本章目录

定位建图与预测规划

第八章 多模态感知与智能决策

8.1 背景意义与图像处理基础

8.2 目标识别、跟踪与光流估计

8.3 定位建图与预测规划

8.4 常用感知传感器

8.5 多模态感知坐标转换

8.6 多模融合与控制决策

8.7 关键工具与框架接口

8.8 实验案例与真机迁移

8.9 习题与课堂讨论

本节 4 个教学单元

8.3.1 联合定位与建图
教材对应：第8.1.2.4节

8.3.2 行为预测与交互建模
教材对应：第8.1.2.5节

8.3.3 路径规划与动态避障
教材对应：第8.1.2.6节

8.3.4 多机融合感知与决策
教材对应：第8.1.2.7节

8.3.1 联合定位与建图：联合估计

目标在哪里之外，还要知道自己在哪里

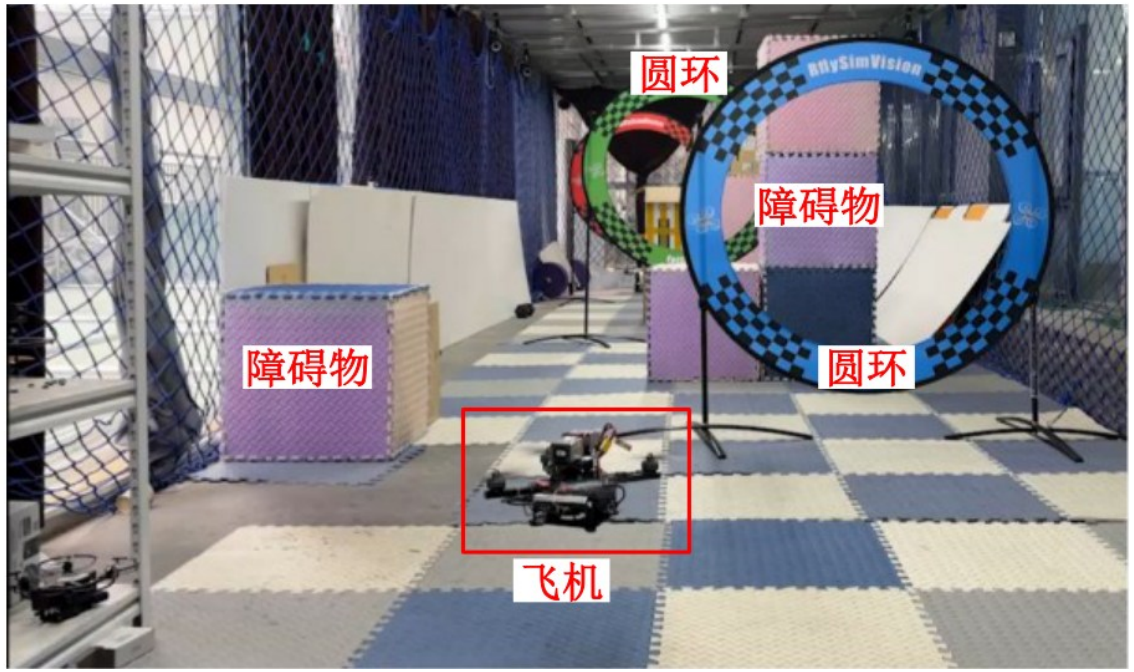
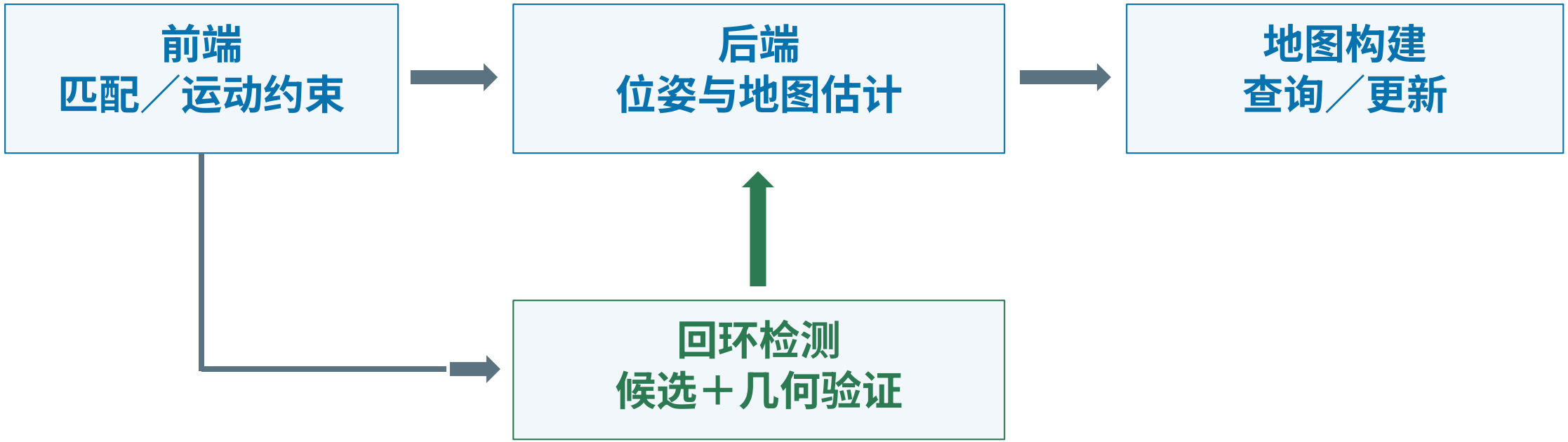


图8.3(a) 目标、障碍与自身

- **SLAM：**同时定位与建图，联合估计自身位姿和环境地图。
- **VO：**视觉里程计，主要输出连续运动；可维护局部特征图。
- **任务区别：**目标框属于当前图像；自身位姿属于约定的地图坐标。

8.3.1 联合定位与建图：四个核心模块

前端给约束，后端维护一致性



前端给局部关系；回环给跨时间关系；后端共同调整历史位姿与地图。

8.3.1 联合定位与建图：传感器路线

视觉匹配特征，激光匹配几何



图8.3(d) ORB-SLAM特征

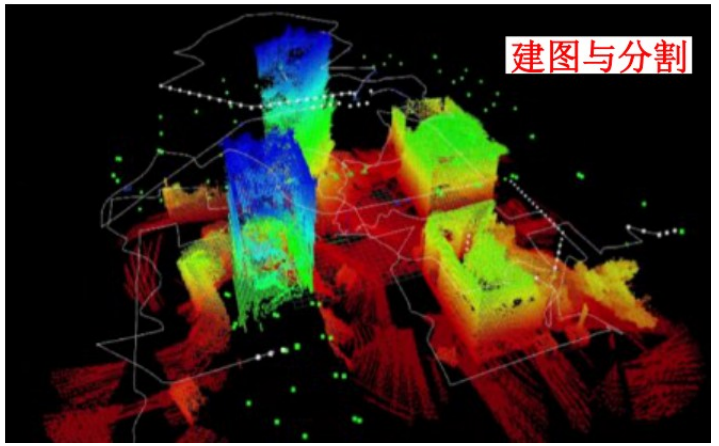


图8.3(f) FAST-LIO地图

视觉SLAM：图像特征／像素匹配。
ORB-SLAM用稀疏特征建立约束；弱纹理与模糊会减少有效观测。

激光SLAM：扫描几何匹配。
Cartographer：扫描与子图。
FAST-LIO：激光惯性里程计。

8.3.1 联合定位与建图：惯性融合

高频传播与几何观测互相校正

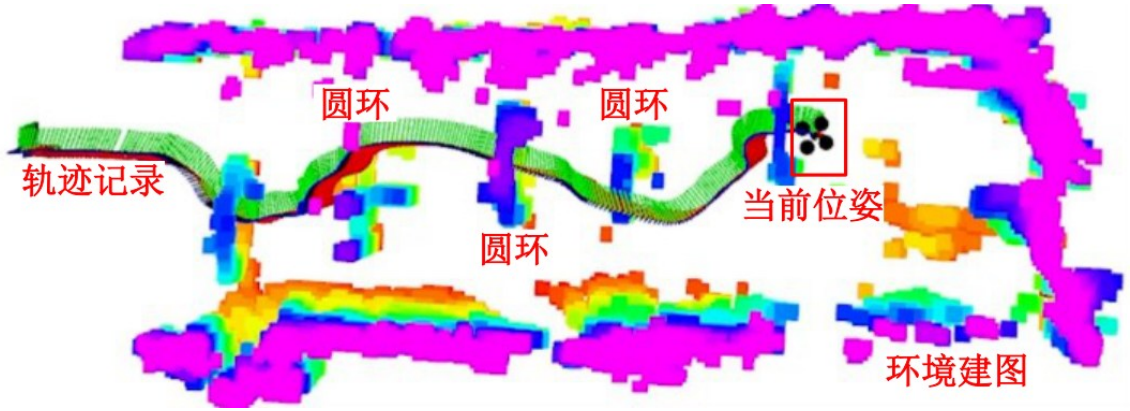


图8.3(c) 位姿、地图与轨迹

组合	含义
VIO／VINS	视觉惯性里程计 视觉惯性导航系统
LIO	激光惯性里程计
LVI	激光、视觉与惯性联合

惯性测量单元（IMU）：高频传播；视觉／激光观测抑制漂移。
VINS-Mono：单目＋IMU；VINS-Fusion：单目／双目＋IMU或纯双目。
频率示例：几何观测10–30Hz，IMU100–1000Hz；配置随算法设置。

8.3.1 联合定位与建图：地图表示

定位、净空和语义需要不同地图

表示	保存的信息	服务的任务
稀疏特征点图	特征位置、描述子与观测关系	匹配定位；不能直接查询净空
稠密占据栅格	空闲、占据、未知及占据概率	碰撞查询；细分辨率增加存储
语义地图	几何位置＋类别／实例／区域	可通行判定、对象级任务

地图表达要和查询匹配：特征匹配、占据检查与对象语义不能互相替代。

8.3.1 联合定位与建图：两类距离表示

TSDF重建表面，ESDF查询净空

TSDF：截断有符号距离函数；ESDF：欧氏有符号距离场。
x为查询点，**q**为障碍边界$\partial\mathcal{O}$上的点；距离单位m。

表示	几何含义	直接用途
TSDF	射线方向的截断表面距离	零交叉面重建／融合深度
ESDF	到最近障碍边界的欧氏距离	净空查询／轨迹避障约束

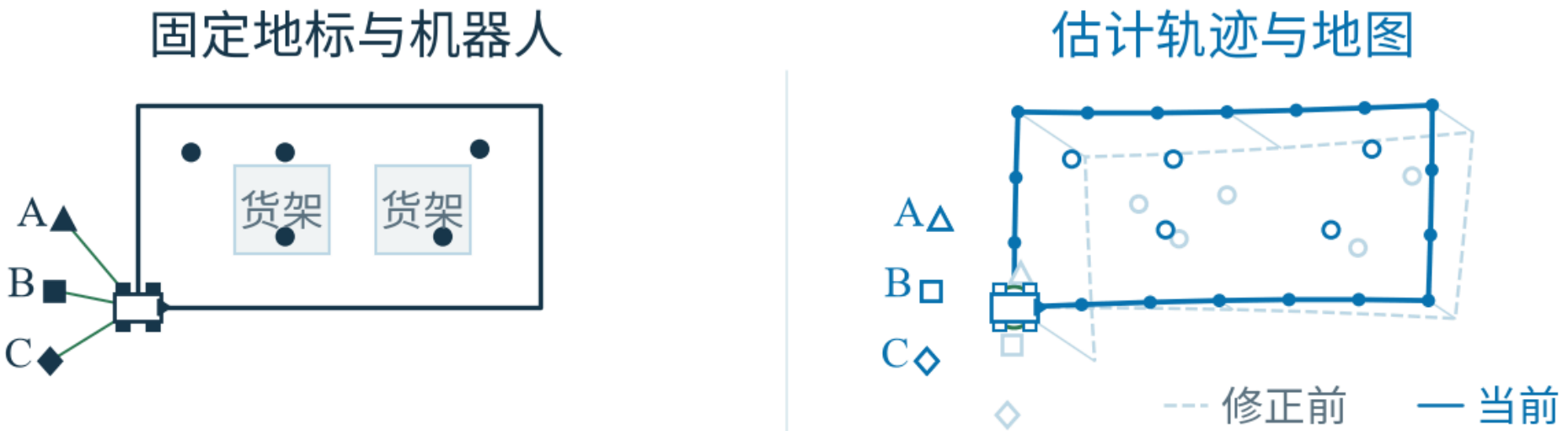
$$d_{\mathrm{ESDF}}(\mathbf{x}) = \pm \min_{\mathbf{q}\in\partial\mathcal{O}} \|\mathbf{x}-\mathbf{q}\|$$

障碍外正、内部负，边界为零；未知空间不直接当成空闲。

8.3.1 联合定位与建图：回环检测

重访约束先核验，再修正整条轨迹

接受回环：轨迹与地图共同修正



图：同位姿重访算例；核验通过后，轨迹与地图共同修正。

8.3.1 联合定位与建图：回环一步算例

四段里程计累计0.4米闭合误差

输入增量：10.1、10.0、−9.8、−9.9m；总和＋0.4m。
固定起点；四边等方差；回访起点作为精确闭合约束。

$$\min_{\delta_1,\ldots,\delta_4} \sum_{i=1}^{4} \delta_i^2 \quad \text{s.t.} \quad \sum_{i=1}^{4} \delta_i = -0.4$$

$$\delta_i = -\frac{0.4}{4} = -0.1\ \text{m}$$

δ_i为第i段增量修正（m）；四段误差相互独立。
修正后增量：10.0、9.9、−9.9、−10.0m；和为0，四边共同分配误差。

8.3.1 联合定位与建图：全轨迹结果

闭合约束同时改变中间位姿

时刻	观测累积位姿／m	闭合后位姿／m
0	0	0
1	10.1	10.0
2	20.1	19.9
3	10.3	10.0
4	0.4	0

$$J = 4 \times (0.1)^2 = 0.04\ \mathrm{m}^2$$

回环改善相对一致性；绝对坐标仍需要参考锚点。

8.3.1 联合定位与建图：SLAM与卫星定位

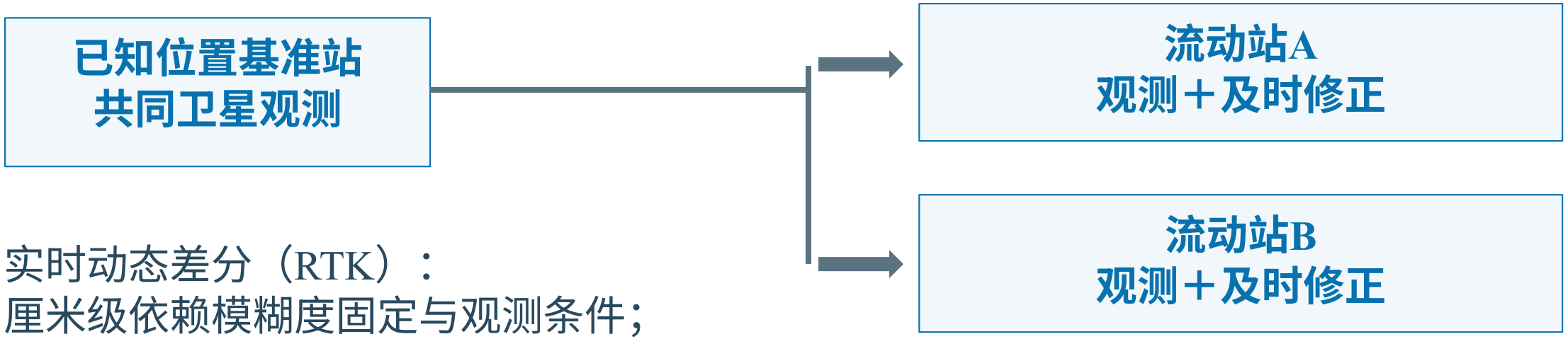
局部一致地图仍需要全局参考

维度	SLAM局部参考	GNSS全球参考
信息来源	场景相对观测与历史约束	卫星观测及坐标基准
误差特点	初始化、尺度与累积漂移	遮挡、多路径与观测几何
适用环境	室内／隧道 需有效纹理或几何	开阔室外 需可用卫星观测

全球导航卫星系统（GNSS）给出全球坐标参考。
map是约定的地图参考；body是当前机体系；共享位置前还要统一时刻。

8.3.1 联合定位与建图：RTK与航向

差分修正与天线基线给出不同信息



实时动态差分（RTK）：
厘米级依赖模糊度固定与观测条件；
一座基准站可支持多个流动站。

基线分量：Δx北向、Δy东向（m）。
北为零，向东为正，单位rad。

$$\psi_{\text{baseline}} = atan2(\Delta y, \Delta x)$$

双天线给基线航向；机体偏航还需安装外参，卫星多路径仍要处理。

8.3.1 联合定位与建图：路线与融合条件

纹理、几何与实时负载决定取舍

路线	可观察信息	退化／代价
视觉为主	颜色、纹理、类别与几何线索	弱纹理、模糊、光照；训练和计算
激光融合	直接距离与空间几何	遮挡、雨雾、几何退化；融合开销

融合检查：时间同步、空间外参、算力负载、故障切换、成本。

$v = 1.0\ \mathrm{m/s}$，$\Delta t = 50\ \mathrm{ms} = 0.050\ \mathrm{s}$；$\Delta p$为位移偏差。

$$\Delta p = v\Delta t = 1.0 \times 0.050 = 0.050\ \mathrm{m}$$

教材路线举例：Tesla视觉；百度、蔚来、小米采用传感器融合。

8.3 定位建图与预测规划

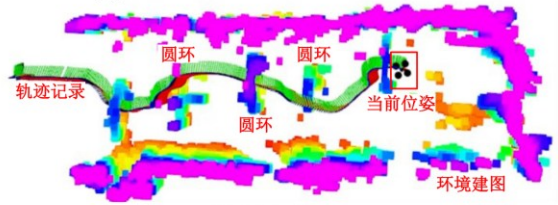
8.3.1 联合定位与建图：教材全图关系

同一场景产生不同层次的算法输出



(a) 真机场景（穿环避障）

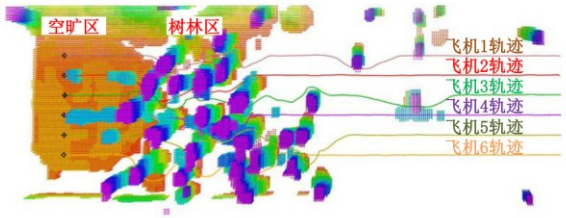
(b) YOLO圆框目标识别



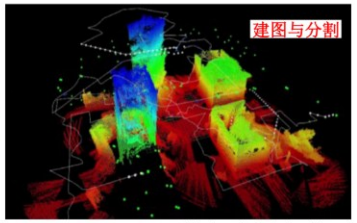
(c) Vins-Fusion视觉定位+Ego-Planner规划



(d) ORB-SLAM特征提取与定位



(e) Ego-Swarm多机避障规划



(f) FAST-LIO激光建图定位

- **a／b：** 真机场景／目标检测
  检测细节见8.2第10页。

- **c／d／f：** 视觉惯性定位与规划／视觉特征／激光惯性地图。

- **e：** 多机规划轨迹
  本节第46页放大讲解。

本章目录

定位建图与预测规划

第八章 多模态感知与智能决策

8.1 背景意义与图像处理基础

8.2 目标识别、跟踪与光流估计

8.3 定位建图与预测规划

8.4 常用感知传感器

8.5 多模态感知坐标转换

8.6 多模融合与控制决策

8.7 关键工具与框架接口

8.8 实验案例与真机迁移

8.9 习题与课堂讨论

本节 4 个教学单元

8.3.1 联合定位与建图
教材对应：第8.1.2.4节

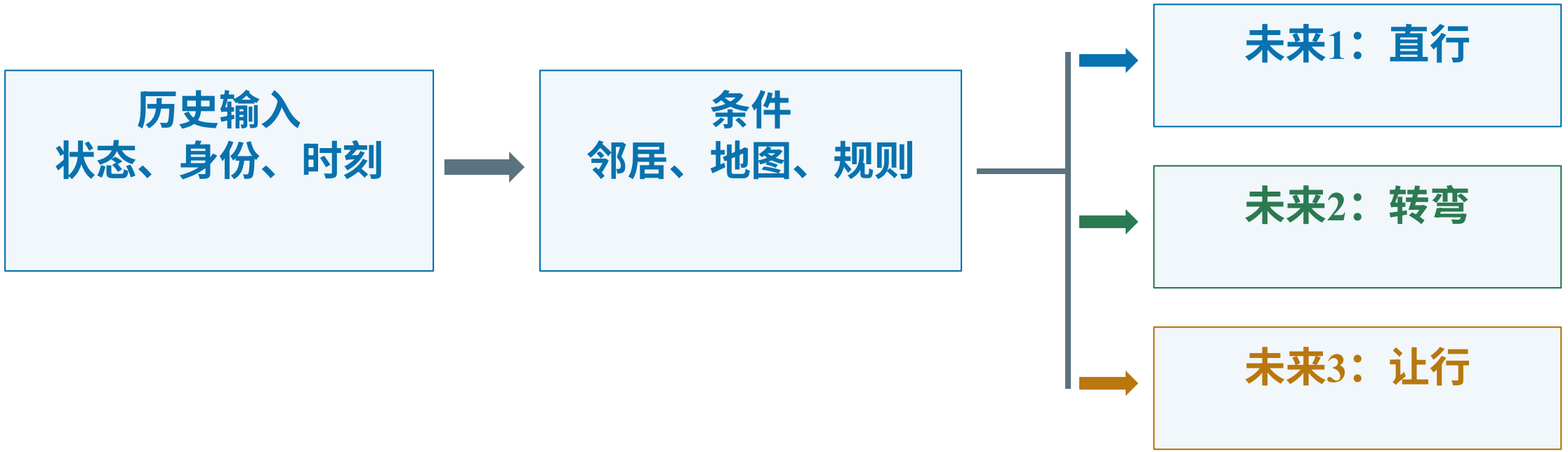
8.3.2 行为预测与交互建模
教材对应：第8.1.2.5节

8.3.3 路径规划与动态避障
教材对应：第8.1.2.6节

8.3.4 多机融合感知与决策
教材对应：第8.1.2.7节

8.3.2 行为预测与交互建模：输入与预测时域

预测输出是一组未来，输入是历史



预测时域说明向前看多久；多种未来表示可能分支，权重需与实际结果比较校准。

8.3.2 行为预测与交互建模：时间序列模型

隐藏状态积累历史，门控控制记忆

模型	中文定义	状态机制
RNN	循环神经网络	上一隐藏状态＋当前观测递推
LSTM	长短期记忆网络	记忆单元及输入／遗忘／输出门
GRU	门控循环单元	更新门／重置门控制状态

$$\mathbf{h}_t = f(\mathbf{h}_{t-1}, \mathbf{x}_t)$$

t为观测序号，$\mathbf{x}_t$为当前观测，$\mathbf{h}_t$为隐藏状态，f为学习的更新函数。

8.3.2 行为预测与交互建模：一步预测算例

匀速基线把最近速度延伸到未来

历史时刻／s	0	1	2
水平位置／m	0	1	2

$$\hat{v} = \frac{2-1}{2-1} = 1\ \mathrm{m/s}$$

$$\hat{x}(2+2) = 2 + 1 \times 2 = 4\ \mathrm{m}$$

完整基线结果：未来1s为3m，未来2s为4m；条件是速度保持不变。

8.3.2 行为预测与交互建模：多目标交互

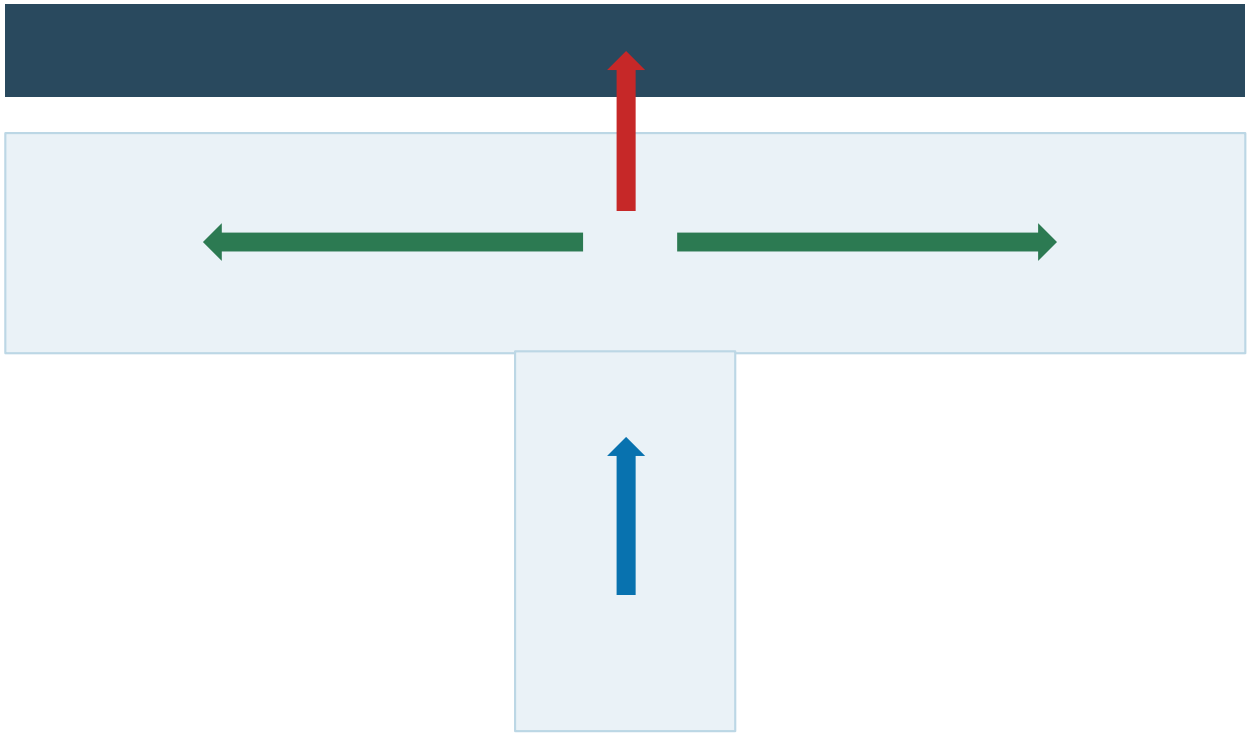
邻居的状态会改变个体的未来

机制	输入关系	汇聚方式
Social-LSTM	空间邻域中的历史状态	社交池化邻居隐藏状态
GNN／GAT	对象节点与交互边	图神经网络／图注意力网络 汇聚邻居，注意力用于加权
GCN	图结构与节点状态	图卷积网络聚合邻居信息

交互模型改变未来分布；邻居权重或预测轨迹不是碰撞安全证明。

8.3.2 行为预测与交互建模：场景约束

可行未来要留在可通行区域内



绿：可通行分支；红：越墙候选

- **几何：** 候选不能穿越障碍。
- **规则：** 区域和通行方向限制行为。
- **动力学：** 速度、加速度与转弯仍需满足约束。

8.3.2 行为预测与交互建模：多未来与不确定性

参数不确定性和行为分叉分别建模

方法	主要机制	输出含义
BDL	贝叶斯深度学习，建模参数分布	模型不确定性与预测分布
GAN	生成对抗网络，学习多样样本	多样轨迹；概率需另行校准
CVAE	条件变分自编码器，条件潜变量	条件下的多个未来模式

多模态＝多种未来模式；多样样本不自动给出可用的风险概率。

8.3.2 行为预测与交互建模：风险算例

低概率分支也可能决定避障动作

行动代价	直行：0.7	转弯：0.3
立即穿行	0	100
等待	2	2

$$\mathbb{E}[C_{\text{cross}}] = 0.7 \times 0 + 0.3 \times 100 = 30$$

$$\mathbb{E}[C_{\text{wait}}] = 0.7 \times 2 + 0.3 \times 2 = 2$$

本例给定分支概率；各项为同尺度无量纲代价，选择等待。

8.3.2 行为预测与交互建模：多时刻结果

重观测更新未来，误差按时域评价

未来模式	3s位置／m	4s位置／m	权重变化
直行	(3,0)	(4,0)	0.7→0.2
左转	(2,1)	(2,2)	0.3→0.8

以2s为预测起点；3s观测转向后更新权重（教学假设）。
平均位移误差（ADE）：逐点误差的均值；最终位移误差（FDE）：末点误差。

给定同一候选在3s、4s的误差0.2m、0.4m。

$$\mathrm{ADE} = \frac{0.2 + 0.4}{2} = 0.3\ \mathrm{m}, \quad \mathrm{FDE} = 0.4\ \mathrm{m}$$

本章目录

定位建图与预测规划

第八章 多模态感知与智能决策

8.1 背景意义与图像处理基础

8.2 目标识别、跟踪与光流估计

8.3 定位建图与预测规划

8.4 常用感知传感器

8.5 多模态感知坐标转换

8.6 多模融合与控制决策

8.7 关键工具与框架接口

8.8 实验案例与真机迁移

8.9 习题与课堂讨论

本节 4 个教学单元

8.3.1 联合定位与建图
教材对应：第8.1.2.4节

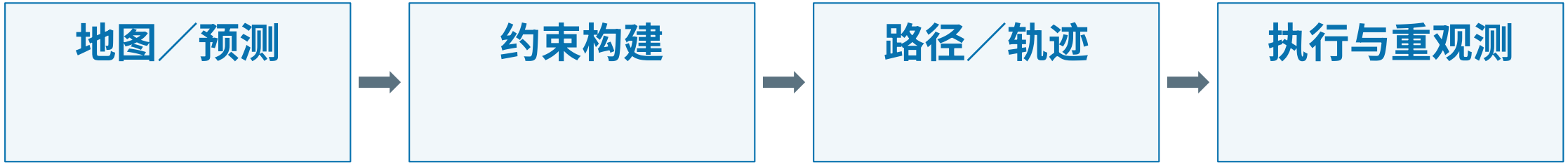
8.3.2 行为预测与交互建模
教材对应：第8.1.2.5节

8.3.3 路径规划与动态避障
教材对应：第8.1.2.6节

8.3.4 多机融合感知与决策
教材对应：第8.1.2.7节

8.3.3 路径规划与动态避障：路径和轨迹

感知输出要变成几何与时间约束



- **路径：**几何位置序列：是否穿越障碍，是否留出足够净空。
- **轨迹：**时间参数化：位置、速度与加速度是否满足执行限值。
- **约束来源：**占据／距离场、目标未来、载具半径、控制限值和更新时延。

8.3.3 路径规划与动态避障：规则与反应式方法

势场、速度禁区和路线图各有机制

方法	作用对象	主要局限
人工势场（APF）	吸引／排斥得到局部运动方向	局部极小、狭窄通道与振荡
速度障碍（VO）	在时域内会碰撞的相对速度	预测、时延和动态可行性
动态运动路线图	碰撞变化后的路线边／节点	需要碰撞几何及更新机制

速度障碍VO≠视觉里程计VO；局部感知不建全局图，也仍要知道障碍几何。

8.3.3 路径规划与动态避障：栅格图搜索

搜索累计代价，启发或修复改变效率

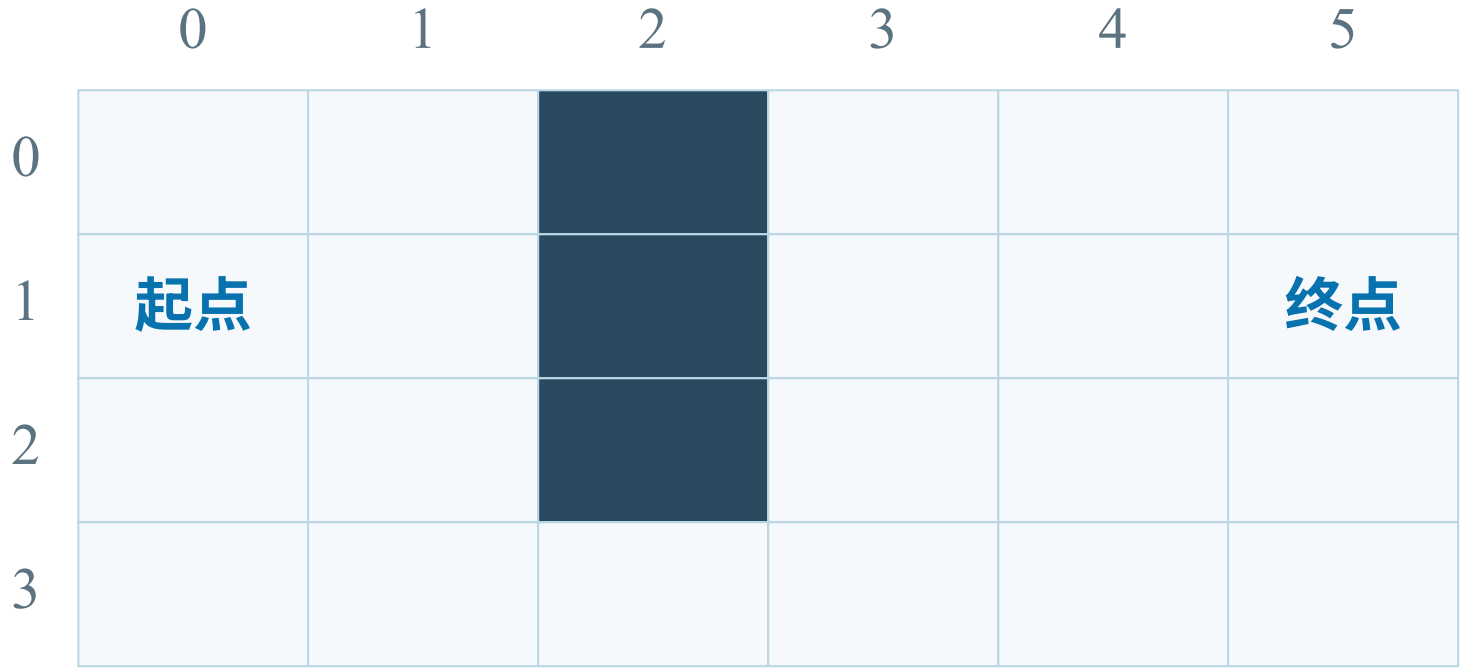
方法	节点优先级／机制	条件或适用
Dijkstra	优先展开最小累计代价g	非负边代价
A*	累计代价＋剩余启发	启发不高估； 不一致时重开节点。
D*类	环境代价变化后增量修复	复用搜索，不是动态安全保证

$$f(n) = g(n) + h(n)$$

n为节点，g为起点至该节点的累计代价，h为剩余估计，f为两者之和。

8.3.3 路径规划与动态避障：A*一步算例

比较已走代价和剩余启发

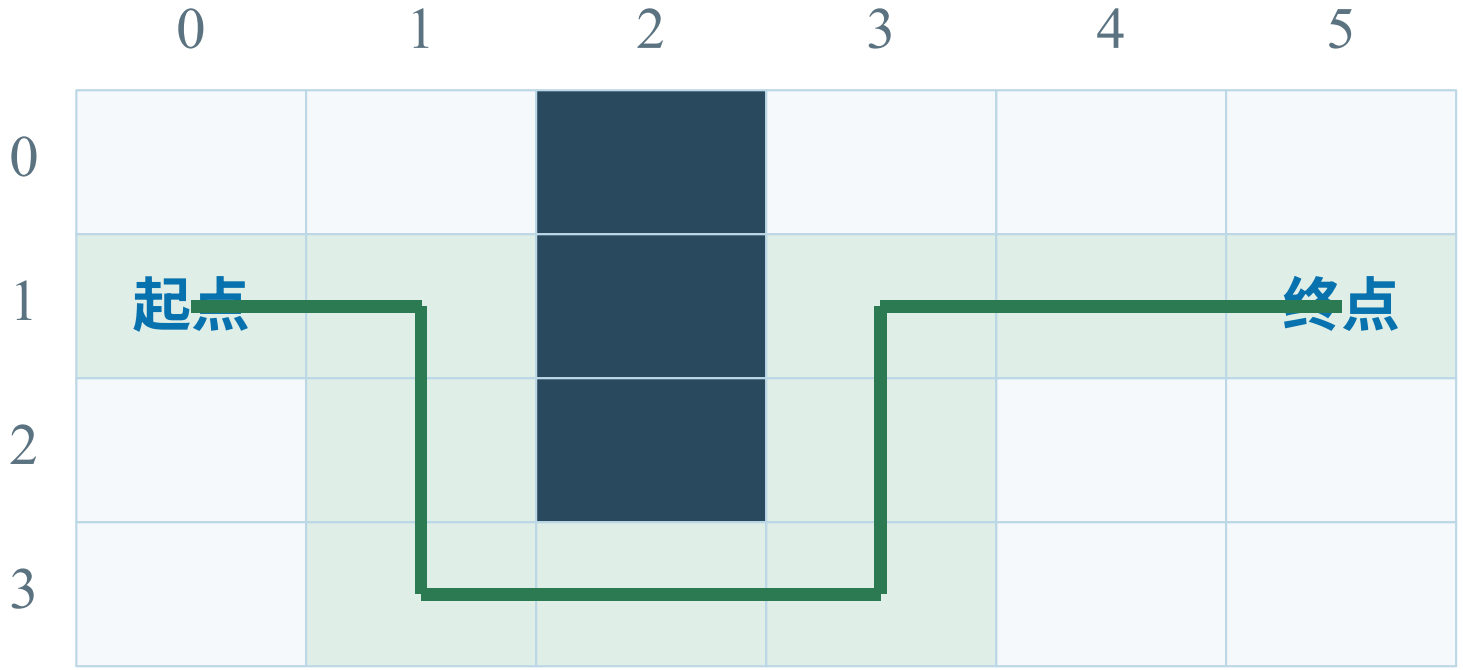


候选	g	h	f
(1,1)	1	4	5
(0,0)	1	6	7
(0,2)	1	6	7

四邻接、边长1m，坐标按(列,行)，深色为障碍。
曼哈顿启发＝横、纵剩余步数之和；先展开总计5的(1,1)。

8.3.3 路径规划与动态避障：完整搜索结果

九步路径绕障，代价模型给出最短性



完整结果：9条边，长度9m。
最少向下再向上4步，横向5步。

$$L_{\min} = 4 + 5 = 9\ \mathrm{m}$$

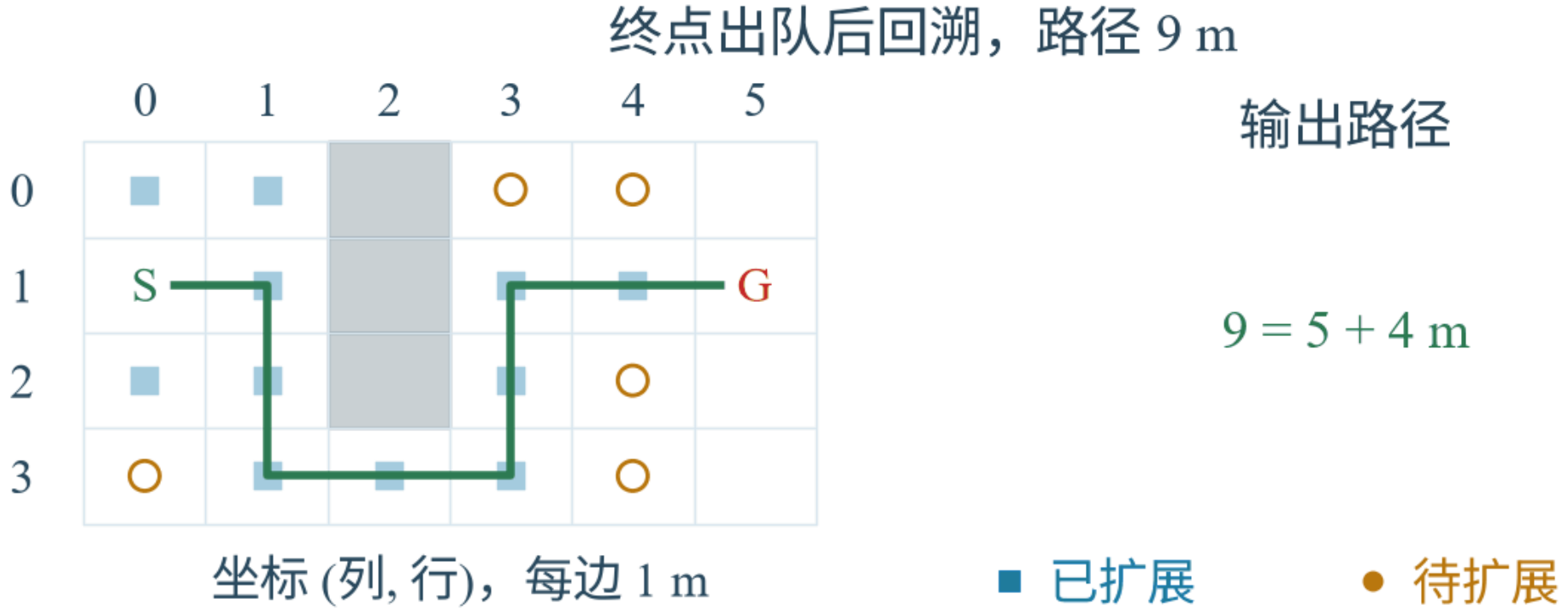
绿线：输出路径；深色：障碍。
最短性只针对该栅格与单位代价，不代表动态可执行。

从下端绕行；载具半径膨胀障碍后需要重新搜索。

8.3.3 路径规划与动态避障：搜索过程

A*从开放集扩展到九步绕障路径

先按累计代价与曼哈顿启发选候选，终点出队后沿父节点回溯。



图：本章6×4栅格的逐步搜索，单位边代价为1 m。

8.3.3 路径规划与动态避障：RRT与RRT*

随机树扩展连接，重连继续改进代价

随机采样 → 最近节点 → 可行扩展 → 加入树

RRT：碰撞检查后加入连接

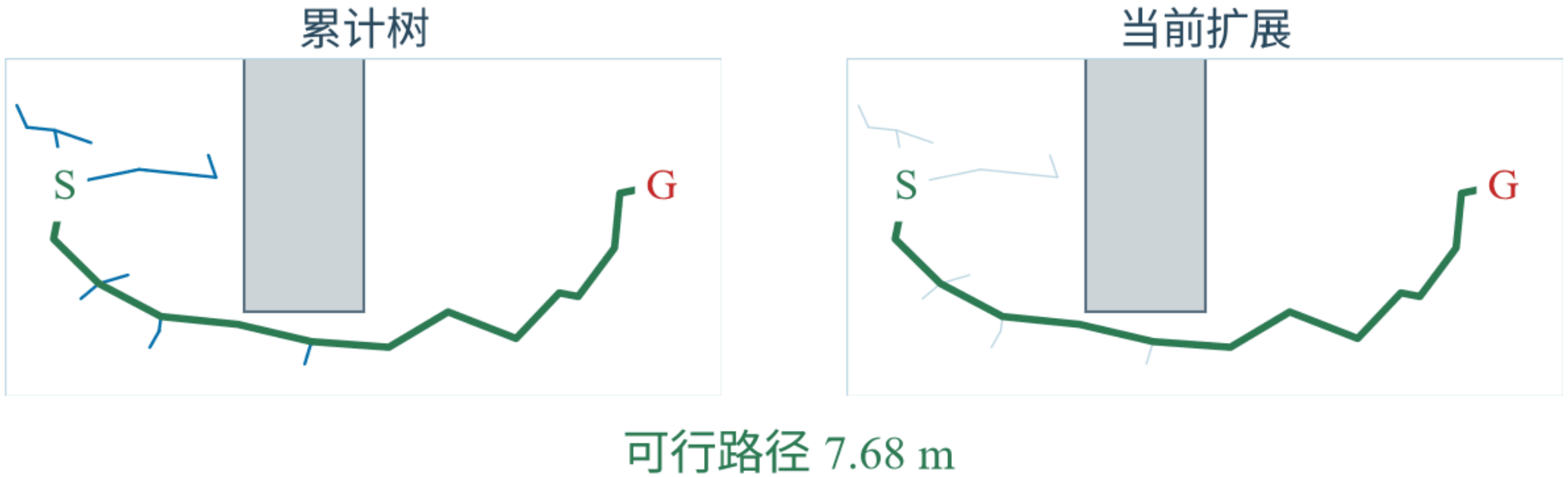
RRT*：选父节点＋重连降代价

快速探索随机树（RRT）反复采样与扩展。
RRT*的“渐近最优”有采样与连接条件，不是有限次数即最优。
树边需要对应状态空间的碰撞／动力学检查，不能只看直线是否穿墙。

8.3.3 路径规划与动态避障：随机树几何

RRT扩展必须通过整段碰撞检查

采样确定方向，最近节点按步长扩展；碰到障碍的整段连接被拒绝。



图：同一障碍地图上的点机器人几何RRT（离线计算）。

8.3.3 路径规划与动态避障：轨迹优化

把平滑、时间和净空放入约束优化

$$\min_{\mathbf{p}(t),T} \int_0^T \|\ddot{\mathbf{p}}(t)\|^2 \, dt + \lambda T$$

$$d_{\mathrm{ESDF}}(\mathbf{p}(t)) \geq r + m, \quad \|\dot{\mathbf{p}}(t)\| \leq v_{\max}, \quad \|\ddot{\mathbf{p}}(t)\| \leq a_{\max}$$

位置单位m、时间单位s；$\dot{\mathbf{p}}$是速度，$\ddot{\mathbf{p}}$是加速度。

$\mathbf{p}(t)$：位置轨迹；T：总时长；r：载具半径；m：安全裕量。

$v_{\max}$、$a_{\max}$：速度、加速度上限；d_{ESDF}：最近障碍距离。

$\lambda \geq 0$：时间代价权重，越大越重视缩短用时；与平滑代价权衡。

给定起终点；局部优化依赖初值，需核验全轨迹净空。

8.3.3 路径规划与动态避障：蚁群与粒子群

群体迭代搜索仍须通过约束检查

蚁群优化（ACO）

构造候选路线
→ 按代价更新信息素
→ 蒸发与再次选择
→ 检验碰撞／限值

粒子群优化（PSO）

候选参数＝粒子位置
→ 个体最好＋群体最好
→ 更新速度／位置
→ 检验碰撞／限值

混合搜索可以改变候选生成；安全条件仍来自几何、时序与动力学模型。

8.3.3 路径规划与动态避障：净空和动态占据

位置安全还要与障碍在时间上错开

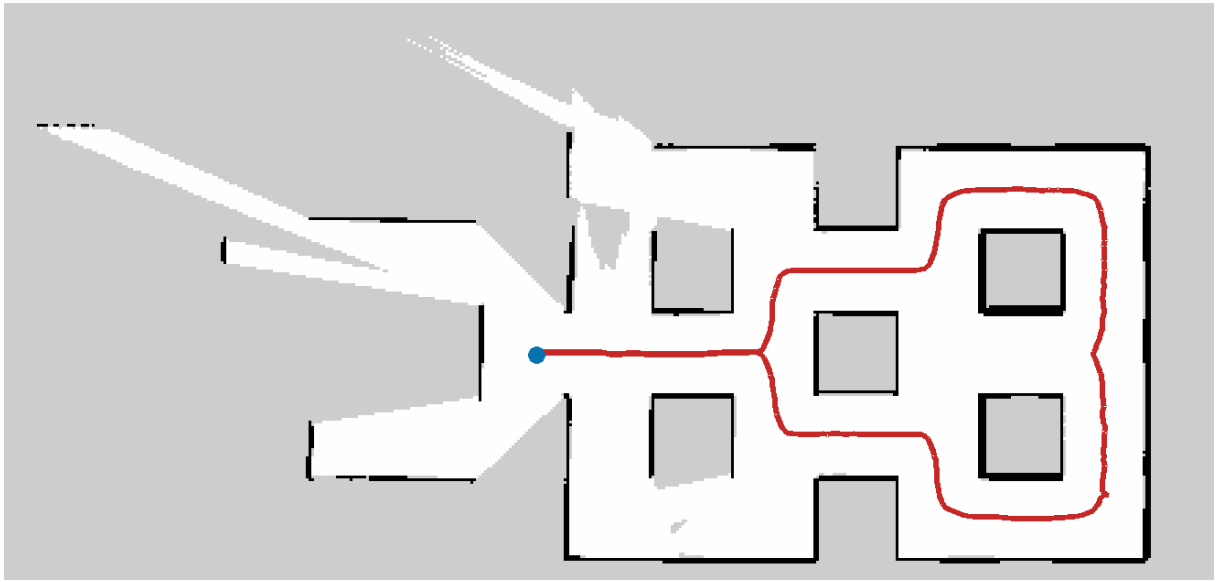
$$d_{\text{safe}} = r + m = 0.25 + 0.15 = 0.40\ \text{m}$$

现在的地图 → 预测到达时刻 → 未来占据检查

- **静态通道：**圆形投影、居中通过时，双侧净宽至少0.80m；空间和误差还需核实。
- **动态障碍：**己方轨迹与他方未来在同一时刻比较；更新时延与制动条件进入检查。

8.3.3 路径规划与动态避障：ESDF例程素材

从地图坐标到可执行路径



已有地图与已存的448点路径

数据：map_cropped.pgm
路径：planned_path.npy
数组顺序：(行,列)
分辨率：0.05m／像素
尚未转为控制参考系。

学生入口：[安装目录]/RflySimAPIs/8.RflySimVision/2.AdvExps/e16_ESDFPathPlan/maps/OGM2ESDF.py

本章目录

定位建图与预测规划

第八章 多模态感知与智能决策

8.1 背景意义与图像处理基础

8.2 目标识别、跟踪与光流估计

8.3 定位建图与预测规划

8.4 常用感知传感器

8.5 多模态感知坐标转换

8.6 多模融合与控制决策

8.7 关键工具与框架接口

8.8 实验案例与真机迁移

8.9 习题与课堂讨论

本节 4 个教学单元

8.3.1 联合定位与建图
教材对应：第8.1.2.4节

8.3.2 行为预测与交互建模
教材对应：第8.1.2.5节

8.3.3 路径规划与动态避障
教材对应：第8.1.2.6节

8.3.4 多机融合感知与决策
教材对应：第8.1.2.7节

8.3.4 多机融合感知与决策：四类协同信息

定位、互感、地图和决策相互支撑

角色	共享信息	解决的问题
共同GNSS参考	全球位置与坐标基准	大范围位置初始共识
视觉互感	邻机相对位姿与观测	短距离关系约束
跨机图优化	关键帧、跨机边、局部地图	地图对齐与一致性
协同决策	障碍、目标与任务成本	分配／时空冲突／控制

8.3.4 多机融合感知与决策：坐标对齐算例

同一个目标先变换到共同坐标

二维约定：x向右、y向上，正旋转逆时针；这里不是NED航向。
B局部原点(10,2)m，局部x轴相对共同x轴逆时针旋转90°。
B局部目标(2,1)m；先旋转，再平移。

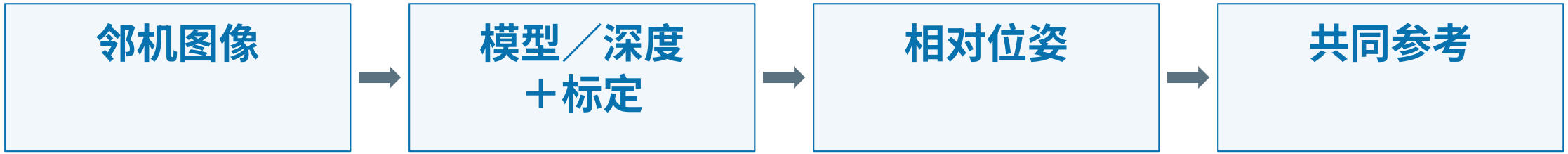
$$\mathbf{p}^{W} = R(90^\circ)\mathbf{p}^{B} + \mathbf{t}_{B}^{W}$$

$$(2,1) \overset{R(90^\circ)}{\rightarrow} (-1,2) \overset{+(10,2)}{\rightarrow} (9,4)\ \mathrm{m}$$

忽略旋转会错为(12,3)m；共同坐标还要配合共同的时间和尺度。

8.3.4 多机融合感知与决策：视觉互感

相对位姿约束连接两架机器



$$T_B^W = T_A^W T_c^A T_B^c$$

T_b^a把b系量变到a系；w为共同参考，A/B为机体系，c为A机相机。

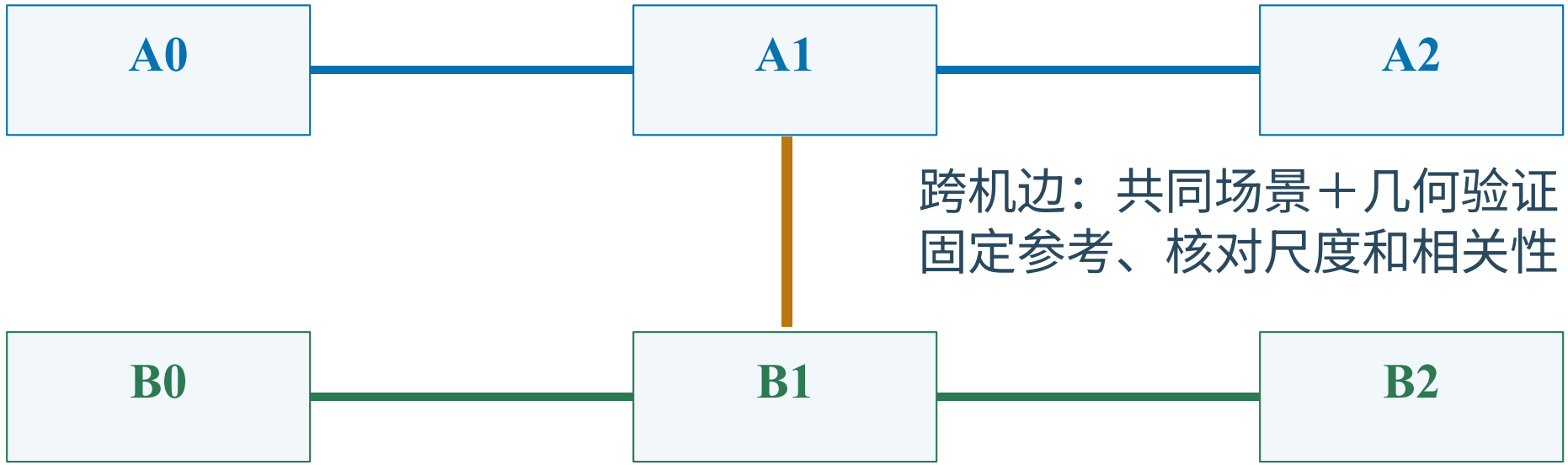
- **条件：**明确模型或尺寸、尺度、外参和时间；视线被遮挡时观测可能中断。

- **输出边界：**二维框中心不是六自由度（3平移＋3旋转）；相对约束需要误差和质量描述。

8.3.4 多机融合感知与决策：多机图优化

跨机边使局部地图获得共同参考

节点＝关键帧位姿；横边＝各自里程计；竖边＝核验后的跨机约束。



共享关键帧不自动保证地图重叠；错误跨机边会影响两机的共同估计。

8.3.4 多机融合感知与决策：融合一步算例

独立估计按误差尺度加权

给定：9.0m，标准差0.2m；9.4m，标准差0.4m。
同一量、同一时刻、无偏且互相独立。

$$\hat{x} = \frac{9.0/0.2^2 + 9.4/0.4^2}{1/0.2^2 + 1/0.4^2} = 9.08\ \mathrm{m}$$

$$\sigma = \sqrt{\frac{1}{1/0.2^2 + 1/0.4^2}} \approx 0.179\ \mathrm{m}$$

同源或已融合的数据存在相关性，不能当作新独立观测重复计数。

8.3.4 多机融合感知与决策：时间与陈旧信息

同一坐标仍需要同一时刻

$$\Delta p = v\Delta t = 2 \times 0.100 = 0.20\ \mathrm{m}$$

例：目标速度2m/s，状态滞后100ms；未传播时有0.20m位置差。



- **消息字段：**坐标／尺度、观测时标、来源／协方差或质量与有效期。
- **检查：**传输延迟、时钟偏差和共享相关性；失联后按有效期处理。

8.3.4 多机融合感知与决策：任务分配算例

整组分配可以优于逐机贪心

任务代价	任务T1	任务T2
A机	2	3
B机	2.5	100

$$C_{\text{greedy}} = 2 + 100 = 102$$

$$C_{\text{joint}} = 3 + 2.5 = 5.5$$

每机各接一项任务，代价无量纲；贪心按A先选，整组比较两种分配。

8.3.4 多机融合感知与决策：连续决策结果

分配目标后仍要处理路径相交

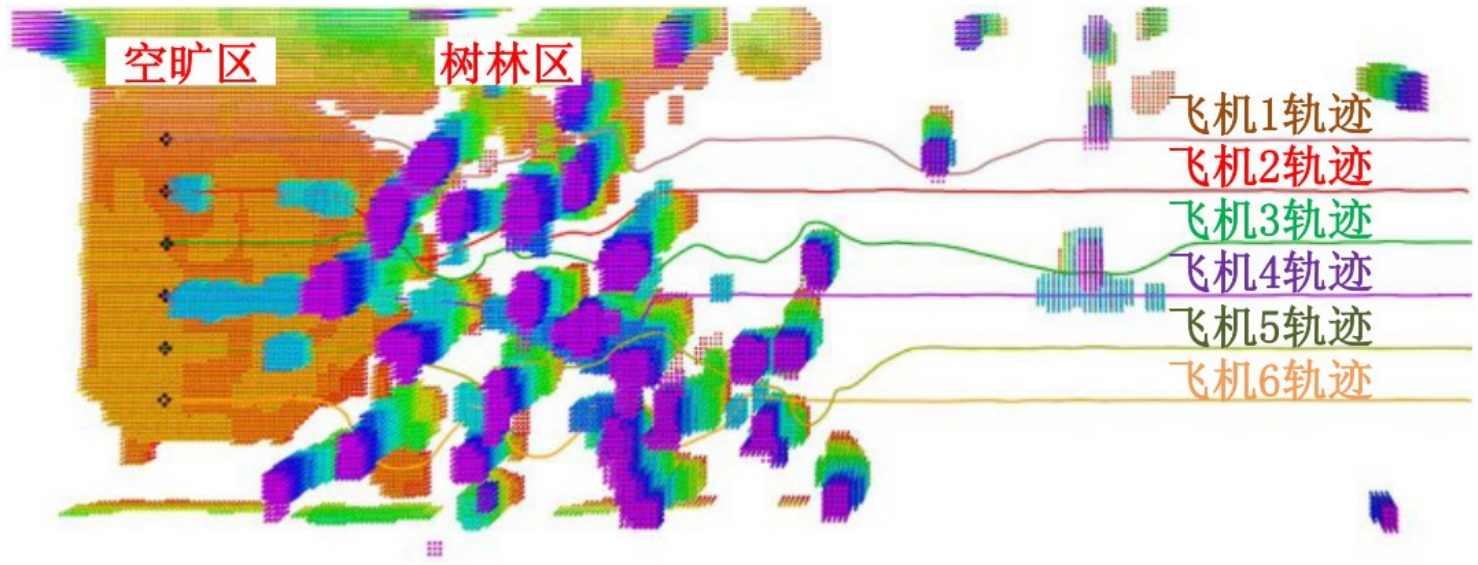
共享状态 → 分配任务 → 规划路径 → 协调时窗 → 反馈更新

时刻／s	A机状态	B机状态
0	接近交点	接近停留点
1	占用交点	等待，不进交点
2	已退出	进入交点
3	继续去目标	已退出

示意顺序：先退出再进入；安全间隔需结合尺寸、制动与信息有效期设置。

8.3.4 多机融合感知与决策：教材多机图

多条轨迹对应多机的避障规划



共享地图：环境几何。
相对状态：机间关系。
共同时间：占用冲突。
限值：各自可执行。

图8.3(e) 多机轨迹，完整关系见第14页

单幅规划图不能证明实测无碰撞；评估还需要执行状态、时间与安全间距。

8.3.4 多机融合感知与决策：集中式与分布式

通信和求解位置改变协同条件

维度	集中式	分布式
求解位置	中心进行全局分配／调度	节点局部求解并交换约束
信息条件	上传状态；中心整合全组信息	邻居信息；异步与不完整观测
主要风险	通信／算力瓶颈、中心故障	一致性、局部冲突与信息陈旧

共同要求：时标、有效期、冲突规则和失联行为；通信条件进入决策模型。

练习与讨论

地图、未来和协同各需要哪些条件

- **问题1：**相似走廊触发回环候选，能否立即修正轨迹？还需什么证据？
- **问题2：**避障需要最近障碍净空，应该查TSDF还是ESDF？
- **问题3：**直行0.7／转弯0.3，穿行成本0／100，等待均为2，应如何选？
- **问题4：**x向右、y向上、正转逆时针（第39页）。
B原点(10,2)m，局部x轴旋转90°，目标(2,1)m。求共同坐标；
按第44页代价表比较任务分配。

参考要点

用观测、坐标和约束解释结果

- **答案1：**先做几何验证再添加回环约束；重复纹理会产生错误候选。
- **答案2：**查ESDF最近障碍距离；TSDF重建表面，不能直接代替全空间净空。
- **答案3：**穿行期望30，等待2；最大概率未来会漏掉低概率高损失分支。
- **答案4：**共同目标(9,4)m；A→T2、B→T1总代价5.5，贪心102；之后仍需协调时序。

本节小结

从局部观测走向共同的可行行动

- **定位与地图：**联合估计、地图表示与回环约束；局部一致性不等于全球坐标。
- **行为预测：**历史、交互、场景、多种未来；样本不自动等于校准概率。
- **规划避障：**几何路径加时间与限值；图搜索、采样、优化的条件各不同。
- **多机协同：**共同坐标、尺度、时基和来源；整组分配后仍要协调路径与占用。