

# 1. 项目技术方案或关键技术解决途径

## 基于单目相机和 3D 激光雷达的自动在线标定技术

三维光探测与测距(LiDAR)和相机的可靠实时外部参数是多模态感知系统的关键组成部分。然而,在操作过程中,外在的转化可能会逐渐漂移,从而导致感知系统的准确性降低。为了解决这一问题,课题组提出了一种基于线的方法,实现了激光雷达和摄像机在真实场景中的自动在线外部标定。首先,从点云和图像中提取和过滤直线特征。然后,利用自适应优化提供准确的外部参数。课题组证明了线特征是稳健的几何特征,可以从点云和图像中提取,从而有助于外部标定。为了证明该方法的优点,在 KITTI 基准上用地面真值对其进行了评价。实验验证了该标定方法的准确性。在数百帧的在线实验中,该方法自动校正了误差,达到了 0.2 度的精度,验证了该方法在各种场景下的适用性。该工作可为感知系统的研究提供依据,并可进一步提高利用这些传感器的其他算法的性能。

### 1.1 自动在线标定技术概述

多摄像机和多摄像机被广泛应用于自动驾驶车辆和机器人,在很多场景中,如自动驾驶,物体分类,分割,SLAM。激光雷达可以提供精确的三维几何信息,但具有稀疏点。相反,相机能够提供丰富的环境表征,但不太准确的距离信息。这些传感器的外部变换矩阵可以使这两个互补的传感器结合在一起,使效果最大化。因此,准确的外部参数对于感知系统提供准确的信息至关重要。

传统的手动标定方法需要特殊设计对象,如棋盘或手动选取点,导致标定过程繁琐,此外,长时间的运行和不同的负载会导致轻微的漂移和对外部参数的偏差。因此,需要一个自动在线算法来调整这些传感器运动。

目前的自动标定工作是利用互信息或人工设计的靶标来标定外部参数。特定目标和强度信息的使用将校准过程限制在实验室设置和特定传感器上。其他一些基于特征的标定方法利用边缘特征来计算外部参数。然而,在某些场景下,这些特性彼此之间并没有很好地对应。本节旨在消除这种局限,简化激光雷达和相机外部参数的标定过程。为此,我们利用鲁棒性和广泛分布的线特征自动计算这些传感器之间的转换矩阵,无需手动标记。作为输入,我们的方法需要一幅图像和之前的几个点云,以及初始的外部参数。在这个作品中使用的线条特征一般都是

散布在户外的场景中，如树木、路灯、汽车等。利用这些线特征，该方法可以根据外部变换矩阵的漂移和变化进行自动调整。

根据标定过程是否需要目标，激光雷达和相机的外部标定方法大致可分为两类：基于目标的标定方法和无目标标定方法。

### 1.1.1 靶向性方法

首先由 Q. Zhang 等人解决了激光雷达和相机之间的标定问题，他们利用校验板对二维激光雷达和相机的外部参数进行标定。该方法利用平面约束和非线性优化方法使重投影误差最小化。R. Unnikrishnan 等人发布了一个 MATLAB 工具包，用于手动选择点校准激光雷达和相机。D. Scaramuzza 等人描述了无人机的校准方法，即从两个传感器中选择几个对应点。S. Kato 等人提出了一种通过选定点来校准激光雷达和摄像机的开放方法。上述方法的主要局限性是需要手动在点云和图像之间分配相应的点，这很费力，尤其是在频繁校准的情况下。它还可能导致人为误差，因为选定的点可能彼此不准确对应。

为了克服上述限制，一些研究人员开发了几种更智能地校准外部参数的方法。他们采用特殊的人工校准目标，如三角板、多边形板、圆形板或球体，来代替人工选择相应点的过程。a. Kassir 等人开发了一种角点检测方法，以便提取棋盘并与点云对齐。Kim 等人还利用多个平面特征来校准激光雷达和相机。这些方法无需选择相应的点，但仍有特殊设计目标和人为目标误差的限制。因此，这些方法在真实环境中是不可行的，因为在传感器的视场中没有人工目标。

### 1.1.2 无目标方法

一些方法已经开发出来，通过利用现实世界的特征来解决基于目标的方法的局限性。例如基于最大化两个传感器测量的表面强度之间的互信息的无目标外部标定方法。这些方法确实释放了需要特定目标的限制。然而，激光雷达可能无法获得必要的强度信息，强度的变化会导致激光雷达与相机之间的变换矩阵的波动。

一些研究者发展了基于运动的方法来计算传感器的外部参数。他们提出了基于手眼标定的自动无目标标定方法，只需要三维点云和相机图像来计算运动信息。除此之外，反射信息和重叠视场是不必要的。然而，由于运动估计误差会影响标定参数的精度，这些方法只能提供粗略的标定结果。

Moghadam 等人提出了一种基于线的标定方法，通过估计平面表面来提取平

面相交线。然而，他们的方法仅限于包含平面相交特征的环境。J. Levinson 等人提出了利用激光雷达点与图像像素之间的边缘对应对齐进行标定的方法。为此，他们假设激光雷达具有较大深度不连续的点会比其他特征更经常地投射到边缘特征上。他们利用不连续点来提取边缘点，并通过对每个对应像素的求和来构造代价函数。最后，利用网格搜索算法对代价进行了优化。但是该方法在边缘聚类特征较多的场景中，如树多的场景中，可能会导致偏差。他们考虑每个波束独立，所以他们只得到垂直边缘特征，不能约束所有方向的误差。受到他们工作的启发，我们利用简洁的线特征和自适应优化方法来搜索最优方向和获得外部参数，使我们能够得到更准确的结果和更高的效率。

为解决上述问题，本算法的主要贡献如下:首先，介绍了一种新的在线自动估计六自由度参数的外部标定方法。该方法利用一般的直线特征来减少传感器间的漂移误差，不需要人工选择点和特殊目标，因此可以应用于任何给定的场景。其次，提出了一种从点云中提取直线特征的点云线提取方法，利用点云处理方法滤除噪声数据，准确提取直线特征。最后，我们引入自适应优化方法和结果置信度估计方法，使优化朝着正确的方向，并有效计算校准结果。

## 1.2 自动在线标定技术算法设计

选择直线特征来精确标定激光雷达与相机系统的外部参数，本方法采用了三个步骤。首先，对图像和点云采用一系列预处理方法进行进一步的特征提取。然后，从图像和点云中提取直线特征，通过特征过滤对直线特征进行细化。最后，通过在初始外部参数上添加小扰动，将点云直线特征投影到像素帧上，计算并优化每个扰动的分值。

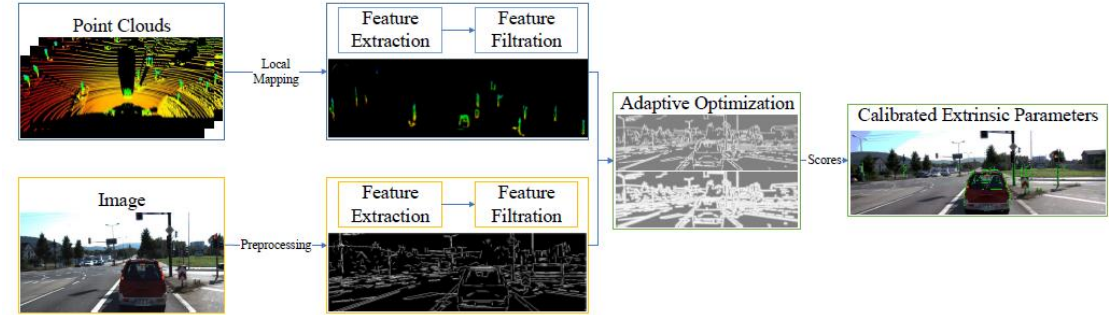


图 1 自动在线标定算法架构

### 1.2.1 问题的描述与假设

激光雷达和摄像机的外部标定问题在于确定正确的变换矩阵。将问题定义为求旋转角向量的 $\theta = (\theta_x, \theta_y, \theta_z)$ ，平移向量 $t = (t_x, t_y, t_z)$ 。对于点云和图像，我们将点云标记为 $P_t$ ，图像标记为 $T_{ij}$ 。我们通过将每个点投影到像素帧上，通过添加每个像素的灰度值计算当前参数的得分来优化 6 自由度参数。将激光雷达点投影到图像上，可以计算成本分数，目标函数定义为：

$$S_t = \sum_{i=n-w}^n I_{ij}^t \left[ \alpha \sum_{p_t \in F_h^t} T_i p_t + (1 - \alpha) \sum_{p_t \in F_v^t} T_i p_t \right] / w$$

每个激光雷达点在哪  $t$  迭代水平特征  $F_h$  和垂直特征  $F_v$  分别。模型的系数对水平和垂直的直线特征赋不同的权重。为了加强对水平误差的约束，将误差分配为 0.65。另外，考虑之前的  $w$  帧来计算帧  $t$  的分数。

这项工作的基本假设是，当校准参数正确时，点云的直线特征对图像的直线特征具有高度的响应性。除此之外，我们还假设摄像机和激光雷达的内在参数已经标定，激光雷达数据和摄像机图像同时采集。

### 1.2.2 数据预处理

在图像处理中，首先将 RGB 图像转换为灰度图像，然后通过<sup>[32]</sup>线检测算法提取直线特征。然后将距离变换模型应用于这些灰度图像。原始图像、直线特征和边缘特征的示例如图 2 所示。图 2(b)中的白边和图 2(c)中的白线分别代表聚类边缘特征和直线特征。如图 2(b)所示，应用距离变换模型后，聚类的边缘特征更加混乱。相反，图 2(c)中的直线特征组织得更好，产生的灰度变化更小。它可以允许更大的搜索步骤，从而防止优化过程进入一个局部解决方案。

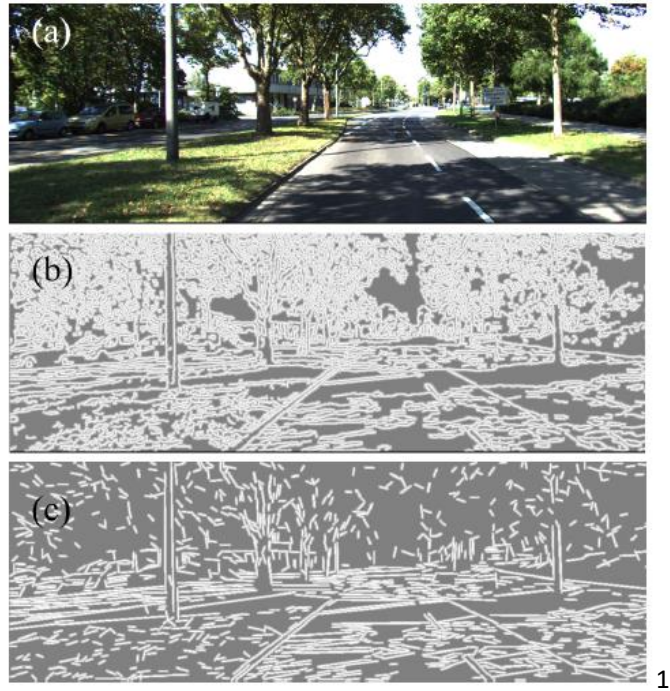


图 2 边缘特征与直线特征的对比

在激光雷达处理中，利用距离不连续性获取更多的边界直线特征是基本原理。为了实现这一目标，采用了局部映射的方法，将三帧的点云合并为一帧，可以在一帧中显示更多的点。利用正态分布变换(NDT)方法计算当前帧与前两帧之间的变换矩阵。单帧和三合一帧提取的边界线点比较如图 3 所示。图 3(a)通过转换三帧的点云  $P$  显示了一个更加密集的点云  $t-2:t$  到一个坐标系  $P_t$ ，与图 3(b)中的另一种方法相比，可以显示更多的点。这样可以提高提取性能，特别是在使用低波束激光雷达时。

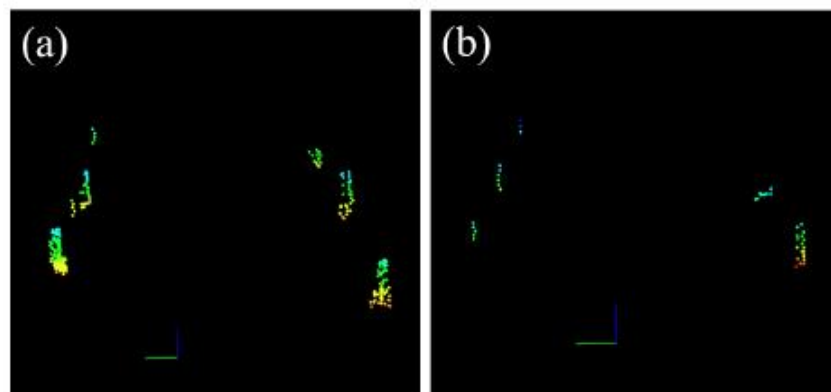


图 3 由三合一帧 (a) 和单帧 (b) 中图区的直线特征对比

然后将更加密集的点云转化为图像形式，每个像素存储对应的激光雷达点的

距离信息。通过比较当前点与相邻点的距离，去除离邻居太远的异常值，可以提取更精确的直线特征。需要注意的是，不同于单独考虑每束光，我们考虑了多束光之间的距离信息。它允许提出的方法提取水平特征，从而最小化水平和垂直误差使用直线特征。水平直线特征  $F_h$  和垂直线  $F_v$  特征分别存储在两个不同的点云中。在这种设置中，忽略了很少出现的平面交直线特征，有利于提高计算效率。

### 1.2.3 过滤特性

从点云中提取的线特征可以在上一节之后进行无序化处理。因此，我们采用两种过滤方法来消除异常值。由于点云已经被转换成图像形式，所以设计了一个卷积核来滤除距离所有 8 个相邻点超过一定阈值的点。过滤前后的线特征如图 4 所示。这种过滤方法可以去除所有的离群值以及与地面对应的点。因此，剩余的特征可以被识别为线特征。

首先过滤后，采用点云聚类算法用于去除相邻点较少的线特征。以上两个过滤步骤可以提供更有条理的点云线特征，保证后续步骤有更好的优化结果。在过滤掉小于 8 像素的线特征后，可以得到图像中更有组织的线特征。

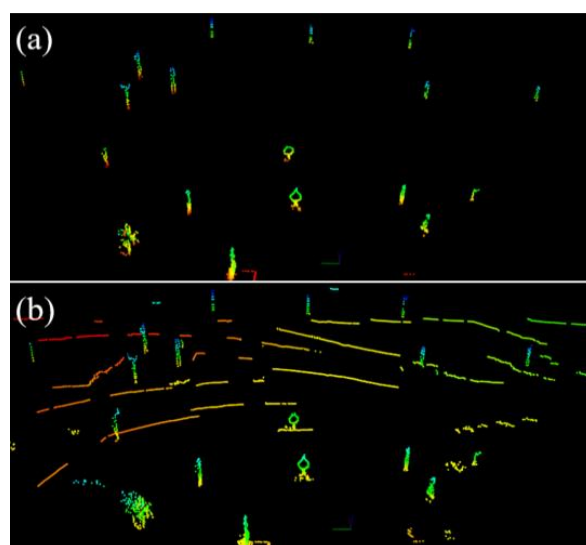


图 4 (a)特征过滤后和(b)特征过滤前的线特征比较



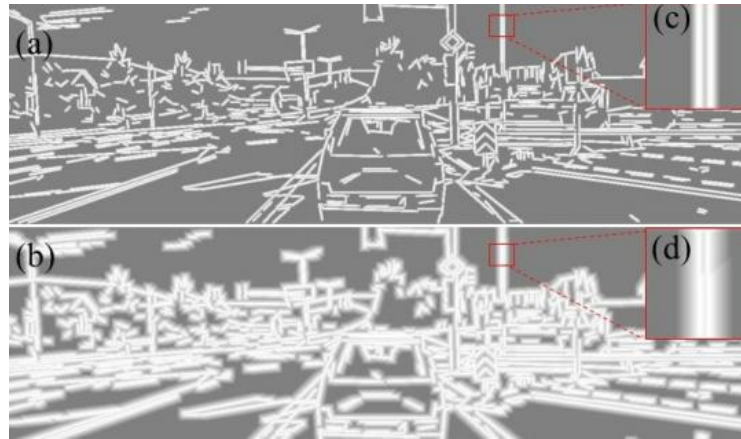


图 5 两种不同的灰度变化

#### 1.2.4 自适应优化

在计算精度方面，如图 5 所示（其中，(a)在精确搜索时显示出较大的灰度变化。(b)在粗搜索中显示出较小的灰度变化。放大视图(c)和(d)更清楚地展示了不同灰度变化的对比。），采用两个搜索步骤来精确求解。首先，为了防止搜索陷入局部解，采用更宽的图像线、更小的灰度变化、更大的步长进行粗搜索，以便快速发现可能包含最优解的区域。然后采用灰度变化较大的图像直线特征，步长较小的图像直线特征，得到更精确的标定结果。当投射到灰度区域的激光雷达点的比例超过一定阈值时，就会发生这两步长和灰度变化的切换。

为了提高计算效率，提出了一种自适应优化方法，使优化方向正确。采用搜索法对代价函数进行优化。它将当前分数与相邻的 728 分进行比较。在此过程中，如果搜索程序发现得分较高的参数，则会停止当前的搜索过程，并在得分较高的位置开始新的搜索过程。此外，当达到设定的迭代次数或找到最佳分数时，搜索过程将停止，从而可以提高计算效率。此外，通过滑动窗口设置优化过程中需要考虑的框架。利用三种框架来防止优化搜索方向错误或陷入局部最优解。因此，最终优化的外部参数应该超过滑动窗口中所有帧的其他参数。综上所述，这两种设置共同作用，可以得到一个可靠而精确的校准结果。终端策略可以实现更快的校准。这个过程可以在算法 1 中看到。

---

**Algorithm 1** Optimization process

---

**Require:** Image line features  $I_t$ , Horizontal  $F_h^t$  and vertical  $F_v^t$  line features at frame  $t$ , Initial extrinsic matrix  $T_t$ , last frame gray rate  $gray\_rate$

**Ensure:** Calibrated extrinsic matrix;

- 1: Initialization:  $score, max\_score \leftarrow 0$ .
- 2: **if**  $gray\_rate > \gamma$  **then**
- 3:      $step\_size = \alpha_1$
- 4: **else**
- 5:      $step\_size = \alpha_2$
- 6: **end if**
- 7: Add disturbance based on  $step\_size$
- 8: **for** each *LiDAR point* in  $F_h^t$  **do**
- 9:      $gray\_value = weight * T_t * p_t$
- 10:     $score += gray\_value$
- 11: **end for**
- 12: **for** each *LiDAR point* in  $F_v^t$  **do**
- 13:      $gray\_value = (1 - weight) * T_t * p_t$
- 14:     $score += gray\_value$
- 15: **end for**
- 16: **if**  $score > max\_score$  **then**
- 17:      $max\_score = score$
- 18:     Update current\_parameters
- 19: **end if**
- 20:  $gray\_rate = score / 255 / points\_num$
- 21: **return** current\_parameters

---

### 1.3 自动在线标定算法实验结果

为了验证所提出的方法，在 KITTI 数据集上进行了实验测试，课题组使用了 Velodyne HDL-64E 激光雷达和高分辨率彩色相机。激光雷达的扫描频率为 10hz。对实验数据进行了同步和校正，地面真实值的外部参数可以从校准文件中得到。

在 KITTI 数据集中对不同的场景进行了两个实验。实验结果如图 6 (a)和(b),和实验二的结果如图 6 所示(c)和(d)。在这两个实验中,设旋转偏差增加了  $x, y, z$  轴和一个 0.05 米的变换偏见到地面真理参数。然后，每 10 帧增加 0.5 度的旋转偏差。需要指出的是，一度旋转偏差是正还是负是随机的。在实验过程中，我们比较了校准误差与地面真相。此外，还测试了误校正的检测能力和校正偏差的速度。



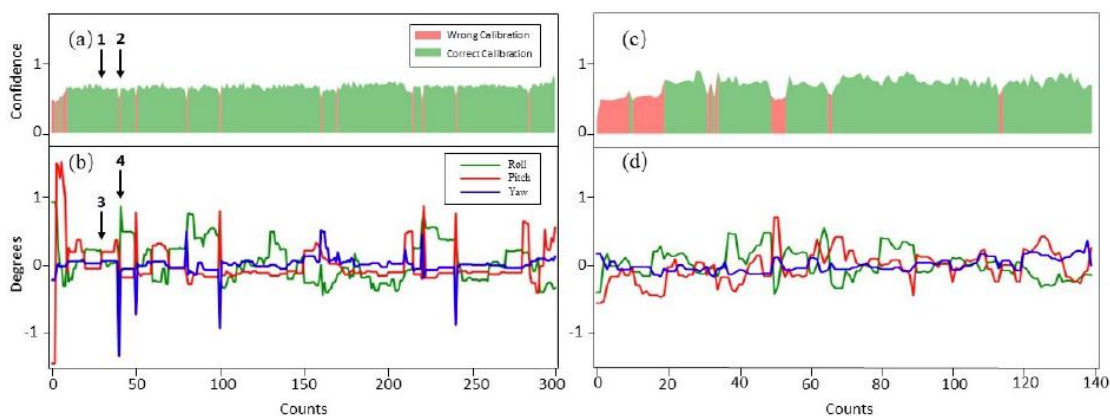


图 6 同时检测并修正横摇、俯仰、偏航的误差

为了评估该方法的精度，在两个实验中计算了校准结果与地面真实值的绝对偏差值。校准偏差的定量分析如图 6 所示。在图 6 (b)中，所提出的方法能够同时校正滚转平均误差为 **0.288** 度，俯仰平均误差为 **0.268** 度，偏航平均误差为 **0.108** 度。如果忽略 **0.7** 以下置信度，横摇、俯仰、偏航的误差将减小到 **0.217** 度、**0.209** 度、**0.082** 度。在图 6 (d)中，横摇、俯仰和偏航的平均误差分别为 **0.217**、**0.228** 和 **0.079** 度。如果忽略 **0.7** 以下置信度，横摇、俯仰和偏航的误差分别减小到 **0.159**、**0.164** 和 **0.070** 度。在不计算人工误差帧数的情况下，横摇、俯仰、偏航的最大误差始终在 **0.5** 度以内。由于激光雷达具有较高的水平分辨率，偏航的校准结果最为精确。尽管激光雷达在垂直方向上的分辨率较低，且该方向上的三维特征较少被呈现，但由于自适应优化算法和在该方向上较高的权重，我们的方法仍能获得较高的精度。总的来说，平均旋转误差 **0.12** 度低于大多数离线校准技术。

KITTI 数据集上更多场景的总体标定结果如图 7 所示，证明了点云和图像的直线特征是校正偏差的鲁棒性特征。人工添加的偏置可以在一帧或两帧内进行校正，比其他方法更快。此外，还说明了可以计算当前校准结果的置信度，并进一步提高了计算效率和精度。



图 7 在公开数据集下多个场景的标定结果