

DOI:10.19344/j. cnki. issn1671-5276.2024.01.039

基于双目视觉的绳驱动飞行机械臂目标识别与抓取

谭荣凯^a, 杨浩秦^b, 王尧尧^a

(南京航空航天大学 a. 机电学院; b. 材料科学与技术学院, 江苏 南京 210016)

摘要:为提高面向飞行机械臂平台的目标姿态估计的精度,为飞行机械臂配备双目视觉系统,对绳驱动机械臂和双目相机进行系统建模,利用 Camshift 算法对目标进行实时追踪,结合目标的几何特征提出一种轻量化的目标姿态估计算法,可用于飞行平台的实时目标追踪与姿态估计并进行飞行机械臂抓取实验。实验证明:该姿态估计算法对于目标追踪和姿态估计具有较好精度,可以实现抓取作业。

关键词:绳驱动飞行机械臂;双目相机;Camshift;目标追踪;姿态估计

中图分类号:TP249 **文献标志码:**A **文章编号:**1671-5276(2024)01-0191-05

Target Tracking and Grabbing of Cable-driven Aerial Manipulator Based on Binocular Vision

TAN Rongkai^a, YANG Haoqin^b, WANG Yaoyao^a

(a. College of Mechanical and Electrical Engineering; b. College of Materials Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: To improve the accuracy of target pose estimation on aerial manipulator platform, the aerial manipulator is equipped with a binocular vision system, and the cable-driven manipulator and the binocular camera are modeled. By the Camshift algorithm and with the geometric characteristics of the target, a lightweight target pose estimation algorithm is proposed, which can be used for real-time target tracking and pose estimating on aerial platform as well as for grasping experiment of aerial manipulator. Experiments show that the proposed pose estimation algorithm has good accuracy for target tracking and pose estimating, and can undertake grasping operation.

Keywords: cable-driven aerial manipulator; binocular camera; Camshift; target recognition; pose estimation

0 引言

目前无人机和多自由度机械臂所组成的空中操作系统由于其特有的优势,受到越来越广泛的关注。它们可以主动地和周围环境进行交互,在一些特殊的场景代替人类进行工作,有着非常广阔的应用前景。为了进一步提高飞行机械臂对周围环境的感知能力,对空中机械臂系统增加了视觉反馈系统,辅助机械臂进行作业。

空中机械臂的视觉系统采用的相机主要有3类:单目相机、双目相机和深度相机。单目相机体积小、价格低,但是其无法获取目标的深度信息;深度相机通过红外可直接获取目标的深度信息,但是其体积大,价格高,不适合用于无人机的实时监测;双目相机通过左右图像的视差来计算目标的深度信息,其体积和价格介于单目相机和深度相机之间,可以满足无人机实时监测的要求。在

机械臂工作过程中,视觉系统承担的主要工作为目标追踪与姿态估计。目标追踪算法经过多年的研究,已经有了长足的发展,传统的算法比如均值漂移算法、基于 Kalman 滤波的目标追踪等以及目前时兴的深度学习算法。深度学习可以识别许多种类的目标,精度较高,但是深度学习依靠大量的计算资源,且需要大量的样本进行训练,而传统的目标追踪算法所需资源较少,适用于无人机的实时追踪任务。

近些年来,有很多学者在无人机的视觉领域作了很多工作。韩国首尔大学的 SEO 等^[1]基于单目相机开发了一种基于图像的圆柱体检测算法,该算法利用 3D 空间中透视投影圆的几何特征,实现了圆柱体的抓取。宾夕法尼亚大学的 THOMAS 等^[2]从快速移动的猛禽获得灵感,基于单目相机在图像空间中研究了无人机的动力学模型,在视觉反馈的基础上生成运动轨迹,为无人机的抓取作了铺垫。

基金项目:国家自然科学基金项目(52175097);浙江大学流体动力与机电系统开放基金项目(GZKF-201915)

第一作者简介:谭荣凯(1998—),男,山东临沂人,硕士研究生,研究方向为飞行机械臂,15151866578@163.com。

日本立命馆大学的 LEEWIWATWONG 等^[3]研究了一种新的图像区域分离方法,通过霍夫变换来实现对棒状物体的识别,从而控制空中机械臂系统自动与棒状物体对齐。塞尔维亚大学的 RAMON-SORIA 等^[4]使用深度相机,利用人工神经网络(ANN)来检测目标,然后通过对齐算法来获取目标的姿态。东南大学王营华等^[5]在目标上贴 ArUco 标签,在标签大小已知的情况下,利用 n 点透视算法求解摄像头位姿,进而利用摄像头位姿对空中机械臂系统进行分级控制,实现对目标的抓取。哈尔滨工业大学 LUO 等^[6]利用 ORB 特征获取目标的特征点,在已知目标物真实大小的情况下,求解目标的姿态信息。

目前许多姿态估计算法采用深度学习的方式或者采用特征点的方式,计算量较大,对无人机机载计算机性能要求较高,且实时性难以保证。本文通过颜色特征结合目标的几何特征对目标进行追踪和姿态估计,采用的目标特征较少且计算方便,适合无人机等实时性要求较高的平台。

本文做的工作主要如下:

1) 提出了一种轻量化的目标姿态估计算法,通过改进 Camshift 算法,结合目标几何特征对目标进行姿态估计,可用于实时性要求较高的飞行作业平台;

2) 对机械臂和相机进行了系统建模,求解各坐标系之间的转换关系,通过逆运动学运算求解已知目标姿态时各关节的抓取角度。

1 系统整体建模

1.1 绳驱动机械臂建模

机械臂采用绳索驱动的形式。绳驱动机械臂将电机和减速器安装在机械臂的底座处,通过绳索传递运动与力,而传统机械臂将电机和减速器安装在机械臂的关节处,在臂的运动过程中,机械臂的质量会导致飞行器本身的质心发生迁移,降低飞行器的控制精度和作业质量,而绳驱动机械臂将电机安装在底座位置可以降低机械臂惯性对飞行器控制的不利影响。

使用 D-H 参数法对绳驱动机械臂进行建模,建立绳驱动连杆坐标系如图 1 所示。设关节 i 绕 z_i 轴正方向的转动角度为 θ_i ,绕 x_i 轴正方向的转动角度为 α_{i-1} ,连杆沿 z_i 轴正方向的移动距离为 d_i ,连杆长度为 A_i 。

由于机械臂只在平面内进行运动,所以 3 个

关节绕 x_i 轴正方向的转动角度 α_{i-1} 以及连杆沿 z_i 轴正方向的移动距离 d_i 均为 0,各连杆长度为: $A_0 = 10 \text{ cm}$, $A_1 = 35 \text{ cm}$, $A_2 = 35 \text{ cm}$, $A_T = 25 \text{ cm}$ 。

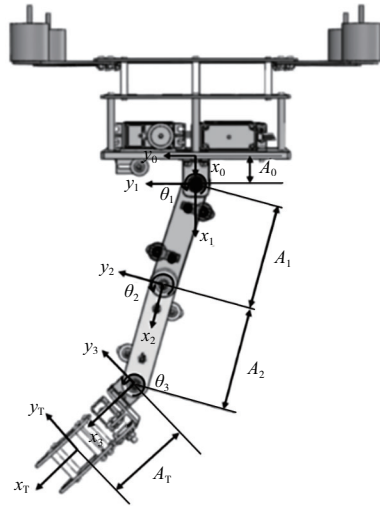


图 1 绳驱动机械臂坐标系

将世界坐标系的原点设于距第一关节 A_0 处底盘的中心位置,末端执行器坐标系 W_T 与世界坐标系 W_0 的转换矩阵如下,其中 ${}^a_b T$ 表示坐标系 a 到坐标系 b 的转换矩阵。

$${}^0_T T = {}^0_1 T {}^1_2 T {}^2_3 T {}^3_T T = \begin{bmatrix} c_{123} & -s_{123} & 0 & A_0 + A_1 c_1 + A_2 c_{12} + A_T c_{123} \\ s_{123} & c_{123} & 0 & A_1 s_1 + A_2 s_{12} + A_T s_{123} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

式中: $c_1 = \cos\theta_1$; $c_{12} = \cos(\theta_1 + \theta_2)$; $c_{123} = \cos(\theta_1 + \theta_2 + \theta_3)$; $s_1 = \sin\theta_1$; $s_{12} = \sin(\theta_1 + \theta_2)$; $s_{123} = \sin(\theta_1 + \theta_2 + \theta_3)$ 。

通过立体视觉可以获得目标在空间的位姿,将位姿作为机械臂末端执行器的期望位姿,通过运动学逆解,求解机械臂各关节的角度,再通过控制机械臂关节运动,即可实现抓取任务。可用末端执行器在底部坐标系 x, y 方向上的坐标以及相当于 z 轴的旋转来描述其在水坐标系下的位姿,其变换矩阵如下式:

$${}^0_T T = \begin{bmatrix} c\theta & -s\theta & 0 & x \\ s\theta & c\theta & 0 & y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

式中: c 为 \cos ; s 为 \sin 。

联立式(1)和式(2),通过代数法进行逆运动学求解,求解过程如下所示。

$$\begin{bmatrix} c_{123} & -s_{123} & 0 & A_0+A_1c_1+A_2c_{12}+A_Tc_{123} \\ s_{123} & c_{123} & 0 & A_1s_1+A_2s_{12}+A_Ts_{123} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} c\theta & -s\theta & 0 & x \\ s\theta & c\theta & 0 & y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

通过对机械臂关节角进行分析以及对上式的联立可得下式,其中 X, Y 为中间变量。

$$\begin{cases} \theta_1 + \theta_2 + \theta_3 = \theta \\ c_1 + c_{12} = X \\ s_1 + s_{12} = Y \\ X = \frac{x - A_0 - A_T c\theta}{A_1} \\ Y = \frac{y - A_T s\theta}{A_1} \end{cases} \quad (4)$$

最终可得:

$$\begin{cases} \theta_1 = \beta - \alpha \\ \theta_2 = \arccos\left(\frac{X^2 + Y^2 - 2}{2}\right) \\ \theta_3 = \theta - \theta_1 - \theta_2 \end{cases}; \begin{cases} \theta_1 = \beta + \alpha \\ \theta_2 = -\arccos\left(\frac{X^2 + Y^2 - 2}{2}\right) \\ \theta_3 = \theta - \theta_1 - \theta_2 \end{cases} \quad (5)$$

式中: $\beta = \arctan 2(Y, X)$; $\alpha = \arccos(\sqrt{X^2 + Y^2}/2)$ 。

由于机械臂采用绳驱动的方式,使得机械臂关节在运动过程中会对其他关节产生影响,即关节之间存在耦合效应,所以关节期望的转角 θ_a 与实际电机的转角 θ_m 并不一致,两者之间存在一个系数 I ,即 $\theta_m = I\theta_a$,限于篇幅,在此直接给出 I 如下所示。

$$I = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \quad (6)$$

1.2 双目相机建模

双目相机相对于单目相机的优势是可以获取物体的深度信息,双目相机的左右两个摄像头可以同时采集信息,通过比较物体在两幅图像中的位置差异,可以获取物体的深度信息。本文所用的双目相机基线长度可变,通过调节基线长度可以实现探测深度的改变,实验中相机采用的基线长度为 3 cm,可探测深度范围为 15 cm ~ 80 cm 之间,满足机械臂抓取要求。

图 2 所示为成像过程中涉及的 4 个坐标系,

分别为世界坐标系 $\{O_W\}$ 、相机坐标系 $\{O_C\}$ 、左右成像坐标系 $\{O_L\}$ 、 $\{O_R\}$ 和位于成像坐标系左上角的像素坐标系,通过各个坐标系之间的转换关系,可以获取二维平面上的物体在三维空间中的位置关系。

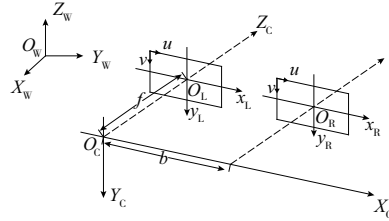


图 2 坐标系之间的关系

像素坐标系与相机坐标系之间的转换关系如下:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_C} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} = \frac{1}{Z_C} \mathbf{K} \mathbf{P}_C \quad (7)$$

式中 \mathbf{K} 称为相机的内参数矩阵,通过标定获得。

相机坐标系与世界坐标系之间的转换关系如下:

$$\mathbf{P}_C = \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = {}^C_W \mathbf{R} \mathbf{P}_W + {}^C_W \mathbf{T} \quad (8)$$

式中: \mathbf{P}_C 是点在相机坐标系下的坐标; ${}^C_W \mathbf{R}$ 和 ${}^C_W \mathbf{T}$ 是世界坐标系转换到相机坐标系的旋转矩阵和平移向量,这两个参数又被称作相机的外参数。

经过标定后,得到相机坐标系到世界坐标的旋转矩阵 ${}^W_C \mathbf{R}$ 和平移向量 ${}^W_C \mathbf{T}$ 分别为:

$$\begin{cases} {}^W_C \mathbf{R} = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ {}^W_C \mathbf{T} = [-30 \quad 120 \quad 180] \end{cases} \quad (9)$$

2 目标追踪与姿态识别

Camshift 算法的主要流程如下:给定 d 维空间 R 的 n 个样本点 $x = [x_i]$, $i = 1, \dots, n$, 选定其中某一点 x_i , 做一个半径为 r 的圆, 其中有 k 个样本点落入圆中, 记这 k 个点的集合为 $\{S_k\}$, 落在圆内的所有点都与 x_i 构成一个向量, 将这些向量相加, 所得的结果称为 meanshift 向量。如图 3(a) 所示, P_1 即为选定的点, P_1 指向 P_2 的向量即这个时刻 meanshift 向量, 然后 P_2 作为新的选定点继续做圆, 如图 3(b) 所示。最终 meanshift 向量可以收敛

到概率密度最大的地方,如图 3(c)所示。

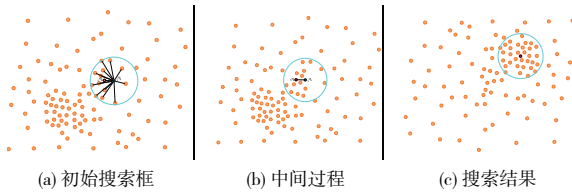


图 3 Camshift 算法

meanshift 向量通过搜索框搜索概率密度最大区域来确定目标对象,但其固有的缺陷,算法的收敛速度与选取的窗口有关,而且窗口一旦选定,窗口的大小不能改变。当目标在视野中发生变化,比如变大或变小时,窗口不能自适应的去调整大小。

为了能让搜索窗口的大小随着目标的变化自适应调整大小,本文利用目标的颜色直方图模型将图像转变为颜色概率分布图,同 meanshift 向量一样,初始化一个搜索窗口的位置和大小,根据搜索窗口得到的结果自适应调整下一帧搜索窗口的位置和大小,最终得到图片中颜色概率密度分布最大的区域。

想要将图像转变为颜色概率分布图,需要对图像进行反向投影。首先将图像从 RGB 颜色空间转换到 HSV 空间,可以降低光照的影响,对其中的 H 分量做直方图,在直方图中包含了不同的 H 分量值出现的概率,然后将图像中每个像素的值得用其颜色出现的概率替换,就得到了颜色概率分布图。

RGB 到 HSV 空间的转换关系如下^[7]:

$$H = \begin{cases} H_1, & B \leq G \\ 360^\circ - H_1, & B > G \end{cases} \quad (10)$$

$$H_1 = \cos^{-1} \left\{ \frac{0.5 [(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\} \quad (11)$$

$$S = \frac{\text{Max}(R, G, B) - \text{Min}(R, G, B)}{\text{Max}(R, G, B)} \quad (12)$$

$$V = \frac{\text{Max}(R, G, B)}{255} \quad (13)$$

假设以绿色作为目标颜色, Camshift 算法利用自适应的搜索框对感兴趣目标颜色区域 (ROI) 进行实时检测并返回 ROI 的相关信息,如 ROI 的外接矩形的 4 个顶点位置以及其中心位置。如此可以利用 Camshift 完成目标追踪的功能,并利用检测目标区域的相关信息进行姿态估计。算法返回的 ROI 外接矩形框的相关信息如图 4 所示。

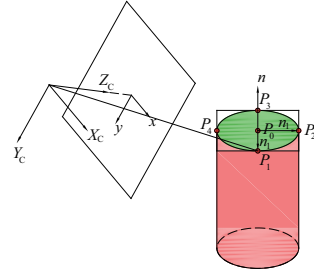


图 4 目标追踪与姿态估计

故而可以得到 P_0-P_4 4 个点在二维图像中的位置信息。设 $p_i (i=0\sim 4)$ 代表 P_0-P_4 在像素坐标系中的坐标, $P_{wi} (i=0\sim 4)$ 为世界坐标系中的坐标。通过式(7)和式(8)可得:

$$Z_c p_i = K \left[{}^c_w R P_{wi} + {}^c_w T \right] \quad (14)$$

由此可求得 P_{wi} , $[X_{wi} Y_{wi} Z_{wi}]$ 为 P_{wi} 的坐标。故圆柱体的半径 r 等价于中心点到任意点间的距离:

$$r = |P_{w0} P_{w3}| = \sqrt{(X_{w0} - X_{w3})^2 + (Y_{w0} - Y_{w3})^2 + (Z_{w0} - Z_{w3})^2} \quad (15)$$

由于圆柱体的几何特性,在求解圆柱体姿态时只考虑圆柱体的法向量。从 P_1-P_4 中任取两点,例如 P_1, P_2 , 设 P_0 指向 P_1 的向量为 n_1 , P_0 指向 P_2 的向量为 n_2 , 则圆柱体的法向量 n 可表示为

$$n = n_1 \times n_2 \quad (16)$$

3 实验结果

实验所用的机械臂为本课题组自行设计的绳驱动机械臂,用 STM32F103 单片机进行控制,相机采用免驱动可变基线双目相机,机载计算机为 NVIDIA JESTON NX,目标为圆柱形物体。主要控制流程为相机检测到目标物体后,通过机载计算机实时计算目标的大小与姿态信息,通过逆运动学解算后得到机械臂各关节的抓取角度,再将此抓取角度信息传递给单片机,单片机控制机械臂完成抓取过程,实验平台如图 5 所示。

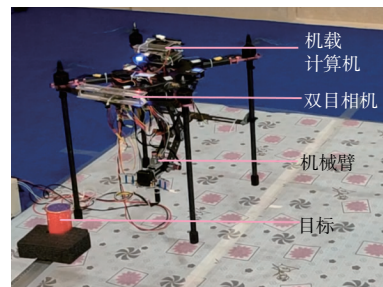


图 5 实验平台

图6所示为机载计算机通过算法估计的圆柱体目标的半径值,直线为圆柱体半径真实值为27.6 mm,曲线为视觉算法估计的半径值,误差为 $[-0.21, +0.25]$ mm。

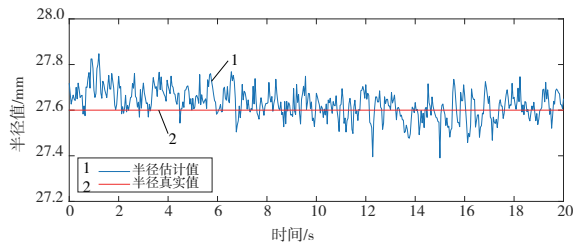


图6 圆柱体半径估计

图7所示为绳驱动飞行机械臂的整个抓取过程。理想状态下,无人机从预定位置起飞,在预定的轨迹上行进,当双目相机检测到目标后通过Camshift算法对目标进行实时追踪并引导无人机靠近目标。当目标处于机械臂的工作空间后,通过机载计算机实时计算目标在三维空间中的姿态信息,然后通过逆运动学解算出机械臂各关节期望运动角度,机载计算机将数据发送到单片机中,然后单片机控制机械臂抓取,抓取完成后无人机退回到预定位置。

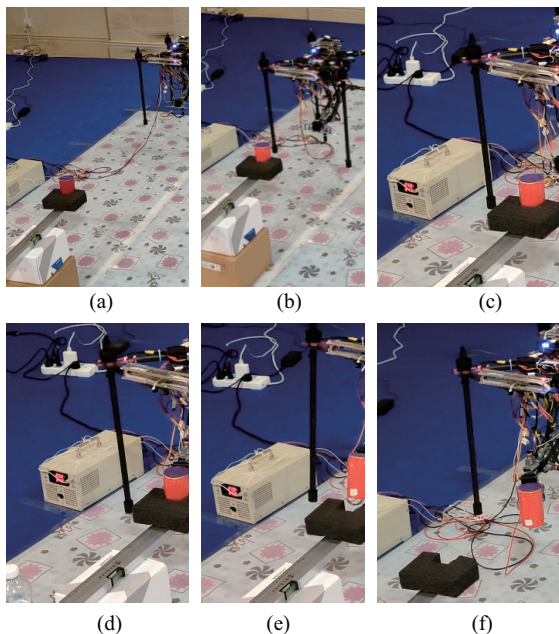


图7 绳驱动飞行机械臂抓取过程

4 结语

1) 针对飞行机械臂实时性要求较高的情况,提出了一种轻量化的目标姿态估计算法。

2) 建立了双目相机模型和绳驱动机械臂运动学模型,得到了目标在相机坐标系与世界坐标系的转换关系,通过逆运动学求解了在已知目标位姿情况下绳驱动机械臂的抓取角度。

3) 进行了目标抓取实验,验证了姿态估计算法的可行性。

参考文献:

- [1] SEO H, KIM S, KIM H J. Aerial grasping of cylindrical object using visual servoing based on stochastic model predictive control [C] // 2017 IEEE International Conference on Robotics and Automation (ICRA). Singapore: IEEE, 2017: 6362-6368.
- [2] THOMAS J, LOIANNIO G, SREENATH K, et al. Toward image based visual servoing for aerial grasping and perching [C] // 2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong, China: IEEE, 2014: 2113-2118.
- [3] LEEWIWATWONG S, SHIMAHARA S, LADIG R, et al. Vision based autonomous orientational control for aerial manipulation via on-board FPGA [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Las Vegas, NV, USA: IEEE, 2016: 854-860.
- [4] RAMON-SORIA P, ARRUE B C, OLLERO A. Grasp planning and visual servoing for an outdoors aerial dual manipulator [J]. Engineering, 2020, 6(1): 77-88.
- [5] 王营华, 宋光明, 刘盛松, 等. 一种视觉引导的作业型飞行机器人设计 [J]. 机器人, 2019, 41(3): 353-361.
- [6] LUO B, CHEN H, QUAN F, et al. Natural feature-based visual servoing for grasping target with an aerial manipulator [J]. Journal of Bionic Engineering, 2020, 17(2): 215-228.
- [7] DU Y C, TARYUDI T, TSAI C T, et al. Eye-to-hand robotic tracking and grabbing based on binocular vision [J]. Microsystem Technologies, 2021, 27(4): 1699-1710.

收稿日期: 2022-07-07