4月 8日	无			
CN	104462590	A	2015年 3月 25日	无			
CN	103559220	A	2014年 2月 5日	无			
CN	103136228	A	2013年 6月 5日	WO	2013078307	A1	2013年 5月 30日
				US	2013138636	A1	2013年 5月 30日
				TW	201322021	A	2013年 6月 1日
				HK	1181481	A0	2013年 11月 8日
				EP	2783310	A1	2014年 10月 1日
				JP	2014533868	A	2014年 12月 15日
EP	1103900	A2	2001年 5月 30日	KR	20010048272	A	2001年 6月 15日
				JP	2001167129	A	2001年 6月 22日
				EP	1103900	A3	2002年 1月 16日
				US	6488627	B1	2002年 12月 3日
				KR	100367932	B1	2003年 1月 14日

表 PCT/ISA/210 (同族专利附件) (2009年7月)