



(12) 发明专利申请

(10) 申请公布号 CN 113781493 A

(43) 申请公布日 2021.12.10

(21) 申请号 202110002321.5

G06N 3/08 (2006.01)

(22) 申请日 2021.01.04

(71) 申请人 北京沃东天骏信息技术有限公司
地址 100176 北京市大兴区北京经济技术
开发区科创十一街18号院2号楼4层
A402室100176

申请人 北京京东世纪贸易有限公司

(72) 发明人 周芳汝 杨玫 安山

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021

代理人 王文思

(51) Int. Cl.

G06T 7/11 (2017.01)

G06T 7/73 (2017.01)

G06N 3/04 (2006.01)

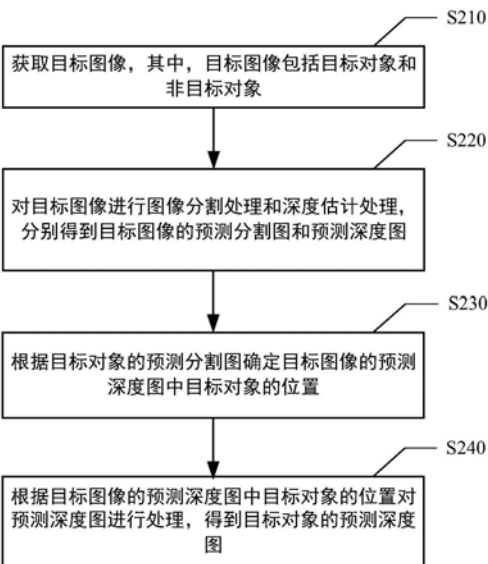
权利要求书2页 说明书13页 附图8页

(54) 发明名称

图像处理方法、装置、电子设备、介质及计算机程序产品

(57) 摘要

本公开实施例提供了一种图像处理方法、装置、电子设备、介质及计算机程序产品。该方法包括：获取目标图像，其中，目标图像包括目标对象和非目标对象；对目标图像进行图像分割处理和深度估计处理，分别得到目标图像的预测分割图和预测深度图；根据目标对象的预测分割图确定目标图像的预测深度图中目标对象的位置；根据目标图像的预测深度图中目标对象的位置对预测深度图进行处理，得到目标对象的预测深度图。



1. 一种图像处理方法, 包括:

获取目标图像, 其中, 所述目标图像包括目标对象和非目标对象;

对所述目标图像进行图像分割处理和深度估计处理, 分别得到所述目标图像的预测分割图和预测深度图;

根据所述目标对象的预测分割图确定所述目标图像的预测深度图中所述目标对象的位置; 以及

根据所述目标图像的预测深度图中所述目标对象的位置对所述预测深度图进行处理, 得到所述目标对象的预测深度图。

2. 根据权利要求1所述的方法, 其中, 所述对所述目标图像进行图像分割处理和深度估计处理, 分别得到目标图像的预测分割图和预测深度图, 包括:

利用图像处理模型处理所述目标图像, 分别得到所述目标图像的预测分割图和预测深度图, 其中, 所述图像处理模型是利用训练样本训练得到的, 其中, 所述训练样本包括样本图像和所述样本图像的深度标签和分割标签。

3. 根据权利要求2所述的方法, 其中, 所述图像处理模型包括特征提取网络、图像分割网络和深度估计网络;

所述利用图像处理模型处理所述目标图像, 分别得到所述目标图像的预测分割图和预测深度图, 包括:

利用所述特征提取网络处理所述目标图像, 得到第一中间特征图;

利用所述图像分割网络处理所述第一中间特征图, 得到第二中间特征图;

利用所述深度估计网络处理所述第一中间特征图, 得到第三中间特征图;

根据所述第二中间特征图和所述第三中间特征图, 生成第四中间特征图;

利用所述深度估计网络处理所述第四中间特征图, 得到所述目标图像的预测深度图;

以及

利用所述图像分割网络处理所述第二中间特征图, 得到所述目标图像的预测分割图。

4. 根据权利要求2所述的方法, 其中, 所述图像处理模型是利用训练样本训练得到的, 包括:

获取所述训练样本; 以及

利用所述训练样本训练全卷积神经网络模型, 得到所述图像处理模型。

5. 根据权利要求4所述的方法, 其中, 所述全卷积神经网络模型包括初始特征提取网络、初始图像分割网络和初始深度估计网络;

利用所述训练样本训练全卷积神经网络模型, 得到所述图像处理模型, 包括:

利用所述初始特征提取网络处理所述样本图像, 得到第五中间特征图;

利用所述初始图像分割网络处理所述第五中间特征图, 得到第六中间特征图;

利用所述初始深度估计网络处理所述第五中间特征图, 得到第七中间特征图;

根据所述第六中间特征图和所述第七中间特征图, 生成第八中间特征图;

利用所述初始深度估计网络处理所述第八中间特征图, 得到所述样本图像的预测深度图;

利用所述初始图像分割网络处理所述第六中间特征图, 得到所述样本图像的预测分割图;

将所述样本图像的深度标签、预测深度图、分割标签和预测分割图输入所述全卷积神经网络模型的损失函数,根据所述损失结果调整所述全卷积神经网络模型的网络参数,直至所述损失函数收敛;以及

将训练后的全卷积神经网络模型作为所述图像处理模型。

6. 根据权利要求5所述的方法,其中,所述对所述样本图像进行图像分割处理,得到所述样本图像的分割标签,包括:

对所述样本图像进行实例分割处理,得到所述样本图像的实例分割标签;

根据所述样本图像的实例分割标签,得到所述样本图像的语义分割标签;以及

将所述样本图像的语义分割标签作为所述样本图像的分割标签。

7. 根据权利要求1所述的方法,其中,所述根据所述目标图像的预测深度图中所述目标对象的位置对所述预测深度图进行处理,得到所述目标对象的预测深度图,包括:

将所述目标图像的预测深度图中除所述目标图像的位置以外的其他位置的像素值设置为预设像素值,得到所述目标对象的预测深度图。

8. 根据权利要求1~7中任一项所述的方法,其中,所述图像分割处理包括语义分割处理或实例分割处理。

9. 一种图像处理装置,包括:

获取模块,用于获取目标图像,其中,所述目标图像包括目标对象和非目标对象;

第一处理模块,用于对所述目标图像进行图像分割处理和深度估计处理,分别得到所述目标图像的预测分割图和预测深度图;

确定模块,用于根据所述目标对象的预测分割图确定所述目标图像的预测深度图中所述目标对象的位置;以及

第二处理模块,用于根据所述目标图像的预测深度图中所述目标对象的位置对所述预测深度图进行处理,得到所述目标对象的预测深度图。

10. 一种电子设备,包括:

一个或多个处理器;

存储器,用于存储一个或多个程序,

其中,当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现权利要求1~8中任一项所述的方法。

11. 一种计算机可读存储介质,其上存储有可执行指令,该指令被处理器执行时使处理器实现权利要求1~8中任一项所述的方法。

12. 一种计算机程序产品,所述计算机程序产品包括计算机程序,所述计算机程序被处理器执行时用于实现权利要求1~8中任一项所述的方法。

图像处理方法、装置、电子设备、介质及计算机程序产品

技术领域

[0001] 本公开实施例涉及计算机技术领域,更具体地,涉及一种图像处理方法、装置、电子设备、介质及计算机程序产品。

背景技术

[0002] 在计算机视觉领域,深度估计属于三维重建的一部分,其要求从二维图像中估计深度信息。针对某些具体任务,例如单目机器人躲避或寻找目标对象,目标对象可以为人,从二维图像中分割出目标对象,并估计目标对象的深度具有极其重要的作用。

[0003] 在实现本公开构思的过程中,发明人发现相关技术中至少存在如下问题:采用相关技术较难以实现针对目标图像中目标对象的深度估计。

发明内容

[0004] 有鉴于此,本公开实施例提供了一种图像处理方法、装置、电子设备、介质及计算机程序产品。

[0005] 本公开实施例的一个方面提供了一种图像处理方法,包括:获取目标图像,其中,上述目标图像包括目标对象和非目标对象;对上述目标图像进行图像分割处理和深度估计处理,分别得到上述目标图像的预测分割图和预测深度图;根据上述目标对象的预测分割图确定上述目标图像的预测深度图中上述目标对象的位置;以及,根据上述目标图像的预测深度图中上述目标对象的位置对上述预测深度图进行处理,得到上述目标对象的预测深度图。

[0006] 本公开实施例的另一个方面提供了一种图像处理装置,包括:获取模块,用于获取目标图像,其中,上述目标图像包括目标对象和非目标对象;第一处理模块,用于对上述目标图像进行图像分割处理和深度估计处理,分别得到上述目标图像的预测分割图和预测深度图;确定模块,用于根据上述目标对象的预测分割图确定上述目标图像的预测深度图中上述目标对象的位置;以及,第二处理模块,用于根据上述目标图像的预测深度图中上述目标对象的位置对上述预测深度图进行处理,得到上述目标对象的预测深度图。

[0007] 本公开实施例的另一个方面提供了一种电子设备,包括:一个或多个处理器;存储器,用于存储一个或多个程序,其中,当上述一个或多个程序被上述一个或多个处理器执行时,使得上述一个或多个处理器实现如上述的方法。

[0008] 本公开实施例的另一个方面提供了一种计算机可读存储介质,其上存储有可执行指令,该指令被处理器执行时使处理器实现如上述的方法。

[0009] 本公开实施例的另一个方面提供了一种计算机程序产品,上述计算机程序产品包括计算机程序,上述计算机程序被处理器执行时用于实现如上所述的方法。

[0010] 根据本公开的实施例,通过获取目标图像,目标图像包括目标对象和非目标对象,对目标图像进行图像分割处理和深度估计处理,分别得到目标图像的预测分割图和预测深度图,根据目标对象的预测分割图确定目标图像的预测深度图中目标对象的位置,并根据

目标图像的预测深度图中目标对象的位置对预测深度图进行处理,得到目标对象的预测深度图。由于将图像分割和深度估计结合,其中,根据预测分割图能够获得预测深度图中目标对象的位置,根据预测深度图中目标对象的位置对预测深度图进行处理能够获得目标对象的预测深度图,因此,至少部分地克服了相关技术中难以实现针对目标图像中目标对象的深度估计的技术问题,进而实现了较为准确地确定目标图像中目标对象的深度,并且方法的泛化性较强。

附图说明

[0011] 通过以下参照附图对本公开实施例的描述,本公开的上述以及其他目的、特征和优点将更为清楚,在附图中:

[0012] 图1示意性示出了可以应用本公开实施例的图像处理方法和装置的示例性系统架构;

[0013] 图2示意性示出了根据本公开实施例的一种图像处理方法的流程图;

[0014] 图3示意性示出了根据本公开实施例的图像处理模型结构图;

[0015] 图4示意性示出了根据本公开实施例的另一种图像处理方法的流程图;

[0016] 图5示意性示出了根据本公开实施例的目标图像的示意图;

[0017] 图6示意性示出了根据本公开实施例的目标图像的预测深度图;

[0018] 图7示意性示出了根据本公开实施例的目标图像的预测分割图;

[0019] 图8示意性示出了根据本公开实施例的一种目标对象的预测深度图;

[0020] 图9示意性示出了根据本公开实施例的另一种目标对象的预测深度图;

[0021] 图10示意性示出了根据本公开实施例的再一种目标对象的示意图;

[0022] 图11示意性示出了根据本公开实施例的再一种图像处理方法的流程图;

[0023] 图12示意性示出了根据本公开实施例的又一种图像处理方法的流程图;

[0024] 图13示意性示出了根据本公开实施例的图像处理装置的框图;以及

[0025] 图14示意性示出了根据本公开实施例的适于图像处理方法的电子设备的框图。

具体实施方式

[0026] 以下,将参照附图来描述本公开的实施例。但是应该理解,这些描述只是示例性的,而并非要限制本公开的范围。在下面的详细描述中,为便于解释,阐述了许多具体的细节以提供对本公开实施例的全面理解。然而,明显地,一个或多个实施例在没有这些具体细节的情况下也可以被实施。此外,在以下说明中,省略了对公知结构和技术的描述,以避免不必要地混淆本公开的概念。

[0027] 在此使用的术语仅仅是为了描述具体实施例,而并非意在限制本公开。在此使用的术语“包括”、“包含”等表明了所述特征、步骤、操作和/或部件的存在,但是并不排除存在或添加一个或多个其他特征、步骤、操作或部件。

[0028] 在此使用的所有术语(包括技术和科学术语)具有本领域技术人员通常所理解的含义,除非另外定义。应注意,这里使用的术语应解释为具有与本说明书的上下文相一致的含义,而不应以理想化或过于刻板的方式来解释。

[0029] 在使用类似于“A、B和C等中至少一个”这样的表述的情况下,一般来说应该按照本

领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B和C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的系统等)。在使用类似于“A、B或C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B或C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的系统等)。

[0030] 本公开的实施例提供了一种图像处理方法、图像处理装置以及应用该方法的电子设备。该方法包括获取目标图像,其中,目标图像包括目标对象和非目标对象。对目标图像进行图像分割处理和深度估计处理,分别得到目标图像的预测分割图和预测深度图。根据目标对象的预测分割图确定目标图像的预测深度图中目标对象的位置。根据目标图像的预测深度图中目标对象的位置对预测深度图进行处理,得到目标对象的预测深度图。

[0031] 图1示意性示出了根据本公开实施例的可以应用图像处理方法或者装置的示例性系统架构100。需要注意的是,图1所示仅为可以应用本公开实施例的系统架构的示例,以帮助本领域技术人员理解本公开的技术内容,但并不意味着本公开实施例不可以用于其他设备、系统、环境或场景。

[0032] 如图1所示,根据该实施例的系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线和/或无线通信链路等。

[0033] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用,例如图像处理类应用、模型构建类应用、搜索类应用、即时通信工具、邮箱客户端和/或社交平台软件等(仅为示例)。

[0034] 终端设备101、102、103可以是具有显示屏并且支持图像处理、网页浏览的各种电子设备,包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等。

[0035] 服务器105可以是提供各种服务的服务器,例如对用户利用终端设备101、102、103所浏览的网站、处理的图片等提供支持的后台管理服务器(仅为示例)。后台管理服务器可以对接收到的图像、摄像信息等数据进行处理、分析以及保存等处理,并将处理结果(例如根据用户请求获取或生成的网页、信息、或数据等)反馈给终端设备。

[0036] 需要说明的是,本公开实施例所提供的图像处理方法一般可以由服务器105执行。相应地,本公开实施例所提供的图像处理装置一般可以设置于服务器105中。本公开实施例所提供的图像处理方法也可以由不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群执行。相应地,本公开实施例所提供的图像处理装置也可以设置于不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群中。或者,本公开实施例所提供的图像处理方法也可以由终端设备101、102或103执行,或者也可以由不同于终端设备101、102或103的其他终端设备执行。相应地,本公开实施例所提供的图像处理装置也可以设置于终端设备101、102或103中,或设置于不同于终端设备101、102或103的其他终端设备中。

[0037] 例如,目标图像可以原本存储在终端设备101、102或103中的任意一个(例如,终端设备101,但不限于此)之中,或者存储在外部存储设备上并可以导入到终端设备101中。然

后,终端设备101可以在本地执行本公开实施例所提供的图像处理方法,或者将目标图像发送到其他终端设备、服务器、或服务器集群,并由接收该目标图像的其他终端设备、服务器、或服务器集群来执行本公开实施例所提供的图像处理方法。

[0038] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器的。

[0039] 图2示意性示出了根据本公开实施例的一种图像处理方法的流程图。

[0040] 如图2所示,该方法包括操作S210~S240。

[0041] 在操作S210,获取目标图像,其中,目标图像包括目标对象和非目标对象。

[0042] 根据本公开的实施例,目标图像可以为单目图像,目标图像可以包括目标对象和非目标对象,其中,目标对象可以是目标图像中的人,而非目标对象可以是目标图像中的背景物体等,例如桌子、树和车等,但是并不局限于此,还可以根据实际需求指定任一目标为目标对象,而与指定的目标对象不同类别的均为非目标对象。目标对象的数量可以包括一个或多个。

[0043] 在操作S220,对目标图像进行图像分割处理和深度估计处理,分别得到目标图像的预测分割图和预测深度图。

[0044] 根据本公开的实施例,对目标图像进行深度估计处理,可以是根据像素值关系反映的深度关系对目标图像上的每个像素点进行深度估计。

[0045] 根据本公开的实施例,图像分割可以是语义分割的方法,但是并不局限于此,还可以是实例分割的方法。语义分割可以为图像中的每个像素点进行类别划分,得到一幅与图像尺寸具有对应关系的语义分割掩码,即图像的预测语义分割图,但语义分割不区分同一类别的不同对象,即不区分实例。实例分割不仅可以实现像素点的类别划分,还可以区分同一类别的不同对象,即区分实例,通过对图像进行实例分割可以得到一幅与图像尺寸具有对应关系的实例分割掩码,即图像的预测实例分割图。可以将图像的预测语义分割图和图像的预测实例分割图统称为图像的预测分割图。在本公开的实施例中,不同类别可以用不同颜色表征,相应的,图像的预测分割图中不同颜色表征不同类别。

[0046] 根据本公开的实施例,可以采用语义分割或实例分割对目标图像进行图像分割处理,得到目标图像的预测分割图。其中,目标图像的预测分割图尺寸可以根据实际情况确定,在此不作具体限定。例如,目标图像的预测分割图尺寸和目标图像尺寸相同。或者,目标图像的预测分割图尺寸是目标图像尺寸的二分之一。

[0047] 根据本公开的实施例,采用语义分割方法进行图像分割处理能够在速度实现实时,从而满足实时任务的处理要求。

[0048] 在操作S230,根据目标对象的预测分割图确定目标图像的预测深度图中目标对象的位置。

[0049] 根据本公开的实施例,可以根据目标图像的预测分割图中目标对象的预测分割图确定目标图像的预测分割图中目标对象的位置,由于目标图像的预测分割图中的目标对象的位置与目标图像的预测深度图中目标对象的位置相对应,因此,可以基于目标图像的预测分割图中的目标对象的位置获得目标图像的预测深度图中目标对象的位置。

[0050] 在操作S240,根据目标图像的预测深度图中目标对象的位置对预测深度图进行处理,得到目标对象的预测深度图。

[0051] 根据本公开的实施例,在实施本公开的相关技术中,可以结合梯度和纹理特征对单目图像进行深度估计。即可以把目标图像的梯度信息和纹理信息作为深度线索,协助深度卷积网络学习目标图像的深度信息,得到目标图像的预测深度图。还可以将目标图像输入基于相机位姿估计得到的深度估计网络,以获得目标图像的预测深度图。

[0052] 但是,以上方法均只能对整个目标图像进行深度估计,难以提取目标图像中目标对象的深度。

[0053] 在实施本公开的相关技术中,还可以采用基于样本的学习方法实现深度估计。基于样本的学习方法是构建数据集,将对目标对象的深度估计问题转换为检索问题,采用特征匹配的方法在数据集中进行检索,得到目标图像中目标对象的深度估计结果。基于样本的学习方法能够估计与数据集中的图像具有匹配关系的目标图像中目标对象的深度。但是如果无法在数据集中检索到与目标图像匹配的图像,则不能实现针对目标图像中的目标对象的深度估计。此外,采用上述方法能够估计的是目标图像中目标对象的深度,并不能得到目标图像中目标对象的位置,该方法泛化性较差且估计精度不高。

[0054] 根据本公开的实施例,由于将图像分割和深度估计进行结合,其中,根据目标图像的预测分割图能够获得目标图像的预测深度图中目标对象的位置,根据目标图像的预测深度图中目标对象的位置对预测深度图进行处理能够获得目标对象的预测深度图,因此,能够实现较为准确地确定目标图像中目标对象的深度。此外,由于不是通过检索匹配的方式实现的深度估计,受样本的影响较小,因此,本公开实施例的方案泛化性较强。

[0055] 需要说明的是,如果目标对象的数量包括至少两个,则针对每个目标对象均可以采用上述操作S210~S240提供的方法得到该目标对象的预测深度图。在呈现上,可以将各个目标对象的预测深度图呈现在一幅预测深度图上,或者,每个目标对象的预测深度图各自呈现。

[0056] 根据本公开实施例的技术方案,通过获取目标图像,目标图像包括目标对象和非目标对象,对目标图像进行图像分割处理和深度估计处理,分别得到目标图像的预测分割图和预测深度图,根据目标对象的预测分割图确定目标图像的预测深度图中目标对象的位置,并根据目标图像的预测深度图中目标对象的位置对预测深度图进行处理,得到目标对象的预测深度图。由于将图像分割和深度估计进行结合,其中,根据预测分割图能够获得预测深度图中目标对象的位置,根据预测深度图中目标对象的位置对预测深度图进行处理能够获得目标对象的预测深度图,因此,至少部分地克服了相关技术中难以实现针对目标图像中目标对象的深度估计的技术问题,进而实现了较为准确地确定目标图像中目标对象的深度,并且方法的泛化性较强。

[0057] 下面参考图3~图10,结合具体实施例对图2所示的方法做进一步说明。

[0058] 图3示意性示出了根据本公开实施例的图像处理模型结构图。

[0059] 如图3所示,图像处理模型包括特征提取网络、图像分割网络和深度估计网络。

[0060] 根据本公开的实施例,本公开构建一种图像处理模型,该图像处理模型是利用训练样本对编解码网络进行训练得到的,即将待预测的目标图像输入至图像处理模型中,分别输出目标图像的预测分割图和预测深度图,进而得到该目标图像的目标对象的预测深度图。根据本公开的实施例,本公开构建的图像处理模型,弥补了相关技术中不能同时输出包含目标图像的预测分割图和预测深度图的技术缺失。

[0061] 图4示意性示出了根据本公开实施例的另一种图像处理方法的流程图。

[0062] 如图4所示,利用图像处理模型处理目标图像,分别得到目标图像的预测分割图和预测深度图可以包括如下操作S410~S460。

[0063] 在操作S410,利用特征提取网络处理目标图像,得到第一中间特征图。

[0064] 在操作S420,利用图像分割网络处理第一中间特征图,得到第二中间特征图。

[0065] 在操作S430,利用深度估计网络处理第一中间特征图,得到第三中间特征图。

[0066] 在操作S440,根据第二中间特征图和第三中间特征图,生成第四中间特征图。

[0067] 在操作S450,利用深度估计网络处理第四中间特征图,得到目标图像的预测深度图。

[0068] 在操作S460,利用图像分割网络处理第二中间特征图,得到目标图像的预测分割图。

[0069] 如图3和图4所示,为了达到实时的目标对象的深度估计,使用MobileNet+ASPP (Atrous Spatial Pyramid Pooling,空洞空间卷积池化金字塔)模块作为特征提取网络,即编码网络。使用深度可分离卷积作为解码网络,即图像分割网络和深度估计网络。

[0070] 根据本公开的实施例,可以记目标图像的高和宽分别为H和W。将目标图像输入到MobileNet模块,输出特征图 f_2 ,其大小为 $\left(\frac{H}{32}, \frac{W}{32}\right)$,在这个过程中,中间层的输出为特征图 f_1 ,其大小为 $\left(\frac{H}{4}, \frac{W}{4}\right)$ 。可以将特征图 f_2 输入到ASPP模块中,得到的输出与特征图 f_2 融合,输出特征图 f_3 。将特征图 f_3 上采样至 $\left(\frac{H}{4}, \frac{W}{4}\right)$ 大小,然后将特征图 f_1 和 f_3 融合,最后以特征图 f_4 作为特征提取网络的输出,特征图 f_4 即为第一中间特征图。

[0071] 根据本公开的实施例,图像分割网络和深度估计网络均以相同的特征图 f_4 (即第一中间特征图)作为输入,分别输出第二中间特征图(即特征图 f_5)和第三中间特征图(即特征图 f_6)。需要说明的是,由于通过特征提取网络处理后的第一中间特征图,分别作为深度估计网络和图像分割网络的输入,因此,特征提取网络综合考虑了目标图像的细节信息和抽象信息。

[0072] 根据本公开的实施例,由于在预测深度图中,同一个目标对象的深度值是比较接近的,并且在目标对象的边界处深度值的梯度可能会较大,因此,为了获得精度更高的预测深度图,可以将图像分割网络输出的第二中间特征图(即特征图 f_5)输入到深度估计网络中,并与第三中间特征图(即特征图 f_6)结合,得到第四中间特征图(即特征图 f_7)。第四中间特征图通过卷积和上采样得到大小为 $\left(\frac{H}{2}, \frac{W}{2}\right)$ 的预测深度图。第二中间特征图通过卷积和上采样得到大小为 $\left(\frac{H}{2}, \frac{W}{2}\right)$ 的预测分割图。

[0073] 根据本公开的实施例,通过将图像分割网络中的第二中间特征图输入到深度估计网络中,来矫正深度估计网络的预测结果,使的深度估计的预测结果更为准确。

[0074] 根据本公开的实施例,根据目标图像的预测深度图中目标对象的位置对预测深度图进行处理,得到目标对象的预测深度图,可以包括如下操作。

[0075] 将目标图像的预测深度图中除目标图像的位置以外的其他位置的像素值设置为

预设像素值,得到目标对象的预测深度图。

[0076] 根据本公开的实施例,可以将目标图像的预测深度图中除目标图像的位置以外的其他位置的像素值设置为预设像素值,得到目标对象的预测深度图。其中,预设像素值可以根据实际情况设定,在此不作具体限定,例如预设像素值可以为0。

[0077] 示例性的,图5示意性示出了根据本公开实施例的目标图像的示意图。图6示意性示出了根据本公开实施例的目标图像的预测深度图。图7示意性示出了根据本公开实施例的目标图像的预测分割图。图8示意性示出了根据本公开实施例的一种目标对象的预测深度图。图8中目标对象为人。图9示意性示出了根据本公开实施例的另一种目标对象的预测深度图。图9中目标对象为冰箱。图10示意性示出了根据本公开实施例的再一种目标对象的示意图。图10中目标对象的数量包括多个。

[0078] 如图5~图10所示,将目标图像的预测分割图作用到目标图像的预测深度图后得到的目标对象的预测深度图。在图7~图10中,黑色表示非目标对象区域。

[0079] 根据本公开的实施例,利用图像处理模型处理目标图像,分别得到目标图像的预测分割图和预测深度图,其中,图像处理模型是利用训练样本训练得到的,其中,训练样本包括样本图像和样本图像的深度标签和分割标签。

[0080] 根据本公开的实施例,图像处理模型是利用训练样本训练得到的,可以包括如下操作。

[0081] 获取训练样本。利用训练样本训练全卷积神经网络模型,得到图像处理模型。

[0082] 根据本公开的实施例,全卷积神经网络模型包括初始特征提取网络、初始图像分割网络和初始深度估计网络。

[0083] 根据本公开的实施例,利用训练样本训练全卷积神经网络模型,得到图像处理模型,可以包括如下操作。

[0084] 利用初始特征提取网络处理样本图像,得到第五中间特征图。利用初始图像分割网络处理第五中间特征图,得到第六中间特征图。利用初始深度估计网络处理第五中间特征图,得到第七中间特征图。根据第六中间特征图和第七中间特征图,生成第八中间特征图。利用初始深度估计网络处理第八中间特征图,得到样本图像的预测深度图。利用初始图像分割网络处理第六中间特征图,得到样本图像的预测分割图。将样本图像的深度标签、预测深度图、分割标签和预测分割图输入全卷积神经网络模型的损失函数,输出损失结果。根据损失结果调整全卷积神经网络模型的网络参数,直至损失函数收敛。将训练后的全卷积神经网络模型作为图像处理模型。

[0085] 根据本公开的实施例,可以结合图3进行理解,在图像处理模型的训练过程中,第五中间特征图可以理解为是图3中的特征图 f_4 ,第六中间特征图可以理解为是图3中的特征图 f_5 ,第七中间特征图可以理解为是图3中的特征图 f_6 ,第八中间特征图可以理解为是图3中的特征图 f_7 。需要说明的是,在训练完成后,初始特征提取网络、初始图像分割网络和初始深度估计网络分别称为特征提取网络、图像分割网络和深度估计网络。

[0086] 图11示意性示出了根据本公开实施例的再一种图像处理方法的流程图。

[0087] 如图11所示,对样本图像进行图像分割处理,得到样本图像的分割标签可以包括如下操作S1110~S1130。

[0088] 在操作S1110,对样本图像进行实例分割处理,得到样本图像的实例分割标签。

[0089] 在操作S1120,根据样本图像的实例分割标签,得到样本图像的语义分割标签。

[0090] 在操作S1130,将样本图像的语义分割标签作为样本图像的分割标签。

[0091] 根据本公开的实施例,实例分割为图像中可能包括属于同一类别的多个实例,需要对其进行区分。例如,对于一目标图像,该目标图像中可能包括属于人这一类别的数量为多个,即包括多个人,在实例分割中,需要对这多个人进行区分,每个人都可以得到相应的实例分割标签。

[0092] 根据本公开的实施例,语义分割为对图像中的每个像素点进行类别划分,但不区分实例。例如,对于一目标图像,该目标图像中可能包括属于人这一类别的数量为多个,即包括多个人,在语义分割中,不需要对这多个人进行区分,多个人得到相同的语义分割标签。

[0093] 根据本公开的实施例,为了制作更精确的语义分割标签,可以使用Mask_RCNN (Recurrent Convolutional Neural Network,循环卷积神经网络)网络在三个深度估计的样本图像数据库CAD_60、CAD_120和EPFL上输出实例分割标签,然后将实例分割标签转换为语义分割标签,将其作为三个深度估计的样本图像数据库上的语义分割标签。但是并不局限于此,还可以单独使用语义分割标签。由于本公开实施例采用的是深度估计的样本图像数据库,因此,可以获得样本图像的深度标签。

[0094] 根据本公开的其他实施例,Mask_RCNN可以检测并分割出82种类别(含背景类)的对象,在实际应用中,深度估计的样本图像数据库中出现的类别可能少于82种,当进行分割处理时,直接使用82种类别作为分割标签,将扩大分割范围,导致增加分割处理的出错概率。

[0095] 根据本公开的实施例,使用的样本图像数据库中出现了其中的59种类别,为了构造稠密的分割标签。如表1所示,可以将Mask_RCNN的类别进行一个映射,将不涉及的类别标记为-1,以满足在对图像分割处理的基础上,降低出错概率,提高分割效果和准确度。

[0096] 表1

[0097]	原类别		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	现类别		0	1	-1	2	-1	-1	-1	-1	-1	-1	-1	3	-1	-1	4	5	6	7	-1
	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
	-1	-1	-1	-1	-1	-1	8	9	10	11	12	13	14	-1	15	16	17	18	19	-1	20
	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
	21	22	23	-1	24	25	26	27	28	29	30	31	-1	32	33	34	35	36	37	38	39
	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81
	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	-1	56	57	58	-1

[0098] 根据本公开的实施例,使用Mask_RCNN网络得到的实例分割标签的结果精度高,有利于图像处理模型的构建。

[0099] 根据本公开的实施例,也可以直接在具有深度标签的样本图像数据库上进行语义分割,得到样本图像的语义分割标签。此外,除了采用在具有深度标签的样本图像数据库的基础上,对样本图像进行图像分割处理,得到样本图像的分割标签,使得样本图像同时具有深度标签和分割标签的方式外。还可以采用已知样本图像的分割标签,对样本图像进行深度估计,得到样本图像的深度标签,使得样本图像同时具有深度标签和分割标签的方式。但是,采用第一种方式得到的目标图像中的目标对象的预测深度图的精确度更高。

[0100] 下面结合具体实施例对本公开的技术方案作进一步说明,图像处理方法的操作可以具体如下。

[0101] 图12示意性示出了根据本公开实施例的又一种图像处理方法的流程图。

[0102] 根据本公开的实施例,如图12所示,图像处理方法的操作可以包括。

[0103] 制作数据集,构建全卷积神经网络模型,训练得到图像处理模型。

[0104] 归一化输入图像,得到RGB图像,作为图像处理模型的输入,分别得到目标图像的预测深度图和预测分割图。

[0105] 判断目标图像的预测分割图中对应的预测分割图是否是目标对象的,若是,则将目标对象的预测分割图作用到目标图像的预测深度图上得到目标对象的预测深度图。

[0106] 此外,根据本公开的实施例,还可以判断是否遍历所有目标对象的预测分割图。若否,遍历下一个类别的目标对象的预测分割图;若是,输出所有目标对象的预测深度图。

[0107] 根据本公开的实施例,可以将图像处理方法应用于单目机器人寻找或躲避特定物体的应用场景中,其中,特征物体即为目标对象。

[0108] 图13示意性示出了根据本公开的实施例的图像处理装置的框图。

[0109] 如图13所示,图像处理装置1300包括获取模块1310、第一处理模块1320、确定模块1330和第二处理模块1340。

[0110] 获取模块1310、第一处理模块1320、确定模块1330和第二处理模块1340通信连接。

[0111] 获取模块1310,用于获取目标图像,其中,目标图像包括目标对象和非目标对象。

[0112] 第一处理模块1320,用于对目标图像进行图像分割处理和深度估计处理,分别得到目标图像的预测分割图和预测深度图。

[0113] 确定模块1330,用于根据目标对象的预测分割图确定目标图像的预测深度图中目标对象的位置。

[0114] 第二处理模块1340,用于根据目标对象的预测深度图中目标对象的位置对预测深度图进行处理,得到目标对象的预测深度图。

[0115] 根据本公开实施例的技术方案,通过获取目标图像,目标图像包括目标对象和非目标对象,对目标图像进行图像分割处理和深度估计处理,分别得到目标图像的预测分割图和预测深度图,根据目标对象的预测分割图确定目标图像的预测深度图中目标对象的位置,并根据目标图像的预测深度图中目标对象的位置对预测深度图进行处理,得到目标对象的预测深度图。由于将图像分割和深度估计进行结合,其中,根据预测分割图能够获得预测深度图中目标对象的位置,根据预测深度图中目标对象的位置对预测深度图进行处理能够获得目标对象的预测深度图,因此,至少部分地克服了相关技术中难以实现针对目标图像中目标对象的深度估计的技术问题,进而实现了较为准确地确定目标图像中目标对象的深度,并且方法的泛化性较强。

[0116] 根据本公开的实施例,第一处理模块1320包括第一处理单元。

[0117] 第一处理单元,用于利用图像处理模型处理目标图像,分别得到分割图和深度图,其中,图像处理模型是利用训练样本训练得到的,其中,训练样本包括样本图像和样本图像的深度标签和分割标签。

[0118] 根据本公开的实施例,图像处理模型包括特征提取网络、图像分割网络和深度估计网络。

[0119] 根据本公开的实施例,第一处理单元包括第一处理子单元、第二处理子单元、第三处理子单元、第四处理子单元、第五处理子单元、和第六处理子单元。

[0120] 第一处理子单元,用于利用特征提取网络处理目标图像,得到第一中间特征图。

[0121] 第二处理子单元,用于利用图像分割网络处理第一中间特征图,得到第二中间特征图。

[0122] 第三处理子单元,用于利用深度估计网络处理第一中间特征图,得到第三中间特征图。

[0123] 第四处理子单元,用于根据第二中间特征图和第三中间特征图,生成第四中间特征图。

[0124] 第五处理子单元,用于利用深度估计网络处理第四中间特征图,得到目标图像的预测深度图。

[0125] 第六处理子单元,用于利用图像分割网络处理第二中间特征图,得到目标图像的预测分割图。

[0126] 根据本公开的实施例,图像处理模型是利用训练样本训练得到的,可以包括如下操作。

[0127] 获取训练样本。利用训练样本训练全卷积神经网络模型,得到图像处理模型。

[0128] 根据本公开的实施例,全卷积神经网络模型包括初始特征提取网络、初始图像分割网络和初始深度估计网络。

[0129] 利用训练样本训练全卷积神经网络模型,得到图像处理模型,可以包括如下操作。

[0130] 利用初始特征提取网络处理样本图像,得到第五中间特征图。利用初始图像分割网络处理第五中间特征图,得到第六中间特征图。利用初始深度估计网络处理第五中间特征图,得到第七中间特征图。根据第六中间特征图和第七中间特征图,生成第八中间特征图。利用初始深度估计网络处理第八中间特征图,得到样本图像的预测深度图。利用初始图像分割网络处理第六中间特征图,得到样本图像的预测分割图。将样本图像的深度标签、预测深度图、分割标签和预测分割图输入全卷积神经网络模型的损失函数,得到损失结果。根据损失结果调整全卷积神经网络模型的网络参数,直至损失函数收敛。将训练后的全卷积神经网络模型作为图像处理模型。

[0131] 根据本公开的实施例,对样本图像进行图像分割处理,得到样本图像的分割标签,可以包括如下操作。

[0132] 对样本图像进行实例分割处理,得到样本图像的实例分割标签。根据样本图像的实例分割标签,得到样本图像的语义分割标签。将样本图像的语义分割标签作为样本图像的分割标签。

[0133] 根据本公开的实施例,第二处理模块1320包括第二处理单元。

[0134] 第二处理单元,用于将目标图像的预测深度图中除目标图像的位置以外的其他位置的像素值设置为预设像素值,得到目标对象的预测深度图。

[0135] 根据本公开的实施例,图像分割处理包括语义分割处理或实例分割处理。

[0136] 根据本公开的实施例的模块、单元中的任意多个、或其中任意多个的至少部分功能可以在一个模块中实现。根据本公开实施例的模块、单元中的任意一个或多个可以被拆分成多个模块来实现。根据本公开实施例的模块、单元中的任意一个或多个可以至少被部

分地实现为硬件电路,例如现场可编程门阵列(Field Programmable Gate Array,FPGA)、可编程逻辑阵列(Programmable Logic Arrays,PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(Application Specific Integrated Circuit,ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式的硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,根据本公开实施例的模块、单元中的一个或多个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0137] 例如,获取模块1310、第一处理模块1320、确定模块1330和第二处理模块1340中的任意多个可以合并在一个模块/单元中实现,或者其中的任意一个模块/单元可以被拆分成多个模块/单元。或者,这些模块/单元中的一个或多个模块/单元的至少部分功能可以与其他模块/单元的至少部分功能相结合,并在一个模块/单元中实现。根据本公开的实施例,获取模块1310、第一处理模块1320、确定模块1330和第二处理模块1340中的至少一个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式等硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,获取模块1310、第一处理模块1320、确定模块1330和第二处理模块1340中的至少一个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0138] 需要说明的是,本公开的实施例中图像处理装置部分与本公开的实施例中图像处理方法是相对应的,图像处理装置部分的描述具体参考图像处理方法部分,在此不再赘述。

[0139] 图14示意性示出了根据本公开实施例的适于实现上文描述的方法的电子设备的框图。图14示出的电子设备仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0140] 如图14所示,根据本公开实施例的电子设备1400包括处理器1401,其可以根据存储在只读存储器(Read-Only Memory,ROM) 1402中的程序或者从存储部分1408加载到随机访问存储器(Random Access Memory,RAM) 1403中的程序而执行各种适当的动作和处理。处理器1401例如可以包括通用微处理器(例如CPU)、指令集处理器和/或相关芯片组和/或专用微处理器(例如,专用集成电路(ASIC)),等等。处理器1401还可以包括用于缓存用途的板载存储器。处理器1401可以包括用于执行根据本公开实施例的方法流程的不同动作的单一处理单元或者是多个处理单元。

[0141] 在RAM 1403中,存储有电子设备1400操作所需的各种程序和数据。处理器1401、ROM 1402以及RAM 1403通过总线1404彼此相连。处理器1401通过执行ROM 1402和/或RAM 1403中的程序来执行根据本公开实施例的方法流程的各种操作。需要注意,所述程序也可以存储在除ROM 1402和RAM 1403以外的一个或多个存储器中。处理器1401也可以通过执行存储在所述一个或多个存储器中的程序来执行根据本公开实施例的方法流程的各种操作。

[0142] 根据本公开的实施例,电子设备1400还可以包括输入/输出(I/O)接口1405,输入/输出(I/O)接口1405也连接至总线1404。系统1400还可以包括连接至I/O接口1405的以下部件中的一项或多项:包括键盘、鼠标等的输入部分1406;包括诸如阴极射线管(CRT)、液晶显

示器(Liquid Crystal Display,LCD)等以及扬声器等的输出部分1407;包括硬盘等的存储部分1408;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分1409。通信部分1409经由诸如因特网的网络执行通信处理。驱动器1410也根据需要连接至I/O接口1405。可拆卸介质1411,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器1410上,以便于从其上读出的计算机程序根据需要被安装入存储部分1408。

[0143] 根据本公开的实施例,根据本公开实施例的方法流程可以被实现为计算机软件程序。例如,本公开的实施例包括一种计算机程序产品,其包括承载在计算机可读存储介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分1409从网络上被下载和安装,和/或从可拆卸介质1411被安装。在该计算机程序被处理器1401执行时,执行本公开实施例的系统中限定的上述功能。根据本公开的实施例,上文描述的系统、设备、装置、模块、单元等可以通过计算机程序模块来实现。

[0144] 本公开还提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例中描述的设备/装置/系统中所包含的;也可以是单独存在,而未装配入该设备/装置/系统中。上述计算机可读存储介质承载有一个或者多个程序,当上述一个或者多个程序被执行时,实现根据本公开实施例的方法。

[0145] 根据本公开的实施例,计算机可读存储介质可以是非易失性的计算机可读存储介质。例如可以包括但不限于:便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(Erasable Programmable Read Only Memory,EPROM)或闪存)、便携式紧凑磁盘只读存储器(Computer Disc Read-Only Memory,CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0146] 例如,根据本公开的实施例,计算机可读存储介质可以包括上文描述的ROM 1402和/或RAM 1403和/或ROM 1402和RAM 1403以外的一个或多个存储器。

[0147] 根据本公开实施例的技术方案,通过获取目标图像,目标图像包括目标对象和非目标对象,对目标图像进行图像分割处理和深度估计处理,分别得到目标图像的预测分割图和预测深度图,根据目标对象的预测分割图确定目标图像的预测深度图中目标对象的位置,并根据目标图像的预测深度图中目标对象的位置对预测深度图进行处理,得到目标对象的预测深度图。由于将图像分割和深度估计进行结合,其中,根据预测分割图能够获得预测深度图中目标对象的位置,根据预测深度图中目标对象的位置对预测深度图进行处理能够获得目标对象的预测深度图,因此,至少部分地克服了相关技术中难以实现针对目标图像中目标对象的深度估计的技术问题,进而实现了较为准确地确定目标图像中目标对象的深度,并且方法的泛化性较强。

[0148] 本公开的实施例还包括一种计算机程序产品,其包括计算机程序,该计算机程序包含用于执行本公开实施例所提供的方法的程序代码,当计算机程序产品在电子设备上运行时,该程序代码用于使电子设备实现本公开实施例所提供的图像处理方法。

[0149] 在该计算机程序被处理器1401执行时,执行本公开实施例的系统/装置中限定的上述功能。根据本公开的实施例,上文描述的系统、装置、模块、单元等可以通过计算机程序

模块来实现。

[0150] 在一种实施例中,该计算机程序可以依托于光存储器件、磁存储器件等有形存储介质。在另一种实施例中,该计算机程序也可以在网络介质上以信号的形式进行传输、分发,并通过通信部分1409被下载和安装,和/或从可拆卸介质1411被安装。该计算机程序包含的程序代码可以用任何适当的网络介质传输,包括但不限于:无线、有线等等,或者上述的任意合适的组合。

[0151] 根据本公开的实施例,可以以一种或多种程序设计语言的任意组合来编写用于执行本公开实施例提供的计算机程序的程序代码,具体地,可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。程序设计语言包括但不限于诸如Java,C++,python,“C”语言或类似的设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中,远程计算设备可以通过任意种类的网络,包括局域网(Local Area Network,LAN)或广域网(Wide Area Networks,WAN),连接到用户计算设备,或者,可以连接到外部计算设备(例如利用因特网服务提供商来通过因特网连接)。

[0152] 附图中的流程图和框图,图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图或流程图中的每个方框、以及框图或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。本领域技术人员可以理解,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合,即使这样的组合或结合没有明确记载于本公开中。特别地,在不脱离本公开精神和教导的情况下,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合。所有这些组合和/或结合均落入本公开的范围。

[0153] 以上对本公开的实施例进行了描述。但是,这些实施例仅仅是为了说明的目的,而并非为了限制本公开的范围。尽管在以上分别描述了各实施例,但是这并不意味着各个实施例中的措施不能有利地结合使用。本公开的范围由所附权利要求及其等同物限定。不脱离本公开的范围,本领域技术人员可以做出多种替代和修改,这些替代和修改都应落在本公开的范围之内。

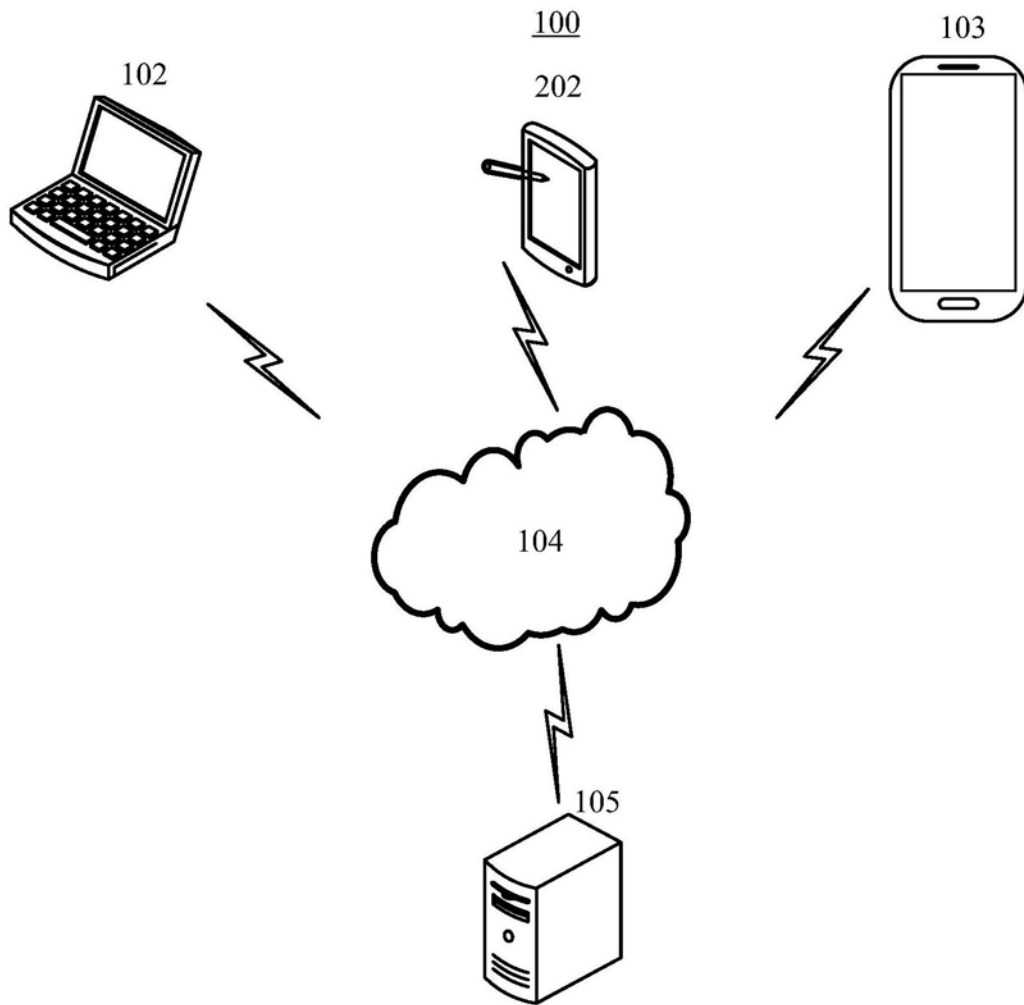


图1

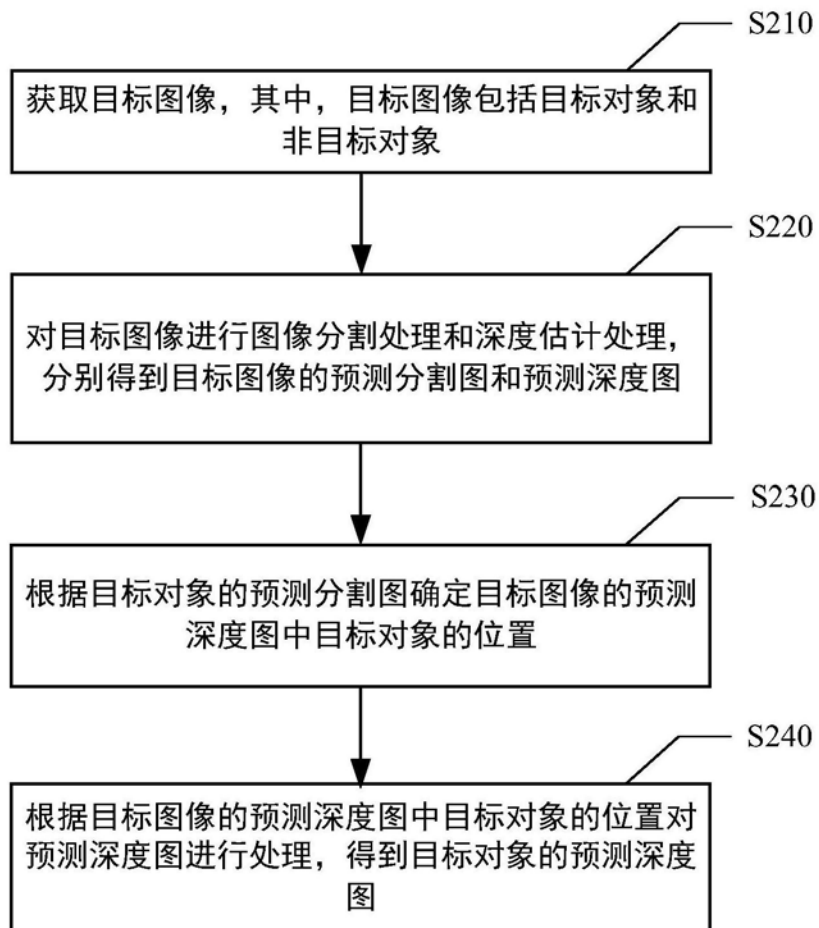


图2

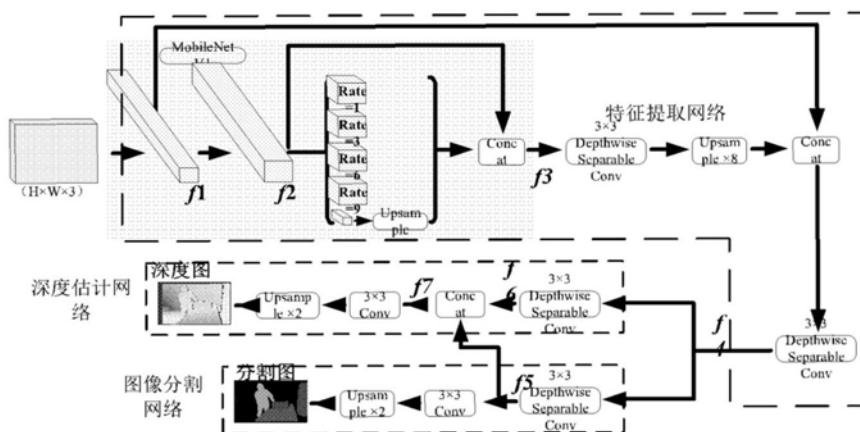


图3

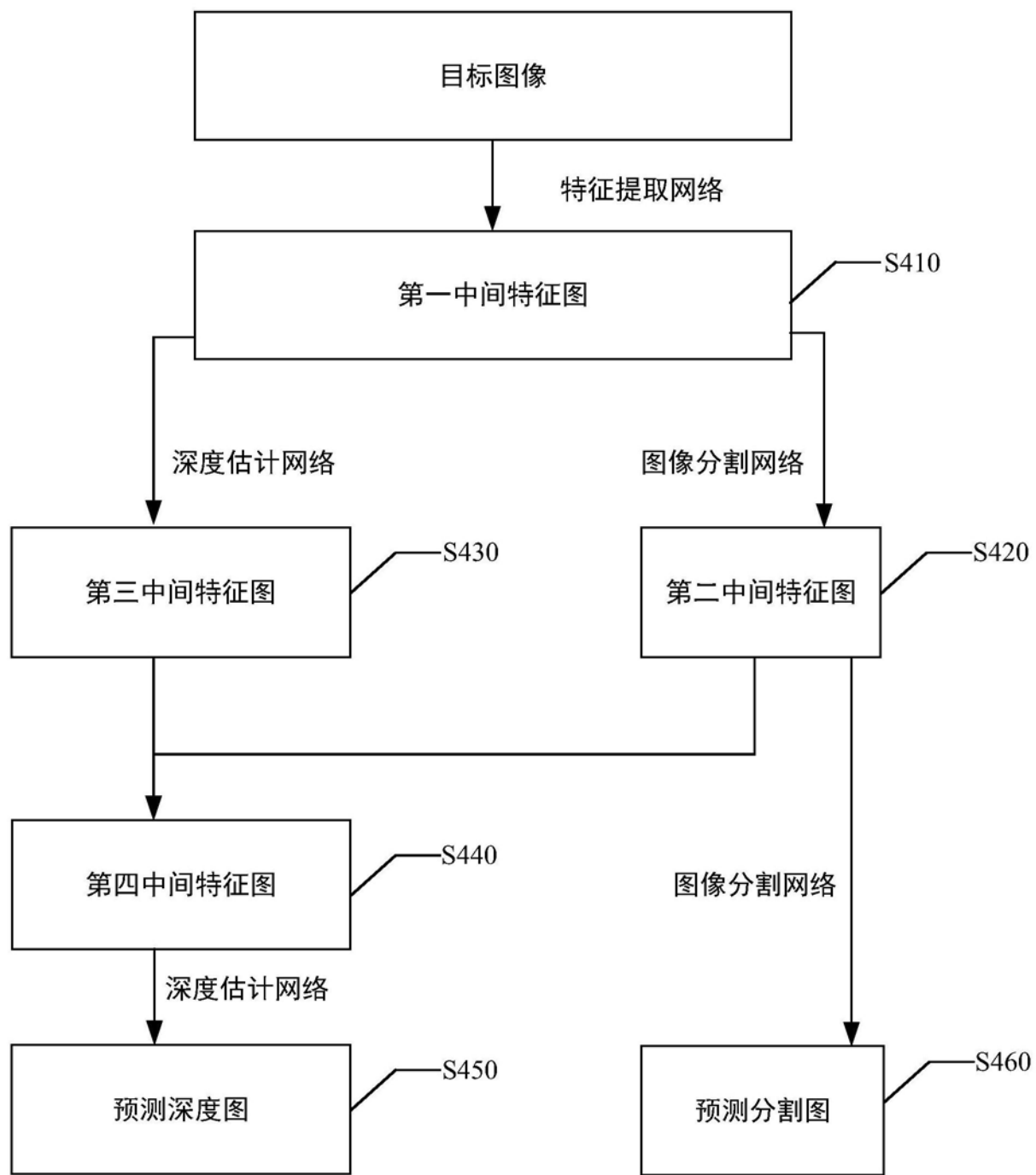


图4



图5



图6



图7

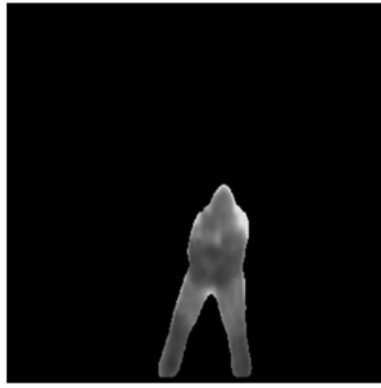


图8



图9



图10

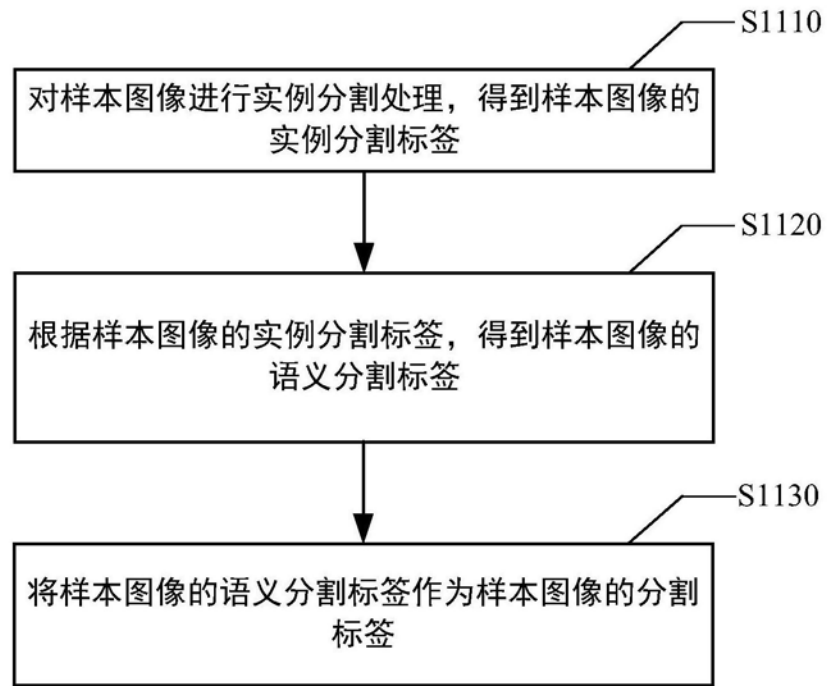


图11

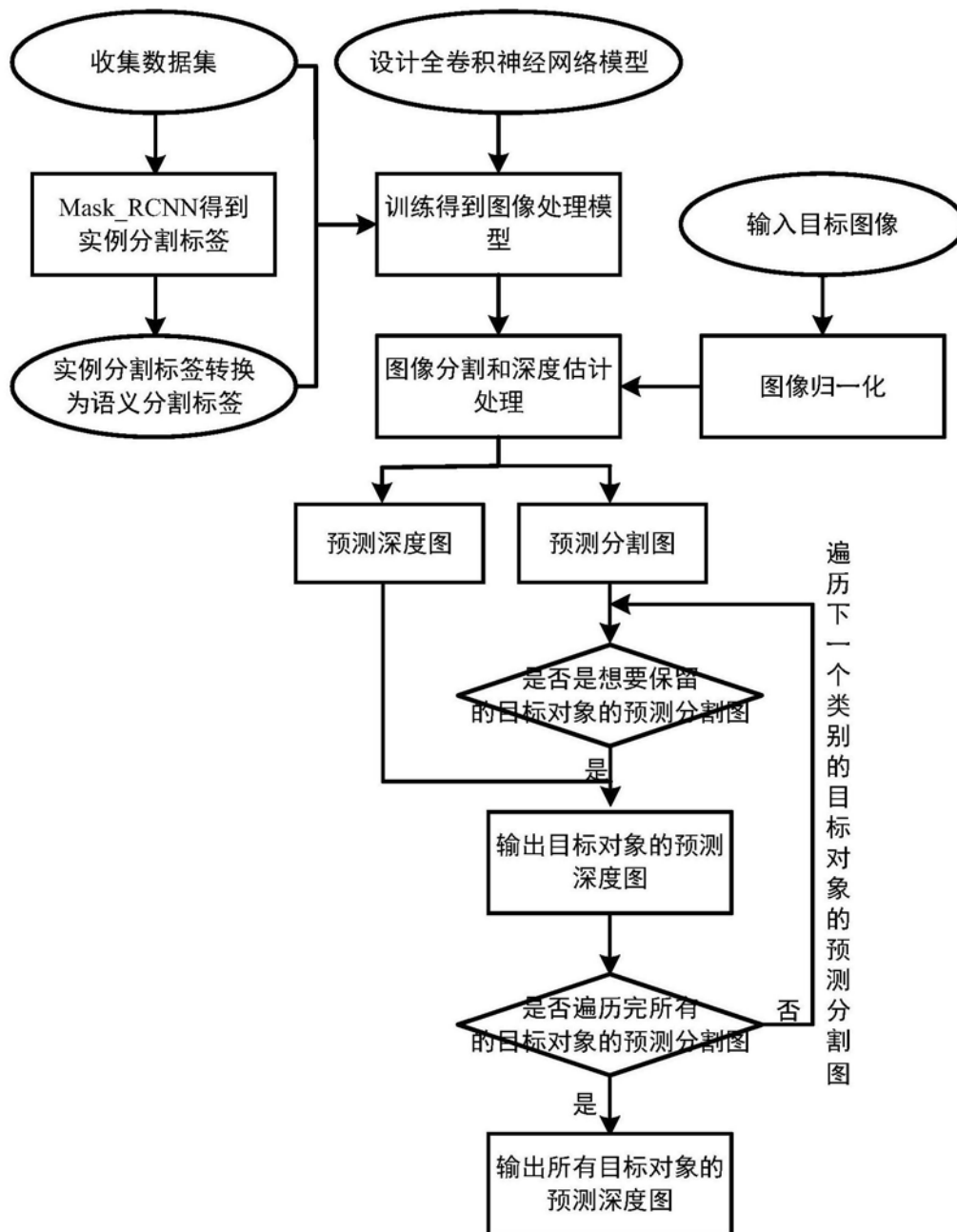


图12

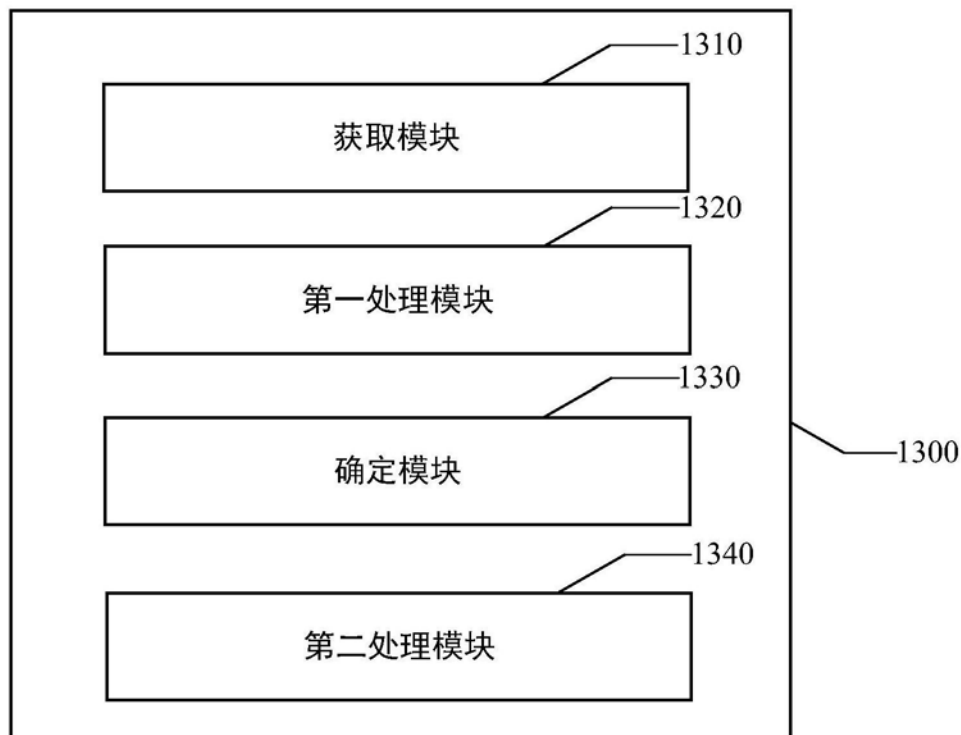


图13

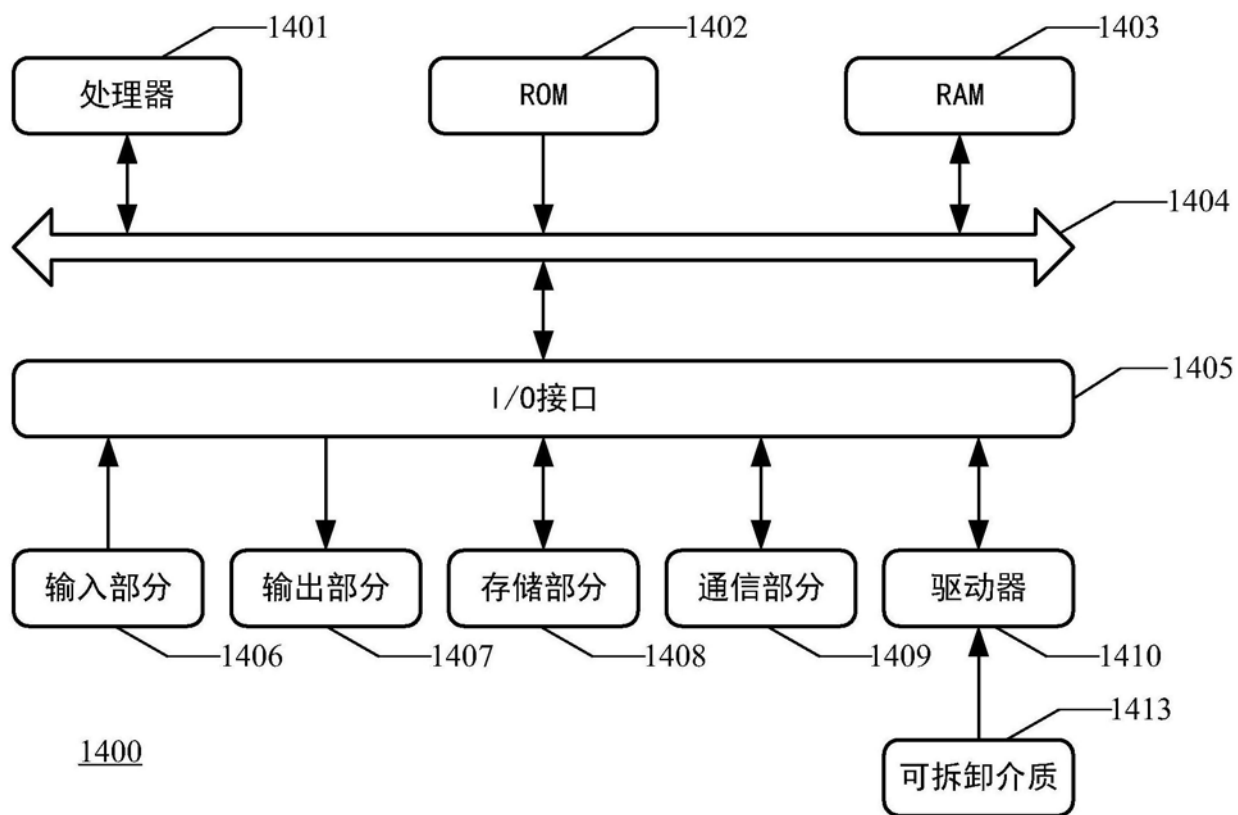


图14