



(12) 发明专利申请

(10) 申请公布号 CN 117635693 A

(43) 申请公布日 2024.03.01

(21) 申请号 202311116315.8

G06T 7/194 (2017.01)

(22) 申请日 2023.08.31

G06N 3/0464 (2023.01)

(30) 优先权数据

G06N 3/08 (2023.01)

22193455.7 2022.09.01 EP

(71) 申请人 皇家飞利浦有限公司

地址 荷兰艾恩德霍芬市

(72) 发明人 N·V·库尔卡尼 R·B·帕蒂尔

F·N·科克斯

(74) 专利代理机构 北京市金杜律师事务所

11256

专利代理师 成城

(51) Int. Cl.

G06T 7/62 (2017.01)

G06T 7/60 (2017.01)

G06T 7/11 (2017.01)

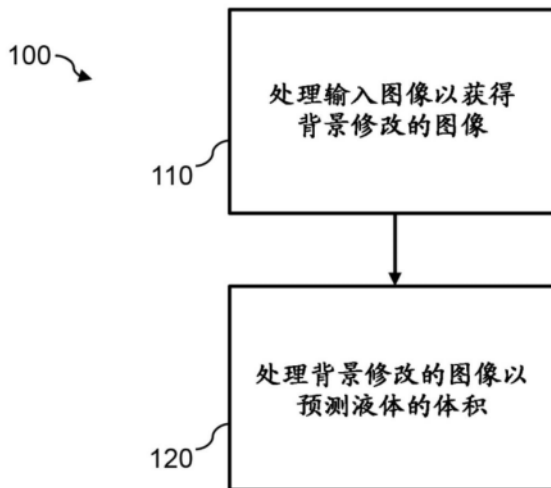
权利要求书2页 说明书12页 附图5页

(54) 发明名称

估计液体的体积

(57) 摘要

提出的概念因此旨在提供与估计至少部分透明的容器中的液体的体积有关的方案、解决方法、概念、设计、方法和系统。这可以通过自动处理包含液体的透明/半透明的容器的输入图像以获得背景修改的图像来实现,其中背景图像数据已经被去除或改变。然后,可以通过利用机器学习算法处理背景修改的图像来确定对容器中的液体的体积的预测。



1. 一种用于估计至少部分透明的容器中的液体的体积的方法,所述方法包括:

利用图像背景去除算法处理 (110) 包含液体的至少部分透明的容器的输入图像,以获得背景修改的图像,所述背景修改的图像是所述输入图像的修改版,其中所述输入图像的背景部分中的像素值已经被修改;

将所述背景修改的图像转换 (240) 成热图,其中所述热图中的像素的值描述了所述像素相对于其邻近的像素的像素强度值变化;以及

利用机器学习算法处理 (120) 所转换的背景修改的图像,以通过将所述所转换的背景修改的图像作为输入提供给卷积神经网络CNN,获得对所述至少部分透明的容器中的所述液体的体积的预测,第一CNN被训练以针对所述所转换的背景修改的图像预测所述至少部分透明的容器中的所述液体的体积。

2. 根据权利要求1所述的方法,其中利用所述机器学习算法处理所述所转换的背景修改的图像之前:

增强 (230) 所述背景修改的图像的对比度。

3. 根据权利要求1所述的方法,其中所述CNN使用训练算法进行训练,所述训练算法被配置接收训练输入的阵列和相应的已知输出,其中所述训练输入包括包含所述液体的所述至少部分透明的容器的图像,并且所述相应的已知输出包括所述至少部分透明的容器中的所述液体的体积。

4. 根据权利要求1所述的方法,其中所述CNN是对所述至少部分透明的容器的图像进一步训练的预训练模型,所述图像已经被手动注释。

5. 根据权利要求1至3中任一项所述的方法,其中利用所述图像背景去除算法处理所述输入图像包括:

将所述输入图像作为输入提供 (210) 给基于神经网络的处理算法,所述基于神经网络的处理算法被训练以针对所述输入图像预测:用于识别所述输入图像中的感兴趣的对象的掩码;以及

基于所预测的掩码,去除或改变 (220) 所述输入图像的背景图像数据,以生成所述背景修改的图像。

6. 根据权利要求5所述的方法,其中所述基于神经网络的处理算法包括:

卷积层,被配置为针对所述输入图像生成特征映射;

区域建议网络,被配置为基于所述特征映射,确定所述图像中的感兴趣的区域;

第一全卷积网络,被配置为基于所述感兴趣的区域,确定用于识别所述输入图像中的感兴趣的对象的掩码;以及

第二全卷积网络,被配置为基于所述感兴趣的区域,确定限制所述输入图像中的所述至少部分透明的容器的框。

7. 根据权利要求6中所述的方法,其中所述第二全卷积网络被配置为基于所述感兴趣的区域,确定所述感兴趣的对象的分类。

8. 根据权利要求6或7所述的方法,其中所述方法还包括:

计算所述框的尺寸;以及

基于所计算的所述框的尺寸与预定阈值的比较,重新调整所述输入图像的尺寸。

9. 根据权利要求5至8中任一项所述的方法,其中去除或改变所述输入图像的背景图像

数据包括:

将所述输入图像和所预测的掩码作为输入提供给Mask R-CNN,所述Mask R-CNN被训练以针对所述输入图像的每个像素预测像素分类。

10. 一种计算机程序,包括代码装置,所述代码装置用于当所述程序在处理系统上运行时,实施前述权利要求中的任一项所述的方法。

11. 一种处理器装置,被配置为:

利用图像背景去除算法处理输入图像,以获得背景修改的图像,所述背景修改的图像是所述输入图像的修改版,其中所述输入图像的背景部分中的像素值已经被修改;

将所述背景修改的图像转换成热图,其中所述热图中的像素的值描述了所述像素相对于其邻近的像素的像素强度值的变化;以及

利用机器学习算法处理所转换的背景修改的图像,以通过将所述所转换的背景修改的图像作为输入提供给卷积神经网络CNN,获得对至少部分透明的容器中的液体的体积的预测,第一CNN被训练以针对所述所转换的背景修改的图像预测所述至少部分透明的容器中的所述液体的体积。

12. 一种用于估计至少部分透明的容器中的液体的体积的系统,所述系统包括:

输入界面(410),被配置为获得包含液体的至少部分透明的容器的输入图像;以及

图像处理单元(420),被配置为:

利用图像背景去除算法处理所述输入图像,以获得背景修改的图像,所述背景修改的图像是所述输入图像的修改版,其中所述输入图像的所述背景部分中的像素值已经被修改;

将所述背景修改的图像转换成热图,其中所述热图中的像素的值描述了所述像素相对于其邻近像素的像素强度值的变化;以及

利用机器学习算法处理所转换的背景修改的图像,以通过将所述所转换的背景修改的图像作为输入提供给卷积神经网络CNN,获得所述对至少部分透明的容器中的所述液体的体积的预测,第一CNN被训练以针对所述所转换的背景修改的图像预测所述至少部分透明的容器中的所述液体的体积。

13. 一种液体体积估计系统,包括:

根据权利要求11或12所述的系统;以及

用户界面(44),被配置为从所述系统接收并显示对所述至少部分透明的容器中的所述液体的体积的预测。

估计液体的体积

技术领域

[0001] 本发明涉及估计液体的体积的领域,并且更具体地涉及估计在至少部分透明的容器中的液体的体积的领域。

背景技术

[0002] 估计在不同形状容器中的液体的体积,特别是半透明的一种,是非常困难的技术问题,这种技术问题的解决方法将有利于各种领域。尤其是,当估计在婴儿奶瓶或者吸乳器容器中的母乳的体积时,估计在半透明或透明的容器中的液体的体积可以是有用的。它也同样可以用于实验室和工业环境。

[0003] 然而,从单纯的图像来看,很难确定在透明或半透明的容器中的液体的体积,例如,由于容器的不同/非标准的形状以及容器由于其透明性或半透明性而与背景混合。

[0004] 因此,需要开发用于可靠地导出至少部分透明的容器中的液体的体积的方法,而不论容器的形状。

[0005] US2016/027173 AI描述了用于估计液体罐中的血液成分的量的方法的变型,包括在罐的图像内识别罐上的参照标记。

[0006] WO 2019/023376 AI描述了用于量化样品的体积的基于神经网络的方法。该方法包括提供样品、捕获样品的图像以及直接分类到多个的体积类或体积中的一个。

发明内容

[0007] 本发明由权利要求限定。

[0008] 根据依照本发明的一个方面的实施例,提供了一种用于估计至少部分透明的容器中的液体的体积的方法。

[0009] 该方法包括:利用图像背景去除算法处理包含液体的至少部分透明的容器的输入图像,以获得背景修改的图像,该背景修改的图像是输入图像的修改版,其中输入图像的背景部分中的像素值已经被修改;将背景修改的图像转换成热图,其中热图中的像素的值描述该像素相对于其邻近的像素的像素强度值的变化;以及利用机器学习算法处理所转换的背景修改的图像,以通过将所转换的背景修改的图像作为输入提供给卷积神经网络CNN,获得对至少部分透明中的容器中的液体的体积的预测,第一CNN被训练以针对所转换的背景修改的图像预测至少部分透明的容器中的液体的体积。

[0010] 因此,提出的概念旨在提供与估计至少部分透明的容器中的液体的体积的有关的方案、解决方法、概念、设计、方法和系统。特别地,实施例旨在提供用于根据处理容器的单个输入图像导出对透明或半透明的容器中的液体的体积的预测或估计的方法。

[0011] 换句话说,提出可以分析包含一定量的液体的至少部分半透明的容器的单个输入图像,以便获得对图像内的容器中的液体的体积的预测。可以通过使用用户通常已经可以访问的设备(例如智能手机)获得这样输入图像。

[0012] 通过仅仅需要单个输入图像,实施例可以支持在没有专门硬件的情况下的至少部

分透明的容器中的液体的体积的估计。因此可以通过提出的概念提供用于估计透明或半透明的容器中的液体的体积的改进方法。

[0013] 例如,通过自动处理包含液体的透明/半透明的容器的输入图像,以便获得背景修改的图像,其中背景图像数据已经被去除或改变,可以通过利用机器学习算法处理背景修改的图像来确定对容器中的液体的体积的预测。提出了当图像的背景(即不包含关于容器的信息的任何图像数据)已经被以某种方式去除或改变时,可以改进对容器中的液体的体积的预测。

[0014] 仅使用图像数据导出对至少部分透明的容器中的液体的体积的预测可以允许例如,被挤出到吸乳器容器中的乳汁的体积可以在没有使用任何专门硬件的情况下(即仅仅通过智能手机)被自动测量。

[0015] 本发明可以特别用于测量吸乳器容器或婴儿奶瓶中的乳汁的体积的情况。这些奶瓶通常是至少半透明的并且具有不同形状和尺寸。因此,一种能够估计吸乳器容器或婴儿奶瓶中的液体(即乳汁)的体积的、与形状/尺寸无关并且仅需要单个图像的方法将特别有益。换句话说,在本文公开的任何方法中,液体可以包括母乳,并且容器可以是吸乳器容器或婴儿奶瓶。

[0016] 最后,用于估计至少部分透明的容器中的液体的体积的改进方法可以通过提出的概念支持。

[0017] 已经发现通过实现液体像素和非液体像素之间增加的区别,热图的使用可以增加预测至少部分透明的容器中的液体的体积的准确性。热图可以是颜色热图,其给出了图像中相对于邻近的像素的颜色变化的强度。除了颜色变化的强度外或者代替颜色变化的强度,热图也可以考虑像素本身的颜色。然而,热图可以适用于灰度/黑白图像(例如,通过分析灰度值变化的强度)。通过增加液体像素和非液体像素之间的区别,基于它们的像素强度值,对液体的体积的预测可以更准确。

[0018] 在一些实施例中,利用机器学习算法处理的所转换的背景修改的图像之前可以:增强背景修改的图像的对比度。通过增强背景修改的图像的对比度,可以增加预测容器中的液体的体积的准确性。例如,对比度增强可以包括对比度有限的自适应直方图均衡化(CLAHE)。

[0019] 在一些实施例中,可以使用训练算法训练CNN,该训练算法被配置为接收训练输入的阵列和相应的已知输出,其中训练输入包括包含液体的至少部分透明的容器的图像,并且相应的已知输出包括至少部分透明的容器中的液体的体积。以这种方式,当被提供有包含液体的至少部分透明的容器的图像时,CNN可以被训练以输出对至少部分透明的容器中的液体的体积的估计。

[0020] 在一些实施例中,CNN可以是对至少部分透明的容器的图像进一步训练的预训练模型,这些图像已经被手动注释。这使得CNN在预测至少部分透明的容器中的液体的体积上变得尤其精通。

[0021] 在一些实施例中,利用图像背景去除算法处理输入图像可以包括:将输入图像作为输入提供给基于神经网络的处理算法,该基于神经网络的处理算法被训练以针对输入图像预测用于识别在输入图像中的感兴趣的对象的掩码;以及基于预测掩码,去除或改变输入图像的背景图像,以生成背景修改的图像。通过预测用于输入图像中的感兴趣的对象

(即,容器)的掩码,可以更简单地去除或改变背景。

[0022] 在一些实施例中,基于神经网络的处理算法可以包括:被配置为生成用于输入图像的特征映射的卷积层;被配置为基于特征映射,确定图像中的感兴趣的区域的第一全卷积网络;被配置为基于感兴趣的区域,确定用于识别输入图像中的感兴趣的对象的掩码的第二全卷积网络。首先通过生成特征映射,经由使用一个卷积层或多个卷积层,可以更简单地确定感兴趣的区域。然后,全卷积网络可以被用于图像的逐像素分割,以便确定用于识别感兴趣的对象(即,容器)的掩码以及限制容器的框。

[0023] 在一些实施例中,第二全卷积网络可以被配置为基于感兴趣的区域,确定感兴趣的对象的分类。例如,对象可以被分类为不相关的对象、容器、或特别是吸乳器容器或婴儿奶瓶。

[0024] 在一些实施例中,方法可以还包括:计算框的尺寸;以及基于所计算的框的尺寸与预定阈值的比较,重新调整输入图像的尺寸。

[0025] 通过计算框的尺寸,可以确定图像中的容器的尺寸的近似值,该近似值可以被用于计算输入图像的尺寸的重新调整,以便它们可以更紧密地匹配用于训练数据的图像,从而可以更简单地预测容器中的液体的体积。

[0026] 在一些实施例中,去除或改变输入图像的背景图像数据可以包括:将输入图像和所预测的掩码作为输入提供给Mask R-CNN,Mask R-CNN被训练以针对输入图像的每个像素预测像素分类。Mask R-CNN已经被证明对图像内的像素的分类特别成功,其具有远低于其它类型的神经网络的误差率。

[0027] 在一些实施例中,提供了计算机程序,该计算机程序包括代码装置,该代码装置用于当程序在处理系统上运行时,实施上述任何方法。

[0028] 根据本发明的另一方面,提供了处理器装置,被配置为:利用图像背景去除算法处理输入图像,以获得背景修改的图像,背景修改的图像是输入图像的修改版,其中输入图像的背景部分中的像素值已经被修改;将背景修改的图像转换成热图,其中热图中的像素的值描述了该像素相对于其邻近的像素的像素强度值的变化;以及利用机器学习算法处理所转换的背景修改的图像,以通过将所转换的背景修改的图像作为输入提供给卷积神经网络CNN,获得至少部分透明的容器中的液体的体积,第一CNN被训练以针对所转换的背景修改的图像预测至少部分透明的容器中的液体的体积。

[0029] 根据本发明的另一方面,提供了用于估计至少部分透明的容器中的液体的体积的系统,该系统包括:输入界面,被配置为获得包含液体的至少部分透明的容器的输入图像;以及图像处理单元,被配置为:利用图像背景去除算法处理输入图像,以获得背景修改的图像,背景修改的图像是输入图像的修改版,其中输入图像的背景部分中的像素值已经被修改;将背景修改的图像转换成热图,其中热图中的像素的值描述了像素相对于邻近的像素的像素强度值的变化;以及利用机器学习算法处理所转换的背景修改的图像,以通过将所转换的背景修改的图像作为输入提供给卷积神经网络CNN,获得至少部分透明的容器中的液体的体积,第一CNN被训练以针对所转换的背景修改的图像预测至少部分透明的容器中的液体的体积。

[0030] 在一些实施例中,上述系统或处理器装置可以还包括用户界面,该用户界面被配

置为从系统或处理器装置接收和显示对至少部分透明的容器中的液体的体积的预测。

[0031] 因此,可以提出用于估计至少部分透明的容器中的液体的体积的概念,并且这可以基于对至少部分透明的容器的输入图像的处理来完成。

[0032] 本发明的这些以及其它方面将从下文描述的实施例中显而易见并加以阐明。

附图说明

[0033] 为了更好地了解本发明,以及更清楚地展示如何实施本发明,现在仅以举例的方式参考附图,其中:

[0034] 图1是根据提出的实施例的用于估计至少部分透明的容器中的液体的体积的方法的简化流程图。

[0035] 图2是根据提出的实施例的用于估计至少部分透明的容器中的液体的体积的方法的更深度的流程图。

[0036] 图3是根据提出的实施例的利用图像背景去除算法处理输入图像的方法的更深度的流程图。

[0037] 图4是根据提出的实施例的用于估计至少部分透明的容器中的液体的体积的系统的简化框图。

[0038] 图5示出了可以在其中使用实施例的一个或多个部分的计算机的示例。

具体实施方式

[0039] 本发明将参考附图被描述。

[0040] 应当理解的是,详细的描述和具体的实施例虽然表明了装置、系统和方法的示例性实施例,但仅用于说明目的,并不旨在限制本发明的范围。本发明的这些和其它特征、方面,以及装置、系统、方法的优点将从下文的描述、所附权利要求书和附图被更好地理解。应当理解的是,附图仅仅是示意图,并不是按比例绘制。还应当理解的是,贯穿附图使用的相同的附图标记指示相同或相似的部件。

[0041] 本领域的技术人员可以从对附图、本公开、和所附权利要求书的研究来理解和实现对本公开实施例的变型。在权利要求书中,词语“包括”不排除其它元件或步骤,并且不定冠词“一”或“一个”不排除复数。

[0042] 应当理解的是,附图仅仅是示意图,并不是按比例绘制。还应当理解的是,贯穿附图使用的相同的附图标记指示相同或相似的部件。

[0043] 根据本公开的实施例,涉及与估计至少部分透明的容器中的液体的体积有关的各种技术、方法、方案和/或解决方法。根据提出的概念,若干可能的解决方法可以被单独或共同地实施。即,尽管这些可能的解决方法可以在下文被单独描述,但这些可能的解决方法的一个或多个可以在一种组合或另一种组合被实施。

[0044] 本发明的实施例旨在提供用于估计至少部分透明的容器中的液体的体积的方法。这可以通过自动处理包含液体的透明/半透明的容器的输入图像以获得背景修改的图像来实现,其中背景图像数据已经被去除或改变(例如,无效或替换为单个预定的像素值)。然后,可以通过利用机器学习算法处理背景修改的图像来确定对容器中的液体的体积的预测。

[0045] 因此,提出的概念旨在仅利用图像数据获得对至少部分透明的容器中的液体的体积的预测。例如,这可以允许在没有利用任何专门硬件(例如仅利用智能手机或配备摄像头的平板电脑)的情况下,自动测量被挤出到吸乳器容器(或婴儿奶瓶)中的乳汁的体积。被自动测量的乳汁的体积还可以被自动传递给用户和/或存储在存储器中,例如,在智能手机的存储器中。

[0046] 因此,实施例可以与测量吸乳器容器或婴儿奶瓶中的乳汁的体积的情况特别相关。这种容器/瓶子通常至少部分半透明的并且具有不同的形状和尺寸。因此,一种能够估计吸乳器容器或婴儿奶瓶中的液体(即乳汁)的体积的、与形状/尺寸无关并且仅需要单个图像的方法将会非常有用。换句话说,在本文公开的任何一种方法中,液体可以是乳汁,并且容器可以是吸乳器容器或婴儿奶瓶。

[0047] 现在参考图1,描述了根据提出的实施例的用于估计至少部分透明的容器中的液体的体积的方法100的流程图。

[0048] 该方法从步骤110开始,在步骤110,利用图像背景去除算法处理包含液体的至少部分透明的容器的输入图像,以获得背景修改的图像,背景修改的是输入图像的修改版,其中背景图像数据已经被去除或改变。可以使用现有的或常规的设备来获得输入图像,该设备包括用户已经拥有的相机。例如,常规的设备可以是以下中的一项:智能手机、平板电脑、笔记本电脑或任何其它合适的设备。之后将更充分地说明此处理步骤的细节。

[0049] 在步骤120中,利用机器学习算法处理背景修改的图像,以获得对至少部分透明的容器中的液体的体积的预测。

[0050] 在一些实施例中,利用机器学习算法处理背景修改的图像(即步骤120)可以包括:将背景修改的图像作为输入提供给卷积神经网络CNN,第一CNN被训练以针对背景修改的图像预测至少部分透明的容器中的液体的体积。CNN已经被证明在分析图像上特别地成功,并且能够以远低于其它类型的神经网络的误差率识别图像内的对象以及因此识别背景。

[0051] 人工神经网络(或简单地说,神经网络)的结构受到人脑的启发。神经网络由层组成,每个层包括多个神经元。每个神经元包括数学运算。特别是,每个神经元可以包括单一类型的变换的不同加权组合(例如,相同类型的变换,s型等,但具有不同的权重)。在处理输入数据的过程中,在输入数据上执行每个神经元的数学运算,以生成数值输出,并且在神经网络中的每个层的输出被依次馈送到下一层。最后一层提供输出。

[0052] 存在数个类型的神经网络,例如卷积神经网络(CNN)和循环神经网络(RNN)。本发明的示例性实施例采用基于CNN的学习算法,这是由于CNN已经被证明在分析图像上特别地成功,并且能够以比其它类型的神经网络低得多的误差率识别图像内的对象。

[0053] CNN通常包含若干层,包括卷积层、池化层以及全连接层。卷积层由一组可学习的过滤器组成,并从输入中提取特征。池化层是非线性下采样的一种形式,通过将一层中的多个神经元的输出组合到下一层中的单个神经元中来减小数据大小。全连接层将一层中的每个神经元连接到下一层中的所有神经元。

[0054] 训练机器学习算法的方法是众所周知的。通常,这种方法包括获得训练数据集,该数据集包括训练输入数据条目和相应的训练输出数据条目。向每个输入数据条目应用初始化机器学习算法,以生成所预测的输出数据条目。所预测的输出数据条目和相应的训练输出数据条目之间的误差被用于修改机器学习算法。这个过程可以被重复,直到误差收敛,并

且所预测的输出数据条目与训练输出数据条目是充分相似的(例如 $\pm 1\%$)。这通常被称为监督学习技术。

[0055] 例如,可以修改每个神经元的数学运算的权重,直到误差收敛。修改神经网络的已知方法包括梯度下降、反向传播算法等。

[0056] 用于可以在步骤120中使用的第二CNN的训练输入数据条目与包含液体的至少部分透明的容器的示例图像相对应。训练输出数据条目与至少部分透明的容器中的已知液体的体积相对应。此外,可以采用数个预处理方法来改进训练样本。换句话说,可以使用训练算法来训练第二CNN,该训练算法被配置为接收训练输入的阵列和相应的已知输出,其中训练输入包括包含液体的至少部分透明的容器的图像,并且相应的已知输出包括至少部分透明的容器中的液体的体积。以这样的方式,当被提供有包含液体的至少部分透明的容器的图像时,CNN可以被训练以输出对至少部分透明的容器中的液体的体积的预测。

[0057] CNN可以是对至少部分透明的容器的图像进一步训练的预训练模型,该图像已经被手动注释。例如,由于许多容器的复杂的形状,因此可以从所有角度和所有可能的体积水平拍摄容器的短视频。然后,可以从视频中提取图像,以便作为具有已知的液体的体积的训练输入。这允许CNN在预测至少部分透明的容器中的液体的体积上变得尤其精通,而不论容器的形状。

[0058] 现在参考图2,描述了用于估计至少部分透明的容器中的液体的体积的方法200的更深度的流程图。该方法从步骤210开始,在步骤210中,将输入图像作为输入提供给基于神经网络的处理算法,该算法可以被训练以针对输入图像预测用于识别输入图像中的感兴趣的对象的掩码。

[0059] 在步骤220中,然后,作为步骤110中所描述的图像背景去除算法的一部分,基于所预测的掩码被去除或改变输入图像的背景图像数据,从而生成背景修改的图像。通过预测用于输入图像(即容器)中的感兴趣的对象的掩码,可以更简单地去除或改变背景。例如,基于所预测的掩码识别输入图像(即容器)中的感兴趣的对象,可以将输入图像分割为背景和前景,其中前景只包含容器。然后,可以从输入图像中减去背景,从而留下只包含容器的图像。

[0060] 在一些实施例中,去除或改变输入图像的背景图像数据包括:将输入图像和所预测的掩码作为输入提供给Mask R-CNN,Mask R-CNN被训练以针对输入图像中的每个像素预测像素分类。Mask R-CNN已经被证明在分类图像内的像素上特别成功,其具有远低于其它类型的神经网络的误差率。Mask R-CNN可以特别地利用FBNET-C Mask RCNN神经网络构架,并且可以在COCO数据集上预训练模型。

[0061] 在训练过程期间,Mask R-CNN算法可以将用于每个采样的感兴趣的区域(RoI)(即半透明或透明的容器)的多任务损失函数定义为:

$$[0062] \quad L' = L_{cls} + L_{box} + L_{mask} \quad (1)$$

[0063] L_{cls} 是分类误差, L_{box} 是检测误差以及 L_{mask} 是分割误差。Mask R-CNN算法可以使用归一化指数函数和多任务函数来获得分类值和回归框参数值。可以使用S型函数来输出掩码值,以实现像素级实例分割。

[0064] 下面提供了用于Mask R-CNN构架的参数值的示例:

	卷积层 1	卷积层 2	卷积层 3	FCN 层 1	FCN 层 2	FCN 层 3
[0065] 输入尺寸	40x40x2	20x20x24	10x10x48	2400	1200	600
过滤	3x3	3x3	3x3	-	-	-
[0066] 器尺寸						
填充	1	1	1	-	-	-
激活	ReLU	ReLU	ReLU	Sigmoid	Sigmoid	Sigmoid
输出尺寸	20x20x24	10x10x48	5x5x96	1200	600	1

[0067] 神经网络的输出可以看作是二级分类,即像素是否属于感兴趣的对象(容器)。通过将问题视为二级分类,可以使用二元交叉熵(BCE)以用于训练:

$$E = -\frac{1}{n} \sum_{i=1}^n (t_i \times \log(y_i) + (1-t_i) \times \log(1-y_i)) \quad (2)$$

[0069] 掩码的标记可给出为如下:

$$\text{Mask} = \begin{cases} 1 & >_{i,j} = OOI \\ 0 & >_{i,j} = BG \end{cases} \quad (3)$$

[0071] 在等式(3)中,OOI表示感兴趣的对象以及BG表示背景。GT表示地面实况。

[0072] 在步骤230中,增强背景修改的图像的对比度。已经发现,通过增强背景修改的图像的对比度,可以增加预测容器中的液体的体积的准确性。对比度增强可以包括对比度有限的自适应直方图均衡化(CLAHE)。

[0073] 在步骤240中,可以将背景修改的图像转换成热图,其中热图中的像素的值描述了

像素相对于其邻近的像素的像素强度值变化。因此,在背景修改的图像已经被转换成热图的实施例中,任何指代背景修改的图像的步骤都可以同样地适用于所转换的背景修改的图像,即热图。例如,CNN可以被训练以针对所转换的背景修改的图像预测至少部分透明的容器中的液体的体积。在一些实施例中,可以在所转换的背景修改的图像或热图上训练第一CNN,其中训练输入可以包括包含液体的至少部分透明的容器的热图。在一些实施例中,CNN可以是对至少部分透明的容器的热图进一步训练的预训练模型,这些热图已经被手动注释。

[0074] 已经发现,通过实现液体像素和非液体像素之间的增加的分,热图的使用增加了预测至少部分透明的容器中的液体的体积的准确性。热图可以是颜色热图,其给出了图像中相对于邻近的像素的颜色变化的强度。除了颜色变化的强度之外或者代替颜色变化的强度,热图还可以考虑像素本身的颜色。热图也可以使用灰度/黑白图像。该技术放大了图像中细微的颜色变化,使其更容易(例如,通过CNN回归模型估计)估计至少部分半透明的容器中的体积,特别是当液体的颜色和容器的颜色相似时。通过增加液体像素和非液体像素之间的区别,基于它们的像素强度值,可以更准确的预测液体的体积,而不论是透明的还是半透明的容器。

[0075] 在步骤120中,利用机器学习算法处理背景修改的图像(或热图),以获得对至少部分透明的容器中的液体的体积的预测。

[0076] 机器学习算法可以是CNN。基于回归的CNN将图像的特征映射到数值。在这种情况下,CNN模型可以学习容器在容纳特定液体的体积时的样子。在一些实施例中,RESNET-50主干网CNN模型可以与具有ReLU激活函数的五个卷积层一起使用。可以采用一周期学习率策略来确定用于随机梯度下降动量(SGDM)优化器的最佳学习率。如下提供了用于训练CNN以获得至少部分透明的容器中的液体的体积的预测的一组示例参数:

	参数	值
[0077]	解算器	SGDM
	动量	0.60
	MaxEpochs	120
[0078]	学习率	0.003
	训练数据	25000
	验证数据	2500

[0079] 现在参考图3,描述了根据提出的实施例的利用图像背景去除算法处理输入图像的方法300的更深度的流程图。换句话说,描述了图2中的步骤210的更深度的描述。

[0080] 在步骤310中,使用包括与步骤210相关的基于神经网络的处理算法的部分的卷积层,生成用于输入图像的特征映射。通过生成特征映射,经由一个使用卷积层或多个卷积层,可以随后更简单地确定感兴趣的区域。

[0081] 在步骤320中,使用包括与步骤210相关的基于神经网络的处理算法的的部分的区域

建议网络,基于特征映射确定图像中的感兴趣的区域。通过确定感兴趣的区域,可以随后更简单地识别感兴趣的对象。可以将感兴趣的区域输出为以X,Y,W和H的形式的一组坐标,其中X和Y表示感兴趣的区域的左上角,以及W和H表示感兴趣的区域的宽和高。

[0082] 在步骤330中,使用包括与步骤210相关的基于神经网络的处理算法的部分的第一全卷积网络,基于所确定的感兴趣的区域,确定用于识别在输入图像中的感兴趣的对象的掩码。在步骤340中,使用包括与步骤210相关的基于神经网络的处理算法的部分的第二全卷积网络,基于所确定的感兴趣的区域,确定在输入图像中限定至少部分透明的容器的框。全卷积网络精通图像的逐像素分割,以便确定掩码和边界框。可以将边界框输出为以X,Y,W和H的形式的一组坐标,其中X和Y表示边界框的左上角,W和H表示边界框的宽和高。

[0083] 在一些实施例中,第二全卷积网络还可以被配置为基于感兴趣的区域确定感兴趣的对象的分类。例如,对象可以被分类为不相关的对象、容器或特别是吸乳器容器或婴儿奶瓶。

[0084] 在步骤350中,计算在输入图像中限定至少部分透明的容器的所确定的框的尺寸。在步骤360中,将所计算的框的尺寸与预定的阈值进行比较,并且基于该比较,可以重新调整输入图像的尺寸。通过计算框的尺寸,可以确定图像中的容器的尺寸的近似值,该近似值可以用于计算输入图像的尺寸的重新调整,以便该图像可以更接近地匹配训练数据中使用的图像。这可以允许更简单地预测容器中的液体的体积。

[0085] 因为用户可以输入容器的来自各种角度和距离的图像,所以相同尺寸的容器可能在图像中显示出不同的尺寸。例如,从较远的地方拍摄的容器的图像将比从较近的地方拍摄的相同的容器的图像显得更小。为了克服这个问题,可以利用所计算的边界框来确定图像是否需要重新调整尺寸。例如,如果边界框的区域低于某个阈值区域,则图像可以被重新调整尺寸直到该区域更接近阈值。例如,以下等式(4)可以被用来确定是否需要调整尺寸:

$$[0086] \quad \mathit{rescale}_{factor} = \begin{cases} (area > Threshold) 0.5 \\ (area < Threshold) 2.0 \end{cases} \quad (4)$$

[0087] 换句话说,如果边界框的区域大于阈值区域,则图像被缩小0.5倍(或任何其它合适的倍数)。类似地,如果该区域小于阈值,则图像被放大2倍(或任何其它合适的倍数)。在重新调整尺寸后(如果有的话),则可以确定用于识别图像(即容器)中的感兴趣的对象的掩码(即步骤330),或者可以根据重新调整尺寸的输入图像和对应的掩码改变或去除背景(即步骤220)。

[0088] 在一些实施例中,提供了计算机程序,该计算机程序包括代码装置,该代码装置用于当程序在处理系统上运行时,实现上述任何方法。

[0089] 现在参考图4,描述了根据提出的实施例的用于估计至少部分透明的容器中的液体的体积的系统400的简化框图。被配置为用于估计的至少部分透明的容器中的液体的体积的系统包括输入界面410,以及图像处理单元420。图像处理单元420还可以被称为处理器装置。

[0090] 所述系统被配置为通过处理包含液体的至少部分透明的容器的输入图像415,估计至少部分透明的容器中的液体的体积。获得包含液体的至少部分透明的容器的输入图像415。该系统输出输出430,输出430包括对至少部分透明的容器中的液体的体积的预测。基于对输入图像的处理,预测输出430,以便获得背景修改的图像。然后利用机器学习算法处

理背景修改的图像,以便获得对液体的体积的预测。

[0091] 更详细地说,输入界面410接收包含液体的至少部分透明的容器的输入图像415。输入界面410将输入图像415提供给图像处理单元420,该图像处理单元420被配置为利用图像背景去除算法处理输入图像415,以便获得背景修改的图像,然后使用机器学习算法处理背景修改的图像,以获得对至少部分透明的容器中的液体的体积的预测的输出430。背景修改的图像是输入图像415的修改版,其中背景图像数据已经被去除或改变。

[0092] 在一些实施例中,用户界面440可以被配置为从系统400接收并显示对至少部分透明容器中的液体的体积的预测的输出430。

[0093] 图5示出了可以在其中使用实施例的一个或多个部分的计算机500的示例。上文讨论的各种操作可以利用计算机500的能力。在这点上,应当理解的是,系统功能块可以在单个计算机上运行或者可以分布在数台计算机和位置上(例如,经由互联网连接)。

[0094] 计算机500包括,但不限于PC、工作站、笔记本电脑、PDA、掌上设备、服务器、存储器等。通常,就硬件体系结构而言,计算机500可以包括一个或多个处理器510、存储器520和一个或多个经由本地接口(未示出)通信耦合的I/O设备530。本地接口可以例如但不排他地是本领域已知的一个或多个总线或其它有线或无线连接。本地接口可以具有额外的元件,例如控制器、缓冲区(缓存)、驱动器、中继器和接收器,以实现通信。此外,本地接口可以包括地址、控制和/或数据连接,以使得能够在上述部件之间实现适当的通信。

[0095] 处理器510是用于执行可以存储在存储器520中的软件的硬件设备。处理器510实际上可以是任何定制的或市售的处理器、中央处理元件(CPU),数字信号处理器(DSP)或者与计算机500相关联的数个处理器之间的辅助处理器,并且处理器510可以是基于微处理器(以微芯片的形式)的半导体或微处理器。

[0096] 存储器520可以包括易失性存储器元件(例如,随机存取存储器(RAM),诸如动态随机存取存储器(DRAM)、静态随机存取存储器(SRAM)等)和非易失性存储器元件(例如,ROM、可擦除可编程只读存储器(EPROM)、电子可擦除可编程只读存储器(EEPROM)、可编程只读存储器(PROM)、磁带、光盘只读存储器(CD-ROM)、磁盘、软盘、磁带盒、盒式磁带或类似物等)的任何一个或组合。此外,存储器520可以包含电子、磁性、光学和/或其它类型的存储介质。注意,存储器520可以具有分布式体系结构,其中各种组件彼此远离,但可以由处理器510访问。

[0097] 存储器520中的软件可以包括一个或多个单独的程序,每个单独的程序包括用于实现逻辑功能的可执行指令的有序列表。存储器520中的软件包括根据示例性实施例的合适的操作系统(O/S)550、编译器560、源代码570和一个或多个应用程序580。如图所示,应用程序580包括用于实现示例性实施例的特征和操作的众多功能部件。根据示例性实施例,计算机500的应用程序580可以表示各种应用程序、计算单元、逻辑、功能单元、过程、操作、虚拟实体和/或模块,但应用程序580并不意味着是限制。

[0098] 操作系统550控制其它计算机程序的执行,并提供调度、输入-输出控制、文件和数据管理、内存管理和通信控制及相关服务。发明人预期,用于实施示例性实施例的应用程序580可以适用于所有市售操作系统。

[0099] 应用程序580可以是源程序、可执行程序(目标代码)、脚本或包括要执行的一组指令的任何其它实体。当应用程序580是源程序时,则通常经由编译器(如编译器560)、汇编

器、解释器等来翻译程序,该程序可以或可以不包括在存储器520中,从而与O/S550连接来正确地操作。此外,可以将应用程序580编写为具有数据和方法类的面向对象的编程语言,或具有例程、子例程和/或函数的程序设计语言,例如但不限于,C、C++、C#、Pascal、Python、BASIC、API calls、HTML、XHTML、XML、ASP scripts、JavaScript、FORTRAN、COBOL、Perl、Java、ADA、.NET等。

[0100] I/O设备530可以包括输入设备,例如但不限于鼠标、键盘、扫描器、麦克风、摄像头等。此外,I/O设备530还可以包括输出设备,例如但不限于打印机、显示器等。最后,I/O设备530可以进一步包括传送输入和输出的设备,例如但不限于NIC或调制器/解调器(用于访问远程设备、其它文件、设备、系统或网络)、射频(RF)或其它收发器、电话接口、桥接器、路由器等。I/O设备530还包括用于在各种网络(诸如互联网或内联网)上进行通信的部件。

[0101] 如果计算机500是PC、工作站、智能设备等,存储器520中的软件可以还包括基本输入输出系统(BIOS)(为了简单起见省略)。BIOS是一组基本的软件例程,该软件例程在启动时初始化和测试硬件,启动O/S550,并支持硬件设备之间的数据传输。BIOS存储在某种类型的只读存储器中,诸如ROM、PROM、EPROM、EEPROM等,以便在计算机800被激活时BIOS可以被执行。

[0102] 当计算机500运行时,处理器510被配置为执行存储在存储器520中的软件,将数据传送到存储器520和从存储器520传送数据,并依据该软件通常地控制计算机500的操作。应用程序580和O/S550全部或部分由处理器510读取,可能在处理器510内缓冲,然后执行。

[0103] 当应用程序580在软件中被实施时,应当注意的是,应用程序580可以虚拟地存储在计算机可读介质上,以供任何计算机相关系统或方法使用或与之连接。在本文的上下文中,计算机可读介质可以是电子、磁性、光学或其它物理设备或装置,其可以包含或存储供计算机相关系统或方法使用或连接的计算机程序。

[0104] 应用程序580可以体现在任何计算机可读介质中,以供指令执行系统、装置、设备使用或连接,诸如基于计算机的系统、包含处理器的系统或能够从指令执行系统、装置或设备获取指令并执行指令的其它系统。在本文的上下文中,“计算机可读介质”可以是任何可以存储、传送、传播或传输程序的装置,以供指令执行系统、装置、或设备使用或连接。计算机可读介质可以例如但不排他地是电子、磁性、光学、电磁、红外或半导体系统、装置、设备或传播介质。

[0105] 图1-3的方法以及图4的系统可以在硬件或软件或二者的混合(例如,作为固件在硬件设备上运行)中被实施。在实施例部分或全部在软件中实现的程度上,过程流程图所示的功能步骤可以通过合适编程的物理计算设备(诸如一个或多个中央处理单元(CPU)或图形处理单元(GPU))来执行。每个过程以及流程图所示的各个组成步骤可以由相同或不同的计算设备执行。根据实施例,计算机可读存储介质存储包括计算机程序代码的计算机程序,该计算机程序代码被配置为当程序在一个或多个物理计算设备上运行时使一个或多个物理计算设备执行如上所述的编码或解码方法。

[0106] 存储介质可能包括易失性和非易失性计算机存储器,如RAM、PROM、EPROM和EEPROM、光盘(如CD、DVD、BD)、磁性存储介质(如硬盘和磁带)。各种存储介质可以固定在计算设备内,或者可以是可移动的,以便存储在其上的一个或多个程序可以加载到处理器中。

[0107] 在实施例部分或全部在硬件中实现的程度上,图4的框图中所示的框可以是单独

的物理组件,或单个物理组件的逻辑细分,或者可以在一个物理组件中以集成的方式全部实现。附图中所示的一个框的功能可以在实施方式中的多个组件之间划分,或者附图中所示的多个框的功能可以在实施方式中的单个组件中组合。适合在本发明的实施例中使用的硬件部件包括但不限于常规的微处理器、专用集成电路(ASICs)和现场可编程门阵列(FPGA)。一个或多个框可以被实施为执行某些功能的专门硬件的组合,以及执行其它功能的一个或多个可编程微处理器和相关的电路。

[0108] 通过研究附图、公开内容和所附权利要求,在实践要求保护的本发明时,可以本领域技术人员理解和实现所公开的实施例的其他变型。在权利要求中,词语“包括”不排除其他元件或步骤,并且不定冠词“一”或“一个”不排除多个。单个处理器或其它单元可以完成权利要求书中列举的数个项目的功能。在相互不同的从属权利要求书中列举某些措施这一事实并不表明这些措施的组合不能被用于有利的目的。如果上面讨论了计算机程序,则计算程序可以存储/分布在合适的介质上,诸如光存储介质或与其他硬件一起或作为其他硬件的一部分提供的固态介质,但也可以以其它形式分布,诸如通过互联网或其他有线或无线电信系统。如果在权利要求书或说明书中使用术语“适于”,则应注意术语“适于”意在等同于术语“被配置为”。权利要求书中的任何参考符号不应解释为限制其范围。

[0109] 图中的流程图和框图说明了根据本发明的各种实施例的系统、方法和计算机程序产品的可能实现的体系结构、功能和操作。在这方面,流程图或框图中的每个框可以表示指令的模块、段或部分,其包括用于实现指定逻辑功能的一个或多个可执行指令。在一些备选的实施方式中,在框中注明的功能可能不按图中注明的顺序出现。例如,连续显示的两个框实际上可以基本上同时执行,或者这些框有时可以以相反的顺序执行,这取决于所涉及的功能。还应注意的是,框图和/或流程图中的每个框、以及框图和/或流程图中的框的组合可以通过基于特殊目的硬件的系统来实现,该系统执行指定的功能或行为,或执行特殊目的硬件和计算机指令的组合。根据本发明的各种实施例的系统、方法和计算机程序产品的可能实现的体系结构、功能和操作。在这方面,流程图或框图中的每个框可以表示指令的模块、段或部分,其包括用于实现指定逻辑功能的一个或多个可执行指令。在一些备选的实施方式中,在框中注明的功能可能不按图中注明的顺序出现。例如,连续显示的两个框实际上可以基本上同时执行,或者这些框有时可以以相反的顺序执行,这取决于所涉及的功能。还应注意的是,框图和/或流程图中的每个框、以及框图和/或流程图中的框的组合可以通过基于特殊目的硬件的系统来实现,该系统执行指定的功能或行为,或执行特殊目的硬件和计算机指令的组合。

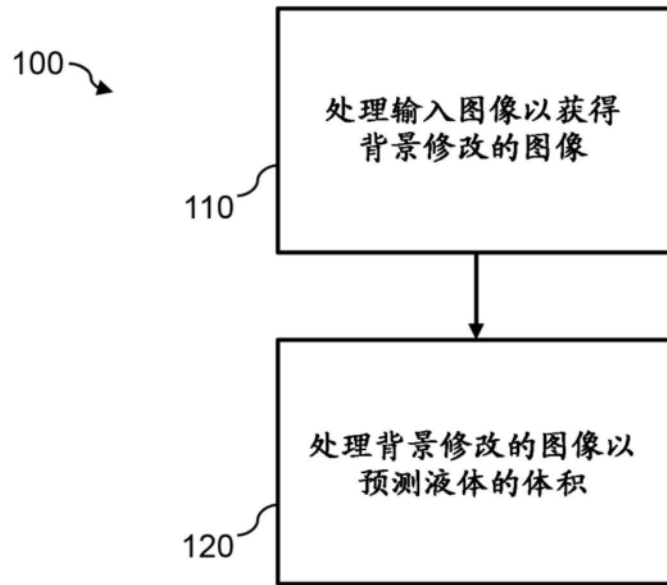


图1

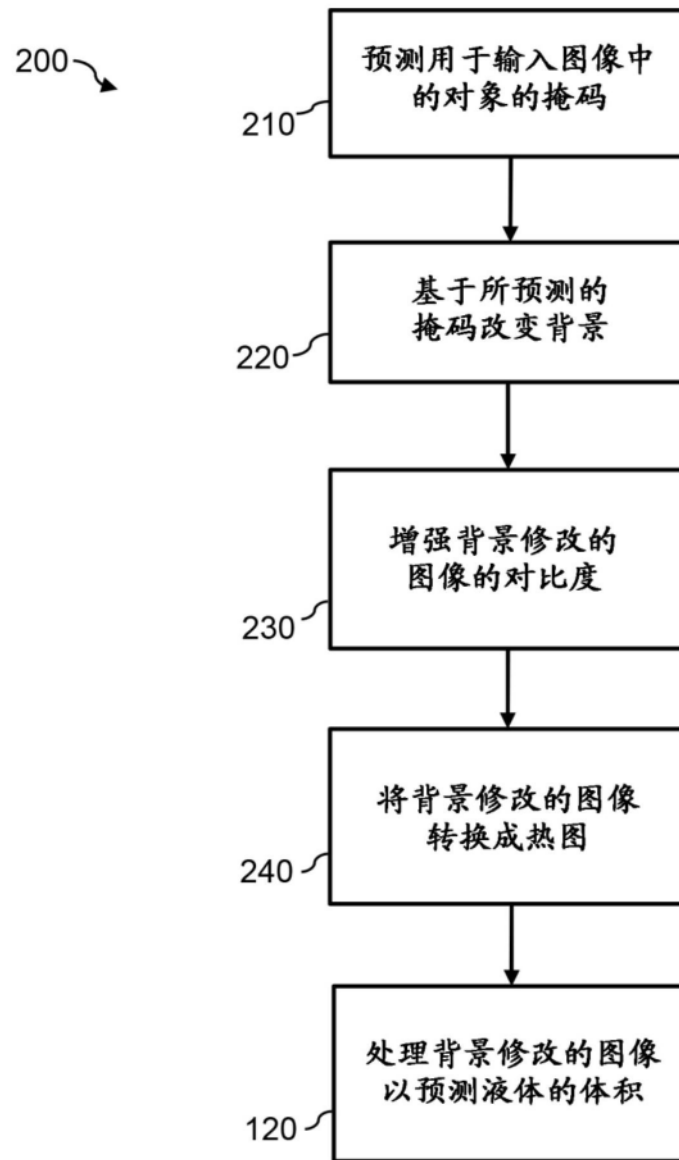


图2

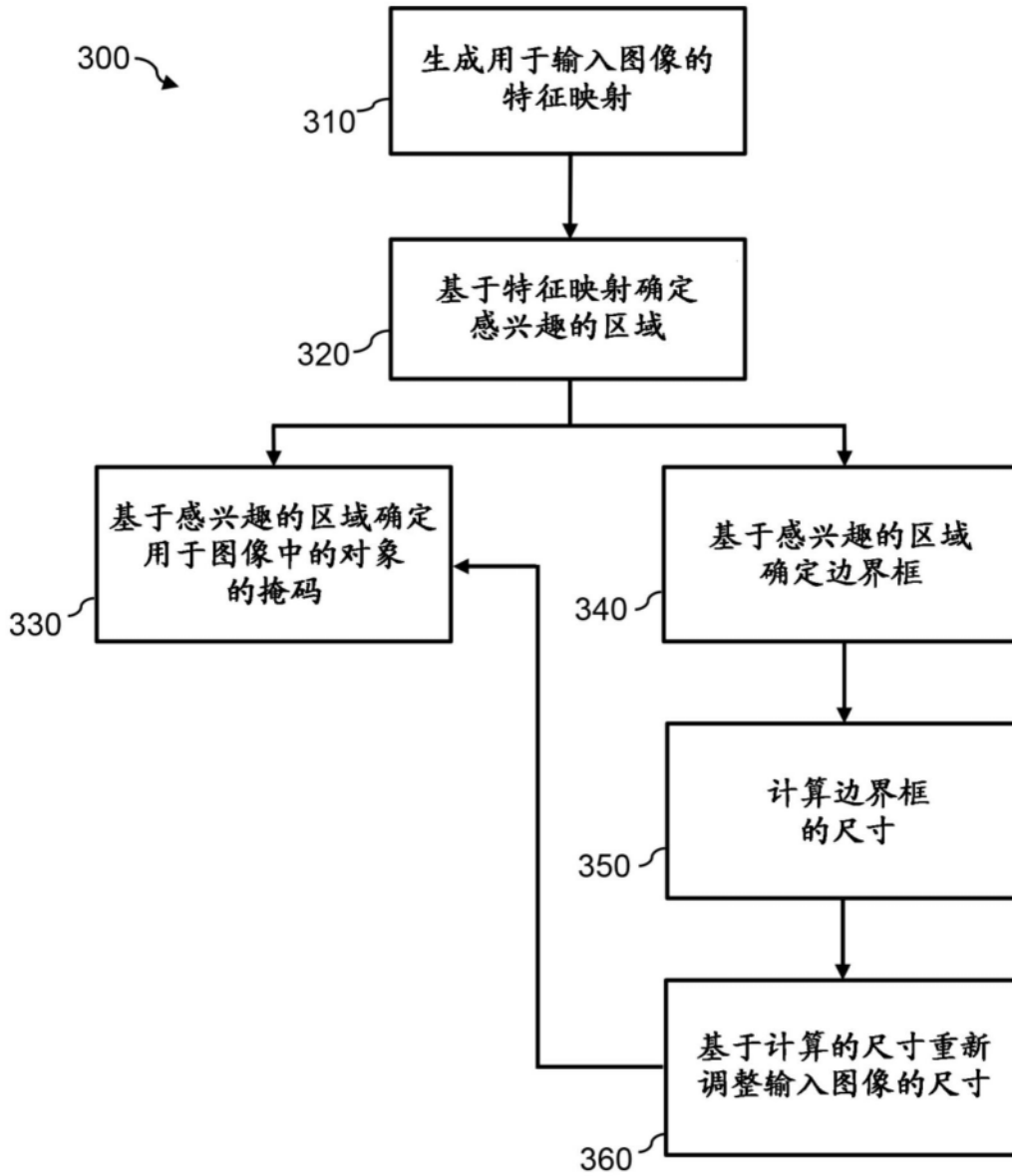


图3

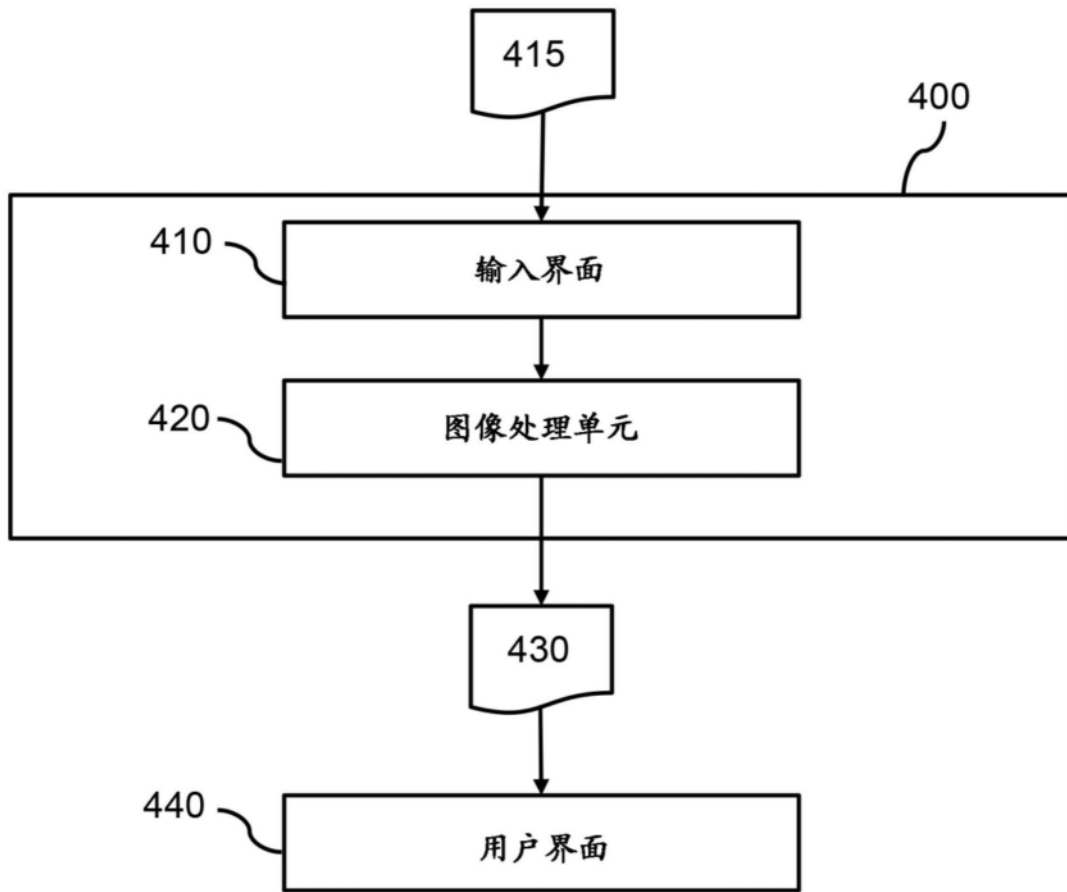


图4

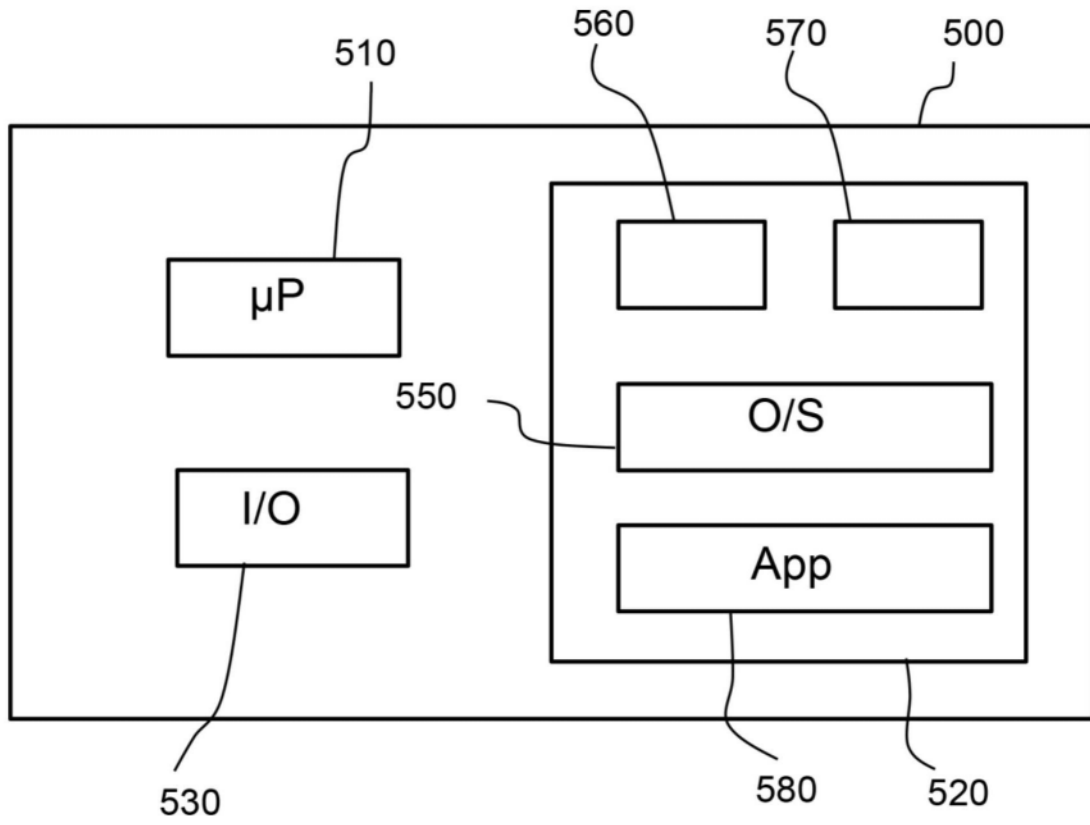


图5