本章目录

前沿技术与控制方法

第六章 外部控制与轨迹规划

6.1 外部控制基础与控制分层
6.2 外部控制原理与控制模式
6.3 轨迹规划建模与生成
6.4 轨迹规划算法与避障
6.5 轨迹跟踪原理与高机动控制
6.6 前沿技术与控制方法
6.7 平台框架与通信接口
6.8 外部控制例程与实验验证
6.9 习题与课堂讨论

本节 4 个教学单元

6.6.1 自动驾驶与外部控制
教材对应：第6.1.4.1节

6.6.2 端到端与强化学习
教材对应：第6.1.4.2节

6.6.3 大模型任务与工具调用原理
教材对应：第6.1.4.3节

6.6.4 VLA动作生成与适配
教材对应：第6.1.4.4节

本节导入：四种技术路线

四种路线改变闭环的不同部分

路线	输入与输出	主要改变
模块化	传感数据 → 环境、轨迹、控制量	显式划分各模块职责
端到端学习	选定观测 → 选定层级动作	联合训练一段输入输出映射
大模型任务调用	自然语言 → 任务或工具参数	任务表达与调度方式
视觉语言动作（VLA）	图像＋语言 → 学习得到的动作	语言条件下的动作策略

四种设计可组合，排列顺序不代表自治等级；例如 VLA 出动作、传统控制器执行。

6.6.1 自动驾驶与外部控制：系统三模块

传感、计算、执行构成完整任务链

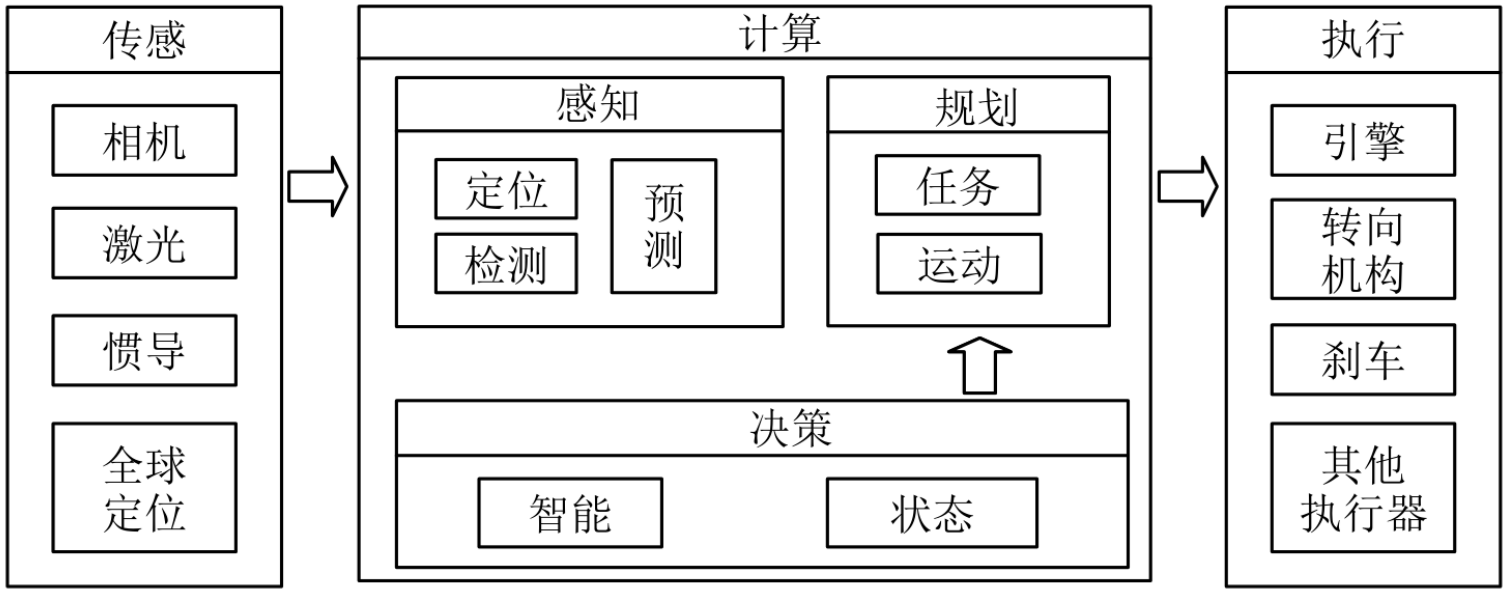


图6.7 开源自动驾驶系统框架

定位： 自身在哪
检测： 周围有什么
预测： 将怎样运动
决策： 状态机选行为
规划： 任务→运动

无地图时，即时定位与地图构建（SLAM）联合估计位姿与地图。

6.6.1 自动驾驶与外部控制：开源软件栈分层

Autoware 将规划与跟踪控制分开

工程模块	主要输入	向外提供
传感；地图	原始图像、点云；道路几何	观测流与地图信息
定位；感知	观测、地图与状态信息	自身位姿；目标与预测
规划	路线目标、状态和环境	当前可执行的轨迹
控制	轨迹与实际状态	转向与纵向控制命令
车辆接口	统一控制命令；车辆状态	底盘命令与车辆反馈

任务规划确定目的地与路线；运动规划生成轨迹；控制模块持续修正跟踪偏差。

6.6.1 自动驾驶与外部控制：跨平台对照

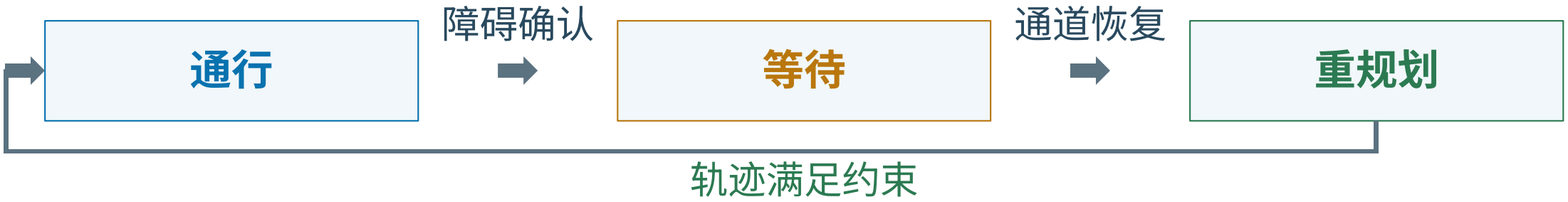
相同接口分工，不同物理动作

比较维度	汽车外部控制	多旋翼外部控制
常见运动目标	路线、速度、转向要求	轨迹、位置、速度或姿态
平台适配	转向与纵向命令 → 底盘协议	外部目标 → 飞控接受的指令
反馈证据	实际车速、转角、控制模式	位姿速度、解锁（arm）与模式
底层职责	驱动、制动、转向的闭环响应	姿态、角速度与执行器响应

接口统一语义，控制器承担具体载具的响应。

6.6.1 自动驾驶与外部控制：任务闭环

发现障碍后，行为与轨迹一起更新



有限状态机：用状态与触发条件组织任务行为；条件标在箭头上。

教学小车：初速 2 m/s，恒定减速度幅值 1 m/s²，忽略延迟。

$$t_{\text{stop}} = \frac{v_0}{a} = \frac{2}{1} = 2\,\text{s} \qquad d_{\text{stop}} = \frac{v_0^2}{2a} = \frac{2^2}{2\times 1} = 2\,\text{m}$$

反馈判据： 实际速度降至停止阈值；通道占用状态与定位数据保持有效。

实际留量： 制动距离之外，还包含延迟行程、状态误差和载具响应余量。

本章目录

前沿技术与控制方法

第六章 外部控制与轨迹规划

6.1 外部控制基础与控制分层
6.2 外部控制原理与控制模式
6.3 轨迹规划建模与生成
6.4 轨迹规划算法与避障
6.5 轨迹跟踪原理与高机动控制
6.6 前沿技术与控制方法
6.7 平台框架与通信接口
6.8 外部控制例程与实验验证
6.9 习题与课堂讨论

本节 4 个教学单元

6.6.1 自动驾驶与外部控制
教材对应：第6.1.4.1节

6.6.2 端到端与强化学习
教材对应：第6.1.4.2节

6.6.3 大模型任务与工具调用原理
教材对应：第6.1.4.3节

6.6.4 VLA动作生成与适配
教材对应：第6.1.4.4节

6.6.2 端到端与强化学习：输入输出边界

端到端首先要说清两端是什么

输入可以是图像或状态；例如“图像→轨迹”，再由传统控制器执行。

输出端	保留的执行链	需要学会或验证的内容
轨迹或位置	传统跟踪控制与稳定内环	轨迹可行性、目标更新及延迟
速度、加速度或姿态	部分底层闭环	动作尺度、坐标、响应与限幅
执行器量	执行器与物理载具	快速稳定性、耦合和硬约束

端到端描述学习边界，接口层级决定控制责任。

6.6.2 端到端与强化学习：收益与训练方式

联合学习的收益来自任务数据

训练方式	数据里提供什么	主要取舍
模仿学习	观测与示范动作配对	采集质量重要； 偏离示范后易缺样本
强化学习（RL）	观测、动作、奖励与后继观测	奖励设计与探索成本重要； 不要求环境可求导
组合方式	先示范，再用交互改进	继承已有技能； 再按任务修正策略

教材观点：联合学习可把系统视为可微分整体，减少手工模块造成的信息损失；此类RL无需对环境动力学求导；收益仍由任务与数据决定。

6.6.2 端到端与强化学习：学习与推理

DDPG 在训练中学策略，执行时用策略

深度确定性策略梯度（DDPG）：面向连续动作的 Actor–Critic 方法。



经验回放：保存交互转移；Critic 学动作价值，Actor 沿价值上升方向更新。

训练回报是优化目标，任务完成由闭环证据判断。

6.6.2 端到端与强化学习：例程结构

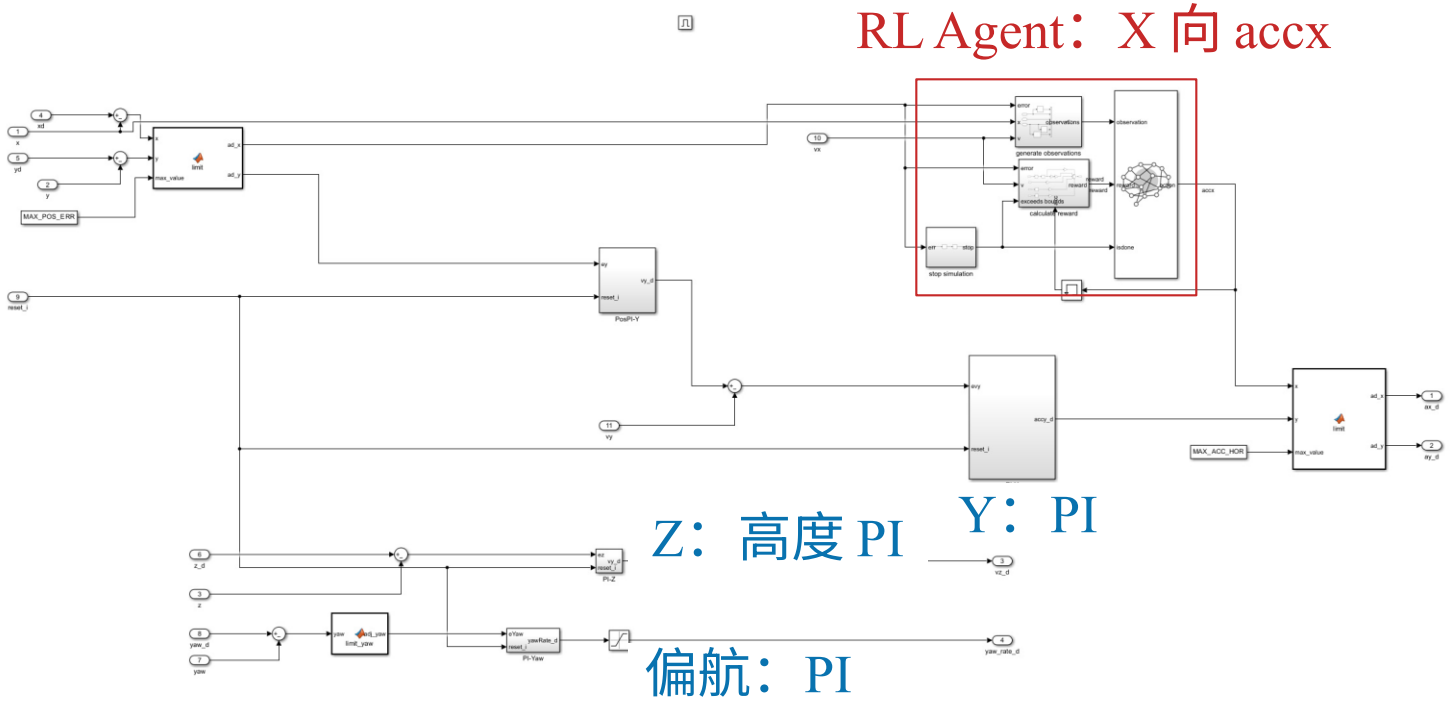
实际例程只用 RL 替换 X 向位控

目录：[安装目录]/RflySimAPIs/6.RflySimExtCtrl/2.AdvExps/e5.RLExps/1.Simulink/
训练入口 train/rluav_init.m；验证模型 rlhighDynamicAcc.slx。

接口	实际变量与含义	规格
观测	误差积分、误差、位置、速度、 速度积分支路、速度微分	6 维；顺序和尺度固定
动作	X 向加速度 accx	1 维；范围 −4～4 m/s^2
执行组合	X 轴策略＋Y/Z/Yaw 常规控制	采样 0.02 s 其余闭环保留

6.6.2 端到端与强化学习：例程模型

RL Agent 块只接管 X 通道



图：位置控制子系统（e5 例程 Readme 截图）

X 通道： RL Agent 输出 accx，经限幅后作期望加速度。

Y、Z、偏航： 仍由常规 PI 环控制。

观测、奖励与终止信号由三个子模块生成；采样 0.02 s。

6.6.2 端到端与强化学习：奖励算一步

奖励如何鼓励到达并抑制运动

$$r_k = 10I_k - (1 - I_k) - 100B_k - 0.5|v_k| - 0.5|a_{k-1}|$$

$I_k = 1$ 表示 $|e_k| < 0.1\,\mathrm{m}$；$B_k = 1$ 表示 $|e_k| > 8\,\mathrm{m}$。
其余时取 0；e_k 为位置误差，v_k 为速度，a_{k-1} 为上一拍加速度动作。

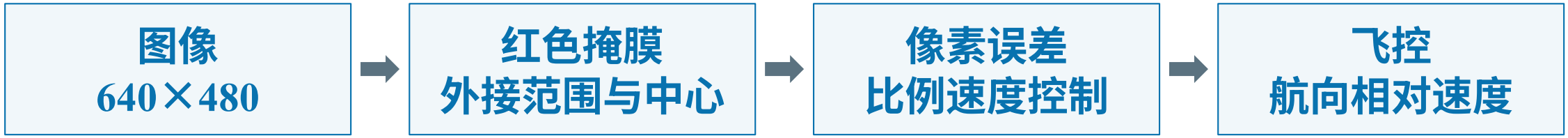
误差 / m	速度 / (m/s)；上一动作 / (m/s²)	单步奖励
0.08	0.2；0.4；未越界	$10 - 0.1 - 0.2 = 9.7$
0.12	0.2；0.4；未越界	$-1 - 0.1 - 0.2 = -1.3$
9.00	0.2；0.4；越界终止	$-1 - 100 - 0.1 - 0.2 = -101.3$

结论：阈值两侧奖励相差 11，阈值主导策略偏好；奖励高也不证明运动平滑。

6.6.2 端到端与强化学习：传统视觉控制对照

例程二：HSV 跟随采用视觉比例闭环

HSV 是色相、饱和度、明度表示；脚本用阈值分割红色目标。



入口：6.RflySimExtCtrl/2.AdvExps/e6.IsaacSimExps/RL_attctrl/circle_follow.py
输出：SendVelFRD(vx, vy, vz, 0)，水平随yaw、NED下向（m/s）；偏航率rad/s。

算法边界：HSV 阈值＋比例反馈，无学习权重；丢目标时只跳过该帧。

同目录 rl.py：学习策略由 12 维状态输出姿态与推力（20 Hz），不读图像。

6.6.2 端到端与强化学习：HSV 视觉控制算一步

像素误差怎样变成速度指令

给定：图像中心 (320, 240) px，目标中心 (360, 220) px；
初始像素高度 100 px，当前 80 px。三项误差均超过 15 px 死区。
源码比例增益：横向 0.009、纵向 0.005、前后 0.015，单位 m/s 每像素。

$$v_y = 0.009(360 - 320) = 0.36 \text{ m/s}$$

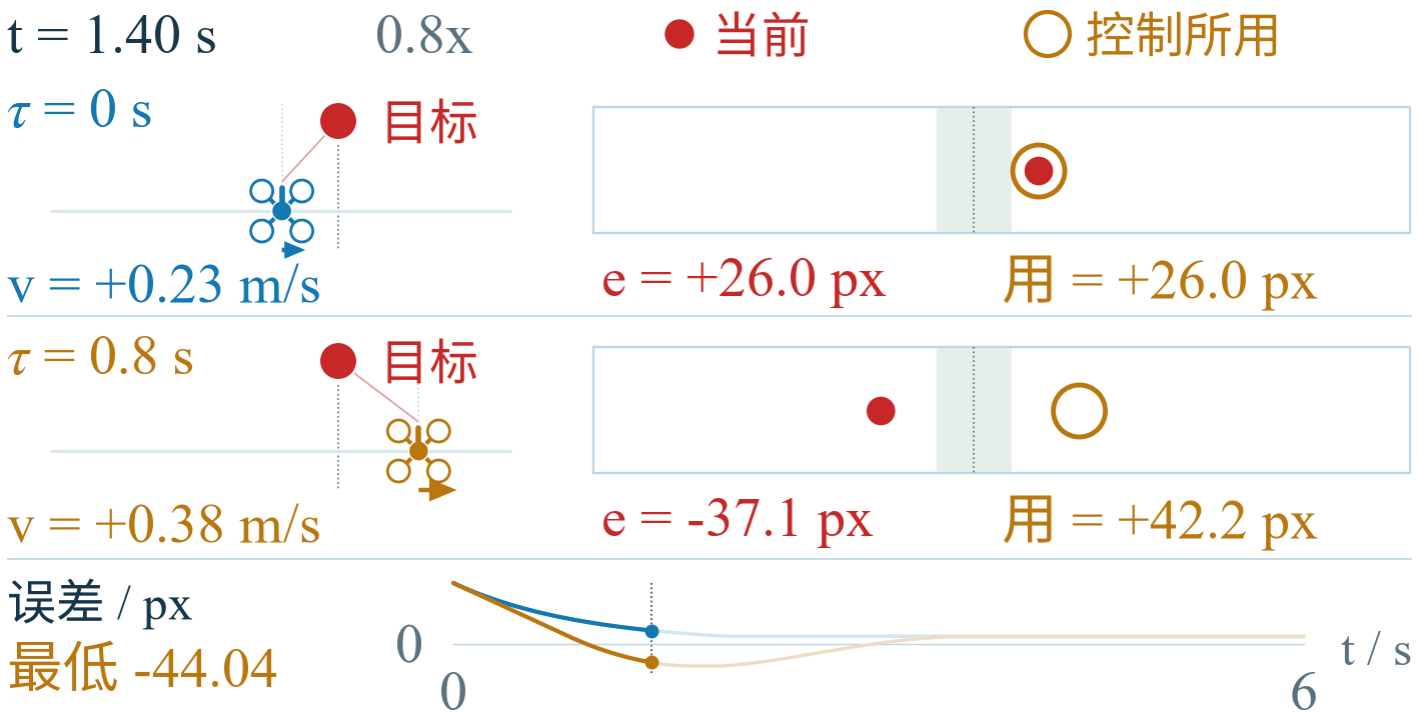
$$v_z = 0.005(220 - 240) = -0.10 \text{ m/s}$$

$$v_x = 0.015(100 - 80) = 0.30 \text{ m/s}$$

近水平时按FRD理解：本配置水平随yaw，z沿NED下向；负值上升。
像素高度作为距离线索，依赖目标尺寸与朝向近似不变。

6.6.2 端到端与强化学习：HSV 跟随的延迟对比

相同控制律，延迟改变完整响应



图：旧图像造成跟随过冲（教学计算）

固定目标、理想响应。
比例增益 0.009。
死区：±15 px。

无延迟：末值 14.75 px。
延迟 0.8 s：
最低 −44.04 px。

$$e_{k+1} = e_k - 120 v_{y,k} \Delta t$$

像素尺度 120 px/m；
步长 0.02 s；初值 120 px。
未包含飞行动力学。

6.6.2 端到端与强化学习：姿态层策略

例程三：rl.py 由状态给出姿态与推力

RL_attctrl/rl.py：加载 Isaac Lab 训练的策略权重，每 0.05 s 推理一次（20 Hz）。

坐标：$v^e = Rv^b$，　$v^b = R^{\mathrm{T}}v^e$（b：机体FRD，e：地面NED）；本安装版反向误乘R。

12 维观测 状态与目标 → **Actor 网络 64-64，ELU** → **tanh 输出 4 维，范围 ±1** → **SendAttPX4 姿态角与推力**

```
roll_cmd = action[3] * 10;  thrust_cmd = hover_thrust - 0.1 * action[0]
mav.SendAttPX4([roll_cmd, pitch_cmd, yaw_cmd], thrust_cmd)  # 角度单位：度
```

观测目标：机体速度、角速度、重力投影、目标相对位置各3维。

本安装版 uavAngVel 与 uavAngRate 不匹配，角速度3维实际补零。

任务：固定目标 [100,50,−30] m，不读图像；训练需 Isaac Lab 与 GPU。

三例分别输出 X 向加速度、航向相对速度、整机姿态与推力。

6.6.2 端到端与强化学习：数据与泛化

学习策略的五类问题各有检查证据

问题	直接检查的证据
数据质量	图像、状态、动作的时间对齐；异常和失败样本覆盖
训练收敛	多次训练回报分布；同任务闭环误差与动作幅度
可解释性	输入观测、策略动作、限幅和接管原因的日志
未见场景	独立的光照、对象、初值与扰动测试集
仿真到现实差距	视觉变化、计算延迟、传感噪声与动力学偏差

本章目录

前沿技术与控制方法

第六章 外部控制与轨迹规划

6.1 外部控制基础与控制分层
6.2 外部控制原理与控制模式
6.3 轨迹规划建模与生成
6.4 轨迹规划算法与避障
6.5 轨迹跟踪原理与高机动控制
6.6 前沿技术与控制方法
6.7 平台框架与通信接口
6.8 外部控制例程与实验验证
6.9 习题与课堂讨论

本节 4 个教学单元

6.6.1 自动驾驶与外部控制
教材对应：第6.1.4.1节

6.6.2 端到端与强化学习
教材对应：第6.1.4.2节

6.6.3 大模型任务与工具调用原理
教材对应：第6.1.4.3节

6.6.4 VLA动作生成与适配
教材对应：第6.1.4.4节

6.6.3 大模型任务与工具调用原理：模型职责

语言理解与视觉理解不等于动作执行

大语言模型（LLM）
文本：理解任务，组织步骤或工具参数

多模态大语言模型（MLLM）
语言＋图像、声音等：结合现场解释任务

视觉语言模型（VLM）
图像＋语言：
定位语义目标、
回答视觉问题

视觉基础模型（VFM）
视觉输入：
提供可复用的
视觉表示

组合关系：VFM提供视觉表示，可与语言模型组合；VFM本身不必含语言模型。

动作输出层
或工具执行层
把理解变成运动

四类模型提供理解与表示；产生运动还要经过执行层。

6.6.3 大模型任务与工具调用原理：两类架构

离线生成与在线响应有不同时间约束

维度	离线任务生成	在线指令响应
产物	行为树：按条件执行动作； 状态机或控制代码	当前任务选择或工具参数
检查时机	部署前检查并修订	每次执行前检查状态与参数
适用任务	边界明确、流程可分解	根据新观测改变任务步骤
主要限制	覆盖异常分支与平台接口	推理时延、网络、观测过期

模型表达任务意图，执行器需要可检验的动作条件。

6.6.3 大模型任务与工具调用原理：源码反例

“向前飞”需要坐标和完成条件

模型回复示例： ["Forward_Backward_Fly", 5, 1]
5是距离（m）；1是约定的无人机编号，本分支未用它选机。

```
target_pos = [current_pos[0] + 5, current_pos[1], current_pos[2], 0]
SendPosNED(x, y, z, yaw)
if pos_error < 0.1 or elapsed_time > command_timeout:  # 原完成条件
```

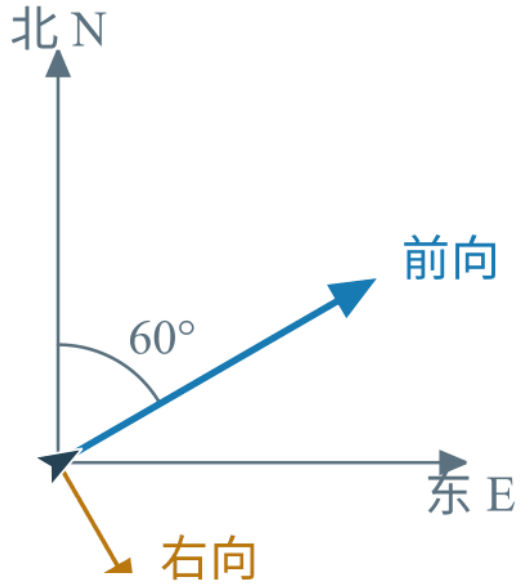
坐标： SendPosNED 使用北东地坐标；北向加 5 m 并不总是沿机头前进。

原条件缺陷： 超时也记为完成；纯转向只看位置误差，转向未完也会判完成。
改法： 位置、偏航、超时分别判断并单独返回；回复文本不作飞行终态。

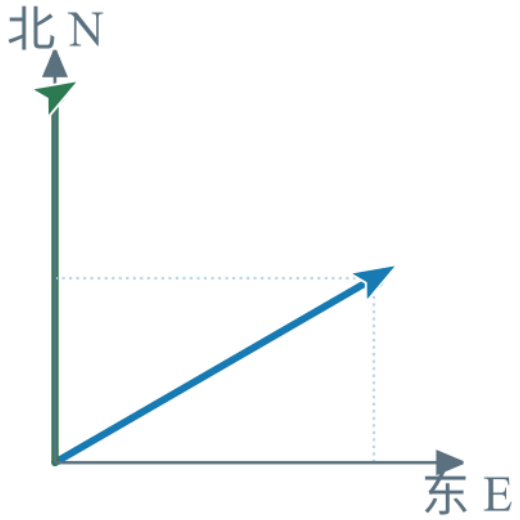
6.6.3 大模型任务与工具调用原理：位移坐标对照

机头前向 5 m，不等于地面北向 5 m

固定偏航 60°



同一起点，不同终点



完成比例 100%

机头前向 5 m
$\Delta N = 2.50$ m
$\Delta E = 4.33$ m

地面北向 5 m
$\Delta N = 5.00$ m
$\Delta E = 0.00$ m

图：固定偏航 60° 比较位移方向；原代码另设偏航目标为 0（坐标示意）
坐标约定：地面北东地（NED）、机体前右下（FRD）；偏航从北向东为正。
前向 5 m：$\Delta N = 5\cos 60^\circ = 2.50$ m，$\Delta E = 5\sin 60^\circ = 4.33$ m。

6.6.3 大模型任务与工具调用原理：工具接口约定

受限工具把任务参数交给执行器

自然语言任务 ➡ 受限工具参数 ➡ 参数与状态检查 ➡ 已验证控制器

教学工具请求：goto_ned(request_id="A7", x_m=5, y_m=0, z_m=-2,
yaw_rad=0, timeout_s=20)
执行前：工具白名单；有限数值；范围；坐标；定位与控制状态；时间戳。

返回 accepted：表示请求通过检查；记录请求编号、目标和接受时刻。

返回 succeeded：需由实际位姿、速度、持续到达时间和新鲜反馈共同支持。

接口约定：模块间约定的输入输出、单位、坐标系、频率与时序。

6.6.3 大模型任务与工具调用原理：工具判定算例

工具输出必须区分成功、超时和过期

教学判据：位置误差 <0.1 m、速度 <0.1 m/s、偏航误差 <5°，持续 1 s；
反馈年龄（当前时刻−反馈时间戳）≤0.2 s，执行期限 20 s。阈值是本算例设定。
判定顺序：过期 → 超时 → 到达；各行独立，未标注时反馈年龄 0.04 s、已执行 5 s。

输入假数据	判定输出	原因
0.08 m；0.04 m/s；偏航 1.7°；持续 1.2 s	succeeded	各项满足且持续 1 s 以上
1.2 m；已经过 21 s	timeout	到期仍未到达
0 m；反馈年龄 0.7 s	stale	反馈太旧，不能确认状态
0.02 m；偏航误差 30°	running	位置到达，方向仍未完成

本章目录

前沿技术与控制方法

第六章 外部控制与轨迹规划

6.1 外部控制基础与控制分层
6.2 外部控制原理与控制模式
6.3 轨迹规划建模与生成
6.4 轨迹规划算法与避障
6.5 轨迹跟踪原理与高机动控制
6.6 前沿技术与控制方法
6.7 平台框架与通信接口
6.8 外部控制例程与实验验证
6.9 习题与课堂讨论

本节 4 个教学单元

6.6.1 自动驾驶与外部控制
教材对应：第6.1.4.1节

6.6.2 端到端与强化学习
教材对应：第6.1.4.2节

6.6.3 大模型任务与工具调用原理
教材对应：第6.1.4.3节

6.6.4 VLA动作生成与适配
教材对应：第6.1.4.4节

6.6.4 VLA动作生成与适配：两阶段机制

VLA 把图像与语言映射为动作

图像＋语言指令 ➡ **视觉语言编码器** ➡ **动作解码器**

观测内容与任务目标
共同作为策略条件。

共享表示保留与动作
任务相关的特征。

离散符号或连续向量
恢复成平台动作。

教材所述优势：统一感知、推理与控制，借大规模数据泛化到新任务与平台。

训练记录：图像、语言、状态与动作按时刻对齐，动作定义与实际执行器一致。

部署链路：动作解码 → 单位与坐标转换 → 约束检查 → 底层控制 → 新观测。

6.6.4 VLA动作生成与适配：OpenVLA 动作接口

OpenVLA 的七维动作不是无人机油门

OpenVLA（2024 年开放研究模型）：图像＋语言 → 动作符号 → 连续机器人动作。

七维动作示例	物理含义	接入时要匹配
3 个平移分量	末端执行器的位置增量	米、坐标系、增量或绝对量
3 个旋转分量	末端旋转增量	角度表示、单位、旋转顺序
1 个夹爪分量	夹爪开合	动作定义和平台执行器

源码参数 unnorm_key 选择动作归一化统计；换平台必须匹配数据与动作接口。

6.6.4 VLA动作生成与适配：适配算一步

动作解码必须恢复单位与坐标

教学假设：策略已用飞行数据训练；输出区间为[−1, 1]，本次输出0.4。
该维训练统计下界−0.05 m、上界0.05 m；线性恢复前向位移。

$$\Delta x_b = \frac{0.4 + 1}{2}(0.05 - (-0.05)) - 0.05 = 0.020\,\mathrm{m}$$

水平机体系前向增量转换到北东地坐标；偏航角 30°：

$$\Delta x_N = 0.020\cos 30^\circ = 0.0173\,\mathrm{m}, \quad \Delta y_E = 0.020\sin 30^\circ = 0.0100\,\mathrm{m}$$

动作接口约定还包含：执行周期、相对/绝对目标、角度单位、限幅与过期处理。

6.6.4 VLA动作生成与适配：动作分块策略的教学时序

动作块覆盖时间不等于允许延迟

动作块：模型一次输出的连续多步动作；块内各步都基于同一帧观测。



覆盖时间：5 步 × 0.1 s = 0.5 s；第1步在 0.18 s 才执行，依据的仍是 0 时刻图像。

等待行程：2 m/s × 0.18 s = 0.36 m，余量要容纳这段状态变化。

闭环更新：0.20 s 采集新图，0.38 s 新动作块接替；按任务设动作有效期与接管条件。

6.6.4 VLA动作生成与适配：评价与趋势

技术趋势落到数据、时延与迁移能力

方向	需要的数据或条件	对照证据
语义到动作	语言、图像、状态、动作对齐	改写指令与新目标测试
高效动作生成	动作块、并行解码和观测更新	总延迟、过期动作与接管次数
平台迁移	动作尺度、相机； 控制器与动力学	新平台微调后的独立测试
任务可靠性	失败样本与运行约束	成功率、误差、碰撞/越界记录

学习策略可与规则规划器并行；约束检查决定接受、重规划或接管。

学习资源

按架构、算法与接口延伸阅读

自动驾驶架构与车辆接口：Autoware 官方文档

https://docs.autoware.org/main/design/

连续动作的策略与价值学习：DDPG 原始论文

https://arxiv.org/abs/1509.02971

语言任务与可执行技能：SayCan 研究项目

https://say-can.github.io/

输入格式、动作解码与迁移：OpenVLA 官方源码

https://github.com/openvla/openvla

课堂练习

用输入、动作和结果辨认控制机制

问题	参考要点
HSV 像素误差 → 速度，是否属于 RL/VLA?	不属于；显式检测与比例反馈，无奖励训练和语言动作映射。
DDPG 只输出 accx，替代了什么?	替代 X 向位控映射；其余轴和底层闭环保留。
模型说“成功”，工具返回 timeout?	任务记为超时；以遥测、期限和终态为准。
机械臂七维动作如何直接接飞行?	重新定义动作、尺度、坐标与周期，并用匹配的飞行数据训练。

本节小结

新算法仍需形成可验证的运动闭环

模块化：用感知、规划、控制与平台接口分担任务，便于替换与逐层检查。

学习控制：先定义观测、动作、奖励或示范；单轴学习不代表整机端到端。

大模型与 VLA：语义与动作能力必须接到有坐标、单位、周期和条件的接口。

验证：用实际反馈、约束与终态评估；工具超时不算成功。

易错点：把目录名当机制、把训练回报当控制效果、把模型文字当工具动作。

前沿方法扩展任务能力，可信执行依赖可验证的接口。