



(12)发明专利申请

(10)申请公布号 CN 106462574 A

(43)申请公布日 2017.02.22

(21)申请号 201580033709.4

(22)申请日 2015.06.19

(30)优先权数据

14/313,670 2014.06.24 US

(85)PCT国际申请进入国家阶段日

2016.12.22

(86)PCT国际申请的申请数据

PCT/US2015/036603 2015.06.19

(87)PCT国际申请的公布数据

W02015/200110 EN 2015.12.30

(71)申请人 谷歌公司

地址 美国加利福尼亚州

(72)发明人 阿恩·莫塞

亚历山大·杰·卡斯贝特

约翰·斯特迪·德内罗

(74)专利代理机构 中原信达知识产权代理有限

责任公司 11219

代理人 李宝泉 周亚荣

(51)Int.Cl.

G06F 17/28(2006.01)

G06K 9/00(2006.01)

G06K 9/32(2006.01)

G06K 9/72(2006.01)

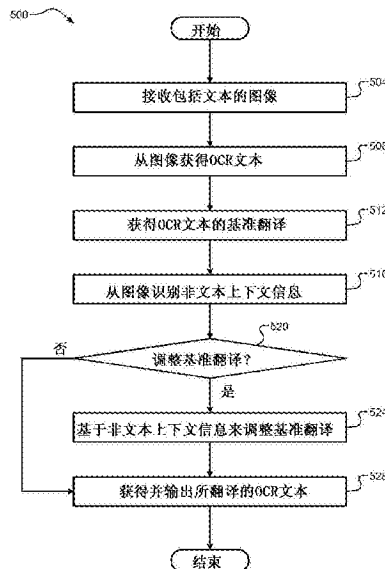
权利要求书2页 说明书8页 附图5页

(54)发明名称

用于基于来自图像的非文本上下文信息的来自图像的文本的机器语言翻译的技术

(57)摘要

一种计算机实现技术可以包括在服务器处从移动计算设备接收包括文本的图像,所述服务器具有一个或多个处理器。所述技术可以包括在服务器处获得对应于文本的光学字符识别(OCR)文本,该OCR文本是通过从图像执行OCR而获得的。所述技术可以包括在服务器处从图像识别非文本上下文信息,所述非文本上下文信息(i)表示除所述文本本身之外的上下文信息以及(ii)指示所述图像的上下文。所述技术可以包括基于所述非上下文信息,在服务器处获得OCR文本到目标语言的翻译以获得所翻译的OCR文本。所述技术可以包括从服务器向移动计算设备输出所翻译的OCR文本。



1. 一种计算机实现的方法,包括:

在服务器处从移动计算设备接收包括文本的图像,所述服务器具有一个或多个处理器;

在所述服务器处获得对应于所述文本的光学字符识别OCR文本,所述OCR文本是通过对所述图像执行OCR而获得的;

在所述服务器处从所述图像识别非文本上下文信息,所述非文本上下文信息 (i) 表示除所述文本本身之外的上下文信息且 (ii) 指示所述图像的上下文;

基于所述非文本上下文信息,在所述服务器处获得所述OCR文本到目标语言的翻译以获得所翻译的OCR文本;以及

从所述服务器向所述移动计算设备输出所翻译的OCR文本。

2. 根据权利要求1所述的计算机实现的方法,还包括:

在所述服务器处获得所述OCR文本到所述目标语言的翻译以获得基准翻译的OCR文本;以及

在所述服务器处基于所述非文本上下文信息来调整所述基准翻译的OCR文本以获得所翻译的OCR文本。

3. 根据权利要求1所述的计算机实现的方法,还包括在所述服务器处基于所述非文本上下文信息来确定所述文本的源语言,其中所翻译的OCR文本进一步基于所述源语言。

4. 根据权利要求1所述的计算机实现的方法,还包括在所述服务器处基于所述非文本上下文信息来确定捕捉所述图像的位置的类型,其中所翻译的OCR文本进一步基于所述位置的类型。

5. 根据权利要求1所述的计算机实现的方法,还包括在所述服务器处确定所述移动计算设备的地理位置,其中所翻译的OCR文本进一步基于所述移动计算设备的所述地理位置。

6. 根据权利要求5所述的计算机实现的方法,还包括:

在所述服务器处基于所述地理位置来确定地图信息;以及

在所述服务器处使用所述地图信息来识别所述地理位置附近的兴趣点,其中,所翻译的OCR文本进一步基于所述地理位置附近的所述兴趣点。

7. 根据权利要求1所述的计算机实现的方法,还包括:在所述服务器处确定对应于所述移动计算设备的用户的用户历史,其中所翻译的OCR文本进一步基于所述用户历史。

8. 根据权利要求1所述的计算机实现的方法,其中,所述非文本上下文信息包括所述文本的字体。

9. 根据权利要求1所述的计算机实现的方法,其中,所述非文本上下文信息包括以下中的至少一个:(i) 所述图像中的对象和 (ii) 所述图像中的所述对象的形状。

10. 根据权利要求9所述的计算机实现的方法,其中,所述非文本上下文信息包括所述对象的颜色中的至少一个,并且进一步包括在所述服务器处基于所述对象的所述颜色来确定所述图像是在室内还是室外捕捉的,其中,所翻译的OCR文本进一步基于所述图像是在室内还是室外捕捉的。

11. 一种具有一个或多个处理器的服务器,所述处理器被配置成执行包括以下的操作:从移动计算设备接收包括文本的图像;

获得对应于所述文本的光学字符识别OCR文本,所述OCR文本是通过对所述图像执行

OCR而获得的；

从所述图像识别非文本上下文信息，所述非文本上下文信息(i)表示除所述文本本身之外的上下文信息且(ii)指示所述图像的上下文；

基于所述非文本上下文信息，获得所述OCR文本到目标语言的翻译以获得翻译的OCR文本；以及

向所述移动计算设备输出所翻译的OCR文本。

12. 根据权利要求11所述的服务器，其中，所述操作还包括：

获得所述OCR文本到所述目标语言的翻译以获得基准翻译的OCR文本；以及

基于所述非文本上下文信息来调整所述基准翻译的OCR文本，以获得所翻译的OCR文本。

13. 根据权利要求11所述的服务器，其中，所述操作还包括基于所述非文本上下文信息来确定所述文本的源语言，其中，所翻译的OCR文本进一步基于所述源语言。

14. 根据权利要求11所述的服务器，其中，所述操作还包括基于所述非文本上下文信息来确定捕捉所述图像的位置的类型，其中，所翻译的OCR文本进一步基于所述位置的类型。

15. 根据权利要求11所述的服务器，其中，所述操作还包括确定所述移动计算设备的地理位置，其中，所翻译的OCR文本进一步基于所述移动计算设备的所述地理位置。

16. 根据权利要求15所述的服务器，其中，所述操作还包括：

基于所述地理位置来获得地图信息；以及

使用所述地图信息来识别所述地理位置附近的兴趣点，

其中，所翻译的OCR文本进一步基于所述地理位置附近的所述兴趣点。

17. 根据权利要求11所述的服务器，其中，所述操作还包括确定对应于所述移动计算设备的用户的用户历史，其中，所翻译的OCR文本进一步基于所述用户历史。

18. 根据权利要求11所述的服务器，其中，所述非文本上下文信息包括所述文本的字体。

19. 根据权利要求11所述的服务器，其中，所述非文本上下文信息包括以下中的至少一个：(i)所述图像中的对象和(ii)所述图像中的所述对象的形状。

20. 根据权利要求19所述的服务器，其中，所述非文本上下文信息包括所述对象的颜色中的至少一个，并且其中，所述操作还包括基于所述对象的所述颜色来确定所述图像是在室内还是室外捕捉的，其中，所翻译的OCR文本进一步基于所述图像是在室内还是室外捕捉的。

用于基于来自图像的非文本上下文信息的来自图像的文本的 机器语言翻译的技术

[0001] 相关申请的交叉引用

[0002] 本申请要求2014年6月24日提交的美国申请号14/313,670的优先权。上述申请的公开被整体地通过引用并入到本文中。

技术领域

[0003] 本公开一般地涉及光学字符识别(OCR)和机器语言翻译,并且更特别地涉及用于基于来自图像的非文本上下文信息的对来自图像的OCR文本的机器语言翻译的技术。

背景技术

[0004] 本文中提供的背景描述是出于一般地呈现本公开的上下文的目的。在已经在背景技术小节中进行描述的程度,本发明人的工作以及以其它方式在提交时未被承认为现有技术之本描述的各方面既未明示地、也为隐含地被认可为针对本公开的现有技术。

[0005] 光学字符识别(OCR)涉及到使用计算设备(例如,服务器)的图像中的文本的检测。OCR可以提供用以在用户设备处以数字形式获得文本的更快方式(例如,与由用户向用户设备手动输入文本相比)。在获得图像中的文本之后,可以以各种方式利用该文本。例如,该文本可被计算设备处理、被存储在存储器处和/或被发送到另一计算设备。处理文本的一个示例是机器语言翻译,其涉及到使用计算设备将文本从源语言翻译成不同的目标语言。

发明内容

[0006] 在一个方面,提出了一种计算机实现的技术。所述技术可以包括在服务器处从移动计算设备接收包括文本的图像,所述服务器具有一个或多个处理器。所述技术可以包括在服务器处获得对应于文本的光学字符识别(OCR)文本,所述OCR文本是通过执行OCR而获得的。被技术可以包括在服务器处从图像识别非文本上下文信息,所述非文本上下文信息(i)表示除文本本身之外的上下文信息且(ii)指示图像的上下文。所述技术可以包括基于非文本上下文信息,在服务器处获得OCR文本到目标语言的翻译以获得所翻译的OCR文本。所述技术还可以包括从服务器向移动计算设备输出所翻译的OCR文本。

[0007] 在某些实施例中,所述技术还包括在服务器处获得OCR文本到目标语言的翻译以获得基准翻译的OCR文本,以及在服务器处基于非文本上下文信息来调整基准翻译的OCR文本以获得所翻译的OCR文本。

[0008] 在其它实施例中,所述技术还包括在服务器处基于非文本上下文信息来确定文本的源语言,其中,所翻译的OCR文本进一步基于源语言。

[0009] 在某些实施例中,所述技术还包括在服务器处基于非文本上下文信息来确定捕捉到图像的位置的类型,其中,所翻译的OCR文本进一步基于位置的类型。

[0010] 在其它实施例中,所述技术还包括在服务器处确定移动计算设备的地理位置,其中,所翻译的OCR文本进一步基于移动计算设备的地理位置。

[0011] 在某些实施例中,所述技术还包括在服务器处基于地理位置来获得地图信息,以及在服务器处使用地图信息来识别地理位置附近的兴趣点,其中,所翻译的OCR文本进一步基于所述地理位置附近的兴趣点。

[0012] 在其它实施例中,所述技术还包括在服务器处确定对应于移动计算设备的用户的用户历史,其中,所翻译的OCR文本进一步基于用户历史。

[0013] 在某些实施例中,所述非文本上下文信息包括文本的字体。在其它实施例中,所述非文本上下文信息包括以下的至少一个:(i)图像中的对象和(ii)图像中的对象的形状。

[0014] 在某些实施例中,所述非文本上下文信息包括对象的颜色中的至少一个,并且其中,所述技术还包括在服务器处基于对象的颜色来确定图像是在室内还是室外捕捉的,其中,所翻译的OCR文本进一步基于图像是在室内还是室外捕捉的。

[0015] 还呈现了一种具有被配置成执行操作的一个或多个处理器的服务器。所述操作可以包括从移动计算设备接收包括文本的图像。所述操作可以包括获得对应于文本的OCR文本,所述OCR文本是通过图像执行OCR而获得的。所述操作可以包括从图像识别非文本上下文信息,所述非文本上下文信息(i)表示除文本本身之外的上下文信息以及(ii)指示图像的上下文。所述操作可以包括基于非文本上下文信息,获得OCR文本到目标语言的翻译以获得所翻译的OCR文本。所述操作还可以包括向移动计算设备输出所翻译的OCR文本。

[0016] 在某些实施例中,所述操作还包括获得OCR文本到目标语言的翻译以获得基准翻译的OCR文本,以及基于非文本上下文信息来调整基准翻译的OCR文本以获得所翻译的OCR文本。

[0017] 在其它实施例中,所述操作还包括基于非文本上下文信息来确定文本的源语言,其中,所翻译的OCR文本进一步基于源语言。

[0018] 在某些实施例中,所述操作还包括基于非文本上下文信息来确定捕捉到图像的位置的类型,其中,所翻译的OCR文本进一步基于位置的类型。

[0019] 在其它实施例中,所述操作还包括确定移动计算设备的地理位置,其中,所翻译的OCR文本进一步基于移动计算设备的地理位置。

[0020] 在某些实施例中,所述操作还包括基于地理位置来获得地图信息,以及使用地图信息来识别地理位置附近的兴趣点,其中,所翻译的OCR文本进一步基于所述地理位置附近的兴趣点。

[0021] 在其它实施例中,所述操作还包括确定对应于移动计算设备的用户的用户历史,其中,所翻译的OCR文本进一步基于用户历史。

[0022] 在某些实施例中,所述非文本上下文信息包括文本的字体。在其它实施例中,所述非文本上下文信息包括以下中的至少一个:(i)图像中的对象和(ii)图像中的对象的形状。

[0023] 在某些实施例中,所述非文本上下文信息包括对象的颜色中的至少一个,并且其中,所述操作还包括基于对象的颜色来确定图像是在室内还是室外捕捉的,其中,所翻译的OCR文本进一步基于图像是在室内还是室外捕捉的。

[0024] 根据在下文中提供的详细描述,本公开的更多适用领域将变得显而易见。应理解的是本详细描述和特定示例仅仅意图用于说明的目的且并不意图限制本公开的范围。

附图说明

[0025] 根据本详细描述和附图将更全面地理解本公开,在所述附图中:

[0026] 图1是根据本公开的某些实施方式的包括示例服务器和示例移动计算设备的计算网络的图;

[0027] 图2是图1的示例服务器的功能框图;

[0028] 图3A-图3B图示出根据本公开的某些实施方式的每一个具有相同文本但不同的非文本上下文信息的示例图像对;

[0029] 图4A-图4B图示出根据本公开的某些实施方式的每一个具有相同文本但不同的非文本上下文信息的另一示例图像对;以及

[0030] 图5是根据本公开的某些实施方式的用于基于来自图像的非文本上下文信息的来自图像的OCR文本的机器语言翻译的示例技术的流程图。

具体实施方式

[0031] 常规系统从图像获得光学字符识别(OCR)文本并执行OCR文本从其源语言到期望语言的机器语言翻译。然而,某些单词和/或短语在某些语言中可能具有多个意义。这些多个意义中的每一个可取决于特定单词或短语的上下文。例如,德语单词“rind”可意指“牛肉”和“奶牛”两者。常规系统将把OCR文本“rind”翻译成英语的“beef(牛肉)”或“cow(奶牛)”中任一个,例如哪个单词在英语中具有较高使用率。并且例如,英语单词“bear right”可以意指“向右”“转向”、“转弯”或“止住”以及“熊”(动物)“在右边”两者。这些系统并未考虑到图像中的可指示其上下文的其它信息,该信息可以用来确定正确的意义并因此获得正确或期望的翻译。因此,呈现了用于来自图像的OCR文本基于来自图像的非文本上下文信息的机器语言翻译的技术。

[0032] 如本文所使用的“非文本上下文信息”指的是来自图像的除文本本身之外的指示图像的上下文的任何信息。非文本上下文信息包括从图像本身提取且指示图像的上下文的信息(特征、对象等)。换言之,非文本上下文信息是从图像提取或导出(例如,在图像内部)的信息而不是关于图像的信息(例如,在图像外部),诸如与图像或图像文件相关联的元数据或其它属性。非文本上下文信息的示例包括形状、颜色以及对象,诸如图像的背景中的特征或对象。还可以从图像的前景获得非文本上下文信息。另外,非文本上下文信息还可以包括文本本身的属性,诸如其字体。

[0033] 本技术可以识别此非文本上下文信息,并且使用该非文本上下文信息来获得从图像获得的OCR文本的翻译。可以使用任何适当的图像处理技术和特征/对象检测技术来获得非文本上下文信息。在一个实施方式中,本技术可以获得OCR文本的基准翻译,并且然后基于非文本上下文信息来确定是否要调整基准翻译。这还可以表示用于基准翻译的验证步骤。还可以使用其它信息来进一步增强这些技术,用户的移动计算设备的地理位置及其过去的行为(或“用户历史”)。此信息可以指示在该处捕捉到图像的位置的类型,其可以用来增强非文本上下文信息识别。还可以使用位置来更容易地识别源语言。

[0034] 在其中这里所讨论的系统可收集关于用户的个人信息或者可利用个人信息的情况下,可为用户提供控制程序或特征是否收集用户信息(例如,关于用户的历史(诸如,其关于社交网络的过去行为)、社交动作或活动、其职业和/或其一般偏好的信息、和/或关于用户当前地理位置的信息)或控制是否和/或如何从内容服务器接收可能与用户更加相关的

信息的机会。另外,某些数据在其被存储或使用之前可以被以一个或多个方式处理,使得个人可识别信息被去除。例如,可处理用户的身份,使得对于用户而言不能确定个人可识别信息,或者可将在该处获得位置信息的用户的地理位置广义化(诸如到城市、ZIP代码或州层级),使得不能确定用户的特定位置。因此,用户可能可控制如何和是否收集关于用户的信息和该信息如何被内容服务器使用。

[0035] 现在参考图1,图示出计算网络100。计算网络100包括服务器104。如本文所使用的术语“服务器”可以指代单个硬件计算机服务器和在并行或分布式架构中操作的多个类似服务器两者。仅仅为了示例,可以实现第一和第二服务器以便分别地执行OCR和机器语言翻译。移动计算设备108被配置成经由网络112与服务器104通信。移动计算设备108的示例包括膝上型计算机、平板计算机、移动电话以及可穿戴技术,诸如智能手表、眼镜、或并入了计算设备的其它可穿戴对象。然而,应认识到的是可以在具有显示器和照相机的任何计算设备(例如,台式计算机)处实现本公开的技术。网络112可以包括局域网(LAN)、广域网(WAN)(例如,因特网)、或其组合。

[0036] 移动计算设备108可以与用户116相关联。例如,用户116可以经由显示器120(例如,触摸显示器)与移动计算设备108相交互。用户116可以使用移动计算设备108来捕捉场景的图像124。图像124可以包括前景128和背景132。如所示,图像124的前景128包括在其上面具有文本140的对象136。在所示示例中,场景是饭店且对象136是菜单。仅仅为了示例,用户116可在捕捉图像124的同时将对象136举在她前面。对象136的其它示例包括文档、标志以及广告。可以对图像执行OCR以检测文本140。如所示,图像124的背景132可以包括非文本上下文信息144,其包括桌子148、和布置在桌子148的上表面上的用餐地点布置152(盘子、叉子、刀、匙、餐巾等)。

[0037] 在获得文本140之后,然后将文本140翻译成目标语言,诸如用户116所理解/说的语言。OCR和/或机器语言翻译可以本地地(在移动计算设备108处)、远程地(在服务器处104)或以其组合的方式执行。例如,移动计算设备108可以将图像发送到服务器104以用于处理。服务器104然后可以获得OCR文本,识别非文本上下文信息144,并且获得OCR文本的翻译。在所示示例中,非文本上下文信息144指示饭店环境,并且因此服务器104可以利用此知识来获得对应于图像124的饭店上下文的OCR文本的翻译。如在本文中先前所讨论的,例如,如果“OCR”是“rind”,则饭店语言将指示到英语单词“beef”而不是英语单词“cow”的翻译。

[0038] 现在参考图2,图示出示例服务器104的功能框图。服务器104可以包括通信设备200、处理器204以及存储器208。通信设备200可以包括被配置成经由网络112与其它设备(诸如移动计算设备108)的任何适当组件(例如,收发机)。存储器208可以是被配置成在服务器104存储信息的任何适当存储介质(闪存、硬盘等)。处理器204可以控制服务器104的操作。由处理器204执行的示例功能包括但不限于加载/执行服务器104的操作系统、控制经由通信设备200进行的信息的发送/接收、以及控制存储器208处的读/写操作。处理器204还可以被配置成执行本公开的技术的至少一部分,现在更详细地对其进行讨论。

[0039] 服务器104可以接收图像以用于处理。该图像可以是从小移动计算设备108接收到的。在某些实施方式中,移动计算设备108可以捕捉图像,但是还可以由移动计算设备108从另一源获得图像。服务器104处的图像的接收可以表示对OCR和机器语言翻译的请求。替选

地,移动计算设备108可连同图像一起向服务器104发送显式请求。服务器104可以对图像执行OCR以获得OCR文本。可以利用任何适当的OCR算法来获得OCR文本。在某些实施方式中,可以在另一设备(诸如,移动计算设备108或另一服务器)处执行OCR的至少一部分。在获得OCR文本之后,服务器104可以然后继续本公开的机器语言翻译技术。

[0040] 在某些实施方式中,服务器104可以获得OCR文本的所翻译的基准或“所翻译的基准OCR文本”。此所翻译的基准OCR文本可以表示在未考虑到来自图像的任何非文本上下文信息的情况下的OCR文本从其源语言至目标语言的机器语言翻译。该目标语言可以由移动计算设备108(诸如由来自用户116的输入或其偏好)指定。服务器104可以从图像识别任何非文本上下文信息。服务器104可以利用任何适当的对象识别和/或特征检测技术来识别非文本上下文信息。在某些实施方式中,服务器104可以执行机器学习以使用标记训练集来训练分类器,并且然后使用所训练的分类器来识别非文本上下文信息。如先前所讨论的,示例非文本上下文信息包括对象类型、形状、颜色以及文本属性(诸如字体)。

[0041] 在识别非文本上下文信息之后,服务器104可以基于非文本上下文信息而获得OCR文本的翻译或“所翻译的OCR文本”。如果先前获得了基准翻译的OCR文本,则服务器104可以根据非文本上下文信息来确定是否要调整基准翻译的OCR文本。仅仅为了示例,可基于各种非文本上下文信息而生成一个或多个附加翻译的OCR文本,并且然后可以诸如使用语料库加权意义将每个所翻译的OCR文本(包括基准OCR文本)相互比较。另外,还可以使用其它信息来增强所翻译的OCR文本的获得。如先前所讨论的,此其它信息的示例包括位置和/或地理位置的类型以及用户历史。另外,服务器104可确定图像是在室内还是室外捕捉的。此确定可以基于非文本上下文信息、其它信息(例如,地理位置)和/或成像参数(环境照明、亮度、闪光灯开/关等)。

[0042] 如果非文本上下文信息指示基准翻译的OCR文本是不正确的或不准确的(例如,不正确的上下文),则服务器104可以基于非文本上下文信息来调整基准翻译的OCR文本以获得所翻译的OCR文本或“调整的所翻译的OCR文本”。否则,可以利用基准翻译的OCR文本。服务器104然后将适当的翻译的OCR文本输出到移动计算设备108以便显示给用户116。服务器104可另外或备选地将适当的所翻译的OCR文本发送到另一设备和/或存储适当的所翻译的OCR文本(例如,在存储器208)以学习和/或用于未来操作的检索。仅仅为了示例,同一图像(诸如知名标志或地标)可能会从不同的用户移动计算设备重复地发送到服务器104,并且因此服务器104可能能够确定这些图像是相同的,并且然后检索适当的所翻译的OCR文本以便更快速地提供给用户移动计算设备。

[0043] 图3A-图3B图示出每个具有相同的文本但不同的上下文和因此具有不同的非文本上下文信息的示例图像对300、350。每个图像300、350包括文本“rind”。如先前所讨论的,德语单词“rind”具有多个意义且对应于英语单词“beef”和英语单词“cow”两者。同样地,图4A-4B还图示出每个具有相同的文本但不同的上下文和因此具有不同的非文本上下文信息的示例图像对400、450。如先前所讨论的,英语短语“bear right”具有多个意义,包括在右边的熊(动物)和向右转向/转弯/止住。这些不同的意义在非英语语言中可以对应于多个翻译。常规系统将为这些图像对300、350和400、450中的每一个输出相同的翻译的OCR文本。然而,由于这些图像对300、350和400、450的上下文是不同的,所以可能需要不同的翻译的OCR文本。

[0044] 在图3A的左侧图像300中,上下文是杂货店,并且罐头304包括文本308。服务器104可以从图像300识别非文本上下文信息,其可以包括但不限于罐头304的形状(圆筒)、货架312、多个(其它)罐头316(例如,成行的)以及多个其它对象320(例如,成行的箱子)。可以指示杂货店上下文的其它示例非文本上下文信息可包括食品项(例如,罐头304)上的标签、文本312的字体、价格标签、以及过道号码/标记。如先前所讨论的,服务器104还可检测到图像300是在室内捕捉的,其可以被用来确定杂货店上下文。杂货店上下文指示意义“beef”而不是“cow”,并且因此可以提供适当的所翻译的OCR文本(“beef”)。

[0045] 在图3B的右侧图像350中,上下文是道路交叉路口,并且标志354包括文本358。服务器104可以从图像350识别非文本上下文信息,其可以包括但不限于标志354的形状(菱形)、标志上的图示362(穿过道路的奶牛)、道路366、道路366的交叉路口370(例如,具有与道路366不同的纹理,从而指示交叉路口)、以及树374。可以指示道路交叉路口上下文的其它示例非文本上下文信息是文本358的字体和一个或多个奶牛。例如,诸如标志358之类的标志可始终使用特定字体或字体集,并且还可使其它对象与特定字体或字体集相关联。服务器104还可检测到图像350是在室外捕捉的,其可以用来确定道路交叉路口上下文。道路交叉路口上下文指示意义“cow”而不是“beef”,并且因此可以提供适当的所翻译的OCR文本(“cow”)。

[0046] 在图4A的左侧图像400中,上下文是施工中的道路或公路,并且标志404包括文本408。服务器104可以从图像400识别非文本上下文信息,其可以包括但不限于标志的形状(菱形,其可指示道路/公路标志)、标志404上的箭头412的类型或形状(例如,箭头412的弯曲性质可指示道路420在分叉/分隔420之后向前)、道路或公路416、车道标记420、道路/公路416中的分叉/分隔424、以及施工路障428。可以指示道路/公路施工上下文的其它示例非文本上下文信息是文本408的字体、其它施工标志、施工工人以及颜色(例如,橙色可指示施工)。服务器104还可检测到图像400是在室外捕捉的,其可以用来确定道路/公路施工上下文。道路/公路施工上下文指示意义向右转向/转弯/止住而不是在右边的熊(动物),并且因此可以提供适当的所翻译的OCR文本。

[0047] 在图4B的右侧图像450中,上下文是动物园且标志454包括文本458。服务器104可以从图像400识别非文本上下文信息,其可以包括但不限于箭头462的类型或形状、通道或人行道466、草地470、树474、具有栅栏的笼子478以及笼子478中的熊482。可以指示动物园上下文的其它示例非文本上下文信息是文本458的字体(例如,有趣或儿童友好的字体)、其它动物园标志、其它动物、其它笼子以及人(例如,其它儿童)。服务器104还可检测到图像400是在室外捕捉的,其可以用来确定动物园上下文。仅仅为了示例,草地470和/或树474的绿色可以指示室外上下文、和更特别地动物园上下文两者。动物园上下文指示在右边的熊(动物)的意义而不是向右转向/转弯/止住,并且因此可以提供适当的所翻译的OCR文本。

[0048] 现在参考图5,图示出用于基于来自图像的非文本上下文信息的对来自图像的OCR文本的机器语言翻译的示例技术500的流程图。在504处,服务器104可以从移动计算设备108接收包括文本的图像。例如,该图像可以是由移动计算设备108捕捉的。在508处,服务器104可以获得对应于图像中的文本的OCR文本。可以在服务器104、在另一服务器、或在其某种组合执行OCR。在512处,服务器104可以可选地识别OCR文本到目标语言的翻译以获得基准翻译的OCR文本。此基准翻译的OCR文本可以表示在未考虑来自图像的非文本上下文信息

的情况下的OCR文本的翻译。

[0049] 在516处,服务器104可以从图像识别非文本上下文信息,该非文本上下文信息(i)表示除文本本身之外的上下文信息和(ii)指示图像的上下文。在520处,服务器104可以基于非文本上下文信息来确定是否要调整基准翻译的OCR文本。如果应根据非文本上下文信息来翻译基准翻译的OCR文本,则技术500可以继续至524。否则,本技术可以继续至528。在524处,服务器104可以基于非文本上下文信息来调整基准翻译的OCR文本以获得所翻译的OCR文本。例如,服务器104可基于非文本上下文信息来获得OCR文本的新翻译。在528处,服务器104可以将所翻译的OCR文本(或基准翻译的OCR)文本输出到移动计算设备108。技术500然后可以结束或返回至504。

[0050] 提供了示例实施例,使得本公开将是透彻的,并且将向本领域的技术人员全面地传达其的范围。阐述了许多特定细节(诸如特定组件、设备以及方法的示例)以提供本公开的实施例的透彻理解。对于本领域的技术人员而言将显而易见的是不需要采用特定细节,可以许多不同的形式体现示例实施例,并且不应将任一者理解成限制本公开的范围。在某些示例实施例中,并未详细地描述众所周知的步骤、众所周知的设备结构以及众所周知的技术。

[0051] 在本文中使用的术语仅仅是出于描述特定示例实施例的目的且并不意图限制本本发明。如本文所使用的单数形式“一”、“一个”和“该”也可意图包括复数形式,除非上下文另外清楚地指明。术语“和/或”包括关联列出项中的一个或多个的任何和所有组合。术语“包括”、“包含”和“具有”是包括性的,并且因此指定所述特征、整体、步骤、操作、元件和/或组件的存在,但不排除一个或多个其它特征、整体、步骤、操作、元件、组件和/或其群组的存在或添加。不应将本文所述的方法步骤、过程以及操作理解为必须要求按照所讨论或所示的特定次序的执行,除非具体地识别为执行次序。还应理解的是可采用附加或备选步骤。

[0052] 虽然在本文中可使用术语第一、第二、第三等来描述各种元件、组件、区域、层和/或区段,但这些元件、组件、区域、层和/或区段不应受到这些术语的限制。这些术语仅可用于将一个元件、组件、区域、层或区段与另一区域、层或区域区别开。诸如“第一”、“第二”之类的术语及其它数值术语当在本文中使用时并不暗示序列或次序,除非上下文清楚地指明。因此,在不脱离示例实施例的教导的情况下,可以将下面所讨论的第一元件、组件、区域、层或区段称为第二元件、组件、区域、层或区段。

[0053] 如本文所使用的,术语模块可指代或包括:专用集成电路(ASIC);电子电路;组合逻辑电路;现场可编程门阵列(FPGA);处理器或处理器的分布式网络(共享、专用或分组)和执行代码或进程的联网集群或数据中心中的储存器;提供所述功能的其它适当组件;或者上述各项中的某些或全部的组合(诸如在片上系统中)。术语模块还可包括存储被一个或多个处理器执行的代码的存储器(共享、专用或分组)。

[0054] 如上文所示用的术语代码可包括软件、固件、字节代码和/或微代码,并且可指代程序、例程、函数、类和/或对象。如上文所示用的术语共享意指可使用单个(共享)处理器来执行来自多个模块的某些或所有代码。另外,来自多个模块的某些或所有代码可被单个(共享)存储器存储。如上文所使用的术语群组意指可使用处理器群组来执行来自单个模块的某些或所有代码。另外,可使用存储器群组来存储来自单个模块的某些或所有代码。

[0055] 本文所述的技术可由一个或多个处理器所执行的一个或多个计算机程序实现。计

计算机程序包括处理器可执行指令,其被存储在非暂时性有形计算机可读介质上。该计算机程序还可包括存储数据。非暂时性有形计算机可读介质的非限制性示例是非易失性存储器、磁存储器以及光学存储器。

[0056] 以上描述中的某些部分根据对信息的操作的算法和符号表示而提出了本文所述的技术。这些算法描述和表示是被数据处理领域的技术人员用来最有效地向本领域的其它技术人员传达其工作的实质的手段。这些操作虽然是在功能上或逻辑上描述的,但被理解成由计算机程序实现。此外,在不失一般性的情况下,也经证明有时作为模块或用功能名称来提及这些操作布置是有利的。

[0057] 除非另外具体地说明,如从以上讨论显而易见的,应认识到的是遍及本描述,利用诸如“处理”或“计算”或“计量”或“确定”或“显示”等术语的讨论涉及计算机系统或类似电子计算设备的动作和过程,其对在计算机系统的寄存器和存储器、或其它此类信息存储、传输或显示设备内被表示为物理(例如,电子)量的数据进行操纵并变换。

[0058] 所述技术的某些方面包括在本文中以算法形式描述的过程步骤及指令。应注意的是可以用软件、固件或硬件来体现所述的过程步骤和指令,并且当用软件来体现时,可以被下载以驻于实时网络操作系统所使用的不同平台上并被从那里操作。

[0059] 本公开还涉及一种用于执行本文中的操作的装置。本装置可出于所需目的而被特别地构造,或者其可包括被存储在可以被计算机访问的计算机可读介质上的计算机程序选择性地激活或重配置的通用计算机。此类计算机程序可被存储于有形计算机可读存储介质中,其诸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、磁卡或光学卡、专用集成电路(ASIC)、或者适合于存储电子指令的任何类型的介质,并且每个被耦合到计算机系统总线。此外,在本说明书中提及的计算机可包括单个处理器,或者可以是采用多个处理器设计以实现增加的计算能力的架构。

[0060] 在本文中提出的算法和操作并非固有地与任何特定计算机或其它装置相关。还可将各种通用系统与根据本文中的教导的程序一起使用,或者构造更专业化的装置以执行所需方法步骤可证明是方便的。用于多种此类系统的所需结构连同等价变型一起对于本领域的技术人员而言将是显而易见的。另外,并未参考任何特定变成语言来描述本公开。应认识到的是可使用多种编程语言来实现如本文所述的本公开的教导,并且可针对本发明的实现和最佳方式的公开而提供对特定语言的任何参考。

[0061] 本公开非常适合于许多拓扑上的多种计算机网络系统。在本领域内,大型网络的配置和管理包括通过网络(诸如因特网)而通信耦合到不同计算机和存储设备的存储设备和计算机。

[0062] 前文对实施例的描述是出于举例说明和描述的目的提供的。其并不意图是穷举的或限制本公开。特定实施例的单独元件或特征一般地不限于该特定实施例,而是在适用的情况下是可互换的,并且可以在所选实施例中使用,即使并未具体地示出或描述。其也可以许多方式改变。不应认为此类变化是偏离本公开,并且所有此类修改意图包括在本公开的范围之内。

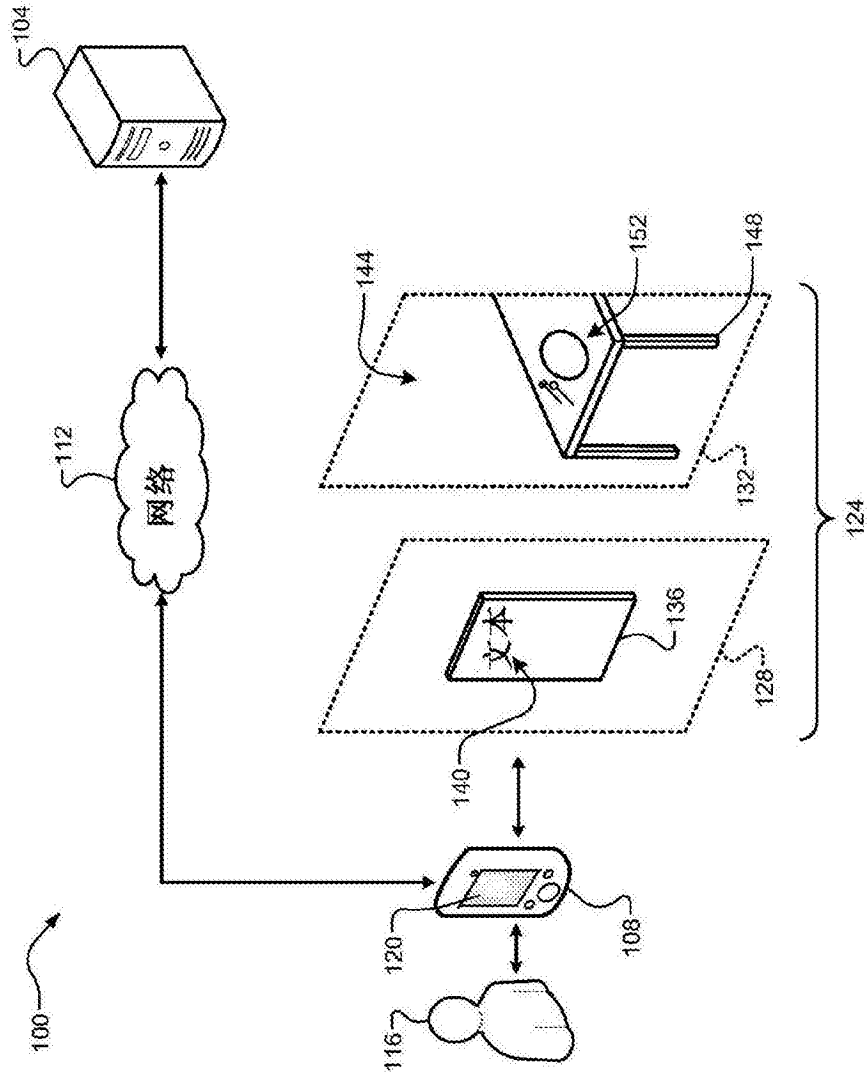


图1

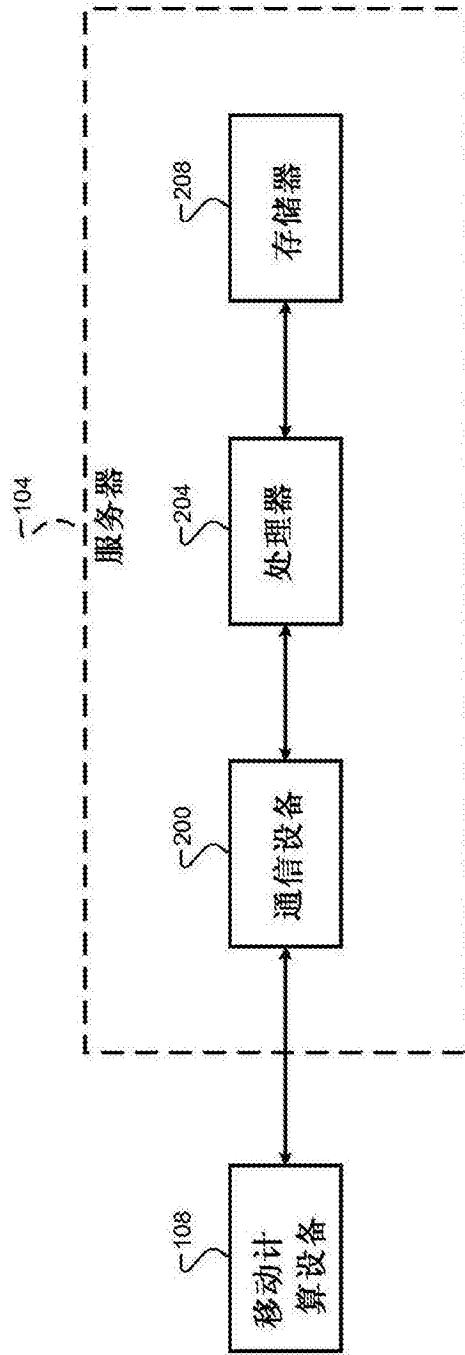


图2

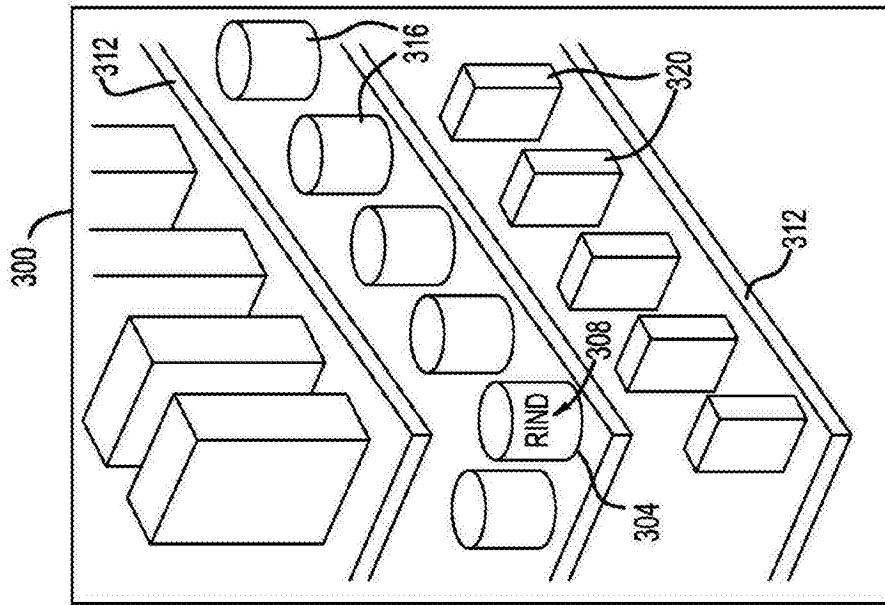


图3A

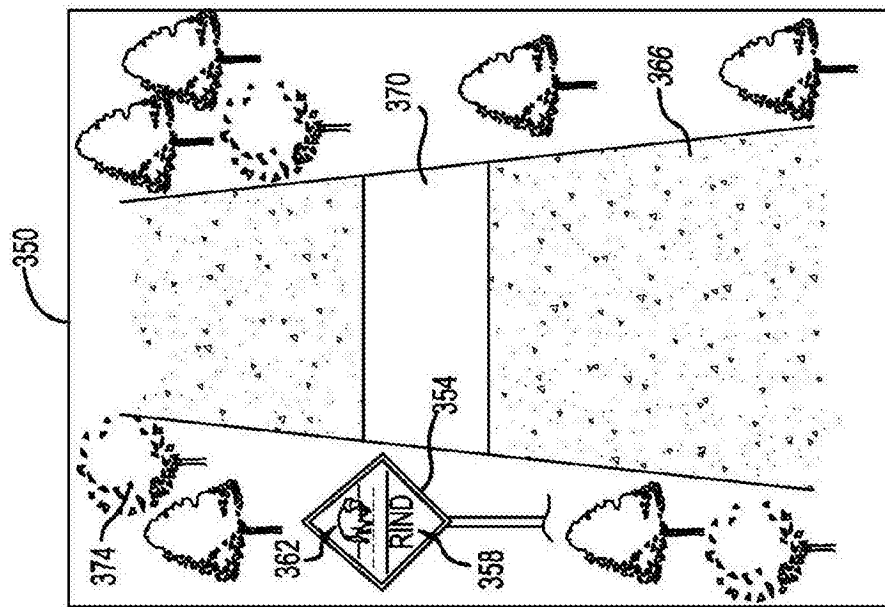


图3B

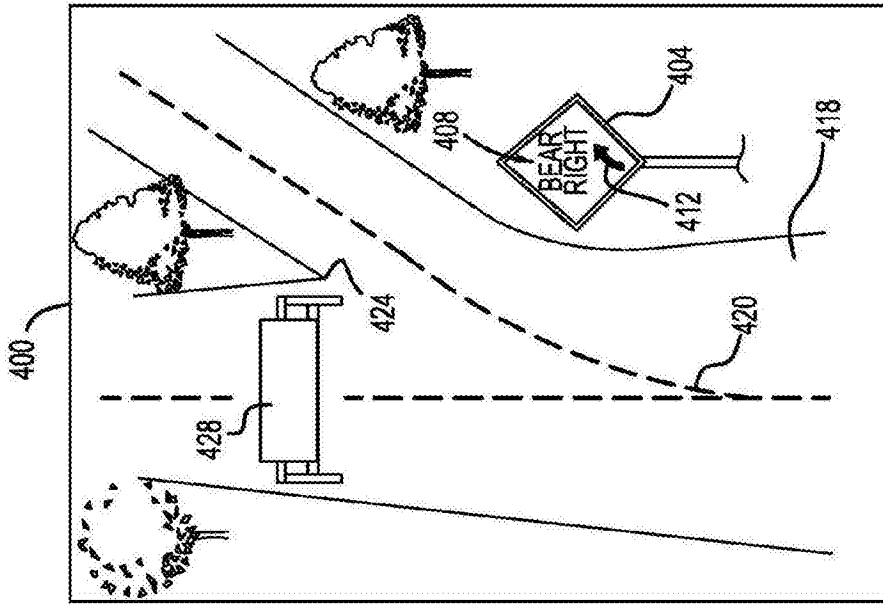


图4A

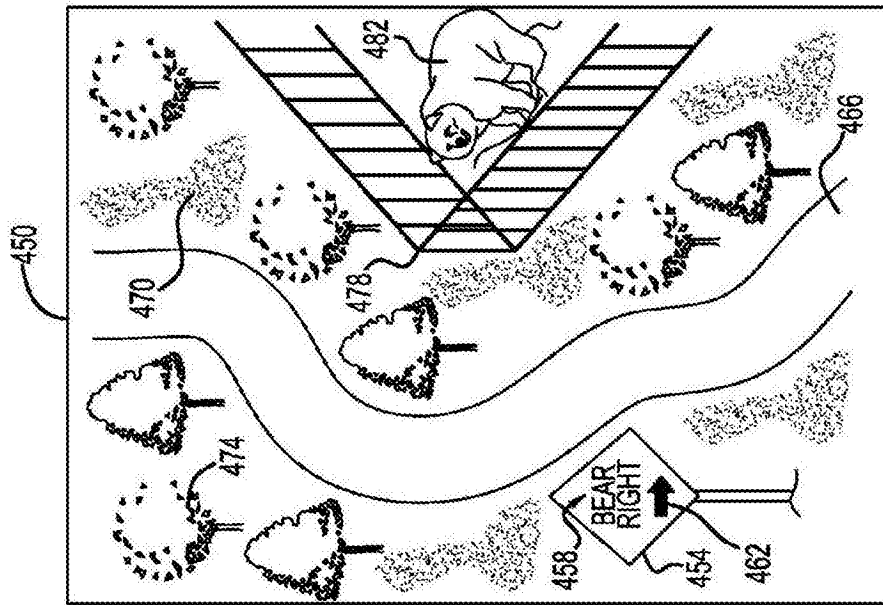


图4B

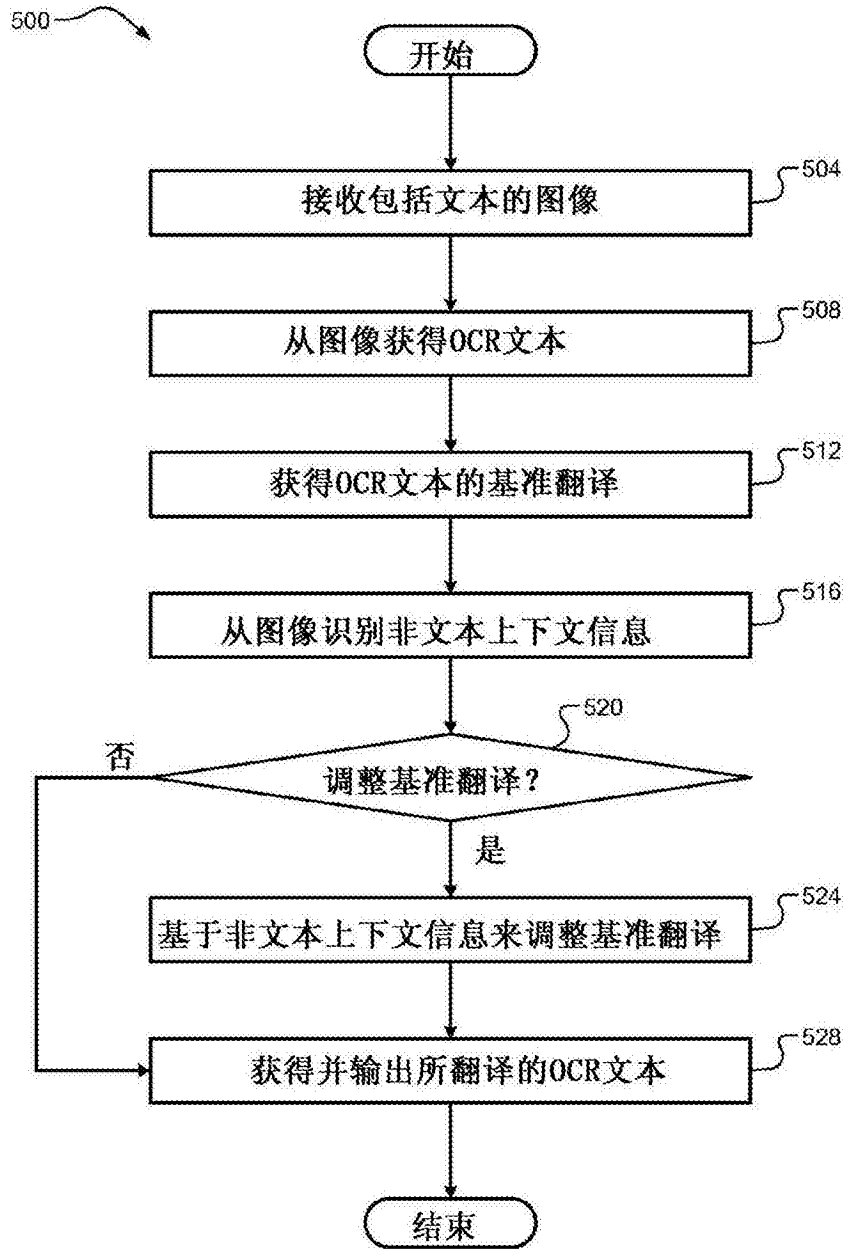


图5