



(12) 发明专利

(10) 授权公告号 CN 114787648 B

(45) 授权公告日 2023. 11. 10

(21) 申请号 202080082466.4

(22) 申请日 2020.08.28

(65) 同一申请的已公布的文献号
申请公布号 CN 114787648 A

(43) 申请公布日 2022.07.22

(30) 优先权数据
62/942,113 2019.11.30 US
63/001,445 2020.03.29 US

(85) PCT国际申请进入国家阶段日
2022.05.26

(86) PCT国际申请的申请数据
PCT/US2020/048604 2020.08.28

(87) PCT国际申请的公布数据
W02021/108002 EN 2021.06.03

(73) 专利权人 波士顿偏振测定公司
地址 美国加利福尼亚州

(72) 发明人 阿加斯塔·卡尔拉
维吉·塔玛兹言
瑟普里斯·K·拉奥
卡提克·文卡塔拉曼
拉米许·拉斯卡 阿丘塔·卡丹比

(74) 专利代理机构 北京泰吉知识产权代理有限公司 11355
专利代理师 史瞳 谢琼慧

(51) Int.Cl.
G01S 3/786 (2006.01)
G03H 1/08 (2006.01)
G06V 20/10 (2022.01)
G06V 10/147 (2022.01)
G06V 10/82 (2022.01)
G06N 3/0464 (2023.01)
G06N 3/08 (2023.01)

(56) 对比文件
CN 110363070 A, 2019.10.22
US 2013293871 A1, 2013.11.07
JP 2017125809 A, 2017.07.20
US 2016261844 A1, 2016.09.08
WO 2007057578 A1, 2007.05.24
JP 2018128265 A, 2018.08.16
WO 2019089049 A1, 2019.05.09
US 2018177461 A1, 2018.06.28

审查员 陈溥

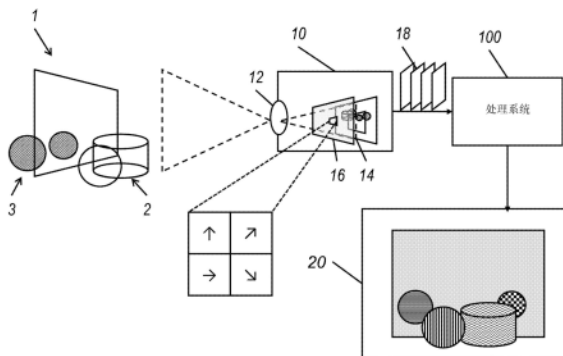
权利要求书3页 说明书20页 附图19页

(54) 发明名称

用于使用偏振提示进行透明对象分段的系统和方法

(57) 摘要

一种用于计算对场景的图像的预测的计算机实施的方法包含:接收场景的一个或多个偏振原始帧,所述偏振原始帧利用偏振滤波器在不同线性偏振角下捕获;从所述偏振原始帧提取一个或多个偏振表示空间中的一个或多个第一张量;以及基于所述一个或多个偏振表示空间中的所述一个或多个第一张量计算关于所述场景中的一个或多个在光学上具有挑战性的对象的预测。



1. 一种用于计算关于场景的图像的预测的计算机实施的方法,所述方法包括:
接收场景的一个或多个偏振原始帧,所述偏振原始帧利用偏振滤波器在不同线性偏振角下捕获;
从所述偏振原始帧提取一个或多个偏振表示空间中的一个或多个第一张量;以及
基于所述一个或多个偏振表示空间中的所述一个或多个第一张量计算关于所述场景中的一个或多个在光学上具有挑战性的对象的预测,
其中所述预测包括分段掩码,
其中所述计算所述预测包括:将所述一个或多个第一张量供应到一个或多个相应卷积神经网络CNN主干,
其中所述一个或多个CNN主干中的每一个被配置成在多个不同比例下计算多个模张量,且
其中所述计算所述预测进一步包括:融合由所述一个或多个CNN主干在相同比例下计算的所述模张量。
2. 根据权利要求1所述的计算机实施的方法,其中所述一个或多个偏振表示空间中的所述一个或多个第一张量包括:
DOLP表示空间中的线性偏振度DOLP图像;以及
AOLP表示空间中的线性偏振角AOLP图像。
3. 根据权利要求1所述的计算机实施的方法,其中所述一个或多个第一张量进一步包括一个或多个非偏振表示空间中的一个或多个非偏振张量,且
其中所述一个或多个非偏振张量包括强度表示空间中的一个或多个强度图像。
4. 根据权利要求3所述的计算机实施的方法,其中所述一个或多个强度图像包括:
第一颜色强度图像;
第二颜色强度图像;以及
第三颜色强度图像。
5. 根据权利要求1所述的计算机实施的方法,其中所述融合所述相同比例下的所述模张量包括:
串接所述相同比例下的所述模张量;
将所述模张量供应到注意力子网络以计算一个或多个注意力地图;以及
基于所述一个或多个注意力地图对所述模张量进行加权以计算所述比例的融合张量。
6. 根据权利要求5所述的计算机实施的方法,其中所述计算所述预测进一步包括将每一比例下计算的所述融合张量供应到被配置成计算所述分段掩码的预测模块。
7. 根据权利要求1所述的计算机实施的方法,其中所述分段掩码被供应到机器人拾取臂的控制器。
8. 根据权利要求1、2、3或4所述的计算机实施的方法,其中所述预测包括基于所述一个或多个在光学上具有挑战性的对象对所述一个或多个偏振原始帧的分类。
9. 根据权利要求1、2、3或4所述的计算机实施的方法,其中所述预测包括所述一个或多个偏振原始帧中描绘的所述一个或多个在光学上具有挑战性的对象的一个或多个检测到的特征。
10. 根据权利要求1所述的计算机实施的方法,其中所述计算所述预测包括将所述一个

或多个偏振表示空间中的所述一个或多个第一张量供应到统计模型,且

其中使用包括所述一个或多个偏振表示空间中的训练第一张量的训练数据和标签训练所述统计模型。

11. 根据权利要求10所述的计算机实施的方法,其中所述训练数据包括:

从由偏振相机捕获的数据计算的所述一个或多个偏振表示空间中的源训练第一张量;
以及

经由包括旋转的仿射变换从所述源训练第一张量生成的额外训练第一张量。

12. 根据权利要求11所述的计算机实施的方法,其中,当所述额外训练第一张量包括线性偏振角AOLP图像时,生成所述额外训练第一张量包括:

使所述额外训练第一张量旋转某一角度;以及

使所述AOLP图像的像素值反向旋转所述角度。

13. 一种计算机视觉系统包括:

偏振相机,其包括偏振滤波器;以及

处理系统,其包括处理器和存储指令的存储器,所述指令在由所述处理器执行时致使所述处理器:

接收场景的一个或多个偏振原始帧,所述偏振原始帧利用偏振滤波器在不同线性偏振角下捕获;

从所述偏振原始帧提取一个或多个偏振表示空间中的一个或多个第一张量;

基于所述一个或多个偏振表示空间中的所述一个或多个第一张量计算关于所述场景中的一个或多个在光学上具有挑战性的对象的预测,其中所述预测包括分段掩码;

通过将所述一个或多个第一张量供应到一个或多个相应卷积神经网络CNN主干来计算所述预测,其中所述一个或多个CNN主干中的每一个被配置成在多个不同比例下计算多个模张量;以及

融合由所述一个或多个CNN主干在相同比例下计算的所述模张量。

14. 根据权利要求13所述的计算机视觉系统,其中所述一个或多个偏振表示空间中的所述一个或多个第一张量包括:

DOLP表示空间中的线性偏振度DOLP图像;以及

AOLP表示空间中的线性偏振角AOLP图像。

15. 根据权利要求13所述的计算机视觉系统,其中所述一个或多个第一张量进一步包括一个或多个非偏振表示空间中的一个或多个非偏振张量,且

其中所述一个或多个非偏振张量包括强度表示空间中的一个或多个强度图像。

16. 根据权利要求15所述的计算机视觉系统,其中所述一个或多个强度图像包括:

第一颜色强度图像;

第二颜色强度图像;以及

第三颜色强度图像。

17. 根据权利要求13所述的计算机视觉系统,其中致使所述处理器融合所述相同比例下的所述模张量的所述指令包括在由所述处理器执行时致使所述处理器进行以下操作的指令:

串接所述相同比例下的所述模张量;

将所述模张量供应到注意力子网络以计算一个或多个注意力地图;以及
基于所述一个或多个注意力地图对所述模张量进行加权以计算所述比例的融合张量。

18. 根据权利要求17所述的计算机视觉系统,其中致使所述处理器计算所述预测的所述指令进一步包括在由所述处理器执行时致使所述处理器进行以下操作的指令:将每一比例下计算的融合张量供应到被配置成计算所述分段掩码的预测模块。

19. 根据权利要求13所述的计算机视觉系统,其中所述分段掩码被供应到机器人拾取臂的控制器。

20. 根据权利要求13、14、15或16所述的计算机视觉系统,其中所述预测包括基于所述一个或多个在光学上具有挑战性的对象对所述一个或多个偏振原始帧的分类。

21. 根据权利要求13、14、15或16所述的计算机视觉系统,其中所述预测包括所述一个或多个偏振原始帧中描绘的所述一个或多个在光学上具有挑战性的对象的一个或多个检测到的特征。

22. 根据权利要求13所述的计算机视觉系统,其中用以计算所述预测的所述指令包括在由所述处理器执行时致使所述处理器将所述一个或多个第一张量供应到统计模型的指令,且

其中使用包括所述一个或多个偏振表示空间中的训练第一张量的训练数据和标签训练所述统计模型。

23. 根据权利要求22所述的计算机视觉系统,其中所述训练数据包括:

从由偏振相机捕获的数据计算的源训练第一张量;以及

经由包括旋转的仿射变换从所述源训练第一张量生成的额外训练第一张量。

24. 根据权利要求23所述的计算机视觉系统,其中,当所述额外训练第一张量包括线性偏振角AOLP图像时,生成所述额外训练第一张量包括:

使所述额外训练第一张量旋转某一角度;以及

使所述AOLP图像的像素值反向旋转所述角度。

用于使用偏振提示进行透明对象分段的系统和方法

[0001] 相关申请的交叉参考

[0002] 本申请要求2020年3月29日在美国专利商标局提交的第63/001,445号美国临时专利申请的优先权及权益,所述美国临时专利申请的全部公开内容以引用的方式并入本文中。

技术领域

[0003] 本公开的实施例的方面涉及计算机视觉的领域以及将图像分段为图像中描绘的相异对象。

背景技术

[0004] 语义分段指代捕获场景的一个或多个二维(2-D)图像以及以算法方式将图像的各个区(例如,图像的每一像素)分类为属于特定类的对象的计算机视觉过程。举例来说,将语义分段应用于花园中的人的图像可将类指派到输入图像的个别像素,其中所述类可包含现实世界对象的类型,例如:人;动物;树;大地;天空;岩石;建筑物;等等。例项分段指代进一步将唯一标签应用于对象的不同例项中的每一个,例如方式是用不同识别符单独地标记输入图像中的每个人和每只动物。

[0005] 语义分段或例项分段过程的一个可能输出是分段地图或分段掩码,其可以是具有与输入图像相同的尺寸的2-D图像,且其中每一像素的值对应于一标签(例如,在语义分段的情况下为特定类,或在例项分段的情况下为特定例项)。

[0006] 透明对象的图像的分段是计算机视觉中的困难的遗留问题。透明对象缺乏纹理(例如,表面颜色信息,例如在“纹理映射”中,正如该术语在计算机图形的领域中使用),实际上采用那些透明对象后方的场景(例如,穿过透明对象可见的场景的背景)的纹理或外观。因此,在一些情况下,所捕获场景中的透明对象(和其它在光学上具有挑战性的对象)是语义分段算法大体上不可见的,或可基于穿过那些透明对象可见的对象来分类。

发明内容

[0007] 本公开的实施例的方面涉及通过使用光偏振(光波的旋转)向语义分段或其它机器视觉过程提供额外信息通道来进行图像的透明对象分段。本公开的实施例的方面还涉及通过使用光偏振进行图像中的其它在光学上具有挑战性的对象的检测和/或分段,其中在光学上具有挑战性的对象可展现一个或多个条件,包含:非朗伯;半透明;多路径诱导;或非反射。在一些实施例中,偏振相机用于捕获偏振原始帧以生成多模态图像(例如,多维偏振信息)。本公开的实施例的一些方面涉及使用深度学习主干来处理多模态偏振输入数据的神经网络架构。相应地,本公开的实施例在各种场景和背景条件下对杂乱、透明和其它在光学上具有挑战性的对象可靠地执行例项分段,借此展现相比于单单基于强度图像的对比方法的改进。

[0008] 根据本公开的一个实施例,一种用于计算对场景的图像的预测的计算机实施的方

法包含:接收场景的一个或多个偏振原始帧,所述偏振原始帧利用偏振滤波器在不同线性偏振角下捕获;从偏振原始帧提取一个或多个偏振表示空间中的一个或多个第一张量;以及基于所述一个或多个偏振表示空间中的所述一个或多个第一张量计算关于场景中的一个或多个在光学上具有挑战性的对象的预测。

[0009] 所述一个或多个偏振表示空间中的所述一个或多个第一张量可包含:线性偏振度(DOLP)表示空间中的DOLP图像;以及线性偏振角(AOLP)表示空间中的AOLP图像。

[0010] 所述一个或多个第一张量可进一步包含一个或多个非偏振表示空间中的一个或多个非偏振张量,且所述一个或多个非偏振张量可包含强度表示空间中的一个或多个强度图像。

[0011] 所述一个或多个强度图像可包含:第一颜色强度图像;第二颜色强度图像;以及第三颜色强度图像。

[0012] 所述预测可包含分段掩码。

[0013] 所述计算所述预测可包含将所述一个或多个第一张量供应到一个或多个相应卷积神经网络(CNN)主干,且所述一个或多个CNN主干中的每一个可被配置成以多个不同比例计算多个模张量。

[0014] 所述计算所述预测可进一步包含:融合通过所述一个或多个CNN主干在相同比例下所计算的模张量。

[0015] 所述融合相同比例下的模张量可包含在相同比例下串接所述模张量;将所述模张量供应到注意力子网络以计算一个或多个注意力地图;以及基于所述一个或多个注意力地图对所述模张量进行加权以计算所述比例的融合张量。

[0016] 所述计算所述预测可进一步包含将每一比例下计算的融合张量供应到被配置成计算分段掩码的预测模块。

[0017] 所述分段掩码可被供应到机器人拾取臂的控制器。

[0018] 所述预测可包含基于所述一个或多个在光学上具有挑战性的对象对所述一个或多个偏振原始帧的分类。

[0019] 所述预测可包含所述一个或多个偏振原始帧中描绘的所述一个或多个在光学上具有挑战性的对象的一个或多个检测到的特征。

[0020] 所述计算所述预测可包含将所述一个或多个偏振表示空间中的所述一个或多个第一张量供应到统计模型,且可使用包含所述一个或多个偏振表示空间中的训练第一张量的训练数据和标签来训练所述统计模型。

[0021] 所述训练数据可包含:从由偏振相机捕获的数据计算的所述一个或多个偏振表示空间中的源训练第一张量;以及经由包含旋转的仿射变换从源训练第一张量生成的额外训练第一张量。

[0022] 当所述额外训练第一张量包含线性偏振角(AOLP)图像时,生成所述额外训练第一张量可包含:将所述额外训练第一张量旋转某一角度;以及将所述AOLP图像的像素值反向旋转所述角度。

[0023] 根据本公开的一个实施例,一种计算机视觉系统包含:偏振相机,其包含偏振滤波器;以及处理系统,其包含处理器和存储指令的存储器,所述指令在由所述处理器执行时致使所述处理器:接收场景的一个或多个偏振原始帧,所述偏振原始帧利用偏振滤波器在不

同线性偏振角下捕获;从偏振原始帧提取一个或多个偏振表示空间中的一个或多个第一张量;以及基于所述一个或多个偏振表示空间中的所述一个或多个第一张量计算关于场景中的一个或多个在光学上具有挑战性的对象的预测。

[0024] 所述一个或多个偏振表示空间中的所述一个或多个第一张量可包含:线性偏振度(DOLP)表示空间中的DOLP图像;以及线性偏振角(AOLP)表示空间中的AOLP图像。

[0025] 所述一个或多个第一张量可进一步包含一个或多个非偏振表示空间中的一个或多个非偏振张量,且其中所述一个或多个非偏振张量包含强度表示空间中的一个或多个强度图像。

[0026] 所述一个或多个强度图像可包含:第一颜色强度图像;第二颜色强度图像;以及第三颜色强度图像。

[0027] 所述预测可包含分段掩码。

[0028] 存储器可进一步存储指令,所述指令在由所述处理器执行时致使所述处理器通过将所述一个或多个第一张量供应到一个或多个相应卷积神经网络(CNN)主干来计算所述预测,其中所述一个或多个CNN主干中的每一个被配置成在多个不同比例下计算多个模张量。

[0029] 存储器可进一步存储指令,所述指令在由所述处理器执行时致使所述处理器:融合通过所述一个或多个CNN主干在相同比例下所计算的模张量。

[0030] 致使所述处理器融合相同比例下的模张量的所述指令可包含在由所述处理器执行时致使所述处理器进行以下操作的指令:在相同比例下串接所述模张量;将所述模张量供应到注意力子网络以计算一个或多个注意力地图;以及基于所述一个或多个注意力地图对所述模张量进行加权以计算所述比例的融合张量。

[0031] 致使所述处理器计算所述预测的所述指令可进一步包含在由所述处理器执行时致使所述处理器进行以下操作的指令:将每一比例下计算的融合张量供应到被配置成计算分段掩码的预测模块。

[0032] 所述分段掩码可被供应到机器人拾取臂的控制器。

[0033] 所述预测可包含基于所述一个或多个在光学上具有挑战性的对象对所述一个或多个偏振原始帧的分类。

[0034] 所述预测可包含所述一个或多个偏振原始帧中描绘的所述一个或多个在光学上具有挑战性的对象的一个或多个检测到的特征。

[0035] 用以计算所述预测的所述指令可包含在由所述处理器执行时致使所述处理器将所述一个或多个第一张量供应到统计模型的指令,且可使用包含所述一个或多个偏振表示空间中的训练第一张量的训练数据和标签来训练所述统计模型。

[0036] 所述训练数据可包含:从由偏振相机捕获的数据计算的源训练第一张量;以及经由包含旋转的仿射变换从所述源训练第一张量生成的额外训练第一张量。

[0037] 当所述额外训练第一张量包含线性偏振角(AOLP)图像时,生成所述额外训练第一张量可包含:将所述额外训练第一张量旋转某一角度;以及将所述AOLP图像的像素值反向旋转所述角度。

附图说明

[0038] 附图连同说明书一起示出本发明的示例性实施例,并连同描述内容一起用以阐释

本发明的原理。

[0039] 图1是根据本发明的一个实施例的系统的示意性框图。

[0040] 图2A是场景的图像或强度图像,其中一个真实透明球体放置在描绘含有两个透明球体(“冒充物”)和一些背景杂波的另一场景的照片的打印件的顶部上。

[0041] 图2B描绘图2A的强度图像,其具有识别透明球体的例项的对比的基于掩码区的卷积神经网络(掩码R-CNN)计算的覆盖分段掩码,其中真实透明球体正确地识别为例项,且两个冒充物不正确地识别为例项。

[0042] 图2C是根据本发明的一个实施例从场景的所捕获的偏振原始帧计算的偏振角图像。

[0043] 图2D描绘根据本发明的一实施例的图2A的强度图像,其具有使用偏振数据计算的覆盖分段掩码,其中真实透明球体正确地识别为例项,且两个冒充物正确地被排除而不作为例项。

[0044] 图3是根据本发明的一个实施例用于基于偏振数据计算分段地图的处理电路的框图。

[0045] 图4是根据本发明的一个实施例用于对输入图像执行分段以计算分段地图的方法的流程图。

[0046] 图5是光与透明对象和不透明(例如,漫射和/或反射)对象的交互的高级描绘。

[0047] 图6A、6B和6C描绘根据本发明的一个实施例由特征提取器计算的实例第一特征地图,所述特征提取器被配置成提取第一表示空间中的导出的特征地图,包含图6A中的强度表示空间中的强度特征地图I、图6B中的DOLP表示空间中的线性偏振度(DOLP)特征地图p和图6C表示空间中的线性偏振角(AOLP)特征地图 Φ 。

[0048] 图7A和7B分别是图6A、6B和6C中的标记为(a)和(b)的区的展开图。图7C是描绘图6A的强度特征地图、图6B的DOLP特征地图和图6C的AOLP特征地图中的图7B中标记的边缘的横截面的曲线图。

[0049] 图8A是根据本发明的一个实施例的特征提取器的框图。

[0050] 图8B是描绘根据本发明的一个实施例的用于从偏振原始帧提取特征的方法的流程图。

[0051] 图9是描绘根据本发明的一个实施例的如应用于基于掩码区的卷积神经网络(掩码R-CNN)主干的偏振CNN架构的框图。

[0052] 图10是根据本发明的一个实施例可与偏振CNN一起使用的注意力模块的框图。

[0053] 图11描绘由根据本发明的一个实施例的注意力模块针对从由偏振相机捕获的偏振原始帧提取的不同模张量(第一表示空间中)计算的注意力权重的实例。

[0054] 图12A、12B、12C和12D描绘由对比的图像分段系统计算的分段地图、由根据本公开的一个实施例的偏振卷积神经网络计算的分段地图,以及地面实况分段地图(例如,手动生成的分段地图)。

具体实施方式

[0055] 在以下详细描述中,借助于说明展示和描述了本发明的仅某些示例性实施例。如所属领域的技术人员将认识到,本发明可以许多不同形式体现且不应将其理解为限于本文

中所阐述的实施例。在说明书通篇，相同参考标号表示相同元件。

[0056] 包含用于制造、生命科学和汽车工业的自动化和分析的计算机视觉或机器视觉系统的许多现实世界应用中出现透明对象。举例来说，在制造领域，计算机视觉系统可用于使以下操作自动化：零件的分类、选择和放置；制造期间组件的放置的校验；以及最终检验和缺陷检测。作为额外实例，在生命科学领域，计算机视觉系统可用于使以下操作自动化：试剂的测量；样本的制备；读取仪器的输出；样本的表征；以及拾取和放置容器样本。汽车工业中的其它实例包含检测街道场景中的透明对象以用于辅助驾驶者或用于操作自动驾驶车辆。额外实例可包含辅助技术，例如能够检测玻璃门和其它透明屏障的自导航轮椅，以及用于辅助视障人士的能够检测透明饮水玻璃杯且区分真实对象和打印冒充物的装置。

[0057] 与不透明对象相比，透明对象缺乏其自身的纹理（例如，表面颜色信息，正如该术语在计算机图形的领域中（例如在“纹理映射”中）使用）。因此，对比系统通常不能正确地识别使用标准成像系统（例如，被配置成捕获单色强度图像或彩色强度图像（例如红色、绿色和蓝色或RGB图像）的相机）捕获的场景中存在的透明对象的例项。这可能是由于透明对象不具有一致的纹理（例如，表面颜色）以供算法锁存到其上或学习以进行检测（例如，机器学习算法的训练过程期间）。类似的问题可能由于部分透明或半透明对象以及一些类型的反射对象（例如，闪亮的金属）和非常深色的对象（例如，亚光黑色对象）引起。

[0058] 相应地，本公开的实施例的方面涉及使用偏振成像来提供信息以供分段算法检测场景中的透明对象。此外，本公开的实施例的方面还适用于检测其它在光学上具有挑战性的对象，例如透明、半透明和反射对象以及深色对象。

[0059] 如本文中所使用，术语“在光学上具有挑战性”指代对象由在足够的阈值水平或程度下满足以下四个特性中的一个或多个的材料制成：非朗伯（例如，非亚光）；半透明；多路径诱导；和/或非反射。在一些情况下，展现四个特性中的仅一个的对象可能检测起来在光学上具有挑战性。此外，对象或材料可同时展现多个特性。举例来说，半透明对象可具有表面反射和背景反射，因此其由于半透明性和多路径两者而具有挑战性。在一些情况下，对象可展现上文所列的四个特性中的一个或多个，但可能检测起来并不在光学上具有挑战性，因为这些条件不是在将造成对比计算机视觉系统的问题的水平或程度下展现的。举例来说，对象可以是半透明的，但仍展现足够的表面纹理以便可以检测且相对于场景中的对象的其它例项分段。作为另一实例，表面必须充分地非朗伯以引起其它视觉系统的问题。在一些实施例中，对象在光学上具有挑战性的程度或水平使用对象的双向反射分布函数(BRDF)的镜面凸瓣的半高全宽(FWHM)来量化。如果此FWHM低于阈值，则材料被视为在光学上具有挑战性的。

[0060] 图1是根据本发明的一个实施例的系统的示意性框图。在图1中展示的布置中，场景1包含放置于不透明亚光对象3（例如，棒球和网球）前方的透明对象2（例如，描绘为例如玻璃弹珠等球体、例如饮水玻璃杯或平底玻璃杯等圆柱体，以及例如透明丙烯酸面板等平面）。偏振相机10具有拥有视场的透镜12，其中透镜12和相机10定向成使得所述视场涵盖场景1。透镜12被配置成从场景1引导光（例如，聚焦光）到例如图像传感器14等光敏介质（例如，互补金属氧化物半导体(CMOS)图像传感器或电荷耦合装置(CCD)图像传感器）上。

[0061] 偏振相机10进一步包含放置于场景1和图像传感器14之间的光学路径中的偏振器或偏振滤波器或偏振遮罩16。根据本公开的各种实施例，偏振器或偏振遮罩16被配置成使

偏振相机10能够在偏振器设定在各种指定角度(例如,45°旋转或60°旋转或不均匀间隔旋转)的情况下捕获场景1的图像。

[0062] 作为一个实例,图1描绘实施例,其中偏振遮罩16是以类似于彩色相机的红绿蓝(RGB)彩色滤光片(例如,拜耳滤波器)的方式与图像传感器14的像素栅格对准的偏振嵌合体。以类似于彩色滤光片嵌合体如何基于波长对入射光进行滤波使得图像传感器14中的每一像素根据嵌合体的彩色滤光片的图案在光谱的特定部分(例如,红、绿或蓝)中接收光的方式,使用偏振嵌合体的偏振遮罩16基于线性偏振对光进行滤波使得不同像素按不同的线性偏振角(例如,0°、45°、90°和135°,或0°、60°度和120°)接收光。相应地,使用例如图1中展示的偏振遮罩16的偏振相机10能够并行地或同时地在四个不同线性偏振下捕获光。偏振相机的一个实例为由俄勒冈州威尔森维尔的 FLIR®Systems公司生产的 Blackfly®S偏振相机。

[0063] 虽然以上描述涉及使用偏振嵌合体的偏振相机的一些可能实施方案,但本公开的实施例并不限于此,而是涵盖能够在多个不同偏振下捕获图像的其它类型的偏振相机。举例来说,偏振遮罩16可具有少于或超过四个不同偏振,或可具有不同角度下的偏振(例如,0°、60°度和120°的偏振角,或者0°、30°、60°、90°、120°和150°的偏振角)。作为另一实例,偏振遮罩16可使用例如电光调制器(例如,可包含液晶层)等电子控制的偏振遮罩实施,其中所述遮罩的个别像素的偏振角可独立地控制,使得图像传感器14的不同部分接收具有不同偏振的光。作为另一实例,电光调制器可被配置成在捕获不同帧时使不同线性偏振的光透射,例如使得相机在整个偏振遮罩循序地设定为不同线性偏振器角度(例如,循序地设定为:0度;45度;90度;或135度)的情况下捕获图像。作为另一实例,偏振遮罩16可包含偏振滤波器,其以机械方式旋转,使得在偏振滤波器以机械方式相对于透镜12旋转以使相对于图像传感器14处于不同偏振角的光透射的情况下由偏振相机10捕获不同偏振原始帧。

[0064] 因此,偏振相机捕获场景1的多个输入图像18(或偏振原始帧),其中偏振原始帧18中的每一个对应于在不同偏振角 Φ_{pol} (例如,0度、45度、90度或135度)下在偏振滤波器或偏振器后方拍摄的图像。相对于场景1以大体上相同的位姿捕获偏振原始帧中的每一个(例如,在偏振滤波器处于0度、45度、90度或135度的情况下捕获的图像均由位于相同位置和定向的相同偏振相机捕获),这与相对于场景从不同的位置和定向捕获偏振原始帧形成对比。偏振相机10可被配置成在电磁光谱的多种不同部分中检测光,所述多种不同部分例如电磁光谱的人类可见部分、人类可见光谱的红、绿和蓝部分,以及电磁光谱的例如红外和紫外等不可见部分。

[0065] 在本公开的一些实施例(例如一些上述实施例)中,不同偏振原始帧由相同偏振相机10捕获,且因此可相对于场景1从大体上相同的位姿(例如,位置及定向)捕获。然而,本公开的实施例并不限于此。举例来说,偏振相机10可相对于场景1在不同偏振原始帧之间移动(例如,当在不同时间捕获对应于不同偏振角的不同原始偏振原始帧时,例如在以机械方式旋转的偏振滤波器的情况下),这是因为偏振相机10已移动或因为场景1中的对象已移动(例如,如果对象定位于移动的传送带上)。相应地,在本公开的一些实施例中,在偏振相机10相对于场景1处于不同位姿的情况下捕获不同偏振原始帧。

[0066] 偏振原始帧18被供应到处理电路100(下文更详细地描述),其基于偏振原始帧18计算分段地图20。如图1中所示,在分段地图20中,场景的透明对象2和不透明对象3全部个

别地标记,其中标签在图1中使用不同颜色或图案(例如,竖直线、水平线、棋盘格图案等)描绘,但其中在实践中,每一标签在分段地图中可以由不同值(例如,整数值,其中图中展示的不同图案对应于不同值)表示。

[0067] 根据本公开的各种实施例,处理电路100使用被配置成执行如下文更详细地描述的各种操作的一个或多个电子电路实施。电子电路的类型可包含中央处理单元(CPU)、图形处理单元(GPU)、人工智能(AI)加速器(例如,向量处理器,其可包含被配置成有效地执行神经网络常见的操作(例如,点积和柔性最大传递函数(softmax))的向量算术逻辑单元)、现场可编程门阵列(FPGA)、专用集成电路(ASIC)、数字信号处理器(DSP)等。举例来说,在一些情况下,本公开的实施例的方面实施于存储在非易失性计算机可读存储器中的程序指令中,其中,所述程序指令当由电子电路(例如,CPU、GPU、AI加速器或其组合)执行时执行本文中描述的从输入偏振原始帧18计算分段地图20的操作。由处理电路100执行的操作可由单个电子电路(例如,单个CPU、单个GPU等)执行,或可在多个电子电路(例如,多个GPU,或CPU结合GPU)之间分配。所述多个电子电路可在彼此的本地(例如,位于同一裸片上、位于同一封装内,或位于同一嵌入式装置或计算机系统内),和/或可彼此远离(例如,经由例如本地个域网(例如 Bluetooth®)等网络、经由例如本地有线和/或无线网络等局域网,和/或经由例如因特网等广域网通信,在此情况下,一些操作在本地执行且其它操作在由云计算服务代管的服务器上执行)。操作以实施处理电路100的一个或多个电子电路在本文中可以被称为计算机或计算机系统,其可包含存储指令的存储器,所述指令当由所述一个或多个电子电路执行时实施本文中所描述的系统和方法。

[0068] 图2A、2B、2C和2D提供背景来说明根据本公开的实施例通过对比方法和语义分段或例项分段计算的分段地图。更详细地,图2A是场景的图像或强度图像,其中一个真实透明球体放置在描绘含有两个透明球体(“冒充物”)和一些背景杂波的另一场景的照片的打印件的顶部上。图2B描绘如由对比的基于掩码区的卷积神经网络(掩码R-CNN)计算的分段掩码使用线的不同图案识别覆叠在图2A的强度图像上的透明球体的例项,其中真实透明球体正确地识别为例项,且两个冒充物不正确地识别为例项。换句话说,掩码R-CNN算法已被欺骗而将两个冒充的透明球体标记为场景中的实际透明球体的例项。

[0069] 图2C是根据本发明的一个实施例从场景的所捕获的偏振原始帧计算的线性偏振角(AOLP)图像。如图2C所示,透明对象具有例如AOLP域等偏振空间中的非常独特的纹理,其中边缘上存在几何结构相依签名,且所述线性偏振角中的透明对象的表面上产生相异或唯一或特定的图案。换句话说,透明对象的本征纹理(例如,相对于获自穿过透明对象可见的背景表面的非本征纹理)在图2C的偏振角图像中比在图2A的强度图像中更可见。

[0070] 图2D描绘根据本发明的一实施例具有如使用偏振数据计算的覆叠分段掩码的图2A的强度图像,其中真实透明球体使用线的覆叠图案正确地识别为例项,且两个冒充物正确地排除而不作为例项(例如,与图2B相比,图2D不包含两个冒充物上方的线的覆叠图案)。虽然图2A、2B、2C和2D示出与在存在冒充的透明对象的情况下检测真实透明对象相关的实例,但本公开的实施例并不限于此,且还可应用于其它在光学上具有挑战性的对象,例如透明、半透明和非亚光或非朗伯对象,以及非反射(例如,亚光黑色对象)和多路径诱导对象。

[0071] 相应地,本公开的实施例的一些方面涉及从偏振原始帧提取待作为输入供应到语

义分段算法或其它计算机视觉算法的表示空间中的张量(或例如偏振特征地图等第一表示空间中的第一张量)。第一表示空间中的这些第一张量可包含偏振特征地图,其编码与从场景接收的光的偏振相关的信息,例如图2C中展示的AOLP图像、线性偏振度(DOLP)特征地图等(例如,来自偏振原始帧中的个别偏振原始帧的斯托克斯向量或变换的其它组合)。在一些实施例中,这些偏振特征地图连同非偏振特征地图(例如,比如图2A中展示的图像等强度图像)一起使用以提供额外信息通道以供由语义分段算法使用。

[0072] 虽然本发明的实施例不限于与特定语义分段算法一起使用,但本发明的实施例的一些方面涉及用于透明或其它在光学上具有挑战性的对象(例如,透明、半透明、非朗伯、多路径诱导对象,和非反射(例如,非常深色)对象)的基于偏振的分段的深度学习架构,其中这些架构可被称为偏振卷积神经网络(偏振CNN)。此偏振CNN框架结构包含适于处理偏振的特定纹理的主干,且可与例如掩码R-CNN等其它计算机视觉架构耦合(例如,以形成偏振掩码R-CNN架构)以产生透明对象的准确且稳健的例项分段的解。此外,此方法可应用于具有透明和非透明(例如,不透明对象)的混合的场景,且可用于识别场景中的透明、半透明、非朗伯、多路径诱导、深色和不透明对象的例项。

[0073] 图3是根据本发明的一个实施例用于基于偏振数据计算分段地图的处理电路100的框图。图4是根据本发明的一个实施例用于对输入图像执行分段以计算分段地图的方法的流程图。如图3所示,在一些实施例中,处理电路100包含特征提取器或特征提取系统800和预测器900(例如,经典计算机视觉预测算法或经训练统计模型),所述预测器被配置成基于特征提取系统800的输出计算关于场景中的一个或多个透明对象的预测输出20(例如,统计预测)。虽然本文在训练用于检测透明对象的系统的上下文中描述本公开的一些实施例,但本公开的实施例并不限于此,且还可应用于针对其它在光学上具有挑战性的对象或由检测起来在光学上具有挑战性的材料制成的对象的技术,所述对象例如半透明对象、多路径诱导对象、并非完全或大体上亚光或朗伯的对象,和/或非常深色的对象。这些在光学上具有挑战性的对象包含通过使用由对光的偏振不敏感的相机系统捕获的图像难以解析或检测的对象(例如,基于由在光学路径中无偏振滤波器或其中不同图像不基于不同偏振角捕获图像的相机捕获的图像)。

[0074] 在图3和4中所展示的实施例中,在操作410中,处理系统100的特征提取系统800从场景的输入偏振原始帧18提取一个或多个第一表示空间中的一个或多个第一特征地图50(包含各个偏振表示空间中的偏振图像或偏振特征地图)。所提取的经导出特征地图50(包含偏振图像)作为输入提供到处理系统100的预测器900,其实施一个或多个预测模型以在操作450中计算检测到的输出20。在其中预测器为图像分段或例项分段系统的情况下,预测可以是例如图3中展示的分段地图,其中每一像素可与所述像素对应于各个可能类(或类型)的对象的一个或多个置信度相关联。在其中预测器为分类系统的情况下,预测可包含多个类和所述图像描绘所述类中的每一个的例项的相应置信度。在其中预测器900为经典计算机视觉预测算法的情况下,预测器可计算检测结果(例如,检测边缘、关键点、基础系数、哈尔小波系数,或者透明对象和/或图像中的例如半透明对象、多路径诱导对象、非朗伯对象和非反射对象等其它在光学上具有挑战性的对象的其它特征,作为输出特征)。

[0075] 在图3中所展示的实施例中,预测器900实施例项分段(或语义分段)系统,且在操作450中基于从输入偏振原始帧18提取的所提取的第一表示空间中的第一张量50计算包含

场景的分段地图的输出20。如上所述,特征提取系统800和预测器900使用被配置成执行其操作的一个或多个电子电路实施,如下文更详细地描述。

[0076] 从偏振原始帧提取第一表示空间中的例如偏振图像和经导出特征地图等第一张量

[0077] 本公开的实施例的一些方面涉及用于在操作410中提取特征的系统和方法,其中这些所提取的特征在操作450中在透明对象的稳健检测中使用。相比之下,仅依赖于强度图像的对比技术可能不能检测透明对象(例如,将图2A的强度图像与图2C的AOLP图像进行比较,上文所论述)。术语“第一表示空间”中的“第一张量”将在本文用于指代从由偏振相机捕获的偏振原始帧18计算(例如,提取)的特征,其中这些第一表示空间至少包含偏振特征空间(例如,含有关于由图像传感器检测到的光的偏振的信息的比如AOLP和DOLP等特征空间),且还可包含非偏振特征空间(例如,不需要关于到达图像传感器的光的偏振的信息的特征空间,例如仅仅基于在无任何偏振滤波器的情况下捕获的强度图像计算的图像)。

[0078] 光与透明对象之间的交互丰富且复杂,但对象的材料决定其在可见光下的透明度。对于许多透明的家用对象,大部分可见光直接通过,且小部分(~4%到~8%,取决于折射率)被反射。这是因为,光谱的可见部分中的光不具有足够的能量来激发透明对象中的原子。因此,透明对象后方(或穿过透明对象可见)的对象的纹理(例如,外观)决定了透明对象的外观。举例来说,当注视桌子上的透明玻璃杯或平底玻璃杯时,平底玻璃杯的另一侧(例如,桌子的表面)上对象的外观通常决定透过杯子看到什么。此性质导致当尝试仅仅基于强度图像进行例项分段时的一些困难:

[0079] 杂波:具有透明对象的密集杂乱场景中难以看到清晰边缘(例如,透明对象的边缘)。在极端状况下,边缘完全不可见(参见例如图6A的区(b),下文更详细地描述),从而导致透明对象的确切形状的模糊。

[0080] 新颖环境:可见光谱中的低反射率致使透明对象在新颖环境(例如,不同于用于训练分段系统的训练数据的环境,例如其中穿过透明对象可见的背景不同于训练数据中的背景)中呈现不同的分布外(out-of-distribution),借此导致较差的一般化。

[0081] 打印冒充物:使用单个RGB图像作为输入的算法通常归因于透视模糊而对打印冒充物(例如,摄影图像的打印件)敏感。虽然存在用于透明对象的语义分段的其它非单眼算法(例如,使用从场景周围的多个不同位姿捕获的图像,例如立体相机),但其范围受限且可能不能够应对例项分段。

[0082] 图5是光与透明对象和不透明(例如,漫射和/或反射)对象的交互的高级描绘。如图5所示,偏振相机10捕获包含不透明背景对象503前方的透明对象502的场景的偏振原始帧。击中偏振相机10的图像传感器14的光线510含有来自透明对象502和背景对象503两者的偏振信息。来自透明对象502的反射光512的小部分很大程度上偏振,且因此与从背景对象503反射出且通过透明对象502的光513相比,对偏振测量值具有较大影响。

[0083] 击中偏振相机10的图像传感器16的光线510具有三个可测量分量:光的强度(强度图像/I)、线性偏振的光的百分比或比例(线性偏振度/DOLP/ ρ),以及所述线性偏振的方向(线性偏振角/AOLP/ Φ)。这些性质编码关于正成像的对象的表面曲率和材料的信息,所述信息可由预测器900使用以检测透明对象,如下文更详细地描述。在一些实施例中,预测器900可基于通过半透明对象的光和/或与多路径诱导对象或非反射对象(例如,亚光黑色对

象)交互的光的类似的偏振性质检测其它在光学上具有挑战性的对象。

[0084] 因此,本发明的实施例的一些方面涉及使用特征提取器800计算一个或多个第一表示空间中的第一张量,其可包含基于强度I、DOLP ρ 和AOLP Φ 导出的特征地图。特征提取器800可通常将信息提取到第一表示空间(或第一特征空间)中,所述第一表示空间包含偏振表示空间(或偏振特征空间),例如“偏振图像”,换句话说,基于原本不可从强度图像(例如,由未包含偏振滤波器或其它用于检测到达其图像传感器的光的偏振的机构的相机捕获的图像)计算的偏振原始帧提取的图像,其中这些偏振图像可包含DOLP ρ 图像(DOLP表示空间或特征空间中)、AOLP Φ 图像(AOLP表示空间或特征空间中)、从斯托克斯向量计算的偏振原始帧的其它组合,以及从偏振原始帧计算的信息的其它图像(或更一般地说,第一张量或第一特征张量)。第一表示空间可包含非偏振表示空间,例如强度I表示空间。

[0085] 在每一像素处测量强度I、DOLP ρ 和AOLP Φ 需要以不同角度 Φ_{pol} 在偏振滤波器(或偏振器)后方拍摄的场景的3个或更多个偏振原始帧(例如,因为存在待确定的三个未知值:强度I、DOLP ρ 和AOLP Φ 。举例来说,上文描述的FLIR® Blackfly® S偏振相机以偏振角(0度、45度、90度或135度下的 Φ_{pol})捕获偏振原始帧,借此产生本文中表示为 I_0 、 I_{45} 、 I_{90} 和 I_{135} 的四个偏振原始帧 $I_{\Phi_{pol}}$ 。

[0086] 每一像素处的 $I_{\Phi_{pol}}$ 与强度I、DOLP ρ 和AOLP Φ 之间的关系可表达为:

$$[0087] \quad I_{\Phi_{pol}} = I \left(1 + \rho \cos \left(2(\phi - \phi_{pol}) \right) \right) \quad (1)$$

[0088] 相应地,在四个不同偏振原始帧 $I_{\Phi_{pol}}$ (I_0 、 I_{45} 、 I_{90} 和 I_{135})的情况下,一组四个等式可用于对强度I、DOLP ρ 和AOLP Φ 求解。

[0089] 偏振形状(Shape from Polarization, SfP)理论(参见例如Gary A Atkinson和Edwin R Hancock。表面定向从漫射偏振的恢复。IEEE图像处理会刊,15(6):1653-1664,2006。)陈述对象的表面法线的折射率(n)、方位角(θ_a)和天顶角(θ_z)与来自所述对象的光线的 Φ 和 ρ 分量之间的关系。

[0090] 当漫反射占主导时:

$$[0091] \quad \rho = \frac{\left(n - \frac{1}{n} \right)^2 \sin^2(\theta_z)}{2 + 2n^2 - \left(n + \frac{1}{n} \right)^2 \sin^2 \theta_z + 4 \cos \theta_z \sqrt{n^2 - \sin^2 \theta_z}} \quad (2)$$

[0092] $\phi = \theta_a$ (3)且当镜面反射占主导时:

$$[0093] \quad \rho = \frac{2 \sin^2 \theta_z \cos \theta_z \sqrt{n^2 - \sin^2 \theta_z}}{n^2 - \sin^2 \theta_z - n^2 \sin^2 \theta_z + 2 \sin^4 \theta_z} \quad (4)$$

$$[0094] \quad \phi = \theta_a - \frac{\pi}{2} \quad (5)$$

[0095] 应注意,在两种情况下, ρ 随 θ_z 增加而按指数规律增加,且如果折射率相同,则镜面反射比漫反射偏振得更多。

[0096] 本公开的实施例的一些方面涉及将从偏振原始帧提取的第一表示空间中的第一张量(例如,经导出特征地图)作为输入供应到预测器以用于计算对于场景的透明对象和/

或其它在光学上具有挑战性的对象(例如,半透明对象、非朗伯对象、多路径诱导对象和/或非反射对象)的计算机视觉预测,例如语义分段系统用于计算分段地图,包含检测场景中的透明对象和其它在光学上具有挑战性的对象的例项。这些第一张量可包含经导出特征地图,其可包含强度特征地图 I 、线性偏振度(DOLP) ρ 特征地图和线性偏振角(AOLP) Φ 特征地图,且其中DOLP ρ 特征地图和AOLP Φ 特征地图是参考编码关于由偏振相机检测到的光的偏振的信息的特征地图的偏振表示空间中的偏振特征地图或张量的实例。相对于图6A、6B、6C、7A、7B和7C更详细地示出偏振特征地图(或偏振图像)的益处。

[0097] 图6A、6B和6C描绘根据本发明的一个实施例作为由特征提取器计算的特征地图的实例第一张量,所述特征提取器被配置成提取第一表示空间中的第一张量,包含图6A中的强度表示空间中的强度特征地图 I 、图6B中的DOLP表示空间中的线性偏振度(DOLP)特征地图 ρ 和图6C中的AOLP表示空间中的线性偏振角(AOLP)特征地图 Φ 。在下文更详细地论述两个关注区-区(a)含有两个透明球体,且区(b)含有饮水玻璃杯的边缘。

[0098] 图7A和7B分别是图6A、6B和6C中的标记为(a)和(b)的区的展开图。图7C是描绘图6A的强度特征地图 I 、图6B的DOLP特征地图 ρ 和图6C的AOLP特征地图 Φ 中的图7B中标记的边缘的横截面的曲线图。

[0099] 参看区(a),如图6A中和图7A的左侧所见,两个透明球体的纹理归因于背景(例如,具有孔的栅格的塑料箱与透明球体静置于的花纹布料)的改变、新颖环境(例如,透过透明对象可见的各种背景)所导致的高光显示问题而在强度图像中不一致。此不一致可能使语义分段或例项分段系统难以辨识图像的这些非常不同外观部分对应于相同类型或类的对象(例如,透明球体)。

[0100] 另一方面,在图6B中和图7A的右侧展示的DOLP图像中,透明对象的形状显而易见,且背景纹理(例如,布料的花纹)在DOLP图像 ρ 中不出现。图7A是图6A中展示的强度图像 I 和图6B中展示的DOLP图像 ρ 的区(a)的放大视图,其展示透明球体的两个不同部分具有强度图像 I 中不一致的(例如,不同外观的)纹理但具有DOLP图像 ρ 中一致的(例如,类似外观的)纹理,借此使语义分段或例项分段系统较可能基于DOLP图像 ρ 辨识这两个类似外观的纹理两者对应于相同类的对象。

[0101] 参看区(b),如图6A和图7B的左侧所见,饮水玻璃杯的边缘实际上在强度图像 I 中不可见(例如,不可与花纹布料区分),但在AOLP图像 Φ 中明亮得多,如图6C和图7B的右侧所见。图7C是图7B中的强度图像 I 和AOLP图像 Φ 中的区识别的框中的边缘的横截面,其展示所述边缘在AOLP Φ 和DOLP ρ 中比在强度图像 I 中具有高得多的对比度,借此使语义分段或例项分段系统较可能基于AOLP Φ 和DOLP ρ 图像检测透明图像的边缘。

[0102] 更正式地讲,本公开的实施例的方面涉及计算第一表示空间中的第一张量50,包含提取偏振表示空间中的第一张量,例如在操作410中基于由偏振相机10捕获的偏振原始帧形成偏振图像(或提取经导出的偏振特征地图)。

[0103] 来自透明对象的光线具有两个分量:包含反射强度 I_r 、反射DOLP ρ_r 和反射AOLP Φ_r 的反射部分,以及包含折射强度 I_t 、折射DOLP ρ_t 和折射AOLP Φ_t 的折射部分。所得图像中的单个像素的强度可写成:

$$[0104] \quad I = I_r + I_t \quad (6)$$

[0105] 当具有线性偏振角 Φ_{pol} 的偏振滤波器放置于相机前方时,给定像素处的值为:

$$[0106] \quad I_{\phi_{pol}} = I_r \left(1 + \rho_r \cos \left(2(\phi_r - \phi_{pol}) \right) \right) + I_t \left(1 + \rho_t \cos \left(2(\phi_t - \phi_{pol}) \right) \right) \quad (7)$$

[0107] 依据 I_r 、 ρ_r 、 Φ_r 、 I_t 、 ρ_t 和 Φ_t 针对DOLP ρ 图像中的像素和AOLP Φ 图像中的像素的值对上述表达式求解：

$$[0108] \quad \rho = \frac{\sqrt{(I_r \rho_r)^2 + (I_t \rho_t)^2 + 2I_t \rho_t I_r \rho_r \cos(2(\phi_r - \phi_t))}}{I_r + I_t} \quad (8)$$

$$[0109] \quad \phi = \arctan \left(\frac{I_r \rho_r \sin(2(\phi_r - \phi_t))}{I_t \rho_t + I_r \rho_r \cos(2(\phi_r - \phi_t))} \right) + \phi_r \quad (9)$$

[0110] 相应地，上述等式(7)、(8)和(9)提供根据本公开的一个实施例用于形成包含强度图像I、DOLP图像 ρ 和AOLP图像 Φ 的第一表示空间中的第一张量50的模型，其中偏振表示空间中的偏振图像或张量(包含DOLP图像 ρ 和AOLP图像 Φ ，基于等式(8)和(9))的使用使得能够可靠地检测透明对象和通常不可由例如掩码R-CNN系统等仅使用强度I图像作为输入的对比系统检测的其它在光学上具有挑战性的对象。

[0111] 更详细地，例如偏振图像DOLP ρ 和AOLP Φ 等偏振表示空间中的第一张量(经导出特征地图50当中)可揭露可能以其它方式在强度I域中无纹理地现呈的对象的表面纹理。透明对象可具有强度域I中不可见的纹理，因为此强度严格地取决于 I_r/I_t 的比率(见等式(6))。不同于不透明对象(其中 $I_t=0$)，透明对象透射大部分入射光，且仅反射此入射光的小部分。

[0112] 另一方面，在偏振域或领域中，透明对象的表面纹理的强度取决于 $\Phi_r - \Phi_t$ 和 $I_r \rho_r / I_t \rho_t$ 的比率(见等式(8)和(9))。假定对于大部分像素来说 $\Phi_r \neq \Phi_t$ 且 $\theta_{zr} \neq \theta_{zt}$ (例如，假定背景和透明对象的几何形状不同)且基于 ρ_r 遵循镜面反射曲线的启示(参见例如Daisuke Miyazaki、Masataka Kagesawa和Katsushi Ikeuchi。从一对偏振图像进行透明表面建模。IEEE图案分析和机器智能会刊(1):73-82,2004。)，意味着其高度地偏振，且在布儒斯特角(近似 60°)处 ρ_r 为1.0(见等式(4))，因而在适当天顶角处， $\rho_r \geq \rho_t$ ，且如果背景为漫射或具有低天顶角，则 $\rho_r \gg \rho_t$ 。此效应可在图2C中看到，其中真实透明球体的纹理在 $\theta_z \approx 60^\circ$ 时占主导。相应地，在许多情况下，以下假设成立：

$$[0113] \quad \frac{I_r}{I_t} \leq \frac{I_r \rho_r}{I_t \rho_t} \quad (10)$$

[0114] 因此，即使透明对象的纹理在强度域I中似乎不可见，透明对象的纹理可在偏振域中，例如在AOLP Φ 中和DOLP ρ 中较可见。

[0115] 返回到导致在对仅强度图像尝试语义分段或例项分段时的困难的情形的三个实例：

[0116] 杂波：杂波中的一个问题在于检测可大体上无纹理的透明对象的边缘(参见例如图6A的区(b)中的饮水玻璃杯的边缘。另一方面，玻璃杯和其边缘的纹理似乎在图6B中展示的DOLP ρ 中较可见，且在图6C中展示的AOLP Φ 中更加可见。

[0117] 新颖环境：除增加透明对象纹理的强度外，例如图6B中展示的DOLP ρ 图像还减少比如纹理化或花纹布料等漫射背景的影响(例如，致使背景布料几乎完全黑色)。这允许透明对象在不同场景中类似地呈现，即使当环境在场景间改变时也如此。参见例如图6B和图7A

中的区(a)。

[0118] 打印冒充物:纸张为平面的,从而产生几乎均匀的AOLP Φ 和DOLP ρ 。透明对象具有一定的表面变化,其将在AOLP Φ 和DOLP ρ 中呈现为非常不均匀(参见例如图2C)。如此,透明对象的打印冒充物可不同于真实透明对象。

[0119] 图8A是根据本发明的一个实施例的特征提取器800的框图。图8B是描绘根据本发明的一个实施例的用于从偏振原始帧提取特征的方法的流程图。在图8A中所展示的实施例中,特征提取器800包含被配置成提取强度表示空间(例如,根据等式(7),作为非偏振表示空间的一个实例)中的强度图像I₅₂的强度提取器820,以及被配置成提取一个或多个偏振表示空间中的特征的偏振特征提取器830。如图8B所示,操作410中的偏振图像的提取可包含在操作411中从偏振原始帧从第一斯托克斯向量提取第一偏振表示空间中的第一张量。在操作412中,特征提取器800进一步从偏振原始帧提取第二偏振表示空间中的第二张量。举例来说,偏振特征提取器830可包含被配置成提取DOLP ρ 图像54(例如,根据等式(8)的第一偏振图像或第一张量,其中DOLP作为第一偏振表示空间)的DOLP提取器840,以及被配置成从所供应的偏振原始帧18提取AOLP Φ 图像56(例如,根据等式(9)的第二偏振图像或第二张量,其中AOLP作为第二偏振表示空间)的AOLP提取器860。作为另一实例,偏振表示空间可包含根据斯托克斯向量的偏振原始帧的组合。作为其它实例,偏振表示可包含根据一个或多个图像处理滤波器(例如,用以增加图像对比度的滤波器或去噪滤波器)修改或变换偏振原始帧。导出的特征地图52、54和56可接着被供应到预测器900以供进一步处理,例如执行推断(例如,生成例项分段地图、对图像进行分类,以及生成图像的文本描述)。

[0120] 虽然图8B示出其中从两个不同表示空间中的偏振原始帧18提取两个不同张量的情况,但本公开的实施例并不限于此。举例来说,在本公开的一些实施例中,从偏振原始帧18提取偏振表示空间中的恰好一个张量。举例来说,原始帧的一个偏振表示空间为AOLP,且另一偏振表示空间为DOLP(例如,在一些应用中,AOLP可足够用于检测透明对象或例如半透明、非朗伯、多路径诱导和/或非反射对象等其它在光学上具有挑战性的对象)。在本公开的一些实施例中,基于相应斯托克斯向量从偏振原始帧18提取两个以上不同张量。举例来说,如图8B所示,可由特征提取器800提取n个不同表示空间中的n个不同张量,其中在操作414中提取第n张量。

[0121] 相应地,从偏振原始帧18提取例如偏振特征地图或偏振图像等特征产生第一张量50,透明对象或例如半透明对象、多路径诱导对象、非朗伯对象和非反射对象等其它在光学上具有挑战性的对象较容易从所述第一张量检测到或与场景中的其它对象分离。在一些实施例中,由特征提取器800提取的第一张量可以是显式地导出的特征(例如,由人类设计者手工制作),其涉及可在偏振原始帧中展现的基本物理现象(例如,AOLP和DOLP图像的计算,如上文所论述)。在本公开的一些额外实施例中,特征提取器800提取其它非偏振特征地图或非偏振图像,例如不同颜色的光(例如,红、绿和蓝光)的强度地图以及强度地图的变换(例如,将图像处理滤波器应用于强度地图)。在本公开的一些实施例中,特征提取器800可被配置成提取经由基于经标记训练数据的端到端监督训练过程自动习得的一个或多个特征(例如,非人类手动指定的特征)。

[0122] 基于从偏振原始帧计算的偏振特征计算例如分段地图等预测

[0123] 如上所述,本公开的实施例的一些方面涉及将例如偏振图像或偏振特征地图(例

如,由特征提取器800提取的DOLP ρ 和AOLP Φ 图像)等偏振表示空间中的第一张量提供到例如语义分段算法等预测器以执行偏振图像的多模态融合从而生成习得特征(或第二张量),且基于习得的特征或第二张量计算例如分段地图等预测。将在下文更详细地描述与语义分段或例项分段相关的特定实施例。

[0124] 通常,存在许多语义分段方法,包含深度例项技术。各种深度例项技术可被分类为基于语义分段的技术(例如以下文献中所描述:Min Bai和Raquel Urtasun。用于例项分段的深度分水岭变换。IEEE计算机视觉和图案识别会议会刊,第5221-5229页,2017; Alexander Kirillov、Evgeny Levinkov、Bjoern Andres、Bogdan Savchynskyy和Carsten Rother。例项切割:利用多重切割从边缘到例项。IEEE计算机视觉和图案识别会议会刊,第5008-5017页,2017;以及Anurag Arnab和Philip HS Torr。具有动态例项化网络的逐像素例项分段。IEEE计算机视觉和图案识别会议会刊,第441-450页,2017。),基于方案的技术(例如以下文献中所描述:Kaiming He、Georgia Gkioxari、Piotr Doll'ar和Ross Girshick。掩码r-cnn。IEEE国际计算机视觉会议会刊,第2961-2969页,2017。))以及基于递归神经网络(RNN)的技术(例如以下文献中所描述:Bernardino Romera-Paredes和Philip Hilaire Sean Torr。递归例项分段。欧洲计算机视觉会议,第312-329页。施普林格,2016,以及Mengye Ren和Richard S Zemel。具有递归注意力的端对端例项分段。IEEE计算机视觉和图案识别会议会刊,第6656-6664页,2017。)。本公开的实施例可应用于这些语义分段技术中的任一个。

[0125] 虽然一些对比方法将串接的偏振原始帧(例如,如上文所描述的图像 I_0 、 I_{45} 、 I_{90} 和 I_{135})直接供应到深度网络中而不自其提取例如偏振图像或偏振特征地图等第一张量,但作为输入的这些偏振原始帧上直接训练的模型通常奋力学习物理先验,这导致较差性能,例如不能检测透明对象或其它在光学上具有挑战性的对象的例项。相应地,本公开的实施例的方面涉及使用偏振图像或偏振特征地图(在一些实施例中,与例如强度特征地图等其它特征地图组合)来执行场景中的透明对象的图像上的例项分段。

[0126] 使用深度例项分段的本公开的一个实施例是基于基于掩码区的卷积神经网络(掩码R-CNN)架构的修改以形成偏振掩码R-CNN架构。掩码R-CNN通过获取输入图像 x 并经由主干网络运行所述输入图像而工作,所述输入图像为图像强度值的 $H \times W \times 3$ 张量(例如,红、绿和蓝通道中,高度乘宽度乘颜色强度): $C=B(x)$ 。主干网络 $B(x)$ 负责从输入图像提取有用的习得特征,并且可为任何标准CNN架构,例如AlexNet(参见例如Krizhevsky、Alex、Ilya Sutskever和Geoffrey E.Hinton。“利用深度卷积神经网络的ImageNet分类。”神经信息处理系统的进展。2012。),VGG(参见例如Simonyan、Karen和Andrew Zisserman。“用于大规模图像辨识的极深卷积网络。”arXiv预印本arXiv:1409.1556(2014。)),ResNet-101(参见例如Kaiming He、Xiangyu Zhang、Shaoqing Ren和Jian Sun。用于图像辨识的深度残差学习。IEEE计算机视觉和图案识别会议会刊,第770-778页,2016。),MobileNet(参见例如Howard、Andrew G.等。“Mobilenet:用于移动视觉应用的有效卷积神经网络。”arXiv预印本arXiv:1704.04861(2017。)),MobileNetV2(参见例如Sandler、Mark等。“MobileNetV2:反相残差和线性瓶颈。”IEEE计算机视觉和图案识别会议会刊,2018。),以及MobileNetV3(参见例如Howard、Andrew等。“针对MobileNetV3的搜索。”IEEE国际计算机视觉会议会刊,2019。))

[0127] 主干网络 $B(x)$ 输出张量的集合,例如 $C=\{C_1, C_2, C_3, C_4, C_5\}$,其中每一张量 C_i 表示不

同分辨率特征地图。这些特征地图接着在特征金字塔网络(FPN)中组合(参见例如Tsung-Yi Lin、Piotr Doll'ar、Ross Girshick、Kaiming He、Bharath Hariharan和Serge Belongie。用于对象检测的特征金字塔网络。IEEE计算机视觉和图案识别会议会刊,第2117-2125页,2017。),利用区方案网络(RPN)处理(参见例如Shaoqing Ren、Kaiming He、Ross Girshick和Jian Sun。较快r-cnn:针对利用区方案网络的实时对象检测。神经信息处理系统的进展,第91-99页,2015。),且最后通过输出子网络(参见例如Ren等和He等,如上所述)以产生类、限界框和逐像素分段。这些利用非最大值删除合并以实现例项分段。

[0128] 本发明的实施例的方面涉及一种用于使用深度学习来利用偏振图像中所含的额外信息的框架结构,其中此额外信息在由对比相机捕获的输入图像中不存在(例如,并非由不使用偏振器或偏振滤波器的标准彩色或单色相机捕获的信息)。根据本公开的实施例的框架结构构建的神经网络架构将在本文中被称作偏振卷积神经网络(CNN)。

[0129] 应用根据本公开的一些实施例的此框架结构涉及对CNN架构的三个改变:

[0130] (1) 输入图像:应用偏振的物理等式以例如通过使用根据本公开的一些实施例的特征提取器800创建到CNN的输入偏振图像。

[0131] (2) 注意力-融合极主干:通过借助于经训练CNN主干融合从偏振图像计算的习得特征而将问题视为多模态融合问题。

[0132] (3) 几何数据扩增:扩增训练数据以表示偏振的物理。

[0133] 然而,本公开的实施例并不限于此。实际上,可对现有CNN架构作出上述三个改变和/或除上述三个改变以外的改变的任何子集以创建本公开的实施例内的偏振CNN架构。

[0134] 根据本公开的一些实施例的偏振CNN可使用被配置成执行下文更详细地描述的操作的一个或多个电子电路实施。在图3中所展示的实施例中,偏振CNN用作用于计算分段地图20的预测器900的组件,如图3所示。

[0135] 图9是描绘根据本发明的一个实施例应用于基于掩码区的卷积神经网络(掩码R-CNN)主干的偏振CNN架构的框图,其中第二张量C(或输出张量,例如习得的特征地图)用于计算例如分段掩码20等输出预测。

[0136] 虽然本公开的一些实施例涉及使用应用于掩码R-CNN主干的偏振CNN架构的语义分段或例项分段,但本公开的实施例并不限于此,且可以类似方式修改例如AlexNet、VGG、MobileNet、MobileNetV2、MobileNetV3等其它主干。

[0137] 在图9中所展示的实施例中,导出的特征地图50(例如,包含输入偏振图像,例如AOLP Φ 和DOLP ρ 图像)作为输入供应到偏振CNN主干910。在图9中所展示的实施例中,输入特征地图50包含三个输入图像:来自等式(1)的强度图像(I)52、AOLP(Φ)56、DOLP(ρ)54作为输入用于检测透明对象和/或其它光学上具有挑战性的对象。这些图像从偏振原始帧18(例如,如上文所描述的图像 I_0 、 I_{45} 、 I_{90} 和 I_{135})计算,归一化到某一范围中(例如,范围[0-255]内的8位值),且变换成三通道灰阶图像以允许基于MSCoCo数据集上预训练的网络容易地进行迁移学习(参见例如Tsung-Yi Lin、Michael Maire、Serge Belongie、James Hays、Pietro Perona、Deva Ramanan、Piotr Doll'ar和C Lawrence Zitnick。Microsoft coco:上下文中的共同对象。欧洲计算机视觉会议,第740-755页。斯普林格出版社,2014。)

[0138] 在图9中所展示的实施例中,输入的经导出特征地图50中的每一个被供应到单独的主干:强度 B_I (I)912、AOLP主干 B_Φ (Φ)914和DOLP主干 B_ρ (ρ)916。CNN主干912、914和916在

不同比例或分辨率*i*下计算每一模式的张量或“模张量”(例如,基于CNN主干的训练或迁移学习期间习得的参数计算的特征地图,在下文更详细地论述) $C_{i,I}$ 、 $C_{i,\rho}$ 、 $C_{i,\Phi}$ 。虽然图9示出具有五个不同比例*i*的实施例,但本公开的实施例并不限于此,且还可应用于具有不同数目的比例的CNN主干。

[0139] 本公开的实施例的一些方面涉及空间感知注意力融合机制以执行多模态融合(例如,从例如强度特征地图I、AOLP特征地图 Φ 和DOLP特征地图 ρ 等不同模式或不同类型的输入特征地图中的每一个计算的特征地图的融合)。

[0140] 举例来说,在图9中所展示的实施例中,使用相应比例的融合层922、923、924、925(统称为融合层920)融合每一比例*i*下从相应主干 B_I 、 B_ρ 、 B_Φ 计算的模张量 $C_{i,I}$ 、 $C_{i,\rho}$ 、 $C_{i,\Phi}$ (每一模式的张量)。举例来说,融合层922被配置成融合比例*i*=2下计算的模张量 $C_{2,I}$ 、 $C_{2,\rho}$ 、 $C_{2,\Phi}$ 以计算融合张量 C_2 。同样,融合层923被配置成融合比例*i*=3下计算的模张量 $C_{3,I}$ 、 $C_{3,\rho}$ 、 $C_{3,\Phi}$ 以计算融合张量 C_3 ,且可由融合层924和925执行类似的计算以基于其比例的相应模式张量分别计算融合特征地图 C_4 和 C_5 。融合张量 C_i (例如, C_2 、 C_3 、 C_4 、 C_5)或由融合层920计算的例如融合特征地图等第二张量接着作为输入供应到预测模块950,所述预测模块被配置成依据融合张量计算预测,其中所述预测可以是例如分段地图20、分类、文本描述等输出。

[0141] 图10是根据本发明的一个实施例可与偏振CNN一起使用的融合层920当中的第*i*融合层的框图。如图10所示,在本公开的一些实施例中,使用注意力模块实施融合层(例如,融合层920中的每一个),其中预测器900串接由CNN主干针对第*i*比例计算的所供应的输入张量或输入特征地图 $C_{i,I}$ 、 $C_{i,\rho}$ 、 $C_{i,\Phi}$ 且生成串接张量1010,其中经由一组经训练卷积层或注意力子网络 Ω_i 针对所述第*i*比例处理串接张量1010。注意力子网络 Ω_i 输出具有相同高度和宽度的3通道图像作为输入张量,且在一些实施例中,柔性最大传递函数应用于3通道图像的每一像素以针对第*i*比例计算逐像素注意力权重 a :

$$[0142] \quad [a_{i,\Phi}, a_{i,\rho}, a_{i,I}] = \text{softmax}(\Omega_i([C_{i,\Phi}, C_{i,\rho}, C_{i,I}])) \quad (11)$$

[0143] 这些注意力权重用于执行每通道的加权平均1020:

$$[0144] \quad C_i = a_{i,\Phi} C_{i,\Phi} + a_{i,\rho} C_{i,\rho} + a_{i,I} C_{i,I} \quad (12)$$

[0145] 相应地,使用注意力模块允许根据本公开的一些实施例的偏振CNN基于其与场景的给定部分的相关程度来对比例*i*下的不同输入(例如,比例*i*下的强度I张量或习得特征地图 $C_{i,I}$ 、DOLP张量或习得特征地图 $C_{i,\rho}$ 和AOLP张量或习得特征地图 $C_{i,\Phi}$)进行加权,其中由经训练注意力子网络 Ω_i 根据用于训练偏振CNN主干的经标记训练数据确定相关性。

[0146] 图11描绘根据本发明的一个实施例由注意力模块针对从由偏振相机捕获的偏振原始帧提取的不同模张量(不同第一表示空间中)计算的注意力权重的实例。如图11所示(参见例如强度图像1152),由偏振相机成像的场景包含放置在打印的照片的顶部上的透明玻璃杯,其中打印的照片描绘透明饮水玻璃杯(饮水玻璃杯的打印冒充物)和一些背景杂波。

[0147] 如图11中所见,习得注意力权重1110在真实饮水玻璃杯周围的区中在DOLP 1114和AOLP 1116上最亮,且避免强度图像1152中模糊的打印冒充物。相应地,预测模块950可计算例如与地面实况1130紧密匹配的分段掩码1120(例如,预测1120展示与场景中的透明对象的形状紧密匹配的形状)。

[0148] 在图9中所展示的实施例中,预测模块950大体上类似于掩码R-CNN架构中使用的

预测模块,且通过使用特征金字塔网络(FPN)和区提议网络(RPN)组合融合特征地图C作为到输出子网络的输入以用于计算输入图像中检测到的对象的每一例项的类、掩码和限界框(Bbox)来计算分段地图,所计算的类、掩码和限界框接着利用非最大值删除合并以计算例项分段地图(或例项分段掩码)20。

[0149] 如上所述,可使用迁移学习基于使用例如MSCoCo数据集训练的现有深度神经网络和例如反向传播和梯度下降等神经网络训练算法来训练偏振CNN架构。更详细地,基于表示到由特征提取器800从偏振原始帧18提取的偏振CNN的输入的额外训练数据进一步训练偏振CNN架构(例如,使用训练偏振原始帧来计算训练导出特征地图50和与训练导出特征地图相关联的地面实况标签)。这些额外训练数据可包含(例如)由偏振相机捕获的含有多种不同环境中的透明对象或其它在光学上具有挑战性的对象的多种场景的偏振原始帧,以及用由偏振相机捕获的图像中描绘的对象的例项和类标记像素的地面实况分段地图(例如,手动生成的分段地图)。

[0150] 在小训练数据集的情况下,仿射变换提供用于扩增训练数据(例如,从现有训练数据生成额外训练数据)以实现良好一般化性能的技术。然而,天真地将仿射变换应用于例如AOLP Φ 图像等一些源训练导出特征地图不能提供对经训练神经网络的性能的显著改进,且在一些情况下会损害性能。这是因为,AOLP是表示电磁波相对于相机坐标系的方向的 0° 到 360° (或 0 到 2π)范围内的角度。如果旋转运算符应用于源训练图像(或源训练导出特征地图),则此等效于将相机围绕其Z轴旋转(例如,沿着透镜12的光轴)。此旋转继而将改变相机的X-Y平面的定向,且因此将改变电磁波的相对方向(例如,线性偏振角)。为了考虑此改变,当通过以某一旋转角执行旋转仿射变换扩增数据时,AOLP的像素值在相反方向中旋转(或反向旋转,或将反向旋转应用于所生成的额外数据)相同角度。此相同原理还应用于训练特征地图或训练第一张量的其它仿射变换,其中应用于训练特征地图或训练第一张量的特定变换可根据训练特征地图表示的内容的基本物理而不同。举例来说,虽然DOLP图像可不受旋转变换影响,但归因于光与透明对象或其它在光学上具有挑战性的对象(例如,半透明对象、非朗伯对象、多路径诱导对象和非反射对象)相互作用的基本物理行为,平移变换将需要相应地改变DOLP。

[0151] 此外,虽然本公开的一些实施例涉及CNN和深度语义分段的使用,但本公开的实施例不限于此。在本公开的一些实施例中,导出的特征地图50作为输入供应(在一些实施例中,与其它特征地图一起)到其它类型的分类算法(例如,在不定位检测到的对象的情况下对图像进行分类)、其它类型的语义分段算法,或经训练以生成场景的自然语言描述的图像描述算法。这些算法的实例包含支持向量机(SVM)、马尔可夫随机场、概率性图形模型等。在本公开的一些实施例中,导出的特征地图作为输入供应到例如特征检测器等经典机器视觉算法(例如,比例恒定特征变换(SIFT)、加速稳健特征(SURF)、梯度位置和定向直方图(GLOH)、定向梯度直方图(HOG)、基础系数、哈尔小波系数等),以输出场景中的检测到的透明对象和/或其它在光学上具有挑战性的对象的检测到的经典计算机视觉特征。

[0152] 图12A、12B、12C和12D描绘由对比的图像分段系统计算的分段地图、由根据本公开的一个实施例的偏振卷积神经网络计算的分段地图,以及地面实况分段地图(例如,手动生成的分段地图)。图12A、12B、12C和12D描绘四个不同测试集上运行的实验的实例以将根据本公开的一个实施例的经训练偏振掩码R-CNN模型的性能与对比的掩码R-CNN模型(在本文

中被称作“强度”掩码R-CNN模型以指示其对强度图像而非对非偏振图像操作)进行比较。

[0153] 用于执行实验的偏振掩码R-CNN模型在含有1,000个图像的训练集上训练,具有来自六个可能类的透明对象:塑料杯、塑料托盘、玻璃杯、摆设和其它的十五个不同环境中的透明对象的超过20,000个例项。例如上文相对于输入图像的仿射变换和基于图像的旋转对AOLP的调整描述的数据扩增技术在训练之前应用于训练集。

[0154] 四个测试集包含:

[0155] (a) 杂波测试集含有类似于不具有打印件的训练集的环境中的杂乱透明对象的200个图像。

[0156] (b) 新颖环境(Env)测试集含有具有训练集中不可用的环境的每图像~6个对象的所获取的50个图像。背景含有刺眼的灯光、花纹布料、闪亮的金属等等。

[0157] (c) 打印冒充物(POS)测试集含有50个图像,每一图像含有1到6个打印对象和1或2个真实对象。

[0158] (d) 机器人箱柜拾取(RBP)测试集含有从我们的机器人臂拾取摆设(例如,装饰性玻璃摆设,适于悬挂在树上)的现场演示获取的300个图像。此集合用于测试现实世界应用中的例项分段性能。

[0159] 对于每一数据集,两个度量用于测量准确性:交并比(Intersection over Union, IoU)0.5-0.7($mAP_{.5:.7}$)范围内的平均精确度(mAP),以及IoU0.75-0.9($mAP_{.75:.9}$)范围内的平均精确度。这两个度量分别测量粗略分段和精细分段。为了进一步测试一般化,所有模型还使用掩码R-CNN的较快R-CNN分量测试对象检测。

[0160] 根据本公开的实施例的偏振掩码R-CNN和强度掩码R-CNN在上文所论述的四个测试集上测试。平均改进为粗略分段中14.3% mAP,以及精细分段中17.2% mAP。杂波问题中的性能改进在进行精细分段时较可见,其中性能差距从~1.1% mAP变为4.5% mAP。因此,偏振数据表现为提供有用的边缘信息从而允许模型更准确地对对象进行分段。如图12A中所见,偏振有助于准确地对其中其在强度图像中为模糊的杂波进行分段。因此,在来自图12A中展示的杂波测试集的实例中,根据本公开的一个实施例的偏振掩码R-CNN正确地检测透明对象的所有六个例项,与地面实况匹配,而对比的强度掩码R-CNN识别透明对象的六个例项中的仅四个。

[0161] 为了一般化到新环境,对于精细分段和粗略分段两者存在大得多的增益,且因此似乎透明对象的本征纹理在偏振图像中对于CNN较可见。如图12B中所展示,强度掩码R-CNN完全不能适应新颖环境,而偏振掩码R-CNN模型成功。虽然偏振掩码R-CNN能够正确地检测透明对象的所有例项,但例项掩码R-CNN不能检测一些例项(参见例如方框的右上角中的例项)。

[0162] 本公开的实施例还展示相对于打印冒充物的稳健性的同样大的改进,从而实现近似90% mAP。如此,本公开的实施例提供对于例如打印冒充物等透视投影问题较稳健的单眼解决方案。如图12C所示,强度掩码R-CNN被打印纸冒充物欺骗。在图12C所示的实例中,一个真实透明球体放置在描绘三个冒充的透明对象的打印件上。强度掩码R-CNN将两个打印冒充物不正确地识别为例项。另一方面,偏振掩码R-CNN较稳健,且仅将真实透明球体检测为例项。

[0163] 所有这些结果有助于阐释针对比如机器人箱柜拾取(RBP)等不受控且杂乱环境展

示的性能的显著改进。如图12D所示,在低照明条件下摆设的机器人拾取的情况下,强度掩码R-CNN模型仅能够检测透明对象的十一个例项中的五个。另一方面,偏振R-CNN模型能够在较差照明的情况下适应此环境,且正确地识别所有十一个例项。

[0164] 更详细地说且作为工业环境中的潜在应用的实例,计算机视觉系统被配置成控制机器人臂以通过将分段掩码供应到机器人臂的控制器来执行箱柜拾取。透明和半透明(非朗伯)对象的箱柜拾取是机器人技术领域的较难的遗留问题。为了展示高质量稳健分段的益处,将对比的强度掩码R-CNN在提供用于控制机器人臂对不同大小的杂乱透明摆设进行箱柜拾取的分段地图的过程中的性能与根据本公开的一个实施例的偏振掩码R-CNN的性能进行比较。

[0165] 箱柜拾取解决方案包含三个组件:用以隔离每一对象的分段组件;深度估计组件;以及位姿估计组件。为了理解分段的效果,简单深度估计和位姿,其中机器人臂移动到分段的中心且在其碰撞表面时停止。这在此实例中起作用,因为对象是完美的球体。稍微不准确的分段可能导致不正确的估计值且因此导致错误拾取。本申请实现偏振掩码R-CNN和强度掩码R-CNN之间的比较。系统在训练集外部的五个环境中测试(例如,在不同于获取训练图像的环境的条件下)。对于每一环境,十五个球体堆叠,且对机器人臂作出的拾取所有15个球体(使用吸杯夹持器)的正确/不正确(遗漏)拾取的数目进行计数,上限为15个不正确拾取。基于强度掩码R-CNN的模型不能够定期清空箱柜,因为机器人臂归因于较差分段质量而不断地遗漏某些拾取。另一方面,根据本公开的一个实施例的偏振掩码R-CNN模型成功地拾取所有90个球体,每6个正确拾取存在大致1个不正确拾取。这些结果验证了~20mAP的改进效果。

[0166] 如上所述,本公开的实施例可用作能够检测透明对象和不透明对象两者的计算机视觉或机器视觉系统的组件。

[0167] 在本公开的一些实施例中,相同预测器或统计模型900经训练以基于含有透明对象和不透明对象两者的经标记实例的训练数据检测透明对象和不透明对象两者(或生成第二表示空间中的第二张量C)。举例来说,在一些此类实施例中,使用偏振CNN架构,例如图9中展示的偏振掩码R-CNN架构。在一些实施例中,通过添加计算一个或多个额外模张量的一个或多个额外CNN主干来进一步修改图9中展示的偏振掩码R-CNN架构。可基于额外第一张量训练额外CNN主干。在一些实施例中,这些额外第一张量包含基于彩色强度图像(例如,不同波长的光的强度,比如红色强度图像或颜色通道、绿色强度图像或颜色通道,和蓝色强度图像或颜色通道)计算的图像地图。在一些实施例中,这些额外第一张量包含基于彩色强度图像的组合计算的图像地图。在一些实施例中,融合模块920融合来自CNN主干(例如,包含额外CNN主干)中的每一个的每一比例下的所有模张量。

[0168] 在本公开的一些实施例中,预测器900包含用于相对于透明对象检测不透明对象的一个或多个单独的统计模型。举例来说,预测器的集合(例如,第一预测器经训练以计算用于透明对象的第一分段掩码,且第二预测器经训练以计算用于不透明对象的第二分段掩码)可计算多个预测,其中单独的预测合并(例如,例如基于与分段掩码的每一像素相关联的置信度得分将第一分段掩码与第二分段掩码合并)。

[0169] 如上文的背景技术中所提及,使机器视觉或计算机视觉系统能够稳健地检测透明对象可在多种情形中应用,包含制造、生命科学、自动驾驶车辆和

[0170] 相应地,本公开的实施例的方面涉及用于通过使用从偏振域提取的特征使用计算机视觉检测透明对象的例项的系统和方法。透明对象在偏振域中比在强度域中具有更明显的纹理。偏振纹理中的此纹理可与根据本公开的实施例的特征提取器和偏振CNN模型一起利用。本公开的实施例对透明对象检测的性能的改进的实例经由与使用掩码R-CNN的例项分段的比较(例如,与使用强度图像而不使用偏振数据的掩码R-CNN的比较)来展现。因此,本公开的实施例

[0171] 虽然已经关于某些示例性实施例描述本发明,但应理解,本发明不限于所公开的实施例,而相反地,本发明希望涵盖包含于所附权利要求书的精神和范围内的各种修改和等效布置以及其等效物。

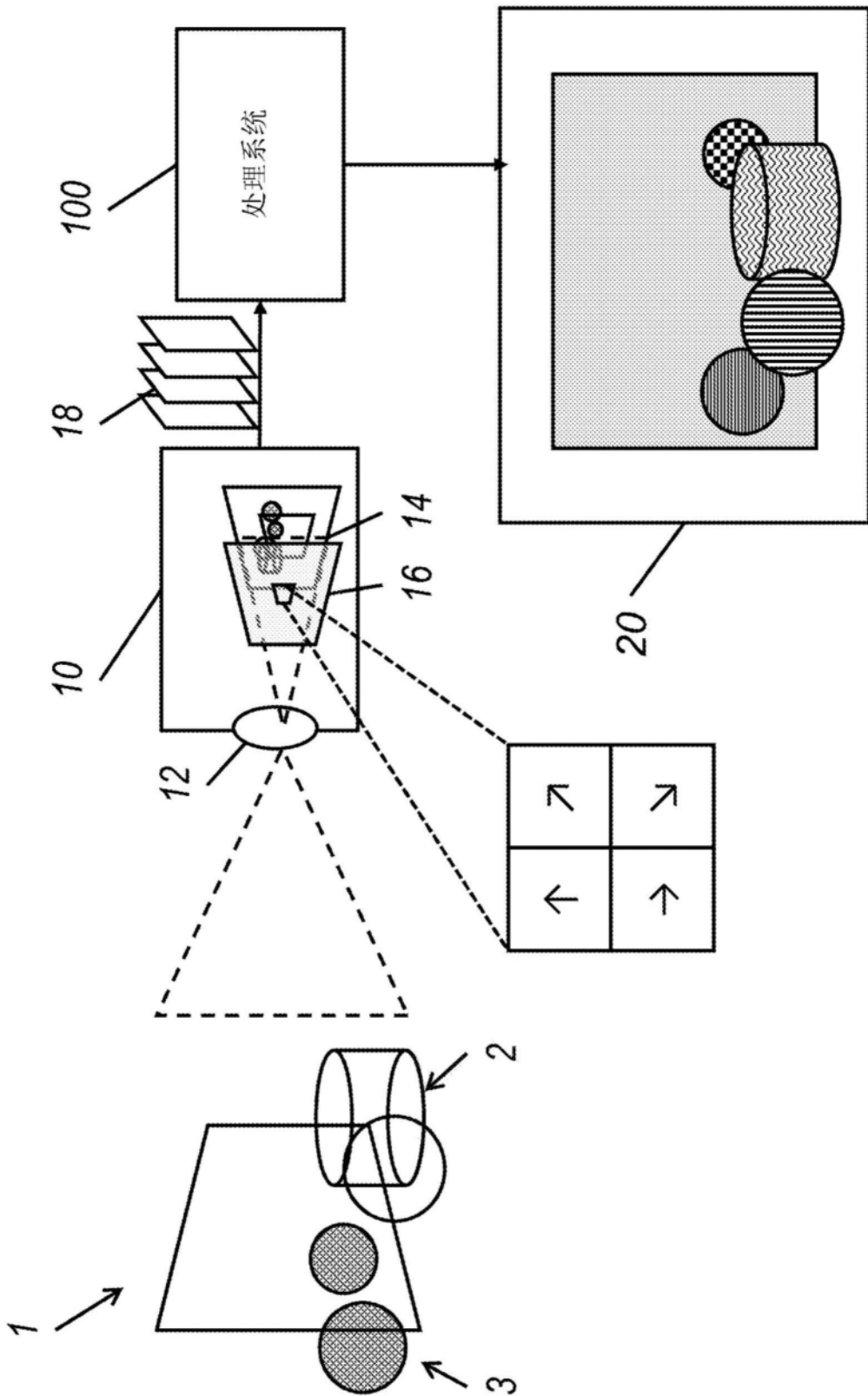


图1



图2A

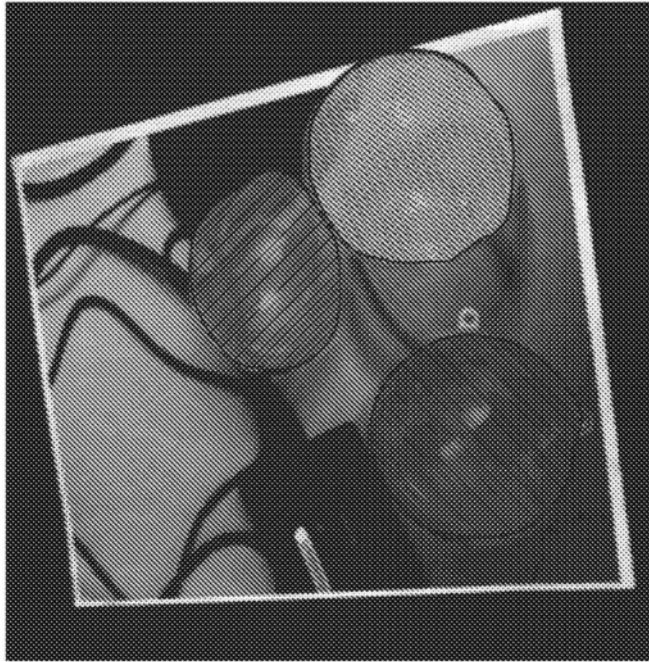


图2B

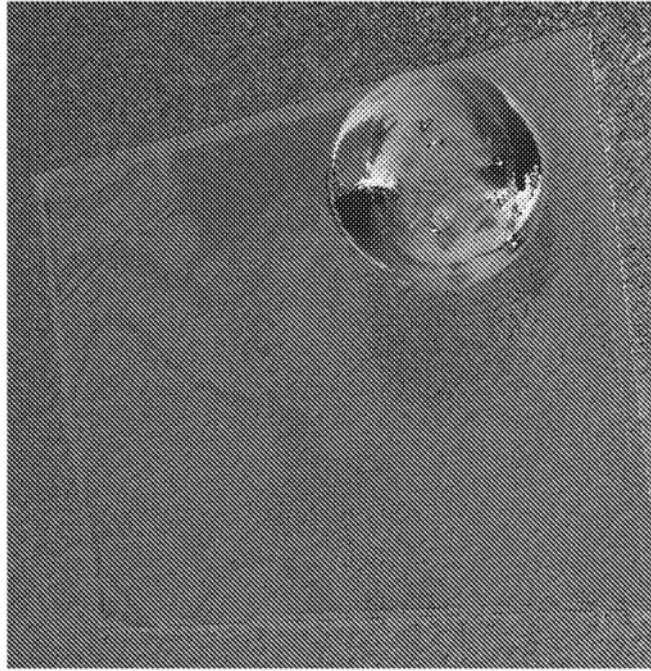


图2C

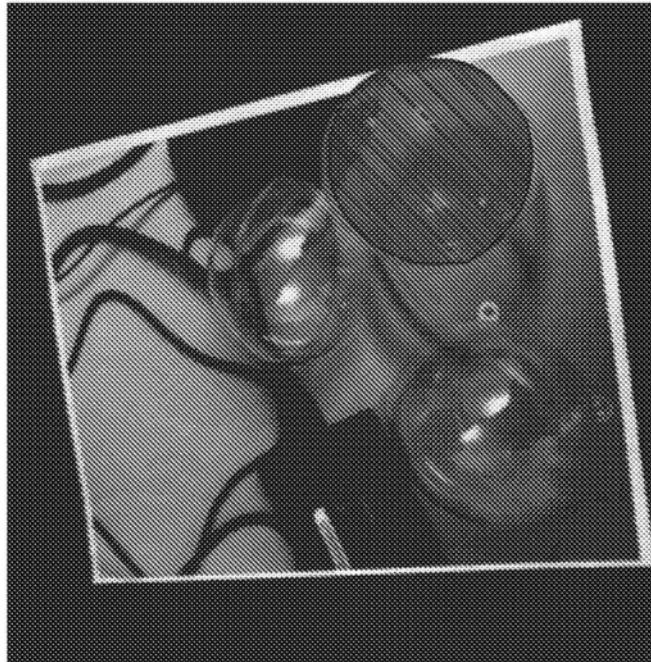


图2D

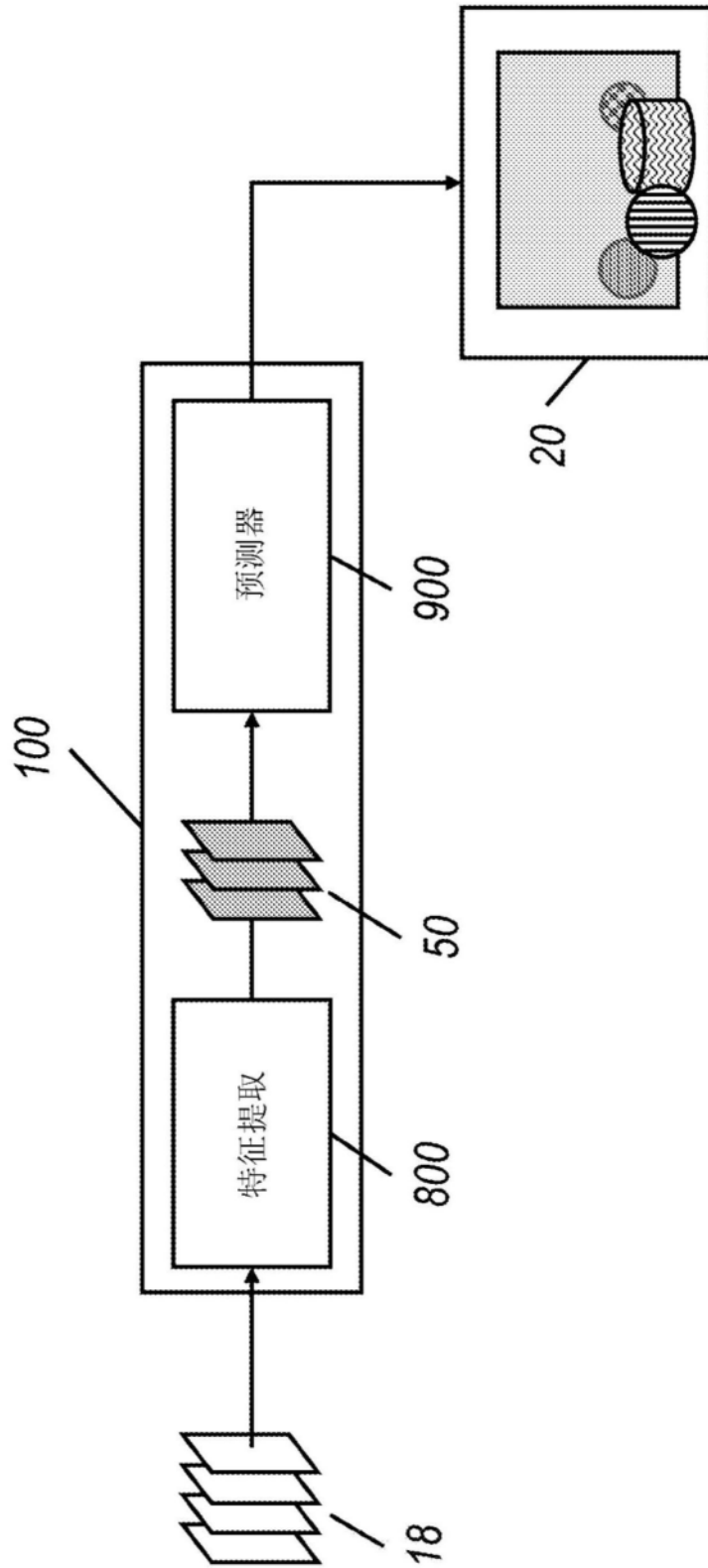


图3

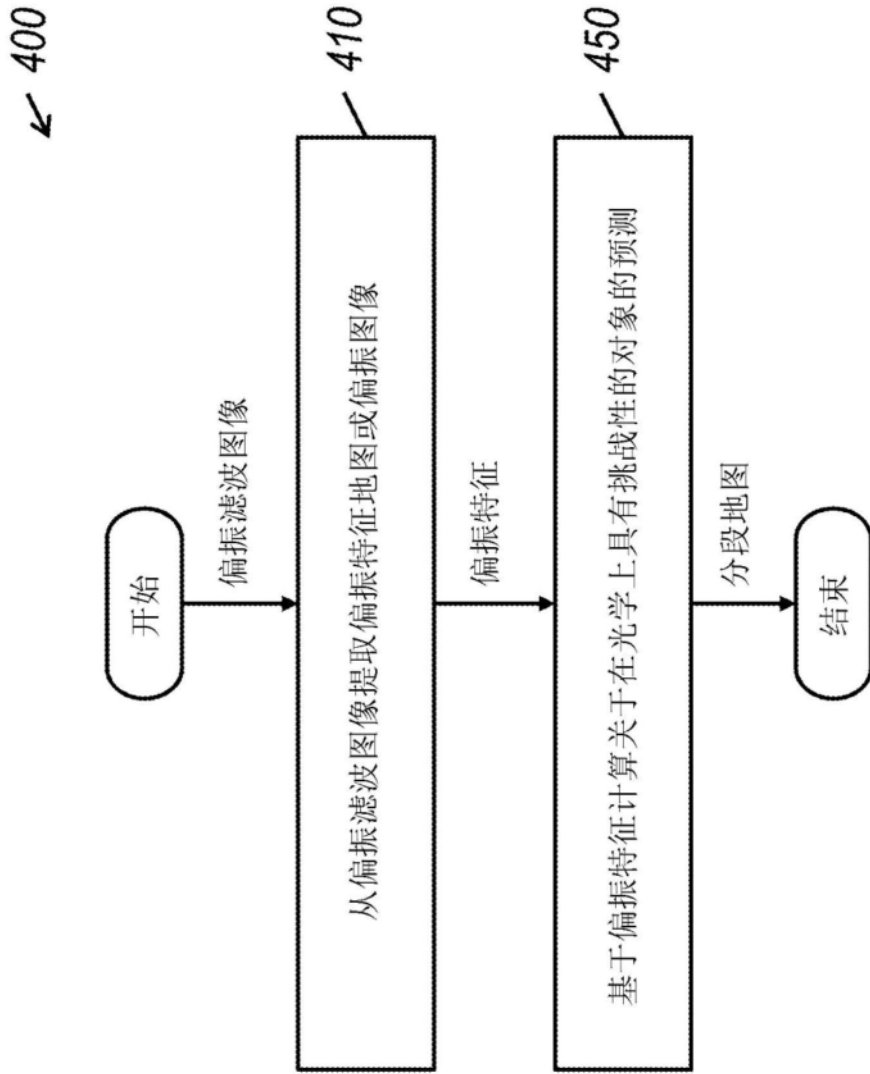


图4

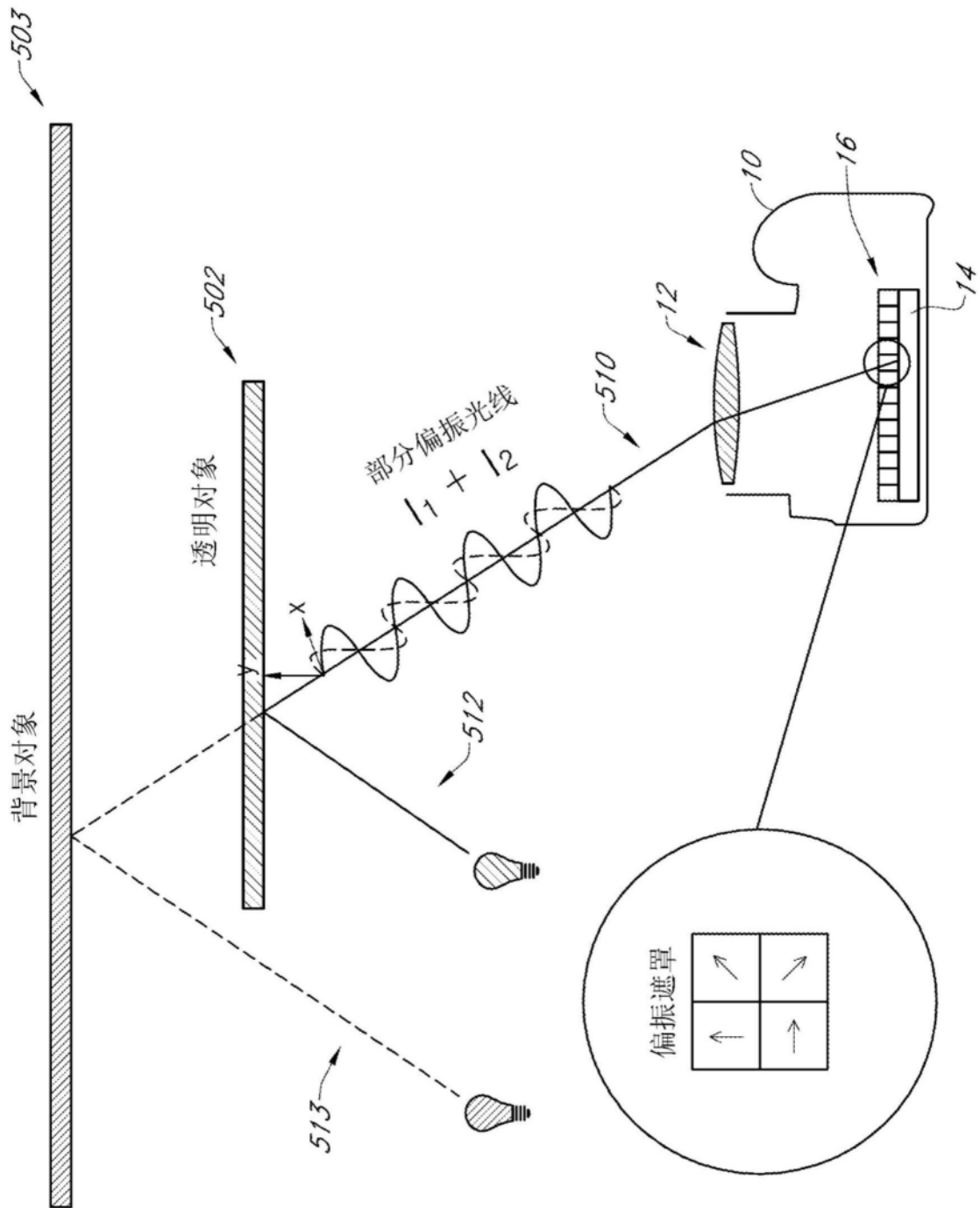


图5

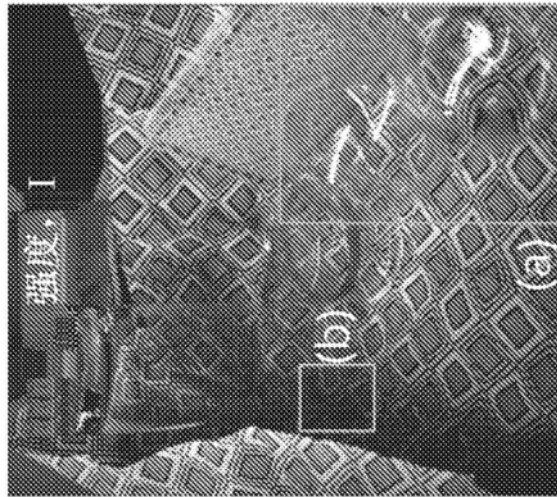


图6A

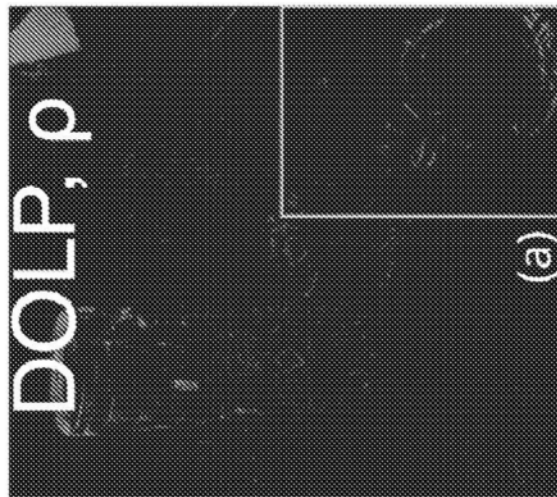


图6B

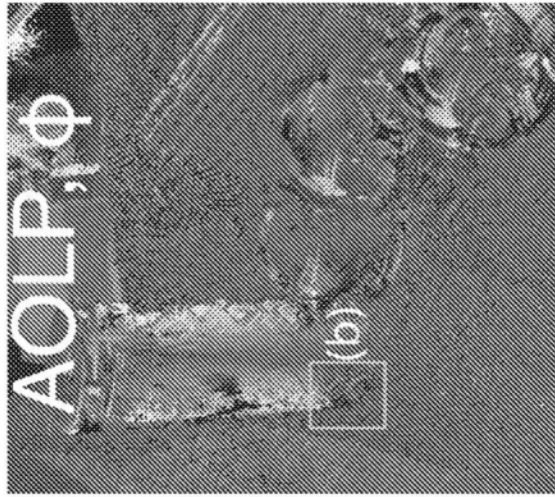
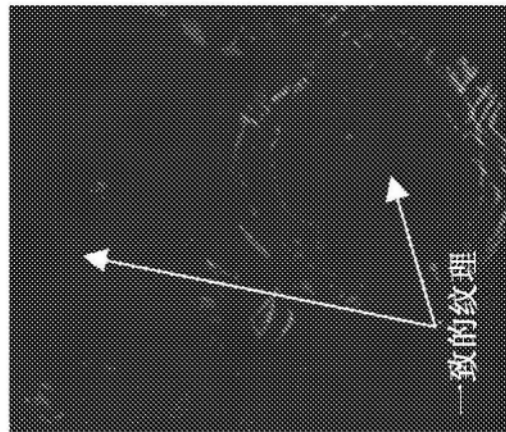
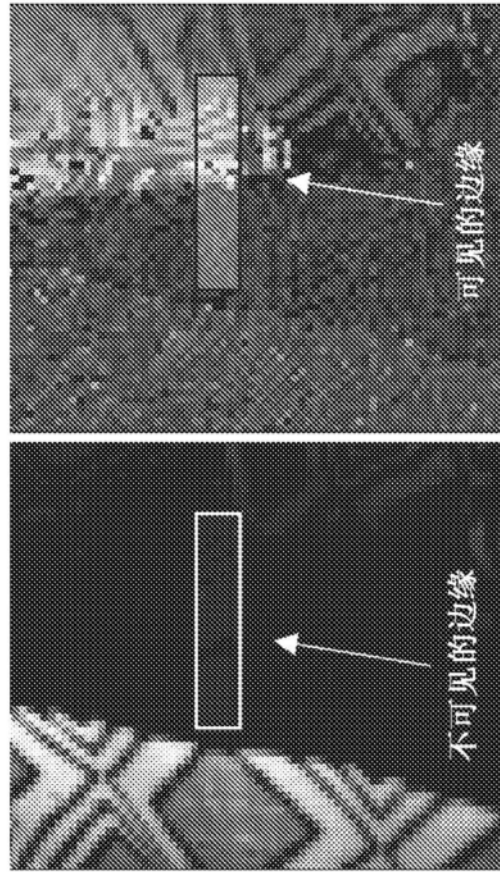


图6C



(a)

图7A



(b)

图7B

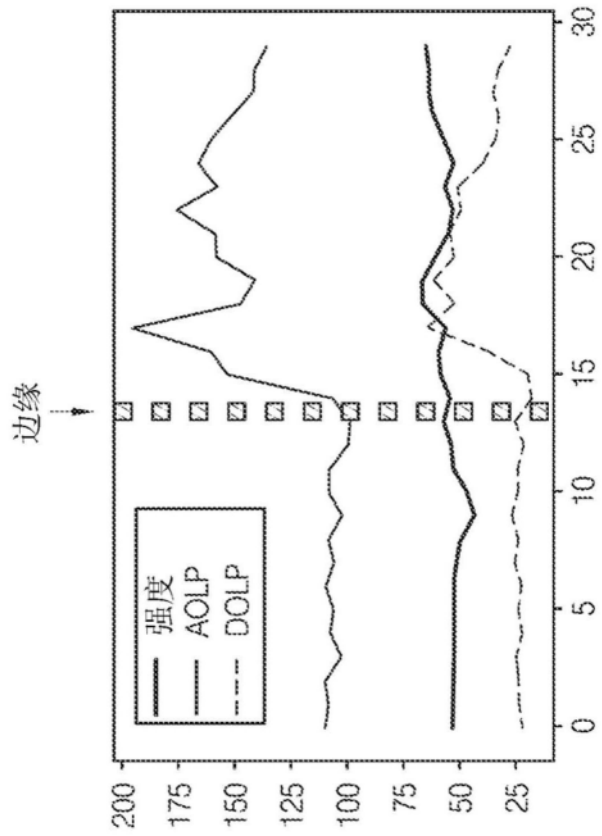


图7C

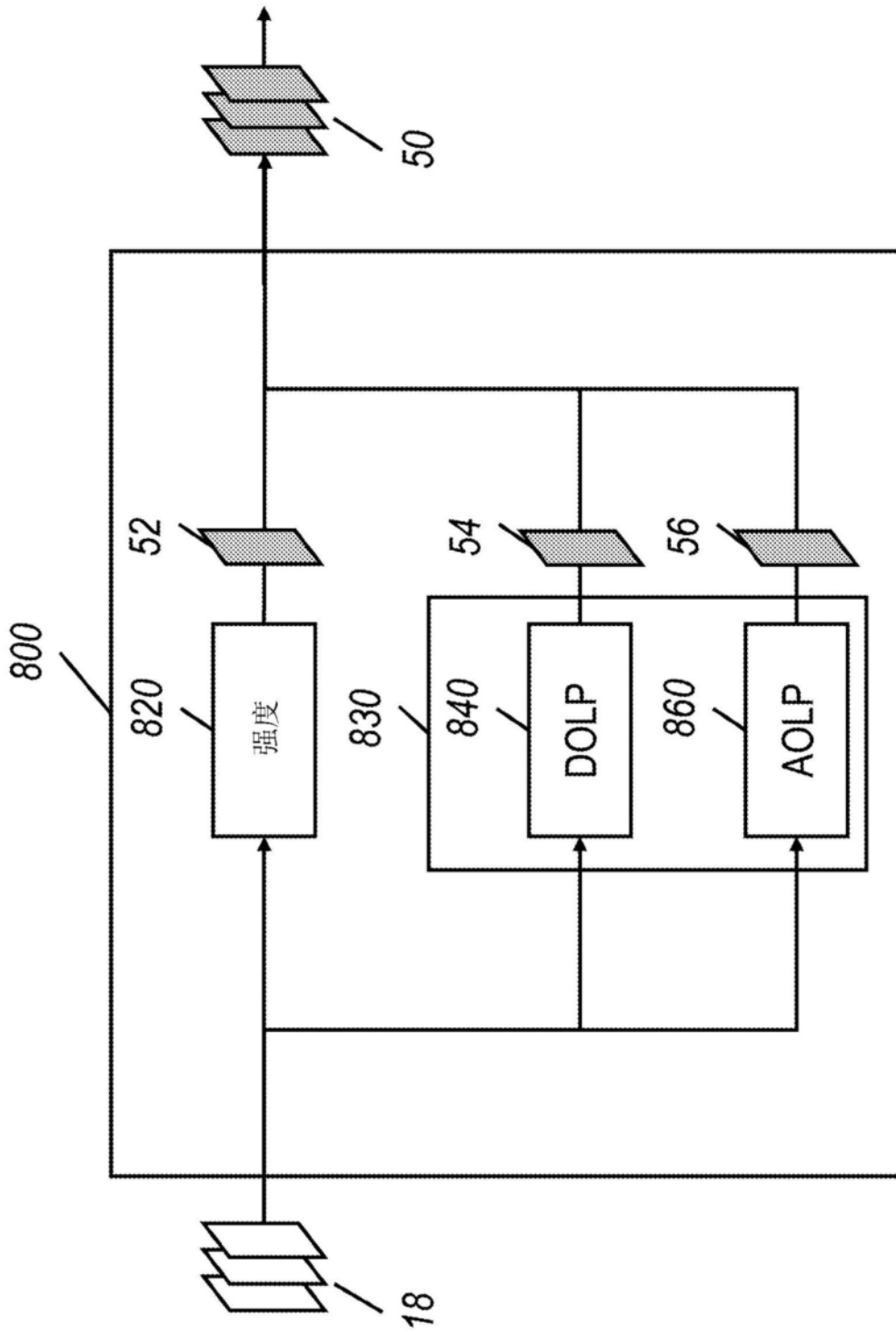


图8A

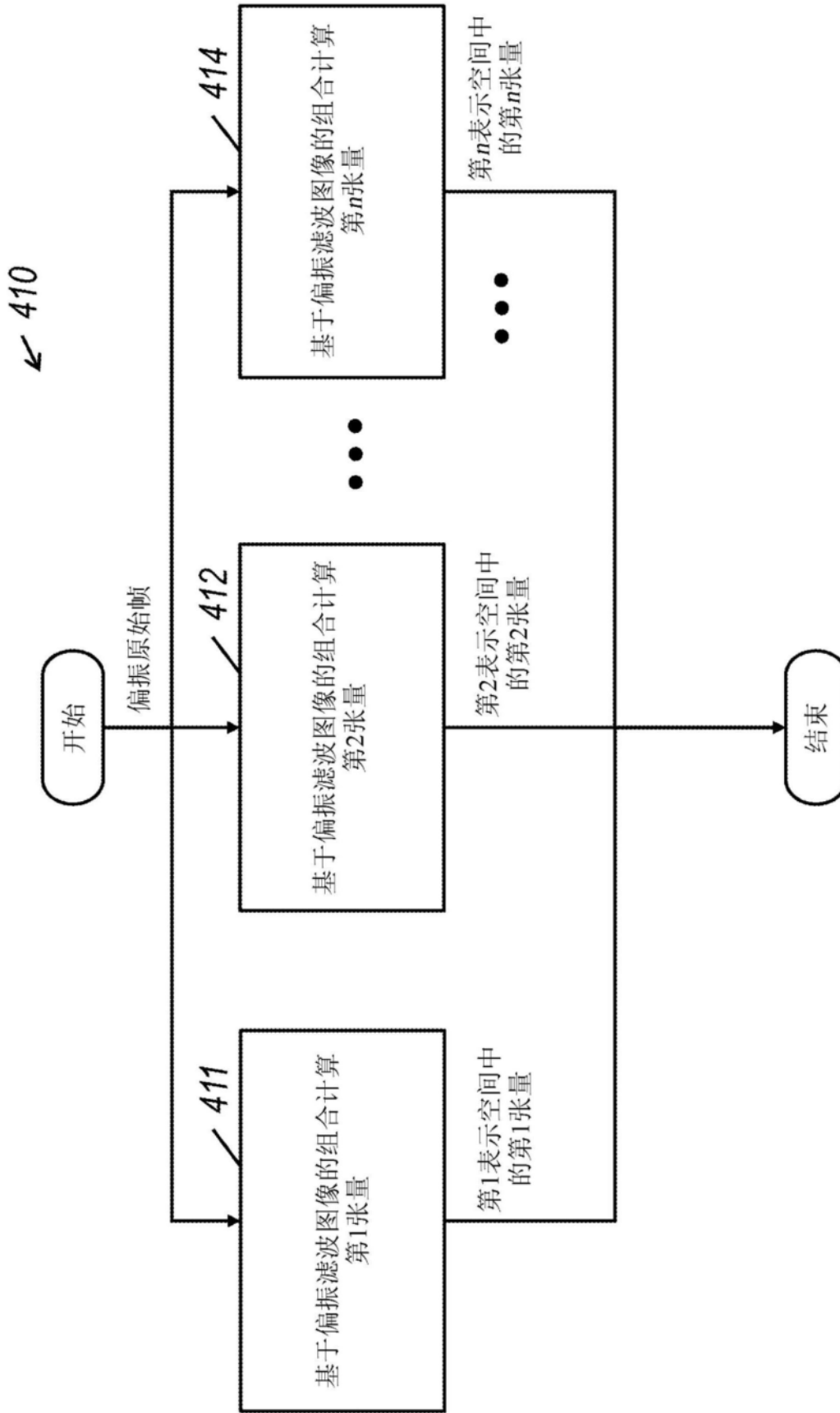


图8B

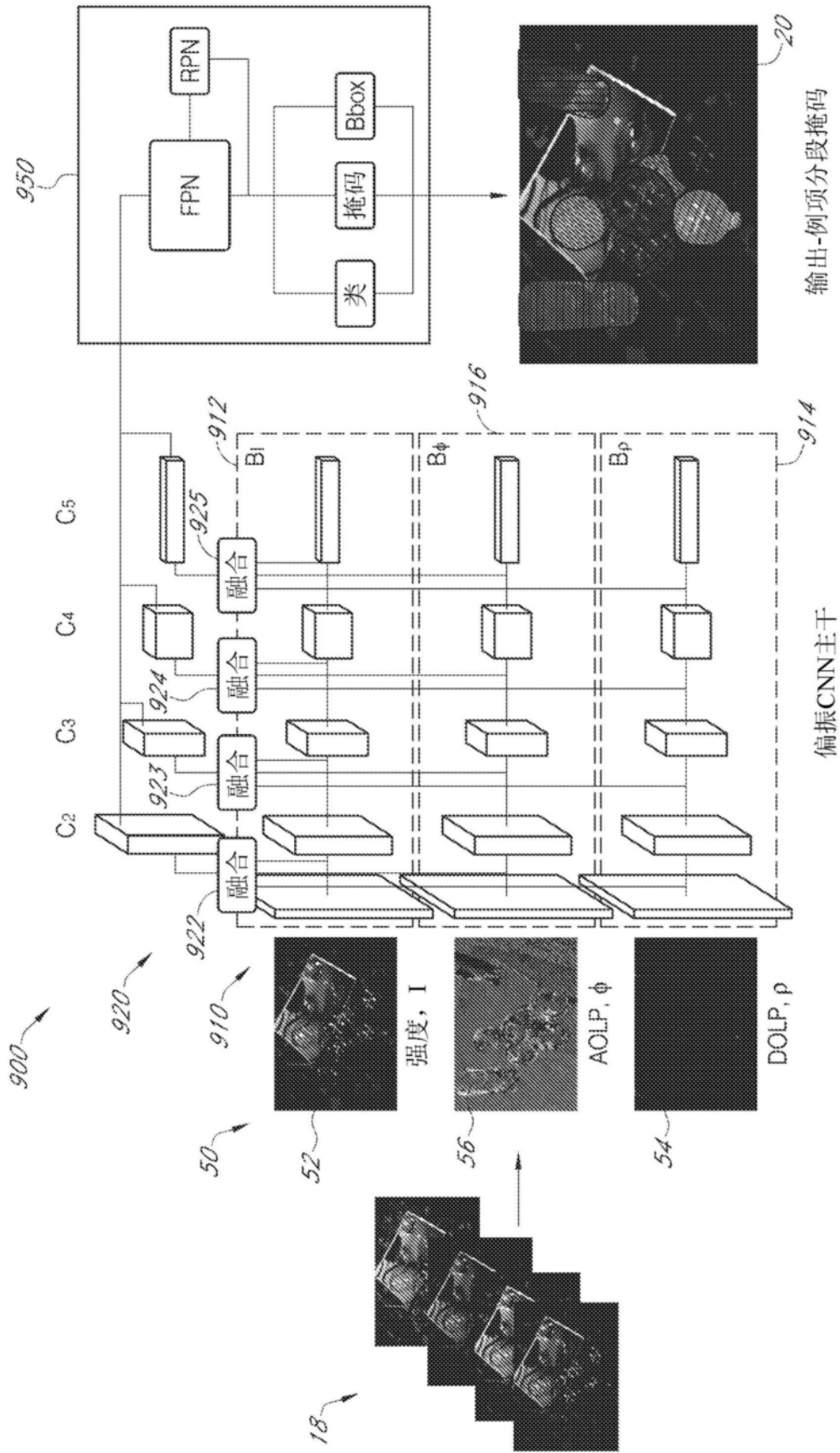


图9

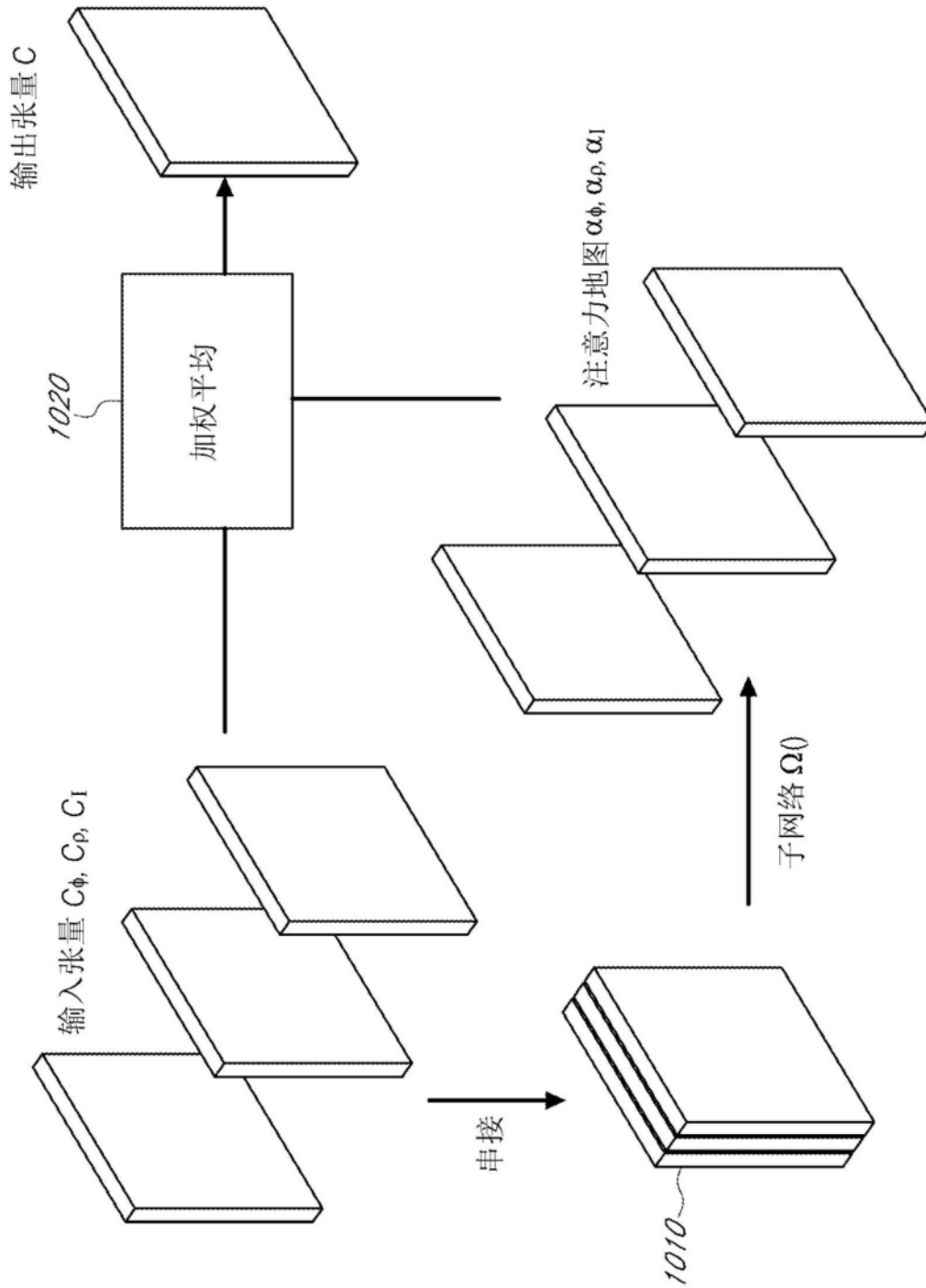


图10

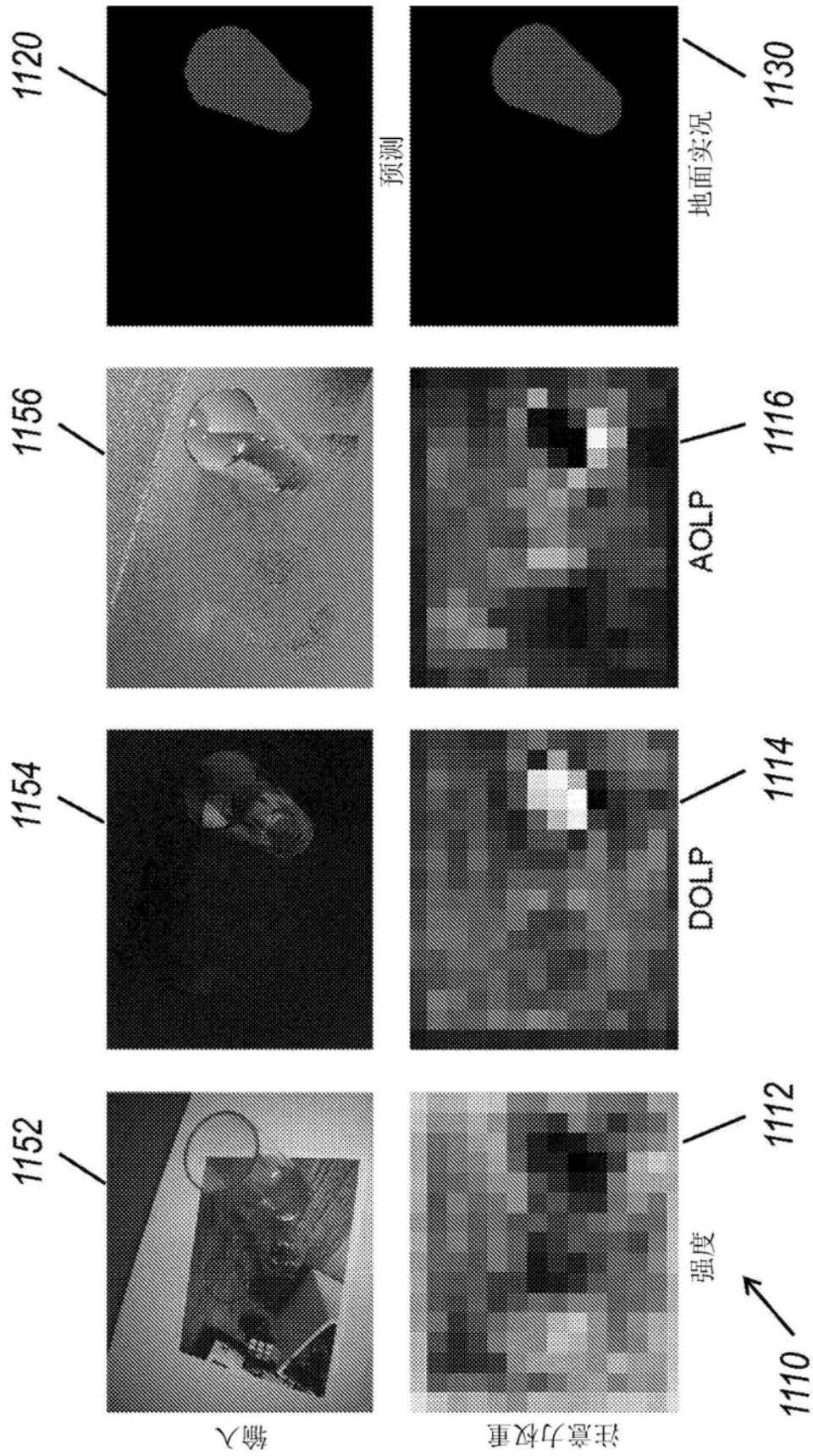
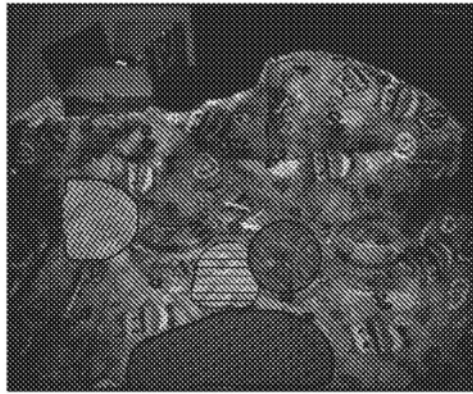
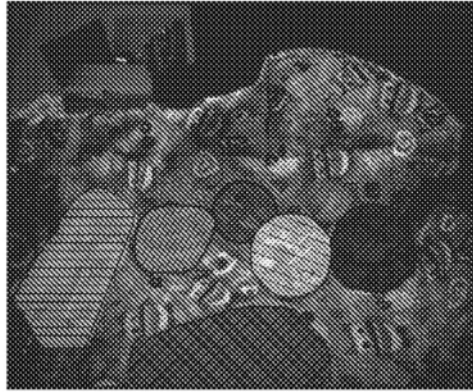


图11

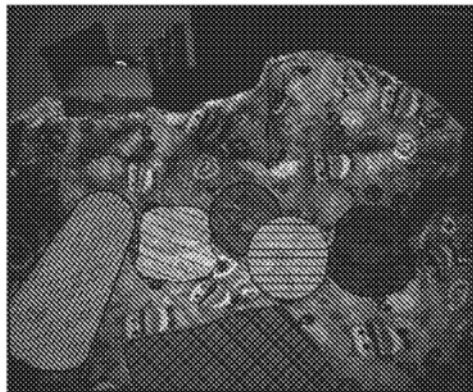
强度掩码
R-CNN



偏振掩码
R-CNN



地面实况



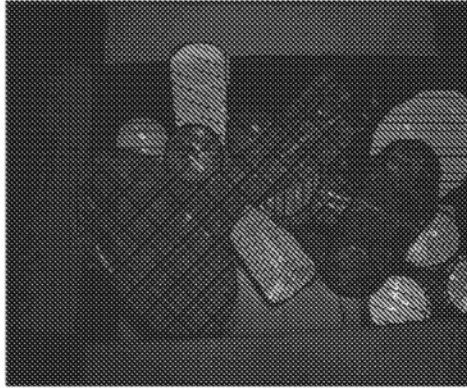
环境数据集

图12A

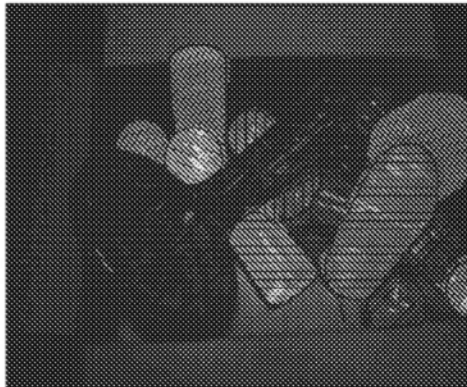
强度掩码
R-CNN



偏振掩码
R-CNN



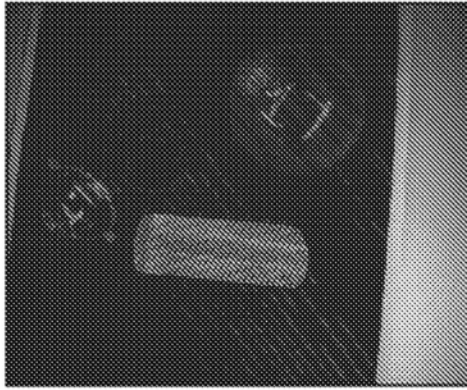
地面实况



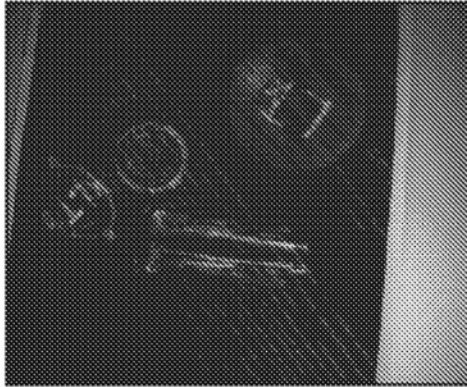
杂波数据集

图12B

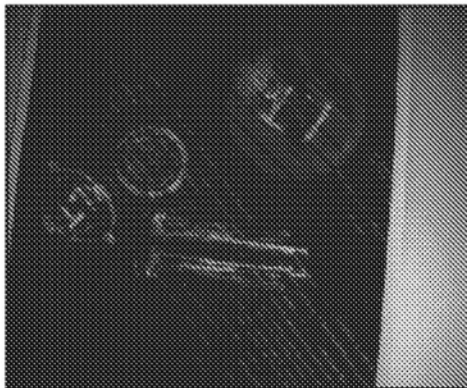
强度掩码
R-CNN



偏振掩码
R-CNN



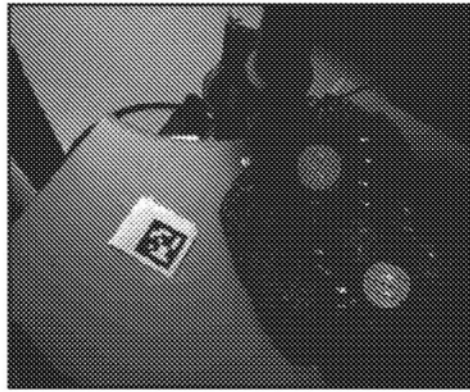
地面实况



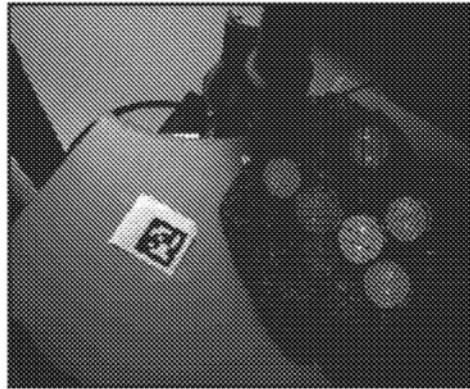
POS数据集

图12C

强度掩码
R-CNN



偏振掩码
R-CNN



地面实况



RBP数据集

图12D