



(12) 发明专利申请

(10) 申请公布号 CN 112200035 A

(43) 申请公布日 2021.01.08

(21) 申请号 202011045690.4

G06T 7/10 (2017.01)

(22) 申请日 2020.09.29

(71) 申请人 深圳市优必选科技股份有限公司

地址 518000 广东省深圳市南山区学苑大道1001号南山智园C1栋16、22楼

(72) 发明人 王阳 赵明国

(74) 专利代理机构 北京超凡宏宇专利代理事务所(特殊普通合伙) 11463

代理人 梁韬

(51) Int. Cl.

G06K 9/00 (2006.01)

G06T 7/246 (2017.01)

G06T 5/00 (2006.01)

G06T 5/50 (2006.01)

G06T 3/40 (2006.01)

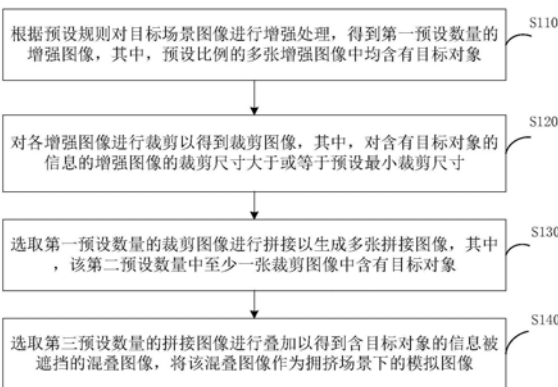
权利要求书2页 说明书13页 附图4页

(54) 发明名称

用于模拟拥挤场景的图像获取方法、装置和视觉处理方法

(57) 摘要

本申请实施例提供一种用于模拟拥挤场景的图像获取方法、装置和视觉处理方法,该方法包括:根据预设规则对目标场景图像进行增强处理,然后按照相应的裁剪、拼接及叠加规则对各所述增强图像依次进行裁剪、拼接及叠加处理,从而得到含目标对象的信息被遮挡的混叠图像以作为用于模拟拥挤场景的模拟图像。该技术方案不仅很好地解决了各种视觉运用场景下的模型训练时实际样本图像的数量有限的问题,尤其是针对目标对象的信息被遮挡的拥挤场景,而且通过该方法所获取的模拟图像还可以兼容全局场景图像的上下文信息及场景图像中目标对象的局部信息,从而可较好地保证模型训练质量等。



1. 一种用于模拟拥挤场景的图像获取方法,其特征在于,所述拥挤场景为存在目标对象的信息被遮挡的场景,所述方法包括:

根据预设规则对目标场景图像进行增强处理,得到第一预设数量的增强图像,其中,预设比例的多张所述增强图像中均含有所述目标对象;

对各所述增强图像进行裁剪以得到裁剪图像,其中,含有所述目标对象的增强图像的裁剪尺寸大于或等于预设最小裁剪尺寸;

选取第二预设数量的所述裁剪图像进行拼接以生成多张拼接图像,其中,所述第二预设数量中至少一张裁剪图像中含有所述目标对象;

选取第三预设数量的所述拼接图像进行叠加以得到含所述目标对象的信息被遮挡的混叠图像,将所述混叠图像作为所述拥挤场景下的模拟图像。

2. 根据权利要求1所述的方法,其特征在于,所述预设规则包括基于所述目标对象的特征属性设定的图像调整类型和/或调整尺寸,以及设定的噪声模拟规则,所述根据预设规则对目标场景图像进行增强,得到第一预设数量的增强图像包括:

根据所述图像调整类型和/或调整尺寸对所述目标场景图像进行调整处理,得到第一预设数量的调整图像;

根据所述噪声模拟规则对各所述调整图像进行噪声模拟处理,得到所述第一预设数量的增强图像。

3. 根据权利要求2所述的方法,其特征在于,所述图像调整类型包括图像的翻转、裁剪、旋转、缩放和亮度调节中的至少一种;

其中,所述翻转包括进行水平或垂直翻转,所述裁剪包括进行部分区域图像截取,所述旋转包括进行角度旋转,所述缩放包括进行尺寸比例缩放,所述亮度调节包括将图像由RGB空间转变为HSV空间并对所述HSV空间中的亮度分量进行大小调节。

4. 根据权利要求1所述的方法,其特征在于,所述第一预设数量的增强图像中还包括不含所述目标对象的背景图像,所述背景图像经过裁剪后得到对应的背景裁剪图像;所述选取第二预设数量的所述裁剪图像进行拼接以生成多张拼接图像包括:

选取第四预设数量的含有所述目标对象的裁剪图像和第五预设数量的所述背景裁剪图像,并按照所述目标场景图像的原始尺寸进行拼接以得到所述拼接图像,其中,所述第四预设数量和所述第五预设数量之和等于第二预设数量。

5. 根据权利要求1至4任一项所述的方法,其特征在于,所述选取第三预设数量的所述拼接图像进行叠加以得到含所述目标对象的局部信息被遮挡的混叠图像包括:

为每张待叠加的所述拼接图像分配各自对应的预设加权值,其中,每一所述预设加权值的取值范围为(0,1);

基于各所述拼接图像的所述预设加权值并按照预设混叠公式进行线性叠加以得到混叠图像,所述混叠图像中的所述目标对象的局部信息被遮挡。

6. 一种图像获取方法,其特征在于,包括:

对目标场景图像进行一次增强,得到多张第一图像;其中,所述一次增强包括对所述目标场景图像进行预设增强类型的调整处理及对调整后的图像进行噪声模拟处理;

对各所述第一图像进行二次增强,得到多张第二图像,将所述第二图像作为相应场景的模拟图像;其中,所述二次增强包括对所述多张第一图像依次进行图像裁剪、图像拼接和

图像混叠处理。

7. 一种视觉处理方法, 其特征在于, 应用于包括拍摄装置的终端设备, 所述方法包括:

通过拍摄装置采集目标场景图像, 并基于所述目标场景图像采用如权利要求1至5中任一项所述的方法生成模拟图像; 其中, 所述预设最小裁剪尺寸根据相应场景下目标对象在所述终端设备的拍摄视野内占有的像素范围设定;

利用所述目标场景图像和所述模拟图像进行预设场景模型训练以得到训练后的场景模型, 所述训练后的场景模型用于对所述终端设备在相应场景下拍摄得到的待处理场景图像进行与目标对象相关的视觉处理操作。

8. 根据权利要求7所述的方法, 其特征在于, 还包括:

生成所述模拟图像之前, 对各所述目标场景图像分别进行目标对象标注以获取对应的标注信息;

生成所述模拟图像之后, 根据所述目标场景图像的所述标注信息以及所述增强、所述裁剪、所述拼接和所述叠加时各自对应的处理参数计算所述模拟图像的标注信息, 以及利用带标注信息的所述目标场景图像和所述模拟图像对所述预设场景模型进行训练。

9. 根据权利要求7或8所述的方法, 其特征在于, 所述终端设备为机器人, 所述预设场景模型为应用于所述机器人中的物体跟踪模型、物体检测模型或物体分割模型。

10. 一种用于模拟拥挤场景的图像获取装置, 其特征在于, 所述拥挤场景为存在目标对象的信息被遮挡的场景, 所述装置包括:

增强模块, 用于根据预设规则对目标场景图像进行增强处理, 得到第一预设数量的增强图像, 其中, 预设比例的多张所述增强图像中均含有所述目标对象;

裁剪模块, 用于对各所述增强图像进行裁剪以得到裁剪图像, 其中, 含有所述目标对象的增强图像的裁剪尺寸大于或等于预设最小裁剪尺寸;

拼接模块, 用于选取第二预设数量的所述裁剪图像进行拼接以生成多张拼接图像, 其中, 所述第二预设数量中至少一张裁剪图像中含有所述目标对象;

叠加模块, 用于选取第三预设数量的所述拼接图像进行叠加以得到含有所述目标对象的局部信息被遮挡的混叠图像, 将所述混叠图像作为所述拥挤场景下的模拟图像。

11. 一种终端设备, 其特征在于, 所述终端设备包括处理器和存储器, 所述存储器存储有计算机程序, 所述处理器用于执行所述计算机程序以实施权利要求1至9中任一项所述的方法。

12. 一种可读存储介质, 其特征在于, 其存储有计算机程序, 所述计算机程序被执行时实施权利要求1至9中任一项所述的方法。

用于模拟拥挤场景的图像获取方法、装置和视觉处理方法

技术领域

[0001] 本申请涉及图像处理技术领域，尤其涉及一种用于模拟拥挤场景的图像获取方法、装置和视觉处理方法。

背景技术

[0002] 视觉信息是人类最直观的输入信息之一。随着智能手机、短视频应用等技术的蓬勃发展，视觉图像数据的规模不断增长。深度学习技术尤其是卷积神经网络技术，随着视觉图像数据集规模的不断增长在不断改进迭代，但是复杂的卷积神经网络模型会带来较高的网络过拟合风险：模型过分拟合了数据集中的特征，导致模型在迁移到真实场景中表现反而变差。因此，视觉图像数据的质量是直接影响计算机视觉模型最终表现的重要因素之一。

[0003] 以机器人应用为例，在实际的机器人的网络模型训练时，常常会遇到数据不足或者数据分布高度相似的情况，导致该机器人对一些特定的场景则表现较差。例如，以一机器人跟踪足球为例，尤其是当出现足球的局部被运动人员等遮挡的情况，则机器人往往容易出现漏检或误检的现象。因此，需要利用相应的技术手段以丰富收集到的训练数据的多样性，尤其是物体信息被遮挡的训练样本，进而提升机器人的网络模型的泛化性能与鲁棒性。

发明内容

[0004] 有鉴于此，本申请的目的是为了克服现有技术中的不足，提供一种用于模拟拥挤场景的图像获取方法、装置、图像获取方法、视觉处理方法、终端设备及可读存储介质。

[0005] 本申请实施例提供一种用于模拟拥挤场景的图像获取方法，所述拥挤场景为存在目标对象的信息被遮挡的场景，所述方法包括：

[0006] 根据预设规则对目标场景图像进行增强处理，得到第一预设数量的增强图像，其中，预设比例的多张所述增强图像中均含有所述目标对象；

[0007] 对各所述增强图像进行裁剪以得到裁剪图像，其中，含有所述目标对象的增强图像的裁剪尺寸大于或等于预设最小裁剪尺寸；

[0008] 选取第二预设数量的所述裁剪图像进行拼接以生成多张拼接图像，其中，所述第二预设数量中至少一张裁剪图像中含有所述目标对象；

[0009] 选取第三预设数量的所述拼接图像进行叠加以得到含所述目标对象的信息被遮挡的混叠图像，将所述混叠图像作为所述拥挤场景下的模拟图像。

[0010] 在一种实施例中，所述预设规则包括基于所述目标对象的特征属性设定的图像调整类型和/或调整尺寸，以及设定的噪声模拟规则，所述根据预设规则对目标场景图像进行增强，得到第一预设数量的增强图像包括：

[0011] 根据所述图像调整类型和/或调整尺寸对所述目标场景图像进行调整处理，得到第一预设数量的调整图像；

[0012] 根据所述噪声模拟规则对各所述调整图像进行噪声模拟处理，得到所述第一预设数量的增强图像。

[0013] 在上述实施例中,所述图像调整类型包括图像的翻转、裁剪、旋转、缩放和亮度调节中的至少一种;

[0014] 其中,所述翻转包括进行水平或垂直翻转,所述裁剪包括进行部分区域图像截取,所述旋转包括进行角度旋转,所述缩放包括进行尺寸比例缩放,所述亮度调节包括将图像由RGB空间转变为HSV空间并对所述HSV空间中的亮度分量进行大小调节。

[0015] 在一种实施例中,所述第一预设数量的增强图像中还包括不含所述目标对象的背景图像,所述背景图像经过裁剪后得到对应的背景裁剪图像;所述选取第二预设数量的所述裁剪图像进行拼接以生成多张拼接图像包括:

[0016] 选取第四预设数量的含有所述目标对象的裁剪图像和第五预设数量的所述背景裁剪图像,并按照所述目标场景图像的原始尺寸进行拼接以得到所述拼接图像,其中,所述第四预设数量和所述第五预设数量之和等于第二预设数量。

[0017] 在一种实施例中,所述选取第三预设数量的所述拼接图像进行叠加以得到含所述目标对象的局部信息被遮挡的混叠图像包括:

[0018] 为每张待叠加的所述拼接图像分配各自对应的预设加权值,其中,每一所述预设加权值的取值范围为(0,1);

[0019] 基于各所述拼接图像的所述预设加权值并按照预设混叠公式进行线性叠加以得到混叠图像,所述混叠图像中的所述目标对象的局部信息被遮挡。

[0020] 本申请实施例还提供一种图像获取方法,包括:

[0021] 对目标场景图像进行一次增强,得到多张第一图像;其中,所述一次增强包括对所述目标场景图像进行预设增强类型的调整处理及对调整后的图像进行噪声模拟处理;

[0022] 对各所述第一图像进行二次增强,得到多张第二图像,将所述第二图像作为相应场景的模拟图像;其中,所述二次增强包括对所述多张第一图像依次进行图像裁剪、图像拼接和图像混叠处理。

[0023] 在一种实施例中,所述目标场景图像为同一场景下的具有相同类型和/或不同类型的目标对象的图像,或者为不同场景下的具有相同类型的目标对象的图像。

[0024] 本申请实施例还提供一种视觉处理方法,应用于包括拍摄装置的终端设备,该方法包括:

[0025] 通过拍摄装置采集目标场景图像,并基于所述目标场景图像采用上述的方法生成模拟图像;其中,所述预设最小裁剪尺寸根据相应场景下目标对象在所述终端设备的拍摄视野内占有的像素设定;

[0026] 利用所述目标场景图像和所述模拟图像进行预设场景模型训练以得到训练后的场景模型,所述训练后的场景模型用于对所述终端设备在相应场景下拍摄得到的待处理场景图像进行与目标对象相关的视觉处理操作。

[0027] 在一种实施例中,生成所述模拟图像之前,对各所述目标场景图像分别进行目标对象标注以获取对应的标注信息;

[0028] 生成所述模拟图像之后,根据所述目标场景图像的所述标注信息以及所述增强、所述裁剪、所述拼接和所述叠加时各自对应的处理参数计算所述模拟图像的标注信息,以及利用带标注信息的所述目标场景图像和所述模拟图像对所述预设场景模型进行训练。

[0029] 在一种实施例中,所述终端设备为机器人,所述预设场景模型为应用于所述机器

人中的物体跟踪模型、物体检测模型或物体分割模型。

[0030] 本申请的实施例还提供一种用于模拟拥挤场景的图像获取装置,所述拥挤场景为存在目标对象的信息被遮挡的场景,所述装置包括:

[0031] 增强模块,用于根据预设规则对目标场景图像进行增强处理,得到第一预设数量的增强图像,其中,预设比例的多张所述增强图像中均含有所述目标对象;

[0032] 裁剪模块,用于对各所述增强图像进行裁剪以得到裁剪图像,其中,含有所述目标对象的增强图像的裁剪尺寸大于或等于预设最小裁剪尺寸;

[0033] 拼接模块,用于选取第二预设数量的所述裁剪图像进行拼接以生成多张拼接图像,其中,所述第二预设数量中至少一张裁剪图像中含有所述目标对象;

[0034] 叠加模块,用于选取第三预设数量的所述拼接图像进行叠加以得到含所述目标对象的局部信息被遮挡的混叠图像,将所述混叠图像作为所述拥挤场景下的模拟图像。

[0035] 在一种实施例中,所述预设规则包括基于所述目标对象的特征属性设定的图像调整类型和/或调整尺寸,以及设定的噪声模拟规则,所述增强模块包括调整处理子模块和噪声模拟子模块;

[0036] 所述调整处理子模块用于根据所述图像调整类型和/或调整尺寸对所述目标场景图像进行调整处理,得到第一预设数量的调整图像;

[0037] 所述噪声模拟子模块用于根据所述噪声模拟规则对各所述调整图像进行噪声模拟处理,得到所述第一预设数量的增强图像。

[0038] 在一种实施例中,所述多张增强图像中还包括不含所述目标对象的背景图像,所述背景图像经过裁剪后得到对应的背景裁剪图像;

[0039] 所述拼接模块还用于选取第四预设数量的含有所述目标对象的裁剪图像和第五预设数量的所述背景裁剪图像,并按照所述目标场景图像的原始尺寸进行拼接以得到所述拼接图像,其中,所述第四预设数量和所述第五预设数量之和等于第二预设数量。

[0040] 在一种实施例中,所述叠加模块包括加权值分配子模块和线性叠加子模块,所述加权值分配子模块用于为每张待叠加的所述拼接图像分配各自对应的预设加权值,其中,每一所述预设加权值的取值范围为(0,1);

[0041] 所述线性叠加子模块用于基于各所述拼接图像的所述预设加权值并按照预设混叠公式进行线性叠加以得到混叠图像,所述混叠图像中的所述目标对象的局部信息被遮挡。

[0042] 本申请实施例还提出一种终端设备,所述终端设备包括处理器和存储器,所述存储器存储有计算机程序,所述处理器用于执行所述计算机程序以实施上述的方法。

[0043] 本申请实施例还提出一种可读存储介质,其存储有计算机程序,所述计算机程序被执行时实施上述的方法。

[0044] 本申请的实施例具有如下优点:

[0045] 本申请的技术方案通过对数量有限的目标场景图像按照预设规则进行增强以得到更多的一次增强图像,进而按照相应的裁剪、拼接及叠加规则对各所述增强图像依次进行裁剪、拼接及叠加的二次增强处理,从而得到含目标对象的信息被遮挡的混叠图像以作为用于模拟拥挤场景的模拟图像。该方法不仅可以很好地解决各种视觉运用场景下的模型训练时实际样本图像的资源受限问题,尤其是针对目标对象信息被遮挡的场景,通过该方

法可以得到更多贴合实际拥挤场景的模拟图像。此外,由于这些模拟图像可以兼容全局场景图像的上下文信息及场景图像中目标对象的局部信息,因此,可较好地保证模型训练的质量等。若将该方法运用于视觉机器人的图像模拟及模型训练,可使该机器人具有更好的泛化性能与鲁棒性等。

附图说明

[0046] 为了更清楚地说明本申请实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本申请的某些实施例,因此不应被看作是对范围的限定,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他相关的附图。

[0047] 图1示出了本申请实施例的用于模拟拥挤场景的图像获取方法的流程示意图;

[0048] 图2示出了本申请实施例的用于模拟拥挤场景的图像获取方法的图像增强流程示意图;

[0049] 图3示出了本申请实施例的用于模拟拥挤场景的图像获取方法的生成混叠图像的流程示意图;

[0050] 图4示出了本申请实施例的用于模拟拥挤场景的图像获取方法的一种应用场景下的目标场景图像和模拟图像;

[0051] 图5示出了本申请实施例的图像获取方法的流程示意图;

[0052] 图6示出了本申请实施例的应用于机器人的视觉目标处理方法的第一流程图;

[0053] 图7示出了本申请实施例的应用于机器人的视觉目标处理方法的第二流程图;

[0054] 图8示出了本申请实施例的机器人的视觉目标处理方法的一种应用场景中目标场景图像中足球标注示意图和混叠图像中的标注示意图;

[0055] 图9示出了本申请实施例的用于模拟拥挤场景的图像获取装置的结构示意图。

具体实施方式

[0056] 下面将结合本申请实施例中附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。

[0057] 通常在此处附图中描述和示出的本申请实施例的组件可以以各种不同的配置来布置和设计。因此,以下对在附图中提供的本申请的实施例的详细描述并非旨在限制要求保护的本申请的范围,而是仅仅表示本申请的选定实施例。基于本申请的实施例,本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0058] 在下文中,可在本申请的各种实施例中使用的术语“包括”、“具有”及其同源词仅意在表示特定特征、数字、步骤、操作、元件、组件或前述项的组合,并且不应被理解为首先排除一个或更多个其它特征、数字、步骤、操作、元件、组件或前述项的组合的存在或增强一个或更多个特征、数字、步骤、操作、元件、组件或前述项的组合的可能性。

[0059] 此外,术语“第一”、“第二”、“第三”等仅用于区分描述,而不能理解为指示或暗示相对重要性。

[0060] 除非另有限定,否则在这里使用的所有术语(包括技术术语和科学术语)具有与本

申请的各种实施例所属领域普通技术人员通常理解的含义相同的含义。术语(诸如在一般使用的词典中限定的术语)将被解释为具有与在相关技术领域中的语境含义相同的含义并且将不被解释为具有理想化的含义或过于正式的含义,除非在本申请的各种实施例中清楚地限定。

[0061] 本申请中,目标场景图像,通常是指在对应真实场景下采集到的图像。例如,对于用于物体识别、跟踪的机器人,该目标场景图像可以是该机器人在对应场景下通过拍摄头拍摄到的包含或不包含该物体的图片。又例如,对于用于识别是否有车辆出现的装置,该目标场景可以是实际现场采集到的可能包含有车辆或不包含车辆的真实图片等。

[0062] 其中,该目标场景图像可以是在同一场景下含有相同类型的目标对象的图像,也可以是不同场景下含有相同或不同的目标对象的图像,具体可根据实际需求来选取。例如,对于机器人跟踪足球的应用场景而言,该目标场景图像可以是机器人获取到的包含不同状态下的足球的图像。对于风景识别的应用场景,该目标场景图像则可包括不同场景下的不同风景的图像等。

[0063] 拥挤场景,主要是指目标对象被遮挡而无法获得目标对象的完整信息的情况。例如,进行物体跟踪时,该物体的部分区域可能会被其他对象遮挡。又例如,进行人头检测时,前面的人可能遮挡了后面的人的部分区域等。

[0064] 对于该目标场景图像,可将这些目标场景图像作为训练样本,进行相应的模型训练。例如,上述的用于物体识别、跟踪的模型,或用于车辆识别的模型等等。然而,在实际运用中,由于受到时间、环境等的诸多限制,采集到的目标场景图像的数量及状态往往是有限的,由于模型训练时没有足够多的训练样本,导致该模型对于一些如拥挤场景等的特殊情况的处理能力往往较差,如会表现出一些误判或漏检等不准确的现象。

[0065] 为此,本申请提出基于该目标场景图像进行相应处理后来得到更多的模拟图像,尤其是出现物体被遮挡的拥挤场景下的图像,这样可以丰富模型训练样本的多样性,进而提升相应模型的泛化能力和鲁棒性等。

[0066] 下面结合具体的实施例对该图像获取方法进行详细说明。

[0067] 实施例1

[0068] 请参照图1,本实施例提出一种用于模拟拥挤场景的图像获取方法,可应用于模拟目标对象的部分信息被遮挡或存在重叠现象等的场景。考虑到目标场景图像的数量有限,为此,本实施例将采用一系列图像处理操作来对有限的目标场景图像进行处理,以得到更多拥挤场景下的模拟图像。下面对该方法进行说明。

[0069] 步骤S110,根据预设规则对目标场景图像进行增强处理,得到第一预设数量的增强图像,其中,预设比例的多张增强图像中均含有目标对象。

[0070] 通常地,该目标场景图像的数量包括多张,利用这些目标场景图像进行数据增强处理,可以得到数量更多的增强图像。该第一预设数量可以根据实际训练需要来设定。可以理解,该增强处理主要是利用数量有限的目标场景图像来得到数量更多的候选图像,进而可基于这些候选图像进行后续的拼接及叠加等处理。

[0071] 在一种实施例中,这些增强图像中存在预设比例的图像均含有目标对象,即从中可以获取到该目标对象的局部或全部等相关信息。示范性地,该预设比例可以选取为小于或等于100%,如50%、70%、80%或100%等。可以理解,若该预设比例小于100%,即表示该

第一预设数量的增强图像中还包括一些不包含目标对象的任何信息而仅包含对应场景的背景图像。

[0072] 在一种实施方式中,上述的预设规则可包括但不限于包括,设定的图像调整类型和/或调整尺寸,以及设定的噪声模拟规则等。其中,上述的图像调整类型和/或调整尺寸主要根据该目标对象的特征属性来设定,示范性地,该特征属性可包括但不限于为该目标对象在对应场景下的呈现状态、尺寸等中的一种或组合。例如,图像调整类型可根据该实际的呈现状态来确定;调整尺寸可根据目标对象在实际场景下的尺寸来确定等。这样可以保证得到的调整图像能够符合实际场景,可避免出现在实际的拍摄图像中不可能出现的情况。以一人额头检测的场景为例,若额头作为该目标对象,在实际的拥挤情况中,前后人群最多只会出现额头的左右遮挡或上下遮挡,即意味着对该目标场景图像进行调整时,预设规则中不应当包括对该目标场景图像进行垂直翻转的调整操作等。

[0073] 对于上述的根据预设规则对目标场景图像进行增强处理的步骤,如图2所示,可包括:

[0074] 步骤S111,根据设定的图像调整类型和/或调整尺寸对目标场景图像进行调整处理,得到第一预设数量的调整图像。

[0075] 示范性地,该图像调整类型主要包括对目标场景图像进行图像调整操作,例如,可包含但不限于包括翻转、裁剪、旋转、缩放、亮度调节等中的一种或多种组合。

[0076] 其中,上述的翻转包括对图像进行水平或垂直翻转,这样可以用于模拟水平对称或者垂直对称场景下的物体分布。裁剪包括对图像进行局部区域图像进行截取,这样可以模拟实际场景中获取到物体的局部信息的情况。旋转包括对图像进行相应角度的旋转,这样可以用于模拟拍摄装置在旋转、晃动等条件下捕捉到的图像。缩放包括对图像进行尺寸比例缩放,即按照一定的缩放比例进行放大或缩小处理,这样可以用于模拟近处与远处场景下的物体分布图像。亮度调节包括将图像由RGB空间转变为HSV空间并对HSV空间中的亮度分量进行大小调节,即通过对图像的亮度信息叠加一些噪声,这样可以用于模拟不同光照条件下捕捉到的图像。

[0077] 值得注意的是,对各张目标场景图像可以每次仅执行其中一种调整操作,也可以是每次执行多种调整操作,这样通过执行不同的调整操作,可以得到更多的图像,实现成倍的扩增数据的多样性等。其中,在进行上述多种的调整操作时,各操作的顺序可不作限定。

[0078] 步骤S112,根据设定的噪声模拟规则对各调整图像进行噪声模拟处理,得到第一预设数量的增强图像。

[0079] 在得到上述的调整图像后,还可以进一步进行噪声添加处理。通过噪声添加处理可以用于模拟因拍摄装置的运动、图像传感器自身存在的噪声等原因引起的图像模糊现象。在一种实施方式中,该噪声模拟规则可以是添加或随机添加如高斯白噪声等。

[0080] 值得注意的是,噪声模拟处理应当是在进行上述的调整操作之后执行。这是由于,该噪声模拟的是拍摄装置的图像传感器在投影成像时产生的噪声,故采用这样的处理顺序可以保证整个增强处理的过程能够符合相机成像的真实流程,进而使得到的模拟图像更接近于真实拍摄得到的图像等。

[0081] 步骤S120,对各增强图像进行裁剪以得到裁剪图像,其中,对含有目标对象的信息的增强图像的裁剪尺寸大于或等于预设最小裁剪尺寸。

[0082] 于是,在得到数量足够多的增强图像后,将对这些增强图像进行裁剪,而得到的裁剪图像将用于后续的拼接操作。可选地,在裁剪时,可以根据实际需求而随时调整裁剪尺寸。

[0083] 值得注意的是,对于那些含有目标对象的信息增强图像,其裁剪时的尺寸应当大于或等于设定的最小裁剪尺寸,这样可以减少无效训练样本的数量。可以理解,若裁剪时的尺寸过小,将导致最终得到的模拟图像可能不会含有目标对象的有用信息,此时将视为无效的模型训练样本。

[0084] 在一种实施方式中,该预设最小裁剪尺寸可根据该目标对象在拍摄装置的拍摄视野内占有的像素范围来设定,通常地,可将占有的最小像素数值作为该最小裁剪尺寸。示范性地,可根据拍摄装置的相关拍摄参数及该目标对象自身的尺寸等来估算占用的像素范围。

[0085] 以一运动机器人为例,为了实现预设距离范围内的足球检测,可利用该足球的平均直径、用于拍摄足球的该机器人的摄像头相关参数,如机器人的摄像头离地高度、摄像头的镜头焦距、镜头成像参数、图像传感器的尺寸等来计算当该足球位于预设距离处时在拍摄视野内的图像像素大小。假设计算得到的最小像素数值约为 $10 \times 12 \text{ pixel}$,则将该最小像素作为最小裁剪尺寸。相应地,在对含有目标对象的增强图像进行裁剪时,应当保证裁剪尺寸不小于该最小裁剪尺寸。

[0086] 步骤S130,选取第一预设数量的裁剪图像进行拼接以生成多张拼接图像,其中,该第二预设数量中至少一张裁剪图像中含有目标对象。

[0087] 在一种实施方式中,用于拼接为一张拼接图像的多张裁剪图像中,应当选取至少一张含有该目标对象的信息。其中,该目标对象的信息可以是全局信息,也可以是经过裁剪后的局部信息。示范性地,该第一预设数量可以根据实际需求设定,如2张~8张等。例如,在一些实施例中,该第一预设数量可选取为4张。

[0088] 可以理解,裁剪包括对图像进行一定区域的图像截取,这样可以用截取的区域来表征原图中的局部信息,进而通过图像拼接可实现在一张图像中呈现物体的多种局部信息,如大小、颜色、纹理、轮廓等。此外,利用图像拼接的方式来使一张图像中呈现不同图像的信息,这样可以丰富数据的多样化,同时还可以保持极小的计算开销,即既强化了图像中物体的局部信息,也保证了在计算资源受限的情况下实现数据的多样性。

[0089] 在另一种实施例中,该目标场景图像还包括不含目标对象的背景图像,同样,背景图像经过裁剪后可得到对应的背景裁剪图像。此时,对于上述步骤S130中的拼接图像的生成包括:

[0090] 选取第四预设数量的含有目标对象的裁剪图像和第五预设数量的背景裁剪图像,并按照目标场景图像的原始尺寸进行拼接以得到拼接图像,即此时拼接图像的尺寸与目标场景图像的原始尺寸相同。其中,第四预设数量和第五预设数量之和等于第二预设数量。

[0091] 例如,若待拼接图像由80%的含有目标对象信息的裁剪图像和20%的不含目标对象信息的背景裁剪图像组成,预设拼接图像设为4张,则对于每次的图像拼接,若选取了3张含有目标对象信息的裁剪图像,则将从得到的背景裁剪图像中选取1张来进行拼接以得到一张拼接图像。

[0092] 步骤S140,选取第三预设数量的拼接图像进行叠加以得到含目标对象的信息被遮

挡的混叠图像,将该混叠图像作为拥挤场景下的模拟图像。

[0093] 在一种实施方式中,如图3所示,对于上述步骤S140中的混叠图像的生成,包括:

[0094] 步骤S141,为每张待叠加的拼接图像分配各自对应的预设加权值,其中,每一预设加权值的取值范围为(0,1)。

[0095] 步骤S142,基于各拼接图像的预设加权值并按照预设混叠公式进行线性叠加以得到混叠图像。其中,该混叠图像中的目标对象的局部信息被遮挡。

[0096] 示范性地,通过为每个拼接图像分配一个相应的加权值,进而将对应拼接图像的图像数据及加权值进行叠加,得到最终的混叠图像。对于每一预设加权值,其取值范围均为(0,1),可选地,所有的预设加权值的总和为1。

[0097] 在一种实施方式中,若该第三预设数量为2,即将2张拼接图像进行线性叠加,此时为2张拼接图像分别分配一个预设加权值,若 $\tilde{x}$ 表示混叠图像,第一张拼接图像 x_i 的预设加权值记为 λ ,第二张拼接图像 x_j 的预设加权值为 $(1-\lambda)$ 于是,采用如下的混叠公式:

$$[0098] \quad \tilde{x} = \lambda x_i + (1 - \lambda) x_j。$$

[0099] 其中, $\lambda \in (0,1)$ 且服从Beta分布。Beta分布是一种连续型概率密度分布,此时可表示为 $\lambda \sim \text{Beta}(a,b)$,参数 a,b 决定该分布函数的形状。本实施例,使参数 a 与 b 相等,即有 $\lambda \sim \text{Beta}(a,a)$, a 用于设置图像混叠的强度,其中 a 越大,混叠效果越明显。

[0100] 在上述的实施方式中,进一步地,若根据目标场景图像的标注信息计算得到各拼接图像中目标对象的标注信息,则混叠图像的标注信息同样会被两个拼接图像的预设加权值调和而变化,若第一张拼接图像 x_i 中的目标对象的标注信息为 y_i ,第二张拼接图像 x_j 中的目标对象的标注信息为 y_j ,此时混叠图像中目标对象的标注信息 $\tilde{y}$ 的计算公式为:

$$[0101] \quad \tilde{y} = \lambda y_i + (1 - \lambda) y_j。$$

[0102] 通常地,当训练样本数据规模较小,数据分布多样性有限、网络训练时间较长时,模型容易呈现出过度拟合的趋势,即表现为过分抽取了图像的局部细节信息而忽视了图像的全局信息。为尽量避免出现上述现象,本实施例通过将呈现有多张局部信息的拼接图像进行多张叠加,这样既可以强化图像的全局上下文信息,又可以避免模型陷入过分抽取局部信息而造成过度拟合的现象。此外,由于多张拼接图像经过混叠处理后,往往会出现目标对象信息被遮挡(即局部信息不可见或模糊)的情况,此时可将其作为用于模拟拥挤场景下的模拟图像。

[0103] 以一实际场景为例,图4的左部分和右部分分别示出了一张目标场景图像和最终的混叠图像。当然,为了得到如右边的混叠图像,还包括其他的目标场景图像(未示出),而为了能够更准确地识别出球场上的运动人员等,通过对多张目标场景图像进行上述的增强、裁剪、拼接及线性叠加处理,从而得到如图4的右边所示的混叠图像。从该混叠图像中可以看出,一些运动人员出现被遮挡的情况,此时即可以视为拥挤场景下的模拟图像。

[0104] 本实施例的用于拥挤场景的图像获取方法通过基于目标对象的特征属性设定的规则对目标场景图像进行增强,以得到数量更多的候选图像;进而对这些候选图像按照相应的要求来进行裁剪、拼接及混叠以最终得到含目标对象的信息被遮挡的模拟图像。利用该模拟图像进行模型训练,可较好地解决现有的训练样本的数量有限、多样性有限的问题,

此外,这些模拟图像可以很好地兼顾图像全局与局部特征,而且在较小的计算开销下,保证一次训练过程可以获得多张图像的全局图像上下文信息与对应场景中的物体局部信息,故也可以保证较高质量的模型训练等。

[0105] 实施例2

[0106] 请参照图5,与上述实施例1的图像获取方法类似,本实施例提出一种图像获取方法,该方法不仅仅局限用于模拟物体信息被遮挡的拥挤场景下的图像,还可以用于其他的场合,如由仿真场景迁移到真实场景时的训练图像扩增等。由于仿真下的数据较粗糙,无法准确模拟真实目标对标的纹理等信息,若基于部分真实场景数据的数据并采用该图像获取方法进行数据扩充,可以弥补仿真数据所存在的缺陷,进而减少算法由仿真场景迁移到真实场景时的时耗等。

[0107] 步骤S210,对目标场景图像进行一次增强,得到多张第一图像。其中,该一次增强包括对目标场景图像进行预设增强类型的调整操作及对调整后的图像进行噪声模拟处理。

[0108] 本实施例中,目标场景图像可以为同一场景下的具有相同类型的目标对象的图像,如办公楼的进出人员的体温检测等;或者可以为不同场景下的具有相同类型的目标对象的图像,如不同场合下拍摄得到的含有车辆在内的图像等;又或者可以为不同场景下的具有不同类型的目标对象的图像,如拍摄得到的不同地方的不同景点的图像等。可以理解,具体可根据实际需求来确定,在此并不作限定。

[0109] 本实施例中,该第一图像即为经过该一次增强处理的图像。关于步骤S210中的一次增强处理,示范性地,可包括但不限于包括先对目标场景图像进行预设增强类型的图像调整操作,再对调整后的图像进行噪声模拟处理等。在一种实施方式中,可对各目标场景图像执行如随机翻转、随机裁剪、随机旋转、随机缩放、随机亮度调节等的一种或多种增强类型的调整操作。进而,可对调整后的图像进行随机添加噪声处理,如高斯白噪声等。

[0110] 应当理解,对于目标场景图像的调整处理,具体可根据实际需求来确定,例如,若用于模拟上述的拥挤场景,则可按照上述实施例1的基于目标对象的特征属性而设定的规则来调整,若是其他的一些场景,也可以是进行上述的随机调整处理。

[0111] 步骤S220,对各第一图像进行二次增强,得到多张第二图像,将该第二图像作为相应场景的模拟图像。其中,该二次增强包括对该多张第一图像依次进行图像裁剪、图像拼接和图像混叠处理。

[0112] 示范性地,对经过一次增强得到的各第一图像进行裁剪,以用于截取各图像中的局部信息,进而用于后续的拼接和叠加处理。可以理解,进行图像裁剪、图像拼接及图像混叠处理时,可根据最终得到的混叠图像的作用来确定是否设定相应的规则来进行处理。

[0113] 其中,进行裁剪处理时,在一种实施方式中,可以是按照随机尺寸进行裁剪;在另一种实施方式中,也可以是按照如上述实施例1中设定的裁剪尺寸规则来裁剪等,具体可根据实际运用场景来选定。

[0114] 其中,进行拼接处理时,可从大量的裁剪图像中随机选取或按照相应规则选取出一定数量的裁剪图像进行多张拼接,从而得到对应的拼接图像。通常地,该拼接图像的尺寸等于目标场景图像的尺寸。

[0115] 其中,进行叠加处理时,可从得到的多张拼接图像中随机选取或按照相应规则选取出一定数量的拼接图像进行多张线性叠加,从而得到对应的混叠图像。可以理解,通过调

整相应的图像叠加参数,可以使图像中的目标对象产生被遮挡或不完全清晰可见等多种效果。

[0116] 本实施例中,该第二图像即为经过二次增强处理后得到的混叠图像。将该第二图像作为相应场景的模拟图像,可较好地解决现有的训练样本的数量有限的问题,另外,由于同一混叠图像中呈现有不同的目标对象的信息,丰富了数据的多样性等。

[0117] 实施例3

[0118] 请参照图6,本实施例提出一种视觉处理方法,主要是将上述的图像获取方法应用于包括拍摄装置的终端设备中,如机器人等。下面对该视觉处理方法进行说明。

[0119] 步骤S310,通过拍摄装置采集目标场景图像,并基于目标场景图像采用上述的图像获取方法生成模拟图像。

[0120] 通常地,当终端设备的用途不同时,终端设备所采用的神经网络模型也不同。考虑到终端设备采集到的实际场景下的目标场景图像的数量可能有限,在训练模型时,可能出现训练数据可能不足或者分布高度相似等现象,进而会影响模型的训练效果,为此,本实施例提出基于采集到的目标场景图像并利用上述的图像获取方法来丰富训练数据的多样性,进而提升神经网络模型的泛化性能与鲁棒性等。

[0121] 其中,关于图像获取方法的具体步骤可参见上述实施例1或2,故在此不再重复描述。值得注意的是,进行裁剪时,含有目标对象的增强图像的裁剪尺寸应当大于或等于预设最小裁剪尺寸。其中,该预设最小裁剪尺寸可根据相应场景下目标对象在终端设备的拍摄视野内占有的像素范围来设定,具体可参见上述步骤S120。

[0122] 步骤S320,利用目标场景图像和模拟图像进行预设场景模型训练,以得到训练后的场景模型,该训练后的场景模型用于对机器人在相应场景下拍摄得到的待处理场景图像进行与目标对象相关的视觉处理操作。

[0123] 示范性地,该预设场景模型主要取决于该终端设备的用途,在不同的运用场合下,该预设场景模型也不现。例如,可包括但不限于为,物体跟踪模型、物体检测模型或物体分割模型等。以一物体跟踪模型为例,该预设场景模型具体可采用深度卷积神经网络等构建得到。其中,关于该深度卷积神经网络的具体结构在此不作过多说明,具体可根据实际需求来确定。

[0124] 利用上述的目标场景图像和模拟图像进行模型训练后,可将训练好的场景模型迁移至实际场景中并对拍摄得到的待处理场景图像执行相应的视觉处理操作。

[0125] 以机器人为例,示范性地,该机器人可以是任何需要进行视觉图像处理的机器人,例如,用于物体跟踪或物体检测的家用或工业用机器人、用于分割物体的工业机器人等。在一种实施方式中,该视觉处理操作可包括但不限于为机器人对目标对象进行识别、跟踪、分割等,当然具体取决于该机器人的作用。

[0126] 可以理解,本实施例的视觉处理方法利用上述的图像获取方法来得到机器人中的场景模型的训练样本图像,不仅可以很好地解决实际训练样本不足的问题,尤其是对于如机器人等这种计算开销受限的终端设备,采用该方法可以在一次训练过程可以获取多张图像的全局图像上下文信息与对应场景中的物体局部信息,从而保证较高质量的模型训练等。

[0127] 在一种实施方式中,如图7所示,对于上述步骤S310中的生成模拟图像之前,该方

法还包括：

[0128] 步骤S330,对各目标场景图像分别进行目标对象标注以获取对应的标注信息。

[0129] 通常地,对于用于训练模型的样本图像,需要先进行目标对象标注,如图8所示,这样可以得到该目标场景图像中的目标对象的标注信息,进而方便进行目标对象的识别训练等。其中,该标注信息主要包括该目标对象所在的标注框的中心点位置以及长、宽等信息。

[0130] 进一步地,生成模拟图像之后,该方法还包括：

[0131] 步骤S340,根据目标场景图像的标注信息以及增强、裁剪、拼接和叠加时各自对应的处理参数计算模拟图像的标注信息。

[0132] 由于图像经过一系列的图像处理,原来的标注框信息会被调和,故需要重新计算目标对象的标注信息。示范性地,通过对目标场景图像进行增强、裁剪、拼接和图像叠加操作后,可利用当时设定的对应参数来计算得到模拟图像中该目标对象的标注信息。

[0133] 例如,若在增强过程中,对目标场景图像进行了角度旋转,相应地,该标注框的位置也将旋转相应角度,根据旋转的角度值及标注框在原始的目标场景图像中的位置信息即可计算得到新的标注信息,其他处理步骤同理计算。

[0134] 步骤S350,利用带标注信息的目标场景图像和模拟图像对预设场景模型进行训练。

[0135] 由于各模拟图像中的标注信息可计算得到,故可以直接利用带标注信息的图像进行上述模型的训练。可以理解,通过对原始的目标场景图像提前标注,再获取模拟图像,这样直接计算得到各模拟图像的标注信息,而不需要再对模拟图像利用标注算法再逐一标注,可大大提高效率。此外,也可以根据计算得到的标注信息来排除一些无效的样本数据等。

[0136] 实施例4

[0137] 请参照图9,基于上述实施例1,本实施例提出一种用于模拟拥挤场景的图像获取装置10,其中,该拥挤场景为定义的存在目标对象的信息被遮挡的场景,该图像获取装置10包括：

[0138] 增强模块110,用于根据预设规则对目标场景图像进行增强处理,得到第一预设数量的增强图像,其中,预设比例的多张所述增强图像中均含有所述目标对象。

[0139] 裁剪模块120,用于对各所述增强图像进行裁剪以得到裁剪图像,其中,含有所述目标对象的增强图像的裁剪尺寸大于或等于预设最小裁剪尺寸。

[0140] 拼接模块130,用于选取第二预设数量的所述裁剪图像进行拼接以生成多张拼接图像,其中,所述第二预设数量中至少一张裁剪图像中含有所述目标对象。

[0141] 叠加模块140,用于选取第三预设数量的所述拼接图像进行叠加以得到含有所述目标对象的局部信息被遮挡的混叠图像,将所述混叠图像作为所述拥挤场景下的模拟图像。

[0142] 在一种实施例中,该预设规则包括基于目标对象的特征属性设定的图像调整类型和/或调整尺寸,以及设定的噪声模拟规则,增强模块110包括调整处理子模块111和噪声模拟子模块112,其中,调整处理子模块111用于根据所述图像调整类型和/或调整尺寸对目标场景图像进行调整处理,得到第一预设数量的调整图像;噪声模拟子模块112用于根据所述噪声模拟规则对各调整图像进行噪声模拟处理,得到第一预设数量的增强图像。

[0143] 在一种实施例中,该多张增强图像中包括不含目标对象的背景图像,所述背景图

像经过裁剪后得到对应的背景裁剪图像,拼接模块130还用于选取第四预设数量的含有目标对象的裁剪图像和第五预设数量的背景裁剪图像,并按照目标场景图像的原始尺寸进行拼接以得到拼接图像,其中,第四预设数量和第五预设数量之和等于第二预设数量。

[0144] 在一种实施例中,叠加模块140包括加权值分配子模块141和线性叠加子模块142,其中,加权值分配子模块141用于为每张待叠加的拼接图像分配各自对应的预设加权值,每一预设加权值的取值范围为(0,1);线性叠加子模块142用于基于各拼接图像的预设加权值并按照预设混叠公式进行线性叠加以得到混叠图像,所述混叠图像中的目标对象的局部信息被遮挡。

[0145] 可以理解,本实施例的各模块对应于上述实施例1的各步骤,上述实施例1的可选项同样适用于本实施例,故在此不再详述。

[0146] 本申请实施例还提出一种终端设备,例如,可以是机器人或其他智能视觉设备等。示范性地,该终端设备包括处理器和存储器,其中,存储器存储有计算机程序,处理器用于执行计算机程序以实施上述实施例的方法或上述装置中的各个模块的功能。

[0147] 存储器可包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序;存储数据区可存储根据移动终端设备的使用所创建的数据。此外,存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

[0148] 本申请实施例还提出一种可读存储介质,其存储有上述终端设备中使用的计算机程序。

[0149] 在本申请所提供的几个实施例中,应该理解到,所揭露的装置和方法,也可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,附图中的流程图和结构图显示了根据本申请的多个实施例的装置、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,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,在作为替换的实现方式中,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,两个连续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,结构图和/或流程图中的每个方框、以及结构图和/或流程图中的方框的组合,可以用执行规定的功能或动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。

[0150] 另外,在本申请各个实施例中的各功能模块或单元可以集成在一起形成一个独立的部分,也可以是各个模块单独存在,也可以两个或更多模块集成形成一个独立的部分。

[0151] 功能如果以软件功能模块的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是智能手机、个人计算机、服务器、或者网络设备等)执行本申请各个实施例方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0152] 以上,仅为本申请的具体实施方式,但本申请的保护范围并不局限于此,任何熟悉

本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本申请的保护范围之内。

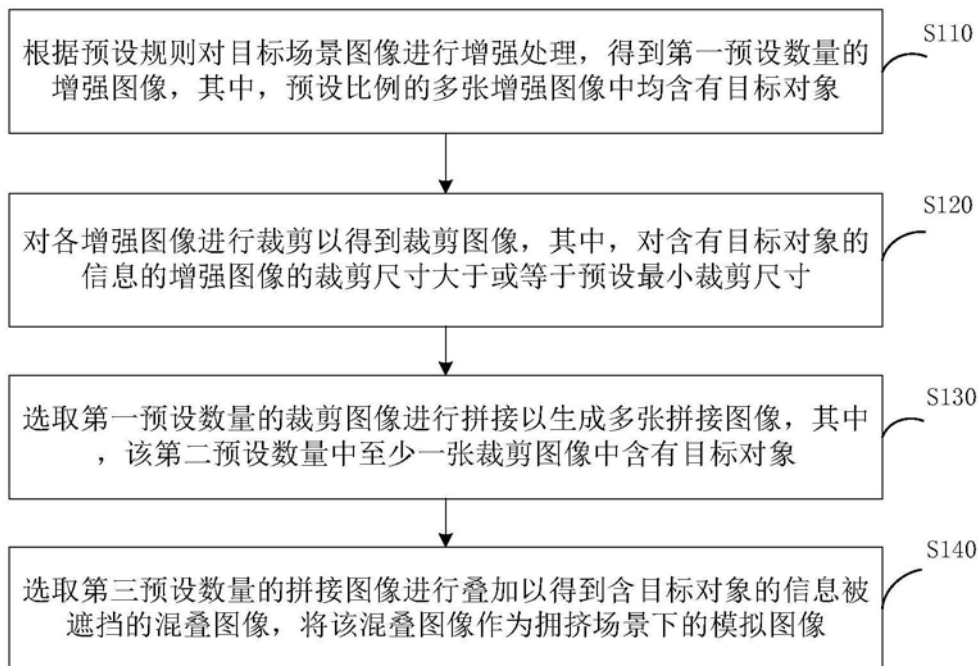


图1

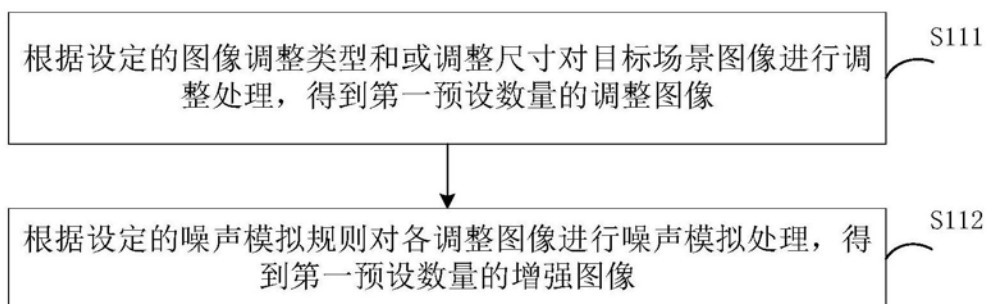


图2

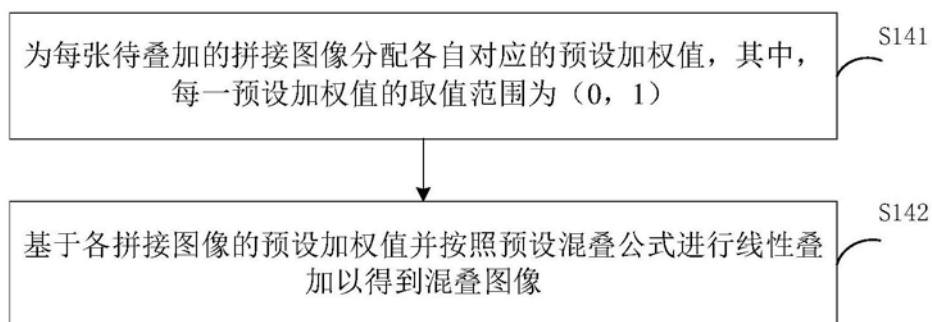


图3

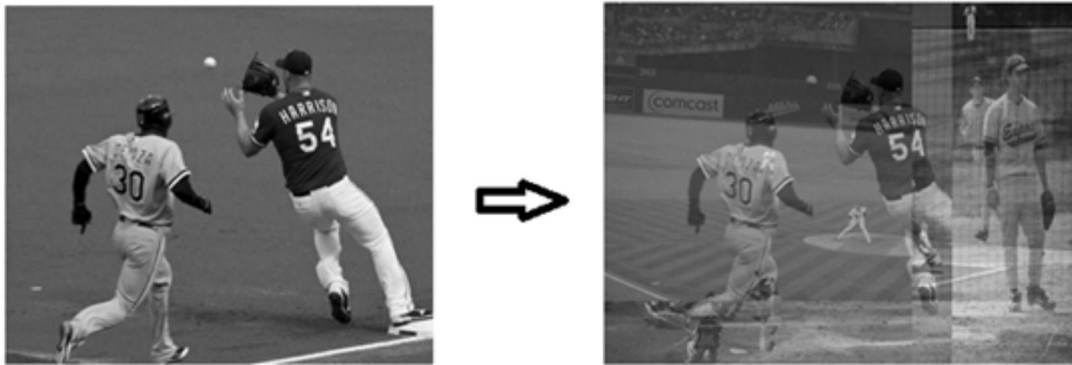


图4

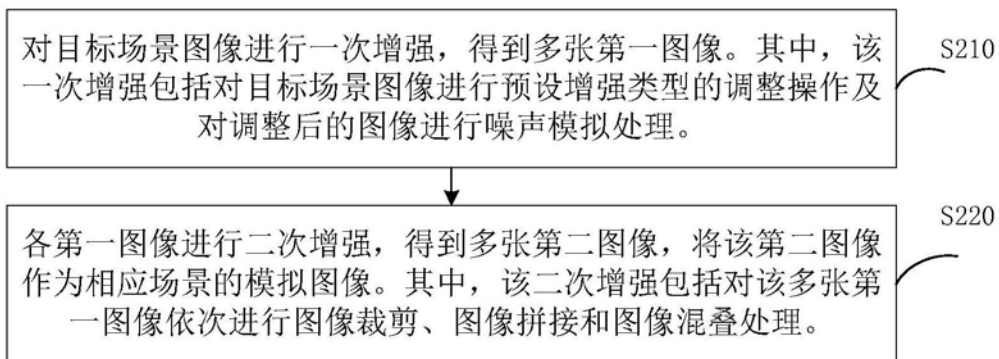


图5

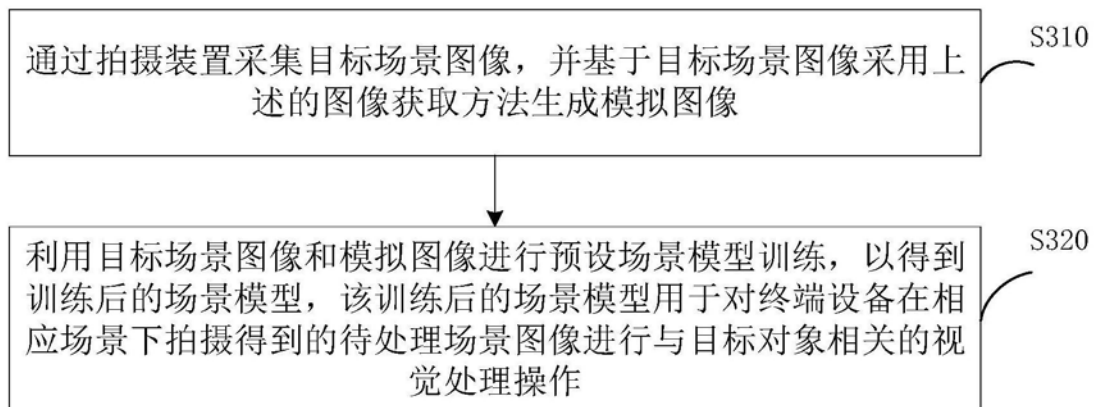


图6

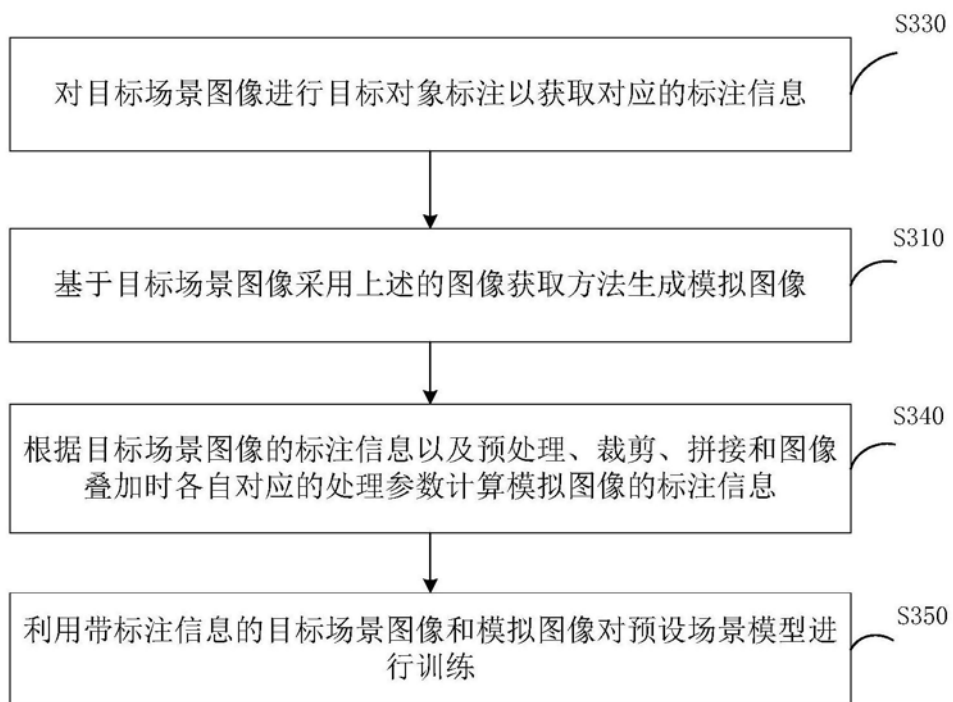


图7

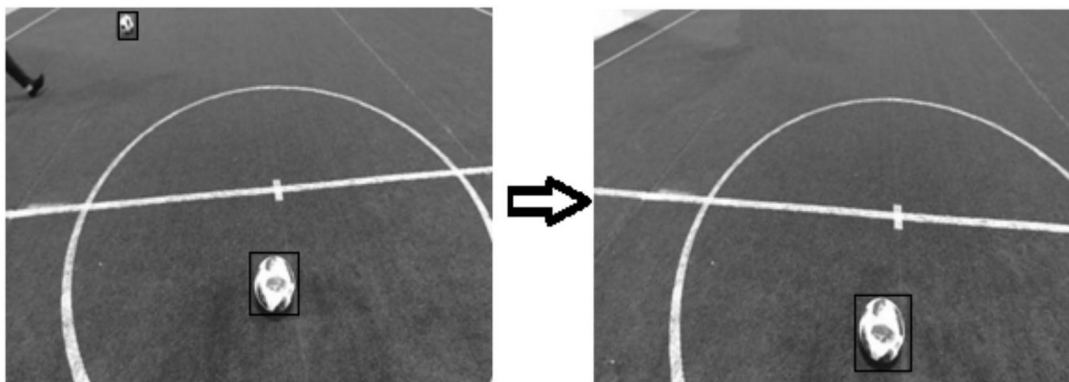


图8

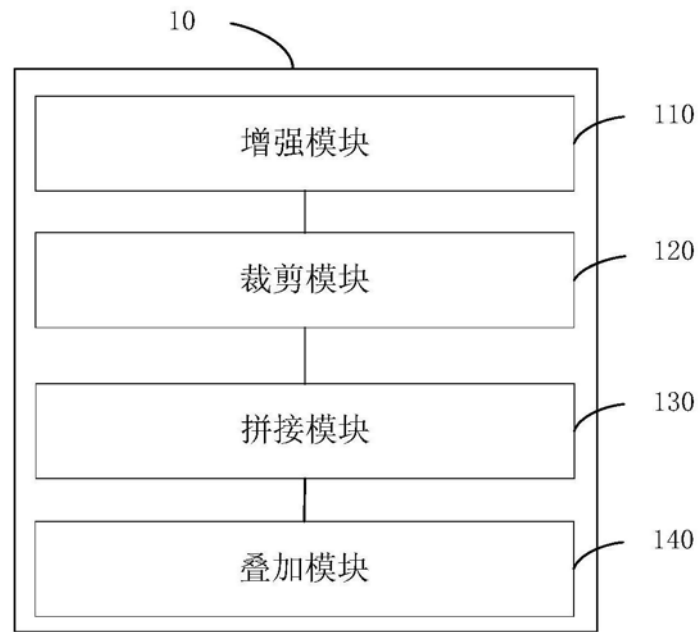


图9