



(12) 发明专利

(10) 授权公告号 CN 110211190 B

(45) 授权公告日 2022. 11. 01

(21) 申请号 201910468496.8

(22) 申请日 2019.05.31

(65) 同一申请的已公布的文献号  
申请公布号 CN 110211190 A

(43) 申请公布日 2019.09.06

(73) 专利权人 北京百度网讯科技有限公司  
地址 100085 北京市海淀区上地十街10号  
百度大厦2层

(72) 发明人 秦硕 李金鹏

(74) 专利代理机构 北京市铸成律师事务所  
11313

专利代理师 邓海鸿 陈建民

(51) Int. Cl.  
G06T 7/80 (2017.01)

(56) 对比文件

CN 109543493 A, 2019.03.29

US 2018157918 A1, 2018.06.07

审查员 苏晓燕

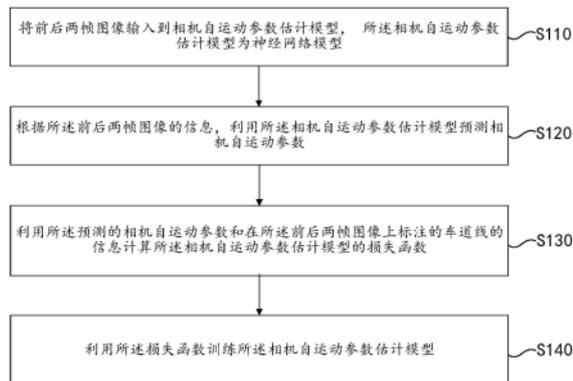
权利要求书3页 说明书15页 附图4页

(54) 发明名称

相机自运动参数估计模型的训练方法、装置及存储介质

(57) 摘要

本发明实施例提出一种相机自运动参数估计模型的训练方法、装置及计算机可读存储介质。其中相机自运动参数估计模型的训练方法包括：将前后两帧图像输入到相机自运动参数估计模型，相机自运动参数估计模型为神经网络模型；根据前后两帧图像的信息，利用相机自运动参数估计模型预测相机自运动参数；利用预测的相机自运动参数和在前后两帧图像上标注的车道线的信息计算相机自运动参数估计模型的损失函数；利用损失函数训练相机自运动参数估计模型。本发明实施例通过使用神经网络模型减小了计算量，使计算速度更快；考虑到图像的全局信息，避免局部信息错误匹配造成的输出错误；训练数据可覆盖各种场景，可准确预测相机自运动参数，鲁棒性较好。



1. 一种相机自运动参数估计模型的训练方法,其特征在于,包括:

将前后两帧图像输入到相机自运动参数估计模型,所述相机自运动参数估计模型为神经网络模型;

根据所述前后两帧图像的信息,利用所述相机自运动参数估计模型预测相机自运动参数;

利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数;

利用所述损失函数训练所述相机自运动参数估计模型;

其中,所述前后两帧图像包括第一图像和第二图像,所述第一图像是所述第二图像之前的一帧图像且所述第一图像与所述第二图像之间间隔N帧图像,其中,N是大于等于0的整数;

所述在所述前后两帧图像上标注的车道线的信息包括:在所述第一图像上标注的第一车道线的标注点和在所述第二图像上标注的第二车道线的标注点,且所述第一车道线对应于所述第二车道线;

利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数,包括:

利用所述预测的相机自运动参数和逆透视映射算法计算出所述第一车道线的标注点投影到所述第二图像上的投影点;

计算所述投影点与所述投影点在所述第二车道线上对应的标注点之间的距离,将所述距离作为所述损失函数的值。

2. 根据权利要求1所述的方法,其特征在于,N的取值范围为 $4 \geq N \geq 0$ 。

3. 根据权利要求1所述的方法,其特征在于,利用所述预测的相机自运动参数和逆透视映射算法计算出所述第一车道线的标注点投影到所述第二图像上的投影点,包括:

利用相机坐标系和图像坐标系的变换关系,将所述第一车道线上的标注点在图像坐标系下的坐标转换为相机坐标系下的坐标;

利用所述预测的相机自运动参数和所述第一车道线上的标注点在相机坐标系下的坐标,计算所述投影点在相机坐标系下的坐标;

利用相机坐标系和图像坐标系的变换关系,将所述投影点在相机坐标系下的坐标转换为所述投影点在图像坐标系下的坐标。

4. 根据权利要求1所述的方法,其特征在于,计算所述投影点与所述投影点在所述第二车道线上对应的点之间的距离,将所述距离作为所述损失函数的值,包括:

利用以下公式计算所述损失函数的值:

$$j = \arg \min(\sqrt{(x'_i - (x_j + x_{j+1})/2)^2 + (y'_i - (y_i + y_{j+1})/2)^2})$$

$$x_{jm} = \frac{y'_i - y_{j+1}}{y_{j+1} - y_j} \times (x_{j+1} - x_j) + x_{j+1}$$

$$\text{Loss} = \sum_i |x'_i - x_{jm}|$$

其中,i和j表示车道线上的标注点的索引编号; $(x'_i, y'_i)$ 表示所述投影点的坐标; $(x_j,$

$y_j$ )表示所述第二车道线上的标注点的坐标; $x_{jm}$ 表示所述投影点在所述第二车道线上对应的点;Loss表示所述损失函数的值。

5.根据权利要求1至4中任一项所述的方法,其特征在于,利用所述损失函数训练所述相机自运动参数估计模型,包括:

利用随机梯度下降法对所述相机自运动参数估计模型进行训练,通过计算所述损失函数对所述相机自运动参数估计模型进行反向传播,更新所述相机自运动参数估计模型的网络参数;

在训练到所述相机自运动参数估计模型收敛的情况下,将所述相机自运动参数估计模型的网络参数固定。

6.根据权利要求1至4中任一项所述的方法,其特征在于,利用所述损失函数训练所述相机自运动参数估计模型之后,还包括:

利用训练好的所述相机自运动参数估计模型估计所述相机自运动参数。

7.一种相机自运动参数估计模型的训练装置,其特征在于,包括:

输入单元,用于将前后两帧图像输入到相机自运动参数估计模型,所述相机自运动参数估计模型为神经网络模型;

预测单元,用于根据所述前后两帧图像的信息,利用所述相机自运动参数估计模型预测相机自运动参数;

计算单元,用于利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数;

训练单元,用于利用所述损失函数训练所述相机自运动参数估计模型;

所述前后两帧图像包括第一图像和第二图像,所述第一图像是所述第二图像之前的一帧图像且所述第一图像与所述第二图像之间间隔N帧图像,其中,N是大于等于0的整数;

所述在所述前后两帧图像上标注的车道线的信息包括:在所述第一图像上标注的第一车道线的标注点和在所述第二图像上标注的第二车道线的标注点,且所述第一车道线对应于所述第二车道线;

所述计算单元包括:

投影点计算子单元,用于利用所述预测的相机自运动参数和逆透视映射算法计算出所述第一车道线的标注点投影到所述第二图像上的投影点;

距离计算子单元,用于计算所述投影点与所述投影点在所述第二车道线上对应的标注点之间的距离,将所述距离作为所述损失函数的值。

8.根据权利要求7所述的装置,其特征在于,N的取值范围为 $4 \geq N \geq 0$ 。

9.根据权利要求7所述的装置,其特征在于,所述投影点计算子单元用于:

利用相机坐标系和图像坐标系的变换关系,将所述第一车道线上的标注点在图像坐标系下的坐标转换为相机坐标系下的坐标;

利用所述预测的相机自运动参数和所述第一车道线上的标注点在相机坐标系下的坐标,计算所述投影点在相机坐标系下的坐标;

利用相机坐标系和图像坐标系的变换关系,将所述投影点在相机坐标系下的坐标转换为所述投影点在图像坐标系下的坐标。

10.根据权利要求7所述的装置,其特征在于,所述距离计算子单元用于利用以下公式

计算所述损失函数的值:

$$j = \arg \min(\sqrt{(x'_i - (x_j + x_{j+1})/2)^2 + (y'_i - (y_j + y_{j+1})/2)^2})$$

$$x_{jm} = \frac{y'_i - y_{j+1}}{y_{j+1} - y_j} \times (x_{j+1} - x_j) + x_{j+1}$$

$$\text{Loss} = \sum_i |x'_i - x_{jm}|$$

其中,  $i$  和  $j$  表示车道线上的标注点的索引编号;  $(x'_i, y'_i)$  表示所述投影点的坐标;  $(x_j, y_j)$  表示所述第二车道线上的标注点的坐标;  $x_{jm}$  表示所述投影点在所述第二车道线上对应的点; Loss 表示所述损失函数的值。

11. 根据权利要求7至10中任一项所述的装置, 其特征在于, 所述训练单元用于:

利用随机梯度下降法对所述相机自运动参数估计模型进行训练, 通过计算所述损失函数对所述相机自运动参数估计模型进行反向传播, 更新所述相机自运动参数估计模型的网络参数;

在训练到所述相机自运动参数估计模型收敛的情况下, 将所述相机自运动参数估计模型的网络参数固定。

12. 根据权利要求7至10中任一项所述的装置, 其特征在于, 所述装置还包括估计单元, 所述估计单元用于:

利用训练好的所述相机自运动参数估计模型估计所述相机自运动参数。

13. 一种相机自运动参数估计模型的训练装置, 其特征在于, 包括:

一个或多个处理器;

存储装置, 用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行时, 使得所述一个或多个处理器实现如权利要求1至6中任一项所述的方法。

14. 一种相机自运动参数估计设备, 其特征在于, 包括:

一个或多个处理器;

存储装置, 用于存储一个或多个程序;

所述设备用于: 利用权利要求13所述的装置训练好的相机自运动参数估计模型估计所述相机自运动参数。

15. 一种相机自运动参数估计系统, 其特征在于, 包括: 如权利要求13所述的相机自运动参数估计模型的训练装置和如权利要求14所述的相机自运动参数估计设备。

16. 一种计算机可读存储介质, 其存储有计算机程序, 其特征在于, 该程序被处理器执行时实现如权利要求1至6中任一项所述的方法。

## 相机自运动参数估计模型的训练方法、装置及存储介质

### 技术领域

[0001] 本发明涉及信息技术领域,尤其涉及一种相机自运动参数估计模型的训练方法、装置及计算机可读存储介质。

### 背景技术

[0002] 相机自运动参数是自动驾驶技术领域中非常重要的一组参数,其主要作用体现在以下几个方面:

[0003] 1. 校正车辆外部参数。

[0004] 通常情况下,车辆和相机是相对静止的。在这种情况下,车辆外部参数就是相机的外参。相机的外参包括旋转矩阵和平移矩阵。在进行环境感知的过程中,图像中感知到的物体需要经过车辆外部参数变换才能得到真实的距离。车辆的外部参数需要经过专门标定,标定的时候车辆是静态的,车辆行驶过程中是动态的,因此车辆和摄像头的自运动会导致车辆外部参数不准,需要对车辆外部参数进行校正。

[0005] 2. 预测和补全车道线。

[0006] 在车道线检测过程中,经常出现车道线模糊的现象。这种情况下车道的划分还是存在的,但是车道线检测算法不能很好地检测到车道线。针对这种情况可以通过历史帧的车道线检测结果,进行车辆位姿变换来预测当前帧的车道线位置。而进行车辆位姿变换需要使用相机自运动参数。以上方法也可以对车道线检测结果进行多重校验,保证结果的稳定性。

[0007] 现有的相机自运动参数估计采用的是关键点检测与匹配的方法。其中,特征提取算法包括: SURF (Speeded Up Robust Features, 加速稳健特征) 算法、FAST (Features from accelerated segment test, 加速段测试的特征) 算法、SIFT (Scale-invariant feature transform, 尺度不变特征变换) 算法、ORB (Oriented FAST and Rotated BRIEF) 算法、Harris (哈里斯) 角点检测算法。特征匹配算法包括: 利用特征提取算法获得关键点的特征, 即一系列的特征向量; 然后通过计算特征向量的距离来进行特征点匹配。

[0008] 现有的相机自运动估计方法存在以下缺陷:

[0009] 1. 现有方法计算速度比较慢。关键点检测通常使用HOG (Histogram of Oriented Gradient, 方向梯度直方图)、HARRIS、SIFT等算法。这类方法对图像中每个像素周围的一片区域进行运算, 得到特征向量。在选取特征向量之后还要对不同帧之间的特征点进行匹配, 需要花费大量的计算资源。

[0010] 2. 现有方法不准确。关键点检测算法通常使用的是局部特征点算法, 并没有考虑到图像的全局信息, 容易造成错误的匹配。比如, 道路上的车道线是连续重复的线段, 图像的边缘或角点特征是相似的, 因此在匹配的时候可能匹配到另一段车道线, 这种情况下计算出来的相机运动参数是错误的。

[0011] 3. 现有方法鲁棒性较差。现有方法基于特征点检测, 然而在某些情况下图像上没有明显的角点或边缘等特征点。比如一些空旷的路段场景比较单一。另外, 在遭遇强光的时候

候,现有方法也不能检测到关键点。比如:示例场景一、对向来车的远光灯让图像过曝;示例场景二、迎着太阳开车的时候路面有大量的反光。在以上示例场景中,使用现有方法不能检测到关键点。

## 发明内容

[0012] 本发明实施例提供一种相机自运动参数估计模型的训练方法、装置及计算机可读存储介质,以解决现有技术中的一个或多个技术问题。

[0013] 第一方面,本发明实施例提供了一种相机自运动参数估计模型的训练方法,包括:

[0014] 将前后两帧图像输入到相机自运动参数估计模型,所述相机自运动参数估计模型为神经网络模型;

[0015] 根据所述前后两帧图像的信息,利用所述相机自运动参数估计模型预测相机自运动参数;

[0016] 利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数;

[0017] 利用所述损失函数训练所述相机自运动参数估计模型。

[0018] 在一种实施方式中,所述前后两帧图像包括第一图像和第二图像,所述第一图像是所述第二图像之前的一帧图像且所述第一图像与所述第二图像之间间隔N帧图像,其中,N是大于等于0的整数;

[0019] 所述在所述前后两帧图像上标注的车道线的信息包括:在所述第一图像上标注的所述第一车道线的信息和在所述第二图像上标注的所述第二车道线的信息,且所述第一车道线对应于所述第二车道线;

[0020] 利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数,包括:

[0021] 利用所述预测的相机自运动参数和逆透视映射算法计算出投影点,所述投影点是所述第一车道线上的标注点投影到所述第二图像上的点;

[0022] 计算所述投影点与所述投影点在所述第二车道线上对应的点之间的距离,将所述距离作为所述损失函数的值。

[0023] 在一种实施方式中,N的取值范围为 $4 \geq N \geq 0$ 。

[0024] 在一种实施方式中,利用所述预测的相机自运动参数和逆透视映射算法计算出投影点,包括:

[0025] 利用相机坐标系和图像坐标系的变换关系,将所述第一车道线上的标注点在图像坐标系下的坐标转换为相机坐标系下的坐标;

[0026] 利用所述预测的相机自运动参数和所述第一车道线上的标注点在相机坐标系下的坐标,计算所述投影点在相机坐标系下的坐标;

[0027] 利用相机坐标系和图像坐标系的变换关系,将所述投影点在相机坐标系下的坐标转换为所述投影点在图像坐标系下的坐标。

[0028] 在一种实施方式中,计算所述投影点与所述投影点在所述第二车道线上对应的点之间的距离,将所述距离作为所述损失函数的值,包括:

[0029] 利用以下公式计算所述损失函数的值:

$$[0030] \quad j = \arg \min \left( \sqrt{(x'_i - (x_j + x_{j+1})/2)^2 + (y'_i - (y_j + y_{j+1})/2)^2} \right)$$

$$[0031] \quad x_{jm} = \frac{y'_i - y_{j+1}}{y_{j+1} - y_j} \times (x_{j+1} - x_j) + x_{j+1}$$

$$[0032] \quad \text{Loss} = \sum_i |x'_i - x_{jm}|$$

[0033] 其中,  $i$  和  $j$  表示车道线上的标注点的索引编号;  $(x'_i, y'_i)$  表示所述投影点的坐标;  $(x_j, y_j)$  表示所述第二车道线上的标注点的坐标;  $x_{jm}$  表示所述投影点在所述第二车道线上对应的点; Loss 表示所述损失函数的值。

[0034] 在一种实施方式中, 利用所述损失函数训练所述相机自运动参数估计模型, 包括:

[0035] 利用随机梯度下降法对所述相机自运动参数估计模型进行训练, 通过计算所述损失函数对所述相机自运动参数估计模型进行反向传播, 更新所述相机自运动参数估计模型的网络参数;

[0036] 在训练到所述相机自运动参数估计模型收敛的情况下, 将所述相机自运动参数估计模型的网络参数固定。

[0037] 在一种实施方式中, 利用所述损失函数训练所述相机自运动参数估计模型之后, 还包括:

[0038] 利用训练好的所述相机自运动参数估计模型估计所述相机自运动参数。

[0039] 第二方面, 本发明实施例提供了一种相机自运动参数估计模型的训练装置, 包括:

[0040] 输入单元, 用于将前后两帧图像输入到相机自运动参数估计模型, 所述相机自运动参数估计模型为神经网络模型;

[0041] 预测单元, 用于根据所述前后两帧图像的信息, 利用所述相机自运动参数估计模型预测相机自运动参数;

[0042] 计算单元, 用于利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数;

[0043] 训练单元, 用于利用所述损失函数训练所述相机自运动参数估计模型。

[0044] 在一种实施方式中, 所述前后两帧图像包括第一图像和第二图像, 所述第一图像是所述第二图像之前的一帧图像且所述第一图像与所述第二图像之间间隔  $N$  帧图像, 其中,  $N$  是大于等于 0 的整数;

[0045] 所述在所述前后两帧图像上标注的车道线的信息包括: 在所述第一图像上标注的所述第一车道线的信息和在所述第二图像上标注的所述第二车道线的信息, 且所述第一车道线对应于所述第二车道线;

[0046] 所述计算单元包括:

[0047] 投影点计算子单元, 用于利用所述预测的相机自运动参数和逆透视映射算法计算出投影点, 所述投影点是所述第一车道线上的标注点投影到所述第二图像上的点;

[0048] 距离计算子单元, 用于计算所述投影点与所述投影点在所述第二车道线上对应的点之间的距离, 将所述距离作为所述损失函数的值。

[0049] 在一种实施方式中,  $N$  的取值范围为  $4 \geq N \geq 0$ 。

[0050] 在一种实施方式中, 所述投影点计算子单元用于:

[0051] 利用相机坐标系和图像坐标系的变换关系,将所述第一车道线上的标注点在图像坐标系下的坐标转换为相机坐标系下的坐标;

[0052] 利用所述预测的相机自运动参数和所述第一车道线上的标注点在相机坐标系下的坐标,计算所述投影点在相机坐标系下的坐标;

[0053] 利用相机坐标系和图像坐标系的变换关系,将所述投影点在相机坐标系下的坐标转换为所述投影点在图像坐标系下的坐标。

[0054] 在一种实施方式中,所述距离计算子单元用于利用以下公式计算所述损失函数的值:

$$[0055] \quad j = \arg \min (\sqrt{(x'_i - (x_j + x_{j+1})/2)^2 + (y'_i - (y_j + y_{j+1})/2)^2})$$

$$[0056] \quad x_{jm} = \frac{y'_i - y_{j+1}}{y_{j+1} - y_j} \times (x_{j+1} - x_j) + x_{j+1}$$

$$[0057] \quad \text{Loss} = \sum_i |x'_i - x_{jm}|$$

[0058] 其中,i和j表示车道线上的标注点的索引编号; $(x'_i, y'_i)$ 表示所述投影点的坐标; $(x_j, y_j)$ 表示所述第二车道线上的标注点的坐标; $x_{jm}$ 表示所述投影点在所述第二车道线上对应的点;Loss表示所述损失函数的值。

[0059] 在一种实施方式中,所述训练单元用于:

[0060] 利用随机梯度下降法对所述相机自运动参数估计模型进行训练,通过计算所述损失函数对所述相机自运动参数估计模型进行反向传播,更新所述相机自运动参数估计模型的网络参数;

[0061] 在训练到所述相机自运动参数估计模型收敛的情况下,将所述相机自运动参数估计模型的网络参数固定。

[0062] 在一种实施方式中,所述装置还包括估计单元,所述估计单元用于:

[0063] 利用训练好的所述相机自运动参数估计模型估计所述相机自运动参数。

[0064] 第三方面,本发明实施例提供了一种相机自运动参数估计模型的训练装置,所述装置的功能可以通过硬件实现,也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。

[0065] 在一个可能的设计中,所述装置的结构中包括处理器和存储器,所述存储器用于存储支持所述装置执行上述相机自运动参数估计模型的训练方法的程序,所述处理器被配置为用于执行所述存储器中存储的程序。所述装置还可以包括通信接口,用于与其他设备或通信网络通信。

[0066] 第四方面,本发明实施例提供了一种相机自运动参数估计设备,所述设备用于:利用上述相机自运动参数估计模型的训练装置训练好的相机自运动参数估计模型估计所述相机自运动参数。

[0067] 第五方面,本发明实施例提供了一种相机自运动参数估计系统,包括上述相机自运动参数估计模型的训练装置和上述相机自运动参数估计设备。

[0068] 第六方面,本发明实施例提供了一种计算机可读存储介质,用于存储相机自运动参数估计模型的训练装置所用的计算机软件指令,其包括用于执行上述相机自运动参数估

计模型的训练方法所涉及的程序。

[0069] 上述技术方案具有如下优点或有益效果：上述技术方案利用神经网络模型估计相机自运动参数。由于神经网络模型的基本算子很少，且输入图像可以很小，因而减小了计算量，使计算速度更快；神经网络的表达能力很强，可以充分考虑到图像的全局信息，避免局部信息错误匹配造成的输出错误；神经网络的训练数据可以覆盖各种各样的场景，例如在过曝、反光、雨雾等场景下也可以准确地预测相机自运动参数，因此本发明实施例的相机自运动参数估计模型具有较好的鲁棒性。

[0070] 上述概述仅仅是为了说明书的目的，并不意图以任何方式进行限制。除上述描述的示意性的方面、实施方式和特征之外，通过参考附图和以下的详细描述，本发明进一步的方面、实施方式和特征将会是容易明白的。

### 附图说明

[0071] 在附图中，除非另外规定，否则贯穿多个附图相同的附图标记表示相同或相似的部件或元素。这些附图不一定是按照比例绘制的。应该理解，这些附图仅描绘了根据本发明公开的一些实施方式，而不应将其视为是对本发明范围的限制。

[0072] 图1示出根据本发明实施例的相机自运动参数估计模型的训练方法的流程图。

[0073] 图2示出根据本发明实施例的相机自运动参数估计模型的训练方法的姿态角示意图。

[0074] 图3示出根据本发明实施例的相机自运动参数估计模型的训练方法的图像坐标示意图。

[0075] 图4示出根据本发明实施例的相机自运动参数估计模型的训练方法的流程图。

[0076] 图5示出根据本发明实施例的相机自运动参数估计模型的训练方法的损失函数计算示意图。

[0077] 图6示出根据本发明实施例的相机自运动参数估计模型的训练方法的流程图。

[0078] 图7示出根据本发明实施例的相机自运动参数估计模型的训练装置的结构框图。

[0079] 图8示出根据本发明实施例的相机自运动参数估计模型的训练装置的计算单元的结构框图。

[0080] 图9示出根据本发明实施例的相机自运动参数估计模型的训练装置的结构框图。

[0081] 图10示出根据本发明实施例的相机自运动参数估计模型的训练装置的结构框图。

### 具体实施方式

[0082] 在下文中，仅简单地描述了某些示例性实施例。正如本领域技术人员可认识到的那样，在不脱离本发明的精神或范围的情况下，可通过各种不同方式修改所描述的实施例。因此，附图和描述被认为本质上是示例性的而非限制性的。

[0083] 图1示出根据本发明实施例的相机自运动参数估计模型的训练方法的流程图。如图1所示，该相机自运动参数估计模型的训练方法包括：

[0084] 步骤S110，将前后两帧图像输入到相机自运动参数估计模型，所述相机自运动参数估计模型为神经网络模型；

[0085] 步骤S120，根据所述前后两帧图像的信息，利用所述相机自运动参数估计模型预

测相机自运动参数；

[0086] 步骤S130,利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数；

[0087] 步骤S140,利用所述损失函数训练所述相机自运动参数估计模型。

[0088] 神经网络模型是由大量处理单元互联组成的非线性、自适应的信息处理系统,具有自组织、自学习能力。本发明实施例采用神经网络模型估计相机自运动参数。由于神经网络模型的基本算子很少,且输入图像可以很小,因而减小了计算量,使计算速度更快。许多硬件平台都支持神经网络模型的加速,比如可使用NVIDIA(英伟达)公司的TensorRT加速库。英伟达TensorRT是一种高性能神经网络推理引擎,用于在生产环境中部署深度学习应用程序。神经网络的表达能力很强,可以充分考虑到图像的全局信息,避免局部信息错误匹配造成的输出错误。神经网络的训练数据可以覆盖各种各样的场景,例如在过曝、反光、雨雾等场景下也可以准确地预测相机自运动参数,因此本发明实施例的相机自运动参数估计模型具有较好的鲁棒性。

[0089] 具体地,步骤S110中,在相机获取的视频信息中选取前后两帧图像,例如选取相邻的两帧图像。将选取的前后两帧图像输入到神经网络模型中。在步骤S120中,利用神经网络模型预测相机自运动参数,其中,相机自运动参数可包括相机的姿态角,也叫欧拉角。欧拉角是用来确定定点转动刚体位置的3个一组的独立角参数,用来描述刚体在三维欧几里得空间的取向。本发明实施例中,相机的姿态角包括pitch、yaw、roll参数。图2示出根据本发明实施例的相机自运动参数估计模型的训练方法的姿态角示意图。在图2中用飞机的模型展示了pitch、yaw、roll这三个角。其中,pitch是围绕X轴旋转的角度,也叫做俯仰角;yaw是围绕Y轴旋转的角度,也叫偏航角;roll是围绕Z轴旋转的角度,也叫翻滚角。

[0090] 在步骤S130中,利用神经网络模型预测的相机自运动参数和前后两帧图像上标注的车道线的信息计算损失函数的值loss。在步骤S140中,通过计算loss对神经网络模型进行反向传播,更新网络参数,以训练所述相机自运动参数估计模型。

[0091] 在一种实施方式中,所述前后两帧图像包括第一图像和第二图像,所述第一图像是所述第二图像之前的一帧图像且所述第一图像与所述第二图像之间间隔N帧图像,其中,N是大于等于0的整数。

[0092] 在一种实施方式中,N的取值范围为 $4 \geq N \geq 0$ 。例如,可选取相邻的两帧图像作为输入神经网络模型的前后两帧图像。再如,也可以是选择的前后两帧图像之间间隔1至4帧图像。

[0093] 在无人车的外部参数标定过程中,可以获得相机的内参、外参,以及车辆的高度。其中,相机的内参包括相机的焦距和主点的位置信息。相机的外参包括旋转矩阵和平移矩阵。旋转矩阵描述了世界坐标系的坐标轴相对于相机坐标轴的方向。平移矩阵描述了在相机坐标系下空间原点的位置。旋转矩阵和平移矩阵共同描述了如何把点从世界坐标系转换到相机坐标系。

[0094] 假设地面是一个固定不变的平面,因此在图像上选取地面上的点,即可通过IPM(inverse perspective mapping,逆透视映射)算法获得图像上选取的地面上的点在相机坐标系下的三维坐标。

[0095] 其中,逆透视映射是透视映射的逆过程,可利用相机的角度、高度等位置信息建立

三维坐标系,消除透视作用,得到场景的俯视图。经过逆透视映射之后,原本有相交趋势的车道线转化为俯视图中的平行线,更便于检测。相机通过实时图像采集能够获取丰富的路面信息,在有合适的光照的前提下,这种信息的获取不受环境其他因素的约束,具有普适性。但由于透视原理所造成的视觉上的畸变导致难以从相机中提取到有效的信息。为了有效地将图像信息利用起来,首先要考虑的是消除这种误差。透视原理所带来的视觉畸变在数学上可以定性地解释为三维景象的不规则投影所造成的。为了消除这种畸变,从数学上将透射过程进行逆变换,将三维景象进行俯视投影。这一逆变换过程称之为逆透视映射。

[0096] 在一种实施方式中,所述在所述前后两帧图像上标注的车道线的信息包括:在所述第一图像上标注的所述第一车道线的信息和在所述第二图像上标注的所述第二车道线的信息,且所述第一车道线对应于所述第二车道线。其中,第一图像上标注的第一车道线和第二图像上标注的第二车道线对应于地面上的同一条车道线。

[0097] 图3示出根据本发明实施例的相机自运动参数估计模型的训练方法的图像坐标示意图。图3中的字母标号含义如下: $u_1$ 和 $v_1$ 表示在第一图像上标注的第一车道线上的标注点在图像坐标系下的坐标; $u_2$ 和 $v_2$ 表示在第二图像上标注的第二车道线上的标注点在图像坐标系下的坐标; $P$ 表示选取的点在世界坐标系下的坐标。参见图3,在理想状态下,投影的车道线与所述第二图像上对应的车道线是重合的,因此,第一车道线上的标注点和第二车道线上的标注点转换到世界坐标系下对应于同一个点 $P$ 。

[0098] 相机图像处理过程中涉及到以下四个坐标系:

[0099] 1) 图像坐标系

[0100] 相机采集的数字图像在计算机内可以存储为数组,数组中的每一个元素可以用像素表示,像素的值即是图像点的亮度(灰度)。在图像上定义直角坐标系 $u-v$ ,每一像素的坐标 $(u,v)$ 分别是该像素在数组中的列数和行数。故 $(u,v)$ 是以像素为单位的图像坐标系的坐标。

[0101] 2) 成像平面坐标系

[0102] 由于图像坐标系只表示像素位于数字图像的列数和行数,并没有用物理单位表示出该像素在图像中的物理位置,因而需要再建立以物理单位(例如厘米)表示的成像平面坐标系 $x-y$ 。用 $(x,y)$ 表示以物理单位度量的成像平面坐标系的坐标。

[0103] 3) 相机坐标系

[0104] 相机坐标系是以相机的聚焦中心为原点,以光轴为 $Z$ 轴建立的三维直角坐标系。相机坐标系的原点为相机的光心,相机坐标系的 $x$ 轴与 $y$ 轴与图像的 $X,Y$ 轴平行,相机坐标系的 $z$ 轴为相机光轴,它与图形平面垂直。

[0105] 4) 世界坐标系

[0106] 在环境中选择一个参考坐标系来描述摄像机和物体的位置,该坐标系称为世界坐标系。相机坐标系和世界坐标系之间的关系可用旋转矩阵 $R$ 与平移向量 $t$ 来描述。

[0107] 可采用以下公式进行坐标变换:

$$[0108] \quad \mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{R} = \begin{bmatrix} r & t \\ 0 & 1 \end{bmatrix}$$

$$[0109] \quad K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

[0110] 其中,K是相机的内参矩阵(Intrinsic matrix); $f_x$ 和 $f_y$ 是表示相机的焦距的参数; $c_x$ 和 $c_y$ 是表示相机的主点的位置信息的参数;R是相机的外参矩阵(Extrinsic matrix);r是相机的旋转矩阵(Camera rotationmatrix);t是相机的平移矩阵。

[0111] 图4示出根据本发明实施例的相机自运动参数估计模型的训练方法的流程图。如图4所示,在一种实施方式中,图1中的步骤S130,利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数,具体可包括:

[0112] 步骤S210,利用所述预测的相机自运动参数和逆透视映射算法计算出投影点,所述投影点是所述第一车道线上的标注点投影到所述第二图像上的点;

[0113] 步骤S220,计算所述投影点与所述投影点在所述第二车道线上对应的点之间的距离,将所述距离作为所述损失函数的值。

[0114] 图5示出根据本发明实施例的相机自运动参数估计模型的训练方法的损失函数计算示意图。在前后两帧图像上,在对应的车道线上标注一系列离散的点。其中,在前后两帧图像上对应的车道线是地面上的同一条车道线。图5中的C点表示投影点,C点所在的弧线表示将第一车道线投影到所述第二图像上得到的车道线,即投影的车道线。图5中的A点和B点表示第二车道线上的标注点,A点和B点所在的弧线表示第二车道线。D点表示所述投影点在所述第二车道线上对应的点。从C点沿着平行于横轴的方向做一条直线,该直线与线段AB的交点就是D点。其中,横轴的方向是与车辆前进方向垂直的方向。

[0115] 具体地,在图4的步骤S210中,利用神经网络模型输出的pitch、roll、yaw角度和IPM算法计算出前后两帧图像中前一帧车道线上的标注点投影到当前帧的点。其中,前后两帧图像中的前一帧也就是第一图像,当前帧也就是第二图像。前一帧车道线上的标注点投影到当前帧的点,也就是第一车道线上的标注点投影到第二图像上的投影点。在步骤S220中,计算投影点与当前帧上第二车道线上对应的点之间的距离,将这个距离作为神经网络模型的loss。在坐标变换的公式中,用 $P_1$ 和 $P_2$ 分别表示在前后二帧图像中选取的点在相机坐标系下的坐标,R表示相机的外参矩阵,则

$$[0116] \quad P_2 = R \cdot P_1$$

[0117] 图像坐标系与相机坐标系之间的换算关系可表示为:

$$[0118] \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{Z} K P_i$$

$$[0119] \quad P_i = Z \times K^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

$$[0120] \quad \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \frac{1}{Z_2} K P_2 = \frac{1}{Z_2} K R P_1 = \frac{Z_1}{Z_2} K R K^{-1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix}$$

[0121] 其中, $P_i$ 表示在各帧图像中选取的点在相机坐标系下的坐标,其具体的坐标值用X、Y、Z表示;Z表示选取的点在相机坐标系下的Z轴坐标值,也是相机坐标系下的物体深度; $Z_1$ 和 $Z_2$ 分别表示在前后二帧图像中选取的点在相机坐标系下的物体深度;(u,v)表示图像坐

标系的坐标。

[0122] 图6示出根据本发明实施例的相机自运动参数估计模型的训练方法的流程图。如图6所示,在一种实施方式中,图4中的步骤S210,利用所述预测的相机自运动参数和逆透视映射算法计算出投影点,具体可包括:

[0123] 步骤S310,利用相机坐标系和图像坐标系的变换关系,将所述第一车道线上的标注点在图像坐标系下的坐标转换为相机坐标系下的坐标;

[0124] 步骤S320,利用所述预测的相机自运动参数和所述第一车道线上的标注点在相机坐标系下的坐标,计算所述投影点在相机坐标系下的坐标;

[0125] 步骤S330,利用相机坐标系和图像坐标系的变换关系,将所述投影点在相机坐标系下的坐标转换为所述投影点在图像坐标系下的坐标。

[0126] 在标注图像中,对于t时刻的图像,选取的第一车道线上的标注点在图像坐标系的坐标用 $P_t^{uv}$ 表示,相机坐标系的坐标用 $P_t^w$ 表示。相机内参矩阵为K,则可知在进行归一化处理之后,相机坐标系和图像坐标系的关系为:

$$[0127] \quad P_t^w = K^{-1}P_t^{uv}$$

[0128] 其中,归一化处理的过程为: $X' = \frac{X}{Z}$ ;  $Y' = \frac{Y}{Z}$ ;  $Z' = 1$ 。

[0129] 相机外参矩阵为R,计算出投影点需要利用外参矩阵对第一车道线上的标注点的坐标做旋转变换。可由神经网络模型预测的相机自运动参数即pitch、yaw、roll参数得到旋转矩阵。旋转变换之前和之后的相机坐标系的坐标分别用 $P_t^w$ 和 $P_t^{w'}$ 表示,旋转变换之前和之后的图像坐标系的坐标分别用 $P_t^{uv}$ 和 $P_t^{uv'}$ 表示,则第一车道线上的标注点通过IPM正变换投射到相机坐标系后,经过旋转变换后的点再通过IPM逆变换投射回图像坐标系的过程为:

$$[0130] \quad P_t^{w'} = RP_t^w$$

$$[0131] \quad P_t^{uv'} = KP_t^{w'} = KRP_t^w = KRK^{-1}P_t^{uv}$$

[0132] 在本发明实施例中,假设在理想状态下,相邻两帧图像中车道线检测的点在图像坐标系中的位置相同。也就是说,在理想状态下,投影的车道线与所述第二图像上对应的车道线是重合的。用 $P_t^{uv'}$ 表示投影点在图像坐标系下的坐标,也就是前一帧车道线上的标注点投影到当前帧的点在图像坐标系下的坐标;用 $P_{t+1}^{uv}$ 表示所述投影点在第二车道线上对应的点在图像坐标系下的坐标,也就是图5中D点的坐标。其中,下标t表示第一图像对应的时刻,也就是就当前帧的前一帧图像对应的时刻;下标“t+1”表示第二图像对应的时刻,也就是就当前帧图像对应的时刻。则上述假设关系可用以下公式表示:

$$[0133] \quad P_t^{uv'} = P_{t+1}^{uv}$$

[0134] 在一种实施方式中,计算所述投影点与所述投影点在所述第二车道线上对应的点之间的距离,将所述距离作为所述损失函数的值,包括:

[0135] 利用以下公式计算所述损失函数的值:

$$[0136] \quad j = \arg \min \left( \sqrt{(x'_i - (x_j + x_{j+1})/2)^2 + (y'_i - (y_i + y_{j+1})/2)^2} \right) \quad (1)$$

$$[0137] \quad x_{jm} = \frac{y'_i - y_{j+1}}{y_{j+1} - y_j} \times (x_{j+1} - x_j) + x_{j+1} \quad (2)$$

$$[0138] \quad \text{Loss} = \sum_i |x'_i - x_{jm}| \quad (3)$$

[0139] 其中,  $i$  和  $j$  表示车道线上的标注点的索引编号;  $(x'_i, y'_i)$  表示所述投影点的坐标;  $(x_j, y_j)$  表示所述第二车道线上的标注点的坐标;  $x_{jm}$  表示所述投影点在所述第二车道线上对应的点;  $\text{Loss}$  表示所述损失函数的值。

[0140] 上述公式 (1) 中的  $\arg$  (argument, 自变量) 是变元的英文缩写。 $\arg \min$  计算的结果就是使等号右边这个式子达到最小值时的变量  $j$  的取值。参见图 5, 公式 (1) 的运算过程也就是在第二车道线上的标注点中找到一个点, 该点和其在第二车道线上的下一个标注点组成一条线段, 使得  $C$  点到该线段的中点距离最小, 则找到的这个点的索引编号就是公式 (1) 的计算结果。找到的这个点对应于图 5 中的  $A$  点。参见图 5,  $A$  点和其在第二车道线上的下一个标注点  $B$  点组成一条线段  $AB$ , 在第二车道线上的标注点中可以找到点  $A$ , 使得  $C$  点到线段  $AB$  的中点取最小值。也就是说, 在第二车道线上的标注点中, 距离  $C$  点最近的两个点是  $A$  点和  $B$  点。

[0141] 上述公式 (2) 是利用插值的方法计算出线段  $AB$  上的一个点  $D$ , 将  $D$  点作为  $C$  点在第二车道线上对应的点。参见图 5, 线段  $CD$  的长度是  $C$  点到线段  $AB$  之间横向的最近距离。也就是说, 从  $C$  点沿着平行于横向的方向做一条直线, 该直线与线段  $AB$  的交点就是  $D$  点。其中, 横向是与车辆前进方向垂直的方向。

[0142] 上述公式 (3) 中的  $|x'_i - x_{jm}|$  计算的是线段  $CD$  的长度, 该长度可表示在  $C$  点计算出的投影的车道线与第二车道线的距离, 也就是投影车道线与实际车道线的距离。参见图 5,  $C$  点是任意一个投影点。公式 (3) 中将在所有的投影点计算出的投影的车道线与第二车道线的距离的和作为相机自运动参数估计模型的损失函数的值。

[0143] 上述公式 (3) 中, 针对多个投影点, 分别计算各投影点对应的投影车道线与实际车道线的距离, 再将各个距离求和, 将求和的结果作为损失函数的值。在另一种实施方式中, 可任取一个投影点, 将该投影点对应的投影车道线与实际车道线的距离作为损失函数的值。在又一种实施方式中, 可针对多个投影点, 按照上述方法分别计算各投影点对应的投影车道线与实际车道线的距离, 再将各个距离求平均值, 将平均值作为损失函数的值。

[0144] 在一个示例中, 可采用一个多输入的卷积神经网络 (CNN, Convolutional Neural Networks) 估计相机自运动参数。例如, 神经网络可包含卷积层, 池化层, 最后通过全卷积层回归出  $\text{pitch}$ 、 $\text{yaw}$ 、 $\text{roll}$  三个参数。

[0145] 另外, 神经网络模型的训练需要训练数据。训练数据包括图片和标注 (label)。本发明实施例中的  $\text{label}$  就是前后两帧图像之间的  $\text{pitch}$ 、 $\text{roll}$ 、 $\text{yaw}$  参数。在一个示例中, 图像标注的方法如下: 在一段视频中, 选取相邻的两帧图像, 或者选取中间间隔为 1 至 4 帧的两帧图像。在一种实施方式中, 两帧图像的  $\text{pitch}$ 、 $\text{roll}$ 、 $\text{yaw}$  参数可以通过车辆的 IMU (Inertial measurement unit, 惯性测量单元) 获得。其中, 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。在另一种实施方式中, 也可以通过手动的方式在两帧图像中选取对应的关键点, 其中, 两帧图像中选取的关键点对应于真实世界中的同一个点。然后根据两帧图像中对应的关键点的信息, 通过矩阵变换的方式计算获得  $\text{pitch}$ 、 $\text{roll}$ 、 $\text{yaw}$  参数。其中, 选取的点可以是角点等纹理分明的点, 例如车道线的边缘, 护栏和地面交接的地方。

[0146] 在一个示例中, 可在选取的两帧图像中以等间隔的点的方式标注出车道线。两帧图像中的对应的车道线必须属于同一条车道。两帧图像中可标注至少 5-10 对线段。利用插值法在线段上选取标注点, 这一过程可由程序自动完成。在一种实施方式中, 利用所述损失

函数训练所述相机自运动参数估计模型,包括:

[0147] 利用随机梯度下降法对所述相机自运动参数估计模型进行训练,通过计算所述损失函数对所述相机自运动参数估计模型进行反向传播,更新所述相机自运动参数估计模型的网络参数;

[0148] 在训练到所述相机自运动参数估计模型收敛的情况下,将所述相机自运动参数估计模型的网络参数固定。

[0149] 梯度下降法是迭代法的一种。在求解机器学习算法的模型参数时,梯度下降法是最常采用的方法之一。在求解损失函数的最小值时,可以通过梯度下降法来一步步的迭代求解,得到最小化的损失函数和模型参数值。在机器学习中,基于基本的梯度下降法发展了两种梯度下降方法,分别为随机梯度下降法和批量梯度下降法。本发明实施例利用随机梯度下降法对所述相机自运动参数估计模型进行训练。

[0150] 反向传播算法是适合于多层神经网络的一种学习算法,它建立在梯度下降法的基础上。反向传播算法的目的是用网络的实际输出与目标矢量之间的误差修改其权值,使输出值与期望的值尽可能地接近,即使网络输出层的误差达到最小。它是通过连续不断地在相对于误差函数斜率下降的方向上计算网络权值和偏差的变化而逐渐逼近目标的。每一次权值和偏差的变化都与网络误差的影响成正比,并以反向传播的方式传递到每一层。

[0151] 反向传播算法由激励传播、权重更新两个环节反复循环迭代,直到网络的响应达到预定的目标范围为止。信息的正向传递与误差的反向传播在正向传播过程中,输入信息从输入经隐含层逐层计算传向输出层,每一层神经元的状态只影响下一层神经元的状态。如果在输出层没有得到期望的输出,则计算输出层的误差变化值,然后转向反向传播,通过网络将误差信号沿原来的连接通路反传回来修改各层神经元的权值直至达到期望目标。

[0152] 在一个示例中,在权重更新的环节中,对于每个神经元突触上的权重,按照以下步骤进行更新:将输入激励和响应误差相乘,从而获得权重的梯度;将这个梯度乘上一个比例并取反后加到权重上。

[0153] 在一个示例中,神经网络模型训练的过程可包括:将两帧图像进行色彩维的拼接。例如,两帧图像的形状是 $1920 \times 1208 \times 3$ ,其中的“3”表示通道数。那么拼接后的图像为 $1920 \times 1208 \times 6$ 。然后将这个6通道的图像输入到一个多层的神经网络中。以颜色通道为例,一个图片被建立或者打开以后是会自动会创建颜色通道的。当编辑图像时,实际上就是在编辑颜色通道。这些通道把图像分解成一个或多个色彩成分。图像的模式决定了颜色通道的数量,如RGB模式有R (Red)、G (Green)、B (Blue) 三个颜色通道,分别代表红色、绿色、蓝色的通道。再如CMYK模式(印刷色彩模式)有C (Cyan)、M (Magenta)、Y (Yellow)、K (black) 四个颜色通道,分别代表青色、洋红、黄色和黑色的通道。又如灰度图只有一个颜色通道。

[0154] 在一种实施方式中,利用所述损失函数训练所述相机自运动参数估计模型之后,还包括:

[0155] 利用训练好的所述相机自运动参数估计模型估计所述相机自运动参数。

[0156] 在训练到神经网络模型收敛之后,将神经网络模型的参数固定,在网络推断的时候使用。其中,网络推断的过程是指神经网络模型训练好之后通过前馈网络(Forward propagation network)得到pitch、yaw、roll参数的过程。神经网络模型导入网络结构和参

数,输入两帧图像即可输出pitch、yaw、roll参数。在前馈神经网络中,各神经元分层排列,每个神经元只与前一层的神经元相连。接收前一层的输出,并输出给下一层。前馈神经网络结构简单,应用广泛,能够以任意精度逼近任意连续函数及平方可积函数,而且可以精确实现任意有限训练样本集。

[0157] 图7示出根据本发明实施例的相机自运动参数估计模型的训练装置的结构框图。如图7所示,本发明实施例的相机自运动参数估计模型的训练装置包括:

[0158] 输入单元100,用于将前后两帧图像输入到相机自运动参数估计模型,所述相机自运动参数估计模型为神经网络模型;

[0159] 预测单元200,用于根据所述前后两帧图像的信息,利用所述相机自运动参数估计模型预测相机自运动参数;

[0160] 计算单元300,用于利用所述预测的相机自运动参数和在所述前后两帧图像上标注的车道线的信息计算所述相机自运动参数估计模型的损失函数;

[0161] 训练单元400,用于利用所述损失函数训练所述相机自运动参数估计模型。

[0162] 图8示出根据本发明实施例的相机自运动参数估计模型的训练装置的计算单元的结构框图。如图8所示,在一种实施方式中,所述前后两帧图像包括第一图像和第二图像,所述第一图像是所述第二图像之前的一帧图像且所述第一图像与所述第二图像之间间隔N帧图像,其中,N是大于等于0的整数;

[0163] 所述在所述前后两帧图像上标注的车道线的信息包括:在所述第一图像上标注的所述第一车道线的信息和在所述第二图像上标注的所述第二车道线的信息,且所述第一车道线对应于所述第二车道线;

[0164] 所述计算单元300包括:

[0165] 投影点计算子单元310,用于利用所述预测的相机自运动参数和逆透视映射算法计算出投影点,所述投影点是所述第一车道线上的标注点投影到所述第二图像上的点;

[0166] 距离计算子单元320,用于计算所述投影点与所述投影点在所述第二车道线上的标注点之间的距离,将所述距离作为所述损失函数的值。

[0167] 在一种实施方式中,N的取值范围为 $4 \geq N \geq 0$ 。

[0168] 在一种实施方式中,所述投影点计算子单元310用于:

[0169] 利用相机坐标系和图像坐标系的变换关系,将所述第一车道线上的标注点在图像坐标系下的坐标转换为相机坐标系下的坐标;

[0170] 利用所述预测的相机自运动参数和所述第一车道线上的标注点在相机坐标系下的坐标,计算所述投影点在相机坐标系下的坐标;

[0171] 利用相机坐标系和图像坐标系的变换关系,将所述投影点在相机坐标系下的坐标转换为所述投影点在图像坐标系下的坐标。

[0172] 在一种实施方式中,所述距离计算子单元320用于利用以下公式计算所述损失函数的值:

$$[0173] \quad j = \arg \min(\sqrt{(x'_i - (x_j + x_{j+1})/2)^2 + (y'_i - (y_i + y_{j+1})/2)^2})$$

$$[0174] \quad x_{jm} = \frac{y'_i - y_{j+1}}{y_{j+1} - y_j} \times (x_{j+1} - x_j) + x_{j+1}$$

$$[0175] \quad \text{Loss} = \sum_i |x'_i - x_{jm}|$$

[0176] 其中,  $i$  和  $j$  表示车道线上的标注点的索引编号;  $(x'_i, y'_i)$  表示所述投影点的坐标;  $(x_j, y_j)$  表示所述第二车道线上的标注点的坐标;  $x_{jm}$  表示所述投影点在所述第二车道线上对应的点;  $\text{Loss}$  表示所述损失函数的值。

[0177] 在一种实施方式中, 所述训练单元 400 用于:

[0178] 利用随机梯度下降法对所述相机自运动参数估计模型进行训练, 通过计算所述损失函数对所述相机自运动参数估计模型进行反向传播, 更新所述相机自运动参数估计模型的网络参数;

[0179] 在训练到所述相机自运动参数估计模型收敛的情况下, 将所述相机自运动参数估计模型的网络参数固定。

[0180] 图 9 示出根据本发明实施例的相机自运动参数估计模型的训练装置的结构框图。如图 9 所示, 在一种实施方式中, 所述装置还包括估计单元 500, 所述估计单元 500 用于:

[0181] 利用训练好的所述相机自运动参数估计模型估计所述相机自运动参数。

[0182] 本发明实施例相机自运动参数估计模型的训练装置中的各单元的功能可以参见上述方法中的对应描述, 在此不再赘述。

[0183] 图 10 示出根据本发明实施例的相机自运动参数估计模型的训练装置的结构框图。如图 10 所示, 该装置包括: 存储器 910 和处理器 920, 存储器 910 内存储有可在处理器 920 上运行的计算机程序。所述处理器 920 执行所述计算机程序时实现上述实施例中的相机自运动参数估计模型的训练方法。所述存储器 910 和处理器 920 的数量可以为一个或多个。

[0184] 该装置还包括:

[0185] 通信接口 930, 用于与外界设备进行通信, 进行数据交互传输。

[0186] 存储器 910 可能包含高速 RAM 存储器, 也可能还包括非易失性存储器 (non-volatile memory), 例如至少一个磁盘存储器。

[0187] 如果存储器 910、处理器 920 和通信接口 930 独立实现, 则存储器 910、处理器 920 和通信接口 930 可以通过总线相互连接并完成相互间的通信。所述总线可以是工业标准体系结构 (ISA, Industry Standard Architecture) 总线、外部设备互连 (PCI, Peripheral Component Interconnect) 总线或扩展工业标准体系结构 (EISA, Extended Industry Standard Architecture) 总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示, 图 10 中仅用一条粗线表示, 但并不表示仅有一根总线或一种类型的总线。

[0188] 可选的, 在具体实现上, 如果存储器 910、处理器 920 及通信接口 930 集成在一块芯片上, 则存储器 910、处理器 920 及通信接口 930 可以通过内部接口完成相互间的通信。

[0189] 本发明实施例提供了一种相机自运动参数估计设备, 所述设备用于: 利用上述相机自运动参数估计模型的训练装置训练好的相机自运动参数估计模型估计所述相机自运动参数。

[0190] 在一种实施方式中, 所述相机自运动参数估计设备与所述相机自运动参数估计模型的训练装置可以是同一设备。利用该设备训练相机自运动参数估计模型, 然后利用训练好的所述相机自运动参数估计模型估计所述相机自运动参数。

[0191] 本发明实施例提供了一种相机自运动参数估计系统, 包括上述相机自运动参数估

计模型的训练装置和上述相机自运动参数估计设备。

[0192] 本发明实施例提供了一种计算机可读存储介质,其存储有计算机程序,该程序被处理器执行时实现上述实施例中任一所述的方法。

[0193] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。而且,描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外,在不相互矛盾的情况下,本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0194] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或隐含地包括至少一个该特征。在本发明的描述中,“多个”的含义是两个或两个以上,除非另有明确具体的限定。

[0195] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为,表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分,并且本发明的优选实施方式的范围包括另外的实现,其中可以不按所示出或讨论的顺序,包括根据所涉及的功能按基本同时的方式或按相反的顺序,来执行功能,这应被本发明的实施例所属技术领域的技术人员所理解。

[0196] 在流程图中表示或在此以其他方式描述的逻辑和/或步骤,例如,可以被认为是用于实现逻辑功能的可执行指令的定序列表,可以具体实现在任何计算机可读介质中,以供指令执行系统、装置或设备(如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统)使用,或结合这些指令执行系统、装置或设备而使用。就本说明书而言,“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。计算机可读介质的更具体的示例(非穷尽性列表)包括以下:具有一个或多个布线的电连接部(电子装置),便携式计算机盘盒(磁装置),随机存取存储器(RAM),只读存储器(ROM),可擦除可编程只读存储器(EPROM或闪速存储器),光纤装置,以及便携式只读存储器(CDROM)。另外,计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质,因为可以例如通过对纸或其他介质进行光学扫描,接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得所述程序,然后将其存储在计算机存储器中。

[0197] 应当理解,本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中,多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。例如,如果用硬件来实现,和在另一实施方式中一样,可用本领域公知的下列技术中的任一项或他们的组合来实现:具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路,具有合适的组合逻辑门电路的专用集成电路,可编程门阵列(PGA),现场可编程门阵列(FPGA)等。

[0198] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,该程序在执行时,包括方法实施例的步骤之一或其组合。

[0199] 此外,在本发明各个实施例中的各功能单元可以集成在一个处理模块中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读存储介质中。所述存储介质可以是只读存储器,磁盘或光盘等。

[0200] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到其各种变化或替换,这些都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以所述权利要求的保护范围为准。

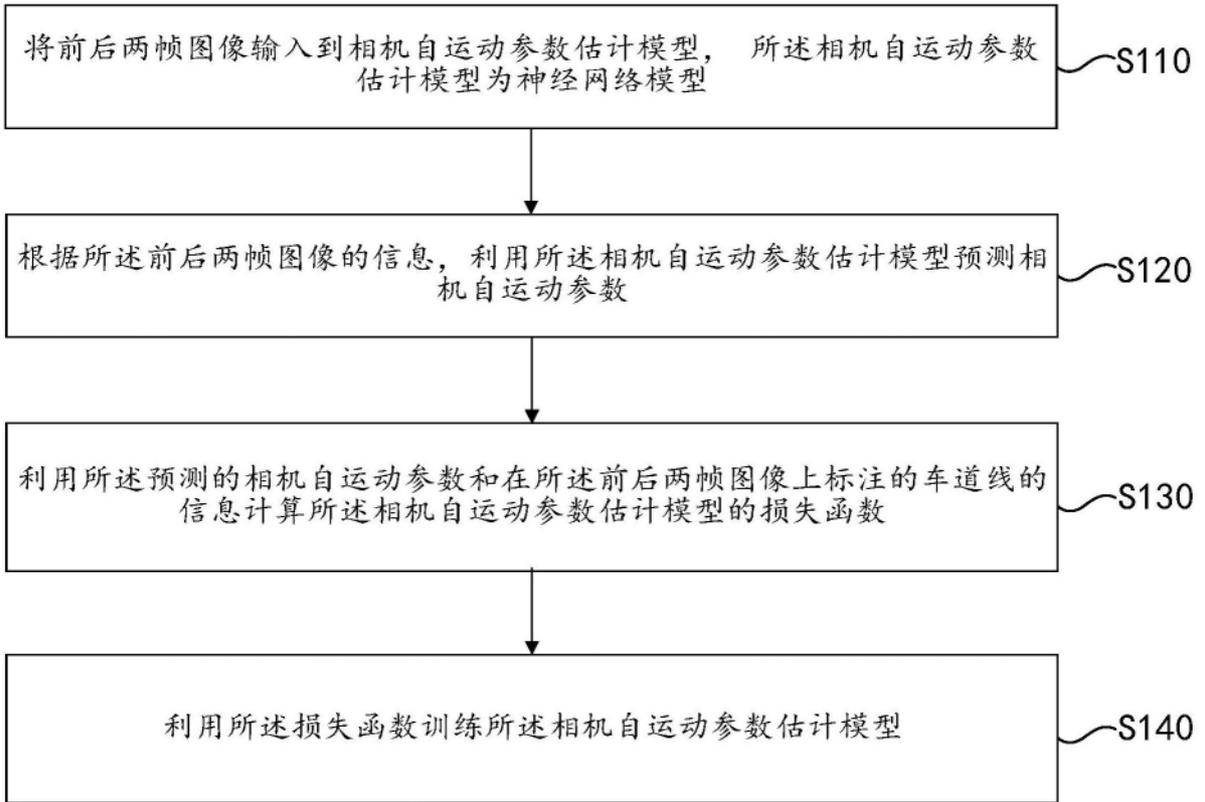


图1

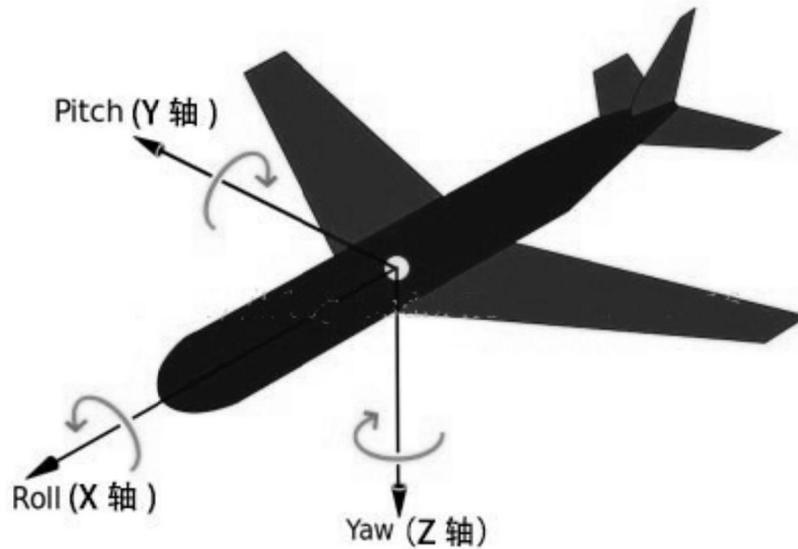


图2

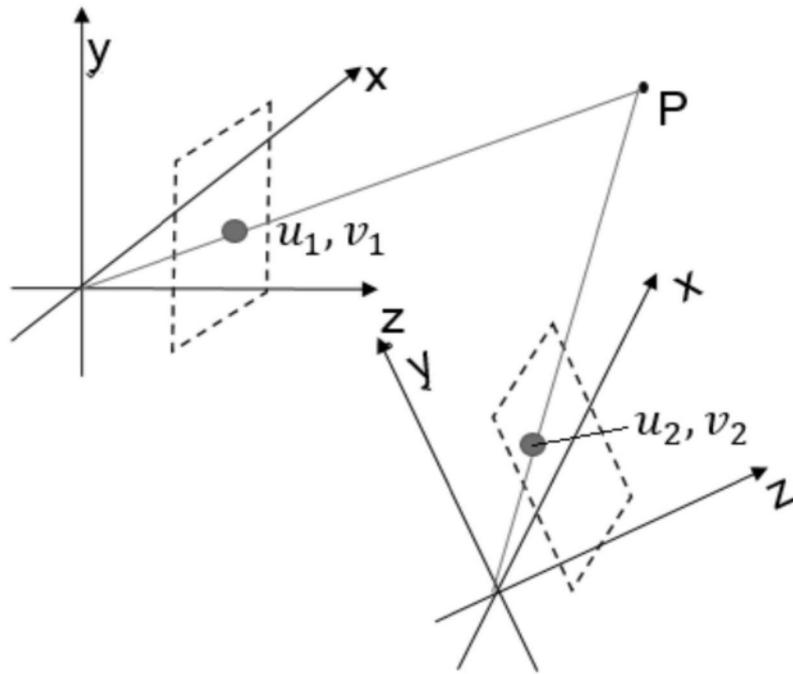


图3

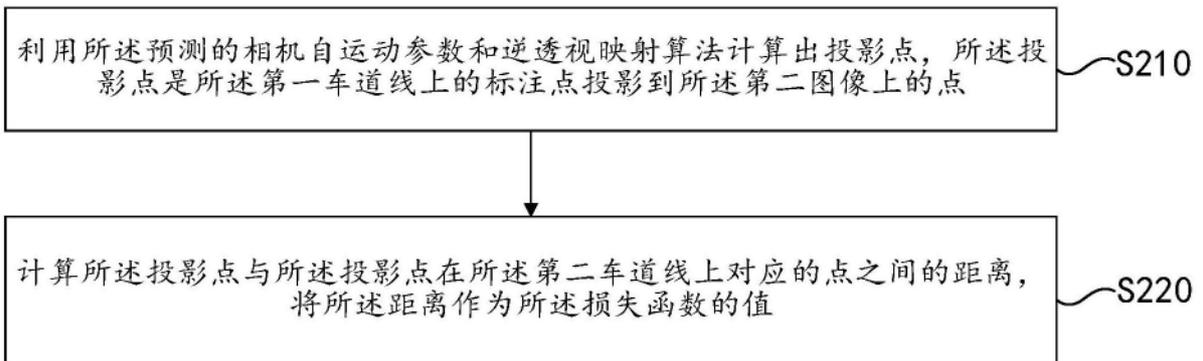


图4

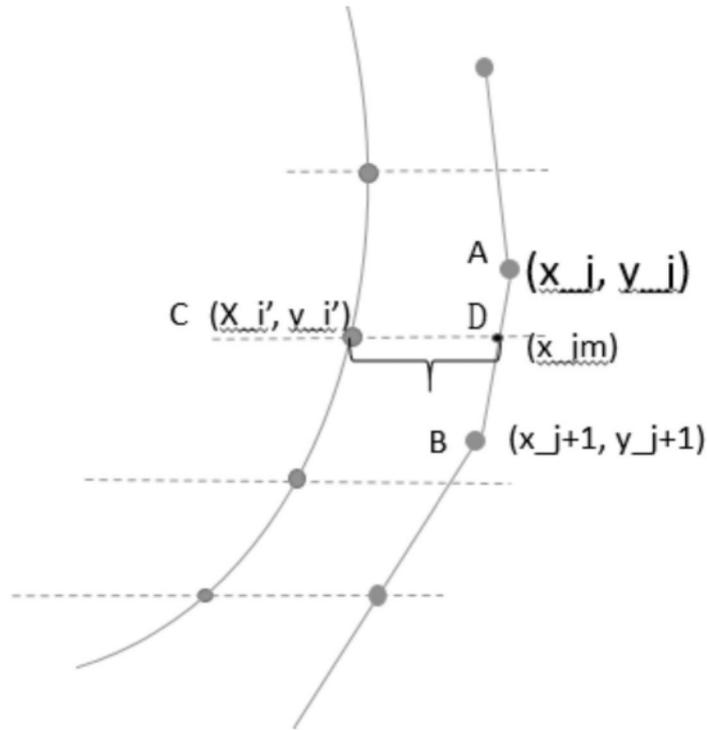


图5

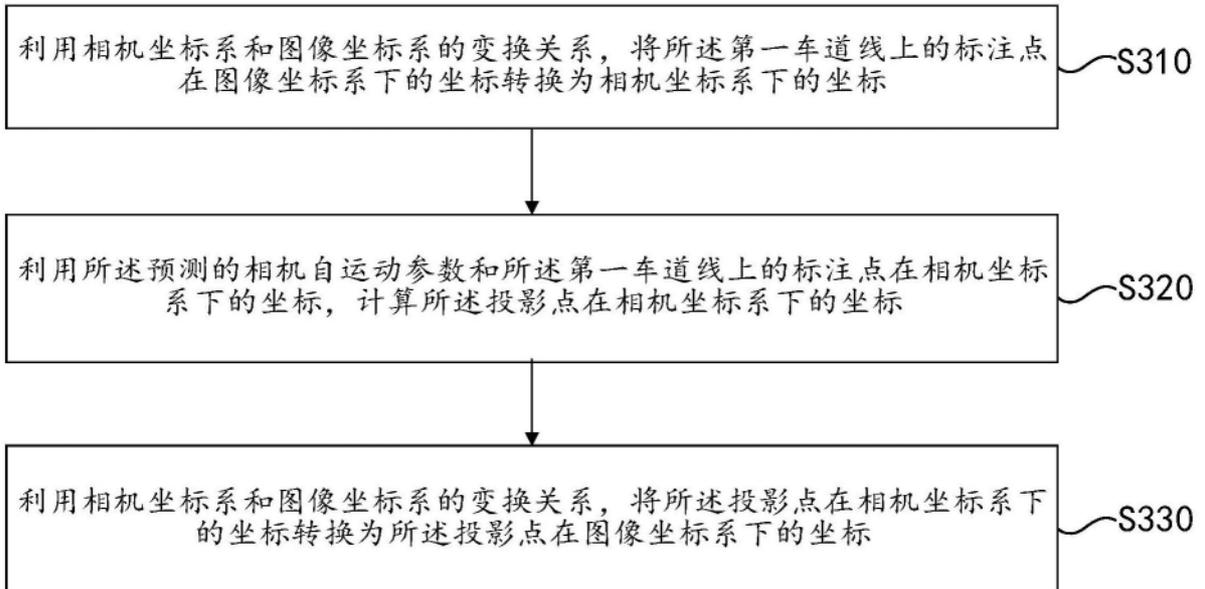


图6

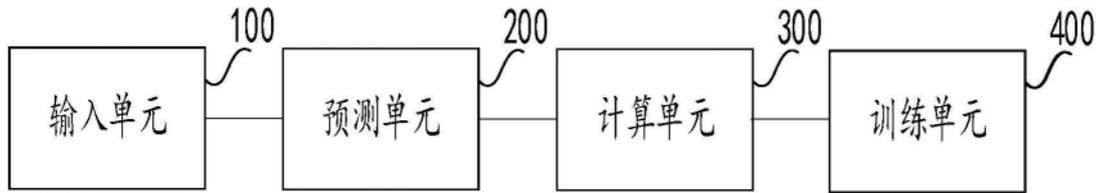


图7

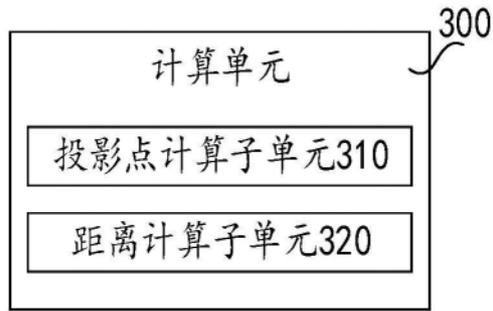


图8

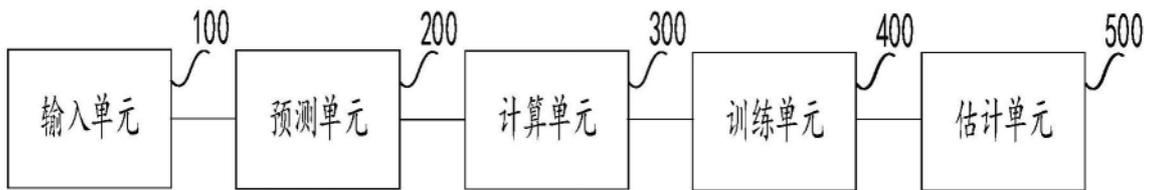


图9

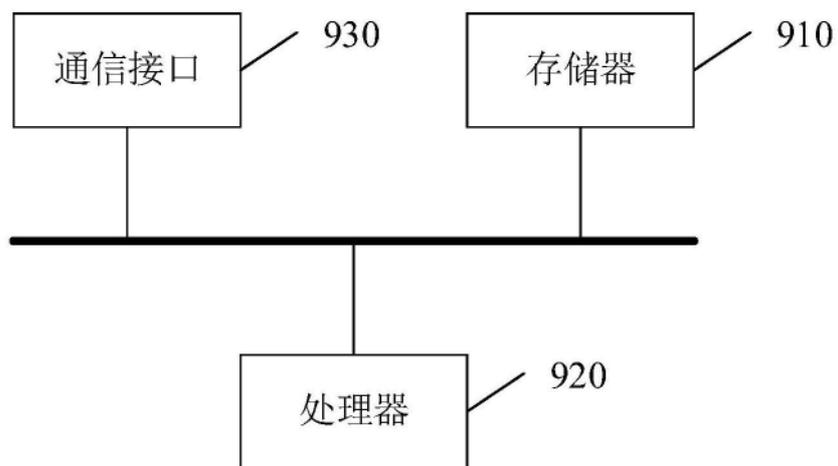


图10