



(21) 申请号 201811550318.1

(22) 申请日 2018.12.18

(65) 同一申请的已公布的文献号
申请公布号 CN 111340922 A

(43) 申请公布日 2020.06.26

(73) 专利权人 北京三星通信技术研究有限公司
地址 100028 北京市朝阳区太阳宫中路12
号楼15层1503
专利权人 三星电子株式会社

(72) 发明人 高天豪 刘志花 金允泰 李炯旭
王强 毛亚敏 马林

(74) 专利代理机构 北京铭硕知识产权代理有限
公司 11286
专利代理师 董钢 韩明星

(51) Int.Cl.

G06T 15/00 (2011.01)

G06N 3/0464 (2023.01)

G06N 3/045 (2023.01)

G06T 7/73 (2017.01)

G06T 7/33 (2017.01)

(56) 对比文件

CN 108416840 A, 2018.08.17

审查员 于滢

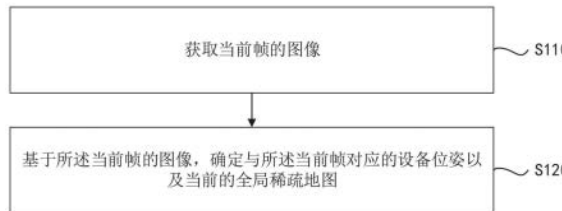
权利要求书4页 说明书20页 附图5页

(54) 发明名称

定位与地图构建的方法和电子设备

(57) 摘要

提供了一种定位与地图构建的方法和电子设备。所述方法可包括：获取当前帧的图像；基于所述当前帧的图像，确定与所述当前帧对应的设备位姿以及当前的全局稀疏地图。本公开的方法和电子设备，可提高确定的设备位姿和地图的准确性，并且具有鲁棒性。



1. 一种获取地图的方法,其中,所述方法包括:
基于包括当前帧的帧,确定与所述当前帧对应的设备位姿和/或全局稀疏地图;
获取与所述当前帧对应的全局稠密地图;
基于所述全局稠密地图,对与所述当前帧对应的设备位姿和/或当前的全局稀疏地图进行优化,
其中,获取与所述当前帧对应的全局稠密地图的步骤,包括:
基于所述当前帧,获取与所述当前帧对应的局部稠密地图,
基于所述局部稠密地图和当前的全局稀疏地图,确定与所述当前帧对应的全局稠密地图,
其中,获取所述局部稠密地图,包括:
通过将所述当前帧中的左目图像和右目图像输入特征提取网络来获取与右目图像相应的右特征图和与左目图像相应的左特征图,
基于左特征图和右特征图获取左目图像和右目图像之间的匹配像素的初始匹配代价体,
通过将初始匹配代价体输入神经网络预测匹配代价体,
基于所预测的匹配代价体确定所述当前帧中的各像素点的深度信息,以及
基于所述确定的深度信息确定所述局部稠密地图。
2. 如权利要求1所述的方法,其中,特征提取网络包括输入左目图像的左特征提取网络和输入右特征图像的右特征提取网络,其中,左特征提取网络和右特征提取网络共享权重。
3. 如权利要求1所述的方法,其中,获取所述初始匹配代价体的步骤包括:通过连接左特征图和右特征图来获取所述初始匹配代价体。
4. 如权利要求1所述的方法,其中,预测所述匹配代价体的步骤包括:基于所述初始匹配代价体和沙漏Hourglass神经网络预测所述匹配代价体。
5. 如权利要求2-4中的任一项权利要求所述的方法,其中,根据所述预测的匹配代价体,确定所述当前帧中的各像素点的深度信息的步骤,包括:
通过卷积神经网络,对所述预测的匹配代价体进行空间卷积处理;
根据空间卷积处理后的所述预测的匹配代价体,确定所述当前帧中的各像素点在左目图像和右目图像之间的视差;
根据各像素点的视差,确定所述当前帧中的各像素点的深度信息。
6. 如权利要求5所述的方法,其中,对所述预测的匹配代价体进行空间卷积处理的步骤,包括:
对所述预测的匹配代价体按照设定的多个方向分别执行如下处理:
按照当前设定方向,对所述预测的匹配代价体进行分割,得到多个匹配代价层,按照当前设定方向依次对各匹配代价层执行卷积处理。
7. 如权利要求6所述的方法,其中,当对任一匹配代价层进行卷积处理时,将所述任一匹配代价层累加上一匹配代价层的卷积结果后进行卷积处理。
8. 如权利要求5所述的方法,其中,确定所述当前帧中的各像素点在左目图像和右目图像之间的视差的步骤,包括:
利用柔性最大传递函数以及空间卷积处理后的所述预测的匹配代价体,获得所述当前

帧中的各像素点在左目图像和右目图像之间的视差概率分布；

基于所述视差概率分布确定所述视差。

9. 如权利要求8所述的方法,其中,基于所述视差概率分布确定所述视差的步骤包括:确定与各像素点对应的视差概率分布的累积值作为对应像素点在左目图像和右目图像之间的视差。

10. 如权利要求1所述的方法,其中,基于所述局部稠密地图和当前的全局稀疏地图,确定与所述当前帧对应的全局稠密地图的步骤,包括:

确定所述当前帧中的属于动态物体的像素点;

从所述当前的全局稀疏地图中剔除属于动态物体的像素点;

将经过像素点剔除的所述当前的全局稀疏地图和所述局部稠密地图融合为全局稠密地图。

11. 如权利要求10所述的方法,其中,确定所述当前帧中的属于动态物体的像素点的步骤,包括:

确定与所述当前帧对应的物体状态信息;

基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

12. 如权利要求11所述的方法,其中,确定与所述当前帧对应的物体状态信息的步骤,包括:

通过特征提取神经网络,对所述当前帧的图像进行特征提取;

根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;

根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

13. 如权利要求12所述的方法,根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息的步骤,包括:

确定所述当前帧与所述当前帧的前一帧之间的光流信息;

根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

14. 如权利要求10-13中的任一项权利要求所述的方法,其中,将经过像素点剔除的所述当前的全局稀疏地图和所述局部稠密地图融合为全局稠密地图的步骤,包括:

对所述局部稠密地图进行网格划分;

将划分的各个网格的顶点更新到经过像素点剔除的全局稀疏地图;

将各个网格的网格内像素点更新到经过像素点剔除的全局稀疏地图。

15. 如权利要求14所述的方法,其中,将任一顶点更新到经过像素点剔除的全局稀疏地图的步骤,包括:

在与所述任一顶点所属的各网格对应的各个顶点中,确定具有全局稀疏地图坐标的顶点;

根据确定的顶点的深度信息、所述任一顶点在所述局部稠密地图中的深度信息、以及确定的顶点与所述任一顶点之间的距离,获得所述任一顶点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述当前帧对应的设备位姿和所述任一顶点在经过像素点剔除的全局稀疏地图中的深度信息,确定所述任一顶点在全局稀疏地图中的坐标,以更新

全局稀疏地图。

16. 如权利要求14所述的方法,其中,将网格内任一像素点更新到经过像素点剔除的全局稀疏地图的步骤,包括:

根据所述任一像素点所在网格的顶点的深度信息、所述任一像素点在所述局部稠密地图中的深度信息、以及所述任一像素点所在网格的顶点与所述任一像素点之间的距离,获得所述任一像素点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述当前帧对应的设备位姿和所述任一像素点在经过像素点剔除的全局稀疏地图中的深度信息,确定所述任一像素点在全局稀疏地图中的坐标,以更新全局稀疏地图。

17. 如权利要求1所述的方法,基于所述全局稠密地图,对与各关键帧分别对应的设备位姿以及当前的全局稀疏地图进行优化的步骤,包括:

基于所述全局稠密地图、与各关键帧分别对应的设备位姿、以及所述当前的全局稀疏地图,对与各关键帧分别对应的设备位姿以及所述当前的全局稀疏地图进行集束优化。

18. 如权利要求10所述的方法,其中,确定所述当前帧中的属于动态物体的像素点的步骤,包括:

确定与所述当前帧对应的物体状态信息;

基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

19. 如权利要求18所述的方法,其中,确定与所述当前帧对应的物体状态信息的步骤,包括:

通过特征提取神经网络,对所述当前帧的图像进行特征提取;

根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;

根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

20. 如权利要求19所述的方法,根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息的步骤,包括:

确定所述当前帧与所述当前帧的前一帧之间的光流信息;

根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

21. 如权利要求10所述的方法,其中,获取所述全局稀疏地图的步骤包括:

获取当前帧的像素点的深度信息;

确定与当前帧对应的设备位姿;以及

基于与当前帧对应的设备位姿以及当前帧的像素点的深度信息,确定当前帧的像素点在全局稀疏地图中的三维坐标。

22. 如权利要求10所述的方法,基于所述全局稠密地图,对与所述当前帧对应的全局稀疏地图进行优化的步骤包括:

基于优化后的与当前帧对应的设备位姿优化所述全局稀疏地图。

23. 如权利要求10所述的方法,其中,获取所述全局稀疏地图的步骤包括:

获取与当前帧的前一帧相应的第一深度信息;

获取与当前帧相应的第二深度信息;

基于第一深度信息和第二深度信息确定与当前帧对应的设备的位姿信息;以及

基于与当前帧对应的设备的位姿信息以及第二深度信息,获取所述全局稀疏地图。

24. 如权利要求10所述的方法,其中,获取第二深度信息的步骤包括:

执行当前帧中的左目图像和右目图像的立体匹配。

25. 如权利要求10所述的方法,其中,获取所述局部稠密地图的步骤包括:

通过将当前帧输入神经网络获得与像素点的深度信息相应的输出;以及
基于所述输出获取局部稠密地图。

26. 如权利要求10所述的方法,其中,获取所述全局稠密地图的步骤包括:

对所述局部稠密地图进行网格划分;

基于经过像素点剔除的全局稀疏地图更新与划分的各个网格的顶点相应的像素点的深度信息;以及

基于更新的与划分的各个网格的顶点相应的像素点的深度信息与和经过像素点剔除的全局稀疏地图更新各个网格的网格内像素点的深度信息。

27. 如权利要求10所述的方法,确定所述当前帧中的属于动态物体的像素点的步骤包括:通过对静态物体进行掩模来确定所述当前帧中的属于动态物体的像素点。

28. 如权利要求1所述的方法,还包括:获取所述当前帧的图像。

29. 如权利要求1所述的方法,其中,所述当前帧为关键帧。

30. 如权利要求10所述的方法,其中,确定所述当前帧中的属于动态物体的像素点的步骤包括:

通过将当前帧输入神经网络来获取神经网络的被分类为静态对象区域和动态对象区域的输出;以及

基于所述输出确定属于所述动态物体的像素点。

31. 如权利要求23所述的方法,其中,所述位姿信息包括:响应于所述设备从第一位置向第二位置移动而变化的旋转信息和平移信息中的任一个或者所述旋转信息和所述平移信息的任何组合。

32. 一种存储有计算机程序的计算机可读存储介质,其中,当所述计算机程序被处理器执行时实现如权利要求1至31之中的任一项权利要求所述的方法。

33. 一种电子设备,其中,所述电子设备包括:

处理器;

存储器,存储有计算机程序,当所述计算机程序被处理器执行时,实现如权利要求1至31之中的任一项权利要求所述的方法。

34. 一种电子设备,其中,所述电子设备包括:

图像获取单元,被配置为获取包括当前帧的帧;

处理器,被配置为执行如权利要求1-31之中的任一项权利要求所述的方法。

定位与地图构建的方法和电子设备

技术领域

[0001] 本公开总体说来涉及即时定位与地图构建 (Simultaneous Localization And Mapping, 简称为SLAM) 领域, 更具体地讲, 涉及一种定位与地图构建的方法和电子设备。

背景技术

[0002] 利用设备上的相机、激光雷达等传感器, 实时的构建出描述设备所处空间的三维地图, 并且确定设备位姿 (包括: 位置和姿态), 这样的技术被称为SLAM。由于相机的价格比激光雷达更低, 并且适用范围更广 (即: 可以在各种天气、场景下使用), 所以现有的SLAM技术的主要研究内容是利用相机作为传感器, 进行实时的地图构建和设备位姿确定。然而, 在现有技术中, 存在确定的位姿和构建的地图的准确性差的问题。

发明内容

[0003] 本公开的示例性实施例在于解决现有技术中存在的问题。

[0004] 根据本公开的示例性实施例, 提供了一种定位与地图构建的方法, 其中, 所述方法包括: 获取当前帧的图像; 基于所述当前帧的图像, 确定与所述当前帧对应的设备位姿以及当前的全局稀疏地图。

[0005] 可选的, 所述方法还包括: 当所述当前帧为关键帧时, 获取与所述当前帧对应的全局稠密地图; 基于所述全局稠密地图, 对与各关键帧分别对应的设备位姿以及当前的全局稀疏地图进行优化。

[0006] 可选的, 获取与所述当前帧对应的全局稠密地图的步骤, 包括: 获取与所述当前帧对应的局部稠密地图; 基于所述局部稠密地图和当前的全局稀疏地图, 确定与所述当前帧对应的全局稠密地图。

[0007] 可选的, 获取与所述当前帧对应的局部稠密地图的步骤, 包括: 通过特征提取神经网络, 对所述当前帧的图像进行特征提取; 通过沙漏Hourglass神经网络, 根据通过特征提取得到的特征图, 获得匹配代价体; 根据所述匹配代价体, 确定所述当前帧中的各像素点的深度信息, 得到与所述当前帧对应的局部稠密地图。

[0008] 可选的, 通过Hourglass神经网络, 根据通过特征提取得到的特征图, 获得匹配代价体的步骤, 包括: 根据所述特征图, 获取初始匹配代价体; 通过Hourglass神经网络, 基于所述初始匹配代价体, 获得所述匹配代价体。

[0009] 可选的, 所述特征图包括左目图像的特征图和右目图像的特征图, 其中, 根据所述特征图, 获取初始匹配代价体的步骤, 包括: 将左目图像的特征图和右目图像的特征图进行连接, 以获得初始匹配代价体。

[0010] 可选的, 根据所述匹配代价体, 确定所述当前帧中的各像素点的深度信息的步骤, 包括: 通过卷积神经网络, 对所述匹配代价体进行空间卷积处理; 根据空间卷积处理后的匹配代价体, 确定所述当前帧中的各像素点在左目图像和右目图像之间的视差; 根据各像素点的视差, 确定所述当前帧中的各像素点的深度信息, 得到与所述当前帧对应的局部稠密

地图。

[0011] 可选的,对所述匹配代价体进行空间卷积处理的步骤,包括:对所述匹配代价体按照设定的多个方向分别执行如下处理:按照当前设定方向,对所述匹配代价体进行分割,得到多个匹配代价层,按照当前设定方向依次对各匹配代价层执行卷积处理,其中,当对任一匹配代价层进行卷积处理时,将所述任一匹配代价层累加上一匹配代价层的卷积结果后进行卷积处理。

[0012] 可选的,确定所述当前帧中的各像素点在左目图像和右目图像之间的视差的步骤,包括:利用柔性最大传递函数以及空间卷积处理后的匹配代价体,获得所述当前帧中的各像素点在左目图像和右目图像之间的视差概率分布;确定与各像素点对应的视差概率分布的累积值作为对应像素点在左目图像和右目图像之间的视差。

[0013] 可选的,基于所述局部稠密地图和当前的全局稀疏地图,确定与所述当前帧对应的全局稠密地图的步骤,包括:确定所述当前帧中的属于动态物体的像素点;从所述当前的全局稀疏地图中剔除属于动态物体的像素点;将经过像素点剔除的所述当前的全局稀疏地图和所述局部稠密地图融合为全局稠密地图。

[0014] 可选的,确定所述当前帧中的属于动态物体的像素点的步骤,包括:确定与所述当前帧对应的物体状态信息;基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

[0015] 可选的,确定与所述当前帧对应的物体状态信息的步骤,包括:通过特征提取神经网络,对所述当前帧的图像进行特征提取;根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

[0016] 可选的,根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息的步骤,包括:确定所述当前帧与所述当前帧的前一帧之间的光流信息;根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

[0017] 可选的,将经过像素点剔除的所述当前的全局稀疏地图和所述局部稠密地图融合为全局稠密地图的步骤,包括:对所述局部稠密地图进行网格划分;将划分的各个网格的顶点更新到经过像素点剔除的全局稀疏地图;将各个网格的网格内像素点更新到经过像素点剔除的全局稀疏地图。

[0018] 可选的,将任一顶点更新到经过像素点剔除的全局稀疏地图的步骤,包括:在与所述任一顶点所属的各网格对应的各个顶点中,确定具有全局稀疏地图坐标的顶点;根据确定的顶点的深度信息、所述任一顶点在所述局部稠密地图中的深度信息、以及确定的顶点与所述任一顶点之间的距离,获得所述任一顶点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述当前帧对应的设备位姿和所述任一顶点在经过像素点剔除的全局稀疏地图中的深度信息,确定所述任一顶点在全局稀疏地图中的坐标,以更新全局稀疏地图。

[0019] 可选的,将网格内任一像素点更新到经过像素点剔除的全局稀疏地图的步骤,包括:根据所述任一像素点所在网格的顶点的深度信息、所述任一像素点在所述局部稠密地图中的深度信息、以及所述任一像素点所在网格的顶点与所述任一像素点之间的距离,获

得所述任一像素点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述当前帧对应的设备位姿和所述任一像素点在经过像素点剔除的全局稀疏地图中的深度信息,确定所述任一像素点在全局稀疏地图中的坐标,以更新全局稀疏地图。

[0020] 可选的,基于所述全局稠密地图,对与各关键帧分别对应的设备位姿以及当前的全局稀疏地图进行优化的步骤,包括:基于所述全局稠密地图、与各关键帧分别对应的设备位姿、以及所述当前的全局稀疏地图,对与各关键帧分别对应的设备位姿以及所述当前的全局稀疏地图进行集束优化。

[0021] 可选的,所述方法还包括:确定所述当前帧中的属于动态物体的像素点;从所述当前的全局稀疏地图中剔除属于动态物体的像素点。

[0022] 可选的,确定所述当前帧中的属于动态物体的像素点的步骤,包括:确定与所述当前帧对应的物体状态信息;基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

[0023] 可选的,确定与所述当前帧对应的物体状态信息的步骤,包括:通过特征提取神经网络,对所述当前帧的图像进行特征提取;根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

[0024] 可选的,根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息的步骤,包括:确定所述当前帧与所述当前帧的前一帧之间的光流信息;根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

[0025] 根据本公开的另一示例性实施例,提供了一种存储有计算机程序的计算机可读存储介质,其中,当所述计算机程序被处理器执行时实现如上所述的方法。

[0026] 根据本公开的另一示例性实施例,提供了一种电子设备,其中,所述电子设备包括:处理器;存储器,存储有计算机程序,当所述计算机程序被处理器执行时,实现如上所述的方法。

[0027] 根据本公开的另一示例性实施例,提供了一种电子设备,其中,所述电子设备包括:图像获取单元,用于获取当前帧的图像;位姿与地图确定单元,用于基于所述当前帧的图像,确定与所述当前帧对应的设备位姿以及当前的全局稀疏地图。

[0028] 可选的,所述电子设备还包括:全局稠密地图获取单元,用于当当前帧为关键帧时,获取与所述当前帧对应的全局稠密地图;优化单元,用于基于所述全局稠密地图,对与各关键帧分别对应的设备位姿以及当前的全局稀疏地图进行优化。

[0029] 可选的,全局稠密地图获取单元包括:局部稠密地图获取单元,用于获取与所述当前帧对应的局部稠密地图;其中,全局稠密地图获取单元,基于所述局部稠密地图和当前的全局稀疏地图,确定与所述当前帧对应的全局稠密地图。

[0030] 可选的,所述局部稠密地图获取单元,通过特征提取神经网络,对所述当前帧的图像进行特征提取;通过沙漏Hourglass神经网络,根据通过特征提取得到的特征图,获得匹配代价体;根据所述匹配代价体,确定所述当前帧中的各像素点的深度信息,得到与所述当前帧对应的局部稠密地图。

[0031] 可选的,所述局部稠密地图获取单元,根据所述特征图,获取初始匹配代价体;通

过Hourglass神经网络,基于所述初始匹配代价体,获得所述匹配代价体。

[0032] 可选的,所述特征图包括左目图像的特征图和右目图像的特征图,其中,所述局部稠密地图获取单元,将左目图像的特征图和右目图像的特征图进行连接,以获得初始匹配代价体。

[0033] 可选的,所述局部稠密地图获取单元,通过卷积神经网络,对所述匹配代价体进行空间卷积处理;根据空间卷积处理后的匹配代价体,确定所述当前帧中的各像素点在左目图像和右目图像之间的视差;根据各像素点的视差,确定所述当前帧中的各像素点的深度信息,得到与所述当前帧对应的局部稠密地图。

[0034] 可选的,所述局部稠密地图获取单元,对所述匹配代价体按照设定的多个方向分别执行如下处理:按照当前设定方向,对所述匹配代价体进行分割,得到多个匹配代价层,按照当前设定方向依次对各匹配代价层执行卷积处理,其中,当对任一匹配代价层进行卷积处理时,将所述任一匹配代价层累加上一匹配代价层的卷积结果后进行卷积处理。

[0035] 可选的,所述局部稠密地图获取单元,利用柔性最大传递函数以及空间卷积处理后的匹配代价体,获得所述当前帧中的各像素点在左目图像和右目图像之间的视差概率分布;确定与各像素点对应的视差概率分布的累积值作为对应像素点在左目图像和右目图像之间的视差。

[0036] 可选的,所述全局稠密地图获取单元包括:动态物体确定单元,用于确定所述当前帧中的属于动态物体的像素点;从所述当前的全局稀疏地图中剔除属于动态物体的像素点,其中,所述全局稠密地图获取单元,将经过像素点剔除的所述当前的全局稀疏地图和所述局部稠密地图融合为全局稠密地图。

[0037] 可选的,所述动态物体确定单元,确定与所述当前帧对应的物体状态信息;基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

[0038] 可选的,所述动态物体确定单元,通过特征提取神经网络,对所述当前帧的图像进行特征提取;根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

[0039] 可选的,所述动态物体确定单元,确定所述当前帧与所述当前帧的前一帧之间的光流信息;根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

[0040] 可选的,所述全局稠密地图获取单元包括融合单元,用于对所述局部稠密地图进行网格划分;将划分的各个网格的顶点更新到经过像素点剔除的全局稀疏地图;将各个网格的网格内像素点更新到经过像素点剔除的全局稀疏地图。

[0041] 可选的,所述融合单元,确定与所述当前顶点相邻的网格的各个顶点中具有全局稀疏地图坐标的顶点;根据确定的顶点的深度信息、所述当前顶点在所述局部稠密地图中的深度信息、以及确定的顶点与所述当前顶点之间的距离,获得所述当前顶点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述当前顶点对应的设备位姿和获得的深度信息,确定所述当前顶点在全局稀疏地图中的坐标,以更新全局稀疏地图。

[0042] 可选的,所述融合单元,根据网格内像素点所在网格的顶点的深度信息、网格内像素点在所述局部稠密地图中的深度信息、以及网格内像素点所在网格的顶点与网格内像素点之间的距离,获得所述当前顶点在经过像素点剔除的全局稀疏地图中的深度信息,并且

根据与所述网格内像素点对应的设备位姿和获得的深度信息,确定网格内像素点在全局稀疏地图中的坐标,以更新全局稀疏地图。

[0043] 可选的,所述优化单元,基于所述全局稠密地图、与各关键帧分别对应的设备位姿、以及所述当前的全局稀疏地图,对与各关键帧分别对应的设备位姿以及所述当前的全局稀疏地图进行集束优化。

[0044] 可选的,所述电子设备还包括:动态物体确定单元,用于确定所述当前帧中的属于动态物体的像素点;从所述当前的全局稀疏地图中剔除属于动态物体的像素点。

[0045] 可选的,所述动态物体确定单元,确定与所述当前帧对应的物体状态信息;基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

[0046] 可选的,所述动态物体确定单元,通过特征提取神经网络,对所述当前帧的图像进行特征提取;根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

[0047] 可选的,所述动态物体确定单元,确定所述当前帧与所述当前帧的前一帧之间的光流信息;根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

[0048] 根据本公开的示例性实施例,提供一种定位与地图构建的方法和电子设备,以提高现有的用于SLAM的方法和设备的准确性,尤其是,在不满足多视角几何理论的应用前提时,可稳定的进行定位与地图构建,因此,具有鲁棒性。

[0049] 本公开改进了传统的用于SLAM的方法和设备,可将多视角几何理论和深度学习理论结合,用于位姿确定和地图构建。本公开可采用以下操作中的至少一种操作:基于多视角几何理论确定位姿和构建全局稀疏地图(可被称作基本SLAM操作),基于深度学习理论来构建局部稠密地图,对构建的地图进行动态物体剔除以降低动态物体对定位精度的影响,对构建的地图进行融合,对设备位姿和地图进行优化(包括集束优化),从而可提高准确性,并且具有鲁棒性。

[0050] 将在接下来的描述中部分阐述本公开总体构思另外的方面和/或优点,还有一部分通过描述将是清楚的,或者可以经过本公开总体构思的实施而得知。

附图说明

[0051] 通过下面结合示例性地示出实施例的附图进行的描述,本公开示例性实施例的上述和其他目的和特点将会变得更加清楚,其中:

[0052] 图1示出根据本公开的示例性实施例的定位与地图构建的方法的流程图;

[0053] 图2示出根据多视角几何理论获取像素点的深度信息的操作的示意图;

[0054] 图3示出根据本公开的另一示例性实施例的定位与地图构建的方法的流程图;

[0055] 图4示出根据本公开的另一示例性实施例的定位与地图构建的方法的示意图;

[0056] 图5示出根据本公开的另一示例性实施例的定位与地图构建的方法的示意图;

[0057] 图6示出根据本公开的示例性实施例的进行深度立体匹配操作的示意图;

[0058] 图7示出根据本公开的示例性实施例的对匹配代价体进行卷积的操作的示意图;

[0059] 图8示出了根据本公开的示例性实施例的优化设备位姿过程和优化地图点深度过程的示意图;

- [0060] 图9示出根据本公开的示例性实施例的获取静态物体掩膜的操作的示意图；
- [0061] 图10示出根据本公开的示例性实施例的地图融合操作的示意图；
- [0062] 图11示出根据本公开的示例性实施例的基于网格划分进行地图融合的操作的示意图；
- [0063] 图12示出根据本公开的示例性实施例的进行集束优化操作的示意图；
- [0064] 图13示出根据本公开的示例性实施例的电子设备的框图。

具体实施方式

[0065] 现将详细参照本公开的实施例,所述实施例的示例在附图中示出,其中,相同的标号始终指的是相同的部件。以下将通过参照附图来说明所述实施例,以便解释本公开。

[0066] 在本公开的示例性实施例中,设备位姿(获取图像的相机等装置所在的设备的位姿)包括位置和姿态,可用坐标表示位置(例如,三维空间中的坐标),用相对角度表示姿态(例如,三维空间中的朝向);稀疏可意指部分像素点,稠密可意指全部或大部分像素点;包含像素点(可以是关键帧的像素点)在三维空间中的坐标(可以称为全局坐标)的地图可以是全局地图,只包含像素点的深度信息的地图可以是局部地图;像素点的全局坐标可以为像素点的深度信息与当前帧的设备位姿的乘积;地图上的像素点可以称为地图点;可使用双目相机(即:包含两个相对位置固定的镜头的相机)获取各帧的图像,当使用双目相机时,一帧的图像包括左目图像和右目图像;可将图像中的有代表性的像素点作为特征点,例如,可将对比度超过预定值的预定区域内的像素点作为特征点。

[0067] 图1示出根据本公开的示例性实施例的定位与地图构建的方法的流程图。如图1中所示,所述方法可包括:步骤S110,获取当前帧的图像;步骤S120,基于所述当前帧的图像,确定与所述当前帧对应的设备位姿以及当前的全局稀疏地图。

[0068] 这里,当前帧可以是关键帧。由于与单目相机(即:只有一个镜头的相机)相比,利用双目相机构建的地图能够更准确地表示真实距离,因此,在本公开中,优选的使用双目相机采集图像。也就是说,当前帧的图像可包括与左目镜头对应的左目图像和与右目镜头对应的右目图像。

[0069] 在这种情况下,可基于多视角几何理论来执行步骤S120,如下面参照图图2描述的操作。

[0070] 图2示出根据多视角几何理论获取像素点的深度信息的操作的示意图。作为示例,以下针对当前帧(当前帧可以是关键帧)中的像素点(该像素点可以是特征点)进行操作,电子设备的图像获取单元可包括左目镜头201和右目镜头202。左目镜头201拍摄左目图像,右目镜头202拍摄右目图像。当前帧的图像包括左目图像和右目图像。

[0071] 对于当前帧,可利用特定的匹配方式(例如,基于OpenCV的匹配方式),可从左目图像和右目图像之一,例如右目图像中,找出与左目图像中的像素点匹配的像素点。即:对应的场景中的同一点在左目图像和右目图像中均出现。

[0072] 左目图像可包括:像素点203、像素点204和像素点205。右目图像可包括:像素点206、像素点207和像素点208。像素点203和像素点206匹配且对应于地图点209,像素点204和像素点207匹配且对应于地图点210,像素点205和像素点208匹配且对应于地图点211。

[0073] 左目镜头201的位置和右目镜头202的位置相对固定,左目镜头201的位置和右目

镜头202之间的距离是已知的,左目图像和右目图像之间的距离也是已知的,左目镜头201与左目图像之间的距离以及右目镜头202与右目图像之间的距离也是已知的。延长左目镜头201和像素点203的连线,延长右目镜头202与像素点206的连线,可获得地图点209,类似地,可获得地图点210和地图点211(地图点209、地图点210和地图点211位于全局稀疏地图)。这样,基于上述距离,可获得当前帧的像素点的深度信息(即:地图点的深度信息)。由于地图包括像素点,因此,在下文中,地图点也称作像素点。

[0074] 可将当前帧之前的帧作为参考,可获得与作为参考的帧(参考帧,可以是关键帧)对应的设备位姿。由此,可通过多视角几何立体匹配方法,确定与当前帧对应的设备位姿,基于与当前帧对应的设备位姿以及当前帧的像素点的深度信息,可确定当前帧的像素点在全局稀疏地图中的坐标。

[0075] 更具体而言,对于第一帧,可默认为关键帧,可基于多视角几何理论获取第一帧中的特定像素点的深度信息,得到局部稀疏地图。第一帧对应的设备位姿为默认位姿。可根据第一帧的深度信息和与第一帧对应的设备位姿确定全局稀疏地图。对于第一帧之后的任一帧,可基于图2所示的操作获取所述任一帧中的像素点的深度信息(像素点在与所述任一帧对应的局部稀疏地图中的深度信息),结合所述任一帧的前一帧中的像素点的深度信息,基于多视角立体匹配方法,确定与所述任一帧对应的设备位姿。基于与所述任一帧对应的设备位姿以及所述任一帧中的像素点的深度信息,可确定所述任一帧中的像素点在全局稀疏地图中的坐标,该坐标可被添加到全局稀疏地图中。

[0076] 在本公开的示例性实施例中,可采用各种确定特征点和关键帧的方法,本公开对此不作限定。

[0077] 图3示出根据本公开的另一示例性实施例的定位与地图构建的方法的流程图,其中,一帧的图像包括左目图像和右目图像。

[0078] 在图3所示示例性实施例中,可选择多个帧为关键帧。在初始化或重定位(一般执行初始化操作,在如下文所述的情况下可执行重定位操作)时,可基于多视角几何理论进行立体匹配,以计算任一关键帧中的像素点的深度信息,并将任一关键帧中的像素点的添加到全局稀疏地图中(可参照针对图2的描述)。

[0079] 作为示例,通过如下步骤确定关键帧:确定每一帧和与所述每一帧在时序上最近的关键帧之间的内容(例如,特征点的数量)变化,在内容变化较多(例如,特征点的数量变化超过阈值)时,将所述每一帧作为关键帧。

[0080] 对于当前帧(当前帧也可以是关键帧),也可利用多视角几何理论获取当前帧中的像素点的深度信息。随后,通过直接法进行匹配,对当前帧和所述当前帧的前一帧(例如,与所述当前帧在时序上最接近的关键帧)进行计算,以获得与当前帧对应的设备位姿。

[0081] 在通过直接法进行匹配的过程中,可将所述前一帧中所包含的地图点(具有全局稀疏地图坐标的像素点)投影到当前帧,由此获得投影到当前帧中的像素点的深度信息,并利用通过投影获得的像素点的深度信息以及当前帧原有的像素点的深度信息,确定与所述当前帧对应的设备位姿。

[0082] 通过上述方式,可持续获得对应的设备位姿,并将获得了深度信息并且不在全局稀疏地图中的像素点添加至全局稀疏地图,例如,针对当前帧,可通过设备位姿和深度信息可获得当前帧的像素点在全局稀疏地图中的坐标,如果获得的坐标未在全局稀疏地图中出

现,则将该像素点添加到全局稀疏地图中。

[0083] 作为示例,在光照或相机运动状态等发生较大变化时,可能会定位失败。此时,需要进行重定位。在重定位时,可从已确定的关键帧中找到内容与定位失败的帧最接近的关键帧,并将与找到的关键帧中的具有全局稀疏地图坐标的像素点的点投影到定位失败的帧,由此获得定位失败的帧中的像素点的深度信息,并将定位失败的帧也作为关键帧,进行设备位姿确定等操作。

[0084] 作为示例,可利用优化的方式进行地图优化和位姿优化,例如,利用全局稀疏地图优化与各个帧对应的设备位姿,并利用优化的设备位姿来优化全局稀疏地图,或者基于全局稠密地图,对各帧分别对应的设备位姿进行优化,以通过优化的设备位姿来优化全局稀疏地图,在这种情况下,需要先获得与当前帧对应的全部稠密地图,这将在下文中进行详细描述。

[0085] 进行优化的原因是:进行像素点匹配(即:找出对应于现实中的同一个点的多个像素点)的操作可能存在误差,这种误差会随着操作的进行而累积,从而导致基于多视角几何理论得到的设备位姿和构建的地图与实际地图之间存在误差。为了消除误差,可采用优化(优选为集束优化)的方式,对与每个关键字对应的设备位姿以及全局稀疏地图进行优化,以获得更准确的结果。

[0086] 另外,在根据多视角几何理论,确定设备位姿并构建地图的情况下,需要满足特定的假设条件。

[0087] 例如,可基于选择的有代表性的像素点(即:特征点),确定位姿,但是,需要在包含特定纹理(例如,便于与周围物体区分开的纹理)的区域或者在物体的边缘或棱角处选择特征点,然后,才能基于选择的特征点确定设备位姿和构建地图。如果区域中的纹理的数量少、或者纹理不明显,则难以选择出特征点,从而无法确定设备位姿和构建地图;如果纹理过多甚至重复,则容易出现误匹配,使确定的设备位姿和构建的地图的准确性降低。集束优化操作无法有效消除误差以提高准确度。

[0088] 又例如,出现超过预定面积的无纹理区域、环境光照变化超过预定限度、或者设备位姿变化(尤其是朝向变化)超过预定变化值时,难以有效地获取特征点,当特征点数量不足时,无法准确计算设备位姿。

[0089] 又如,在构建地图时,需要假设区域中的物体是静态的。如果选择的有代表性的像素点对应于非静态的物体,则会使确定设备位姿和构建地图所依据的关系(例如,几何约束)不成立,从而影响确定的设备位姿和构建的地图的准确性。

[0090] 考虑到上述问题,在本公开的示例性实施例中,可获取与所述当前帧对应的全局稠密地图,优选的,当当前帧为关键帧时,获取与所述当前帧对应的全局稠密地图;基于全局稠密地图,对各帧(优选为,各关键帧)分别对应的设备位姿以及当前的全局稀疏地图进行优化(优选为,进行集束优化);在本公开的另一示例性实施例中,可确定当前帧中的属于动态物体的像素点;从当前的全局稀疏地图中剔除属于动态物体的像素点。

[0091] 作为示例,获取与当前帧对应的全局稠密地图的步骤,包括:获取与当前帧对应的局部稠密地图;基于局部稠密地图和当前的全局稀疏地图,确定与当前帧对应的全局稠密地图。

[0092] 作为示例,确定当前帧中的属于动态物体的像素点的步骤,包括:确定与当前帧对

应的物体状态信息;基于所述物体状态信息,确定当前帧中的属于动态物体的像素点。

[0093] 作为示例,基于所述局部稠密地图和当前的全局稀疏地图,确定与所述当前帧对应的全局稠密地图的步骤,包括:确定所述当前帧中的属于动态物体的像素点;从所述当前的全局稀疏地图中剔除属于动态物体的像素点;将经过像素点剔除的所述当前的全局稀疏地图和所述局部稠密地图融合为全局稠密地图。

[0094] 图4示出根据本公开的另一示例性实施例的定位与地图构建的方法的示意图。如图4中所示,本公开的示例性实施例的方法可包括以下步骤:定位与地图构建(步骤S411和步骤S415)、深度立体匹配(步骤S420)、静态物体检测(步骤S430)、地图融合(步骤S440)、位姿和地图优化(步骤S460)。另外,在执行上述步骤之前,可通过视频采集设备(例如,双目相机)采集图像(例如,包括左目图像和右目图像的双目图像),并从采集的图像中确定关键帧。

[0095] 在即时定位(步骤S411)和地图重建(步骤S115)的步骤中,基于多视角几何理论对采集的图像(例如,当前关键帧的图像)进行计算,获得当前关键帧中的像素点的深度信息,并基于当前关键帧之前的帧(也可以是关键帧)与当前关键帧之间的位置关系,获得与当前关键帧对应的设备位姿,由此获得当前帧的像素点的在全局稀疏地图中的坐标。

[0096] 具体而言,在步骤S411,进行即时定位操作。利用多视角几何理论确定当前关键帧中的像素点的深度信息。第一帧为初始关键帧(时间顺序上的第一个帧),该帧所对应的设备位姿是默认的,可基于该帧中的像素点建立全局坐标系,以构建全局稀疏地图。对于第一帧之后的关键帧,可根据不同帧之间的深度等的信息的变化来确定与任一关键帧对应的设备位姿。在步骤S415,进行地图构建。通过第一帧中的像素点建立全局坐标系,在获得第一帧之后的关键帧中的像素点的坐标之后,可将获得坐标的像素点更新到全局坐标系下,作为全局稀疏地图中的地图点,由此更新全局稀疏地图。

[0097] 在深度立体匹配(步骤S420)的步骤中,预测当前关键帧中的全部像素点的深度信息,并由此构建基于所述当前关键帧的局部稠密地图。可通过预训练的深度学习网络对当前关键帧的图像进行计算以预测像素点的深度信息。

[0098] 具体而言,在步骤S420,进行深度立体匹配。利用预训练的神经网络来获取左目图像和右目图像之间的视差(具体为,像素点在左目图像和右目图像之间的视差),基于获取的视差,计算获取了视差的像素点的深度信息。对于当前帧,可创建以当前帧的左目图像的中心为原点的坐标系,并确定获得深度的像素点(可以是当前帧中的所有像素点)在该坐标系下的坐标,从而获得稠密的局部地图。这里,可通过多种方法,获得与像素点在左目图像和右目图像之间的视差对应的深度信息,这里不再赘述。

[0099] 在静态物体检测(步骤S430)的步骤中,确定像素点是否属于动态物体,例如,基于采集的图像获得用于表示物体的状态(包括静态或动态)的掩膜(例如,静态物体掩膜),其中,静态物体掩膜可表示像素点为静态物体;基于所述掩膜,可从全局稀疏地图中剔除与动态物体对应的像素点(即,全局稀疏地图中与动态物体对应的地图点)。可利用预训练的神经网络来预测像素点是否对应于动态物体。

[0100] 具体而言,在步骤S430,进行静态物体检测。利用预训练的深度学习网络,预测当前帧中每一个像素点是否对应于静态物体,例如,获得用于指示像素点是否对应于静态物体的静态物体掩膜;可利用是否对应于静态物体的预测结果从全局稀疏地图中剔除与静态

物体对应的像素点。

[0101] 在地图融合(步骤S440)的步骤中,对局部稠密地图以及剔除了与动态物体对应的像素点的全局稀疏地图中进行融合,以获得可信地图。这里,以全局稀疏地图为基础,将局部稠密地图映射到全局稀疏地图(优选为,剔除了与静态物体对应的像素点的全局稀疏地图),从而得到全局稠密的地图。在步骤S450,可进行辅助定位,即:将高精地图、雷达地图融合到可信地图中,以增加地图精度。

[0102] 在位姿和地图优化(步骤S460)的步骤中,以第一帧的位姿作为初始值进行集束优化以获得与各个关键帧对应的优化的设备位姿,利用优化的设备位姿对全局稀疏地图进行优化。在集束优化(步骤S460)的过程中,可将与初始关键帧对应的设备位姿作为集束优化的位姿初始值。

[0103] 具体而言,在步骤S160,进行设备位姿和地图的集束优化。通常,以三维地图(例如,全局稠密地图或剔除了与动态物体对应的像素点的全局稀疏地图)中的地图点在各个关键帧中的投影(投影为关键帧中的像素点)以及各个关键帧所对应的设备位姿为约束,调整设备位姿和三维地图中的地图点的坐标,以降低误差;在一方面,可利用地图点的坐标来优化设备位姿;在另一方面,也可利用设备位姿来优化地图点的坐标,这样的操作被称为“集束优化”。

[0104] 在本公开的示例性实施例中,如果执行深度立体匹配,则获得的全局稠密地图中的地图点的数量会多于在不执行深度立体检测的情况下构建的全局稀疏地图中的地图点的数量。考虑到优化效率,可将可信地图(全局稠密地图)中的点分为两类,即:来自全局稀疏地图的地图点和来自局部稠密地图的地图点。在位姿优化过程中,可基于任意地图点进行设备位姿的集束优化;在地图优化的过程中,仅对来自全局稀疏地图的地图点进行集束优化,以提高优化的效率。在集束优化结束后,可再次执行步骤S440,使优化后的全局稀疏地图参与到地图融合操作。可将集束优化结果存储到存储设备中。

[0105] 另外,在本公开的示例性实施例中,省略上述步骤中的一个或更多个步骤也是可行的。

[0106] 图5示出根据本公开的另一示例性实施例的定位与地图构建的方法的示意图,其中,可参照上文进行初始化、重定位、基于多视角几何理论的立体匹配、将像素点的坐标添加至全局稀疏地图、地图点投影、位姿优化、地图优化等操作。

[0107] 通过初始化或重定位,可初步获得设备位姿和全局稀疏地图。通过位姿优化和地图优化,可提高获得的设备位姿和地图的精度。然而,基于多视角几何理论来确定设备位姿和构建地图会出现上文描述的结果不准确等问题,为了提高准确性,并且使方法和设备具有鲁棒性,本示例性实施例进行如下操作:

[0108] 针对每个关键帧(例如,当前帧),通过深度立体匹配,确定所述每个关键帧中的每个像素点的深度信息,获得局部稠密地图,以获得更多深度信息。针对每个关键帧(例如,当前帧),确定与像素点对应的物体的运动状态(例如,通过静态物体检测,获得静态物体掩膜),以便从全局稀疏地图中剔除与静态物体对应的像素点。

[0109] 作为示例,通过与各个关键帧分别对应的静态物体掩膜对全局稀疏地图进行过滤,以从全局稀疏地图中剔除与动态物体对应的像素点。将经过动态物体剔除的全局稀疏地图与局部稠密地图融合,可获得可信地图(全局稠密地图)。可信地图可用于优化各个关

键帧所对应的设备位姿。

[0110] 作为示例,利用预训练的一个多任务卷积网络(用于深度立体匹配的神经网络),可预测与关键帧对应的视差图,并且通过视差图来获得深度信息以及构建局部稠密地图。利用预训练的另一个多任务卷积网络(用于动态物体检测的神经网络),可预测与关键帧对应的静态物体检测框。利用又一个多任务卷积网络(用于光流计算的神经网络),可预测与关键帧对应的光流图。通过静态物体检测框和光流图可获得静态物体掩膜。

[0111] 作为示例,上述三个多任务卷积网络可具有相同的用于特征提取的网络层。

[0112] 通过深度立体匹配,可预测当前帧(例如,一个关键帧)的任意一个像素点的深度信息,即使当前帧的图像的部分区域不含纹理,或者图像的部分区域的纹理过多和重复,也不会影响像素点的深度信息的预测结果的准确性。可利用预测了深度信息的像素点构建更稠密的被称作局部稠密地图的三维地图(点云),在进行对于设备位姿和地图的优化时利用了通过局部稠密地图和全局稀疏地图融合出的全局稠密地图,利用包含更多深度信息的全局稠密地图进行设备位姿和地图的优化,可获得更准确的设备位姿和全局地图。

[0113] 通过静态物体检测,可预测当前帧中的每一个像素点是否属于静态物体,并可从构建的全局稀疏地图中移除与非静态物体对应的像素点,从而可提高构建的地图在时间维度上的一致性(即:同一物体在任一时刻在三维空间中的位置保持不变),可降低动态物体对设备位姿确定和地图构建的不利影响。

[0114] 由于通过地图融合操作获得的全局稠密地图(全局稠密地图可被用于优化各个帧所对应的设备位姿,优化的设备位姿又可用于优化全局稀疏地图)中,既包括根据多视角几何理论的获取深度信息的像素点(来自全局稀疏地图的像素点),又包括利用深度立体匹配操作获取深度信息的像素点(来自局部稠密地图的像素点),因此,可利用的深度信息更多,从而更容易减少误匹配带来的误差,使得融合出的全局稠密地图可信度更高。因此,可将融合出的地图称为可信地图。利用可信地图进行集束优化,可进一步消除误差,以提高定位和地图构建的准确性。

[0115] 在本公开的示例性实施例,可根据如下步骤来获取与当前帧对应的局部稠密地图:通过特征提取神经网络,对所述当前帧的图像进行特征提取;通过沙漏(Hourglass)神经网络,根据通过特征提取得到的特征图,获得匹配代价体;根据所述匹配代价体,确定所述当前帧中的各像素点的深度信息,得到与所述当前帧对应的局部稠密地图。

[0116] 作为示例,通过Hourglass神经网络,根据通过特征提取得到的特征图,获得匹配代价体的步骤,包括:根据所述特征图,获取初始匹配代价体;通过Hourglass神经网络,基于所述初始匹配代价体,获得所述匹配代价体。

[0117] 作为示例,所述特征图包括左目图像的特征图和右目图像的特征图,其中,根据所述特征图,获取初始匹配代价体的步骤,包括:将左目图像的特征图和右目图像的特征图进行连接,以获得初始匹配代价体。

[0118] 作为示例,根据所述匹配代价体,确定所述当前帧中的各像素点的深度信息的步骤,包括:通过卷积神经网络,对所述匹配代价体进行空间卷积处理;根据空间卷积处理后的匹配代价体,确定所述当前帧中的各像素点在左目图像和右目图像之间的视差;根据各像素点的视差,确定所述当前帧中的各像素点的深度信息,得到与所述当前帧对应的局部稠密地图。

[0119] 作为示例,对所述匹配代价体进行空间卷积处理的步骤,包括:对所述匹配代价体按照设定的多个方向分别执行如下处理:按照当前设定方向,对所述匹配代价体进行分割,得到多个匹配代价层,按照当前设定方向依次对各匹配代价层执行卷积处理,其中,当对任一匹配代价层进行卷积处理时,将所述任一匹配代价层累加上一匹配代价层的卷积结果后进行卷积处理,具体地,可不改变第一个匹配代价层,从第二个匹配代价层开始,将任一匹配代价层更新为所述任一匹配代价层与上一卷积结果(即:上一匹配代价层与卷积核的乘积)之和。

[0120] 作为示例,确定所述当前帧中的各像素点在左目图像和右目图像之间的视差的步骤,包括:利用柔性最大传递函数以及空间卷积处理后的匹配代价体,获得所述当前帧中的各像素点在左目图像和右目图像之间的视差概率分布;确定与各像素点对应的视差概率分布的累积值作为对应像素点在左目图像和右目图像之间的视差。

[0121] 可通过图6所示实施例来执行深度立体匹配。图6示出根据本公开的示例性实施例的进行深度立体匹配操作的示意图,在执行图6所示示例性实施例时,可对左目图像和右目图像进行校正(例如,光路校正)以获得准确的图像。

[0122] 如图6中所示,可通过如下步骤进行深度立体匹配:将当前帧的左目图像输入到串联的预训练的二维卷积神经网络和预训练的空间金字塔池化网络以获取左目图像的特征图(例如,以矩阵形式表示的特征图,也被称作左特征图);将当前帧的右目图像输入到另一对串联的二维卷积神经网络和空间金字塔池化网络以获取右目图像的特征图(也被称作右特征图),其中,两个二维卷积神经网络共享权值,和/或两个空间金字塔池化网络共享权值;沿着视差方向对获取的左特征图和右特征图的进行连接,以获得初始匹配代价体(匹配代价体也可以被视作匹配代价矩阵),其中,可使用任何已知的方法连接两幅特征图,例如,假定两幅特征图的维度均是 $m \times n \times c$ (m 、 n 和 c 为自然数),若视差为0,则直接连接两幅特征图,得到大小为 $m \times n \times 2c$ 的矩阵,若视差为 d (d 为自然数),则先将右特征图沿着预先确定的 x 轴方向平移 d 列,然后,再将左特征图与右特征图连接,这样会得到 d 个大小为 $m \times n \times 2c$ 的矩阵;利用预训练的Hourglass网络(例如,Hourglass神经网络)对初始匹配代价体进行计算以获得预测的匹配代价体;对预测的匹配代价体进行卷积以获得经过卷积的匹配代价体;利用包括柔性最大传递函数(例如,softmax函数)的回归函数对经过卷积的匹配代价体进行计算以获得与左目图像和右目图像中多对匹配像素点(一对匹配像素点对应场景中的同一点)中的每对匹配像素点对应的视差概率分布(即:与所述每对匹配像素点对应各个可能的视差值及对应的概率);根据获得的视差概率分布计算所述每对匹配像素点的视差(对于每对匹配像素点,可利用计算累积值(例如,数学期望)的方法来计算视差,例如,将每个可能的视差与对应于所述每个可能的视差的概率相乘,然后将各个乘积相加即可得到与所述每对匹配像素点对应的视差)。

[0123] 执行对预测的匹配代价体进行卷积以获得经过卷积的匹配代价体的步骤是因为考虑到:图像上相邻点之间的深度变化不显著(可认为相邻点的深度连续),像素点的视差可以是与像素点在左目图像和右目图像之间的视差,为了使得相邻像素点之间的视差变化不显著(例如,表现为视差更光滑或空间上连续),可利用卷积神经网络(例如,空间卷积神经网络),对匹配代价体进行卷积(例如,空间卷积),这样,可以增加相邻像素点的视差的空间连续性,并且可以去掉一些噪声点。在本公开中,通过卷积,还可将可信度比较高的深度

信息传递到深度信息的可信度比较低的像素点。

[0124] 图7示出根据本公开的示例性实施例的对匹配代价体进行卷积的操作的示意图。

[0125] 如图7中所示,对匹配代价体进行卷积以获得经过卷积的匹配代价体的步骤包括:

[0126] 利用与匹配代价体的高度方向垂直的多个第一平面,将匹配代价体分割为沿着平行于高度方向的第一方向(图中所示的从上到下的方向)依次排列的多个第一匹配代价层;

[0127] 利用预定卷积核沿着第一方向从第二个第一匹配代价层开始依次对各个第一匹配代价层进行更新,以将每个第一匹配代价层更新为所述每个第一匹配代价层加上前一个第一子匹配代价层与预定卷积核的乘积,并且在最后一个第一子匹配代价层进行更新后,将各个第一子匹配代价层合并为第一卷积匹配代价体;

[0128] 利用与所述高度方向垂直的多个第二平面,将第一卷积匹配代价体分割为沿着与第一方向相反的第二方向(图中所示的从下到上的方向)依次排列的多个第二子匹配代价层;

[0129] 利用预定卷积核沿着第二方向从第二个第二匹配代价层开始依次对各个第二匹配代价层进行更新,以将每个第二匹配代价层更新为所述每个第二匹配代价层加上前一个第二子匹配代价层与预定卷积核的乘积,并且在最后一个第二子匹配代价层进行更新后,将各个第二子匹配代价层合并为第二卷积匹配代价体;

[0130] 利用与匹配代价体的长度方向或宽度方向垂直的多个第三平面,将第二卷积匹配代价体分割为沿着垂直于第三平面的第三方向(图中所示的从左到右的方向)依次排列的多个第三子匹配代价层;

[0131] 利用预定卷积核沿着第三方向从第二个第三匹配代价层开始依次对各个第三匹配代价层进行更新,以将每个第三匹配代价层更新为所述每个第三匹配代价层加上前一个第三子匹配代价层与预定卷积核的乘积,并且在最后一个第三子匹配代价层进行更新后,将各个第三子匹配代价层合并为第三卷积匹配代价体;

[0132] 利用与第三平面平行的多个第四平面,将第三卷积匹配代价体分割为沿着与第三方向相反的第四方向(图中所示的从右到左的方向)依次排列的多个第四子匹配代价层;

[0133] 利用预定卷积核沿着第四方向从第二个第四匹配代价层开始依次对各个第四匹配代价层进行更新,以将每个第四匹配代价层更新为所述每个第四匹配代价层加上前一个第四子匹配代价层与预定卷积核的乘积,并且在最后一个第四子匹配代价层进行更新后,将各个第四子匹配代价层合并为第四卷积匹配代价体,作为经过卷积的匹配代价体。

[0134] 如上文参照图7的描述,可从匹配代价体的高度和宽度(或长度)两个维度进行空间滤波(即,进行匹配代价累积),可具体分为按照如下四个方向的四个过程:从上到下,从下到上,从左到右和从右到左。这四个过程计算方法是类似的。既进行从上到下的过程又进行从下到上的过程(或者既进行从左到右的过程又进行从右到左的过程)是为了使得代价平衡。通过对子匹配代价层进行更新,可以增加所述任意一个方向的匹配代价的连续性。

[0135] 传统的对像素点的深度的预测方法通常采用端到端的训练,没有将一些立体匹配方法和卷积神经网络结合,例如,未进行上述空间滤波,这使得传统的方法在处理物体的边界时,往往存在边界断开的情况。然而,当经过上述空间滤波(卷积)后的匹配代价体减少或不存在断开的边界。

[0136] 图8示出了根据本公开的示例性实施例的优化设备位姿(例如,相机位姿)过程和

优化地图点深度过程的示意图,其中,地图点可被理解为关键帧中的具有全局地图坐标的像素点。

[0137] 如图8中所示,在优化相机位姿的过程中,针对当前帧,可获得(例如,通过地图点投影来获得)新匹配的像素点,可用新匹配的像素点作为原有匹配的像素点的替代,参照当前帧之前的关键帧,将原有的相机位姿更新为新的相机位姿;在优化地图点深度的过程中,针对当前帧,新匹配的像素点具有新的深度信息(更新后的深度),可结合所述新的相机位姿,获得新匹配的像素点的坐标,以更新全局地图。

[0138] 在图8所示的示例性实施例中,需要确定的深度信息,可基于多视角几何理论确定深度信息。多视角几何理论的应用前提是满足如下假设,这样的假设包括:假定相机所处场景中的与将要计算深度信息的像素点对应的物体是静态的、不运动的。

[0139] 然而,在实际应用中,场景中的物体往往是动态的。例如,在自动驾驶应用中,路面上的车是动态的,并且当将相机所在的车上的点作为地图点参与上述优化过程时,会为优化结果带来误差。

[0140] 为了满足物体静态的假设,可确定当前帧中的属于动态物体的像素点,并且基于确定的结果从地图中剔除与动态物体对应的像素点。

[0141] 作为示例,确定当前帧中的属于动态物体的像素点的步骤,包括:确定与所述当前帧对应的物体状态信息;基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

[0142] 作为示例,确定与所述当前帧对应的物体状态信息的步骤,包括:通过特征提取神经网络,对所述当前帧的图像进行特征提取;根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

[0143] 作为示例,根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息的步骤,包括:确定所述当前帧与所述当前帧的前一帧之间的光流信息;根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

[0144] 采用静态物体掩膜来确定物体状态信息是因为:在实践中,申请人发现,基于图像分割方法,可直接获得不同种类物体的掩膜,从而直接根据通过对像素点进行语义分割而获取的语义信息,将动态属性很弱(被识别为动态物体的可能性低)的物体,如交通标示、路面、门窗等作为静态物体。然而,图像分割方法通常根据单帧进行语义分割以识别静态物体,难以识别难以与周围环境区分的物体边界属于静态物体还是动态物体。

[0145] 为了避免针对单帧进行识别所带来的问题,考虑了利用用于表示帧间变化(例如,位置变化)的“光流”帮助物体识别。

[0146] 图9示出根据本公开的示例性实施例的获取静态物体掩膜的操作的示意图。

[0147] 如图9中所示,可输入双目视频流,输入的双目视频流包括连续的多个帧(关键帧)的图像,每个图像包括左目图像和右目图像。可利用参数共享的多个特征提取网络来分别提取各个关键帧的特征图。

[0148] 随后,利用训练的用于光流检测的神经网络(为了描述方便,称作第一卷积网络),针对当前关键帧的特征图和当前帧的前一帧的特征图进行预测,以获得当前关键帧与当前

关键帧的前一帧之间的光流图;利用预训练的用于动静检测的神经网络(为了描述方便,称作第二卷积网络),针对当前帧的特征图进行预测,以获得与当前关键帧对应的静态物体包围框。

[0149] 将静态物体包围框与流光图连接;利用预训练的用于动态状态预测的神经网络(为了描述方便,称作第三卷积网络),针对连接的静态物体包围框和流光图进行预测,以获得静态物体掩膜。

[0150] 作为示例,图6中的二维卷积神经网络、图9的特征提取网络可共享权值。

[0151] 在本示例性实施例中,经过训练的第二卷积网络,可以用于预测物体类别属性信息,例如,将物体划分为多个类别,属于第一类别的车、人、动物等,具有较强的动态属性(被识别为动态物体的可能性高),而属于第二类别的交通标志、桌子椅子、路面等,具有较弱的动态属性。将动态属性较弱的物体,作为潜在的静态物体。

[0152] 另外,可将关键帧的特征图和关键帧的前一帧的特征图进行连接,并利用第一卷积网络对连接后的特征图进行预测,获得前后两帧之间的光流信息(例如,光流图)。将光流信息与静态物体包围框作为第三卷积神经网络的输入,以预测静态物体掩膜。静态物体掩膜可以是二值图,可表示当前关键帧中每个像素点的类别(该类别包括静态物体类别和动态物体类别)。

[0153] 为了进一步提高定位的准确性,可对地图进行融合,以便综合各种地图的优点。

[0154] 作为示例,可将经过像素点剔除的当前的全局稀疏地图和局部稠密地图融合为全局稠密地图,具体包括:对所述局部稠密地图进行网格划分;将划分的各个网格的顶点更新到经过像素点剔除的全局稀疏地图;将各个网格的网格内像素点更新到经过像素点剔除的全局稀疏地图,其中,全局稠密地图包括:经过像素点剔除、网格顶点更新、以及网格内像素点更新的全局稀疏地图。

[0155] 图10示出根据本公开的示例性实施例的地图融合操作的示意图。

[0156] 如图10中所示,针对输入的图像,通过深度立体匹配操作可获得局部稠密地图,通过静态物体检测可获得静态物体掩膜,通过基本SLAM操作(例如,与图2或图3有关的描述)可获得全局稀疏地图。可利用静态物体掩膜从全局稀疏地图中剔除与动态物体对应的像素点。对局部稠密地图进行网格划分,将划分网格后的局部稠密地图与进行像素点剔除的全局稀疏地图进行融合,以获得全局稠密地图。融合过程可包括:更新作为网格顶点的像素点的深度信息;更新网格内像素点的深度信息,获得的全局稠密地图是可信的,可用于上文描述的集束优化。

[0157] 在上述操作中,进行网格划分是为了提高地图融合速度。在进行网格划分后,可先更新作为网格顶点的像素点的深度信息,后更新网格内像素点的深度信息,从而从粗到细的更新像素点的深度信息。

[0158] 对于已有的深度图拼接方法,需要两幅深度图中存在较多重合区域(重合区域中的像素点的深度信息相同)才能进行拼接,这造成约束多,从而导致效率低的问题。

[0159] 此外,还可以利用离线的高精地图和/或雷达地图等辅助定位,即:将高精地图和/或雷达地图与全局稠密地图进行融合,并将融合后的地图用于集束优化。

[0160] 作为示例,在全局稀疏地图中,不但含有大量噪声点,而且像素点(地图点)较为稀疏,难以满足实际应用的需求(例如,根据地图了解环境状况)。针对地图点稀疏的问题,可

通过深度立体匹配操作获取稠密地图,即:与关键帧对应的局部稠密点云地图;针对存在噪声点的问题,通过静态物体检测操作,消除由动态物体在时间维度上的位置变化引入的噪声。可信的场景三维信息是准确进行定位和环境感知的基础,可通过有效利用多种地图来获得可信的场景三维信息。例如,将全局稀疏地图(尤其是剔除了与动态物体对应的像素点的全局稀疏地图)与局部稠密地图融合为可信地图。

[0161] 作为示例,在将局部稠密地图和全局稀疏地图融合时,考虑到全局稀疏地图中的地图点的深度信息是基于严格的几何关系(基于多视角几何理论的集合关系)计算出的,该深度信息更为准确,因此,可将局部稠密地图中的地图点向全局稀疏地图中更新。可先更新部分地图点(例如,网格顶点)的深度信息,然后,插值出其余像素点的深度信息。

[0162] 本公开不限制划分网格的方式,任何在地图中划分网格的方式均是可行的。

[0163] 更具体地,将任一顶点更新到经过像素点剔除的全局稀疏地图的步骤,包括:在与所述任一顶点所属的各网格对应的各个顶点中,确定具有全局稀疏地图坐标的顶点;根据确定的顶点的深度信息、所述任一顶点在所述局部稠密地图中的深度信息、以及确定的顶点与所述任一顶点之间的距离,获得所述任一顶点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述当前帧对应的设备位姿和所述任一顶点在经过像素点剔除的全局稀疏地图中的深度信息,确定所述任一顶点在全局稀疏地图中的坐标,以更新全局稀疏地图。

[0164] 将网格内任一像素点更新到经过像素点剔除的全局稀疏地图的步骤,包括:根据所述任一像素点所在网格的顶点的深度信息、所述任一像素点在所述局部稠密地图中的深度信息、以及所述任一像素点所在网格的顶点与所述任一像素点之间的距离,获得所述任一顶点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述当前帧对应的设备位姿和所述任一像素点在经过像素点剔除的全局稀疏地图中的深度信息,确定所述任一像素点在全局稀疏地图中的坐标,以更新全局稀疏地图。

[0165] 作为示例,基于所述全局稠密地图,对与各帧分别对应的设备位姿以及当前的全局稀疏地图进行优化的步骤,包括:基于所述全局稠密地图、与各关键帧分别对应的设备位姿、以及所述当前的全局稀疏地图,对与各关键帧分别对应的设备位姿以及所述当前的全局稀疏地图进行集束优化。

[0166] 图11示出根据本公开的示例性实施例的基于网格划分进行地图融合的操作的示意图。

[0167] 在图11所示的示例性实施例中,对局部稠密地图进行网格划分,在将各个网格顶点依次更新到全局稀疏地图的过程中,可通过如下公式计算作为网格顶点的像素点p在全局稠密地图中的深度信息 d_p :

$$[0168] \quad d_p = \frac{d_{l,p}}{N_p \times \sum_{q \in \Omega_p \setminus M} D(p, q)} \sum_{q \in \Omega_p \setminus M} \frac{d_{g,q} \times D(p, q)}{d_{l,q}}$$

[0169] 通过如下公式计算网格内像素点p'在全局稠密地图中的深度信息 $d_{p'}$:

$$[0170] \quad d_{p'} = \frac{d_{l,p'}}{4} \sum_{k \in [1,4]} \frac{d_{q_k} \times D(p', q_k)}{d_{l,q_k}}$$

[0171] 其中, $d_{l,p}$ 表示像素点 p 在局部稠密地图中的深度信息, $d_{g,q}$ 表示像素点 q 在全局稀疏地图中的深度信息, q 表示集合 $\Omega_p \setminus M$ 中的像素点, Ω_p 表示局部稠密地图中与像素点 p 相邻的四个网格中的像素点的集合, $\Omega_p \setminus M$ 表示从 Ω_p 中去除了与动态物体对应的像素点之后的集合, N_p 表示 Ω_p 中的各个像素点中具有全局稀疏地图坐标的像素点的数量, $D(p, q)$ 表示像素点 p 与像素点 q 之间的距离, $d_{l,q}$ 表示像素点 q 在局部稠密地图中的深度信息, $d_{l,p'}$ 表示像素点 p' 在局部稠密地图中的深度信息, q_k 表示像素点 p' 所在的网格的顶点, d_{q_k} 表示 q_k 在全局稀疏地图中的深度信息, d_{l,q_k} 表示顶点 q_k 在局部稠密地图中的深度信息, $D(p', q_k)$ 表示像素点 p' 与顶点 q_k 之间的距离。

[0172] 另外, 需要说明的, 地图融合的方式包括但不限于: 将高精地图和/或雷达地图, 与全局稀疏地图融合或剔除了与动态物体对应的像素点的全局稀疏地图融合。

[0173] 作为示例, 基于所述全局稠密地图, 对各帧分别对应的设备位姿以及当前的全局稀疏地图进行优化的步骤, 包括: 基于所述全局稠密地图、与各关键帧分别对应的设备位姿、以及所述当前的全局稀疏地图, 对各关键帧分别对应的设备位姿以及所述当前的全局稀疏地图进行集束优化。

[0174] 图12示出根据本公开的示例性实施例的进行集束优化操作的示意图。在图12所示的示例性实施例中, 可将全局稠密地图中的地图点分为两类: 来自全局稀疏地图的地图点和来自局部稠密地图 (且不在全局稀疏地图中, 即: 不具有全局稀疏地图的地图点) 的地图点。在集束优化操作中, 可基于全局稠密地图对各关键帧分别对应的设备位姿进行优化 (对应图12中的位姿优化), 具体的, 基于全局稠密地图中不在全局稀疏地图中的地图点, 对各关键帧分别对应的设备位姿进行优化, 即可利用局部稠密地图中的地图点进行设备位姿的优化, 此外, 基于优化后的各关键帧分别对应的设备位姿, 对当前的全局稀疏地图进行优化 (对应图12中的地图优化), 使得全局稀疏地图中的地图点将随着优化的设备位姿而优化, 但是不改变局部稠密地图中的地图点。每进行一次集束优化后, 可利用地图融合方法, 得到更新的全局稠密地图 (将优化后的全局稀疏地图和局部稠密地图进行融合, 得到更新的全局稠密地图), 用于下一次的集束优化。

[0175] 根据本公开的另一示例性实施例, 当基于所述当前帧的图像, 确定了与所述当前帧对应的设备位姿以及当前的全局稀疏地图时, 可确定所述当前帧中的属于动态物体的像素点; 从所述当前的全局稀疏地图中剔除属于动态物体的像素点, 其中, 可针对优化的或未优化的全局稀疏地图进行像素点剔除。上文已经描述了, 从全局稀疏地图中剔除属于动态物体的像素点的具体操作, 这里不再赘述。

[0176] 根据本公开, 在针对存在动态物体、大面积的无纹理区、剧烈的光照变化等场景进行定位时, 也可具有较强的鲁棒性。通过深度立体匹配操作, 可根据单帧立体图像预测深度信息, 从而为地图构建提供大量有效的地图点。通过静态物体检测操作, 可识别出场景中的动态物体, 并将与动态物体对应的地图点移除, 以保证设备位姿和地图点的几何约束的正确性。通过地图融合操作, 了将不同的地图融合, 可获得稠密的全局地图, 用于集束优化, 从而可提高针对设备位姿和地图的优化的鲁棒性。

[0177] 图13示出根据本公开的示例性实施例的电子设备的框图。如图13中所示, 根据本公开的示例性实施例的电子设备包括: 图像获取单元1301, 用于获取当前帧的图像; 位姿与

地图确定单元1302,用于基于所述当前帧的图像,确定与所述当前帧对应的设备位姿以及当前的全局稀疏地图。

[0178] 作为示例,所述电子设备还包括:全局稠密地图获取单元,用于当当前帧为关键帧时,获取与所述当前帧对应的全局稠密地图;优化单元,基于所述全局稠密地图,对与各关键帧分别对应的设备位姿以及当前的全局稀疏地图进行优化。

[0179] 作为示例,全局稠密地图获取单元包括:局部稠密地图获取单元,用于获取与所述当前帧对应的局部稠密地图;其中,全局稠密地图获取单元,基于所述局部稠密地图和当前的全局稀疏地图,确定与所述当前帧对应的全局稠密地图。

[0180] 作为示例,局部稠密地图获取单元,通过特征提取神经网络,对所述当前帧的图像进行特征提取;通过沙漏Hourglass神经网络,根据通过特征提取得到的特征图,获得匹配代价体;根据所述匹配代价体,确定所述当前帧中的各像素点的深度信息,得到与所述当前帧对应的局部稠密地图。

[0181] 作为示例,局部稠密地图获取单元,根据所述特征图,获取初始匹配代价体;通过Hourglass神经网络,基于所述初始匹配代价体,获得所述匹配代价体。

[0182] 作为示例,所述特征图包括左目图像的特征图和右目图像的特征图,其中,局部稠密地图获取单元,将左目图像的特征图和右目图像的特征图进行连接,以获得初始匹配代价体。

[0183] 作为示例,局部稠密地图获取单元,通过卷积神经网络,对所述匹配代价体进行空间卷积处理;根据空间卷积处理后的匹配代价体,确定所述当前帧中的各像素点在左目图像和右目图像之间的视差;根据各像素点的视差,确定所述当前帧中的各像素点的深度信息,得到与所述当前帧对应的局部稠密地图。

[0184] 作为示例,局部稠密地图获取单元,对所述匹配代价体按照设定的多个方向分别执行如下处理:按照当前设定方向,对所述匹配代价体进行分割,得到多个匹配代价层,按照当前设定方向依次对各匹配代价层执行卷积处理,其中,当对任一匹配代价层进行卷积处理时,将所述任一匹配代价层累加上一匹配代价层的卷积结果后进行卷积处理。

[0185] 作为示例,局部稠密地图获取单元,利用柔性最大传递函数以及空间卷积处理后的匹配代价体,获得所述当前帧中的各像素点在左目图像和右目图像之间的视差概率分布;确定与各像素点对应的视差概率分布的累积值作为对应像素点在左目图像和右目图像之间的视差。

[0186] 作为示例,全局稠密地图获取单元包括:动态物体确定单元,用于确定所述当前帧中的属于动态物体的像素点;从所述当前的全局稀疏地图中剔除属于动态物体的像素点,其中,全局稠密地图获取单元,将经过像素点剔除的所述当前的全局稀疏地图和所述局部稠密地图融合为全局稠密地图。

[0187] 作为示例,动态物体确定单元,确定与所述当前帧对应的物体状态信息;基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

[0188] 作为示例,动态物体确定单元,通过特征提取神经网络,对所述当前帧的图像进行特征提取;根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

[0189] 作为示例,动态物体确定单元,确定所述当前帧与所述当前帧的前一帧之间的光

流信息;根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

[0190] 作为示例,全局稠密地图获取单元包括融合单元,用于对所述局部稠密地图进行网格划分;将划分的各个网格的顶点更新到经过像素点剔除的全局稀疏地图;将各个网格的网格内像素点更新到经过像素点剔除的全局稀疏地图。

[0191] 作为示例,融合单元,确定与所述当前顶点相邻的网格的各个顶点中具有全局稀疏地图坐标的顶点;根据确定的顶点的深度信息、所述当前顶点在所述局部稠密地图中的深度信息、以及确定的顶点与所述当前顶点之间的距离,获得所述当前顶点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述当前顶点对应的设备位姿和获得的深度信息,确定所述当前顶点在全局稀疏地图中的坐标,以更新全局稀疏地图。

[0192] 作为示例,融合单元,根据网格内像素点所在网格的顶点的深度信息、网格内像素点在所述局部稠密地图中的深度信息、以及网格内像素点所在网格的顶点与网格内像素点之间的距离,获得所述当前顶点在经过像素点剔除的全局稀疏地图中的深度信息,并且根据与所述网格内像素点对应的设备位姿和获得的深度信息,确定网格内像素点在全局稀疏地图中的坐标,以更新全局稀疏地图。

[0193] 作为示例,所述优化单元,基于所述全局稠密地图、与各关键帧分别对应的设备位姿、以及所述当前的全局稀疏地图,对与各关键帧分别对应的设备位姿以及所述当前的全局稀疏地图进行集束优化。

[0194] 作为示例,所述电子设备还包括:动态物体确定单元,用于确定所述当前帧中的属于动态物体的像素点;从所述当前的全局稀疏地图中剔除属于动态物体的像素点。

[0195] 作为示例,动态物体确定单元,确定与所述当前帧对应的物体状态信息;基于所述物体状态信息,确定所述当前帧中的属于动态物体的像素点。

[0196] 作为示例,动态物体确定单元,通过特征提取神经网络,对所述当前帧的图像进行特征提取;根据通过特征提取得到的特征图,确定所述当前帧中的各物体的类别属性信息;根据所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的物体状态信息。

[0197] 作为示例,动态物体确定单元,确定所述当前帧与所述当前帧的前一帧之间的光流信息;根据所述光流信息和所述当前帧中的各物体的类别属性信息,获取与所述当前帧对应的静态物体掩膜,得到与所述当前帧对应的物体状态信息。

[0198] 根据本公开的另一示例性实施例,提供了一种电子设备,其中,所述电子设备包括:处理器;存储器,存储有计算机程序,当所述计算机程序被处理器执行时,实现如上所述的方法。

[0199] 应该理解,根据本公开的示例性实施例的电子设备的具体实现方式可参照结合图1至图12描述的相关具体实现方式来实现,在此不再赘述。

[0200] 根据本公开的另一示例性实施例,提供了一种存储有计算机程序的计算机可读存储介质,其中,当所述计算机程序被处理器执行时实现如上所述的方法。

[0201] 计算机可读存储介质是可存储由计算机系统读出的数据的任意数据存储装置。计算机可读存储介质的示例包括:只读存储器、随机存取存储器、只读光盘、磁带、软盘、光数据存储装置和载波(诸如经有线或无线传输路径通过互联网的数据传输)。

[0202] 此外,应该理解,根据本公开示例性实施例的电子设备的各个单元可被实现硬件

组件和/或软件组件。本领域技术人员根据限定的各个单元所执行的处理,可以例如使用现场可编程门阵列(FPGA)或专用集成电路(ASIC)来实现各个单元。

[0203] 此外,根据本公开示例性实施例的方法可以被实现为计算机可读存储介质中的计算机代码。本领域技术人员可以根据对上述方法的描述来实现所述计算机代码。当所述计算机代码在计算机中被执行时实现本公开的上述方法。

[0204] 虽然已表示和描述了本公开的一些示例性实施例,但本领域技术人员应该理解,在不脱离由权利要求及其等同物限定其范围的本公开的原理和精神的情况下,可以对这些实施例进行修改。

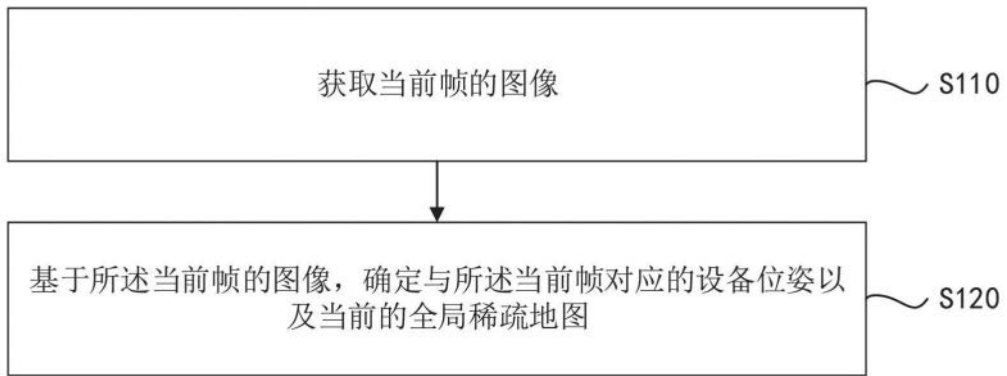


图1

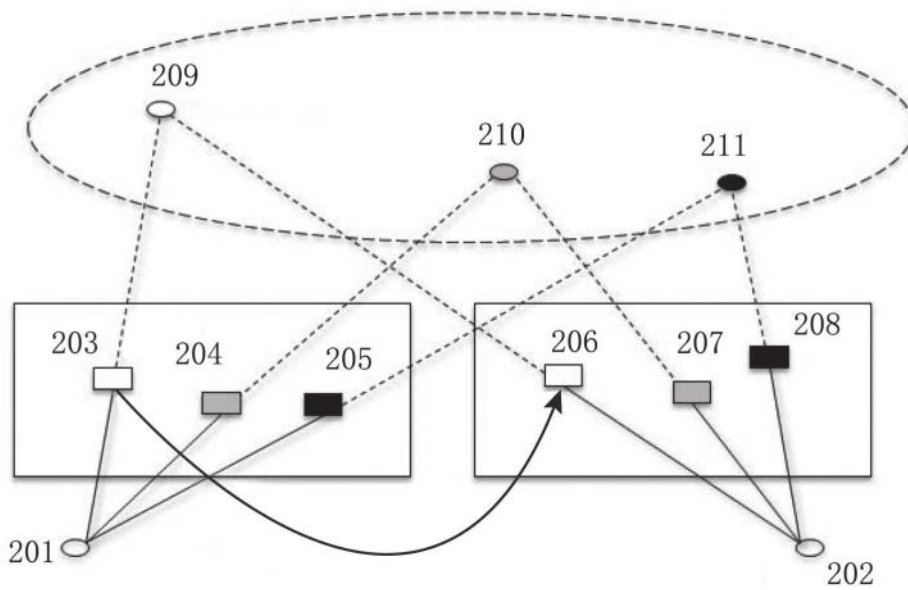


图2

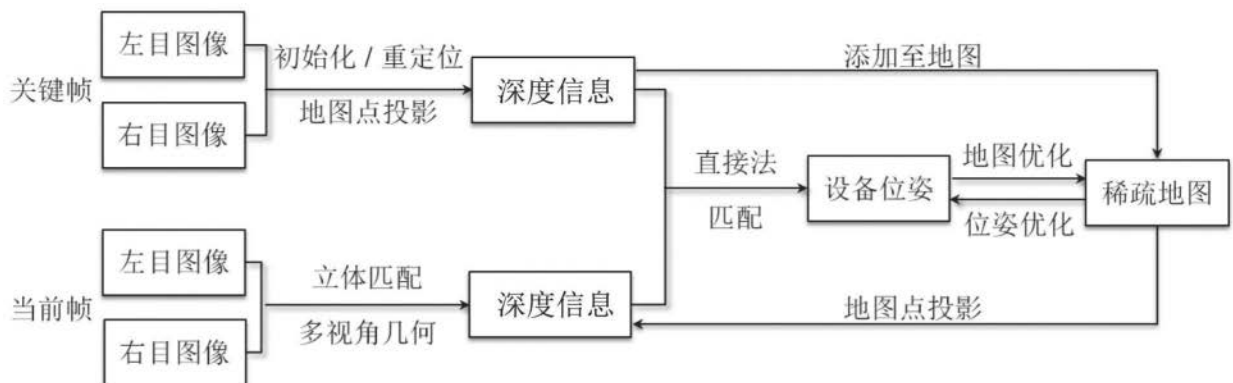


图3

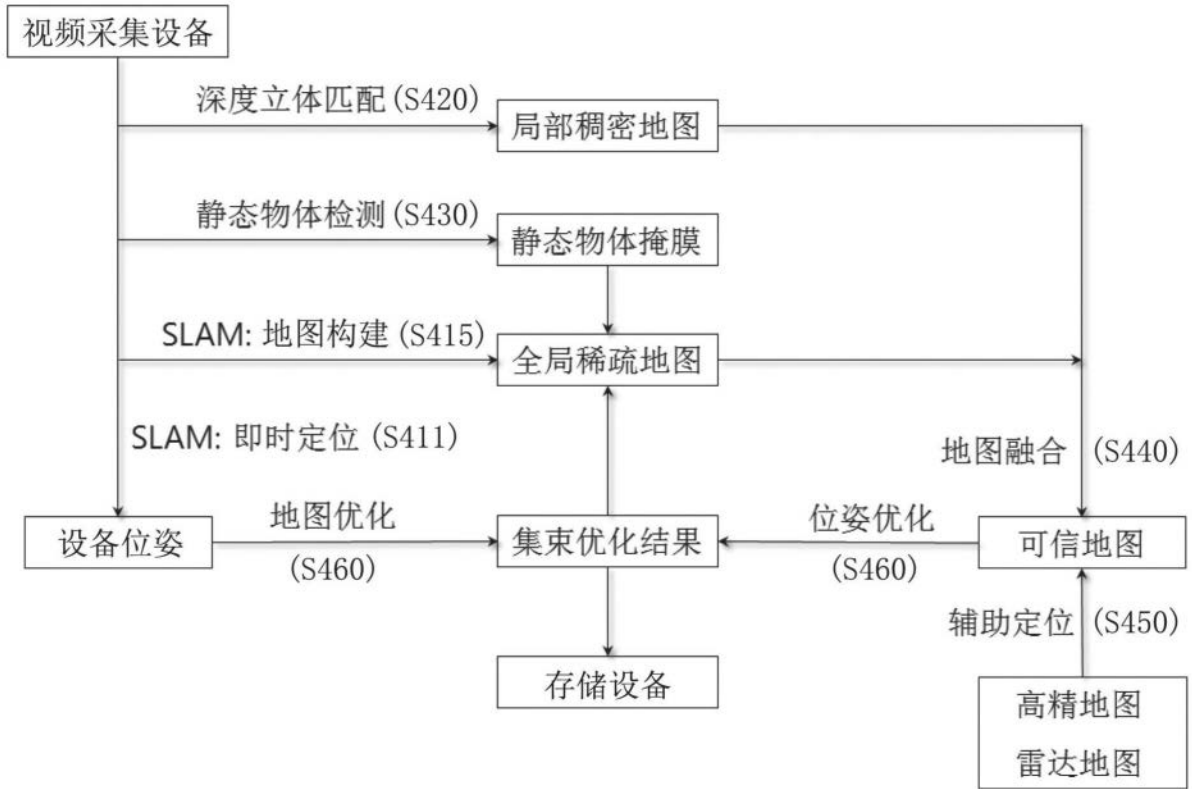


图4

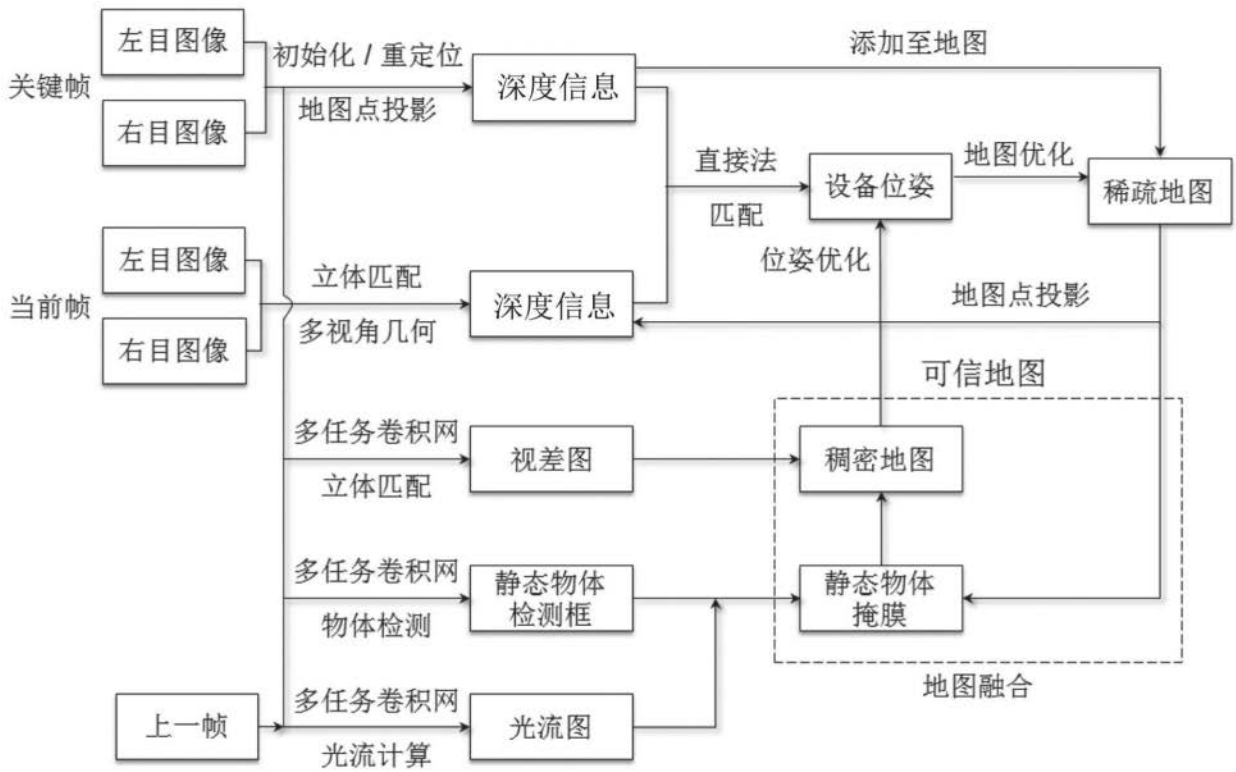


图5

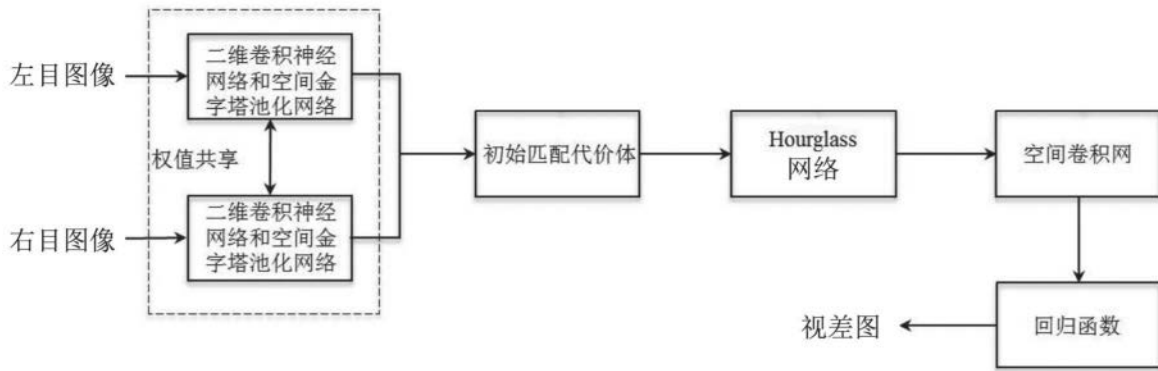


图6

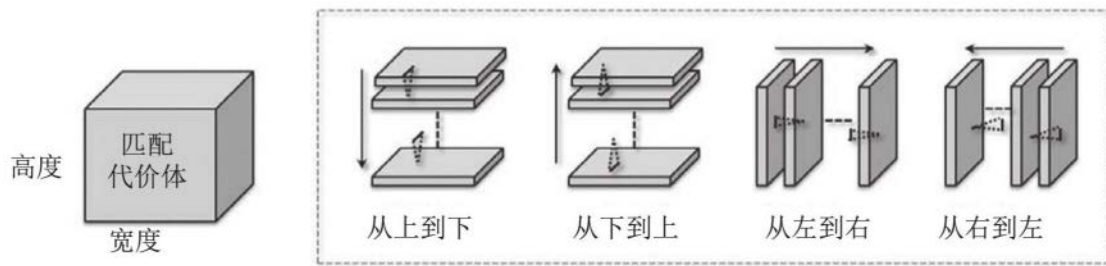


图7

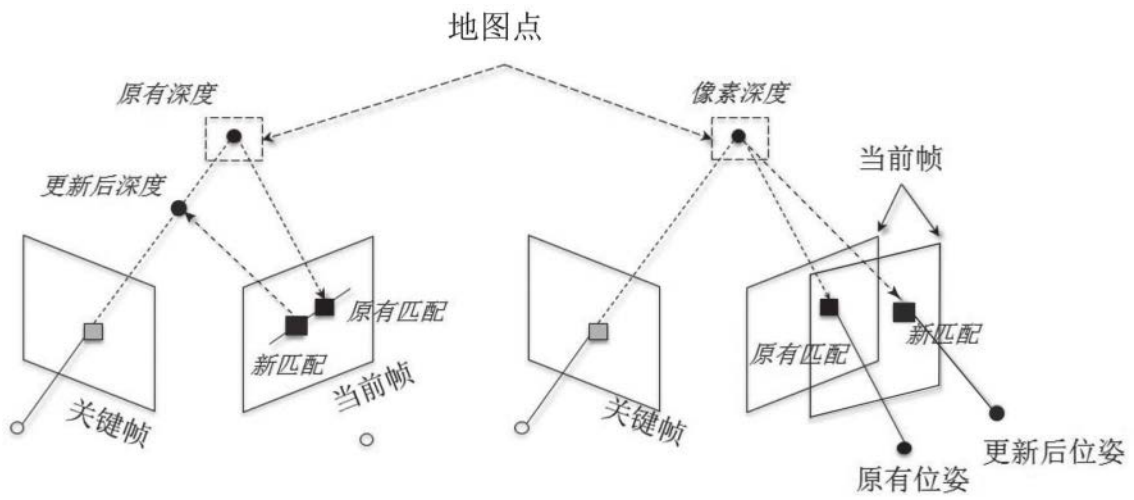


图8

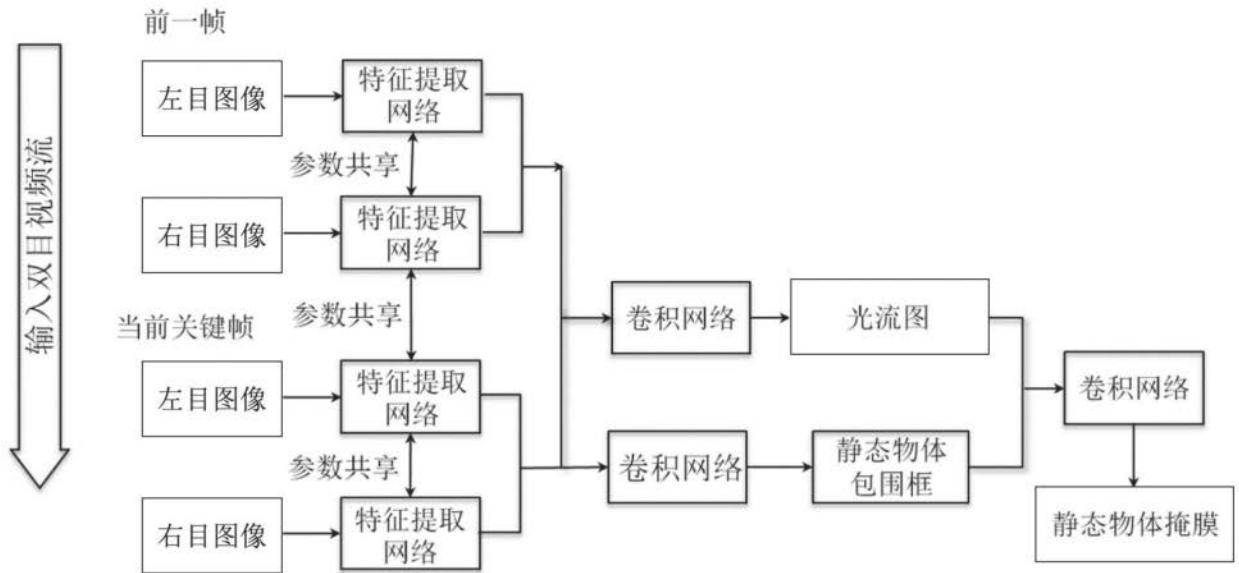


图9

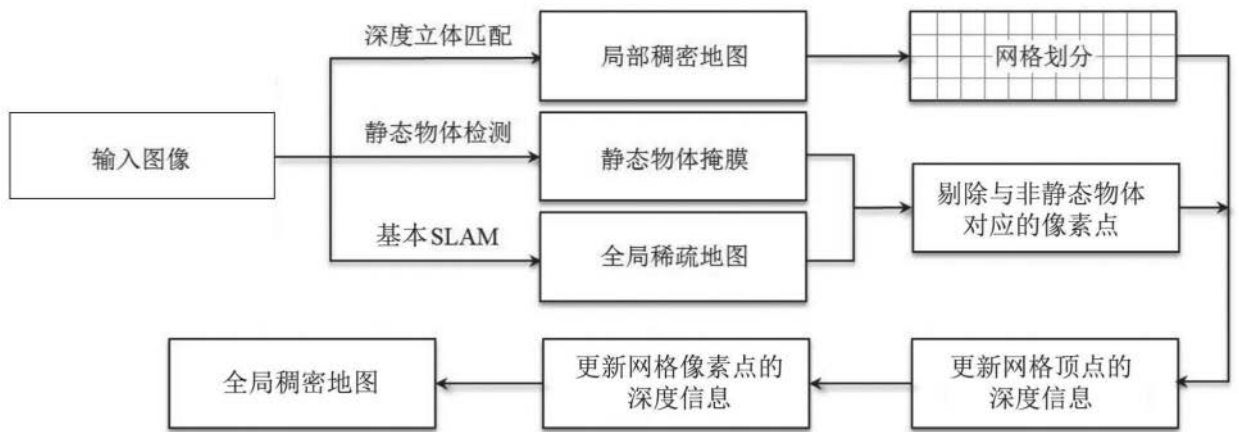


图10

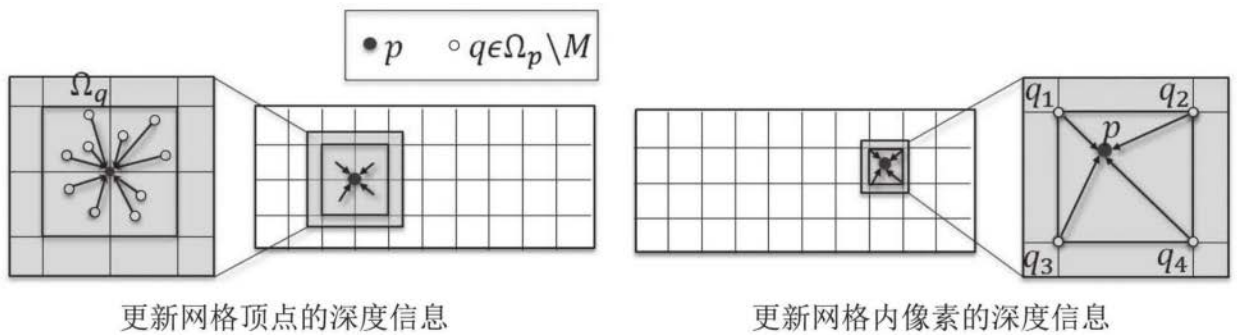


图11

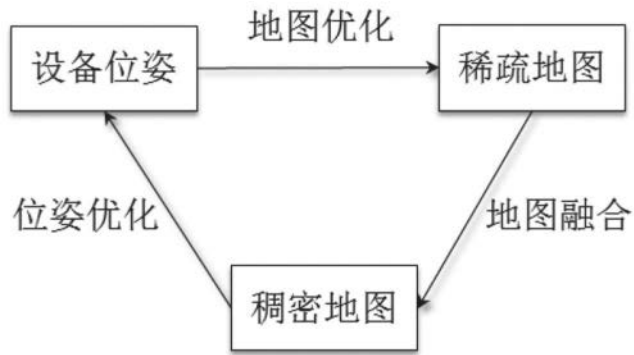


图12

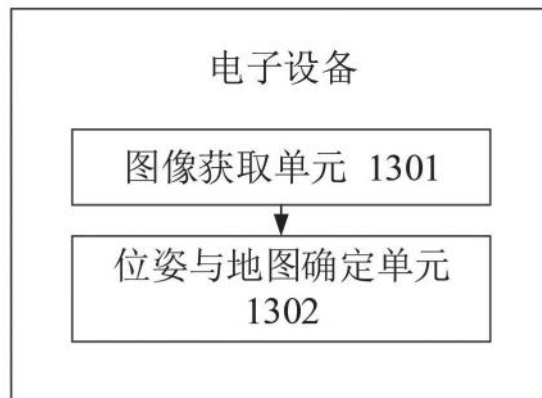


图13