

基于 PnP 解算器的空间目标位姿估算与追踪

冯唐正

南京理工大学

电子工程与光电技术学院

版本：0.10

日期：2023 年 12 月 15 日

摘 要

3D 目标位姿估算是机器视觉领域的重要研究方向之一。本文介绍了一种基于 PnP 解算器的目标位姿估算方法，通过针孔相机获取目标图像，经过畸变矫正后，使用 IPPE 模式的 PnP 解算器计算每一帧图像中目标表面中心在相机参考系的坐标信息以及目标的表面在三维空间中的姿态信息。通过估算结果，得到航空器三主轴旋角，以此驱动云台使得步兵机器人发射机构跟踪目标。研究结果表明，IPPE 模式下的 PnP 解算器能够使得步兵机器人具有视觉上的感知能力。

关键词：位姿估算；PnP 解算；机器视觉系统；针孔相机；目标追踪

Abstract

3D target pose estimation is one of the important research directions in the field of machine vision. This article introduces a target pose estimation method based on a PnP solver. The target image is obtained through a pinhole camera, and after distortion correction, the PnP solver in IPPE mode is used to calculate the coordinate information of the target surface center in the camera reference frame and the pose information of the target surface in three-dimensional space for each frame of the image. By estimating the results, the three-axis rotation angle of the aircraft is obtained, which drives the gimbal to enable the infantry robot launch mechanism to track the target. The research results indicate that the PnP solver in IPPE mode can enable infantry robots to have visual perception ability.

Keywords: Pose estimation; PnP calculation; Machine vision system; Pinhole camera; Target tracking

1 引言

在机器视觉系统中，目标位姿估算是一个非常重要的任务，它涉及到对目标物体在三维空间中的位置和姿态的解算。这种估算可以帮助机器视觉系统更好地理解和分析目标物体的特征和属性，从而更好地进行识别、追踪、定位等任务。目标位姿估计能够提高机器视觉系统的感知能力和交互能力。通过估算，机器视觉系统可以更好地了解目标物体的状态和运动轨迹，同时，更好地与人类或者其他智能系统进行交互和协作。

目标位姿估算已经成为机器视觉领域的重要研究方向之一。近年来，随着深度学习技术的快速发展，目标位姿估算也取得了很大的进展。许多学者提出了各种不同的目标位姿估算算法，如基于卷积神经网络的目标位姿估算算法 [5]、基于回归器的目标位姿估算算法 [4] 等。这些算法在精度和计算效率上都有所提高，同时也能够更好地适应不同的场景和环境。目前，许多算法已经能够实现高精度、高效率的目标位姿估算。然而，在实际应用中，目标位姿估算仍然存在一些挑战和难点。例如，对于遮挡、过曝、噪声等问题，目标位姿估算的性能可能会受到影响。因此，研究者们仍然在不断探索新的方法和技术，以进一步提高目标位姿估算的准确性和鲁棒性。

本文通过 PnP 解算器，使用 OpenCV 提供的 SOLVEPNP_IPPE[1] 模式对图像中的 3D 物体表面进行位姿估算。如图 1 所示，设若空间中存在待检测 3D 目标。其 $w_X * h_X$ 的表面在图像中投影为尺寸为 $w * h$ 的四边形。通过 PnP 解算器，我们可以得到该表面中心到相机光学中心的距离以及目标物体该表面在三维空间中的姿态矩阵。

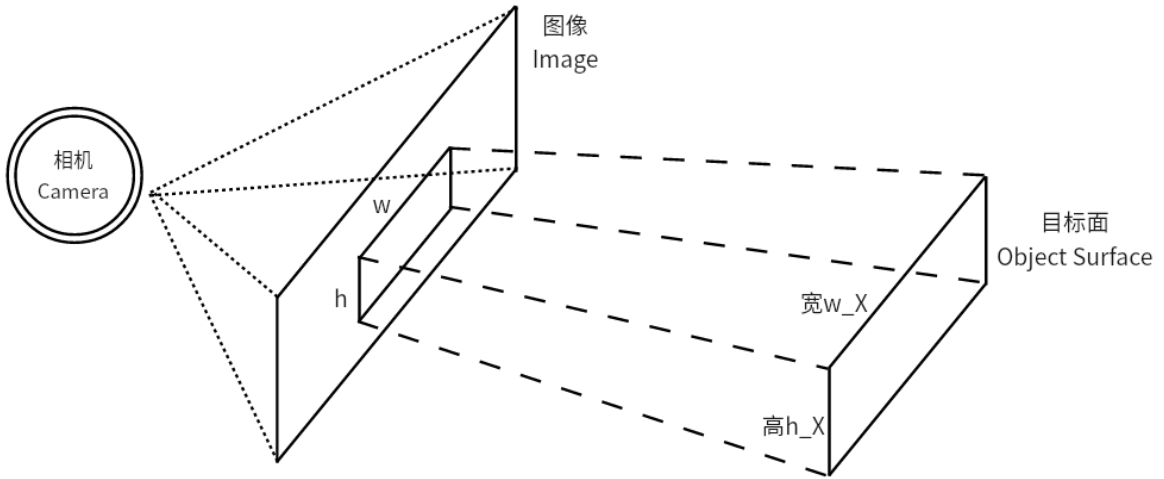


图 1: 问题场景

本文将详细阐述针孔相机模型及其物象变换过程，并介绍 OpenCV 提供的 PnP 解算器 IPPE 模式的使用方法。随后，将深入探讨相机畸变矫正的原理，并给出相应的矫正公式。最后，我们将进行实验以验证 PnP 解算的效果以及目标追踪的准确性。

本研究表明，IPEE 模式的 PnP 解算器在步兵机器人的视觉处理中具有重要作用，能够提高步兵机器人的观察、感受和追踪目标的准确性。后续研究将进一步优化估算信息，以提高步兵机器人的视觉准确性和鲁棒性。

2 数学建模

2.1 针孔相机模型

针孔相机模型 (Pinhole Camera Model, PCM) 如图 2 所示，其中，黑色标注的 \mathcal{F}_c 为相机参考系 (Camera Frame of Reference, CFR)，黄色参考系为像素参考系 (Pixel Frame of Reference, PFR)，蓝色参考系为图像参考系 (Image Frame of Reference, IFR)。在光轴 (Optical Axis) 上

距离光轴 $z = f$ 处为图像原点 (c_x, c_y) ，世界参考系 (*World Frame of Reference, WFR*) 下的点 $P = (X, Y, Z)^T$ 在图像上的投影为 (u, v) ，世界参考系在图中未标出。

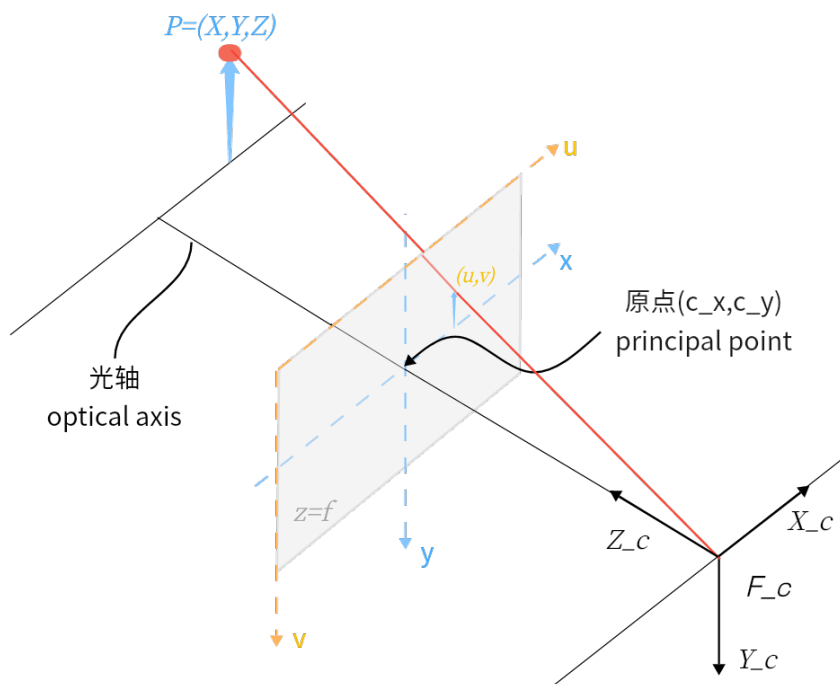


图 2: 针孔相机模型

2.2 参考系变换

从世界参考系变换到相机参考系，给出其刚体变换矩阵 \mathcal{T}

$$\mathcal{T} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

令 \mathcal{R} 与 \mathbf{t} 写成如下形式，

$$\mathcal{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad (2)$$

则变换矩阵 \mathcal{T} 可以写成

$$\mathcal{T} = \begin{bmatrix} \mathcal{R}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ O & 1 \end{bmatrix} \quad (3)$$

写出点 P 的其次坐标

$$P_W = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4)$$

则在相机参考系下点 P_C 为

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \mathcal{T} P_W \quad (5)$$

从相机参考系变换到图像参考系的变换为透视变换，为了保证计算的连续性，将变换矩阵写为

$$\mathcal{S} = \begin{bmatrix} \frac{f}{Z_C} & 0 & 0 & 0 \\ 0 & \frac{f}{Z_C} & 0 & 0 \\ 0 & 0 & \frac{1}{Z_C} & 0 \end{bmatrix} \quad (6)$$

其中， f 为焦距，它们的单位是 mm 。图像参考系和像素参考系的变换包括平移和缩放，其变换矩阵为

$$\mathcal{M} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

其中， dx 与 dy 分别每个感光器的物理尺寸，即单位像素的物理尺寸大小，它们的单位是 $mm/pixel$ 。写出像素参考系中目标投影的齐次坐标 $p = \begin{bmatrix} u & v & 1 \end{bmatrix}^T$ ，假设相机已经经过畸变校正，则有

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathcal{M} \mathcal{S} \mathcal{T} P_W \quad (8)$$

即

$$Z_C \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \quad (9)$$

其中， f_x 与 f_y 分别为 x 与 y 方向上的像素焦距，它们的单位是 $pixel$ ，且

$$f_x = \frac{f}{dx}$$

$$f_y = \frac{f}{dy}$$

而 Z_C 为

$$[Z_C] = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathcal{R}_{3 \times 3} & \mathbf{t}_{3 \times 1} \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \quad (10)$$

2.3 Solve PnP IPPE

将目标面中心作为世界坐标系原点，目标平面为 x - y 平面，则目标表面的四个角点坐标为

$$\begin{aligned} p_1 &= \begin{bmatrix} -\frac{h}{2} \\ -\frac{w}{2} \\ 0 \end{bmatrix} & p_2 &= \begin{bmatrix} +\frac{h}{2} \\ -\frac{w}{2} \\ 0 \end{bmatrix} \\ p_3 &= \begin{bmatrix} -\frac{h}{2} \\ +\frac{w}{2} \\ 0 \end{bmatrix} & p_4 &= \begin{bmatrix} +\frac{h}{2} \\ +\frac{w}{2} \\ 0 \end{bmatrix} \end{aligned} \quad (11)$$

OpenCV 提供的 SolvePnP IPPE 模式要求输入 4 个及以上的三维点，且点均在同一个平面，它将计算出 \mathcal{R} 与 \mathbf{t} 。其中 \mathbf{t} 即为装甲板中心在相机参考系的坐标。

2.4 畸变矫正

相机经过透镜成像存在畸变。我们需要对其进行畸变矫正。[2]

2.4.1 径向畸变

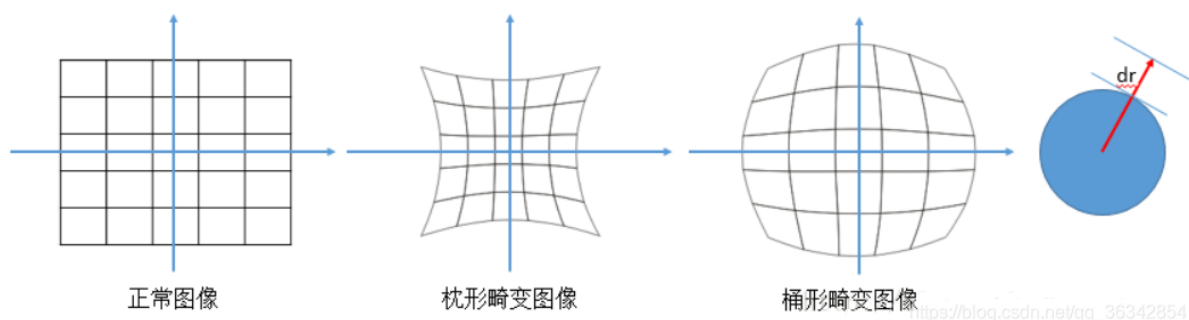


图 3: 径向畸变

径向畸变（如图 3）产生的原因是光线在远离透镜中心的地方偏折更大。其矫正公式如下（其中取带皮亚诺余项的 6 阶泰勒级数展开式）

$$\begin{cases} x_{dr} = x(1 + k_1r^2 + k_2r^4 + k_3r^6 + o(r^6)) \\ y_{dr} = y(1 + k_1r^2 + k_2r^4 + k_3r^6 + o(r^6)) \end{cases} \quad (12)$$

式中， (x, y) 是理想坐标， x_{dr} 与 y_{dr} 是畸变后的像素点坐标，且： $r^2 = x^2 + y^2$ 。

2.4.2 切向畸变

切向畸变（如图 4）产生的原因是透镜不完全平行于图像平面。其矫正公式为

$$\begin{cases} x_{corrected} = x + [2p_1xy + p_2(r^2 + 2x^2)] \\ y_{corrected} = y + [p_1(r^2 + 2y^2) + 2p_2xy] \end{cases} \quad (13)$$

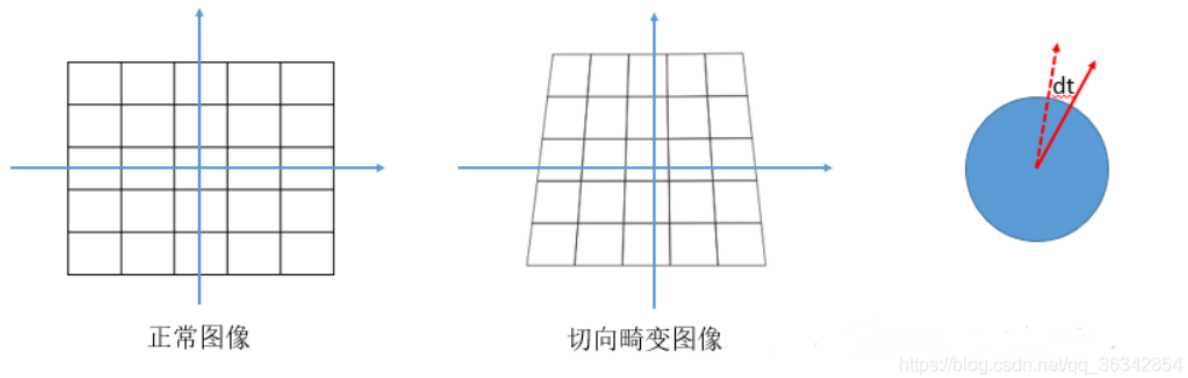


图 4: 切向畸变

由此，我们得到相机的 5 个畸变参数

$$D = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ p_1 \\ p_2 \end{bmatrix} \quad (14)$$

在实际应用中，需要对相机进行标定 [3, 6]，以此得到一系列畸变坐标 (x, y) 与矫正坐标 (x', y') ，最后通过这些数据计算得到畸变矫正的相机内参 D 。

3 实验

3.1 目标装甲板

要追踪的目标如图 5 所示，所以在公式 11 中 h 与 w 分别为 125 与 135。

3.2 惯用系变换

我们将相机参考系更改为右手坐标系（如图 6）。在这个新的坐标系中，我们将原系的 z 轴作为变换系的 x 轴，原系的 $-x$ 轴作为变换系的 y 轴，原系的 $-y$ 轴作为变换系的 z 轴。

并给出其变换矩阵 15:

$$\mathcal{F}_{PTZ} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \quad (15)$$

3.3 云台参考系

云台的航空器主轴（Aircraft Principal Axes, APA, 如图 7）丢失了 Roll 轴，通过 Yaw 轴与 Pitch 轴控制云台。在实际应用中 Pitch 轴通常存在机械约束。

将惯用系中的坐标信息变换到 Pitch、Yaw 轴上的旋角。

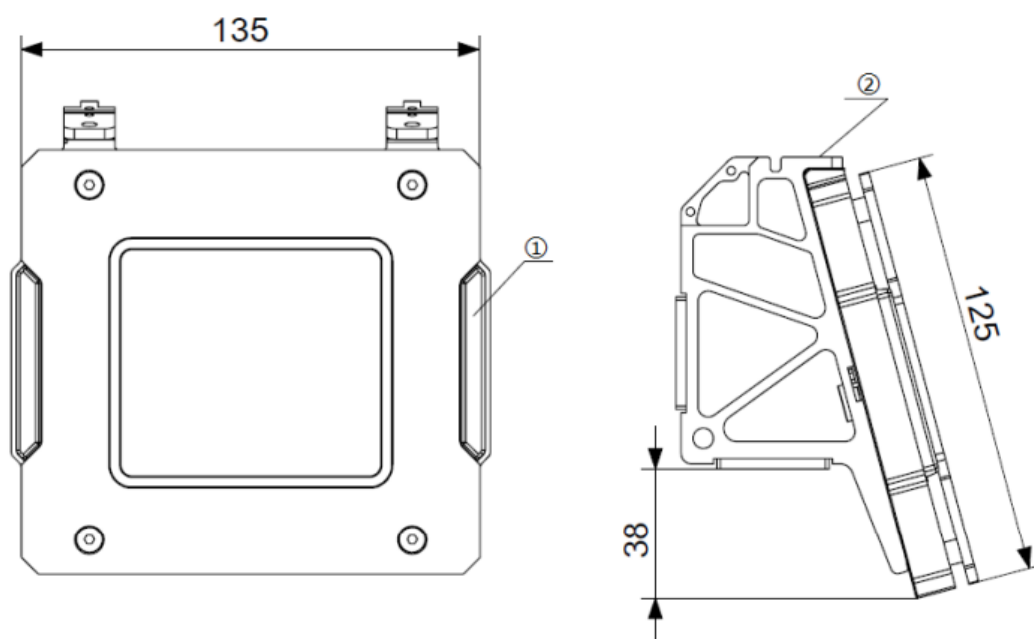


图 5: 目标装甲板

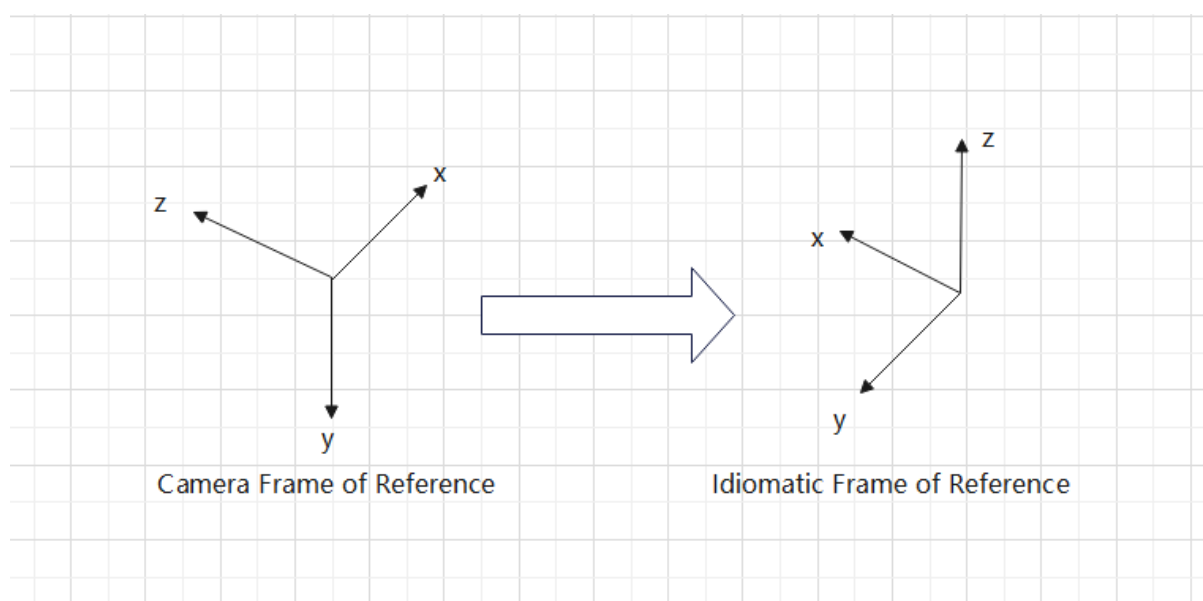


图 6: 惯用系变换

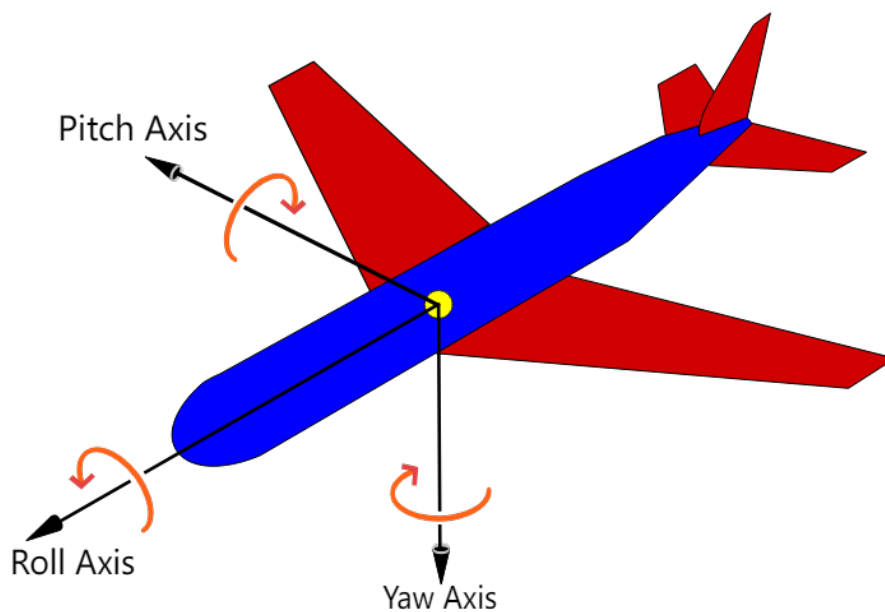


图 7: 航空器三主轴

$$\Delta yaw = \arctan -\frac{y}{x} \quad (16)$$

$$\Delta pitch = \arctan \frac{z}{x} \quad (17)$$

由于云台采用右手系，上述旋角在编程中应取负值。

3.4 实验结果

使用云台追踪目标装甲板。在机械约束内，步兵机器人能够即时解算出 $\Delta pitch$ 与 Δyaw ，并通过云台使得发射机构追踪目标装甲板（如图 8 所示）。

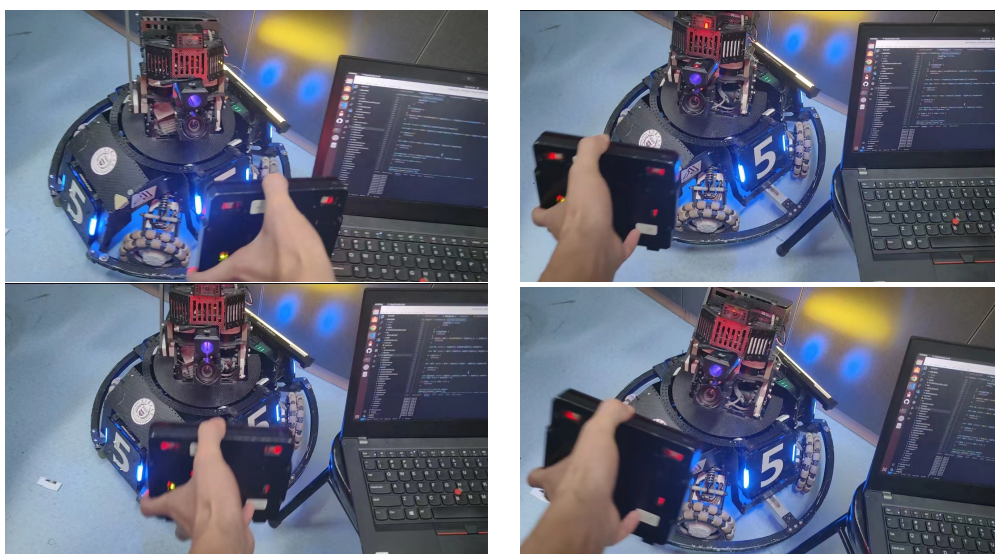


图 8: 云台追踪结果

4 结论

本文探讨了一种基于 PnP 解算的空间目标位姿估计及其追踪方法。首先,阐述了针孔相机成像与相机畸变矫正的机制。然后,使用 IPEE 模式的 PnP 解算器将目标从图片参考系变换到相机参考系。在此过程中,分别以刚体变换的旋转矩阵、平移矩阵作为空间目标的姿态信息与位置信息。接着,将位姿信息变换为航空器三主轴旋角,并使用云台追踪空间目标,以此验证 PnP 计算结果的准确性。

本研究的结果表明,IPEE 模式的 PnP 解算器能够作为步兵机器人在视觉上的处理器,从平面信息中得到空间信息。通过这种方法,步兵机器人能够更准确地观察、感受并追踪目标。在后续的研究中,基于拓展卡尔曼滤波的研究将对该估算信息做进一步的修正与预测,以提高步兵机器人在其视觉上的准确性与鲁棒性。

参考文献

- [1] Eric Marchand, Hideaki Uchiyama, and Fabien Spindler. "Pose Estimation for Augmented Reality: A Hands-On Survey". In: *IEEE transactions on visualization and computer graphics* 22.12 (Jan. 2016), pp. 2633–2651.
- [2] 张静, 秦文健, and 贾亮. "摄像机畸变矫正研究的综述". In: *科技创新导报* 7 (Jan. 2011), pp. 5, 7.
- [3] 林慧英 et al. "基于图像畸变矫正的摄像机标定方法". In: *吉林大学学报 (工学版)* 37.2 (Jan. 2007), pp. 433–437.
- [4] 王凯祺. "基于单张 RGB 图像的三维物体重建和位姿估计". 硕士论文. 华南理工大学, 2020.
- [5] 王静 et al. "基于深度学习的相机位姿估计方法综述". In: *计算机工程与应用* 59.7 (Apr. 2023), pp. 1–14.
- [6] 胡海秀 et al. "智能空间中基于畸变矫正的改进视觉标定". In: *机械与电子* 11 (Jan. 2011), pp. 67–70.