



(12) 发明专利申请

(10) 申请公布号 CN 113838129 A

(43) 申请公布日 2021. 12. 24

(21) 申请号 202110925485.5

(22) 申请日 2021.08.12

(71) 申请人 高德软件有限公司

地址 102200 北京市昌平区科技园区昌盛路18号B1座1-5层

(72) 发明人 陶醉

(74) 专利代理机构 北京清源汇知识产权代理事务所(特殊普通合伙) 11644

代理人 冯德魁 张艳梅

(51) Int. Cl.

G06T 7/73 (2017.01)

G06F 16/29 (2019.01)

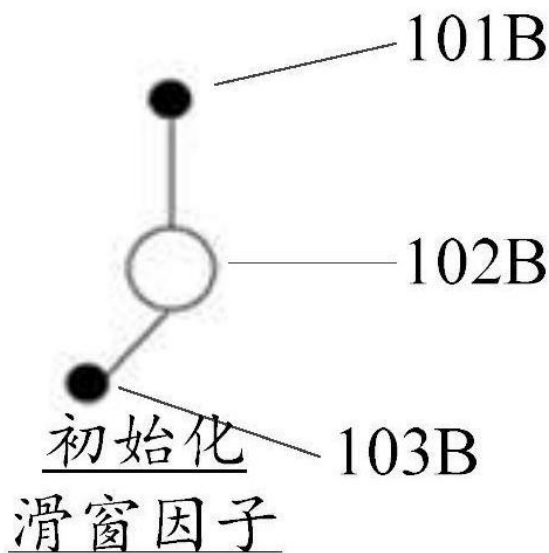
权利要求书2页 说明书19页 附图4页

(54) 发明名称

一种获得位姿信息的方法、装置以及系统

(57) 摘要

本申请公开了一种获得位姿信息的方法,包括:获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;获得包含该实景图像所反映地理位置的三维矢量地图;在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X。



1. 一种获得位姿信息的方法,其中,包括:

获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;

获得包含该实景图像所反映地理位置的三维矢量地图;

在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;

建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;

根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X。

2. 根据权利要求1所述的获得位姿信息的方法,其中,所述建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系,包括:

提取所述矢量地图元素的关键特征,以及识别所述图像元素的关键特征;

将关键特征吻合的所述矢量地图元素与所述图像元素作为相互匹配的矢量地图元素,彼此建立匹配关系。

3. 根据权利要求1所述的获得位姿信息的方法,其中,所述根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X,包括:

获得所述实景图像中的与所述矢量地图元素匹配的图像元素的特征点;

根据所述特征点在所述实景图像中的像素坐标信息、所述特征点在所述三维矢量地图中对应的空间坐标信息、获取所述实景图像的设备参数,获得所述位姿信息的第一观测位姿值;

根据获取实景图像时的先验位姿信息,以及获取先验位姿的设备的位姿噪声,获得位姿信息的第二观测位姿值;

根据所述第一观测位姿值、第二观测位姿值,使用预设的位姿优化方程,推算所述实景图像的位姿信息X。

4. 根据权利要求1所述的获得位姿信息的方法,其中,所述先验位姿信息采用如下方式获得:

采用获得实景图像的视觉惯导里程计定位信息和GPS定位信息融合,获得对应所述实景图像的先验位姿信息。

5. 根据权利要求1所述的获得实景图像位姿信息的方法,其中,所述实景图像为连续的视频帧;

所述根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X,包括:

建立包含W个视频帧的时间窗口;

通过对应当前待处理的实景图像的先验位姿信息提供先验位姿约束,其中,所述当前待处理的实景图像是准备进入滑窗因子的时间窗口的实景图像;

若所述实景图像中存在与三维矢量地图中的矢量地图元素匹配的图像元素,则根据所述矢量地图元素中的空间坐标信息以及所述实景地图中的像素坐标信息,为该视频帧加入

视觉特征重投影约束；

根据视觉惯导约束的相关信息，建立当前待处理实景图像与前一个实景图像之间的视觉惯导约束；

根据上述约束关系，对时间窗口中的各个视频帧采用滑窗因子优化的方式确定所述实景图像位姿信息X。

6. 根据权利要求5所述的获得位置信息的方法，其中，所述根据视觉惯导约束的相关信息，建立当前待处理实景图像与前一个实景图像之间的视觉惯导约束，包括：

根据所述当前待处理实景图像的先验位姿信息和所述前一个实景图像对应的先验位姿信息，建立所述当前待处理实景图像与前一个实景图像之间的视觉惯导约束。

7. 根据权利要求5所述的获得位姿信息的方法，其中，对于所述每帧实景图像，当所述时间窗口中进入第K+1个实景图像时，将超出所述时间窗口的第K-W+1个实景图像进行边缘化处理，并根据所述约束关系，获得所述第K+1个实景图像的位姿信息X。

8. 一种获得位姿信息的装置，其中，包括：

信息获得单元，用于获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息；

地图获得单元，用于获得包含该实景图像所反映地理位置的三维矢量地图；

元素获得单元，用于在所述三维矢量地图中，获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素；

匹配单元，用于建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系；

推算单元，根据所述矢量地图元素的空间坐标信息，以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息，对所述先验位姿信息进行优化，推算获得所述实景图像的位姿信息X。

9. 一种电子设备，其中，包括：

处理器；

存储器，用于存储方法的程序，所述程序在被所述处理器读取执行时，执行以下步骤：获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息；获得包含该实景图像所反映地理位置的三维矢量地图；在所述三维矢量地图中，获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素；建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系；根据所述矢量地图元素的空间坐标信息，以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息，对所述先验位姿信息进行优化，推算获得所述实景图像的位姿信息X。

10. 一种计算机存储介质，其中，所述计算机存储介质存储有计算机程序，所述程序被执行时执行以下步骤：获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息；获得包含该实景图像所反映地理位置的三维矢量地图；在所述三维矢量地图中，获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素；建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系；根据所述矢量地图元素的空间坐标信息，以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息，对所述先验位姿信息进行优化，推算获得所述实景图像的位姿信息X。

一种获得位姿信息的方法、装置以及系统

技术领域

[0001] 本申请涉及图像处理方法,具体涉及一种获得位姿信息的方法。

背景技术

[0002] 当前,电子地图已经成为城市管理、交通出行等领域的基础数据,对交通出行、解决拥堵等问题具有重要的意义。随着技术的发展,电子地图对现实世界的表达越来越丰富和精确,充分反映各种地理细节的高精电子地图已经成为智能驾驶、自动驾驶或者数字城市等领域必不可少的基础数据;随着人们对电子地图依赖程度的加深,对于电子地图的准确性和实时性要求也不断提高。

[0003] 在实际生活中,地理情况尤其是道路情况是动态变化的,如何快速更新电子地图数据,以使其与实际情况相互吻合,成为电子地图提供商的重要课题。

[0004] 在高精地图制图领域,目前一般是采用搭载了高精度的组合惯导和激光雷达等传感器的采集车采集高精地图生产资料;这种技术需要采用专门的传感器,使用专用车辆,沿道路进行生产资料采集。这种方案其缺点在于,由于受到设备成本过高的限制,无法大规模部署,不能满足高精地图生产资料进行大范围、高频次信息采集的需求。

[0005] 为克服上述方案的缺陷,低成本的高精地图生产资料的采集成为一种重要的技术路线。这种技术采用社会车辆搭载被一般公众广泛使用的设备进行高精地图生产资料的采集,这类设备一般集成了消费级的视觉传感器(如相机)、定位模块、惯导模块等,结合这些社会车辆的日常出行进行资料采集,实现了低成本、高频度的资料采集。

[0006] 然而,采用上述方案也存在一些明显的缺陷;一个主要的问题是,由于消费级的传感器精度有限,通过该类传感器采集的数据精度不足,解决上述问题,成为这一技术路线充分发挥作用的关键。

发明内容

[0007] 本申请提供一种获得位姿信息的方法;该方法通过给获得的实景图像提供准确的位姿标识,可以有效提高采用普通设备获得的影像对高精地图更新的数据价值,使高精地图的众包更新成为可能。

[0008] 本申请提供的获得位姿信息的方法,包括:

[0009] 获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;

[0010] 获得包含该实景图像所反映地理位置的三维矢量地图;

[0011] 在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;

[0012] 建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;

[0013] 根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景

图像的位姿信息X。

[0014] 可选的,所述建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系,包括:

[0015] 提取所述矢量地图元素的关键特征,以及识别所述图像元素的关键特征;

[0016] 将关键特征吻合的所述矢量地图元素与所述图像元素作为相互匹配的矢量地图元素,彼此建立匹配关系。

[0017] 可选的,所述根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X,包括:

[0018] 获得所述实景图像中的与所述矢量地图元素匹配的图像元素的特征点;

[0019] 根据所述特征点在所述实景图像中的视觉测量值、所述特征点在所述三维矢量地图中对应的空间坐标信息、获取所述实景图像的设备参数,获得所述获取所述实景图像的设备观测位姿信息;

[0020] 根据所述先验位姿信息和获取所述实景图像的设备位姿噪声,获得所述先验位姿的观测信息;

[0021] 根据所述获取所述实景图像的设备观测位姿信息、所述先验位姿的观测信息以及预设的位姿优化方程,推算所述实景图像的位姿信息X。

[0022] 可选的,所述先验位姿信息采用如下方式获得:

[0023] 采用获得实景图像的视觉惯导里程计定位信息和GPS定位信息融合,获得对应所述实景图像的先验位姿信息。

[0024] 可选的,所述实景图像为连续的视频帧;

[0025] 所述根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X,包括:

[0026] 建立包含W个视频帧的时间窗口;

[0027] 通过对应当前待处理的实景图像的先验位姿信息提供先验位姿约束,其中,所述当前待处理的实景图像是准备进入滑窗因子的时间窗口的实景图像;

[0028] 若所述实景图像中存在与三维矢量地图中的矢量地图元素匹配的图像元素,则根据所述矢量地图元素中的空间坐标信息以及所述实景地图中的像素坐标信息,为该视频帧加入视觉特征重投影约束;

[0029] 根据视觉惯导约束的相关信息,建立当前待处理实景图像与前一个实景图像之间的视觉惯导约束;

[0030] 根据上述约束关系,对时间窗口中的各个视频帧采用滑窗因子优化的方式确定所述实景图像位姿信息X。

[0031] 可选的,所述根据视觉惯导约束的相关信息,建立当前待处理实景图像与前一个实景图像之间的视觉惯导约束,包括:

[0032] 根据所述当前待处理实景图像的先验位姿信息和所述前一个实景图像对应的先验位姿信息,建立所述当前待处理实景图像与前一个实景图像之间的视觉惯导约束。

[0033] 可选的,对于所述每帧实景图像,当所述时间窗口中进入第K+1个实景图像时,将

超出所述时间窗口的第 $K-W+1$ 个实景图像进行边缘化处理,并根据所述约束关系,获得所述第 $K+1$ 个实景图像的位姿信息 X 。

[0034] 本申请还提供一种获得位姿信息的装置,包括:

[0035] 信息获得单元,用于获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;

[0036] 地图获得单元,用于获得包含该实景图像所反映地理位置的三维矢量地图;

[0037] 元素获得单元,用于在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;

[0038] 匹配单元,用于建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;

[0039] 推算单元,根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息 X 。

[0040] 本申请还提供一种电子设备,包括:

[0041] 处理器;

[0042] 存储器,用于存储方法的程序,所述程序在被所述处理器读取执行时,执行以下步骤:获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;获得包含该实景图像所反映地理位置的三维矢量地图;在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息 X 。

[0043] 本申请还提供一种计算机存储介质,所述计算机存储介质存储有计算机程序,所述程序被执行时执行以下步骤:获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;获得包含该实景图像所反映地理位置的三维矢量地图;在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息 X 。

[0044] 与现有技术相比,本申请具有如下优点:

[0045] 本申请提供的获得位姿信息的方法,包括获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;获得包含该实景图像所反映地理位置的三维矢量地图;在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息 X 。

[0046] 本技术方案通过提取实景图像中的图像元素,将这些图像元素与三维矢量地图相

互比对,获得对实景图像的位姿信息的标注。从标注后的实景图像中,获得其中所包含的图像元素的准确位置,有效提高众包更新所搜集信息的精度,从而可以采用这些具有准确位置信息的图像元素,对现有高精地图进行更新。通过本申请提供的上述方法,可以实现通过搜集一般日常出行车辆所采集的图像信息,对高精地图进行可靠的高频次更新。

[0047] 在本申请的进一步优选实施方式中,将上述基本方法与时间窗口分析方法相互结合,实现了对连续视频帧的位姿信息标注;能够为高精地图的众包更新提供更为可靠的数据支持。

附图说明

[0048] 图1是本申请第一实施例提供的获得实景图像位姿信息的系统;

[0049] 图2是本申请场景实施例提供的滑窗因子结构示意图;

[0050] 图3是本申请场景实施例提供的滑窗优化过程示意图;

[0051] 图4是本申请第二实施例提供的获得位姿信息的方法的流程图;

[0052] 图5是本申请第三实施例提供的获得视频帧位姿信息的方法的流程图;

[0053] 图6是采用模块方式表达的使用滑窗优化方式获得视频帧位姿信息的具体技术方案;

[0054] 图7本申请第四实施例提供的获得位姿信息的装置的框图;

[0055] 图8是本申请第五实施例提供的一种电子设备的功能框图。

具体实施方式

[0056] 在下面的描述中阐述了很多具体细节以便于充分理解本申请。但是本申请能够以很多不同于在此描述的其它方式来实施,本领域技术人员可以在不违背本申请内涵的情况下做类似推广,因此本申请不受下面公开的具体实施的限制。

[0057] 本申请实施例提供一种获得位姿信息的方法,以提高信息采集设备获得的影像对高精地图更新的数据价值。为了更清楚的展示本申请实施例提供的获得位姿信息的方法,首先对本申请实施例提供的技术方案的应用场景进行介绍。

[0058] 本申请的典型应用场景是,采用移动交通设备(包括无人机)中搭载的视觉惯导里程计(Visual Inertial Odometry,英文缩写为VIO)对已经获得过高精地图的道路交通区域进行数据更新;本申请提供的技术方案解决在获得道路交通区域的实景图像,尤其是连续的视频帧后,如何对每一个实景图像(对连续的视频帧而言,就是每个视频帧)标注位姿信息。所谓实景图像的位姿信息,实际上就是获得该实景图像的视觉设备在获取时刻的位置和角度信息。本申请不涉及在获得每个实际图像的位姿信息后如何利用这些信息对高精地图进行更新的问题。所述移动交通设备,典型的,可以采用一般社会车辆;也不排除使用无人机等设备。

[0059] 除了上述典型应用场景,本申请的技术方案也可以用于对机器人的室内导航等其他场景。

[0060] 所述VIO是融合视觉SLAM技术(Simultaneous Localization and Mapping,中文称作“即时定位与地图创建”)和IMU技术(Inertial measurement unit,惯性测量单元)后获得相得益彰效果的技术方案。

[0061] 所述SLAM技术的英文全称是Simultaneous Localization and Mapping,中文即“即时定位与地图创建”,所谓视觉SLAM技术就是用摄像头(通常为单目摄像头)来完成环境感知,实现即时定位与地图创建;由于单目相机只能得到特征的相对尺度,不能得到绝对尺度,因此难以直接用于导航。

[0062] 所述IMU,英文全称为Inertial measurement unit,中文称作“惯性测量单元”;是测量物体三轴姿态角(或角速率)以及加速度的装置;该技术可以用于弥补视觉SLAM的不足。将上述视觉SLAM技术和IMU技术相互融合,能够显著提高基于单目摄像头的SLAM技术算法的性能,是一种低成本、高性能的导航方案。

[0063] 在使用上述VIO设备时,还可以同时获得GPS定位信息,所述GPS定位信息利用GPS信号采用实时动态载波相位差分技术(Real-time kinematic,简称RTK)实现定位。

[0064] 在获取实景图像时,利用上述两种定位技术相互融合获得的对应所述实景图像的位姿信息可以作为先验位姿信息,可以记为RTK-VIO先验位姿信息。所谓位姿信息,是由一组参数表征的表示获得该实景图像时摄像设备所处的空间位置(位置信息)和摄像角度(姿态信息)的信息;一个典型的位姿信息包括六个参数(x,y,z,ox,oy,oz),即分别表示在空间三个方向上的坐标位置,以及与各个坐标轴的夹角;当然,根据不同的情况,可以采用不同的位姿信息表达参数。

[0065] 以下结合上述典型应用场景,介绍本申请的具体实施例。

[0066] 请参考图1,其为本申请第一实施例提供的获得位姿信息的系统;以下结合该示意图,对本申请第一实施例提供的获得位姿信息的系统进行介绍;该系统也是后续获得位姿信息的方法的运行环境。

[0067] 所述系统包括:视觉惯导里程计101(图像采集器)和计算机系统102(服务器端)。

[0068] 所述视觉惯导里程计101为一种图像采集器,主要包括惯性测量元件以及与所述惯性测量元件耦接的相机,也包含用于获得GPS信号的GPS元件。在车辆行驶的过程中,所述视觉惯导里程计101通过所述相机拍摄到图片和/或视频,并计算相机在拍摄所述图片和/或视频时的位姿。在本申请提供的场景实施例中,将所述视觉惯导里程计101采集到的反映真实地理空间的图像和/或视频中的视频帧的集合称为实景图像。

[0069] 所述视觉惯导里程计101通过相机拍摄到所述实景图像后,将所述实景图像发送至计算机系统102中,所述计算机系统对所述实景图像进行分析,获得所述相机拍摄所述实景图像时的精确的位姿信息,从而为利用这些实景图像进行高精地图更新提供基础信息。

[0070] 尽管本实施例中采用视觉惯导里程计作为图像采集器,但不排除使用其它可以在摄取图像时获得位置信息的设备作为图像采集器。

[0071] 所述视觉惯导里程计101在获取实景图像时,记录所述实景图像获取位置的位置参数。

[0072] 所述实景图像获取位置的位置参数,是所述实景图像成像时摄像设备所处地理位置的位置参数,在具体应用过程中,所述实景图像获取位置的位置参数包括以下位置参数中的至少一种:

[0073] 1、获取该所述实景图像时的GPS定位信息;

[0074] 2、获取所述实景图像时的惯性导航定位信息。

[0075] 上述信息通过所述摄像设备中设置的惯性导航设备以及GPS设备获得;所述GPS定

位信息也可以从搭载所述摄像设备的车辆获得;这些信息和实景图像具有对应关系。

[0076] 上述GPS定位信息和惯性导航定位信息相互融合获得的先验位姿信息可以记为RTK-VIO先验位姿信息;该信息仅仅反映所述实景图像的初步位姿信息,很可能存在很大误差,因此需要进行优化。具体获得RTK-VIO先验位姿信息方法在本技术领域已经有多种方案;例如,通过非线性优化的方法把视觉信息,惯导信息和GNSS信息(即全球导航卫星系统信息,包括RTK)融合在一起,得到照片(或相机)的初始轨迹;具体实现方案并非本发明重点,在此不再赘述。

[0077] 所述计算机系统102,包括:

[0078] 信息获取单元1021,用于获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息。其中,所述实景图像就是视觉惯导里程计101发送的实景图像。获取实景图像时,包括获取所述实景图像时的先验位姿信息。所述图像元素是经过对实景图像进行图像识别后获得的道路交通中存在的典型元素;例如,从实景图像中识别出的车道线、路牌等各种交通标识,也不排除一些典型的道路交通建筑物,例如道路上的桥梁;也不排除路边的典型房屋。

[0079] 地图获得单元1022,用于获得包含该实景图像所反映地理位置的三维矢量地图。所述三维矢量地图是指,根据相应的规范和标准对地图上的各种内容进行编码和属性的定义,具备地图要素的类别、等级和特征的地图。这样的地图在内容上更具具备动态性,地图的内容和表示效果可以实时修改;在结构上,有特定的组织形式和数据结构,有关于地图中图像元素对应的编码、属性、位置、名称及其相互之间的拓扑关系等方面的内容;在组织上,可以分层、分类、分级,能够快速的对地图中的图像元素进行检索和查询。所述三维矢量地图一般是预先已经建立好的,当然,其中包含的内容和实际情况之间的关系可能由于道路交通实际情况的变化而与实际情况已经不再相符;本申请技术方案希望解决的就是快速反映最新的情况。

[0080] 在本申请提供的场景实施例中,所述三维矢量地图是包含所述相机拍摄的实景图像所反映的具体的地理位置的三维矢量地图;所述实景图像在获取过程中可以根据获取时的GPS信号以及惯性导航信息得到该实景图像的先验位姿信息;根据该先验位姿信息确定的位置,就可以在三维矢量地图的图库中选择包含该实景图像所反映地理位置的部分。

[0081] 在所述地图获得单元1022,所述获得包含该实景图像所反映地理位置的三维矢量地图,包括:根据所述实景图像的GPS定位信息或惯性导航定位信息中的一种,或者GPS定位信息和惯性导航定位信息的融合信息,获得该实景图像所反映的地理位置,并根据该地理位置,获得相关的所述三维矢量地图。

[0082] 元素获得单元1023,用于在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素。

[0083] 所述与所述图像元素相关的矢量地图元素,可以包括:车道线、路牌等。选择这些图像元素的原因是其中包含的信息有很强的独特性,便于唯一确认,以及具有稀疏性,在同一个图像中一般不可能出现很多,从而便于和三维矢量地图中的矢量地图元素准确匹配。具体识别的方法,在本申请第二实施例中予以介绍。

[0084] 匹配单元1024,用于建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系。

[0085] 具体的,所述建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系,包括:提取所述矢量地图元素的关键特征,以及识别所述图像元素的关键特征;将所述关键特征吻合的所述矢量地图元素与所述图像元素作为相互匹配的矢量地图元素,彼此建立匹配关系。

[0086] 其中,所述矢量地图的关键特征,可以理解为三维矢量地图的中各个矢量地图元素对应的语义特征、图像特征等;对应的,所述图像元素的关键特征可以理解为,所述实景图像中的各个图像元素的类别特征(例如:该图像元素为车道线或者是道路标识牌)、语义特征(例如:车道线的具体含义为左转、直行等,或者是道路标识牌中标识的文字具体体现的是XX村庄等)、图形特征(例如长方形、三角形、圆形)等。

[0087] 另外,所述将所述关键特征吻合的所述矢量地图元素与所述图像元素作为相互匹配的矢量地图元素,彼此建立匹配关系可以通过以下方式实现。

[0088] 首先,根据所述矢量地图元素的空间坐标信息以及所述实景地图的获取位置的预设范围,确定所述预设范围内的所述矢量地图元素;比如说,所述预设范围为10米,可以根据所述实景图像的先验位姿信息预估所述实景图像在所述三维矢量地图中的大致位置(a, b, c),进而确定以三维矢量地图的位置(a, b, c)为中心,半径为10米的范围内的矢量地图元素。

[0089] 进一步的,基于前述获得的所述预设范围内的各个所述矢量地图元素的关键特征,以及所述实景图像中的各个图像元素的关键特征,建立所述各个矢量地图元素以及各个图像元素之间的匹配关系。例如:假设所述实景图像中包含的需要确定位置的图像元素为路牌,路牌中指示的内容为街道名称,则将所述矢量地图与所述实景地图比对后,就可以根据矢量地图中的类型为路牌的矢量元素,查找预定范围的路牌,根据文字信息,找到与实景图像中的路牌文字信息吻合的矢量地图中的路牌。如果在预定范围内有多个文字内容匹配的图像元素,则选取距离最近的图像元素;当然,也可以不考虑预定范围,直接选取距离最近的图像元素;或者直接选取距离最近的图像元素,再排除超过预先选定的距离范围阈值的图像元素,即如果最近的图像元素也超过了所述距离范围阈值,则判定没有找到匹配的图像元素。上述几种选择方案在结果上存在理论上的差别,但实际应用中,由于选择的图像元素均为具有一定稀疏性的图像元素,如本例子中的路牌这种图像元素,在反映道路交通的实景图像中一般不会以相同的内容重复出现,所以可以准确的实现一对一的精确匹配。

[0090] 推算单元1025,用于根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X

[0091] 实景地图中的图像元素与三维矢量地图中的矢量地图元素的匹配结果,就可以进一步得到所述实景图像中的与所述三维矢量地图中的图像元素相对应的特征点。

[0092] 所述实景图像中的与所述三维矢量地图中的矢量地图元素相对应的特征点,是指矢量地图中某个便于识别的点的位置信息。例如,矢量地图元素的类型为路牌,如果确定了实景图像中的路牌A和三维矢量地图中的路牌A' 相互匹配,可以进一步确定实景图像中路牌A的左上角角点和三维矢量地图中的路牌A' 的左上角角点为相互对应的特征点。由于特征点具有确定的坐标,便于进行位姿关系的推算。在本申请中,上述实景图像和三维矢量地

图中的相互对应的特征点称为匹配对。

[0093] 采用所述匹配对在所述实景图像中的像素坐标信息,以及所述特征点在所述三维矢量地图中的矢量地图元素的空间坐标信息可以对所述先验位姿信息进行优化。所述地图要素主要包括车道线,交通标识,杆状物等;每个地图要素都会有空间坐标信息;这里不具体到具体的要素,主要是强调相应要素(特征)的坐标信息。对于空间中具有确定空间坐标位置的某个特征点,其在实景图像中的位置,取决于获取该图像时摄像设备所处的位姿,因此,在分别获得了相互对应的实景图像和三维地图坐标中的同一个特征点的坐标之后,就可以推算出获取该实景图像时摄像设备的位姿信息,由于该位姿信息和实景图像对应,也称为该实景图像的位姿信息。上述方式获得的能够优化先验位姿信息的信息称为视觉特征重投影约束信息。

[0094] 上述推算过程中,没有考虑以视频方式获得的视频帧作为实景图像时,各个视频帧之间的关系,实际上,一段视频的各个视频帧之间必然存在一定的联系,因此,在基于上述获得实景图像的位姿信息的方法获得各个视频帧的单独的位姿信息之后,还可以根据各个视频帧之间的联系对各视频帧的位姿信息进行优化以获得准确的各个视频帧(即,各个连续的实景图像)的位姿信息。

[0095] 具体的,可以结合已经获得的相邻的实景图像的先验位姿信息(即,先验位姿约束)、三维矢量地图中的矢量地图元素与实景地图中的图像元素之间的位置约束关系(即,视觉特征重投影约束)以及当前待处理实景图像与前一个实景图像之间的视觉惯导里程计的惯导数据(视觉惯导约束),对的位姿信息进行优化。

[0096] 在实际应用中,基于上述方式推算获得的连续的实景图像的位姿信息实际上是若干个的候选位姿信息,在推算获得所述位姿信息后之后,还需要排除明显不符合先验知识的结果。该步骤的主要目的是避免明显不合理的推算结果。

[0097] 所述先验知识是指关于搭载所述视觉惯导里程计的移动设备的先验知识,例如,假设所述移动设备为社会车辆,则所述社会车辆的先验知识为涉及车辆的横滚角、俯仰角以及航向角与路面之间夹角的相关内容,如果所述横滚角、俯仰角以及航向角与路面之间夹角较小,超过了预设的阈值,则认为推算结果不符合先验知识,该推算结果应当排除。由于同一个实景图像可能存在多个可以与三维矢量地图中的矢量元素进行匹配的图像元素,因此,在采用视觉特征重投影的方法时,可以用多个方式推导实景图像的位姿信息 X ,如果某个推导结果被排除,则可以采用其它匹配关系的推导结果。

[0098] 具体的,对于连续的视频帧,所述推算单元1025还包括:滑窗优化子单元。

[0099] 滑窗优化(Sliding Window Optimization),是利用滑窗通过边缘化的手段去除超过一定时间的变量的过程,在涉及视觉惯性里程计的问题上,状态估计问题被建模为最大后验证概率问题,通常在假设满足高斯分布的情况下,代价函数的优化为一个最小二乘问题,可通过泰勒展开转化为线性系统迭代求解。在测算视觉惯性里程计获得视频帧对应的位姿的问题上,随着视觉惯性里程计的运行,采集的视频帧逐渐变多,状态变量规模不断扩大,需要利用滑窗通过边缘化的手段去除一定时间的变量,限制计算量以保障计算效率和运行效率。

[0100] 所述滑窗优化子模块,用于建立包含 W 个滑窗因子的时间窗口,每个滑窗因子对应连续视频中的一个视频帧;每个滑窗因子包含先验位姿约束、视觉特征重投影约束、以及反

映连续视频帧之间相互约束关系的视觉惯导约束(简称为VIO约束),根据上述约束条件,推算确定所述时间窗口中每个视频帧的位姿信息X。

[0101] 所述先验位姿约束,是指根据获得每个视频帧时的GPS定位信息和惯性导航定位信息,初步估算出的与各个视频帧对应的位姿信息,即前述RTK-VIO先验位姿信息。该估算结果比较粗略,只能作为该视频帧位姿信息的初步结果,并在后续的视觉特征重投影约束中起作用。

[0102] 所述视觉特征重投影约束,即前述说明的利用三维矢量地图的特定种类矢量元素投影到视频帧后获得投影结果,找到匹配对,并根据匹配对对视频帧的位姿信息进行估算;具体方式参照本发明各个实施例中对于获得视觉特征重投影约束信息的说明。

[0103] 所述VIO约束,即根据视频帧之间的时间间隔,根据视觉惯导里程计的数据,建立的视频帧相互之间的位姿约束关系;在本申请的具体实施方式中,是指滑窗优化时间窗口内的各个视频帧之间的位姿约束关系。

[0104] 如图2所示,其为本申请场景实施例提供的初始化滑窗因子结构示意图,所述滑窗因子包括:先验位姿约束101B、视觉特征重投影约束103B,通过上述两重约束,获得初始位姿信息102B。该初始化滑窗因子对应第一个视频帧,还没有考虑其他视频帧产生的VIO约束。

[0105] 所述先验位姿约束101B,是指获取视频帧时各个实景图像对应的先验位姿信息之间的约束关系,即,约束了各个连续的实景图像之间位姿信息的变化范围。

[0106] 所述视觉特征重投影约束103B,是指根据所述矢量地图元素在所述三维矢量地图中的位置信息,以及与所述矢量地图元素相匹配的所述图像元素在所述实景图像中的像素信息推算获得的所述视频帧的位姿信息,此处,可以将根据上述推算过程获得的视频帧的位姿信息看做一个约束条件,即,基于滑窗优化的方式获得的位姿信息X与上述推算获得的位姿信息之间的差值应小于预设的变化范围,上述推算获得的位姿信息的获得方式见前述对视觉特征重投影约束信息获得过程的描述。

[0107] 所述视觉惯导里程计约束102B,所述视觉惯导里程计约束是指当前待处理实景图像与前一个实景图像之间的视觉惯导约束。

[0108] 如图3所示,其为本申请场景实施例提供的滑窗优化过程示意图。该图示出了当时时间窗口设置为包括W个视频帧的位姿变量,当前视频帧由第K个变为第K+1个时,对所述视频帧进行滑窗优化的过程。

[0109] 可以理解的,相邻的若干个视频帧存在一定的约束关系,同样的,相邻的视频帧位姿信息也存在一定的约束关系,且距离越近的视频帧这种约束关系越大,在本申请实施例中,认为W个相邻的视频帧之间的存在这种关系。

[0110] 在利用滑窗优化的方法处理视频帧的位姿信息X时,首先基于前文计算实景图像的位姿信息的方法获得第1个图像帧的位姿信息,之后利用所述第1个视频帧对下一个视频帧的惯性导航约束,结合所述下一个视频帧的先验位姿约束,以及所述下一个视频帧的视觉特征重投影约束,获得下一个视频帧的位姿信息。在该下一个视频帧的位姿信息进入所述滑窗优化的时间窗口后,再将其位姿信息与处于时间窗口中的所有视频帧的位姿信息一起进行滑窗优化,在该视频帧离开时间窗口时获得的最终位姿信息就是最终确定的该视频帧的位姿信息确定值。

[0111] 对于滑窗优化中的滑窗设置为W个视频帧时,假设当前帧为第K帧,则滑窗中的第一个帧为(K-W+1)帧,作为当前帧的第K帧需要考虑在其前面的各个帧直到第(K-W+1)帧对其产生的惯性导航约束;当前帧变为第(K+1)帧时,则第(K-W+1)帧被边缘化,即不再考虑第(K-W+1)帧对第(K+1)帧产生的惯性导航约束。

[0112] 与上述场景实施例相对应的,本申请第二实施例提供一种获得实景图像位姿信息的方法,请参考图4,其为本申请第二实施例提供的获得位姿信息的方法的流程图,以下对该方法的实现步骤和依据进行详细说明。由于该方法实施例相似与上述场景实施例,以下仅对上述场景实施例中未提及的部分进行详细说明,其他相关之处参见上述场景实施例的部分说明即可。

[0113] 本方法的实施主体一般为服务端,所述服务端包括但不限于服务器、服务器集群。

[0114] 本申请第二实施例提供的获得实景图像位姿信息的方法,包括如下步骤:

[0115] 步骤S201,获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息。

[0116] 所述实景图像,是通过视觉惯导里程计等图像采集设备采集到的反映真实地理空间的图像和/或视频中的视频帧的集合。在获得实景图像的同时,能够同时取得对应该实景图像的先验位姿信息。

[0117] 其中,所述实景图像获取时的先验位姿信息包括以下位置参数中的至少一种:

[0118] 获取该所述实景图像时的GPS定位信息;视觉惯导里程计定位信息。

[0119] 在具体实施中,所述先验位姿信息通常采用上述两种定位信息融合获得;具体的融合方法在本领域已经有多种方案;通过融合获得的先验位姿信息记为RTK-VIO。

[0120] 步骤S202,获得包含该实景图像所反映地理位置的三维矢量地图。

[0121] 所述三维矢量地图指的是该实景图像所反映地理位置的三维矢量地图,包括:根据所述实景图像的GPS定位信息或视觉惯导里程计定位信息中的一种,或者两者的融合信息,获得该实景图像所反映的地理位置,并根据该地理位置,获得相关的所述三维矢量地图;该相关的三维矢量地图,一般是需要更新的高精三维矢量地图中对应所述实景图像所在位置的某个区域范围。

[0122] 步骤S203,在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素。

[0123] 所述实景图像中的图像元素与三维矢量地图中的反映同一个实物的矢量元素对应后,能够更好的反映拍摄所述实景图像时相机的位姿信息,比如说,所述实景图像中包括某一路牌,则可以根据所述路牌在图像中的位置和成像大小,结合三维矢量地图中该路牌的位置和高度信息,得到相机拍摄所述实景图像时的拍摄角度和拍摄距离,根据所述拍摄角度和拍摄距离可以进一步得到相机在拍摄实景图像的这一时刻所处的位置和方向,即获得所述相机拍摄的实景图像的先验位姿信息;从而优化所述实景图像的先验位姿信息。

[0124] 其中,所述实景图像获取位置是指,所述实景图像的位姿信息反映至所述实景图像中后的位置,即,根据待优化的先验位姿信息,获得所述实景图像获取位置对应的三维空间位置,进而根据该三维空间位置划定三维矢量地图的预设范围,并根据三维矢量地图中标注的内容,获得上述矢量地图元素。

[0125] 在本步骤中,为了获得所述矢量地图元素,首先需要先获取所述实景图像的图像

元素,以便用于后续的匹配步骤,例如:从实景图像中提取路牌或者车道线等易于识别并且足够稀疏的图像元素。

[0126] 所述实景图像中的的图像元素可以选择具有较强识别特征的图像元素,一般为:车道线和/或路牌,这些元素一般配合有相关的文字,而文字内容一般具有明显的特征,不会混淆;在实际应用的过程中,从所述实景图像中提取所述图像元素的方法有多种。

[0127] 当所述图像元素为路牌时,所述提取过程包括如下步骤:

[0128] 根据所述实景图像,提取路牌轮廓;

[0129] 对路牌轮廓中的像素点进行直线拟合;

[0130] 对经过直线拟合后的像素点,提取具有角点特征的像素点作为路牌角点。

[0131] 所述路牌角点是指路牌中的极值点,即在某一方面特别突出的点。比如说:路牌四角上的点,所述路牌角点也可以理解为路牌中具有一定特征的点。

[0132] 当所述图像元素为车道线时,所述提取过程包括:

[0133] 对所述实景图像进行逆透视变化;可以理解的拍摄得到的实景图像是二维图像,在拍摄过程中,必然存在将拍摄实体的三维点云数据转换为二维图像的转换关系,所述逆透视变化就是指根据所述转换关系对所述二维图像进行反转换,得到所述实景图像对应的三维点云数据,所述三维点云数据是指拍摄实体的所有采样点的数据集合,其中,每个采样点对应一个点云数据。一般情况下,所述点云数据中包括实体的各个采样点的对应的坐标以及所述采样点的表面特征,例如:反射率。

[0134] 对逆透视变化后的反射率进行聚类;所述聚类的过程就是指根据点云数据对拍摄实景图像时对应的实体进行还原的过程。所述聚类的方式有多种,例如:计算点云数据之间的欧氏距离,以完成聚类。

[0135] 将所述实景图像中的拍摄实体还原后,对聚类后的三维端点进行提取;由于是经过逆透视变化后,最终获得了三维图像,因此本步骤是对三维端点进行提取。

[0136] 根据所述三维端点,进一步提取二维像素端点。二维像素端点,即在所述实景图像中的像素点的坐标。

[0137] 步骤S204,建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系。

[0138] 对该步骤可以参见前面对匹配单元1024的说明;通过本步骤,将获得的矢量地图元素根据其包含的特征参数,例如形状、位置、尺寸等,与实景图像中的图像元素匹配;由于所选择的矢量地图元素和实景图像的图像元素均具稀疏性,并且矢量地图和实景图像预先已经做了空间上的大致对应,因此,能够比较方便的确定矢量地图元素与实景图像中的图像元素的对应关系;上述对应关系的确定过程,实际上可以想象是将所述矢量地图元素分别投影到所述实景图像中,与相匹配的图像元素相互贴合。相互匹配的矢量地图元素与图像元素显然应当具备共同的空间位置。

[0139] 步骤S205,根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X。

[0140] 具体的,所述推算获得所述实景图像的位姿信息X的过程具体包括:

[0141] 获得在所述实景图像中的与所述矢量元素匹配的图像元素的相互对应的特征点;

[0142] 采用所述特征点在所述实景图像中的像素坐标信息,以及所述特征点在所述三维矢量地图中的矢量地图元素的空间坐标信息,对所述先验位姿信息进行优化,获得所述实景图像的位姿信息 x 。具体推算过程说明如下。

[0143] 首先,根据上述匹配对,可以形成对相机位姿的观测;所谓观测,是指一个能够跟要估计的状态(此处为相机位姿 X)建立函数关系的测量值,通常状态估计的过程是要估计的状态不能直接得到,需要一系列的观测值进行求解。

[0144] 通过上述特征点的对应关系,就可以形成对相机位姿的第一观测位姿值:

$$[0145] \quad \text{pixel}_{\text{measure}} = \text{projection}(x, \text{landmark}_{\text{map}}) + \text{noise}_{\text{pixel}}$$

[0146] 其中, $\text{pixel}_{\text{measure}}$ 是所述特征点在所述实景图像中的像素坐标信息, $\text{landmark}_{\text{map}}$ 是特征点在三维矢量地图中的三维空间坐标信息, X 是位姿信息, $\text{noise}_{\text{pixel}}$ 是指检测出的实景图像中的视觉噪声特征,该值为经验值或者由特征提取算法直接给出,可以视为获取实景图像的设备的设备内部参数; Projection 是实现 $\text{landmark}_{\text{map}}$ 结合位姿信息 X 后,转换为二维实景地图的视觉测量值(即像素坐标信息) $\text{pixel}_{\text{measure}}$ 的转换函数,这个函数表达了矢量地图元素在三维矢量地图的特征点的空间坐标信息(即 $\text{landmark}_{\text{map}}$)已知的情况下,将三维矢量地图中的特征点投影到图像像素平面,获得空间坐标信息对应的像素坐标信息的过程;该 Projection 函数的具体形式取决于获取实景图像的设备,可以视为获取实景图像的设备的参数。。

[0147] 通过上述公式,可以获得所述实景图像的第一观测位姿值。

[0148] 此外,还可以根据获取实景图像时的GPS定位信息和/或视觉惯导里程计定位信息获得对位姿信息 X 的第二观测位姿值:

$$[0149] \quad X_{\text{vio}} = X + \text{noise}_{\text{vio}}$$

[0150] 其中, $\text{noise}_{\text{vio}}$ 是指视觉惯导里程计的位姿噪声,或者叫视觉惯导里程计位姿的不确定因素, noise 反映观测值的精度,是视觉惯导里程计的参数。 x_{vio} 是先验位姿信息(可以是视觉惯导里程计先验位姿信息,也可以是RTK-VIO融合的先验位姿信息);这是所述实景图像的第二观测位姿值。

[0151] 最后,利用上述第一观测位姿值和第二观测位姿值,使用位姿优化方程,就可以获得准确的位姿信息 X :

$$[0152] \quad \text{error}_{\text{reprojection}} = \text{projection}(x, \text{landmark}_{\text{map}}) - \text{pixel}_{\text{measure}}$$

$$[0153] \quad \text{error}_{\text{vio}} = x_{\text{vio}} - X$$

$$[0154] \quad X = \underset{x \in \text{se}(3)}{\text{argmin}} \{ \text{error}_{\text{reprojection}}^T \Omega_{\text{pixel}} \text{error}_{\text{reprojection}} + \text{error}_{\text{vio}}^T \Omega_{\text{vio}} \text{error}_{\text{vio}} \}$$

[0155] 其中, Ω 为测量噪声对应的信息矩阵。 error 是指预测值跟观测值的差别。 x 是属于 $\text{se}(3)$ 的群元素; $\text{se}(3)$ 是李代数的一种,李代数属于SLAM常用的数学工具。 argmin 表示使后面这式子达到最小值时 X 的取值。

[0156] 基于上述推算获得所述实景图像的位姿信息之后,排除明显不符合先验知识的结果,以完成对所述位姿信息的筛选,最终获得准确的位姿信息 X 。该位姿信息 X 是获得具体的实景图像时摄像设备的位姿信息,自然的,也就是对应该实景图像的位姿信息。

[0157] 综上所述所述,本申请提供的获得实景图像位姿信息的方法,通过提取实景图像

中具有特殊特征而便于识别的图像元素,将这些图像元素与已经获得高精地图相互比对,获得对实景图像的位姿信息的标注。标注后的实景图像中,就可以获得其中所包含的图像元素的准确位置,有效提高众包更新所搜集信息的精度,从而可以对现有高精地图的众包更新。通过本申请提供的上述方法,可以实现搜集一般日常出行车辆所采集的图像信息,对高精地图进行可靠的高频次更新。

[0158] 本申请第三实施例仍提供一种实景图像位姿信息的方法。由于该方法实施例基本相似与上述场景实施例和第一实施例,所述以下仅对上述场景实施例中未提及的部分进行详细说明,其他相关之处参见上述场景实施例的部分说明即可。

[0159] 与上述第二实施例不同的,本申请第三实施例中的获得实景图像为连续的视频帧。可以理解的,此时需要获取的实景图像的位姿信息为对应每个视频帧的位姿信息。

[0160] 本实施的获得实景图像位姿信息的方法首先包括对最初的视频帧执行前述步骤S201至步骤S205,并将通过上述步骤获得的初始帧的位姿信息作为初始帧初始值。可以参见图2的初始化滑窗因子。在初始值的基础上,对于后续视频帧采取如下三方面的信息相互约束,获得后续视频帧的优化后的位姿信息:

[0161] A、先验位姿约束,即通过视觉惯导里程计获得视频图像时记录的GPS信息和惯性导航信息,获得的RTK-VIO先验位姿约束;

[0162] B、视觉特征重投影约束,即采用如步骤S201至步骤S205的方法获得的视频帧中的图像元素与三维矢量地图中的矢量元素之间的匹配对,获得的约束关系;

[0163] C、视觉惯导约束,即根据视频帧之间的时间间隔,以及根据视频惯导里程计中记录的惯性导航信息,获得滑窗范围内的前面视频帧对后面的当前视频帧(即当前需要计算位姿信息的视频帧)的约束关系。

[0164] 采用上面的三方面约束共同优化计算视频帧的位姿信息时,需要采用前面提到的滑窗优化的方法,将距离超过滑窗范围的视频帧的影响边缘化,以减少计算量。滑窗因子的示意图可以参见图3。

[0165] 请参考图5,其为本申请第三实施例所述的获得视频帧位姿信息的流程图。该实施例提供了实现图3所述滑窗优化方式的一种具体方案。

[0166] 其中,在对初始帧使用上述步骤S201到所述步骤S205获得初始帧初始位姿信息之后,采用如下方法实现后续视频帧的进入滑窗优化过程的步骤,其流程图如图5所示,请同时参考图3的滑窗因子示意图。

[0167] 步骤S301,建立包含W个视频帧的时间窗口。

[0168] 与本申请第二实施例相同的,所述通过对应当前待处理视频帧的位置参数提供先验位姿约束也是指将视觉惯导里程计提供的位置参数作为先验位姿约束条件;在本实施例中,具体是使用前面提到的RTK-VIO先验位姿信息。所述当前待处理视频帧是准备进入滑窗因子的时间窗口的视频帧,即图3中(K+1)视频帧。

[0169] 由于本申请第三实施例涉及的是对视频帧的处理,而连续视频帧之间的位姿关系是很容易通过视觉惯性导航里程计进行推算的,而在本实施例中,默认在获取视频的过程中,针对每个视频帧均有对应的惯性导航数据,相邻视频帧的时间间隔也是固定的,因此,在处理确定视频帧对应的位姿信息的问题时,还需要加入相邻视频帧之间的视觉惯导约束条件以进一步提高对视频帧的位姿标定的准确性。

[0170] 为了便于处理相邻视频帧的视觉轨道约束问题,在记录每个视频帧产生时的位置信息时,记录时间戳,以便确定视频帧之间的时间间隔,进而可以根据时间间隔推算视觉惯导约束的结果。

[0171] 步骤S302,若所述实景图像中存在与三维矢量地图中的矢量地图元素匹配的图像元素,则根据所述矢量地图元素中的空间坐标信息以及所述实景地图中的像素坐标信息,为该视频帧加入视觉特征重投影约束。

[0172] 所述视觉特征重投影约束,即与第二实施例类似的过程,在所述视频帧中(相当于实景图像)中识别预定类型的图像元素,例如车道线、路牌;在所述三维矢量地图中,获得所述视频帧获取位置的预定范围内的与所述预定类型的图像元素相关的矢量元素,并将其投影到所述实景图像;根据投影情况,在所述实景图像中获得与所述矢量元素匹配的图像元素形成,并在其中选择特殊点形成匹配对,所述获得与所述矢量元素匹配的图像元素包括以下匹配条件:图像特征相匹配,并且距离矢量元素投影到所述实景地图后的投影位置在预定范围内或者距离最近;获得匹配对后,采用所述特征点匹配对在所述实景图像中的像素坐标信息,以及所述特征点在所述三维矢量地图中的矢量地图元素的空间坐标信息,对所述先验位姿信息进行优化。

[0173] 具体的,所述视觉特征重投影的计算过程包括:

$$[0174] \quad \text{pixel}_{\text{prediction}} = K \cdot T_W^C \cdot \text{Landmark}$$

$$[0175] \quad P_{\text{pixel}} = \frac{\partial \text{Pixel}}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \hat{P} \left(\frac{\partial \text{Pixel}}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \right)^T + \frac{\partial \text{Pixel}}{\partial \text{Landmark}} P_{\text{Landmark}} \left(\frac{\partial \text{Pixel}}{\partial \text{Landmark}} \right)^T$$

[0176] 其中, $\text{pixel}_{\text{prediction}}$ 表示预测的三维矢量地图中的矢量地图元素由空间坐标系转换至二维实景图像坐标系后的坐标位置, T_W^C 表示由三维矢量地图转换至的位姿矩阵, Landmark 表示三维矢量地图提供的矢量地图元素的空间坐标系的坐标位置;

[0177] K 为相机内参, P_{Landmark} 是以协方差标识的地图3D点坐标精度, P_{pixel} 是地图3D点投影得到的像素坐标的协方差矩阵;该协方差矩阵在地图匹配以及优化的匹配代价 Cost 中起到作用。 \hat{P} 表示位姿信息 x 对应的协方差矩阵。

[0178] $\left(\frac{\partial \text{Pixel}}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \right)$ 表示观测值 Pixel 对变量 x 的偏导。计算的时候是通过链式法则,先对 T_C^W 求偏导数,再使用 T_C^W 对 x 求偏导数。

[0179] $\frac{\partial \text{Pixel}}{\partial \text{Landmark}}$ 表示观测值对 landmark 坐标变量的偏导数,上角标 T 表示转置, ∂ 表示偏导。

[0180] 其中,在进行矢量元素与图像元素的匹配时,可依据马氏距离进行距离计算,并选取距离最近的对象作为匹配结果。具体计算公式如下:

$$[0181] \quad D_{ML} = \frac{\partial \text{Pixel}}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \hat{P} \left(\frac{\partial \text{Pixel}}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \right)^T + \frac{\partial \text{Pixel}}{\partial \text{Landmark}} P_{\text{Landmark}} \left(\frac{\partial \text{Pixel}}{\partial \text{Landmark}} \right)^T$$

[0182] 通过当前状态 x 的协方差矩阵 \hat{P} 和 landmark 坐标的协方差矩阵 P_{landmark} , 以及 pixel

观测对 x 和 $landmark$ 坐标的偏导数计算得来;马氏距离是概率/统计学里的一个常用概念,在此不予详述。

[0183] 应当说明,并非每一个视频帧都可以找到合适的匹配对,从而实现视觉特征重投影约束;对于没有找到合适匹配对的视频帧,则不考虑该约束。

[0184] 步骤S303,根据视觉惯导约束的相关信息,建立当前待处理实景图像与前一个实景图像之间的视觉惯导约束。

[0185] 上述各个步骤是获得各视频帧的位姿信息的所需准备的约束条件;其中,本步骤S303用于确定惯性导航参数约束;上述各个步骤获取约束条件的步骤采用不同的顺序。具体到本步骤相关的获得视觉惯导约束的过程说明如下。

[0186] 根据上述约束条件获得所述每个视频帧的位姿信息的过程中,需要先获得视频的第一个时间窗口所对应的第一个视频帧的初始位姿信息作为初始条件,也就是获得图2所示的初始化滑窗因子,该初始化滑窗因子只考虑了先验位姿约束以及视觉特征重投影约束;从第二个视频帧开始,准备进入所述滑窗因子分析的时间窗口的视频帧——本实施例中所述的当前待处理视频帧,即图3的第 $(K+1)$ 视频帧——则除了使用对应该视频帧的先验位姿约束、视觉特征重投影约束之外,还需要根据与前一个视频帧——即图3的第 K 视频帧——之间的间隔时间关系,根据所述第 K 视频帧的位姿信息,加入视觉惯导约束,形成该当前待处理视频帧第 $(K+1)$ 视频帧的初始位姿信息,该初始位姿信息是进入所述时间窗口进行滑窗因子分析时使用的位姿信息;在进入所述滑窗因子分析过程之后,还会根据时间窗口中的所有滑窗因子的彼此关系进行整体的调整。

[0187] 步骤S304,通过对应当前待处理的实景图像的先验位姿信息提供先验位姿约束,其中,所述当前待处理的实景图像是准备进入滑窗因子的时间窗口的实景图像。

[0188] 其中,所述先验位姿约束,是指获取视频帧时各个实景图像对应的先验位姿信息之间的约束关系,即,约束了各个连续的实景图像之间位姿信息的变化范围。

[0189] 步骤S305,根据上述各个约束关系,推导当前待处理视频帧的初始位姿信息。

[0190] 上述各个步骤获得了各视频帧的位姿信息的所需准备的约束条件;其中,步骤S302获得了视觉特征重投影约束,步骤S303获得了来自前一个视频帧的视觉惯导约束,步骤S304获得了先验位姿约束;根据上述各个约束的数据,可以推导出当前待处理视频帧的最合理的初始位姿信息。

[0191] 通过本实施例提供的方法,可以获得连续视频帧的初始位姿信息,并依次加入滑窗因子分析的时间窗口中,在其中进行滑窗因子分析;在滑窗因子分析过程中,根据各个视频帧相互之间的关系不断进行调整,直到所述当前待处理视频帧成为图3中的第 $(k-w+1)$ 视频帧,即被边缘化掉,离开所述时间窗口,此时获得的第 $(k-w+1)$ 视频帧的位姿信息为经过滑窗因子优化的位姿信息,作为该视频帧的最终位姿信息。

[0192] 本申请第二实施例和第三实施例分别提供一种实景图像的位姿信息获取方法以及一种针对视频帧的实景图像的位姿信息获取方法,在实际应用的过程中,本申请第三实施例所述的针对视频帧的实景图像的位姿信息获取方法实际上是在第二实施例的基础上进行的拓展,即,将所述视频中的第一帧图像当做实景图像,获取第一帧图像的位姿信息,再根据本申请第三实施例中提供的约束条件和计算方法依次获得整个视频中各视频帧对应的初始位姿信息,并将这些初始位姿信息进入滑窗优化的时间窗口,展开滑窗优化。获取

视频帧的位姿信息时可以将本申请第二实施例所述的实景图像的位姿信息的获取方法看做对视频帧位姿信息初始化的过程,即,获取获取视频中第一帧图像(初始帧)对应的初始位姿信息;本申请第三实施例对初始帧之后的视频帧进行多个约束条件的约束,获得进入滑窗优化的初始位姿信息,然后通过后续的滑窗优化过程,最终获取连续视频帧中各个视频帧对应的位姿信息。

[0193] 上述第三实施例重点说明了对连续视频帧的初始位姿信息的获得方法进行了说明;本申请第四实施例提供一种包含滑窗优化过程的获得视频帧位姿信息的方法;以下结合图6说明一种采用模块方式表达的使用滑窗优化方式获得视频帧位姿信息的更具体的技术方案;该方案实现了对视频流的连续处理,以获得视频流中各个视频帧的位姿信息。

[0194] 如图6,所述滑窗优化的流程包括7个相互衔接并循环的处理模块。

[0195] 模块301,滑窗优化模块,该模块引入滑窗因子以及对应各个滑窗因子的相关约束;所述优化,优化变量为W个滑窗因子,每个滑窗因子可以理解为对应W个连续视频帧中的一个视频帧的位姿信息的位姿变量(因为尚未确定位姿信息具体数值,在此称为位姿变量);W的具体数值由设置的时间窗口确定;每个位姿变量(滑窗因子)存在一个RTK-VIO先验位姿约束,由视觉惯导里程计的视觉惯导定位信息,以及GPS设备获得的GPS定位信息,两者相互融合获得;同时,也要考虑反映相邻视频帧之间影响的视觉惯导约束,以及反映视觉投影的视觉特征重投影约束。在滑窗优化中,首先是对于位于时间窗口中的W个视频帧进行整体优化,即考虑各个视频帧的时间间隔与视觉惯导约束等关系,使时间窗口中的W个视频帧的位姿信息具有最合理的数据;其次,是将超出时间窗口的视频帧去除,即边缘化。该视频帧获得最终确定的位姿信息。滑窗优化的具体方法和解释在前述第一实施例已经有相关说明,在此不再赘述。

[0196] 所述视觉惯导约束由RTK-VIO提供的绝对位姿计算两视频帧之间的相对运动作为测量值得出。

[0197] 在所述滑窗优化中,当第k+1帧照片的观测到来时,边缘化掉虚线框内到k+1帧的时间超过时间窗口的变量,即第(K-W+1)视频。此时,如图3的因子图中,加入实线框内第k+1帧变量和相关的约束因子,记录新增变量的时间戳。

[0198] 模块302边缘化处理获得滑窗优化时间窗口中当前帧的状态,包括位姿信息和协方差矩阵。所述位姿信息,就是本实施例希望获得的对应当前视频帧的位姿结果。该模块执行对模块301的滑窗因子的计算,获得对应当前视频帧的位姿信息和协方差矩阵。所述当前帧为边缘化处理后的视频帧,即图3中第(K-W+1)视频帧。

[0199] 模块303进行位姿预测,具体是通过视觉惯导约束的方式,使用时间窗口中最后视频帧的状态信息,预测下一个视频帧(当前待处理视频帧,即将进入时间窗口的视频帧)的信息。在本例中,前一个帧为第K帧,当前待处理视频帧为第K+1帧。

[0200] 模块304将矢量元素投影到图像;具体是获取相关的高精三维矢量地图,并通过索引获得特定类型的矢量元素,例如,路牌或者路标元素,并将矢量元素投影到第k+1视频帧,同时计算像素的协方差矩阵。由于本例中,是形成一个对视频帧进行处理的循环,对于模块304,第K+1帧已经成为当前帧。

[0201] $pixel_{prediction} = K \cdot T_W^C \cdot Landmark$

$$[0202] \quad P_{pixel} = \frac{\partial Pixel}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \hat{P} \left(\frac{\partial Pixel}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \right)^T + \frac{\partial Pixel}{\partial Landmark} P_{Landmark} \left(\frac{\partial Pixel}{\partial Landmark} \right)^T$$

[0203] 其中,K是相机内参, $P_{Landmark}$ 是以协方差表示的三维地图点坐标精度, P_{Pixel} 是三维地图点投影到视频帧中得到的像素坐标的协方差矩阵。 $pixel_{prediction}$ 是投影到图像中的像素坐标。

[0204] 模块305,对特征匹配成功的元素进行马氏距离计算;距离最近的元素作为匹配对。

[0205] 所述马氏距离计算公式如下:

$$[0206] \quad D_{ML} = \frac{\partial Pixel}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \hat{P} \left(\frac{\partial Pixel}{\partial T_C^W} \frac{\partial T_C^W}{\partial x} \right)^T + \frac{\partial Pixel}{\partial Landmark} P_{Landmark} \left(\frac{\partial Pixel}{\partial Landmark} \right)^T$$

[0207] 马氏距离的意义以及公式中变量含义在前面实施例已经描述,在此不再赘述。

[0208] 模块306进行局部因子图优化,具体是使用模块305获得的视觉重投影约束、以及RTK-VIO先验位姿约束、视觉惯量约束对模块305获得的匹配关系进行局部因子图优化;即获得该当前待处理视频帧的初始位姿信息,或者称为局部优化位姿信息。

[0209] 模块307,判断位姿变化量以及优化后的重投影误差是否正常,也就是排除明显不符合先验知识的结果。该步骤的主要目的是避免明显不合理的推算结果;排除之后,将获得的当前待处理视频帧约束信息输出到模块301进行滑窗优化。

[0210] 上述方式实现了对连续视频帧的循环处理,能为连续视频帧进行位姿信息标定。

[0211] 由此可见,本申请第三实施例提供的上述方法,将基本方法与时间窗口分析方法相结合,实现了对连续视频帧的位姿信息标注;能够为高精地图的众包更新提供更为可靠的数据支持。

[0212] 本申请第四实施例提供一种获得实景图像位姿信息的装置,请参考图7,其为本申请第四实施例提供的获得位姿信息的装置的框图。由于该装置实施例基本相似与上述场景实施例和方法实施例,所以描述的比较简单,相关之处参见上述方法实施例的部分说明即可。

[0213] 本申请提供的获得实景图像位姿信息的装置,包括:

[0214] 信息获得单元401,用于获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;

[0215] 地图获得单元402,用于获得包含该实景图像所反映地理位置的三维矢量地图;

[0216] 元素获得单元403,用于在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;

[0217] 匹配单元404,用于建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;

[0218] 推算单元406,根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X。

[0219] 可选的,所述建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系,包括:

- [0220] 提取所述矢量地图元素的关键特征,以及识别所述图像元素的关键特征;
- [0221] 将关键特征吻合的所述矢量地图元素与所述图像元素作为相互匹配的矢量地图元素,彼此建立匹配关系。
- [0222] 可选的,所述根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X,包括:
- [0223] 获得所述实景图像中的与所述矢量地图元素匹配的图像元素的特征点;
- [0224] 根据所述特征点在所述实景图像中的视觉测量值、所述特征点在所述三维矢量地图中对应的空间坐标信息、获取所述实景图像的设备参数,获得所述获取所述实景图像的设备观测位姿信息;
- [0225] 根据所述先验位姿信息和获取所述实景图像的设备位姿噪声,获得所述先验位姿的观测信息;
- [0226] 根据所述获取所述实景图像的设备观测位姿信息、所述先验位姿的观测信息以及预设的位姿优化方程,推算所述实景图像的位姿信息X。
- [0227] 可选的,所述先验位姿信息采用如下方式获得:
- [0228] 采用获得实景图像的视觉惯导里程计定位信息和GPS定位信息融合,获得对应所述实景图像的先验位姿信息。
- [0229] 可选的,所述实景图像为连续的视频帧;
- [0230] 所述根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X,包括:
- [0231] 建立包含W个视频帧的时间窗口;
- [0232] 通过对应当前待处理的实景图像的先验位姿信息提供先验位姿约束,其中,所述当前待处理的实景图像是准备进入滑窗因子的时间窗口的实景图像;
- [0233] 若所述实景图像中存在与三维矢量地图中的矢量地图元素匹配的图像元素,则根据所述矢量地图元素中的空间坐标信息以及所述实景地图中的像素坐标信息,为该视频帧加入视觉特征重投影约束;
- [0234] 根据视觉惯导约束的相关信息,建立当前待处理实景图像与前一个实景图像之间的视觉惯导约束;
- [0235] 根据上述约束关系,对时间窗口中的各个视频帧采用滑窗因子优化的方式确定所述实景图像位姿信息X。
- [0236] 可选的,所述根据视觉惯导约束的相关信息,建立当前待处理实景图像与前一个实景图像之间的视觉惯导约束,包括:
- [0237] 根据所述当前待处理实景图像的先验位姿信息和所述前一个实景图像对应的先验位姿信息,建立所述当前待处理实景图像与前一个实景图像之间的视觉惯导约束。
- [0238] 可选的,对于所述每帧实景图像,当所述时间窗口中进入第K+1个实景图像时,将超出所述时间窗口的第K-W+1个实景图像进行边缘化处理,并根据所述约束关系,获得所述第K+1个实景图像的位姿信息X。
- [0239] 本申请第五实施例提供一种电子设备,其特征在于,包括:

[0240] 处理器501;

[0241] 存储器502,用于存储获得实景图像位姿信息的方法的程序,所述程序在被处理器处理后执行以下步骤:获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;获得包含该实景图像所反映地理位置的三维矢量地图;在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X。

[0242] 本申请同时提供一种计算机存储介质,所述计算机存储介质存储有计算机程序,所述程序被执行时执行以下步骤:获得一帧以上的实景图像的图像元素及每帧实景图像的先验位姿信息;获得包含该实景图像所反映地理位置的三维矢量地图;在所述三维矢量地图中,获得所述实景图像获取位置的预定范围内的与所述图像元素相关的矢量地图元素;建立与各个所述图像元素相关的矢量地图元素与各个所述图像元素之间的匹配关系;根据所述矢量地图元素的空间坐标信息,以及与所述矢量元素匹配的所述图像元素在所述实景图像中的像素坐标信息,对所述先验位姿信息进行优化,推算获得所述实景图像的位姿信息X。

[0243] 在一个典型的配置中,上述实施例配置在计算设备中,所述计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0244] 内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0245] 1、计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带,磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读介质不包括非暂存电脑可读媒体(transitory media),如调制的数据信号和载波。

[0246] 2、本领域技术人员应明白,本申请的实施例可提供为方法、系统或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本申请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、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

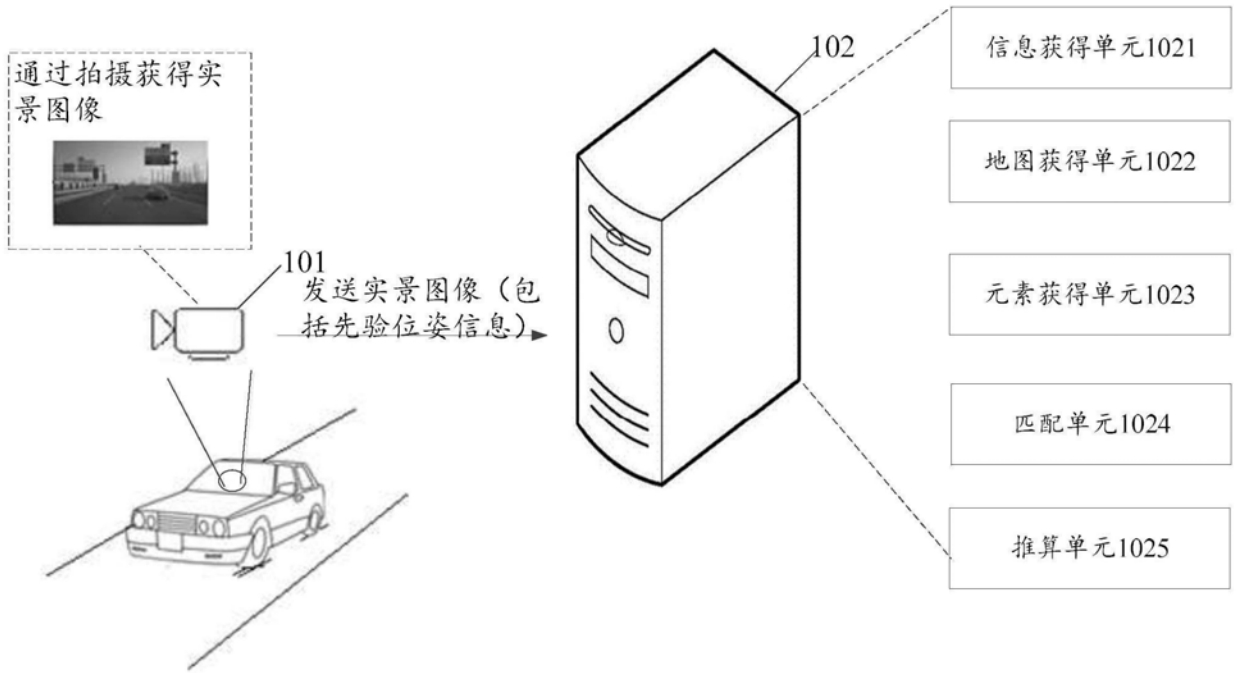


图1

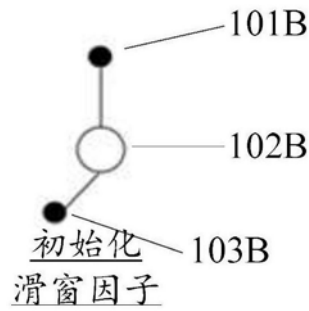


图2

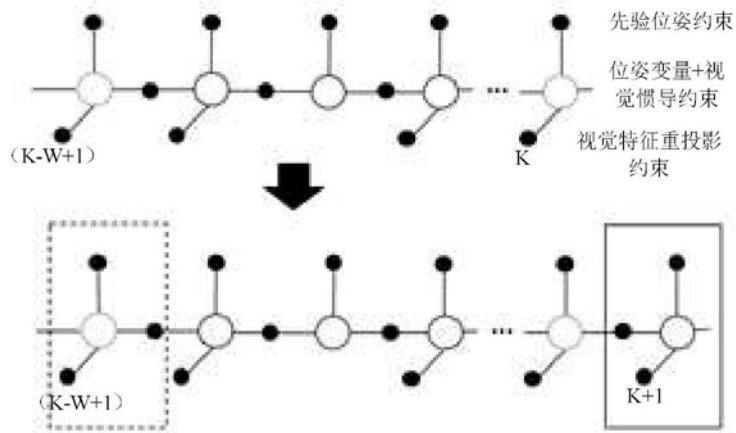


图3

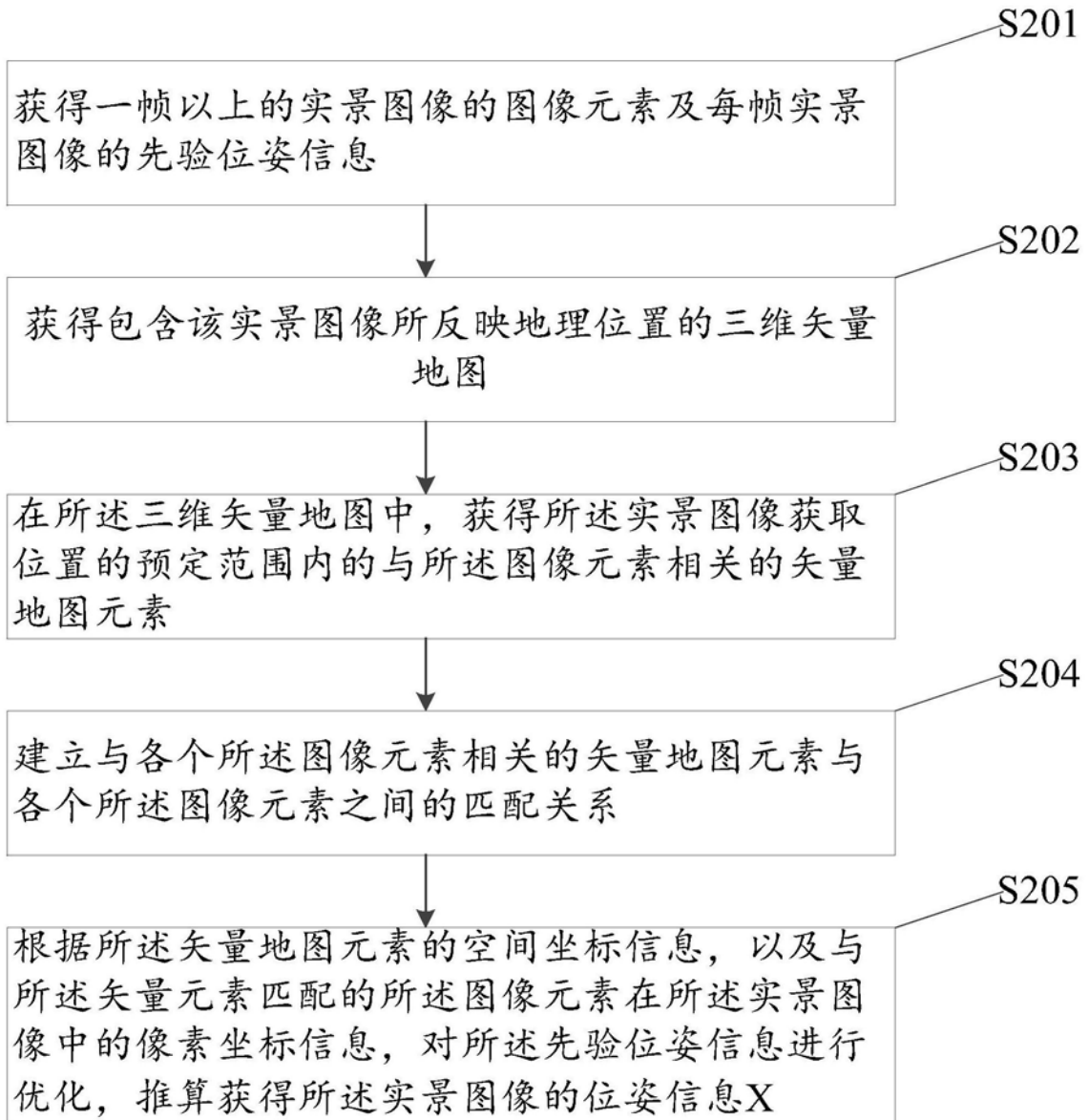


图4

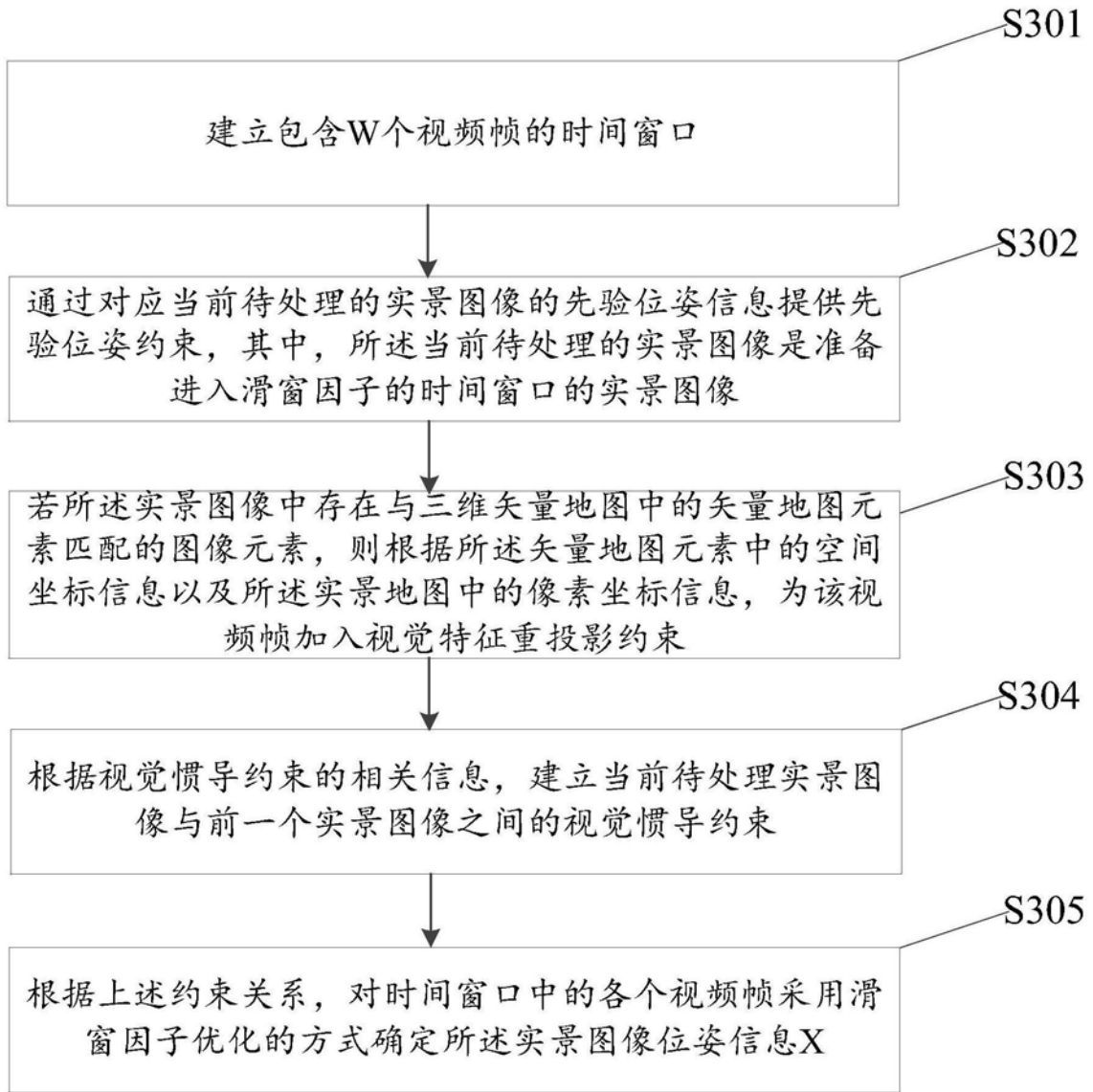


图5

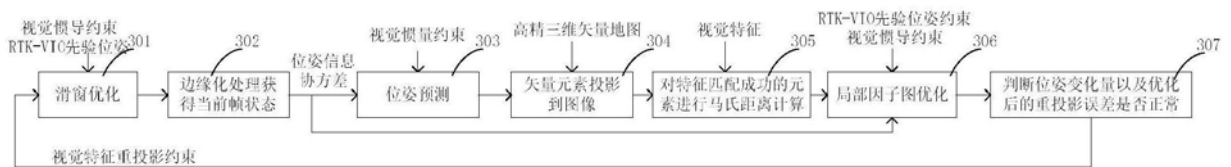


图6

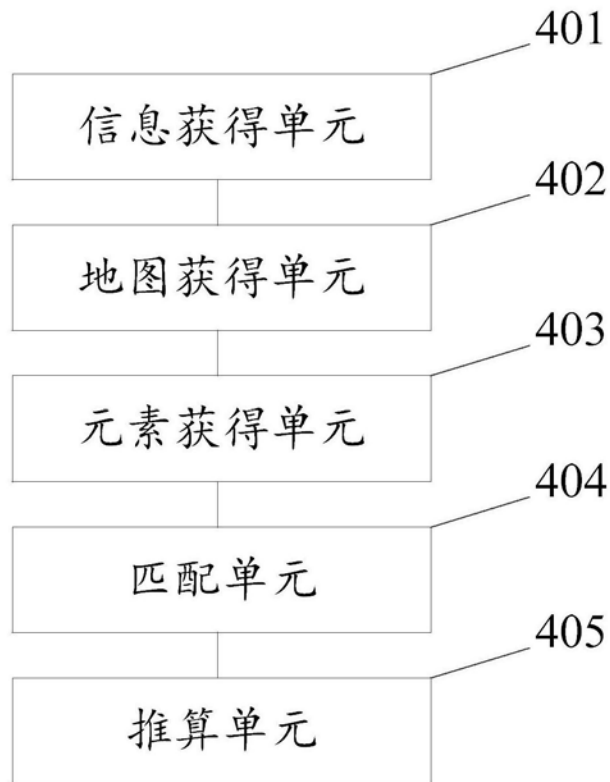


图7



图8