

基于神经网络的标定算法

目录

- 当前标定现状
 - EOL标定
 - 售后标定
 - 标定指标要求
- why:
 - 为什么需要售后（在线）标定
 - 为什么需要探索深度学习的方法
- 综述文章解读
 - 深度学习方法分类
 - 相机内参标定
 - 相机外参标定
 - 相机内外参联合标定
 - 文章整理和初步解读
- CFT: Multi-camera Calibration Free Transformer
- CAMERA CALIBRATION THROUGH CAMERA PROJECTION LOSS
 - 总结概述:
 - 论文贡献:
 - 亮点工作 Separating sources of loss errors:
 - 训练细节
- 其他资料:
 - 内部标定总结及技术路线:
 - 理论知识复习:
 - 标定工具库:
 - 灭点外参标定:

当前标定现状

EOL标定

EOL标定

- 棋盘格：标定4鱼眼相机。
- 二维码+apriTag：标定bev周视相机
- 圆形标定板：标定长焦相机

售后标定

售后动态标定

- 标定方法：长直车道线
- 特斯拉售后标定方案：calibration completes quicker when driving on a straight road with multiple lanes (such as a controlled-access highway), with highly-visible lane markings
- https://service.tesla.com/docs/Public/diy/modely/en_us/GUID-D6F7D1BC-193D-4A2E-99B0-E3BA3D41BDE2.html

标定指标要求

使用方	成功率	精度	速度	泛化性
集度	99%	以百度智驾标准为准	3min, 且支持多路并行标定	希望适配大部分场景, 包括阴影, 路沿, 车道宽度, 部分不规则车道线等等。
ANP		bev模型要求车道线完全对齐, 几乎不允许误差。其他模块要求误差小于0.2度		
AVP		12.8m*12.8m范围内误差不超过一个线宽。侧向不超过0.1度		

why:

为什么需要售后（在线）标定

由于相机组件磨损、温度波动或外部干扰（如碰撞），相机的校准参数容易随时间变化。

为什么需要探索深度学习的方法

- 传统标定方法的问题：
 - 基于诸如棋盘格等标定靶标（不适用于售后标定）
 - 基于真实世界中的特定几何机构，例如车道线、灭点等（依赖特定环境鲁棒性差，且用户难以理解）
 - 基于序列图片采用多视几何理论（SFM）进行标定（在车辆载体下，前视相机帧间运动不充足）
- 上下游与泛化性角度：
 - BEV模型的输入仅是纠正后的图像，并没有对标定参数的显示需求
 - 深度学习方法可以实现不绑定车型，不依赖特定相机模型，不依赖标定靶，泛化性更高

深度学习方法的目标：achieve camera model-free and label-free calibration

综述文章解读

<https://github.com/KangLiao929/Awesome-Deep-Camera-Calibration>

<https://arxiv.org/pdf/2303.10559.pdf>

深度学习方法分类

学习范式角度：

- regression-based calibration
 - 使用卷积提取语义特征
 - 使用全链接层回归参数（例如：畸变、扭曲等）
- reconstruction-based calibration
 - 直接学习像素到像素的映射
 - 不需要相机模型的假设

学习策略角度：

- 监督学习
- 半监督
- 弱监督
- 无监督
- 强化学习

相机内参标定

Deepfocal是一个开创性的工作，旨在估计任意一张图像的焦距。MisCaliDet 给出了相机是否需要重新标定的判别。

相机外参标定

PoseNet 首先提出了6-Dof的标定工作

神经网络逐渐被引导去感知与几何相关的特征，这对于外方位估计至关重要。

相机内外参联合标定

DeepVP 单帧图片基于灭点进行标定

文章整理和初步解读

黄色高亮是初步判断更贴切的解决方案

1	Year	Publication	Title	Abbreviation	Objective	Platform	Network	初读备注
2	2022	PAMI	Content-Aware Unsupervised Deep Homography Estimation and Beyond	Liu et al.	Projection matrixs	PyTorch	ResNet	
3	2022	ECCV	Rethinking generic camera models for deep single image camera calibration to recover rotation and fisheye distortion	GenCaliNet	Intrinsics + Extrinsics + Distortion coefficients		DenseNet	<ul style="list-style-type: none">recovering camera rotationrecovering camera rotation and fisheye distortionproposed generic camera model that has an adaptivity for off-the-shelf fish cameras <div></div>
4	2022	TIM	Keypoint-Based LiDAR-Camera Online Calibration With Robust Geometric Network	RGKNet	Camera + LiDAR	PyTorch	CNNs+Point Net	设计激光标定 skip
5								
6	2022	CVPR	Focal Length and Object Pose Estimation via Render and Compare	FocalPose	Intrinsics + Extrinsics	PyTorch	CNNs	<ul style="list-style-type: none">通过可微分的方式推导和更新了焦距更新规则，提出了一种于6D姿态估计的方法主要研究3D目标向图像投影
7								
8	2022	ICASSP	Camera calibration through camera projection loss	CPL	Intrinsics + Extrinsics	TensorFlow	Inception-V3	<ul style="list-style-type: none">可以同时估计相机内参和外参，包括相机基线、视差、仰角、平移、焦距和主点位置。

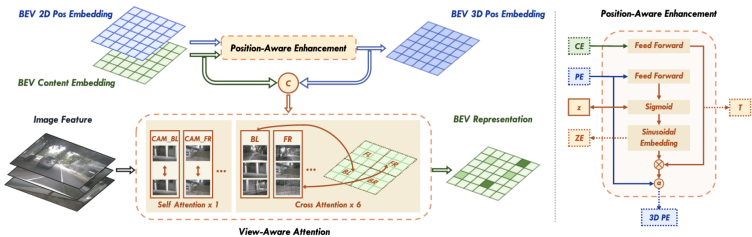
9	2022	ICRA	Self-supervised camera self-calibration from video	Fang et al.	Intrinsics + Extrinsics	PyTorch	CNNs	<ul style="list-style-type: none">使用自我监督的深度和姿态学习作为代理目标，从原始视频中自校准各种通用相机模型，从而首次提供对纯粹从自我监督中学习的相机模型参数的准确评估。在具有挑战性和完全不同的数据集上展示了我们的框架的通用性，无需改变架构即可处理透视、鱼眼和反射图像的深度和姿态。
10								
11	2021	ICCV	CTRL-C: Camera calibration TRansformer with Line-Classification	CTRL-C	Intrinsics + Extrinsics	PyTorch	Transformer	<ul style="list-style-type: none">输出相机的pitch、roll等（看到内参相关）通过line_segment检测实现图像几何的理解同时考虑语义信息和几何特征（line_segment） 
12								

CFT: Multi-camera Calibration Free Transformer

Learning BEV representation for 3D object detection without any geometric guidance.

主要贡献:

- We propose CFT to learn robust BEV representation for multi-camera 3D object detection **without camera parameters**.
- We design PA in obtaining richer 3D information of BEV, which could **fully mine height features from content embedding**, promoting the establishment of relationships between image views and BEV.



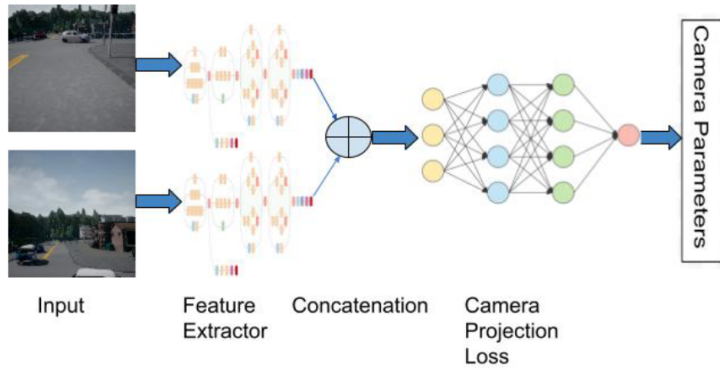
BEV 3D检测的各技术路线

- Depth-based method : infer depth in image views and project them to the BEV plane with the extrinsics and intrinsics
- BEVformer : emphasizes directly learning the transformation relationship between image view and BEV based on the attention mechanism.

CAMERA CALIBRATION THROUGH CAMERA PROJECTION LOSS

总结概述:

给出了一种双目相机标定的方法，神经网络主要用于特征提取工作，双目相机有可用深度，Loss设计遵循3D-2D投影来监督
介绍双目相机标定，其标定方法能否应用到bev相机？



论文贡献：

- 第一个同时完成内参、外参、基线、视差标定的端到端方法
- 使用CARLA制作了模拟数据集
- proposed a new representation that represents camera model equations as a neural network in a multi-task learning (MTL) framework.
- We proposed a novel camera projection loss (CPL) that combines analytical equations in learning framework.

亮点工作 Separating sources of loss errors:

问题点：然而，使用多个相机参数来预测3D点在学习过程中引入了一个新的问题：从一个点到其理想投影的偏差可归因于一个以上的参数。换句话说，一个参数的错误可以通过相机投影损失反向传播到其他参数。

为了解决这个问题，设计了各参数分离的loss计算方式

$$\begin{aligned}
 L_{f_x} &= L((f_x, f_y^{GT}, u_0^{GT}, v_0^{GT}, b^{GT}, d^{GT}, \theta_p^{GT}, \\
 &\quad t_x^{GT}, t_y^{GT}, t_z^{GT}, X^{GT}, Y^{GT}, Z^{GT}), \omega) \\
 L_{f_y} &= L((f_x^{GT}, f_y, u_0^{GT}, v_0^{GT}, b^{GT}, d^{GT}, \theta_p^{GT}, \\
 &\quad t_x^{GT}, t_y^{GT}, t_z^{GT}, X^{GT}, Y^{GT}, Z^{GT}), \omega) \\
 &\dots \\
 L_Z &= L((f_x^{GT}, f_y^{GT}, u_0^{GT}, v_0^{GT}, b^{GT}, d^{GT}, \theta_p^{GT}, \\
 &\quad t_x^{GT}, t_y^{GT}, t_z^{GT}, X^{GT}, Y^{GT}, Z), \omega) \\
 L^* &= \frac{L_{f_x} + L_{f_y} + L_{u_0} + \dots + L_Z}{13}
 \end{aligned}$$

训练细节

- GeForce GTX 1050 Ti GPU for 200 epochs
- ADAM optimizer with Mean Absolute Error (MAE) loss function
- a base learning rate η of 10^{-3} with a batch size of 16

其他资料：

内部标定总结及技术路线：



camera_online_calib_td.pdf (11MB)

500

理论知识复习：

<https://webthesis.biblio.polito.it/16698/1/tesi.pdf>

标定工具库：

<https://github.com/PJLab-ADG/SensorsCalibration>

<https://github.com/OpenCalib> （鱼眼相机：<https://arxiv.org/pdf/2305.16840.pdf>）

灭点外参标定：

<https://thomasfermi.github.io/Algorithms-for-Automated-Driving/CameraCalibration/VanishingPointCameraCalibration.html>

- 给出了标定参数对感知结果的影响分析方法
- 给出了从灭点计算俯仰角的公式

