

一种图像深度估计方法及装置、电子设备、存储介质

相关申请的交叉引用

本申请基于申请号为 201910621318.4、申请日为 2019 年 07 月 10 日，申请名称为“一种图像深度估计方法及装置、电子设备、存储介质”的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的全部内容在此以引入方式结合在本申请中。

技术领域

本公开涉及计算机视觉技术领域，尤其是一种图像深度估计方法及装置、电子设备、存储介质。

背景技术

图像的深度估计是计算机视觉领域的重要问题。在无法直接获得图像的深度信息时，只有通过深度估计方法才能完成场景的三维重建，进而为增强现实和游戏等应用服务。

目前，基于计算机视觉的深度估计方法可以分为主动视觉方法和被动视觉方法两类。其中，主动视觉方法是指向被测物体发射可控制光束，然后拍摄光束在物体表面上形成的图像，通过几何关系计算出被测物体距离的方法，被动视觉方法包括立体视觉、聚焦法，以及散焦法等，主要是通过一个或多个摄像装置获取的二维图像信息确定深度信息。

发明内容

本公开实施例期望提供一种图像深度估计方法及装置、电子设备、存储介质。

本公开实施例的技术方案是这样实现的：

本公开实施例提供了一种图像深度估计方法，所述方法包括：

获取当前帧对应的参考帧和所述当前帧的逆深度空间范围；

对所述当前帧和所述参考帧分别进行金字塔降采样处理，获得所述当前帧对应的 k 层当前图像，以及所述参考帧对应的 k 层参考图像；k 为大于等于 2 的自然数；

基于所述 k 层参考图像和所述逆深度空间范围，对所述 k 层当前图像进行逆深度估计迭代处理，获得所述当前帧的逆深度估计结果。

可以理解的是，在本公开的实施例中，对当前帧和当前帧对应的参考帧进行降采样处理，将获得的多层当前图像结合多层参考图像进行逆深度估计迭代处理，以确定当前帧的逆深度估计结果。由于在确定逆深度估计结果的过程中，逐层减少逆深度搜索空间，从而减少了逆深度估计的计算量，提升了估计速度，能够实时获得逆深度估计结果。

在上述图像深度估计方法中，所述获取当前帧对应的参考帧，包括：

获取至少两个待筛选帧；

从所述至少两个待筛选帧中，选取与所述当前帧之间满足预设角度约束条件的至少一帧，将所述至少一帧作为所述参考帧。

可以理解的是，在本公开的实施例中，按照预设角度预设条件从至少两个待筛选帧中选取参考帧，可以在一定程度上选取出质量较佳，适合与当前帧进行匹配的帧，从而在后续深度估计过程中提高估计的准确性。

在上述图像深度估计方法中，所述预设角度约束条件包括：

所述当前帧对应的位姿中心和所述参考帧对应的位姿中心，与目标点的连线形成的夹角处于第一预设角度范围；所述目标点为所述当前帧对应的平均深度点与所述参考帧对应的平均深度点连线的中点；

所述当前帧和所述参考帧对应的光轴夹角处于第二预设角度范围；

所述当前帧和所述参考帧对应的纵轴夹角处于第三预设角度范围。

可以理解的是，在本公开的实施例中，第一个角度条件限定了当前场景到两个相机的距离，角

度过大说明场景过近,两帧重合度会较低,角度过小,则说明场景过远,视差较小,误差会比较大,当相机特别接近时也可能发生角度过小的情况,此时误差同样较大。第二角度条件是为了保证两个相机有足够的共视区域。第三个角度条件是为了避免相机绕着光轴旋转,影响后续深度估计计算过程。同时满足上述三个角度条件的帧作为参考帧有利于提高当前帧深度估计的精度。

- 5 在上述图像深度估计方法中,所述基于所述k层参考图像和所述逆深度空间范围,对所述k层当前图像进行逆深度估计迭代处理,获得所述当前帧的逆深度估计结果,包括:

 基于所述k层当前图像和所述逆深度空间范围,确定第i层采样点中每一个采样点对应的逆深度候选值;所述第i层采样点为对所述k层当前图像中第i层当前图像采样获得的像素点,i为大于等于1且小于等于k的自然数;

- 10 根据所述第i层采样点中每一个采样点对应的逆深度候选值和所述k层参考图像中第i层参考图像,确定所述第i层采样点中每一个采样点的逆深度值,获得第i层逆深度值;

 令 $i=i+1$,继续对所述k层当前图像中分辨率高于所述第i层当前图像的第 $i+1$ 层当前图像进行逆深度估计,直至 $i=k$ 为止,获得第k层逆深度值;

 将所述第k层逆深度值确定为所述逆深度估计结果。

- 15 可以理解的是,在本公开的实施例中,基于k层参考图像和逆深度空间范围,对k层当前图像进行逆深度估计迭代处理,例如可以从顶层(第1层)当前图像(像素最少的图像)开始,依次向底层进行逆深度估计迭代,逐层缩小逆深度搜索空间,从而有效的减少逆深度估计的计算量。

 在上述图像深度估计方法中,所述基于所述k层当前图像和所述逆深度空间范围,确定第i层采样点中每一个采样点对应的逆深度候选值,包括:

- 20 对所述逆深度空间范围进行区间划分,并在每个划分区间中选择一个逆深度值,得到多个初始逆深度值;

 将所述多个初始逆深度值确定为第1层采样点中每一个采样点对应的逆深度候选值;

 当i不等于1的情况下,从所述k层当前图像中获取第i-1层采样点,以及第i-1层逆深度值;

- 25 基于所述第i-1层逆深度值、第i-1层采样点,以及所述多个初始逆深度值,确定所述第i层采样点中每一个采样点对应的逆深度候选值。

 可以理解的是,在本公开的实施例中,针对逆深度空间范围进行区间划分,从而在不同区间内选取逆深度值,可以使得每一个区间内都存在一个逆深度值作为逆深度候选值。也就是说,每一个采样点在不同逆深度范围内都存在一个逆深度候选值,在后续进行确定采样点的逆深度值,可以保证不同逆深度范围的逆深度值都可以进行逆深度值估计确定,保证估计过程覆盖整个逆深度空间范围,从而最终可以估计出准确的逆深度值。

- 30 在上述图像深度估计方法中,所述基于所述第i-1层逆深度估计值、第i-1层采样点,以及所述多个初始逆深度值,确定所述第i层采样点中每一个采样点对应的逆深度候选值,包括:

 从所述第i-1层采样点中,确定与第一采样点距离最近的第二采样点,以及与所述第二采样点相邻的至少两个第三采样点;所述第一采样点为所述第i层采样点中任意一个采样点;

- 35 根据所述第i-1层逆深度值,获取所述至少两个第三采样点中每一个采样点的逆深度值,以及所述第二采样点的逆深度值,得到至少三个逆深度值;

 从所述至少三个逆深度值中,确定最大逆深度值和最小逆深度值;

 从所述多个初始逆深度值中,选取处于所述最大逆深度值和所述最小逆深度值范围内的逆深度值,对所述多个等分逆深度值进行选取,将选取出的逆深度值确定为所述第一采样点对应的逆深度候选值;

- 40 继续确定所述第i层采样点中非所述第一采样点的采样点对应的逆深度候选值,直至确定出所述第i层采样点中每一个采样点对应的逆深度候选值。

 可以理解的是,在本公开的实施例中,利用i-1层采样点对应的逆深度值从多个初始逆深度值中确定第i层采样点的逆深度候选值,可以更为准确的获得第i层采样点的逆深度候选值,并且,减少了逆深度候选值的数量,相应的,减少了逆深度估计的计算量。

- 45 在上述图像深度估计方法中,所述根据所述第i层采样点中每一个采样点对应的逆深度候选值和所述k层参考图像中第i层参考图像,确定所述第i层采样点中每一个采样点的逆深度值,获得第i层逆深度值,包括:

 对所述第i层采样点中每一个采样点,分别按照对应的逆深度候选值中的每一个逆深度值,将所述第i层采样点中每一个采样点投影到所述第i层参考图像中,获得所述第i层采样点中每一个采样点对应的第i层投影点;

根据所述第 i 层采样点和所述第 i 层投影点进行块匹配, 获得所述第 i 层采样点中每一个采样点对应的第 i 层匹配结果;

根据所述第 i 层匹配结果, 确定所述第 i 层采样点中每一个采样点的逆深度值, 获得所述第 i 层逆深度值。

- 5 可以理解的是, 在本公开的实施例中, 针对的第 i 层采样点, 分别与相应的第 i 层投影点进行匹配, 从而确定与采用不同逆深度值投影的投影点的差异程度, 因此, 可以准确选取出第 i 层采样点的逆深度值。

在上述图像深度估计方法中, 所述根据所述第 i 层采样点和所述第 i 层投影点进行块匹配, 获得所述第 i 层采样点中每一个采样点对应的第 i 层匹配结果, 包括:

- 10 利用预设窗口, 从所述第 i 层当前图像中选取以待匹配采样点为中心的第一图像块, 并从所述第 i 层参考图像中选取以所述待匹配采样点对应的第 i 层投影点中的每一个投影点分别为中心的多个第二图像块; 所述待匹配采样点为所述第 i 层采样点中任意一个采样点;

将所述第一图像块分别与所述多个第二图像块中每一个图像块进行比较, 获得多个匹配结果, 并将所述多个匹配结果确定为所述待匹配采样点对应的第 i 层匹配结果;

- 15 继续确定所述第 i 层采样点中与所述待匹配采样点不同的采样点对应的第 i 层匹配结果, 直至获得所述第 i 层采样点中每一个采样点对应的第 i 层匹配结果。

可以理解的是, 在本公开的实施例中, 利用块匹配的方式进行采样点和投影点的匹配, 得到的匹配结果实际上就是匹配的惩罚值, 其表征了该投影点与采样点的差异度, 相应的, 也体现了投影该投影点的逆深度值可作为采样点逆深度值的程度, 因此, 可以利用其结果后续较为准确的选取采样点的逆深度值。

- 20 在上述图像深度估计方法中, 所述根据所述第 i 层匹配结果, 确定所述第 i 层采样点中每一个采样点的逆深度值, 获得所述第 i 层逆深度值, 包括:

从目标采样点对应的第 i 层匹配结果中选取目标匹配结果; 所述目标采样点为所述第 i 层采样点中任意一个采样点;

- 25 将所述目标采样点对应的第 i 层投影点中, 所述目标匹配结果对应的投影点确定为目标投影点;

将所述逆深度候选值中, 所述目标投影点对应的逆深度值确定为所述目标采样点的逆深度值;

继续确定所述第 i 层采样点中与所述目标采样点不同的采样点的逆深度值, 直至确定出所述第 i 层采样点中每一个采样点的逆深度值, 获得所述第 i 层逆深度值。

- 30 可以理解的是, 在本公开的实施例中, 上述针对采样点匹配的过程, 实际上就是针对一个采样点, 分别确定与采用不同逆深度值投影的投影点的差异程度。选取匹配结果值最小结果, 表征对应的投影点与采样点差异度最小, 因此, 可以将该投影点采用的逆深度值确定为采样点的逆深度值, 从而得到采样点准确的逆深度值。

在上述图像深度估计方法中, 所述获得第 k 层逆深度值之后, 所述方法还包括:

对所述第 k 层逆深度值进行插值优化, 获得优化后的第 k 层逆深度值;

- 35 将所述优化后的第 k 层逆深度值确定为所述逆深度估计结果。

可以理解的是, 在本公开的实施例中, 上述过程中估计的深度为离散值, 因此, 还可以进行二次插值, 调整每个采样点的逆深度, 从而获得更为准确的逆深度值。

在上述图像深度估计方法中, 所述对所述第 k 层逆深度值进行插值优化, 获得优化后的第 k 层逆深度值, 包括:

- 40 对所述第 k 层逆深度值中每一个逆深度值, 分别从第 k 层采样点对应的采样点的候选逆深度值中, 选取相邻逆深度值; 所述第 k 层采样点为对所述 k 层当前图像中第 k 层当前图像采样获得的像素点;

获取所述相邻逆深度值对应的匹配结果;

- 45 基于所述相邻逆深度值和所述相邻逆深度值对应的匹配结果, 对所述第 k 层逆深度值中的每一个逆深度值进行插值优化, 获得所述优化后的第 k 层逆深度值。

可以理解的是, 在本公开的实施例中, 利用确定的采样点的逆深度值, 其相邻逆深度值和相邻逆深度值对应匹配结果, 可以更为精确的对采样点的逆深度值进行插值调整, 且调整方式简单快速。

本公开实施例提供了一种图像深度估计装置, 包括:

获取部分, 配置为获取当前帧对应的参考帧和所述当前帧的逆深度空间范围;

- 50 降采样部分, 配置为对所述当前帧和所述参考帧分别进行金字塔降采样处理, 获得所述当前帧对应的 k 层当前图像, 以及所述参考帧对应的 k 层参考图像; k 为大于等于 2 的自然数;

估计部分,配置为基于所述k层参考图像和所述逆深度空间范围,对所述k层当前图像进行逆深度估计迭代处理,获得所述当前帧的逆深度估计结果。

在上述图像深度估计装置中,所述获取部分,具体配置为获取至少两个待筛选帧;从所述至少两个待筛选帧中,选取与所述当前帧之间满足预设角度约束条件的至少一帧,将所述至少一帧作为所述参考帧。

在上述图像深度估计装置中,所述预设角度约束条件包括:

所述当前帧对应的位姿中心和所述参考帧对应的位姿中心,与目标点的连线形成的夹角处于第一预设角度范围;所述目标点为所述当前帧对应的平均深度点与所述参考帧对应的平均深度点连线的中点;

所述当前帧和所述参考帧对应的光轴夹角处于第二预设角度范围;

所述当前帧和所述参考帧对应的纵轴夹角处于第三预设角度范围。

在上述图像深度估计装置中,所述估计部分,具体配置为基于所述k层当前图像和所述逆深度空间范围,确定第i层采样点中每一个采样点对应的逆深度候选值;所述第i层采样点为对所述k层当前图像中第i层当前图像采样获得的像素点,i为大于等于1且小于等于k的自然数;根据所述第i层采样点中每一个采样点对应的逆深度候选值和所述k层参考图像中第i层参考图像,确定所述第i层采样点中每一个采样点的逆深度值,获得第i层逆深度值;令 $i=i+1$,继续对所述k层当前图像中分辨率高于所述第i层当前图像的第i+1层当前图像进行逆深度估计,直至 $i=k$ 为止,获得第k层逆深度值;将所述第k层逆深度值确定为所述逆深度估计结果。

在上述图像深度估计装置中,所述估计部分,具体配置为对所述逆深度空间范围进行区间划分,并在每个划分区间中选择一个逆深度值,得到多个初始逆深度值;将所述多个初始逆深度值确定为第1层采样点中每一个采样点对应的逆深度候选值;在i不等于1的情况下,从所述k层当前图像中获取第i-1层采样点,以及第i-1层逆深度值;基于所述第i-1层逆深度估计值、第i-1层采样点,以及所述多个初始逆深度值,确定所述第i层采样点中每一个采样点对应的逆深度候选值。

在上述图像深度估计装置中,所述估计部分,具体配置为从所述第i-1层采样点中确定与第一采样点距离最近的第二采样点,以及与所述第二采样点相邻的至少两个第三采样点;所述第一采样点为所述第i层采样点中任意一个采样点;根据所述第i-1层逆深度值,获取所述至少两个第三采样点中每一个采样点的逆深度值,以及所述第二采样点的逆深度值,得到的至少三个逆深度值;从所述至少三个逆深度值中,确定最大逆深度值和最小逆深度值;从所述多个初始逆深度值中,选取处于所述最大逆深度值和所述最小逆深度值范围内的逆深度值,将选取出的逆深度值确定为所述第一采样点对应的逆深度候选值;继续确定所述第i层采样点中非所述第一采样点的采样点对应的逆深度候选值,直至确定出所述第i层采样点中每一个采样点对应的逆深度候选值。

在上述图像深度估计装置中,所述估计部分,具体配置为对所述第i层采样点中每一个采样点,分别按照对应的逆深度候选值中的每一个逆深度值,将所述第i层采样点中每一个采样点投影到所述第i层参考图像中,获得所述第i层采样点中每一个采样点对应的第i层投影点;根据所述第i层采样点和所述第i层投影点进行块匹配,获得所述第i层采样点中每一个采样点对应的第i层匹配结果;根据所述第i层匹配结果,确定所述第i层采样点中每一个采样点的逆深度值,获得所述第i层逆深度值。

在上述图像深度估计装置中,所述估计部分,具体配置为利用预设窗口,从所述第i层当前图像中选取以待匹配采样点为中心的第一图像块,并从所述第i层参考图像中选取以所述待匹配采样点对应的第i层投影点中的每一个投影点分别为中心的多个第二图像块;所述待匹配采样点为所述第i层采样点中任意一个采样点;将所述第一图像块分别与所述多个第二图像块中每一个图像块进行比较,获得多个匹配结果,并将所述多个匹配结果确定为所述待匹配采样点对应的第i层匹配结果;继续确定所述第i层采样点中与所述待匹配采样点不同的采样点对应的第i层匹配结果,直至获得所述第i层采样点中每一个采样点对应的第i层匹配结果。

在上述图像深度估计装置中,所述估计部分,具体配置为从目标采样点对应的第i层匹配结果中选取目标匹配结果;所述目标采样点为所述第i层采样点中任意一个采样点;将所述目标采样点对应的第i层投影点中,所述目标匹配结果对应的投影点确定为目标投影点;将所述逆深度候选值中,所述目标投影点对应的逆深度值确定为所述目标采样点的逆深度值;继续确定所述第i层采样点中与所述目标采样点不同的采样点的逆深度值,直至确定出所述第i层采样点中每一个采样点的逆深度值,获得所述第i层逆深度值。

在上述图像深度估计装置中,所述估计部分,还配置为对所述第k层逆深度值进行插值优化,

获得优化后的第 k 层逆深度值；将所述优化后的第 k 层逆深度值确定为所述逆深度估计结果。

在上述图像深度估计装置中，所述估计部分，具体配置为对所述第 k 层逆深度值中每一个逆深度值，分别从第 k 层采样点中对应的采样点的候选逆深度值中，选取相邻逆深度值；所述第 k 层采样点为对所述 k 层当前图像中第 k 层当前图像采样获得的像素点；获取所述相邻逆深度值对应的匹配结果；基于所述相邻逆深度值和所述相邻逆深度值对应的匹配结果，对所述第 k 层逆深度值中的每一个逆深度值进行插值优化，获得所述优化后的第 k 层逆深度值。

本公开实施例提供了一种电子设备，所述电子设备包括：处理器、存储器和通信总线；其中，所述通信总线，配置为实现所述处理器和所述存储器之间的连接通信；

所述处理器，配置为执行所述存储器中存储的图像深度估计程序，以实现上述图像深度估计方法。

在上述电子设备中，所述电子设备为手机或平板电脑。

本公开实施例提供了一种计算机可读存储介质，所述计算机可读存储介质存储有一个或者多个程序，所述一个或者多个程序可以被一个或者多个处理器执行，以实现上述图像深度估计方法。

本公开实施例提供了一种计算机程序，包括计算机可读代码，所述计算机可读代码被处理器执行时，实现上述图像深度估计方法对应的步骤。

由此可见，本公开实施例的技术方案中，获取当前帧对应的参考帧和当前帧的逆深度空间范围；对当前帧和参考帧分别进行金字塔降采样处理，获得当前帧对应的 k 层当前图像，以及参考帧对应的 k 层参考图像；k 为大于等于 2 的自然数；基于 k 层参考图像和逆深度空间范围，对 k 层当前图像进行逆深度估计迭代处理，获得当前帧的逆深度估计结果。也就是说，本公开提供的技术方案，采取了对多层当前图像结合多层参考图像进行逆深度估计迭代处理，以逐层减少逆深度搜索空间，确定当前帧的逆深度估计结果，该逆深度估计结果为当前帧的像素点在相机坐标系下的 z 轴坐标值的倒数，不需要额外进行坐标变换，且逐层减少逆深度搜索空间有助于减少逆深度估计的计算量，提升估计速度，从而能够实时获得图像的逆深度估计结果，且逆深度估计结果的精确度较高。

附图说明

图 1 为本公开实施例提供的一种图像深度估计方法的流程示意图；
图 2 为本公开实施例提供的一种示例性的相机位姿夹角的示意图；
图 3 为本公开实施例提供的一种逆深度估计迭代处理的流程示意图一；
图 4 为本公开实施例提供的一种示例性的 3 层当前图像的示意图；
图 5 为本公开实施例提供的一种确定逆深度候选值的流程示意图；
图 6 为本公开实施例提供的一种示例性的采样点投影示意图；
图 7 为本公开实施例提供的一种逆深度估计迭代处理的流程示意图二；
图 8 为本公开实施例提供的一种图像深度估计装置的结构示意图；
图 9 为本公开实施例提供的一种电子设备的结构示意图。

具体实施方式

下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述。

本公开实施例提供了一种图像深度估计方法，其执行主体可以是图像深度估计装置，例如，图像深度估计方法可以由终端设备或服务器或其它电子设备执行，其中，终端设备可以为用户设备（User Equipment, UE）、移动设备、用户终端、终端、蜂窝电话、无绳电话、个人数字处理（Personal Digital Assistant, PDA）、手持设备、计算设备、车载设备、可穿戴设备等。在一些可能的实现方式中，该图像深度估计方法可以通过处理器调用存储器中存储的计算机可读指令的方式来实现。图 1 为本公开实施例提供的一种图像深度估计方法的流程示意图。如图 1 所示，主要包括以下步骤：

S101、获取当前帧对应的参考帧和当前帧的逆深度空间范围。

在本公开的实施例中，执行主体以图像深度估计装置为例进行说明。首先，图像深度估计装置对当前帧进行深度估计时，需要先获取到当前帧对应的参考帧和当前帧的逆深度空间范围。

需要说明的是，在本公开的实施例中，当前帧为需要进行深度估计的图像，而参考帧为在对当前帧进行深度估计时，用于进行参考匹配的图像，参考帧的数量可以为多个，考虑到深度估计的速度和鲁棒性的平衡，选取 5 个左右的参考帧较为合适，具体的当前帧的参考帧本公开实施例不作限

定。

具体的，在本公开的实施例中，图像深度估计装置获取当前帧对应的参考帧包括以下步骤：获取至少两个待筛选帧；从至少两个待筛选帧中，选取与当前帧之间满足预设角度约束条件的至少一帧，将该至少一帧作为参考帧。

- 5 需要说明的是，在本公开的实施例中，图像深度估计装置还可以以其它方式获取参考帧，例如，接收用户发送的针对至少两个待筛选帧的选择指令，将选择指令指示的至少一帧作为参考帧。具体的参考帧获取方式本申请实施例不作限定。

- 10 需要说明的是，在本公开的实施例中，图像深度估计装置从至少两个待筛选帧中，选取出的当前帧对应的参考帧可以为多个，而每一个参考帧均与当前帧之间满足预设角度约束条件。待筛选帧即为与当前帧针对同一场景，但是不同角度下获取到的图像。图像深度估计装置可以配置有摄像模块，通过该摄像模块可以获取到待筛选帧，当然，也可以通过其它独立的摄像设备先获取待筛选帧，图像深度估计装置再从摄像设备中进一步获取待筛选帧。具体的预设角度约束条件可以根据实际深度估计需求预先设置在图像深度估计装置中，也可以存储在其他装置中，需要进行深度估计时从其他装置中获取，又或者可以通过接收用户输入的角度约束条件获取等，本公开实施例不作限定。

- 15 具体的，在本公开的实施例中，预设角度约束条件包括：当前帧对应的位姿中心和参考帧对应的位姿中心，与目标点的连线形成的夹角处于第一预设角度范围；目标点为当前帧对应的平均深度点与参考帧对应的平均深度点连线的中点；当前帧和参考帧对应的光轴夹角处于第二预设角度范围；当前帧和参考帧对应的纵轴夹角处于第三预设角度范围。其中，纵轴即为三维空间中相机坐标系的Y轴。

- 20 在本公开的一些实施例中，当前帧对应的位姿中心，实际上就是相机在处于获取当前帧时的位置和姿态下，相机的中心（光心）。参考帧对应的位姿中心，实际上就是相机在处于获取参考帧时的位置和姿态下，相机的中心（光心）。

- 25 示例性的，在本公开的实施例中，如图2所示，定义获取当前帧时相机的位姿为位姿1，获取参考帧时相机的位姿为位姿2，位姿1时相机的中心（光心）到对应场景的平均深度点为点P1，位姿2时相机的中心（光心）到对应场景的平均深度点为点P2，P1和P2的连线中点为点P，预设角度预设条件具体包括三个角度条件：第一个角度条件为，位姿1和位姿2时相机中心与P点的连线形成的视角 α 在 $[5^\circ, 45^\circ]$ 之间；第二个角度条件为相机处于位姿1和位姿2时的光轴夹角在 $[0^\circ, 45^\circ]$ 之间；第三个角度条件为相机处于位姿1和位姿2时的Y轴夹角在 $[0^\circ, 30^\circ]$ 之间，只有同时满足了这三个角度条件的帧才能作为参考帧。以上角度区间实际中都可以进行调整。

- 30 需要说明的是，在本公开的实施例中，获取当前帧和参考帧的相机可以配置有定位装置，从而在获取当前帧和参考帧时直接获取到相应的位姿，图像深度估计装置可以获取定位装置中获得的相关位姿，当然，图像深度估计装置还可以按照位姿估计算法，结合获得的当前帧和参考帧中的一些特征点，计算出相应的位姿。

- 35 可以理解的是，在本公开的实施例中，第一个角度条件限定了当前场景到两个相机的距离，角度过大说明场景过近，两帧重合度会较低，角度过小，则说明场景过远，视差较小，误差会比较大，当相机特别接近时也可能发生角度过小的情况，此时误差同样较大。第二角度条件是为了保证两个相机有足够的共视区域。第三个角度条件是为了避免相机绕着光轴旋转，影响后续深度估计计算过程。同时满足上述三个角度条件的帧作为参考帧有利于提高当前帧深度估计的精度。

- 40 需要说明的是，在本公开的实施例中，图像深度估计装置可以根据当前帧直接获取到对应的逆深度空间范围，逆深度空间范围为当前帧中的像素点的逆深度值可取的空间范围，当然，图像深度估计装置还可以接收用户的设置指令，根据设置指令获取用户指示的逆深度空间范围。具体的逆深度空间范围本公开实施例不作限定。例如，逆深度空间范围为 $[d_{\min}, d_{\max}]$ ， $d_{\min}$ 为逆深度空间范围内最小的逆深度值， $d_{\max}$ 为逆深度空间范围内最大的逆深度值。

- 45 S102、对当前帧和参考帧分别进行金字塔降采样处理，获得当前帧对应的k层当前图像，以及参考帧对应的k层参考图像；k为大于等于2的自然数。

在本公开的实施例中，图像深度估计装置在获取到当前帧对应的参考帧之后，可以对当前帧和参考帧分别进行金字塔降采样处理，从而获得当前帧对应的k层当前图像，以及参考帧对应的k层参考图像。

- 50 需要说明的是，在本公开的实施例中，由于参考帧可以为多个，因此，图像深度估计装置对于每一个参考帧图像分别进行金字塔降采样处理，从而获得的k层参考图像实际上为多组，具体的k层参考图像的数量本公开实施例不作限定。

需要说明的是,在本公开的实施例中,图像深度估计装置对当前帧和参考帧分别进行金字塔降采样处理,获得的当前图像金字塔和参考图像金字塔的层数是相同的,采用的尺度因子也是相同的。例如,图像深度估计装置对当前帧和参考帧分别进行尺度因子为2的降采样,形成三层当前图像和三层参考图像,在这两组三层图像中,顶层图像的分辨率最低,中间层图像的分辨率高于顶层图像的分辨率,底层图像的分辨率最高,实际上,底层图像也就是原图像,即对应的当前帧和参考帧。具体的图像层数 k ,以及降采样的尺度因子可以根据实际需求预先设置,本公开实施例不限定。

示例性的,在本公开的实施例中,图像深度估计装置获取到了当前帧 I_t 对应的5个参考帧,分别为:参考帧 I_1 、参考帧 I_2 、参考帧 I_3 、参考帧 I_4 ,以及参考帧 I_5 ,图像深度估计装置对这些帧分别进行尺度因子为2的降采样,从而获取当前帧 I_t 对应的3层当前图像,以及参考帧 I_1 、参考帧 I_2 、参考帧 I_3 、参考帧 I_4 和参考帧 I_5 各自对应的三层参考图像。

S103、基于 k 层参考图像和逆深度空间范围,对 k 层当前图像进行逆深度估计迭代处理,获得当前帧对应的逆深度估计结果。

在本公开的实施例中,图像深度估计装置在获得 k 层当前图像和 k 层参考图像之后,可以基于 k 层参考图像和逆深度空间范围,对 k 层当前图像进行逆深度估计迭代处理,例如可以从顶层(第1层)当前图像(像素最少的图像)开始,依次向底层进行逆深度估计迭代,逐层缩小逆深度搜索空间,直到最底层第 k 层,获得当前帧对应的逆深度估计结果。

图3为本公开实施例提供的一种逆深度估计迭代处理的流程示意图一。如图3所示,图像深度估计装置基于 k 层参考图像和逆深度空间范围,对 k 层当前图像进行逆深度估计迭代处理,获得当前帧对应的逆深度估计结果,包括如下步骤:

S301、基于 k 层当前图像和逆深度空间范围,确定第 i 层采样点中每一个采样点对应的逆深度候选值;第 i 层采样点为对 k 层当前图像中第 i 层当前图像采样获得的像素点, i 为大于等于1且小于等于 k 的自然数。

在本公开的实施例中, k 层当前图像按照分辨率由低到高依次包括:第1层当前图像、第2层当前图像、第3层当前图像,……,第 k 层当前图像,第1层当前图像为 k 层当前图像中的顶层图像,第 k 层当前图像为当前图像金字塔中的底层图像,同样的, k 层参考图像中按照分辨率由低到高依次包括:第1层参考图像、第2层参考图像、第3层参考图像,……,第 k 层参考图像,第1层参考图像为参考图像金字塔中的顶层图像,第 k 层参考图像为参考图像金字塔中的底层图像。

需要说明的是,在本公开的实施例中,图像深度估计装置可以对 k 层当前图像中第 i 层当前图像进行像素点采样,采样获得的像素点即为第 i 层采样点,具体的 i 的取值为大于1且小于等于 k 的自然数,本公开实施例不限定。

需要说明的是,在本公开的实施例中,图像深度估计装置对第 i 层当前图像进行像素点采样,可以按照预设的采样步长来实现。具体的采样步长可以根据实际需求确定,本公开实施例不限定。

图4为本公开实施例提供的一种示例性的3层当前图像的示意图。如图4所示,图像深度估计装置可以预先对当前帧,在 x 轴和 y 轴坐标按照采样步长为2进行像素点采样,共获得3层当前图像,其中,第1层当前图像分辨率最低,第2层当前图像分辨率高于第1层当前图像,第3层当前图像分辨率高于第2层当前图像,第3层当前图像实际上就是当前帧原图。

具体的,在本公开的实施例中,图像深度估计装置基于 k 层当前图像和逆深度空间范围,确定第 i 层采样点中每一个采样点对应的第 i 层逆深度候选值,包括:当 i 等于1时,对逆深度空间范围进行区间等分,获得划分区间的多个等分逆深度值;将多个等分逆深度值确定为第1层采样点中每一个采样点对应的逆深度候选值;当 i 不等于1时,从 k 层当前图像中获取第 $i-1$ 层采样点,以及第 $i-1$ 层逆深度估计值;基于第 $i-1$ 层逆深度估计值、第 $i-1$ 层采样点,以及多个等分逆深度值,确定第 i 层采样点中每一个采样点对应的逆深度候选值。

可以理解的是,在本公开的实施例中,图像深度估计装置针对逆深度空间范围进行区间划分,从而在不同区间内选取逆深度值,可以使得每一个区间内都存在一个逆深度值作为逆深度候选值。也就是说,每一个采样点在不同逆深度范围内都存在一个逆深度候选值,在后续进行确定采样点的逆深度值,可以保证不同逆深度范围的逆深度值都可以进行逆深度值估计确定,保证估计过程覆盖整个逆深度空间范围,从而最终可以估计出准确的逆深度值。

可以理解的是,在本公开的实施例中,当 i 等于1时,即图像深度估计装置需要确定第1层采样点中每一个采样点对应的逆深度候选值,其中,第1层采样点为 k 层当前图像中分辨率最低的第1层当前图像中的采样点,图像深度估计装置获取到当前帧对应的逆深度空间范围为 $[dmin, dmax]$,可以将其进行等分,获得进行划分区间的 q 个等分逆深度值 $d1$ 、 $d2$,……, dq ,可以将这 q 个等分

逆深度值均确定为初始逆深度值,也就是第1层采样点中每一个采样点对应的逆深度候选值,当然,逆深度候选值还可以包括 $d_{\min}$ 和 $d_{\max}$ 。即对于第1层采样点中每一个采样点而言,其对应的逆深度候选值是完全相同的。图像深度估计装置对逆深度空间范围的等分区间可以根据实际需求进行设置,本公开实施例不作限定。

5 需要说明的是,在本公开的实施例中,图像深度估计装置如果按照上述等分的方式对逆深度空间范围进行区间划分,并将划分区间的逆深度值作为逆深度候选值,可以保证逆深度候选值均匀覆盖整个逆深度空间范围,保证后续从逆深度候选值中确定的逆深度值更加准确。

需要说明的是,在本公开的实施例中,在 i 等于1的情况下,除了针对逆深度空间范围进行等分的方式进行划分,还可以以非等分方式进行划分。例如,以预先设置的多个不同间隔依次对逆深度空间范围进行划分,或者,基于预设的初始划分间隔,结合间隔变化规则,每划分一次进行间隔的调整,再利用调整后的间隔进行下一区间的划分。当然,初始逆深度值的选取也可以直接在划分的区间中随机选取一个逆深度值,也可以选取每个划分区间的中间的逆深度值。具体的区间划分方式和初始逆深度值选取方式本公开实施例不作限定。

需要说明的是,在本公开的实施例中,在 i 不等于1的情况下,图像深度估计装置需要从 k 层当前图像中获取第 $i-1$ 层采样点,也就是对 k 层当前图像中,第 $i-1$ 层采样点进行采样所获得的像素点,并且,还需要获取第 $i-1$ 层逆深度值。每一层当前图像都可以以不同的采样步长进行采样。其中,在确定第 i 层采样点中每一个采样点对应的逆深度候选值之前,在 $i=i-1$ 的情况下,图像深度估计装置已经按照上述逆深度估计步骤获得了第 i 层逆深度值,也就是第 $i-1$ 层采样点中每一个采样点的逆深度值。因此,图像深度估计装置可以直接获取到第 $i-1$ 层逆深度值,并进一步的根据第 $i-1$ 层逆深度值、第 $i-1$ 层采样点,以及多个等分逆深度值,确定第 i 层采样点中每一个采样点对应的逆深度候选值。

图5为本公开实施例提供的一种确定逆深度候选值的流程图。如图5所示,图像深度估计装置基于第 $i-1$ 层逆深度估计值、第 $i-1$ 层采样点,以及多个初始逆深度值,确定第 i 层采样点中每一个采样点对应的逆深度候选值,包括:

25 S501、从第 $i-1$ 层采样点中确定与第一采样点距离最近的第二采样点,以及与第二采样点相邻的至少两个第三采样点;第一采样点为第 i 层采样点中任意一个采样点。

S502、根据第 $i-1$ 层逆深度值,获取至少两个第三采样点中每一个采样点的逆深度值,以及第二采样点的深度值,得到至少三个逆深度值。

S503、从至少三个逆深度值中,确定最大逆深度值和最小逆深度值。

30 S504、从多个初始逆深度值中,选取处于最大逆深度值和最小逆深度值范围内的逆深度值,将选取出的逆深度值确定为第一采样点对应的逆深度候选值。

S505、继续确定第 i 层采样点中非第一采样点的采样点对应的逆深度候选值,直至确定出第 i 层采样点中每一个采样点对应的逆深度候选值。

需要说明的是,在本公开的实施例中,在 i 等于1的情况下,第 i 层采样点,即第1层采样点中每一个采样点对应的逆深度候选值,均相同,而在 i 不等于1的情况下,第 i 层采样点中,每一个采样点对应的第 i 层逆深度候选值可以根据第 $i-1$ 层采样点和第 $i-1$ 层逆深度值,从多个初始逆深度值中进行选取,确定出范围较小的逆深度候选值,且第 i 层采样点中每一个采样点对应的逆深度候选值可能均不相同。

示例性的,在本公开的实施例中, x_i^j 为第 i 层采样点中任意一个采样点,图像深度估计装置可以在第 $i-1$ 层采样点中查找出距离 x_i^j 最近的采样点 x_{i-1}^{j-1} ,从而从第 $i-1$ 层采样点中以 x_{i-1}^{j-1} 为中心,确定其相邻的多个(例如8个)采样点,之后,根据第 $i-1$ 层逆深度值,获取 x_{i-1}^{j-1} ,以及与其相邻的8个采样点中每一个采样点的逆深度值,即获得9个逆深度值,进一步的,将9个逆深度值中以最大的逆深度值 d_1 和最小的逆深度值 d_2 为界限,将多个初始逆深度值中 d_1 和 d_2 之间的深度值选取出来,包括 d_1 和 d_2 ,均确定为 x_i^j 对应的候选逆深度值。

45 需要说明的是,在本公开的实施例中,图像深度估计装置从第 $i-1$ 层采样点中确定与第二采样点相邻的第三采样点,可以将其周围的8个采样点均确定为第三采样点,当然,也可以将其左右相邻的2个采样点,或者上下相邻的2个采样点确定为第三采样点,还可以将其上下左右相邻的4个采样点均确定为第三采样点,具体的第三采样点的数量本公开实施例不作限定。

需要说明的是,在本公开的实施例中,图像深度估计装置还可以按照其它的规则确定第 i 层采样点中每一个采样点对应的逆深度候选值。例如,接收用户设置的针对不同层采样点设置的不同逆深度候选值,同一层采样点中每一个采样点对应的逆深度候选值相同。具体的逆深度候选值确定方

式本公开实施例不作限定。

S302、根据第 i 层采样点中每一个采样点对应的逆深度候选值和 k 层参考图像中第 i 层参考图像，确定第 i 层采样点中每一个采样点的逆深度值，获得第 i 层逆深度值。

具体的，在本公开的实施例中，图像深度估计装置根据第 i 层采样点中每一个采样点对应的逆深度候选值和 k 层参考图像中第 i 层参考图像，确定第 i 层采样点中每一个采样点的逆深度值，获得第 i 层逆深度值，包括：对第 i 层采样点中每一个采样点，分别按照对应的逆深度候选值中的每一个逆深度值，将第 i 层采样点中每一个采样点投影到第 i 层参考图像中，获得第 i 层采样点中每一个采样点对应的第 i 层投影点；根据第 i 层采样点和第 i 层投影点进行块匹配，获得第 i 层采样点中每一个采样点对应的第 i 层匹配结果；根据第 i 层匹配结果，确定第 i 层采样点中每一个采样点的逆深度值，获得第 i 层逆深度值。

需要说明的是，在本公开的实施例中，图像深度估计装置对第 i 层采样点中的每一个采样点，均按照对应的逆深度候选值中的每一个逆深度值投影到第 i 层参考图像中。当然，如果有多个参考帧，相应的，有多个第 i 层参考图像，那么图像深度估计装置是将第 i 层采样点中的每一个采样点，分别按照对应的逆深度候选值中的每一个逆深度值，分别投影到每一个第 k 层参考图像中。

具体的，在本公开的实施例中，对于当前帧 t 和参考帧 r ，图像深度估计装置对于第 i 层采样点中的任意一个采样点 $x_t^i = [u, v]$ ， u 和 v 为该采样点的 x 轴和 y 轴坐标，对于 x_t^i 对应的逆深度候选值中的任意一个逆深度值 d_z ，按照以下公式 (1) 和公式 (2) 进行投影到第 k 层参考图像中：

$$X_r = KR_r \left(d_z^{-1} K^{-1} [x_t^i, 1]^T \right) + KT_r, \quad K = \begin{bmatrix} f_x^i & 0 & c_x^i \\ 0 & f_y^i & c_y^i \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

$$x_r^i = [X_r(0)/X_r(2), X_r(1)/X_r(2)] \quad (2)$$

需要说明的是， K 为获取当前帧 t 和参考帧 r 的相机对应的相机内参矩阵， f_x^i 和 f_y^i 为第 i 层当前图像对应的焦距在 x 轴和 y 轴上的基于像素度量的尺度因子， f_x^i 为使用像素来描述的 x 轴方向焦距的长度， f_y^i 为使用像素来描述的 y 轴方向焦距的长度。 (c_x^i, c_y^i) 为第 i 层当前图像的主点位置， R_r 为 3×3 的旋转矩阵， T_r 为 3×1 的平移向量。公式 (1) 最终获得的 X_r 是一个 3×1 的矩阵，其中，第一行元素为 $X_r(0)$ ，第二行元素为 $X_r(1)$ ，第三行元素为 $X_r(2)$ ，按照公式 (2) 进一步计算，即可获得采样点 x_t^i 按照对应的逆深度候选值中的逆深度值 d_z ，投影到参考帧 r 中第 i 层参考图像中的投影点 x_r^i 。

可以理解的是，在本公开的实施例中，对于第 i 层采样点中的每一个采样点，均可以通过公式 (2) 和公式 (3)，按照对应逆深度候选值中的每一个逆深度值投影到第 i 层参考图像中，如果是多个第 i 层参考图像，重复执行即可。

需要说明的是，在本公开的实施例中，图像深度估计装置在获得第 i 层投影点之后，可以根据第 i 层采样点和第 i 层投影点进行块匹配，具体是对第 i 层采样点中的每一个采样点，与对应的第 i 层投影点中的每一个投影点分别进行块匹配，从而获得每一个采样点对应的第 i 层匹配结果。

具体的，在本公开的实施例中，图像深度估计装置根据第 i 层采样点和第 i 层投影点进行块匹配，获得第 i 层采样点中每一个采样点对应的第 i 层匹配结果，包括：利用预设窗口，从第 i 层当前图像中选取以待匹配采样点为中心的第一图像块，并从第 i 层参考图像中选取以待匹配采样点对应的第 i 层投影点中的每一个投影点分别为中心的多个第二图像块；待匹配采样点为第 i 层采样点中任意一个采样点；将第一图像块分别与多个第二图像块中每一个图像块进行比较，获得多个匹配结果，并将多个匹配结果确定为待匹配采样点对应的第 i 层匹配结果；继续确定第 i 层采样点中与待匹配采样点不同的采样点对应的第 i 层匹配结果，直至获得第 i 层采样点中每一个采样点对应的第 i 层匹配结果。例如，采用一个 3×3 的窗口，在第 i 层当前图像和第 i 层参考图像中，分别以第 i 层采样点中的每一个采样点和其对应的投影点为中心，获取采样点和投影点的邻域点，得到两个图像块，然后对获取的图像块中对应位置的像素点的像素值进行比较，得到两个图像块的匹配的惩罚值（如像素差值的绝对值之和）。针对同一逆深度值，每个第 i 层参考图像，可以得到一个惩罚值；存在多个第 i 层参考图像时，对得到的多个惩罚值进行融合（例如多个惩罚值取平均），即可以得到每个采样点对应一个逆深度值的第 i 层匹配结果。针对每个采样点的多个逆深度值，均可以得到每个逆深度值对应的一个惩罚值，即得到每个采样点对应的第 i 层匹配结果。

具体的，在本公开的实施例中，如图 6 所示，对于当前帧 t 和 m 个参考帧， m 为大于等于 1 的自然数，图像深度估计装置对于第 i 层采样点中的任意一个采样点 $x_t^i = [u, v]$ ，按照以下公式 (3)

所示,与对应的第 i 层投影点中以逆深度值为 d_z 投影获得的投影点 x_t^i 进行块匹配,从而获得第 i 层匹配结果中逆深度值为 d_z 的匹配结果:

$$C(x_t^i, d_z) = \sum_{r=1}^m f(x_t^i, x_r^i) / m \quad (3)$$

其中, x_t^i 为 x_t^i 按照自身对应的候选逆深度值中的逆深度值 d_z , 分别投影到 m 个参考帧中每一帧分别对应的第 i 层参考图像中的投影点, 共计 m 个。 $f(x_t^i, x_r^i)$ 为 x_t^i 和 x_r^i 的邻域像素值比较函数, 该比较函数可以是 x_t^i 和 x_r^i 的邻域灰度值的零均值归一化协方差 (Zero-mean Normalized Cross Correlation, ZNCC), 也可以使用绝对差之和 (Sum of absolute differences, SAD) 或差方和 (Sum of Squared Differences, SSD) 两种方法。 $C(x_t^i, d_z)$ 即为 x_t^i 对应的第 i 层匹配结果中, 逆深度值为 d_z 的匹配结果。

需要说明的是, 在本公开的实施例中, 第 i 层采样点中, 每一个采样点对应的第 i 层匹配结果均包括了自身对应的逆深度候选值中, 不同逆深度值的匹配结果, 例如, 对于第 i 层采样点中的任意一个采样点 x_t^i , 对应的逆深度候选值包括 $d_1, d_2, \dots, d_q$, 获得的第 i 层匹配结果包括每一个逆深度值的匹配结果, 具体的第 i 层匹配结果本公开实施例不限定。

示例性的, 在本公开的实施例中, 当前帧对应的参考帧包括 2 个帧, 每一个帧对应有一组 2 层参考图像, 即有两个第 1 层参考图像, 图像深度估计装置将当前帧中第 1 层当前图像的一个采样点 x_t^1 , 按照其对应的逆深度候选值 d_1, d_2 和 d_3 分别投影到两个第 1 层参考图像中, 分别在两个第 1 层参考图像中获得三个投影点, 共 6 个投影点, 作为其对应的第 1 层投影点。 其中, 按照 d_1 投影到一个第 1 层参考图像的投影点为 x_1^1 , 按照 d_1 投影到另一个第 1 层参考图像的投影点为 x_2^1 , 因此, 可以将 x_t^1, x_1^1 和 x_2^1 代入公式 (3) 中, 即 m 等于 2, 获得 x_t^1 对逆深度值为 d_1 的匹配结果, 同样的, 也可以获得逆深度候选值为 d_2 和 d_3 的匹配结果, 组成 x_t^1 对应的第 i 层匹配结果。

具体的, 在本公开的实施例中, 图像深度估计装置根据第 i 层匹配结果, 确定第 i 层采样点中每一个采样点的逆深度值, 获得第 i 层逆深度值, 包括: 从目标采样点对应的第 i 层匹配结果中选取目标匹配结果; 目标采样点为第 i 层采样点中任意一个采样点; 将目标采样点对应的第 i 层投影点中, 目标匹配结果对应的投影点确定为目标投影点; 将逆深度候选值中, 目标投影点对应的逆深度值确定为目标采样点的逆深度值; 继续确定第 i 层采样点中与目标采样点不同的采样点的逆深度值, 直至确定出第 i 层采样点中每一个采样点的逆深度值, 获得第 i 层逆深度值。

需要说明的是, 在本公开的实施例中, 图像深度估计装置在获得第 i 层采样点中每一个采样点对应的第 i 层匹配结果之后, 可以按照以下公式 (4) 确定第 i 层采样点中任意一个采样点 x_t^i 的逆深度值:

$$d(x_t^i) = \operatorname{argmin}_{\arg=d_z} C(x_t^i, d_z) \quad (4)$$

其中, 由于 x_t^i 对应的第 i 层匹配结果中逆深度值为 d_z 的匹配结果 $C(x_t^i, d_z)$, 相比与其它逆深度值的匹配结果值最小, 因此, 将对应的逆深度值 d_z 实际上确定为 x_t^i 的逆深度值。

可以理解的是, 在本公开的实施例中, 上述针对采样点匹配的过程, 实际上就是针对一个采样点, 分别确定与采用不同逆深度值投影的投影点的差异程度, 而采用公式 (4) 进行逆深度值的确定, 实际上就是选取匹配结果值最小结果, 表征对应的投影点与采样点差异度最小, 因此, 可以将该投影点采用的逆深度值确定为采样点的逆深度值, 从而得到采样点准确的逆深度值。

需要说明的是, 在本公开的实施例中, 图像深度估计方法还可以以其它方式确定第 i 层采样点中每一个采样点的逆深度值。 例如, 从每一个采样点对应的匹配结果中选取处于特定范围的部分结果, 之后, 从部分结果中随机选取一个匹配结果, 将随机选取出的匹配结果对应的投影点采用的逆深度值确定为采样点的逆深度值。

S303、令 $i=i+1$, 继续对 k 层当前图像中分辨率高于第 i 层当前图像的第 $i+1$ 层当前图像进行逆深度估计, 直至 $i=k$ 为止, 获得第 k 层逆深度值。

在本公开的实施例中, 图像深度估计装置获得第 i 层逆深度值之后, 令 $i=i+1$, 从而进一步继续对第 i 层当前图像的第 $i+1$ 层当前图像进行逆深度估计, 其过程与获取第 i 层逆深度值相同, 在此不再赘述, 在不断迭代估计过程中, 直至 $i=k$ 时, 图像深度估计装置获得了第 k 层逆深度值, 即 k 层当前图像中分辨率最高的图像, 实际就是当前帧原图中每一个采样点的逆深度值, 则停止令 $i=i+1$ 。

S304、将第 k 层逆深度值确定为逆深度估计结果。

在本公开的实施例中, 图像深度估计装置在获得第 k 层逆深度值之后, 即可将第 k 层逆深度值确定为逆深度估计结果。

可选地, 上述过程中估计的深度为离散值, 为获得更为准确的逆深度, 还可以进行二次插值,

调整每个采样点的逆深度。具体地,如图7所示,在步骤S303之后还可以包括S305~S306:

S305、对第k层逆深度值进行插值优化,获得逆深度估计结果。

在本公开的实施例中,图像深度估计装置在获得第k层逆深度值之后,第k层逆深度值包括第k层采样点中每一个采样点对应的逆深度值,而为了获得更准确的第k层逆深度值,可以对第k层逆深度值进行插值优化,也就是将第k层采样点中每一个采样点的逆深度值分别进行调整优化,从而获得优化后的第k层逆深度值。

具体的,在本公开的实施例中,图像深度估计装置对第k层逆深度值进行插值优化,获得优化后的第k层逆深度值,包括:对第k层逆深度值中每一个逆深度值,分别从第k层采样点对应的采样点的候选逆深度值中,选取逆深度值的相邻逆深度值;第k层采样点为对k层当前图像中第k层当前图像采样获得的像素点;获取相邻逆深度值对应的匹配结果;基于相邻逆深度值和相邻逆深度值对应的匹配结果,对第k层逆深度值中的每一个逆深度值进行插值优化,获得优化后的第k层逆深度值。

具体的,在本公开的实施例中,第k层逆深度值包括第k层采样点中每一个采样点对应的逆深度值,图像深度估计装置需要对第k层采样点中每一个采样点对应的逆深度值进行插值优化,从而获得插值优化结果,作为当前帧的逆深度估计结果。其中,对第k层采样点中任意一个采样点 x_a^k ,若其对应的逆深度值为 d_z ,可以按照公式(5)进行插值优化:

$$d_{opt} = d_z + 0.5 \times (d_z - d_{z-1}) \times (C_{z+1} - C_{z-1}) / (C_{z+1} + C_{z-1} - 2 \times C_z) \quad (5)$$

其中, d_{z-1} 为采样点 x_a^k 对应的逆深度候选值中,与 d_z 相邻的前一个逆深度值。 C_{z+1} 为 $C(x_a^k, d_{z+1})$, C_{z-1} 为 $C(x_a^k, d_{z-1})$, C_z 为 $C(x_a^k, d_z)$,均可在计算 x_a^k 的逆深度值时,通过公式(3)计算得到, d_{z+1} 和 d_{z-1} 为 x_a^k 对应的候选逆深度值中 d_z 相邻的两个逆深度值,在此不再赘述。

可以理解的是,在本公开的实施例中,图像深度估计装置按照公式(5)对第k层逆深度值进行插值优化,由于k层当前图像中,第k层当前图像实际上就是当前帧,即实际上在获得了当前帧中每一个采样点的逆深度值之后,进一步对其进行了优化,从而获得了当前帧中每一个采样点更为准确的逆深度值,即获得了当前帧的逆深度估计结果。在本公开的实施例中,图像深度估计装置还可以获取三个或者更多个相邻逆深度值及其对应的匹配结果,利用与公式(5)类似的多项式进行插值优化。此外,图像深度估计装置还可以针对第k层采样点中每一个采样点的逆深度值,获取其对应逆深度候选值中与确定的逆深度值相邻的两个深度值,并将这三个逆深度值的均值作为采样点最终的逆深度值,实现逆深度值的优化。

S306、将优化后的第k层逆深度值确定为逆深度估计结果。

在本公开的实施例中,图像深度估计装置在获得优化后的第k层逆深度值之后,即可将优化后的第k层逆深度值确定为逆深度估计结果。

可选的,在本公开的实施例中,图像深度估计装置在确定出逆深度估计结果之后,即步骤S103之后,还可以执行以下步骤:

S104、根据逆深度估计结果,确定当前帧的深度估计结果。

在本公开的实施例中,图像深度估计装置在获得当前帧的逆深度估计结果之后,即可根据逆深度估计结果,确定当前帧的深度估计结果;该深度估计结果可用于实现基于当前帧的三维场景构建。

需要说明的是,在本公开的实施例中,对于一个采样点而言,其逆深度值和深度值互为倒数,因此,图像深度估计装置在获得当前帧的逆深度估计结果,即当前帧中每一个采样点插值优化后的逆深度值之后,分别取其倒数即可获得对应的深度值,从而获得当前帧的深度估计结果。例如,当前帧中某一个采样点插值优化后的逆深度值为A,则其深度值为 $1/A$ 。

需要说明的是,在本公开的实施例中,相比于现有技术中需要进行三角化反求解等计算才能获得相机坐标系下的z轴坐标值,上述图像深度估计方法所确定最终的深度估计结果为当前帧的采样点在相机坐标系下的z轴坐标值,不需要额外进行坐标变换。

需要说明的是,在本公开的实施例中,上述图像深度估计方法可以应用在实现基于当前帧的三维场景构建过程中。例如,用户利用移动设备摄像头拍摄某个场景时,可以利用上述图像深度估计方法获得当前帧的深度估计结果,进而重建视频场景的3D结构;用户点击移动设备中视频的当前帧中的某个位置时,可以利用上述图像深度估计方法确定的当前帧的深度估计结果,进行点击位置的视线求交找到锚点摆放虚拟物体,从而实现虚拟物体和真实场景几何一致性融合的增强现实效果;单目视频中可以利用上述图像深度估计方法恢复出三维场景结构,计算真实场景和虚拟物体之间的遮挡关系,从而实现虚拟物体和真实场景遮挡一致性融合的增强现实效果;单目视频中可以利用上述图像深度估计方法恢复出场景三维结构,获得具有真实感的阴影效果,从而实现虚拟物体和真实

场景光照一致性融合的增强现实效果；单目视频中可以利用上述图像深度估计方法恢复出场景三维结构，与虚拟动画角色之间的物理碰撞，从而实现虚拟动画角色和真实场景物理一致性融合的真实感动画效果。

- 5 此外，本公开实施例中，也可以不执行上述步骤 S104，该逆深度估计结果可以用于非三维场景建立的其它图像处理。例如，直接输出图像采样点的深度信息变化值，至其它设备进行目标识别或三维点距离计算等数据处理。

- 10 本公开实施例提供了一种图像深度估计方法，获取当前帧对应的参考帧和当前帧的逆深度空间范围；对当前帧和参考帧分别进行金字塔降采样处理，获得当前帧对应的 k 层当前图像，以及参考帧对应的 k 层参考图像；k 为大于等于 2 的自然数；基于 k 层参考图像和逆深度空间范围，对 k 层当前图像进行逆深度估计迭代处理，获得当前帧的逆深度估计结果。也就是说，本公开提供的技术方案，采取了对多层当前图像结合多层参考图像进行逆深度估计迭代处理，以逐层减少逆深度搜索空间，确定当前帧的深度估计结果，且最终的深度估计结果为当前帧的像素点在相机坐标系下的 z 轴坐标值，不需要额外进行坐标变换，从而能够实时获得图像的深度估计结果，且深度估计结果的精确度较高。

- 15 本公开实施例还提供了一种图像深度估计装置，图 8 为本公开实施例提供的一种图像深度估计装置的结构示意图。如图 8 所示，包括：

获取部分 801，配置为获取当前帧对应的参考帧和所述当前帧的逆深度空间范围；

降采样部分 802，配置为对所述当前帧和所述参考帧分别进行金字塔降采样处理，获得所述当前帧对应的 k 层当前图像，以及所述参考帧对应的 k 层参考图像；k 为大于等于 2 的自然数；

- 20 估计部分 803，配置为基于所述 k 层参考图像和所述逆深度空间范围，对所述 k 层当前图像进行逆深度估计迭代处理，获得所述当前帧的逆深度估计结果；

可选地，本公开实施例的图像深度估计装置还可以包括：确定部分 804，配置为根据所述逆深度估计结果，确定所述当前帧的深度估计结果；所述深度估计结果可用于实现基于所述当前帧的三维场景构建。

- 25 可选的，所述获取部分 801，具体配置为获取至少两个待筛选帧；从所述至少两个待筛选帧中，选取与所述当前帧之间满足预设角度约束条件的至少一帧，将所述至少一帧作为所述参考帧。

可选的，所述预设角度约束条件包括：

所述当前帧对应的位姿中心和所述参考帧对应的位姿中心，与目标点的连线形成的夹角处于第一预设角度范围；所述目标点为所述当前帧对应的平均深度点与所述参考帧对应的平均深度点连线的中点；

- 30 所述当前帧和所述参考帧对应的光轴夹角处于第二预设角度范围；

所述当前帧和所述参考帧对应的纵轴夹角处于第三预设角度范围。

- 35 可选的，所述估计部分 803，具体配置为基于所述 k 层当前图像和所述逆深度空间范围，确定第 i 层采样点中每一个采样点对应的逆深度候选值；所述第 i 层采样点为对所述 k 层当前图像中第 i 层当前图像采样获得的像素点，i 为大于等于 1 且小于等于 k 的自然数；根据所述第 i 层采样点中每一个采样点对应的逆深度候选值和所述 k 层参考图像中第 i 层参考图像，确定所述第 i 层采样点中每一个采样点的逆深度值，获得第 i 层逆深度值；令 $i=i+1$ ，继续对所述 k 层当前图像中分辨率高于所述第 i 层当前图像的第 i+1 层当前图像进行逆深度估计，直至 $i=k$ 为止，获得第 k 层逆深度值；将所述第 k 层逆深度值确定为所述逆深度估计结果。

- 40 可选的，所述估计部分 803，具体配置为当对所述逆深度空间范围进行区间划分，并在每个划分区间中选择一个逆深度值，得到多个初始逆深度值；将所述多个初始逆深度值确定为第 1 层采样点中每一个采样点对应的逆深度候选值；在 i 不等于 1 的情况下，从所述 k 层当前图像中获取第 i-1 层采样点，以及第 i-1 层逆深度值；基于所述第 i-1 层逆深度估计值、第 i-1 层采样点，以及所述多个初始逆深度值，确定所述第 i 层采样点中每一个采样点对应的逆深度候选值。

- 45 可选的，所述估计部分 803，具体配置为从所述第 i-1 层采样点中确定与第一采样点距离最近的第二采样点，以及与所述第二采样点相邻的至少两个第三采样点；所述第一采样点为所述第 i 层采样点中任意一个采样点；根据所述第 i-1 层逆深度值，获取所述至少两个第三采样点中每一个采样点的逆深度值，以及所述第二采样点的逆深度值，得到至少三个逆深度值；从所述至少三个逆深度值中，确定最大逆深度值和最小逆深度值；从所述多个初始逆深度值中，选取处于所述最大逆深度值和所述最小逆深度值范围内的逆深度值，将选取出的逆深度值确定为所述第一采样点对应的逆深度候选值；继续确定所述第 i 层采样点中非所述第一采样点的采样点对应的逆深度候选值，直至确定

出所述第 i 层采样点中每一个采样点对应的逆深度候选值。

可选的,所述估计部分 803,具体配置为对所述第 i 层采样点中每一个采样点,分别按照对应的逆深度候选值中的每一个逆深度值,将所述第 i 层采样点中每一个采样点投影到所述第 i 层参考图像中,获得所述第 i 层采样点中每一个采样点对应的第 i 层投影点;根据所述第 i 层采样点和所述第 i 层投影点进行块匹配,获得所述第 i 层采样点中每一个采样点对应的第 i 层匹配结果;根据所述第 i 层匹配结果,确定所述第 i 层采样点中每一个采样点的逆深度值,获得所述第 i 层逆深度值。

可选的,所述估计部分 803,具体配置为利用预设窗口,从所述第 i 层当前图像中选取以待匹配采样点为中心的多个第一图像块,并从所述第 i 层参考图像中选取以所述待匹配采样点对应的第 i 层投影点为中心的多个第二图像块;所述待匹配采样点为所述第 i 层采样点中任意一个采样点;将所述第一图像块分别与所述第二图像块中每一个图像块进行比较,获得多个匹配结果,并将所述多个匹配结果确定为所述待匹配采样点对应的第 i 层匹配结果;继续确定所述第 i 层采样点中与所述待匹配采样点不同的采样点对应的第 i 层匹配结果,直至获得所述第 i 层采样点中每一个采样点对应的第 i 层匹配结果。

可选的,所述估计部分 803,具体配置为从目标采样点对应的第 i 层匹配结果中选取目标匹配结果;所述目标采样点为所述第 i 层采样点中任意一个采样点;将所述目标采样点对应的第 i 层投影点中,所述目标匹配结果对应的投影点确定为目标投影点;将所述逆深度候选值中,所述目标投影点对应的逆深度值确定为所述目标采样点的逆深度值;继续确定所述第 i 层采样点中与所述目标采样点不同的采样点的逆深度值,直至确定出所述第 i 层采样点中每一个采样点的逆深度值,获得所述第 i 层逆深度值。

可选的,所述估计部分 803,还配置为对所述第 k 层逆深度值进行插值优化,获得优化后的第 k 层逆深度值;将所述优化后的第 k 层逆深度值确定为所述逆深度估计结果。

可选的,所述估计部分 803,具体配置为对所述第 k 层逆深度值中每一个逆深度值,分别从第 k 层采样点对应的采样点的候选逆深度值中,选取相邻逆深度值;所述第 k 层采样点为对所述 k 层当前图像中第 k 层当前图像采样获得的像素点;获取所述相邻逆深度值对应的匹配结果;基于所述相邻逆深度值和所述相邻逆深度值对应的匹配结果,对所述第 k 层逆深度值中的每一个逆深度值进行插值优化,获得所述优化后的第 k 层逆深度值。

本公开实施例提供了一种图像深度估计装置,获取当前帧对应的参考帧和当前帧的逆深度空间范围;对当前帧和参考帧分别进行金字塔降采样处理,获得当前帧对应的 k 层当前图像,以及参考帧对应的 k 层参考图像;k 为大于等于 2 的自然数;基于 k 层参考图像和逆深度空间范围,对 k 层当前图像进行逆深度估计迭代处理,获得当前帧的逆深度估计结果。也就是说,本公开提供的图像深度估计装置,采取了对多层当前图像结合多层参考图像进行逆深度估计迭代处理,以逐层减少逆深度搜索空间,确定当前帧的深度估计结果,且最终的深度估计结果为当前帧的像素点在相机坐标系下的 z 轴坐标值,不需要额外进行坐标变换,从而能够实时获得图像的深度估计结果,且深度估计结果的精确度较高。

本公开实施例还提供了一种电子设备,图 9 为本公开实施例提供的一种电子设备的结构示意图。如图 9 所示,所述电子设备包括:处理器 901、存储器 902 和通信总线 903;其中,

所述通信总线 903,配置为实现所述处理器 901 和所述存储器 902 之间的连接通信;

所述处理器 901,配置为执行所述存储器 902 中存储的图像深度估计程序,以实现上述图像深度估计方法。

需要说明的是,在本公开的实施例中,所述电子设备为手机或平板电脑,当然,也可以为其它类型设备,本公开实施例不作限定。

本公开实施例还提供了一种计算机可读存储介质,所述计算机可读存储介质存储有一个或者多个程序,所述一个或者多个程序可以被一个或者多个处理器执行,以实现上述图像深度估计方法。计算机可读存储介质可以是易失性存储器(volatile memory),例如随机存取存储器(Random-Access Memory, RAM);或者非易失性存储器(non-volatile memory),例如只读存储器(Read-Only Memory, ROM),快闪存储器(flash memory),硬盘(Hard Disk Drive, HDD)或固态硬盘(Solid-State Drive, SSD);也可以是包括上述存储器之一或任意组合的各自设备,如移动电话、计算机、平板设备、个人数字助理等。

本公开实施例还提供了一种计算机程序,包括计算机可读代码,所述计算机可读代码被处理器执行时,实现上述图像深度估计方法对应的步骤。

本领域内的技术人员应明白,本公开的实施例可提供为方法、系统、或计算机程序产品。因此,

本公开可采用硬件实施例、软件实施例、或结合软件和硬件方面的实施例的形式。而且，本公开可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器和光学存储器等）上实施的计算机程序产品的形式。

本公开是参照根据本公开实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程信号处理设备的处理器以产生一个机器，使得通过计算机或其他可编程信号处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程信号处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程信号处理设备，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

以上所述，仅为本公开的一些实施例而已，并非用于限定本公开的保护范围。在不违背逻辑的情况下，本申请不同实施例之间可以相互结合，不同实施例描述有所侧重，为侧重描述的部分可以参见其他实施例的记载。

工业实用性

在本公开实施例的技术方案中，获取当前帧对应的参考帧和当前帧的逆深度空间范围；对当前帧和参考帧分别进行金字塔降采样处理，获得当前帧对应的k层当前图像，以及参考帧对应的k层参考图像；k为大于等于2的自然数；基于k层参考图像和逆深度空间范围，对k层当前图像进行逆深度估计迭代处理，获得当前帧的逆深度估计结果。也就是说，本公开提供的技术方案，采取了对多层当前图像结合多层参考图像进行逆深度估计迭代处理，以逐层减少逆深度搜索空间，确定当前帧的逆深度估计结果，该逆深度估计结果为当前帧的像素点在相机坐标系下的z轴坐标值的倒数，不需要额外进行坐标变换，且逐层减少逆深度搜索空间有助于减少逆深度估计的计算量，提升估计速度，从而能够实时获得图像的深度估计结果，且深度估计结果的精确度较高。