本章目录

多模态感知坐标转换

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

8.5.1 像素坐标与图像坐标
教材对应：第8.1.5节

8.5.2 图像到相机坐标
教材对应：第8.1.5节

8.5.3 相机到机体及地面坐标
教材对应：第8.1.5节

8.5.4 地理坐标与完整转换算例
教材对应：第8.1.5节

8.5 多模态感知坐标转换

8.5.1 像素坐标与图像坐标：坐标转换链

同一目标的五种坐标表示

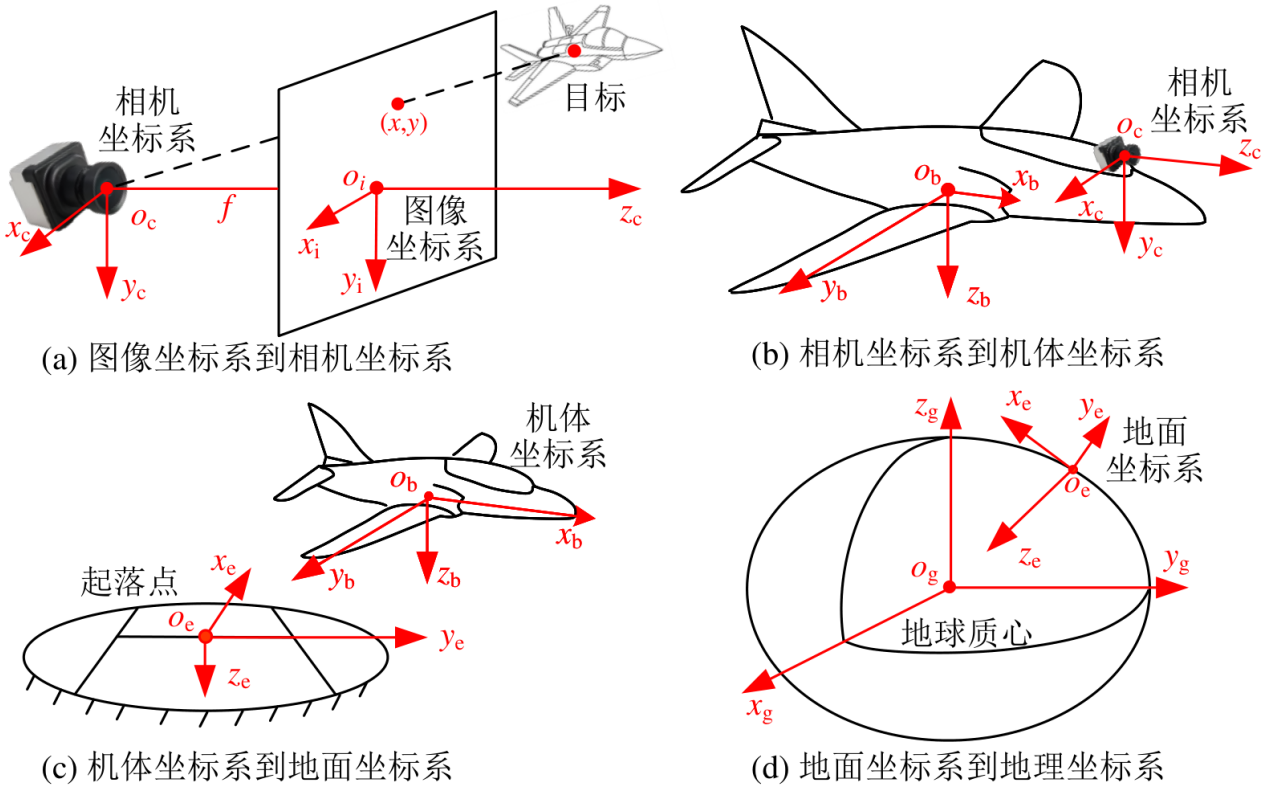


图8.5 坐标转换示意图

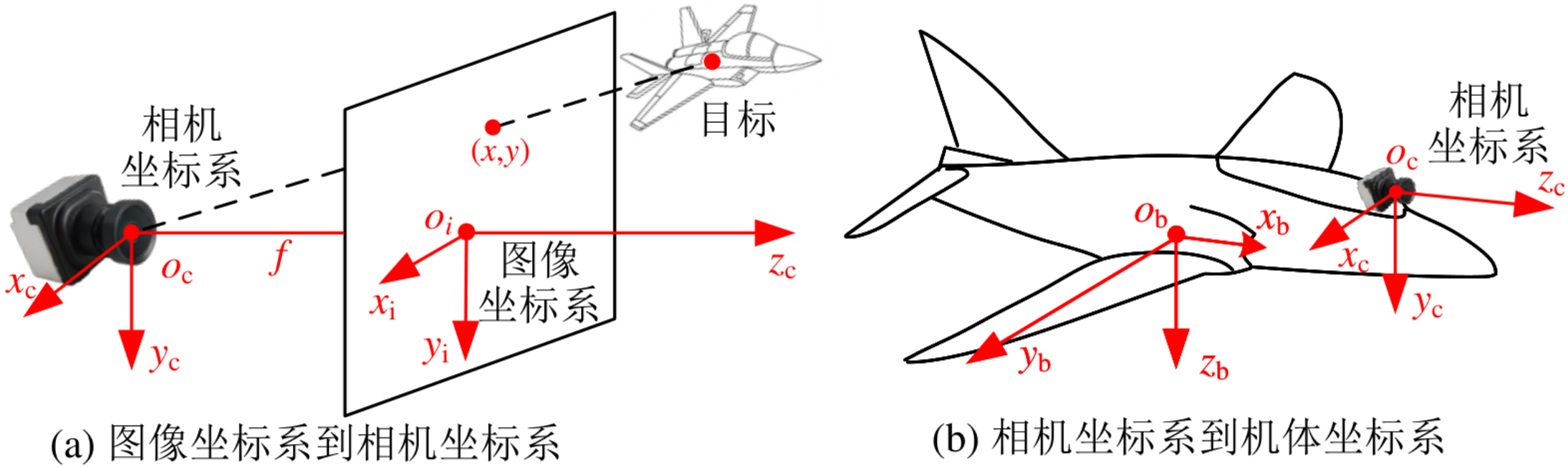
像素与相机：像素给出视线，深度给出沿视线的尺度。

机体与地面：安装外参和载具位姿把点放入任务场景。

共同地理基准：载具和目标使用同一原点、坐标约定与时间基准进行协同。

8.5.1 像素坐标与图像坐标：图像与相机

像素定位与安装位置的几何关系



(a) 图像坐标系到相机坐标系　　(b) 相机坐标系到机体坐标系

图8.5（上半部）成像几何与传感器安装

左图需要像素、内参和深度。右图需要相机安装位置及朝向。

8.5.1 像素坐标与图像坐标：像素坐标系

分辨率规定数据规模与坐标范围

原点与方向：左上角为像素原点，u向右，v向下，单位px。

图像尺寸	零基整数索引范围	像素数量
640×480	$0 \le u \le 639,\ 0 \le v \le 479$	307 200
1280×720	$0 \le u \le 1279,\ 0 \le v \le 719$	921 600（3倍）

参考系记号： ${}^{\mathrm{i}}\mathbf{p}$图像， ${}^{\mathrm{c}}\mathbf{p}$相机， ${}^{\mathrm{b}}\mathbf{p}$机体， ${}^{\mathrm{e}}\mathbf{p}$地面， ${}^{\mathrm{g}}\mathbf{p}$地理。

分辨率是尺寸；主点由标定获得，目标中心可采用亚像素坐标。

8.5.1 像素坐标与图像坐标：图像坐标系

主点平移把像素差换成物理长度

$[u_0, v_0]^{\mathrm{T}}$为主点；d_x, d_y为物理像素尺寸。本例零倾斜，图像轴右、下为正。

$$x_{\mathrm{i}} = (u - u_0)d_x \tag{8.1}$$

$$y_{\mathrm{i}} = (v - v_0)d_y$$

给定像素(480,320)、主点(320,240)，$d_x = d_y = 0.01\ \mathrm{mm/px}$。

$$x_{\mathrm{i}} = 160(0.01) = 1.6\ \mathrm{mm}, \quad y_{\mathrm{i}} = 80(0.01) = 0.8\ \mathrm{mm}$$

两个分量为正，表示目标投影位于主点的右下方。

8.5.1 像素坐标与图像坐标：相机内参

内参连接物理焦距与像素焦距

$$[x_{\mathrm{i}}, y_{\mathrm{i}}, f]^{\mathrm{T}} = f\mathbf{K}^{-1}[u, v, 1]^{\mathrm{T}} \tag{8.2}$$

$$\mathbf{K} = \begin{bmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8.3}$$

$$f_x = f/d_x, \quad f_y = f/d_y$$

物理焦距f与d_x, d_y用相同长度单位；f_x, f_y单位为px，s是像素轴倾斜参数，本例取0。

8.5.1 像素坐标与图像坐标：分辨率变化

缩放图像后同步缩放内参

项目	640×480原图	320×240缩放图
目标像素	(480,320)	(240,160)
主点	(320,240)	(160,120)
像素焦距	320px	160px
归一化射线	$[0.5,0.25,1]^{\mathrm{T}}$	$[0.5,0.25,1]^{\mathrm{T}}$

$$\mathbf{q} = \mathbf{K}^{-1}[u, v, 1]^{\mathrm{T}}$$

错误保留旧内参得到$[-0.25, -0.25,1]^{\mathrm{T}}$，方向已改变。

8.5.1 像素坐标与图像坐标：标定与畸变

真实相机由多视角标定估计参数

参数/工具	输入与产物
标定板与多视角	已知几何尺寸和图像角点，估计内参与畸变。
径向/切向畸变	桶形或枕形径向偏移；装配偏差导致切向偏移。
Camera Calibrator **OpenCV / Kalibr**	可视化标定、函数标定、相机/惯性联合标定。

配套阅读：2.CameraCalcDemo / Readme.pdf
采集OneCameraCal.py；离线bord-Calibration.py输出标定结果。

标定后的射线还需要深度约束。

本章目录

多模态感知坐标转换

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

8.5.1 像素坐标与图像坐标
教材对应：第8.1.5节

8.5.2 图像到相机坐标
教材对应：第8.1.5节

8.5.3 相机到机体及地面坐标
教材对应：第8.1.5节

8.5.4 地理坐标与完整转换算例
教材对应：第8.1.5节

8.5.2 图像到相机坐标：相机坐标系

针孔模型把三维点投影到成像平面

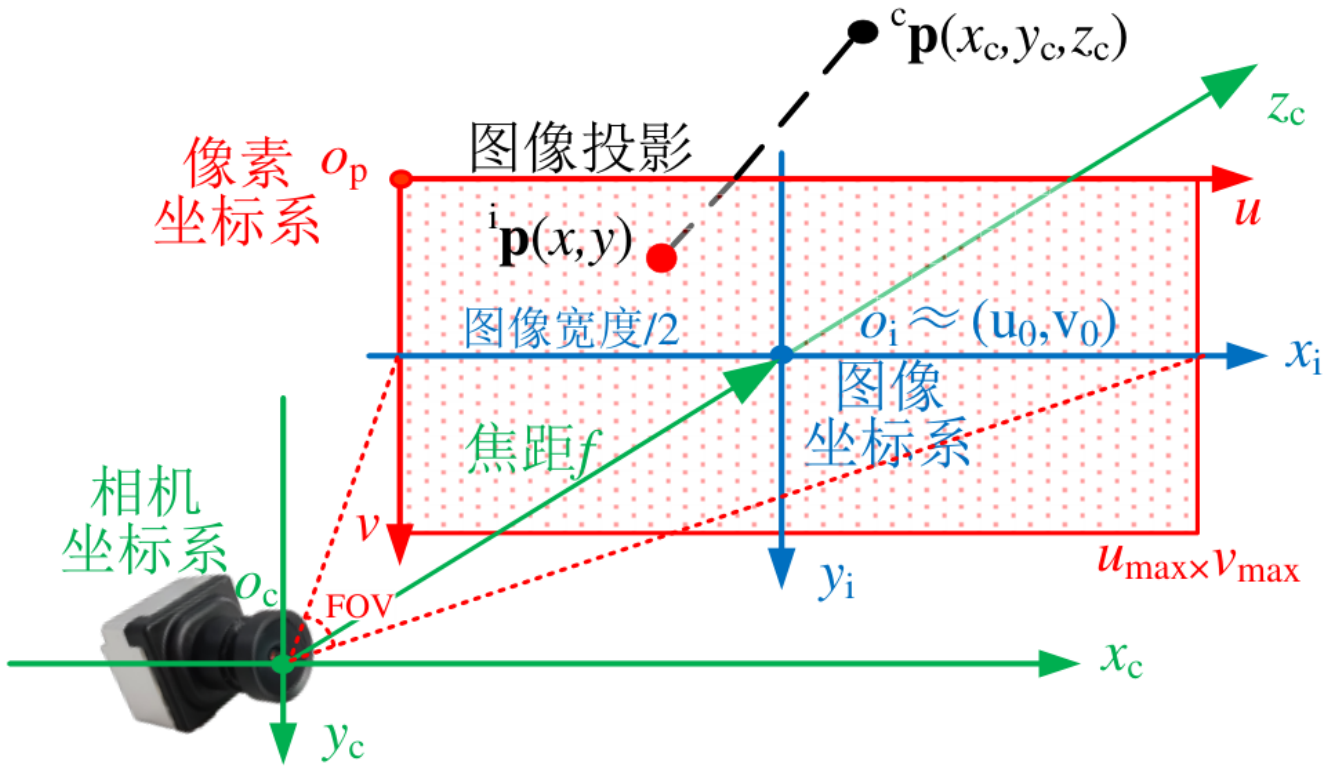


图8.6 像素坐标系到相机坐标系的转换

光心原点：x_c向右，y_c向下，z_c沿光轴向前。

焦距与深度：焦距是光心到像面的距离；深度是目标沿光轴的分量。

针孔条件：本例使用已校正像素、正交方形像素和零倾斜模型。

8.5.2 图像到相机坐标：视场角

横向视场角确定像素焦距

理想投影、无畸变、方形像素；FOV默认指水平方向。

$$\frac{\mathrm{FOV}}{2} = \arctan\left(\frac{d_x u_{max}/2}{f}\right) \tag{8.4}$$

$$f_x = f_y = \frac{f}{d_x} = \frac{u_{max}}{2\tan(\mathrm{FOV}/2)} \tag{8.5}$$

$$u_{max} = 640\ \mathrm{px}, \quad \mathrm{FOV} = 90^\circ \implies f_x = f_y = 320\ \mathrm{px}$$

90°/2=45°，tan45°=1。宽度640px得到焦距320px。

8.5.2 图像到相机坐标：理想相机内参

640×480、90°相机的完整内参

$$\mathbf{K}=\begin{bmatrix}\dfrac{u_{max}}{2\tan(\mathrm{FOV}/2)} & 0 & u_{max}/2\\ 0 & \dfrac{u_{max}}{2\tan(\mathrm{FOV}/2)} & v_{max}/2\\ 0 & 0 & 1\end{bmatrix} \tag{8.6}$$

$$\mathbf{K}=\begin{bmatrix}320 & 0 & 320\\ 0 & 320 & 240\\ 0 & 0 & 1\end{bmatrix}$$

像素焦距320px，主点(320,240)；真实主点需要标定。

$$\mathrm{FOV}_y = 2\arctan(240/320) \approx 73.74^\circ$$

8.5.2 图像到相机坐标：反投影

深度为相机射线确定尺度

$$[x_c, y_c, z_c]^{\mathsf{T}} = z_c \mathbf{K}^{-1} [u, v, 1]^{\mathsf{T}} \tag{8.7}$$

输入为已校正像素(480,320)、内参$\mathbf{K} \in \mathbb{R}^{3\times3}$和轴向深度$z_c = 4$ m。

$$\mathbf{q} = [0.5, 0.25, 1]^{\mathsf{T}}, \qquad {}^{c}\mathbf{p} = 4\mathbf{q} = [2, 1, 4]^{\mathsf{T}} \text{ m}$$

结果表示相机右方2m、下方1m、前方4m。相机点是3×1向量。

8.5.2 图像到相机坐标：斜距与轴向深度

斜距沿视线量测，深度沿光轴量测

$$\mathbf{q} = \mathbf{K}^{-1}[u, v, 1]^{\mathsf{T}}, \qquad {}^{c}\mathbf{p} = r\frac{\mathbf{q}}{\|\mathbf{q}\|}$$

$$z_c = \frac{r}{\|\mathbf{q}\|}, \quad \|\mathbf{q}\| = \sqrt{1.3125} \approx 1.14564$$

量测处理	轴向深度/m	恢复相机点/m
正确：$r = \sqrt{21}$ m换算	4.0000	(2.0000,1.0000,4.0000)
错误：斜距直接作深度	4.5826	(2.2913,1.1456,4.5826)

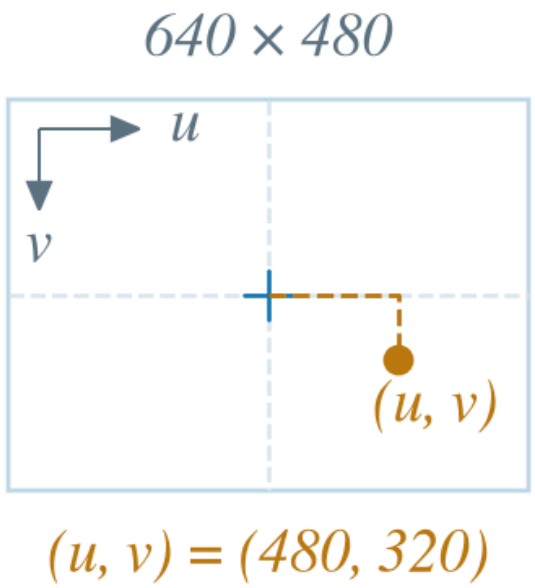
只有目标恰在光轴上时，斜距才等于轴向深度。

8.5.2 图像到相机坐标：针孔成像与反投影

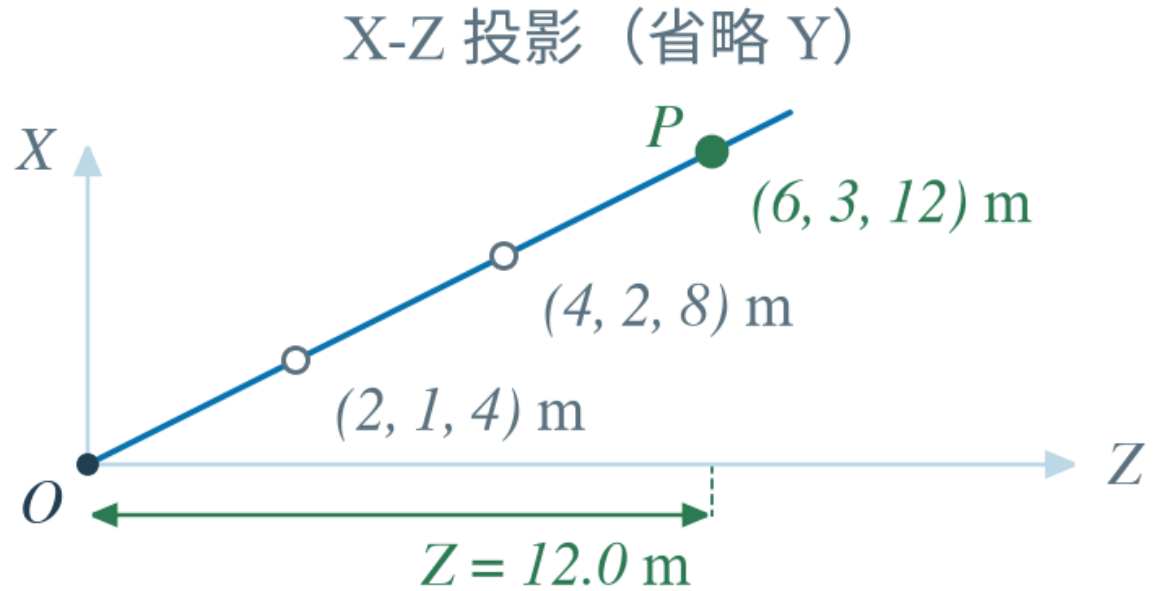
像素确定一条射线，深度确定空间位置

图中记号：$c_x = u_0$，　$X = x_c$，　$Z = z_c$，　$\mathbf{P} = {}^{c}\mathbf{p}$

同一像素，多个点



给定深度选点



$f_x = f_y = 320$ px　　$c_x = 320$，　$c_y = 240$

图：同一像素对应一条空间射线，给定轴向深度才能选定一点。

8.5.2 图像到相机坐标：反投影核验

正反投影一致用于检查单位与尺度

$$u = f_x x_c / z_c + u_0, \quad v = f_y y_c / z_c + v_0$$

轴向深度/m	反投影三维点/m	再次投影/px
4	(2,1,4)	(480,320)
8	(4,2,8)	(480,320)
12	(6,3,12)	(480,320)

回算相同像素只说明几何自洽。深度的正确性依赖测距、立体或已知尺寸约束。

几何回算能发现方向和尺度错误。

本章目录

多模态感知坐标转换

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

8.5.1 像素坐标与图像坐标
教材对应：第8.1.5节

8.5.2 图像到相机坐标
教材对应：第8.1.5节

8.5.3 相机到机体及地面坐标
教材对应：第8.1.5节

8.5.4 地理坐标与完整转换算例
教材对应：第8.1.5节

8.5.3 相机到机体及地面坐标：机体坐标系

右下前相机与前右下机体的轴置换

约定	正轴方向	本例选择
相机右下前	x右，y下，z前	相机光心为原点
机体FRD	x前，y右，z下	机体中心为原点
机体FLU	x前，y左，z上	与FRD的y、z同时变号

$$\mathbf{R}_c^b = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{R}_c^b[2,1,4]^\mathsf{T} = [4,2,1]^\mathsf{T}$$

本例光轴对准机头且无相对滚转；安装旋转变化时须重算矩阵。

8.5.3 相机到机体及地面坐标：相机外参

旋转与安装偏移共同决定机体点

$$ {}^{b}\mathbf{p} = \mathbf{R}_{c}^{b} \quad {}^{c}\mathbf{p} + \quad {}^{b}\mathbf{t}_{c} \tag{8.8} $$

$$ {}^{b}\tilde{\mathbf{p}} = \mathbf{T}_{c}^{b} \quad {}^{c}\tilde{\mathbf{p}} $$

${}^{b}\mathbf{t}_{c} = [0.2,0.1,0.05]^{\mathrm{T}}$ m表示相机光心相对机体原点的安装偏移。

$$ {}^{b}\mathbf{p} = [4,2,1]^{\mathrm{T}} + [0.2,0.1,0.05]^{\mathrm{T}} = [4.2,2.1,1.05]^{\mathrm{T}}\ \mathrm{m} $$

旋转无量纲，平移与点坐标均使用米。变换方向是相机到机体。

8.5.3 相机到机体及地面坐标：齐次变换

四维齐次点统一旋转和平移

$$\mathbf{T}_c^b = \begin{bmatrix} \mathbf{R}_c^b & {}^b\mathbf{t}_c \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \tag{8.9}$$

$${}^c\tilde{\mathbf{p}} = [\, z_c(\mathbf{K}^{-1})^{\mathsf{T}} \quad \mathbf{e}_3 \,]^{\mathsf{T}}[u, v, 1]^{\mathsf{T}}, \quad \mathbf{e}_3 = [0,0,1]^{\mathsf{T}}$$

$$\mathbf{T}_c^b = \begin{bmatrix} 0 & 0 & 1 & 0.2 \\ 1 & 0 & 0 & 0.1 \\ 0 & 1 & 0 & 0.05 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

三维点补1成为齐次点；刚体矩阵4×4，像素提升矩阵4×3。

8.5.3 相机到机体及地面坐标：外参方向

已知机体到相机变换时，先取逆再用

$$\mathbf{T}_c^b = (\mathbf{T}_b^c)^{-1} = \begin{bmatrix} (\mathbf{R}_b^c)^\mathsf{T} & -(\mathbf{R}_b^c)^\mathsf{T} & {}^c\mathbf{t}_b \\ \mathbf{0}_{1\times 3} & & 1 \end{bmatrix}$$

求逆获得相机到机体安装外参；下面反向回算原相机点作核验。

$${}^c\mathbf{p} = (\mathbf{R}_c^b)^\mathsf{T}({}^b\mathbf{p} - {}^b\mathbf{t}_c) = [2,1,4]^\mathsf{T}\ \mathrm{m}$$

计算方式	机体点/m
正确相机到机体	(4.2,2.1,1.05)
误用转置旋转、仍加原偏移	(1.2,4.1,2.05)

OpenCV标定的旋转/平移先核对源坐标系和目标坐标系。

8.5.3 相机到机体及地面坐标：地面坐标系

局部NED与ENU保留同一物理位置

局部原点取起飞点、基站或任务中心。近场平面模型适用于有限范围的任务场景。

坐标约定	三轴正方向	垂直位置含义
NED：北东地	北、东、下	负Z表示原点上方
ENU：东北天	东、北、上	正Z表示原点上方

$$\mathbf{p}_{\mathrm{ENU}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \mathbf{p}_{\mathrm{NED}}$$

8.5.3 相机到机体及地面坐标：机体到地面

90°偏航把机头方向旋转到东向

$$ {}^{e}\mathbf{p} = \mathbf{R}_b^e \ {}^{b}\mathbf{p} + {}^{e}\mathbf{t}_b \Leftrightarrow {}^{e}\tilde{\mathbf{p}} = \mathbf{T}_b^e \ {}^{b}\tilde{\mathbf{p}} \tag{8.10} $$

$$ \mathbf{R}_b^e = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad {}^{e}\mathbf{t}_b = [10,20,-30]^{\mathrm{T}}\ \mathrm{m} $$

$$ {}^{e}\mathbf{p} = [-2.1,4.2,1.05]^{\mathrm{T}} + [10,20,-30]^{\mathrm{T}} = [7.9,24.2,-28.95]^{\mathrm{T}}\ \mathrm{m} $$

本例载具水平、偏航90°。结果为北7.9m、东24.2m、原点上方28.95m。

本章目录

多模态感知坐标转换

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

8.5.1 像素坐标与图像坐标
教材对应：第8.1.5节

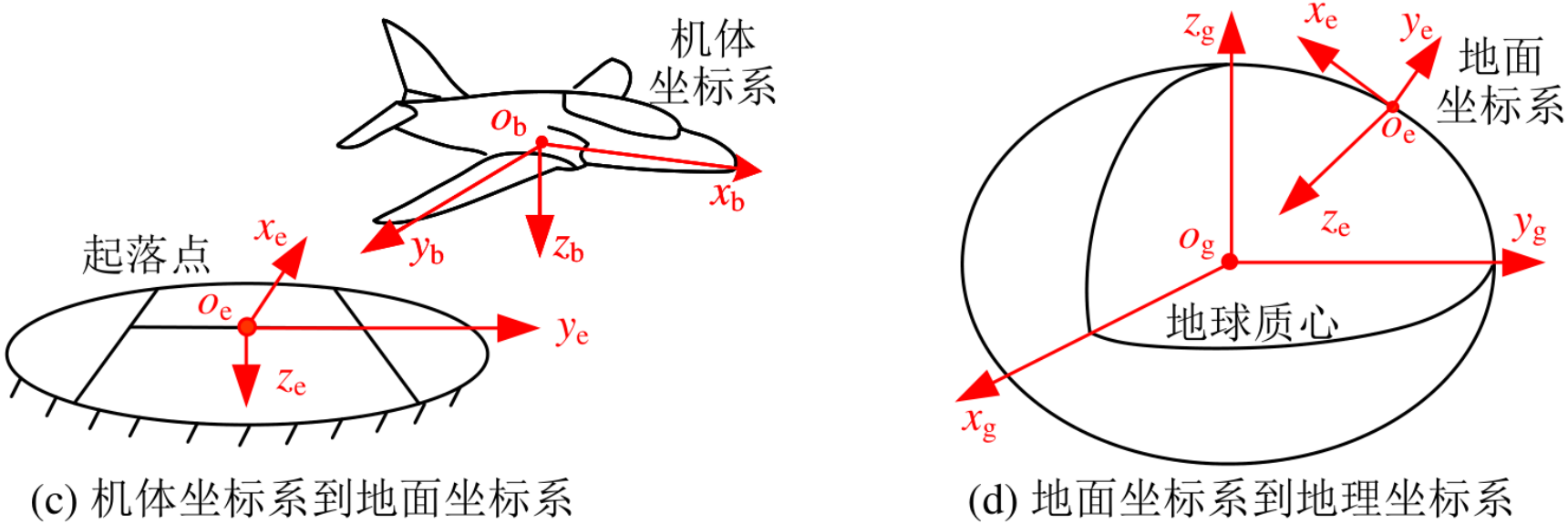
8.5.2 图像到相机坐标
教材对应：第8.1.5节

8.5.3 相机到机体及地面坐标
教材对应：第8.1.5节

8.5.4 地理坐标与完整转换算例
教材对应：第8.1.5节

8.5.4 地理坐标与完整转换算例：地理坐标系

纬经高与地心地固是不同位置表示



(c) 机体坐标系到地面坐标系　　(d) 地面坐标系到地理坐标系

图8.5（下半部）局部地面系与地球基准

LLA为纬度、经度、椭球高：纬度±90°，经度±180°，高度m。
ECEF三轴均为米；椭球高与海拔正高采用不同高度基准。

8.5.4 地理坐标与完整转换算例：局部到地理坐标

NED转换以WGS84参考点为基准

$$ {}^{g}\mathbf{p}_{\mathrm{LLA}} = ned2lla(\ \ {}^{e}\mathbf{p}_{\mathrm{NED}}, \mathrm{origin}_{\mathrm{LLA}}) \tag{8.11}$$

本例地理原点： $\mathrm{origin}_{\mathrm{LLA}} = [30°, 114°, 50\ \mathrm{m}]^{\mathsf{T}}$，高度是WGS84椭球高。

$$\tilde{\mathbf{p}} = [n, e, d, w]^{\mathsf{T}} \implies \mathbf{p} = [n/w, e/w, d/w]^{\mathsf{T}}$$

教材w=1，示例显式取前三项。当前SDK不检查或归一化一般w；MATLAB原生函数还需指定flat/ellipsoid方法。

8.5.4 地理坐标与完整转换算例：近场地理计算

纬度、经度和高度逐项核验

在纬度30°处，WGS84子午圈曲率半径$M_0 \approx 6351377.104$ m，卯酉圈曲率半径$N_0 \approx 6383480.918$ m。

$$\phi \approx \phi_0 + \frac{n}{M_0 + h_0}, \quad \lambda \approx \lambda_0 + \frac{e}{(N_0 + h_0)\cos\phi_0}, \quad h \approx h_0 - d$$

公式中的角增量是弧度，输出纬度和经度转为度。输入NED=(7.9,24.2,-28.95)m。

计算方式	纬度/°	经度/°	椭球高/m
近场一阶	30.000071265	114.000250811	78.950000
经ECEF精确转换	30.000071265	114.000250810	78.950051

8.5.4 地理坐标与完整转换算例：总体坐标转换

完整链的矩阵顺序与维数

$$ {}^{g}\mathbf{p}_{\mathrm{LLA}} = ned2lla(\pi_h(\mathbf{T}_b^e \mathbf{T}_c^b \mathbf{B}_c [u, v, 1]^{\mathsf{T}}), \quad \mathrm{origin}_{\mathrm{LLA}}) \tag{8.12} $$

$$ \mathbf{B}_c = [\, z_c(\mathbf{K}^{-1})^{\mathsf{T}} \quad \mathbf{e}_3 \,]^{\mathsf{T}} \in \mathbb{R}^{4\times 3}, \quad \mathbf{e}_3 = [0,0,1]^{\mathsf{T}} $$

$$ \pi_h([n, e, d, w]^{\mathsf{T}}) = [n/w, e/w, d/w]^{\mathsf{T}} \in \mathbb{R}^3, \quad w = 1 $$

$$ (4\times 4)(4\times 4)(4\times 3)(3\times 1) = (4\times 1) \overset{\pi_h}{\rightarrow} (3\times 1) $$

输入：校正像素、内参、深度、安装外参、同一时刻机体位姿、共同地理原点。

8.5.4 地理坐标与完整转换算例：完整转换算例

一个像素点得到可核验的全球位置

阶段	输入/运算	结果
像素到相机	像素(480,320)，Z4m 640×480，水平FOV90°	(2,1,4)m
相机到机体	轴置换，安装偏移(0.2,0.1,0.05)m	(4.2,2.1,1.05)m
机体到NED	偏航90°，位置(10,20,-30)m	(7.9,24.2,-28.95)m
NED到LLA	原点(30°,114°,50m)	(30.000071265°, 114.000250810°,78.950051m)

将LLA转回局部NED，再逆刚体变换并投影，回算像素为(480,320)。

8.5.4 地理坐标与完整转换算例：误差分析

角度、深度和时延误差沿链传播

误差源	本例一阶计算	影响量级
像素误差1px	$z_c/f_x = 4/320$	横向0.0125m
深度误差0.2m	$0.2[0.5,0.25,1]^{\mathrm{T}}$	(0.10,0.05,0.20)m
偏航误差1°	$\sqrt{4.2^2 + 2.1^2}\,\pi/180$	横向约0.082m
时间错位50ms	10m/s×0.05s	位置约0.5m

内/外参、安装变化、姿态、GNSS原点、地理模型和配准都会影响全局位置。

共享坐标也要共享误差与时间基准。

8.5.4 地理坐标与完整转换算例：误差分析与融合

多机共享保留观测时间和不确定度

图像/点云观测 → 共同坐标与时间 → 位置、目标ID、不确定度

共享字段：地理原点及高度基准、观测时刻、目标标识、位置协方差和来源标识。

融合方法	在坐标链中的角色
EKF/贝叶斯滤波	利用运动预测与新观测递推校正状态。
因子图优化	组合跨时刻、跨传感器或跨载具的几何约束。
深度融合网络	学习多模态特征关系，同时检查数据分布和置信校准。

练习与讨论

改变深度后重算全部局部坐标

条件：640×480、水平FOV90°，像素(480,320)，轴向深度改为8m。

$$\mathbf{R}_c^b = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad {}^b\mathbf{t}_c = [0.2, 0.1, 0.05]^{\mathrm{T}}\ \mathrm{m}$$

载具水平、偏航90°，局部位置(10,20,-30)m，地理原点(30°,114°,50m)。

1. 求相机点、机体点、局部NED点，并标出方向。
2. 判断安装偏移、斜距和纬经高顺序是否需要改变。

参考要点

相同像素、双倍深度的坐标结果

步骤	参考结果	方向解释
相机点	(4,2,8)m	右4m、下2m、前8m
机体点	(8.2,4.1,2.05)m	前8.2m、右4.1m、下2.05m
局部NED点	(5.9,28.2,-27.95)m	北5.9m、东28.2m、上27.95m

平移保持不变：安装与载具位置不随目标深度翻倍。

斜距/顺序：$r = \sqrt{84}\ \mathrm{m}$，LLA为纬度、经度、椭球高。

本节小结

几何、尺度、位姿和基准共同决定位置

几何与尺度：像素和内参确定视线，深度或其他几何约束确定三维点。

刚体位姿：安装外参和机体位姿分别处理相机到机体、机体到地面。

共同基准：局部原点、NED/ENU、WGS84椭球高和观测时间共同定义位置。

易错点：斜距当深度、混用单位、误用外参方向、LLA直接相加和错用旧位姿。