

## 3.1 自动驾驶感知与定位系统概述

自动驾驶三大关键技术分别是感知、决策和执行。感知与定位技术在自动驾驶中起着至关重要的作用,使汽车能够监控周围环境、探测障碍物并规划行驶路径。在感知阶段,车辆通过摄像机、雷达等传感装置感知行驶环境并根据所感知的环境做出操纵指令;在定位阶段,车辆通过高精地图、GPS、BDS 等静态数据,进行感知层面的多信源数据融合,规划合理的行驶路径。本章首先介绍自动驾驶传感器类型与工作原理,如多线激光雷达、组合惯导、相机、毫米波雷达、超声波雷达等,然后介绍常用的传感器数据融合方法与自动驾驶定位技术,如 GPS 定位、RTK 定位、多传感器融合定位和辅助定位等定位方法。

## 3.2 自动驾驶感知传感系统

自动驾驶感知传感系统主要包括视觉传感器、激光雷达传感器、毫米波雷达传感器、超声波雷达传感器等。自动驾驶感知传感系统常用传感器及特点如图 3-1 所示。

### 3.2.1 视觉传感器

#### 1. 摄像机

作为视觉传感器最常用的设备——摄像机,根据镜头和布置方式的不同主要可以分为单目摄像机、双目摄像机、三目摄像机和环视摄像机。

(1) 单目摄像机。如图 3-2 所示,单目摄像机模组只包含一个摄像机和一个镜头。由于很多图像算法的研究都是基于单目摄像机开发的,因此相对于其他类别的摄像机,单目摄像机的算法成熟度更高,但它有两个先天缺陷:一是其视野完全取决于镜头;二是单目测距精度低。



图 3-1 自动驾驶感知传感系统常用传感器及其特点

(2) 双目摄像机。为克服单目测距的缺陷，双目摄像机应运而生。由于相近的两个摄像机拍摄物体时，会得到同一物体在摄像机的成像平面的像素偏移量，根据像素偏移量、相机焦距和两个摄像机的物理距离等信息，经过数学计算即可得到物体的距离，如图 3-3 所示。双目摄像机能得到较高精度的测距结果，但是它与单目摄像机一样，视野完全依赖于镜头。而且双目测距摄像机对其两个镜头的安装位置、距离要求严格，所以此类相机的标定通常比较复杂。



图 3-2 单目摄像机



图 3-3 无人车双目摄像机

(3) 三目摄像机。由于单目和双目摄像机存在上述缺陷，目前广泛应用于无人驾驶的摄像机为三目摄像机。三目摄像机是三个不同焦距单目摄像机的组合，根据焦距不同每个摄像机所感知的范围也不尽相同，三目摄像机能较好地弥补感知范围不足的问题，因此在业界被广泛应用，如图 3-4 所示。

(4) 环视摄像机。上述三种摄像机所用的镜头都是非鱼眼的，环视摄像机的镜头是鱼镜头，而且安装位置是朝向地面的。当前部分车型上的“360°全景显示”功能，所用到的就是环视摄像机。安装在车辆前方、车辆左、右后视镜下和车辆后方的四个鱼镜头采集图像，如图 3-5 所示。环视摄像机可获取更大的视野，但代价是图像畸变严重，但也可以通过标定值进行图像的投影变换，可将图像还原成俯视图。



图 3-4 三目摄像机模组



图 3-5 环视摄像机采集图像

## 2. 摄像机的功能

摄像机在自动驾驶无人车上的应用,主要包括感知能力和定位能力两大类功能。

### 1) 感知能力

摄像机可以提供的感知能力如下:

(1) 车道线识别: 车道线是摄像机能够感知的最基本的信息,拥有车道线检测功能即可实现高速公路的车道保持功能。

(2) 障碍物识别: 自动驾驶中能识别多种障碍物,比如摩托车、卡车,甚至动物都是可以检测到。有了障碍物信息,无人车即可完成车道内的跟车行驶。

(3) 交通标志牌和地面标志(Traffic Sign and Road Sign)识别: 此类信息的识别,主要是作为道路特征与高精度地图做匹配后的辅助定位,也可以基于这些感知结果进行地图更新。

(4) 可通行空间识别: 可通行空间表示无人车可以正常行驶的区域,可以让车辆不再局限于车道内行驶,实现更多跨车道的超车功能等。

(5) 交通信号灯识别: 城区行驶的无人驾驶汽车,对交通信号灯状态的感知能力非常重要,这也是百度 Apollo 2.0 实现“简单路况自动驾驶”所必须开放的功能之一。

### 2) 定位能力

视觉同步定位与建图技术(Simultaneous Localization and Mapping, SLAM)是自动驾驶中常用的定位技术,采用预先建好的地图与实时感知结果进行匹配,从而确定当前位置。但由于 SLAM 地图容量过大,对存储设备容量要求很高,如何制作出足够轻量化的地图,成为 SLAM 技术商业化的关键。

## 3.2.2 激光雷达

### 1. 激光雷达简介

激光雷达是激光探测及测距(Light Detection and Ranging)系统的简称,也称 Laser Radar 或 LADAR(Laser Detection and Ranging)。其工作原理是向目标发射探测信号(激光束),然后将接收到的从目标反射回来的信号(目标回波)与发射信号进行对比分析,通过

相关算法处理后可获得目标的距离、方位、高度、速度、姿态甚至形状等参数,从而实现目标的探测、跟踪与识别。激光雷达传感器分为单线、多线(多者可达 64 线),相应的检测效果也逐步提升。目前每增加一线,成本大约上涨 1 万元。

## 2. 激光雷达测距原理

(1) 激光雷达三角测距: 激光雷达发射器先发射激光, 经过物体(Object)反射后被 CMOS(一种图像传感器, 即图 3-6 中 Imager)捕捉, 设捕捉点为  $X_2$ 。现过焦点  $O$  作一条虚线平行于入射光线, 与 Imager 相交于  $X_1$ , 由于  $\beta$  已知, 所以可得到  $X_1$  的位置, 即  $X_1$ 、 $X_2$  之间的距离为  $X$ , 易得左右两个三角形相似, 所以有  $q/f = s/x$ , 又有  $\sin\beta = q/d$ , 二者联立可得  $d = sf/(x \sin\beta)$ , 原理图如图 3-6 所示。

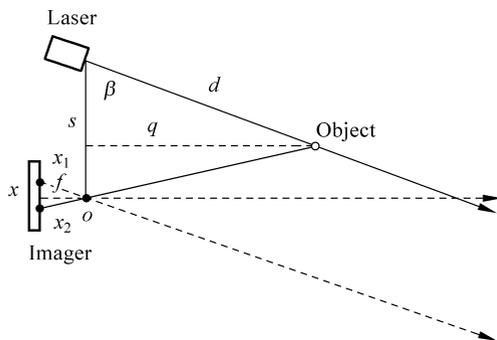


图 3-6 激光雷达三角测距原理

(2) TOF(Time of Flight, 飞行时差)测距: 传统的测距技术分为双向测距(Two Way Ranging)和单向测距(One Way Ranging)。TOF 测距主要利用信号在两个异步收发机(Transceiver)之间往返的飞行时间来测量节点间的距离。激光器发射一个激光脉冲, 并由计时器记录下出射的时间, 返回光经接收器接收, 并由计时器记录下回返的时间。两个时间相减即得到了光的“飞行时间( $t$ )”, 而光速  $c$  是一定的, 因此在已知速度和时间后很容易就可以计算出距离  $d = c \times t/2$ , 测距原理如图 3-7 所示。

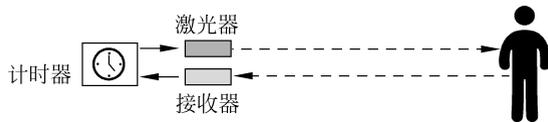


图 3-7 TOF 测距原理

## 3. 激光雷达扫描方式

(1) 传统的机械式旋转: 旋转 LiDAR(激光雷达)在地图领域的使用较为成熟, 但是由于价格极其昂贵, 目前量产的可能性较低。

(2) MEMS 微振镜: 将原本激光雷达的机械结构通过微电子技术集成到硅基芯片上, 本质上 MEMS 激光雷达是一种混合固态激光雷达, 并没有做到完全取消机械结构。

(3) 相控阵原理：通过调节发射阵列中每个发射单元的相位差来改变激光的出射角度，采用相控阵原理实现固态激光雷达。生活中最常见的干涉例子是水波，两处振动产生的水波相互叠加，有的方向两列波互相增强，有的方向正好抵消。与之类似，相控阵原理是采用多个光源组成阵列，通过控制各光源发射的时间差，合成角度灵活、精密可控的主光束。

(4) Flash LiDAR：Flash LiDAR 在短时间内发射出一大片覆盖探测区域的激光，通过高度灵敏的接收器进行接收，从而实现对周围环境图像的绘制。

#### 4. 激光雷达实例

知名的激光传感器供应商包括美国的 Velodyne 和 Quanergy，德国的 Ibeo 等，国内包括速腾（Robosense）聚创、北科天绘等公司。速腾激光雷达 RS-LiDAR-16（见图 3-8）是一种自动驾驶常用的激光雷达，其参数见表 3-1。



图 3-8 速腾激光雷达 RS-LiDAR-16

表 3-1 RS-LiDAR-16 激光雷达参数

类别	参数值	类别	参数值
线数	16 个	出点数	320000pts/s（单回波模式）
激光波长	905nm	出点数	640000pts/s（双回波模式）
激光安全等级	级别 1 人眼安全	以太网输出	100Mb/s
测距能力	200m	输出数据协议	UDP
盲区	$\leq 0.4\text{m}$	UDP 数据包内容	三维空间坐标、反射强度、时间戳等
精度（典型值）	$\pm 2\text{cm}$	工作电压	9~36V
水平视场角	$360^\circ$	产品功率	9W
垂直视场角	$30^\circ$	重量（不包含数据线）	840g
水平角分辨率	$0.09^\circ \sim 0.36^\circ$	尺寸	$\phi 109\text{mm} \times H 80.7\text{mm}$
垂直角分辨率	$\pm 2^\circ$	工作温度	$-40^\circ\text{C} \sim +60^\circ\text{C}$
帧率	5/10/20Hz	存储温度	$-40^\circ\text{C} \sim +105^\circ\text{C}$
转速	300/600/1200rpm（5/10/20Hz）	防护等级	IP67

RS-LiDAR-16 采用传统的机械旋转方式，通过 16 个激光发射组件快速旋转的同时发射高频率激光束对外界环境进行持续性的扫描，经过 TOF 测距算法提供三维空间点云数据及物体反射率，可以让机器看到周围的世界，为定位、导航、避障等提供有力的保障。自动驾驶激光雷达扫描如图 3-9(a)和图 3-9(b)所示。

激光雷达传感器相比于其他自动驾驶传感器，其优点在于可以生成三维的位置信息，快速确定物体的位置、大小、外貌和材质，同时还能形成精确的数字模型。相比摄像头等传感

器,其探测距离更远、精确度更高、响应速度更灵敏,且不受环境光的影响,但其最大的缺陷是不能在雨雪等极端天气下工作,天气会影响激光雷达的效果,影响三维地图的构建,导致无人车的“位置晕眩”。此外,昂贵的价格也是激光雷达难以像超声波雷达那样在民用汽车领域普及的重要原因。目前一个激光雷达的价格就有可能超过普通小汽车整车价格,现阶段尚没有大规模量产的可能。图 3-10 和图 3-11 为激光雷达扫描形成的点云图像。

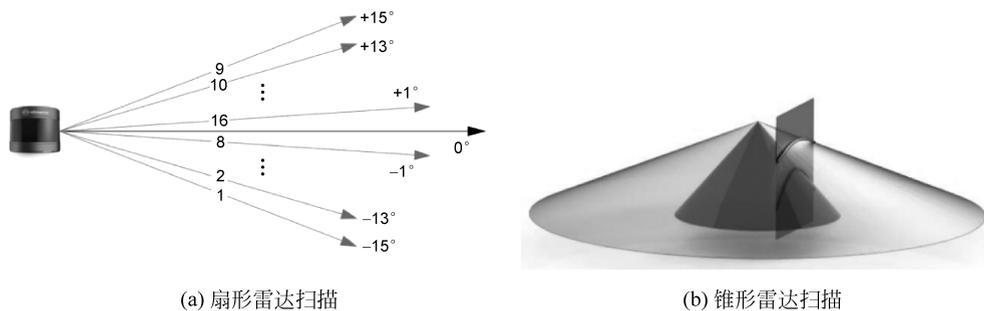


图 3-9 自动驾驶激光雷达扫描

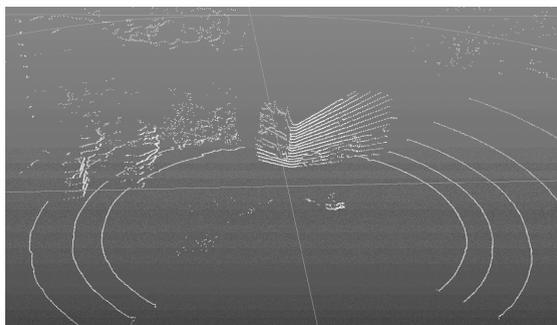


图 3-10 激光雷达扫描点云图像

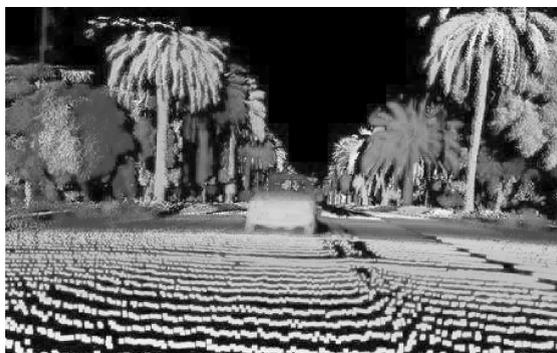


图 3-11 激光雷达成像三维视图

### 3.2.3 毫米波雷达

汽车毫米波雷达是自动驾驶辅助系统(ADAS)的核心部件之一,工作频段为 21.65~26.65GHz 和 76~81GHz。比较常见的汽车毫米波雷达工作频率在 24GHz、77GHz、79GHz 这 3 个频率附近。毫米波雷达具有远距离探测、全天候工作、车速测量、温度稳定性强等能力;不受天气和光线条件影响,在雨雪、烟雾等恶劣环境下优势显著。选择性价比更高的毫米波雷达作为测距和测速的传感器,可以解决摄像机测距、测速不够精确的问题。毫米波雷达不仅成本适中,而且能够较好地处理沙尘天气下的成像问题,是激光雷达替代不了的。图 3-12 为 Continental ARS-408 毫米波雷达。



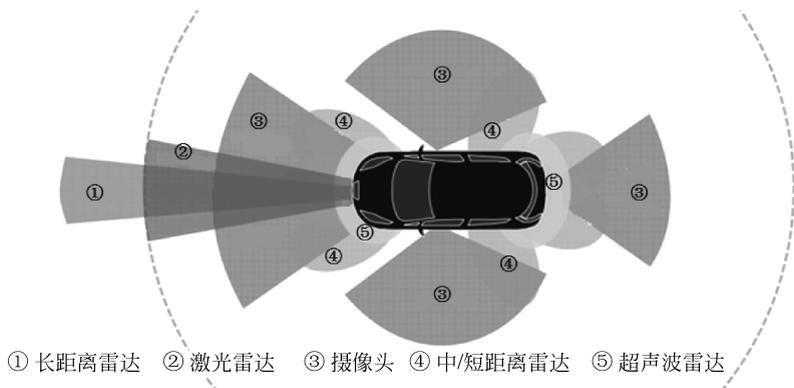
图 3-12 Continental ARS-408 毫米波雷达

#### 1. 毫米波雷达的分类

频率在 10~200GHz 的电磁波因其波长在毫米量级,故称为毫米波。应用在自动驾驶领域的毫米波雷达主要有 3 个频段,分别是 24GHz、77GHz 和 79GHz。不同频段的毫米波雷达性能和成本不同。

(1) 短距离雷达: 24GHz 频段。

处在该频段上的雷达的检测距离有限,因此常用于检测近处的障碍物(车辆)。图 3-13 中的标注为④的雷达,能够实现高级驾驶辅助系统(Advanced Driving Assistance System, ADAS)功能有盲点检测、变道辅助等;在自动驾驶系统中常用于感知车辆近处的障碍物,为换道决策提供感知信息。



① 长距离雷达 ② 激光雷达 ③ 摄像头 ④ 中/短距离雷达 ⑤ 超声波雷达

图 3-13 毫米波雷达和激光雷达自动驾驶方案

(2) 中长距离雷达: 77GHz 频段。

如图 3-13 所示,车辆正前方被标注为①的长距离雷达,即为频段在 77GHz 左右的雷达,其最大检测距离可以达到 160m 以上,因此常被安装在前保险杠上正对汽车的行驶方

向,如特斯拉 AutoPilot 2.0 中所配备的长距离毫米波雷达,能够用于实现紧急制动、高速公路跟车等 ADAS 功能;同时也能满足自动驾驶领域对障碍物距离、速度和角度的测量需求。

### (3) 长距离雷达: 79GHz 频段。

该频段的传感器能够实现的功能和 77GHz 类似,用于更长距离的测量。根据公式: 光速=波长×频率,频率更高的毫米波雷达,其波长越短,意味着分辨率越高;而分辨率越高,意味着在距离、速度、角度上的测量精度更高,因此 79GHz 的毫米波雷达必然是未来的发展趋势。

毫米波雷达相比于激光雷达有更强的穿透性,能够轻松地穿透保险杠上的塑料,因此常被安装在汽车的保险杠内,这也是为什么很多具备 ACC(自适应巡航)的车上装有毫米波雷达,却很难从外观上发现它们的原因。

## 2. 毫米波雷达的挑战

毫米波雷达在自动驾驶领域的实际开发过程中,主要面临以下挑战:

(1) 数据稳定性差: 数据的不稳定性对后续的软件算法提出了较高的要求。

(2) 对金属敏感: 由于毫米波雷达发出的电磁波对金属极为敏感,在实际测试过程中会发现近处路面上突然出现的钉子、远距离外的金属广告牌都会被认为是障碍物。一旦车辆高速行驶,被这些突然跳出的障碍物干扰时,会刹车不断,从而导致汽车的舒适性下降。

(3) 高度信息缺失: 毫米波雷达的数据只能提供距离和角度信息,不能像激光雷达那样能提供高度信息,没有高度信息的障碍物会给技术开发带来很多挑战。

## 3.2.4 超声波雷达

### 1. 超声波雷达工作原理

超声波雷达是一种利用超声波定位的雷达,是一款极其常见的传感器。它是汽车倒车时的安全辅助装置,能以声音或者更为直观的显示器告知驾驶员周围障碍物的情况,解除了驾驶员驻车、倒车和起动车辆时探视障碍物所引起的困扰,并帮助驾驶员扫除了视野死角和视线模糊的缺陷。

超声波雷达的工作原理是通过超声波发射装置向外发出超声波和接收器接收到超声波的时间差来测算距离。目前,常用探头的工作频率有 40kHz、48kHz 和 58kHz 3 种。一般来说,频率越高灵敏度越高,但水平与垂直方向的探测角度也就越小,故一般采用 40kHz 的探头。超声波雷达防水、防尘,即使有少量的泥沙遮挡也不影响。探测范围在 0.1~3m,而且精度较高,因此非常适用于泊车。其工作原理如图 3-14 所示。

车载的超声波雷达一般安装在汽车的保险杠上方,隐藏在保险杠的某个位置,如新石器自动驾驶无人车采取的超声波雷达共有 14 个,如图 3-15 所示。超声波雷达探头如图 3-16 所示。

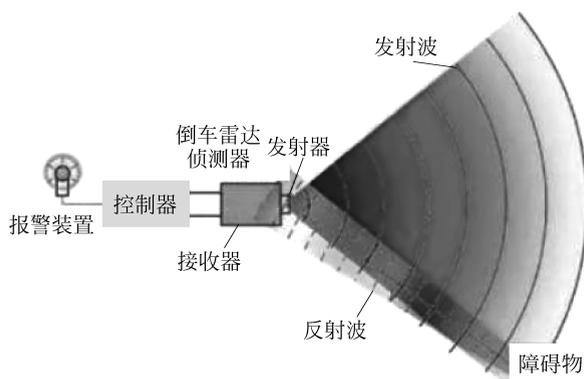


图 3-14 超声波雷达传感器工作原理



图 3-15 新石器自动驾驶无人车超声波雷达



图 3-16 超声波雷达探头

## 2. 超声波雷达的类型

常见的超声波雷达有两种：一种是安装在汽车前后保险杠上方的，用于测量汽车前后障碍物的倒车雷达，称为 UPA (Ultrasonic Parking Assistant, 超声波驻车辅助)；另一种是安装在汽车侧面的，用于测量侧方障碍物距离的超声波雷达，称为 APA (Automatic Parking Assistant, 自动泊车辅助)。

(1) UPA 超声波雷达。UPA 超声波雷达的探测距离一般在 15~250cm，主要用于测量汽车前后方的障碍物。

(2) APA 超声波雷达。APA 超声波雷达的探测距离一般在 30~500cm，APA 的探测范围更远，因此成本更高，功率也更大。

目前大部分车型搭载的超声波雷达都为倒车雷达 UPA，而随着近年来自动驾驶概念的兴起，基于超声波的自动泊车功能，逐渐进入大众视野，APA 的市场也逐渐打开，如图 3-17 和图 3-18 所示。

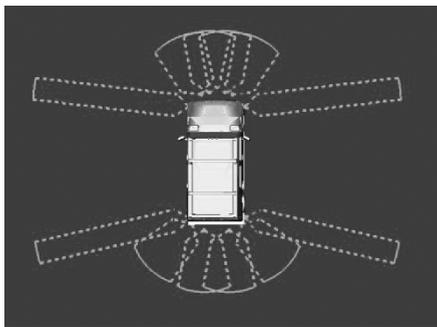


图 3-17 超声波雷达探测区域

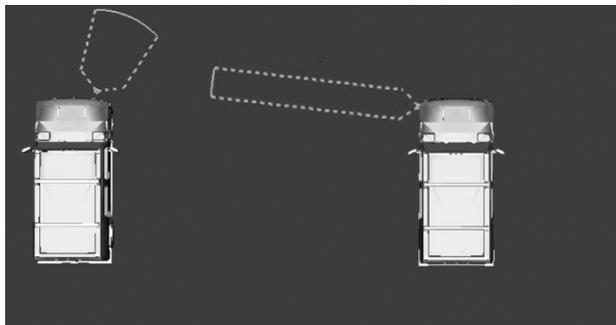


图 3-18 UPA(左)和单个 APA(右)超声波探测区域

### 3. 超声波雷达的特点

(1) 温度敏感。超声波雷达的测距原理和之前介绍的激光雷达、毫米波雷达类似,即距离=传播速度 $\times$ 传播时间/2。不同的是激光雷达和毫米波雷达的波速都为光速,而超声波雷达的波速跟温度有关。近似关系如下:

$$C = C_0 + 0.607 \times T \quad (3-1)$$

其中, $C_0$ 为零度时的声波速度 332m/s, $T$ 为温度(单位:°C)。例如温度在 0°C时,超声波的传播速度为 332m/s;温度在 30°C时,超声波的传播速度为 350m/s。对传感器精度要求极高的自动驾驶系统来说,选择将超声波雷达的测距进行保守计算,并将温度信息引入自动驾驶系统,以提升测量精度。超声波的能量消耗较缓慢,在介质中传播的距离比较远,穿透性强,测距的方法简单,成本低。

(2) 无法精确描述障碍物位置。超声波雷达在速度很高情况下测量距离有一定的局限性,这是因为超声波的传输速度很容易受天气情况的影响,在不同的天气情况下,超声波的传输速度不同,而且传播速度较慢。当汽车高速行驶时,使用超声波测距无法跟上汽车的车距实时变化,误差较大。另一方面,超声波散射角大、方向性较差,在测量较远距离的目标时回波信号会比较弱,影响测量精度。但是在短距离测量中,超声波测距传感器具有很大优势。

## 3.3 传感器的标定

### 3.3.1 IMU 的标定

IMU 中文叫作惯性测量单元,它能够获取自身的加速度、角速度信息,有的 IMU 还能够获得地磁量。在对 IMU 标定介绍之前,需要先选定车辆坐标系,如图 3-19 所示。

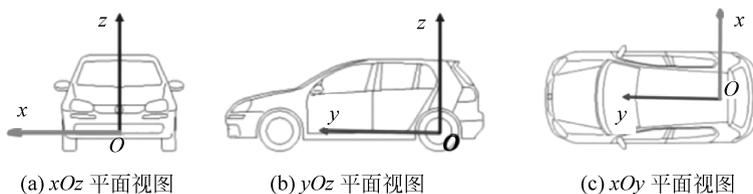


图 3-19 车辆坐标系的定义

车辆坐标系的定义为:

$z$  轴——通过车顶垂直于地面指向上方;

$y$  轴——在行驶的方向上指向车辆前方;

$x$  轴——面向前方时,指向车辆右侧。

车辆坐标系的原点在车辆后轮轴的中心。

IMU 传感器的标定主要是标定 IMU 和车体坐标系的外参,由于两个坐标系位置重合,所以外参指的是绕坐标轴 3 个方向的角度偏移量,从算法的角度考虑,将车体的中心转移到附近的 IMU 所在的位置。

IMU 标定时,选择一条笔直的道路,让车辆沿着道路直行一段距离,采集一段 IMU 的数据,通过分析 IMU 的输出轨迹和车辆沿  $Y$  轴前进的轨迹,计算 IMU 在车上的外参。

### 3.3.2 相机的标定

相机的标定主要包括内参和外参标定。相机的内参标定包括相机的焦距、主点位置、畸变系数等参数。相机的外参标定,主要包括相机坐标系和世界坐标系的旋转矩阵和平移向量。下面简要介绍相机成像原理。

#### 1) 世界坐标系

世界坐标系(World Coordinate)也称为测量坐标系,是一个三维直角坐标系,以其为基准可以描述相机和待测物体的空间位置,世界坐标系的位置可以根据实际情况自由确定。

#### 2) 相机坐标系

相机坐标系(Camera Coordinate)也是一个三维直角坐标系,原点位于镜头光心处, $x$ 、 $y$  轴分别与像平面的两边平行, $z$  轴为镜头光轴,与像平面垂直。

## 3) 相机坐标系转换为世界坐标系

相机坐标系转换为世界坐标系的转换方程为：

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3-2)$$

其中  $\mathbf{R}$  为  $3 \times 3$  的旋转矩阵,  $\mathbf{t}$  为  $3 \times 1$  的平移矢量, 等式右边为相机坐标系的齐次坐标, 左边为世界坐标系的齐次坐标。

## 4) 像素坐标系、图像坐标系

如图 3-20 所示, 像素坐标系是一个二维直角坐标系, 反映了相机 CCD/CMOS 芯片中像素的排列情况。原点位于图像的左上角,  $u$  轴、 $v$  轴分别与像面的两边平行。像素坐标系中坐标轴的单位是像素(整数)。

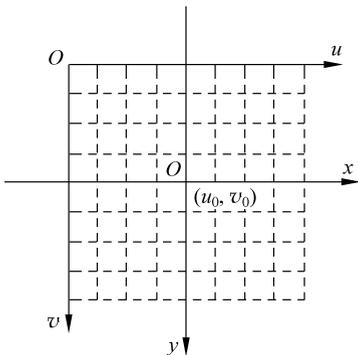


图 3-20 像素坐标系示意图

像素坐标系不利于坐标变换, 因此需要建立图像坐标系, 其坐标轴的单位通常为毫米 (mm), 原点是相机光轴与像平面的交点(称为主点), 即图像的中心点,  $x$  轴、 $y$  轴分别与  $u$  轴、 $v$  轴平行, 故两个坐标系实际是平移关系, 即通过平移就可得到。常用的相机标定方法, 可以选择 MATLAB 上面的标定工具箱, 或者开源的相机标定算法, 主要是利用棋盘格进行标定, 算法采用的是张正友提出的标定算法, 感兴趣的同学可以查阅相关论文。

### 3.3.3 激光雷达和组合惯导标定

在自动驾驶领域, 多线激光雷达和组合惯导是高精度地图制作、激光点云定位以及点云物体检测任务中常用的传感器配置。因此, 标定两者之间的外参数具有非常重要的意义。在许多应用中, 激光雷达被水平放置于车顶以获得  $360^\circ$  的感知范围, 而组合惯导(惯性导航)也常与车的对称轴以及水平面对齐以简化坐标系的定义。

激光雷达和组合惯导的标定可以通过手、眼标定获得外参初始值, 由于多线激光雷达与

组合惯导两者的测量之间没有直接的对应,因此需要使用手眼标定的方法对外参的初值进行求解。假设在  $t_i$  时刻组合惯导的位置姿态为  $T_i^{\text{ins}}$ ,使用激光雷达里程计或者 SLAM 技术得到多线激光雷达的位置为  $T_i^{\text{lidar}}$ 。则经典的手、眼标定问题为求解  $T_{\text{lidar}}^{\text{ins}}$ ,使得

$$T_{i,i+1}^{\text{ins}} T_{\text{lidar}}^{\text{ins}} = T_{\text{lidar}}^{\text{ins}} T_{i,i+1}^{\text{lidar}} \quad (3-3)$$

其中,  $T_{i,i+1}^{\text{ins}} = T_{i+1}^{\text{ins}} T_i^{\text{ins}} - 1$ ,  $T_{i,i+1}^{\text{lidar}} = T_{i+1}^{\text{lidar}} T_i^{\text{lidar}} - 1$  是两个传感器的相对运动。由于车辆在近似平面内运动,将问题简化为二维的手、眼标定问题,即

$$(\mathbf{R}_{\text{ins}} - \mathbf{I})\mathbf{t} = \mathbf{R}\mathbf{t}_{\text{lidar}} - \mathbf{t}_{\text{ins}} \quad (3-4)$$

其中,  $\mathbf{R}_{\text{ins}}$  和  $\mathbf{t}_{\text{ins}}$  分别是组合惯导相对运动的旋转和平移部分,  $\mathbf{t}_{\text{lidar}}$  是激光雷达相对运动的平移部分,  $\mathbf{R}$  和  $\mathbf{t}$  是外参的旋转和平移矢量。

令

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (3-5)$$

则有

$$\left( \mathbf{R}_{\text{ins}} - \mathbf{I} \begin{bmatrix} -t_{\text{lidar}}^x & t_{\text{lidar}}^y \\ -t_{\text{lidar}}^y & -t_{\text{lidar}}^x \end{bmatrix} \right) \begin{bmatrix} t^x \\ t^y \\ \cos\theta \\ \sin\theta \end{bmatrix} = -\mathbf{t}_{\text{ins}} \quad (3-6)$$

因此,一次相对运动能构造两个约束条件,当有 3 个以上不同位置朝向的运动时,方程满秩,可线性求解。为了保证初值求解以及优化过程中对外参构成足够的约束,要求车辆以 8 字形状的轨迹行驶,如图 3-21 所示。

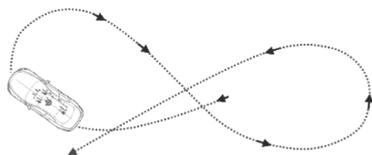


图 3-21 标定车辆行驶轨迹

通过组合惯导的位置姿态信息,结合外参初值,可以对激光雷达采集的点云进行拼接。由于初值求解误差的存在,不同位置采集的点云拼接后存在对齐误差,具体表现为拼接点云视觉效果模糊、场景中边缘结构不锐利。为此,利用优化拼接的方法对外参进行优化,使用经典的 GICP 以及 Entropy 的代价函数即可,标定过程如图 3-22 所示。

### 3.3.4 相机和激光雷达标定

相机和激光雷达的联合标定,互相弥补了两种传感器各自的不足,将两者的感知作用发挥到极致。相机和激光雷达联合标定的基本原理:通常是选择一组参照物,可以使用棋盘格标定板,或者是自然场景,然后获取这组参照物在相机和激光雷达各自坐标系下的位置和

姿态信息,反向计算出相机和激光雷达的外参矩阵。标定分为在线标定和离线标定:离线标定通常选择棋盘格作为参照物进行标定,因为标定过程中,需要对棋盘格上的角点进行人为的选取,这期间会用到第三方的软件进行点云的提取操作,耗时较长,精度也无法保证;在线标定是指车辆在采集完数据后,自动对数据进行处理,没有人为干预,输出外参结果。

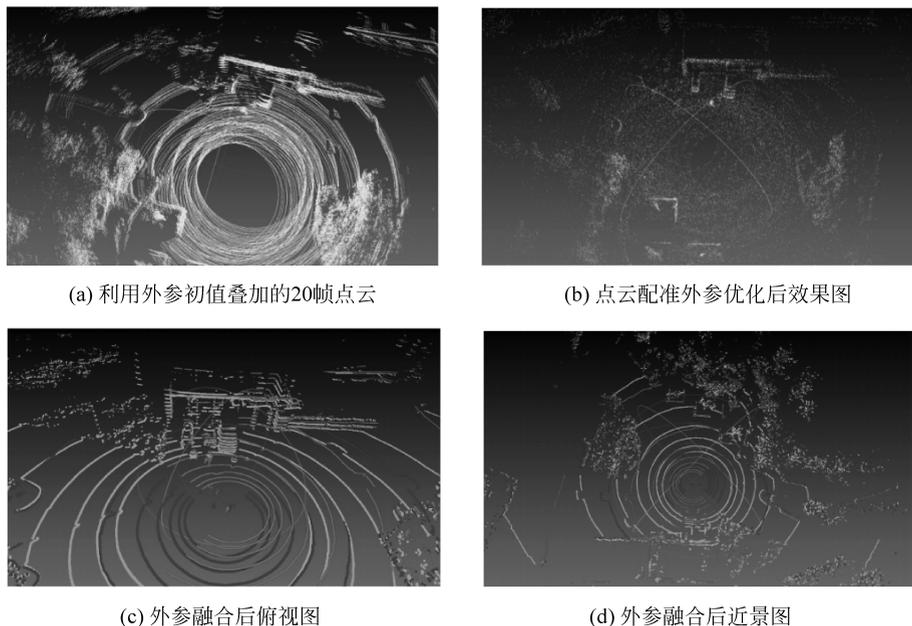


图 3-22 组合惯导标定过程

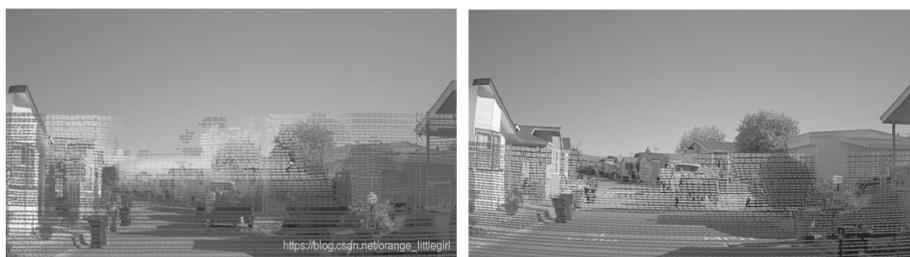
相机到激光雷达的标定方法比较依赖于初始外参值的选取,一个偏差较大的外参,有可能导致标定失败。所以,在条件允许的情况下,尽可能提供更加精准的初始外参值。由于标定方法是基于自然场景的,所以一个理想的标定场地可以显著地提高标定结果的准确度。建议选取一个纹理丰富的场地,如有树木、电线杆、路灯、交通标志牌、静止的物体和清晰车道线的场地。图 3-23 是一个标定场地示例。

在采集标定数据时,为避免时间不同步,尽量将车辆进行慢速行驶,可以有效地缓解因时间差异所引起的标定问题。相机中需看到一定数量的投影点云,否则该工具将无法进行外参标定运算。因此,建议使用短焦距相机来进行相机-激光雷达的标定。

当标定完成后,会生成相应的标定结果验证图片。如果 50m 以内的目标,点云边缘和图像边缘能够重合,则可以证明标定结果的精度很高。反之,若出现错位现象,则说明标定结果存在误差。当误差大于一定范围时(范围依据实际使用情况而定),该外参不可用。标定结果示例如图 3-24(a)和图 3-24(b)所示。



图 3-23 良好的标定场地示例



(a) 准确外参的点云投影效果

(b) 外参误差较大的点云投影效果

图 3-24 标定结果对比示例

## 3.4 传感器数据融合

### 3.4.1 传感器数据融合介绍

#### 1. 多传感器数据融合

多传感器数据融合(Multi-Sensor Data Fusion, MSDF)就是利用计算机技术将来自多传感器的信息和数据,在一定的准则下加以自动分析综合,以完成所需要的决策和估计而进行的信息处理过程。传感器数据融合是用于包含处于不同位置的多个或者多种传感器的信息处理技术。随着传感器应用技术、数据处理技术、计算机软硬件技术和工业化控制技术的发展成熟,多传感器数据融合技术已形成一门新兴学科。我国对多传感器数据融合技术的研究在工程上已应用于信息的定位和识别等方面。而且相信随着科学的进步,多传感器数据融合技术会成为一门智能化、精细化数据信息图像等综合处理和研究的专门技术。一般地,多传感器数据融合处理过程包括 6 个步骤,如图 3-25 所示。首先是多源传感系统搭建

与定标,进而采集数据并进行数字信号转换,再进行数据预处理和特征提取,接着是融合算法的计算分析,最后输出稳定的目标特征信息。

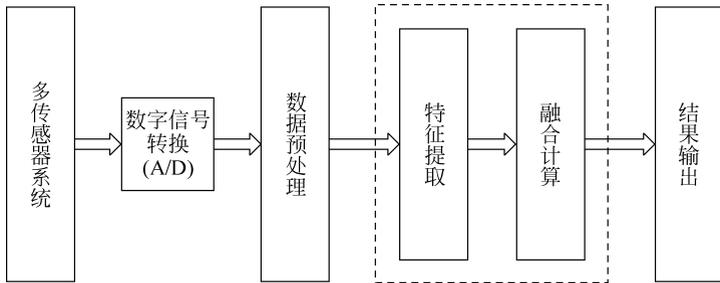


图 3-25 多传感器数据融合过程

作为多传感器融合的研究热点之一,融合方法一直受到人们的重视,国内外已经存在大量的研究成果,存在各类融合方法。常用方法大致可分为两大类:随机和人工智能方法。信息融合的不同层次对应不同的算法,包括加权平均融合、卡尔曼滤波法、Bayes 估计、统计决策理论、概率论方法、模糊逻辑推理、人工神经网络和 D-S(Dempster-Shafer)证据理论等。

## 2. 传感器数据融合定义

传感器数据融合可概括为把分布在不同位置的多个同类或不同类传感器所提供的局部数据资源加以综合,采用计算机技术对其进行分析,消除多传感器信息之间可能存在的冗余和矛盾,加以互补,降低其不确实性,获得被测对象的一致性解释与描述,从而提高系统决策、规划、反应的快速性和正确性,使系统获得更充分的信息。其信息融合在不同信息层次上出现,包括数据层融合、特征层融合、决策层融合。

(1) 数据层融合。针对传感器采集的数据,依赖于传感器类型,进行同类数据的融合。数据层的融合要处理的数据都是在相同类别的传感器下采集,所以数据层融合不能处理异构数据。

(2) 特征层融合。特征层融合指的是提取所采集数据包含的特征向量,用来体现所监测物理量的属性,这是面向监测对象特征的融合。如在图像数据的融合中,可以采用边沿的特征信息,来代替全部数据信息。

(3) 决策层融合。决策层融合指的是根据特征层融合所得到的数据特征,进行一定的判别、分类,以及简单的逻辑运算,根据应用需求进行较高级的决策,是高级的融合。决策层融合是面向应用的融合。比如在森林火灾的监测监控系统中,通过对于温度、湿度和风力等数据特征的融合,可以断定森林的干燥程度及发生火灾的可能性等。这样需要发送的数据就不是温度、湿度的值以及风力的大小,而只是发送发生火灾的可能性及危害程度等。在传感网络的具体数据融合实现中,可以根据应用的特点来选择融合方式。

### 3.4.2 自动驾驶传感器数据融合

自动驾驶感知和定位中传感器数据融合按照实现原理分为硬件层融合、数据层融合、任

务层融合三部分,自动驾驶汽车环境感知传感器布局如图 3-26 所示。

硬件层融合:如禾赛和 Mobileye 等传感器厂商,利用传感器的底层数据进行融合。

数据层融合:利用各种传感器得到的后期数据,即每个传感器各自独立生成目标数据,再由主处理器融合这些特征数据实现感知任务。

任务层融合:先由各传感器完成感知或定位任务,如障碍物检测、车道线检测、语义分割和跟踪以及车辆自身定位等,然后添加置信度进行融合。

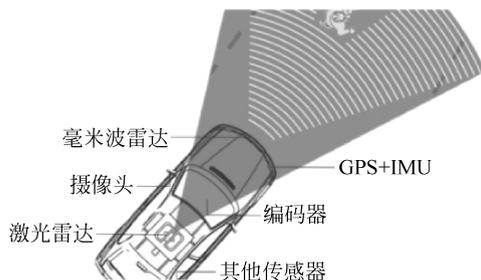


图 3-26 自动驾驶汽车环境感知传感器布局

有些传感器之间很难在硬件层融合,比如摄像头或者激光雷达和毫米波雷达之间,因为毫米波雷达的目标分辨率很低(无法确定目标大小和轮廓),但可以在数据层上探索融合,比如目标速度估计、跟踪轨迹等。摄像头和激光雷达的数据融合实际是激光雷达点云投影,在摄像头图像平面形成的深度和图像估计的深度进行结合。采用的传统方法主要包括 Bayes(贝叶斯)滤波、自适应引导图滤波、传统的形态学滤波等,目前采用较多的融合方法为基于深度学习的传感器数据融合算法,感兴趣的同学可以进行相关资料的查阅,在此就不再详细介绍。

传感器数据融合在于获得不同传感器的输入内容,并且使用组合在一起的信息来更加准确地感知周围的环境。相对于独立系统,这样可以做出更准确、更安全的决策。雷达也许不具有光传感器所具有的分辨率,不过它在测距和穿透雨、雪和浓雾方面具有很大优势。这些天气条件或光照不足的恶劣情况不利于摄像头发挥作用,不过摄像头能够分辨颜色(可以想一想街道指示牌和路标),并且具有很高的分辨率。目前路面上图像传感器的分辨率已经达到 100 万像素。在未来几年内,图像传感器的发展趋势将是 200 万像素甚至 400 万像素。

雷达和摄像头是两项传感器技术完美融合、互为补充的典范。采用这种方法的融合系统所实现的功能要远超这些独立系统能够实现的功能总和。使用不同的传感器种类可以在某一种传感器全都出现故障的环境条件下,额外提供一定冗余度。这种错误或故障可能是由自然原因(诸如一团浓雾)或是人为现象(例如对摄像头或雷达的电子干扰或人为干扰)导致的。即使是在一个传感器失效的情况下,这样的传感器融合系统也可以保持某些基本或紧急的功能。

### 3.4.3 自动驾驶传感器数据融合示例

不同传感器类型,其数据融合的复杂程度有所不同,以下为两个基本的传感器融合示例:①后视摄像头加上超声波测距;②前视摄像头加上多模式前置雷达。如图 3-27 所示,可通过对现有系统进行轻微更改或通过增加一个单独的传感器融合控制单元来实现。如将前置雷达与前视摄像头融合在一起,以实现自适应巡航控制加车道保持辅助,或者将后视摄像头与超声波测距报警组合在一起来实现自动泊车。

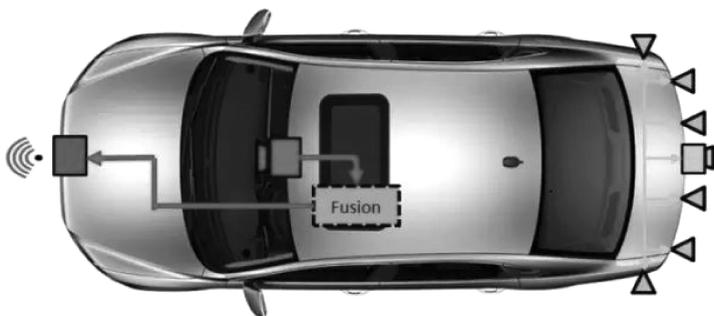


图 3-27 前视摄像头和超声波雷达传感器数据融合

将前视摄像头的功能与前置雷达组合在一起,能够在任何天气条件下测量高达 150m 的物体的速度和距离,摄像头在探测和区分物体方面(包括读取街道指示牌和路标)十分出色。通过使用具有不同视场角(FoV)和不同光学元件的多个摄像头传感器,系统可以识别车前通过的行人和自行车,以及 150m 甚至更远范围内的物体,同时还能可靠实现自动紧急制动和城市启停巡航控制等功能。传感器数据融合除了能自主实现复杂的功能外,还可以在现有功能中实现更少的误报和漏报。

### 3.4.4 多传感器融合数据处理

自动驾驶多传感器融合系统中,如何将传感器的数据发送到中央电子控制单元(ECU)直接影响着自动驾驶性能。当对遍布车身的多个传感器进行融合时,需要特意考虑传感器和中央融合 ECU 之间的连接,对于数据处理的位置也是如此,它也会影响整个系统的实现,传感器融合数据处理方式主要有以下 3 种。

(1) 集中式处理。所有的数据处理和决策制定都是在同一个位置完成,数据是来自不同传感器的“原始数据”,原理图如图 3-28 所示。

集中处理方式优点为:传感器模块体积小小巧、成本低、功耗也低,这是因为其只需要执行检测和数据传输任务。传感器的安装位置也很灵活,并且所需安装空间很小,替换成本低。通常情况下,由于无须处理或做决策,传感器模块具有较低的功能安全要求。ECU 可以获取全部数据,这是因为数据不会因为传感器模块内的预处理或压缩而丢失。由于传感

器成本较低,并且外形尺寸较小,因此可以部署更多的传感器。其缺点为:实时处理传感器数据需要提供宽带通信(数据传输速度高达数 Gb/s),因此可能出现较高电磁干扰(EMI),中央 ECU 需要有高处理能力和速度来处理所有输入数据。对于很多高带宽 I/O 和高端应用处理器来说,这意味着更高的电能需求和更大的散热量,同时传感器数量增加需要性能更好的中央 ECU。FPD-Link III 接口,可以在一根同轴电缆上传送传感器及功耗、控制和配置等多种数据(双向反向通道),这样便可极大降低系统的接线要求。

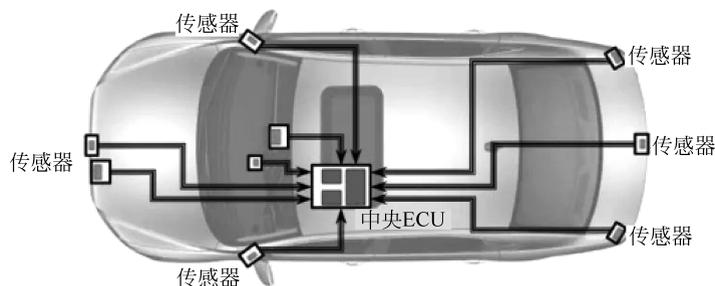


图 3-28 传感器融合数据集中处理示意图

(2) 全分布式处理。由本地传感器模块进行高级数据处理,并在一定程度上进行决策制定的。全分布式系统只将对象数据或元数据(描述对象特征或识别对象的数据)发回到中央 ECU。ECU 将数据组合在一起,并最终决定如何执行或做出反应,如图 3-29 所示。

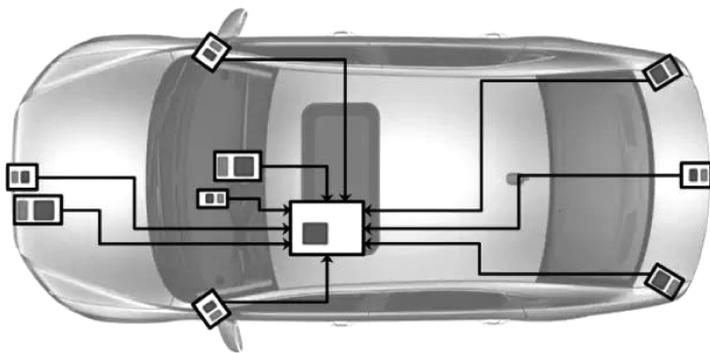


图 3-29 传感器融合数据分布式处理示意图

分布式处理优点为:传感器模块与中央 ECU 之间可以使用更低带宽、更加简单且便宜的接口。在很多情况下,小于 1Mb/s 的 CAN 总线就足够用了。中央 ECU 只需将对象数据融合在一起,因此其所需处理能力更低。对于某些系统来说,用一个高级的安全微控制器就能满足要求,同时模块更小,所需功耗也就更低。由于很多处理都是在传感器内部完成的,传感器数量增加不会大幅增加对中央 ECU 的性能需求。其缺点为:传感器模块需要应用处理器,使得体积更大、价格更高且功耗更大。由于本地处理和决策制定,传感器模块的功能安全要求也就更高。当然增加更多的传感器,成本也会大幅上升。中央决策制定 ECU

只能获取对象数据,而无法访问实际的传感器数据。因此很难实现感兴趣的区域“放大”。

(3) 集中式和分布式处理相结合。根据系统中所使用传感器的数量与种类,针对不同车型和升级选项的可扩展性要求,将两种拓扑混合在一起得到的一种优化解决方案。一个全分布式系统可以使用现有的传感器模块与数据融合 ECU 组合在一起,如图 3-30 所示。

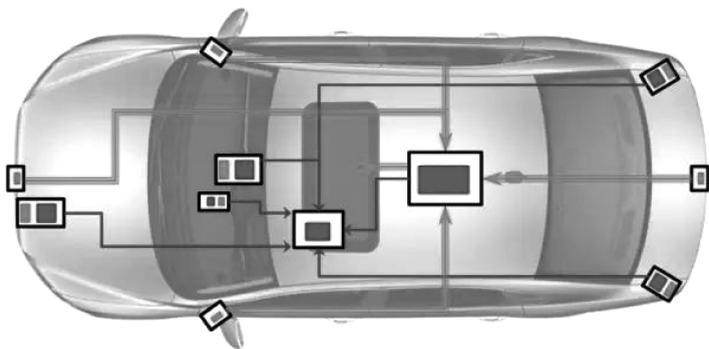


图 3-30 集中式和分布式处理相结合

### 3.4.5 传感器数据融合算法

利用多个传感器所获取的关于对象和环境全面、完整的信息,主要体现在融合算法上,多传感器系统的核心问题是选择合适的融合算法。对于多传感器系统来说,信息具有多样性和复杂性,因此对信息融合算法的基本要求是具有鲁棒性和并行处理能力,以及较高的运算速度和精度。下面简要介绍 3 种常用的数据融合算法:贝叶斯统计理论、神经网络技术以及卡尔曼滤波方法。

#### 1) 贝叶斯统计理论

贝叶斯统计理论是一种统计学方法,用来估计统计量的某种特性,是关于随机事件  $A$  和  $B$  的条件概率的一则定理。所谓“条件概率”(Conditional Probability),就是指在事件  $B$  发生的情况下,事件  $A$  发生的概率,用  $P(A|B)$  来表示。根据图 3-31 容易推导得到:

$$P(A|B) = P(A) \frac{P(B|A)}{P(B)} \quad (3-7)$$

由此可以推导出条件概率的公式,其中  $P(A)$  称为先验概率(Prior Probability),即在事件  $B$  发生之前,对事件  $A$  发生概率的认识。

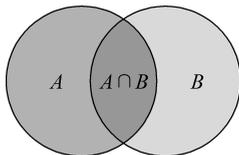


图 3-31 贝叶斯统计理论文氏图

视觉感知模块中交通限速标志识别(Traffic Sign Recognition, TSR)是智能驾驶的重要一环。在识别过程中,交通限速标志牌被树木、灯杆等遮挡是影响识别的主要干扰。一般定义事件  $A$  为交通信号标志正确识别,事件  $B$  为交通信号标志未能识别,例如事件  $A$  为限速标志未被遮挡,事件  $B$  为限速标志被遮挡。限速标志未被遮挡完全暴露出来,识别率是相当高的,但如果限速标记牌被阻挡住,识别率是比未遮挡的低很多。这两个指标的融合使用,可以用于作为评价目前图像处理算法识别限速标志性能的重要参考,实际的融合过程比这复杂得多。

### 2) 卡尔曼滤波方法

卡尔曼滤波(Kalman Filtering)是一种利用线性系统状态方程,通过系统输入/输出观测数据,对系统状态进行最优估计的算法。简单来说,卡尔曼滤波是一个最优化自回归数据处理算法,卡尔曼滤波在测量方差已知的情况下能够从一系列存在测量噪声的数据中,估计动态系统的状态,而且便于计算机编程实现,并能够对现场采集的数据进行实时的更新和处理。卡尔曼滤波是目前应用最为广泛的滤波方法,在通信、导航、制导与控制等多领域得到了较好的应用。卡尔曼滤波是多源传感数据融合应用的重要手段之一,目前高级辅助驾驶系统上,搭载有毫米波雷达和超声波雷达模块,两者均能对障碍物车辆进行有效的位置估计判别。雷达利用主动传感原理,发射毫米波,接收障碍物回波,根据波传播时间计算角度距离。两者均能识别出车辆位置,那么如何融合信息?如何计算出具体的车辆位置呢?卡尔曼滤波利用目标的动态信息,去掉噪声的影响,得到关于目标位置的估计,这个估计可以是对当前目标位置的估计(滤波),也可以是对将来位置的估计(预测),还可以是对过去位置的估计(插值或平滑)。卡尔曼滤波就是这样一个根据当前时刻目标的检测状态,预测估计目标下一时刻目标检测状态的一个动态迭代循环过程。

### 3) 神经网络技术

神经网络(Neural Network, NN)是机器学习(Machine Learning, ML)的一种方式,是人工智能、认知科学、神经生理学、非线性动力学、信息科学和数理科学的“热点”。神经网络由一层一层的神经元构成,层数越多其深度越深。所谓深度学习(Deep Learning)就是用很多层神经元构成的神经网络达到机器学习的功能,神经细胞通过输入层、隐含层和输出层的连接,形成一个复杂的神经网络系统,通过有效的学习训练,使输出层的结果与现实越来越靠近,误差越来越小,当其精度满足一定的功能需求时,神经网络训练完毕,此刻构建的神经网络系统即能解决众多机器学习上的图像识别、语音识别和文字识别上的问题,神经网络结构如图 3-32 所示。

在智能驾驶目前的发展历程上看,神经网络技术乃至现在最新的深度学习技术,广泛用于视觉感知模块的车辆识别、车道线识别和交通标志识别。通过对中国路况工况的数据采集和处理,广泛获取国内不同天气状况(雨天、雪天、晴天等),不同路况(城市道路、乡村道路、高速公路等)的真实的的环境数据,为深度学习提供了可靠的数据基础。神经网络的输入层数据,也即是传感器获取的数据,是多源多向的,可以是前挡风玻璃片上视觉感知模块的障碍物位置、形状、颜色等信息,也可以是毫米波雷达、超声波雷达检测的障碍物距离、角度、

速度、加速度等信息,还可以是 360°环视系统上采集的车位数据、地面减速带数据。

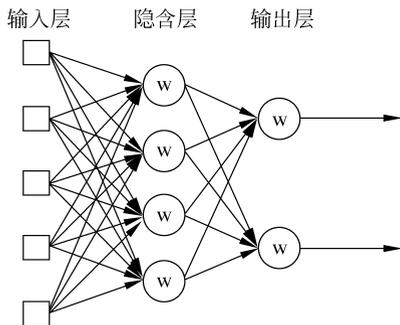


图 3-32 神经网络结构

## 3.5 环境感知与预测

### 3.5.1 环境感知理论概述

环境感知也被称为 MODAT (Moving Object Detection and Tracking), 是自动驾驶四大核心技术(环境感知、精确定位、路径规划和线控执行)中被研究最多的部分。环境感知主要包括路面、静态物体和动态物体 3 个方面。对于动态物体,不仅要检测,还要对其轨迹进行追踪,并根据追踪结果预测该物体下一步的轨迹(位置)。这在北京市区必不可少,最典型场景就是北京五道口:如果见到行人就停,那就永远无法通过五道口,行人几乎是从不停歇地从车前走过。驾驶员会根据行人的移动轨迹大概评估其下一步的位置,然后根据车速计算出安全空间(路径规划)。多个移动物体的轨迹的追踪与预测,难度比单一物体要高得多,也是无人车最具难度的技术。

### 3.5.2 环境感知算法与应用实例

#### 1. 车道线识别

车道线识别技术是指从自动驾驶车辆装载的摄像头获取的视频及图像中识别出路面的车道线及其位置的技术。由于车道线的视觉特征比较明显,所以车道线识别一般采用摄像机获取图像,应用相应的图像处理技术来分析图片中的车道线,按照摄像头的标定信息将图片中的车道线位置转换为车体坐标系中的位置,用于自动驾驶车辆后续的规划和控制,其应用的图像处理技术主要可分为传统图像处理和深度学习图像处理两种。

(1) 传统图像处理。通过提取车道线的图像特征(包括边缘、色差、直线、曲线等),直接标记车道线位置区域,最后将位置区域拟合为车道线,具体步骤如图 3-33 所示。

车载摄像机一般为广角摄像机,获取的原始图像存在一定的失真现象。图像预处理就

是根据摄像机的内部参数及标定的外部参数将原始图像调整为正常比例图像。灰度化处理是将图像中每个点 RGB 三维的数据整合为一维的灰度数据,这样可以简化后续处理的复杂度,缩短整体处理时间,如图 3-34(a)和图 3-34(b)所示。

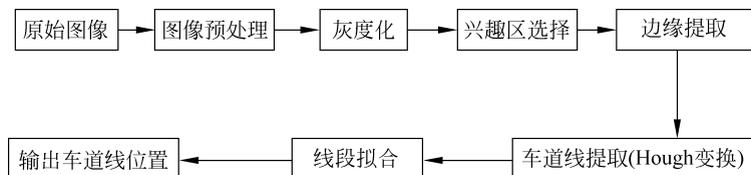


图 3-33 传统图像处理步骤



(a) 车道线原图

(b) 灰度图像车道线识别图

图 3-34 传统灰度化车道线识别

一般情况下车道线与路面有较明显的颜色差异(车道线为白色或黄色,路面为灰色或黑色),因此车道线的边缘处会有明显的颜色变化,在一维的灰度数据中表现为明暗变化。边缘提取就是标记出这样变化明显的位置,常用算法有 Canny 算法和 Sobel 算法,可以根据实际效果选取效果好的算法,如图 3-35(a)和图 3-35(b)所示。



(a) 车道线原图

(b) 边缘提取车道线识别图

图 3-35 边缘提取车道线识别

对于固定安装的摄像机,车道线在图中的位置相对固定,如平视安装的摄像机,车道线一般位于生成图片的下半部分,兴趣区域选择就是截取图片的对应区域,以减少计算和增加

识别精度。

Hough 变换是提取图片中直线、圆形、椭圆特征的常用方法。选定兴趣区域以后，Hough 变换可以帮助我们提取图片中的直线段，即车道线中的直线段，最后将提取的线段拟合为连续的直线或者曲线，即可生成车道线，如图 3-36(a)和图 3-36(b)所示。

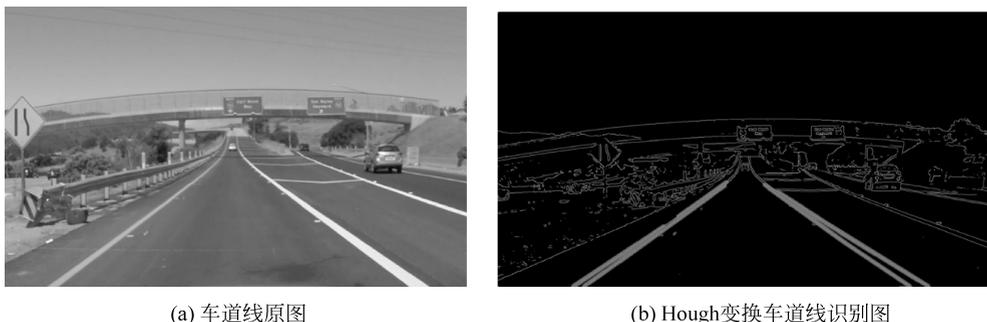


图 3-36 Hough 变换车道线识别

(2) 深度学习图像处理。传统方法使用固定图像特征，很容易在道路场景变化时出现准确度下降的情况，深度学习图像处理训练深度神经网络来提取图像中车道线的特征，进而对图像中每个像素点分类，判断像素点是否属于车道线的一部分，最后再将像素点聚类并拟合为车道线，其识别过程如图 3-37 所示。



图 3-37 基于深度学习的车道线识别过程

除了传统图像处理中还原失真的预处理以外，深度学习图像预处理往往还包括数据增强和数据归一化等，这些操作可以使深度神经网络在训练时收敛更快。车道线提取中常用的神经网络有卷积神经网络(CNN)和循环神经网络(RNN)。CNN 善于在复杂场景中分类，可以定位属于车道线的像素点。RNN 可以应用于车道线边界的识别。根据深度神经网络设计的不同，像素点的聚类拟合也不尽相同。一般情况下，深度神经网络输出各像素点为车道线像素点的准确度，之后依据模型拟合出车道线曲线。常用模型有三次多项式，样条曲线等。也可以使用支持向量机(SVM)或者 RNN 来识别车道线边界。

深度神经网络也可以学习并提取图像特征，并分类识别每个像素点是否为车道线。但是由于自动驾驶场景的多变，获取的图片会有明暗变化、阴影遮挡、上下坡造成的扭曲、晃动造成的图像模糊等问题，这些问题都是车道线识别技术在实际应用中的挑战。

在自动驾驶领域，除了实时识别车道线，同样可以在高精地图中定义好车道线的位置。在车辆获得精确的定位以后，依据地图的定义，自动驾驶车辆就可以依据自身在地图中的位置确定车道线的位置，从而知道可以行驶的区域。此方法需要提前制定高精地图，同时需要

精确定位辅助。

## 2. 车辆行人识别

在自动驾驶领域,车辆行人识别是车辆完成自动驾驶能力的关键一环。自动驾驶车辆不仅要识别出道路中移动的其他车辆、行人、摩托车以及未知障碍物等,还要对其位置、速度和运动轨迹做出判断,从而在规划控制中采取不同策略应对不同场景,识别步骤如图 3-38 所示。

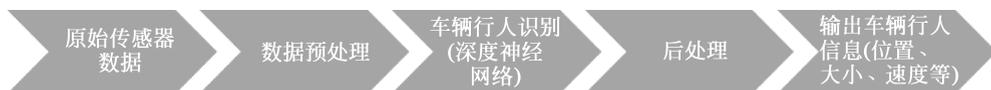


图 3-38 自动驾驶车辆行人识别过程

常见的自动驾驶车辆传感器以摄像机、毫米波雷达和激光雷达为主。由于毫米波雷达分辨率较低,一般作为辅助传感器,车辆行人识别主要应用摄像机和激光雷达。其中摄像机输出为二维图像,富含颜色、强弱等视觉信息,在识别车辆行人等功能上有一定优势。但是在识别物体的定位上有所不足,需要多个摄像机协同或者其他测距传感器辅助才能完成。激光雷达输出为三维点云,虽然没有色彩属性,但包含识别物体的位置信息、反光强弱等。即便分类识别精度不如摄像头,但是可以一次性输出自动驾驶系统后续所需的障碍物位置、大小、速度等信息。车辆行人识别主要分为传统图像处理和神经网络学习两种。

(1) 传统图像处理。以摄像头为主的自动驾驶方案中,图像处理是车辆行人识别技术的基础,摄像机捕捉到的图片在预处理以后,输入深度神经网络进行处理。根据摄像机方案的不同,预处理不仅限于归一化等深度学习标准操作,也包括多摄像头的拼接、其他传感器的数据融合等,如图 3-39 所示。



图 3-39 传统图像处理车辆行人识别

(2) 神经网络学习。常用的深度神经网络为卷积神经网络(CNN),经过调整和训练,神经网络可以将对应方案输入图片中的像素点一一分类,标记为车辆、行人和骑手等,并输出每个分类的可信度。神经网络的输出结果需要进行处理和分析,筛选出最可能的标记像素点。这些像素点会被组合成完整的物体,从而得到该物体的位置和大小,速度等动态信息可以根据历史记录追踪计算。

以激光雷达为主的自动驾驶方案中,点云是车辆行人识别的基础。与图像不同,点云本身就是三维数据,传统的用于图像处理的 CNN 无法直接作用于点云数据。所以需要三

维点云降到二维或者调整 CNN 使之可以应用三维数据。点云降维处理方式是将点云数据压缩整合,并投影到某一个视角下的平面中,常用视角包括车前视角和鸟瞰图视角。其中鸟瞰图视角的投影效果更好,应用更加方便。投影过的点云可以看作一张二维图像,可作为 CNN 的输入。由于点云数据包含三维位置信息,可以很容易得到鸟瞰图视角下各个物体的位置信息。

另一种方法是将点云数据压缩整合为一个个三维像素点,整个点云组成一个三维图像,调整 CNN 使之可以输入三维像素点,之后 CNN 直接分析每个三维像素点,并对其分类。分类后的三维像素点包含固有的三维数据。同图像处理类似,分类后的二维或者三维像素点需要聚类组合后,才能最终确定识别到的物体的大小及位置,点云处理输出图像如图 3-40 所示。



图 3-40 自动驾驶点云车辆行人识别

### 3. 交通标志与交通灯识别

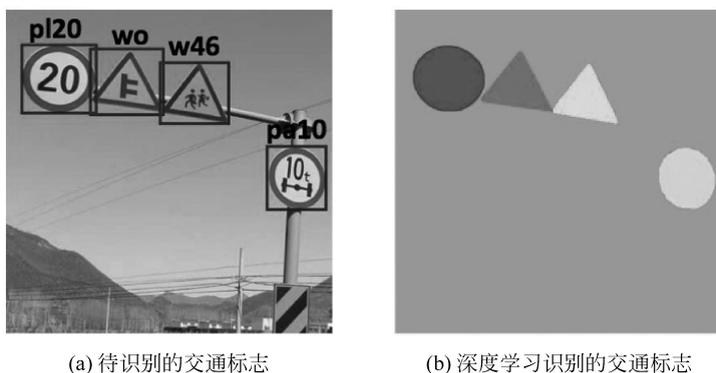
交通标志与交通灯识别技术是检测并识别道路交通标识和交通灯的技术,用来让自动驾驶车辆读懂道路行驶规则。

(1) 交通标志识别。由于交通标志的图像特征(形状、颜色等)明显,一般使用摄像机获取图像数据用于识别。识别方法大致分为模型识别和深度学习识别。早期的交通标志识别以模型识别为主,主要应用交通标志的颜色、形状、边缘等特征,在图像中找寻特定标志并识别具体内容。随后,模式特征、表面特征、模板等也被引入交通标志识别,并使用常见的机器学习技术(如支持向量机 SVM 和各种分类器)来分类识别标志,但是这些方法不具备很好的泛化能力。随着大量数据集的建立,深度学习技术也被应用到交通标志识别技术中,获得了不错的效果,如图 3-41(a)和图 3-41(b)所示。

(2) 交通灯识别。发现并识别出自动驾驶车辆前方交通灯的位置及显示状态(红、黄、绿),通常情况下只需要识别一个交通灯的状态,复杂情况下则需要识别多个交通灯的状态并判断出哪个交通灯对应车辆行驶的车道。

由于交通灯的结构特征明显,通常选用图像中的颜色(红、黄、绿)、形状(圆形或箭头)和结构(三个相连)等作为判断的特征,创建模型进行模型识别。然而,这类方法在复杂的情形(过曝、遮挡、阴影、交通灯安装不标准等)中效果不好,因此模型识别的方法逐渐被机器学习和深度学习方法所替代。SVM、CNN 等都被应用到交通灯的识别当中。交通标志作为每

条道路的固定规则,往往变化情况较少,但是不同国家地区应用的标准不同。因此,用一套系统实时识别道路中的所有交通标志信息是充满挑战的任务。在配备高精地图的自动驾驶车辆中可以在地图中定义对应的规则。这样,结合自动驾驶车辆的实时定位,可以从高精地图中获取相应的道路行驶规则,省略了识别的过程。人行道等也可以定义在高精地图中。



(a) 待识别的交通标志

(b) 深度学习识别的交通标志

图 3-41 交通标志识别

另一方面,车联网领域车路协同的发展也使得自动驾驶车辆实时获取交通标志信息和交通灯信息成为可能。配备车联网通信功能的自动驾驶车辆可以通过车路协同与行驶道路设施通信,准确实时地获取所需信息,交通灯识别效果如图 3-42 所示。



图 3-42 交通灯识别

## 3.6 定位系统

### 3.6.1 定位理论概述

自动驾驶的三大工作内容分别是感知、决策和执行,感知中很重要的一个部分就是定位,如果没有精确的定位结果,后续的决策和执行将无从谈起。自动驾驶中的汽车定位要达到安全可靠,需满足以下性能指标:

- (1) 精度: 测量值和真实值之间的重合度。
- (2) 完好性: 服务不可用时提出告警的能力。

(3) 连续性：告知客户系统正常工作的持续能力。

(4) 可用性：提供符合指标定位服务的百分比。

常用的定位方式可以划分为 GPS(Global Position System, 全球定位系统)定位、DGPS(Differential Global Position System, 差分全球定位系统)定位、RTK(Real Time Kinematic, 实时动态)定位、惯导定位和高精(地图)定位,下面对定位原理进行介绍。

### 3.6.2 定位传感器

#### 1. GPS 接收天线

GPS 接收天线是接收全球定位系统卫星信号并确定地面空间位置的仪器。GPS 卫星发送的导航定位信号,是一种可供用户共享的信息资源。对于陆地、海洋和空间的广大用户,只要用户拥有能够接收、跟踪、变换和测量 GPS 信号的接收设备,就可以在任何时候进行导航定位测量。新石器自动驾驶汽车 GPS 接收无线如图 3-43 所示。



(a) GPS接收天线模块正面视图

(b) GPS接收天线侧面视图

图 3-43 新石器自动驾驶汽车 GPS 接收无线

GPS 接收天线特点：①天线部分采用多馈点设计方案,实现相位中心与几何中心的重合,将天线对测量误差的影响降低到最小；②天线单元增益高,方向图波束宽,确保低仰角信号的接收效果,在一些遮挡较严重的场合仍能正常搜星；③带有抗多径扼流板,有效降低多径对测量精度的影响；④防水、防紫外线外罩,为天线能长期在野外工作提供保障。

#### 2. IMU

IMU 是测量物体三轴姿态角(或角速度)以及加速度的装置,如图 3-44 所示。IMU 包含了三个单轴的加速度计和三个单轴的陀螺,加速度计检测物体在载体坐标系独立三轴的加速度信号,而陀螺检测载体相对于导航坐标系的角速度信号,测量物体在三维空间中的角速度和加速度,并以此计算出物体的姿态,其在导航中有重要应用价值。

卫星组合导航系统由高精度测绘级卫星接收板卡、三轴 MEMS 陀螺仪、三轴 MEMS 加速度计组成,可在星况良好的环境下提供厘米级定位精度,并在卫星信号遮挡、多路径等环境下长时



图 3-44 星宇达的 M2 IMU

间保持位置、速度、姿态精度。设计轻便小巧、简单易用,适用于辅助驾驶、无人驾驶、车载定位定向、AGV 车等环境,主要特点如下:

- (1) 产品支持后处理解算。
- (2) 适合车载/无人驾驶行业的接口设计。
- (3) 宽电压输入 9~36V。
- (4) 最高 200Hz 数据更新速率。
- (5) IP65 防水防尘等级。
- (6) 工作温度  $-30^{\circ}\text{C}\sim 70^{\circ}\text{C}$ 。

无人驾驶系统依靠 GPS 和 IMU 就可以知道经纬度信息和行驶方向(航向),当然 IMU 还能提供诸如横摆角速度、角加速度等更丰富的信息,这些信息有助于自动驾驶汽车的定位和决策控制。

### 3.6.3 定位技术

#### 1. GPS 定位技术

(1) GPS 的构成。定位技术已经成为越来越热的一门技术,以后也将成为所有移动设备的标配,而定位技术中,目前精度最高、应用最广泛的就是 GPS 定位。

GPS 主要包括空间部分、控制部分和用户部分三个部分,其构成如图 3-45 所示。

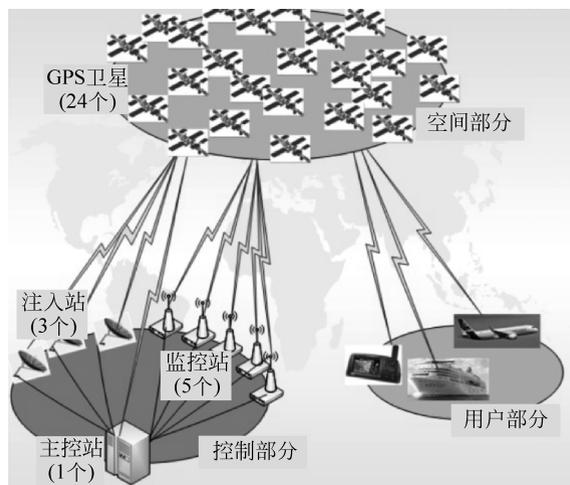


图 3-45 GPS 的构成

GPS 空间部分主要由 24 颗 GPS 卫星构成,其中 21 颗工作卫星,3 颗备用卫星。24 颗卫星运行在 6 个轨道平面上,运行周期为 12 个小时。保证在任一时刻、任一地点高度角  $15^{\circ}$  以上都能够观测到 4 颗以上的卫星,主要发送导航定位的卫星信号,GPS 导航星座图如图 3-46 所示。

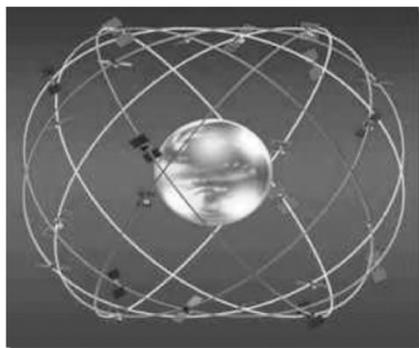


图 3-46 GPS 导航星座图

GPS 控制部分由 1 个主控站、5 个监控站和 3 个注入站组成,其作用是监测和控制卫星运行、编算卫星星历(导航电文)和保持系统时间。

(2) GPS 定位原理。GPS 定位实际上就是通过 4 颗已知位置的卫星来确定 GPS 接收器的位置。实际解算过程非常复杂,而且涉及误差的评估和计算,本章节不深入讨论。GPS 位置解算如图 3-47 所示。

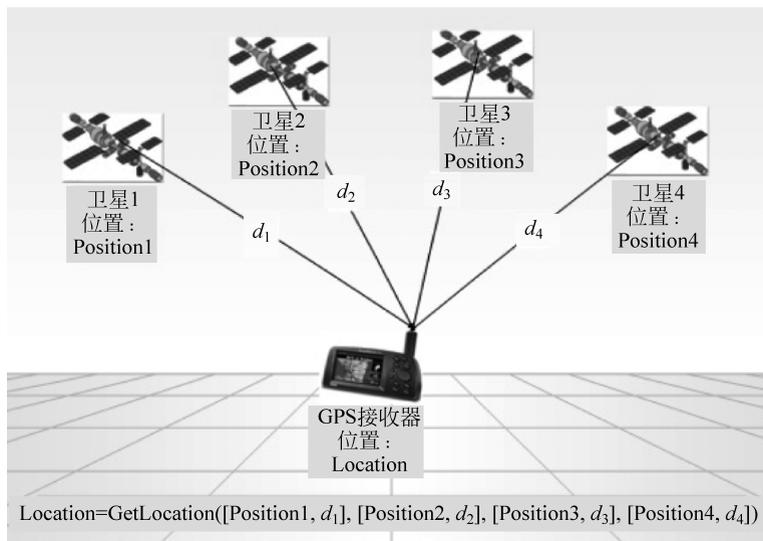


图 3-47 GPS 位置解算

图中的 GPS 接收器为当前要确定位置的设备,卫星 1、2、3、4 为本次定位要用到的 4 颗卫星; Position1、Position2、Position3、Position4 分别为 4 颗卫星的当前位置(空间坐标);  $d_1$ 、 $d_2$ 、 $d_3$ 、 $d_4$  分别为 4 颗卫星到要定位的 GPS 接收器的距离; Location 为要定位的卫星接收器的位置。

每个 GPS 卫星都在广播自己的位置,在发送位置信息的同时,也会附加上该数据包发出时的时间戳。GPS 接收器收到数据包后,用当前时间减去时间戳上的时间,就是数据包在空中传输所用的时间了。时间乘上它的传输速度,就是数据包在空中传输的距离,也就是该卫星到 GPS 接收器的距离了。理想速度就是光速  $c$ ,把传播时间记为  $T_i$ ,用公式表示就是:

$$d_i = c \times T_i \quad (i = 1, 2, 3, 4) \quad (3-8)$$

根据立体几何知识,三维空间中三对  $[\text{Position}_i, d_i]$  数据就可以确定一个点了(实际上可能是两个,可以通过逻辑判断舍去一个),为什么这里需要四对呢?理想情况下,的确三对就够了,但是事实上,因  $c$  值是很大的(理想速度即光速),那么对于时间  $T_i$  而言,一个极小的误差都会被放大很多倍从而导致整个结果无效。在 GPS 定位中,对时间的精度要求是极高的,会出现一个卫星上的时钟和接收机时钟时间同步的误差,这个误差作为一个待求参数,所以四个未知参数就需要 4 对方程来求解。这就是 `GetLocation()` 函数必须用四组数据的原因,也就是为什么必须有四颗卫星才能定位的原因。进行位置计算时都是用空间坐标形式表示,但是对 GPS 设备及应用程序而言,通常需要的是经度、纬度和高度的信息,进行从空间坐标形式到经纬度形式的转换就可以得到经纬度坐标。

## 2. DGPS 定位技术

DGPS 即差分全球定位系统。目前 GPS 能够提供的定位精度是优于 10m,为得到更高的定位精度,可以采用差分 GPS 技术。将一台 GPS 接收机安置在基准站上进行观测,根据基准站已知的精确坐标,计算出基准站到卫星的距离和由于误差的存在基准站接收机观测的伪距离之间存在一个差值,这个差值(改正数)由基准站实时地发送出去。用户接收机在进行 GPS 观测的同时,也接收到基准站发出的正数,并对其定位结果进行修正,从而提高定位精度。绝对单点定位有多种误差,导致精度受限。而差分定位就是为了消除或者减弱这种误差设计的,所以定位精度可以提高。DGPS 定位的优点主要包括:

- (1) 消除卫星时钟差。
- (2) 消除卫星星历误差。
- (3) 消除电离层延迟。
- (4) 消除对流层延迟。
- (5) 将接收机时钟差作为未知数求出。

以上措施将有效地提高 GPS 定位精度。

## 3. RTK 定位技术

RTK 载波相位差分技术是实时处理两个测量站载波相位观测量的差分方法。如图 3-48 所示,将基准站采集的载波相位发给用户接收机进行求差解算坐标。这是一种新的常用的卫星定位测量方法,以前的静态、快速静态、动态测量都需要进行解算才能获得厘米级的精度,而 RTK 是能够在野外实时得到厘米级定位精度的测量方法。

RTK 系统由基准站子系统、管理控制中心子系统、数据通信子系统、用户数据中心子系

统和用户应用子系统组成,可以在很短的时间内获得厘米级的定位精度,被广泛应用于控制测量、施工放样、工程测量及地形测量等领域,但 RTK 也有一些缺点,主要表现在需要架设本地参考站,误差随流动站到基准站距离的增加而变大。

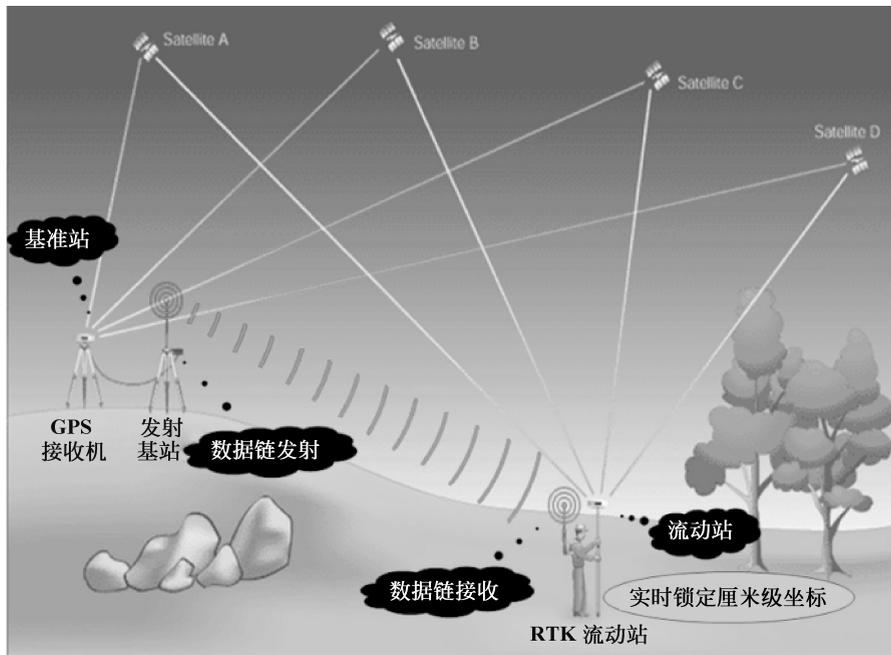


图 3-48 RTK 定位原理图

#### 4. 多传感器融合定位技术

GNSS、IMU、视觉定位、激光雷达、轮速计、高精地图等都是传感器的一种,多种传感器可采集不同形式的数据,用于生成不同类型和不同权重的约束。多传感器融合定位技术,可以规避单一定位方法的缺陷,获得更准确的定位结果。

(1) 惯性导航原理。惯性测量装置包括加速度计和陀螺仪,又称惯性测量单元(IMU)。3个自由度陀螺仪用来测量运动载体的3个转动运动的角速度;3个加速度计用来测量运动载体的3个平移运动的加速度。计算机根据测得的加速度信号计算出运动载体的速度和位置数据,控制显示器显示各种导航参数,其原理如图3-49所示。

惯性导航的优点主要有:不依赖于任何外部信息,也不向外部辐射能量,故隐蔽性好且不受外界电磁干扰的影响;可全球、全天候工作于空中、地表乃至水下,能提供位置、速度、航向和姿态角数据,产生的导航信息连续性好而且噪声低;数据更新率高、短期精度和稳定性好。

惯性导航的缺点主要有:由于导航信息经过积分而产生,定位误差随时间而增大,长期精度差;每次使用之前需要较长的初始对准时间;设备的价格较昂贵,而且不能输出时间信息。



一定的顺序交错发射避免相互干扰,光源和接收器组件旋转起来之后,即可获得对周围环境的多线扫描结果,形成一个三维坐标系中点的集合,称为点云。

激光雷达定位可分为有图定位和无图定位,有图定位分为建图和用图两个步骤。建图时将点云逐帧叠加在车辆的运行轨迹上即可获得点云地图。运行轨迹可以是高精度组合惯导系统输出的轨迹,也可以是点云 vSLAM 输出的轨迹。点云直接建图通常会生成比较大的点云文件,原始点云地图并不适合大范围使用。建立点云地图后,在该点云地图上对每帧点云用匹配算法就能获得当前车辆的准确位置,如图 3-51 和图 3-52 所示。



图 3-51 激光雷达点云图

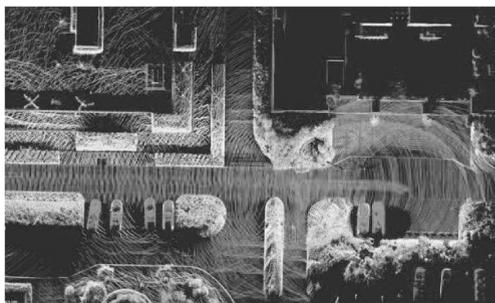


图 3-52 激光雷达构建某园区点云地图

无图定位和视觉里程计类似,将点云两帧匹配并组合后构造一个点云里程计以实现相对定位,例如开源软件 Autoware 里的点云定位模块,也可以提取点云的平面特征和角点特征做匹配,构建点云特征里程计实现定位。目前很多自动驾驶公司采用的都是点云地图定位,少数采取的是视觉 vSLAM 定位,例如北京新石器公司采取的是百度高精地图点云匹配定位。

## 本章小结

自动驾驶的三大工作内容分别是感知与定位、决策和执行,其中感知与定位技术在自动驾驶中起着至关重要的作用,它们使汽车能够监控周围环境,探测障碍物并规划道路。感知与定位包括三个阶段:感知阶段、定位阶段和规划阶段。感知阶段通过摄像头、雷达等传感装置感知行驶环境并根据所感知到的环境做出操纵指令;定位阶段通过高精度地图、GPS 等静态数据固化,降低感知层面的计算量;规划阶段则是要与各个信息源互动,实时规划出车辆的最佳反应。本章主要介绍自动驾驶传感器类型与工作原理,如多线激光雷达、组合惯导、相机、毫米波雷达、超声波雷达等;然后介绍环境感知与传感器数据融合方法,最后介绍了常用的自动驾驶定位技术,如 GPS 定位、RTK 定位、多传感器融合定位等定位方法。