

· 计算机工程与应用 ·

基于强化学习的旋翼无人机智能追踪方法

史豪斌*, 徐 梦

(西北工业大学计算机学院 西安 710129)

【摘要】针对旋翼无人机追踪场景中常用的PID控制方法与视觉伺服控制方法的不足, 该文尝试将视觉伺服控制与强化学习结合, 提出了一种基于强化学习的旋翼无人机智能追踪方法。首先使用基于图像的视觉伺服实现旋翼无人机的闭环控制, 然后建立使用Sarsa学习算法调节伺服增益的强化学习模型, 通过训练可以使得旋翼无人机自主选择视觉伺服增益。该文设计了旋翼无人机在实物场景与仿真场景下的运动目标追踪实验, 实验结果论证了该方法相对于PID控制与基于图像的视觉伺服控制方法具有更好的追踪效果。

关键词 智能追踪; 强化学习; 旋翼无人机; 视觉伺服

中图分类号 TP39 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2019.04.012

An Intelligent Tracking Method of Rotor UAV Based on Reinforcement Learning

SHI Hao-bin* and XU Meng

(School of Computer Science, Northwestern Polytechnical University Xi'an 710129)

Abstract Aiming at the deficiencies of PID control method and visual servo control method commonly used in the tracking scene of Rotor UAV(unmanned aerial vehicle), this paper attempts to combine visual servo control with reinforcement learning, and proposes an intelligent tracking method for Rotor UAV based on reinforcement learning. Firstly, image-based visual servo is used to track the closed-loop control of the Rotor UAV, and then a reinforcement learning model is established to adjust the servo gain with Sarsa learning algorithm. After many training sessions, the Rotor UAV can choose its own visual servo gain. In this paper, the experiment of tracking the moving target of Rotor UAV in physical and simulation scenarios is designed. The experimental results demonstrate that the proposed method has better tracking effect than PID control and classical image-based visual servo control method.

Key words intelligent tracking; reinforcement learning; rotorcraft UAV; visual servo

近年来, 随着旋翼无人机相关技术的发展, 广大学者对旋翼无人机的追踪与控制问题进行了广泛的研究^[1-2]。文献[3]针对旋翼无人机追踪地面目标车辆过程中存在的精度不高以及参数固定的问题进行研究, 提出了一种模糊PID控制方法, 但是PID控制器在处理非线性不确定系统时的抗干扰能力差, 控制精度也不高。文献[4]应用了一种基于位置的视觉伺服(position-based visual servoing, PBVS)的控制方法实现旋翼无人机的追踪控制, 但是基于位置的视觉伺服的误差定义在三维笛卡尔空间, 对初始条件、噪声、摄像机参数误差和目标位姿的估计精度都非常敏感。文献[5]提出了一种新型的基于图像的视觉伺服旋翼无人机最优路径规划方法, 提高了旋翼无人机的控制效果。但是基于图像的视觉伺服控制对

于伺服增益的选取大多通过人工赋值的方式, 选取合适的伺服增益值往往依靠经验, 因此该方法不能很好地在复杂的非线性环境中实现精确的控制。

针对经典的PID控制与基于图像的视觉伺服控制在旋翼无人机追踪与控制过程中存在的收敛性差、控制精度不高等问题, 本文设计了一种基于强化学习的旋翼无人机追踪控制方法, 通过基于图像的视觉伺服形成旋翼无人机的闭环反馈控制, 结合强化学习调节伺服增益值, 可以实现旋翼无人机智能系统对于环境的适应能力。

1 旋翼无人机视觉伺服模型

1.1 基于图像的视觉伺服

旋翼无人机的视觉传感器中的 M 个特征点的

收稿日期: 2018-10-24; 修回日期: 2018-12-24

基金项目: 陕西省重点研发计划(2018GY-187)

作者简介: 史豪斌 (1978-), 副教授, 主要从事人工智能方面的研究。E-mail: shihaobin@nwpu.edu.cn

位置坐标为: $F = \{f_i^{\text{cp}} = (x_i^{\text{cp}}, y_i^{\text{cp}})^{\text{T}} \mid i = 1, 2, \dots, M\}$,
 设定 \boldsymbol{e} 为基于图像的视觉伺服的特征误差, 即:

$$\boldsymbol{e} = \begin{bmatrix} f_1^{\text{cp}} - f_1^{\text{des}} \\ f_2^{\text{cp}} - f_2^{\text{des}} \\ \vdots \\ f_M^{\text{cp}} - f_M^{\text{des}} \end{bmatrix} = \begin{bmatrix} x_1^{\text{cp}} - x_1^{\text{des}} \\ y_1^{\text{cp}} - y_1^{\text{des}} \\ x_2^{\text{cp}} - x_2^{\text{des}} \\ y_2^{\text{cp}} - y_2^{\text{des}} \\ \vdots \\ x_M^{\text{cp}} - x_M^{\text{des}} \\ y_M^{\text{cp}} - y_M^{\text{des}} \end{bmatrix} \in R^{2M \times 1} \quad (1)$$

式中, $f_i^{\text{des}} = (x_i^{\text{des}}, y_i^{\text{des}})^{\text{T}}$ 为第 i 个特征点的目标位置坐标, 该值是被唯一确定的。

对于旋翼无人机动力学模型, 有6个自由度的运动参数, 即 $\hat{\boldsymbol{v}} = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)^{\text{T}}$, 但是旋翼无人机是一种线速度与角速度之间强耦合的动力学系统。根据文献[6], 旋翼无人机在低速运动的状态下, 微小的姿态倾斜角度对视觉的影响可以忽略。本文设定, 考虑当旋翼无人机与地面近乎平行飞行时, 不会由于视觉计算而产生滚转角与俯仰角, 同时在实际应用场景中, 旋翼无人机的滚转角与俯仰角值几乎为零, 因此考虑 $\boldsymbol{v} = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)^{\text{T}}$ 。

根据特征误差与视觉传感器运动之间的关系为 $\boldsymbol{de}/\boldsymbol{dt} = \boldsymbol{J}\boldsymbol{v}$, 其中 \boldsymbol{J} 就是图像关联矩阵。为了计算方便, 将视觉坐标系统 ΣC 与旋翼无人机的惯性坐标系统 ΣE 统一, 旋翼无人机的前进与后退方向在 X 轴, 左右移动方向为 Y 轴。在一个理想的情景下, 设定 A 为视觉坐标系统 ΣC 空间中的一个三维坐标点, 坐标记为 (X_a, Y_a, Z_a) , 单位为 mm。设定 A 点在底部视觉传感器的图像平面上的成像点 a 的坐标为 (x_a, y_a) , 单位为 mm, 在数字图像中的坐标为 $(x_a^{\text{px}}, y_a^{\text{px}})$, 单位为 pixel, 成像模型如图1所示。

在底部视觉模型中引入放缩比例为 (d_x, d_y) , d_x 表示 x 方向上的每一个像素, 在成像图像平面上表示为 d_x 个单位长度, 相机的焦距为 f 。图像平面的中心坐标在 ΣC 中的坐标记为 $(0, 0, -f)$, (x_a, y_a) 为点 A 在图像平面中以图像中心点为原点的坐标, $(x^{\text{px}}, y^{\text{px}})$ 是数字图像中以其中心点为原点的坐标。在理想情况下, 视觉坐标系统与图像平面坐标系、像素平面坐标系的原点重合, 根据几何学得到数学关系如式(2)所示:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} -fa & 0 & 0 \\ 0 & -fa & 0 \\ 0 & 0 & a \end{bmatrix} \begin{bmatrix} X_a \\ Y_a \\ Z_a \end{bmatrix} \quad a \neq 0 \quad (2)$$

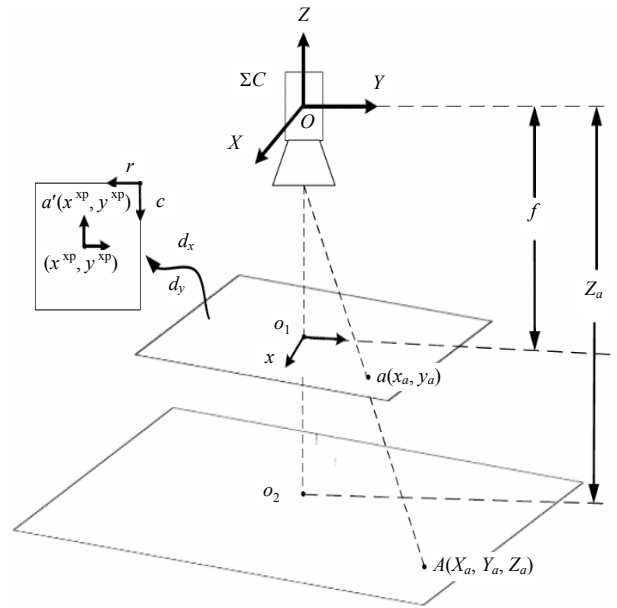


图1 旋翼无人机的视觉模型

设定数字图像中, r, c 分别表示行像素值与列像素值, 数字图像的中心点所在的行数为 $2r_d$, 列数为 $2c_d$, 设定 A 点在数字图像中的像素值行数为 a_r , 列数为 a_c , 可以得到:

$$\begin{cases} x_a^{\text{px}} = r_d - a_r \\ y_a^{\text{px}} = c_d - a_c \end{cases} \quad (3)$$

由数字图像与成像平面的映射关系可以得到:

$$\begin{cases} x_a = d_x x_a^{\text{px}} \\ y_a = d_y y_a^{\text{px}} \end{cases} \quad (4)$$

建立像素平面到三维视觉坐标系中任意空间点的映射关系:

$$\begin{bmatrix} x_a^{\text{px}} \\ y_a^{\text{px}} \\ 1 \end{bmatrix} = \begin{bmatrix} -fa/d_x & 0 & 0 \\ 0 & -fa/d_y & 0 \\ 0 & 0 & a \end{bmatrix} \begin{bmatrix} X_a \\ Y_a \\ Z_a \end{bmatrix} \quad (5)$$

由此旋翼无人机的底部视觉模型建立完成。其中 d_x, d_y, f 为旋翼无人机中视觉传感器的固有参数, 可以通过相机标定等方式来获得。

1.2 旋翼无人机的动力学扩展

旋翼无人机通过改变4个推进器的转动速度, 可以控制旋翼无人机上下运动。如果控制第一和第三推进器的旋转速度, 会引起俯仰运动; 如果控制第二和第四推进器的旋转速度, 会引起翻滚运动; 通过共同调节第一、第二、第三和第四推进器的旋转速度, 可使旋翼无人机产生偏航运动。旋翼无人机具体的推进器转子的旋转速度是由给定的加速度参数来实现控制的。其函数关系如式(6)所示:

$$[\omega_1, \omega_2, \omega_3, \omega_4]^{\text{T}} = \boldsymbol{J}_a [a_x^v, a_y^v, a_z^v, a_x^\omega, a_y^\omega, a_z^\omega]^{\text{T}} \quad (6)$$

式中, $\{\omega_1, \omega_2, \omega_3, \omega_4\}$ 为对应的推进器转子的转动速度; $\{a_x^v, a_y^v, a_z^v, a_x^o, a_y^o, a_z^o\}$ 为6自由度的加速度; J_a 是推进器转动速度到加速度的转换矩阵。

假定旋翼无人机保持水平飞行并且传入4个自由度的控制参数, 为了实现旋翼无人机x或y方向的运动, 需要一个滚转角 θ_x 或者俯仰角 θ_y 。当旋翼无人机在较低速的条件下飞行时, 滚转角 θ_x 或俯仰角 θ_y 由旋翼无人机内部控制对 v_x 和 v_y 进行运动学解析后计算得到。该方法为: 通过旋翼无人机的动力学原理计算出x、y方向的线加速度 a_x^v 、 a_y^v , 具体计算如式(7)所示:

$$\begin{cases} a_x^v = dv_x/dt \\ a_y^v = dv_y/dt \end{cases} \quad (7)$$

以x方向为例, 当给定了一个预定的加速度 a_x^v , 旋翼无人机的运动状态的简化模型如图2所示。在图2中, 设定 θ_x^* 为旋翼无人机在x方向的期望滚转角, 可计算获得:

$$\tan \theta_x^* = \frac{ma_x^v}{mg} \Rightarrow \theta_x^* = \arctan \frac{a_x^v}{g} \quad (8)$$

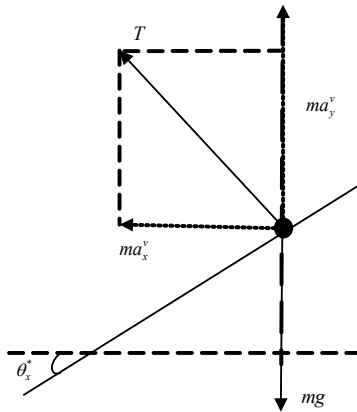


图2 x方向运动受力分析

当前旋翼无人机在x方向实际滚转角 θ_x^c 由旋翼无人机自身携带的陀螺仪获取。x方向的角速度 ω_x 通过公式 $\omega_x = (\theta_x^* - \theta_x^c)/\Delta t$ 计算得到。其中, Δt 为设定的运动时间, 同理可以计算出y方向的角速度 ω_y , 计算公式为:

$$\begin{cases} \theta_y = \arctan(a_y^v/g) \\ \omega_y = (\theta_y^* - \theta_y^c)/\Delta t \end{cases} \quad (9)$$

式中, θ_y^* 为旋翼无人机在y方向的期望俯仰角; θ_y^c 为陀螺仪获取的y方向当前角度。

通过该方法可以计算获得x和y方向的角速度, 4自由度的速度扩展为 $v = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)^T$ 。

1.3 图像关联矩阵

图像关联矩阵建立了从特征点误差变化率到旋

翼无人机的线速度、角速度之间的转换关系。为了保证误差的指数解耦, 设定 $de/dt = -\lambda e$, 其中 λ 为参数, 称其为伺服增益。为了保证在各个维度的变化相同进而保证系统的稳定性, 设定伺服增益值为非负值, 特征点误差到旋翼无人机运动参数之间的函数关系可以表示为:

$$-\lambda e = Jv \quad (10)$$

式中, J 为图像关联矩阵, 可以得到旋翼无人机的运动参数:

$$v = -\lambda J^+ e \quad (11)$$

式中, $J^+ \in R^{4 \times 2M}$ 为关联矩阵的广义逆矩阵。如果设定 J^+ 为方阵, $J^+ = J^{-1}$, 但这种方法能处理的特征点的数量有限, 文献[6]提出了使用摩尔-彭若思广义逆矩阵计算 $J^+ = (J^T J)^{-1} J^T$, 本文使用 $J^+ = \frac{1}{2} [(J^T J)^{-1} J^T + (J_*^T J_*)^{-1} J_*^T]$, 其中 J_* 是当 $e = 0$ 时的关联矩阵 J 。文献[7]证明了后者的效果优于使用摩尔-彭若思广义逆矩阵计算关联矩阵。当旋翼无人机与地面近乎直线飞行时, 可以通过姿态调节去除俯仰角与滚转角的影响[8]。

结合旋翼无人机的动力学特性, 根据运动学公式可得 (X_a, Y_a, Z_a) 随时间的变化率与线速度, 角速度的关系如下所示:

$$\begin{pmatrix} dX_a/dt \\ dY_a/dt \\ dZ_a/dt \end{pmatrix}^T = \begin{bmatrix} -1 & 0 & 0 & Y \\ 0 & -1 & 0 & -X \\ 0 & 0 & -1 & 0 \end{bmatrix} (v_x, v_y, v_z, \omega_z)^T \quad (12)$$

综合上述公式可以得到图像的关联矩阵:

$$J = \begin{bmatrix} -\frac{f}{Z \cdot d_x} & 0 & \frac{x}{Z \cdot d_x} & \frac{y}{d_x} \\ 0 & -\frac{f}{Z \cdot d_y} & \frac{y}{Z \cdot d_y} & -\frac{x}{d_y} \end{bmatrix} \quad (13)$$

v 具有4个自由度, 因此至少需要2个特征点才能求解, 但在计算过程中选取2个特征点容易陷入局部最小值的困境, 所以在实际应用中本文选取多个特征点。设定选取 K 个特征点, 可得到对应于 K 个特征点的图像关联矩阵 $J_K = [J_1, J_2, \dots, J_K]^T$ 。

由此, 本文建立了将 $(v_x, v_y, v_z, \omega_z)$ 作为旋翼无人机控制参数的旋翼无人机的视觉伺服模型。在实际应用中, 滚转角与俯仰角特别小但是不为零, 利用旋翼无人机的动力学原理, 通过 $v = (v_x, v_y, v_z, \omega_z)^T$ 来计算出预计的滚转角和俯仰角, 然后使用动力学扩展的方法计算出俯仰角速度 ω_x 和滚转角速度 ω_y , 即

将 $\mathbf{v} = (v_x, v_y, v_z, \omega_z)^T$ 扩展为 $\mathbf{v} = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)^T$ 再进行计算。

1.4 旋翼无人机视觉特征提取算法

本文使用基于轮廓的图像特征提取算法, 并使用费尔曼链码^[9]来描述并提取出目标的轮廓。具体过程为: 首先, 建立标准轮廓库, 并且将获取到的目标轮廓在标准轮廓库中找到其对应的形状, 标准轮廓库中的轮廓表示为 $D = \{D_i | i=1, 2, \dots, N\}$, 其次, 对目标进行轮廓特征提取得到的目标轮廓链码为: $C = \{C_i | i=1, 2, \dots, N\}$, 其中 K 为费尔曼链码的方向数, 最后, 使用一阶差分法对费尔曼链码进行旋转归一化, 旋转归一化的方法如公式(14)所示:

$$d_i = \begin{cases} C_i - C_N & i=1 \\ C_i - C_{i+1}, i > 1, & C_i - C_{i-1} > 0 \\ C_i - C_{i-1} & i > 1 \end{cases} \quad (14)$$

在旋转归一化费尔曼链码后, 计算提取出的轮廓与标准轮廓之间的Levenshtein距离^[10], 记为 L , 设置阈值, 如果 L 小于阈值则目标轮廓被正确识别, 反之, 没有被成功识别。轮廓像素使用费尔曼链码表示为 $A = \{P_i | i=1, 2, \dots, N\}$ 。

在识别出目标的轮廓之后, A_i 表示第 i 个目标所识别的轮廓像素集合, $p_j^i(x_j^i, y_j^i)$ 为 A_i 中的某一个点坐标, n_i 为 A_i 中轮廓像素点个数。设定 $f_i(x_i, y_i)$ 表示第 i 个目标的中心特征点, 而中心特征点是所需提取出的目标特征点, 也就是边缘轮廓像素的平均值。中心特征点根据式(15)计算获得:

$$f_i = \sum_{j=1}^{n_i} p_j^i / n_i \quad p_j^i \in A_i \quad (15)$$

2 基于强化学习的增益调节方法

Sarsa学习算法^[11]相比于Q学习算法^[12]具有更快的收敛特性, 因此在一些学习与控制问题的应用中Sarsa算法具有优于Q学习算法的性能。本文使用Sarsa算法来进行旋翼无人机视觉伺服增益值的调节。

2.1 状态划分

状态空间的划分是建立强化学习模型的关键部分, 本文使用的状态划分的方法如下所示:

1) 在 640×360 的成像平面中将图像沿 X 轴划分为 n_x 段, 沿 Y 轴划分为 n_y 段, 其中 n_x, n_y 满足 $640\%n_x = 0$ 与 $360\%n_y = 0$, 而状态空间的大小为

$$W = \sum_{i=1}^{n_x/2} \left(\frac{n_y}{2} - i + 1 \right)。$$

2) 设定当前特征点的像素坐标为 (r_c, c_c) , 目标

特征点的像素坐标为 (r_a, c_a) , 像素坐标距离 \hat{S} 为 $\hat{S} = \text{Sgn}^2(r_c - r_a) + \text{Sgn}^2(c_c - c_a)$, 其中 $\text{Sgn}()$ 为向上取整函数。本文设定如果 $\text{Sgn}(a) = 0$, 则 $\text{Sgn}(a) = \text{Sgn}(a) + 1$ 。如图3所示, 像素距离在圆A的圆周上表示状态1, 在圆B的圆周上处于状态7。如果依次选取 $R = \{r_c | c=1, 2, \dots, N\}$, 并且依次计算对应的 \hat{S} , 则 \hat{S} 构成了一个状态空间, 大小为 N 。

3) 状态空间记为 $S = \{s_i | i=0, 1, 2, \dots, K\}$, 在某种情形下, 确定当前状态 s 的策略为: 在计算出像素坐标距离之后, 如果 $s_i < \hat{S} \leq s_{i+1}$, 则 $s = s_{i+1}$, 否则 $s = s_i$ 。

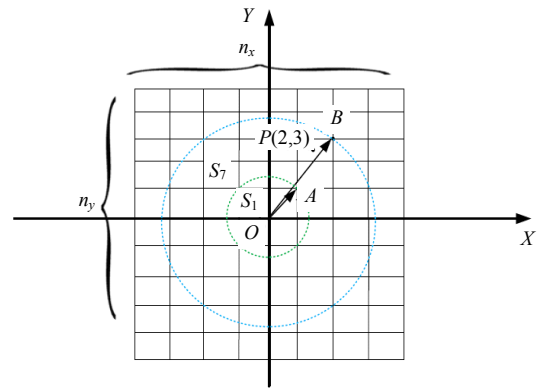


图3 状态划分

2.2 动作划分

本文使用伺服增益的一阶差分表示动作, 初始值 λ^* 作为伺服增益的初始值。设定动作集的大小为 $2n^a + 1$, 使用等差数列表示动作集, 设定公差为 d^a , 动作集可以表示为 $A = \{a_i | i=1, 2, 3, \dots, 2n^a\}$, 展开后表示为: $\{-n^a d^a, -(n^a - 1)d^a, \dots, -d^a, 0, d^a, 2d^a, \dots, (n^a - 1)d^a, n^a d^a\}$ 。伺服增益调整公式为:

$$\lambda_{i+1}^a - \lambda_i^a = \Delta \lambda_i^a = a_i \quad (16)$$

2.3 奖励函数

奖励函数分为3个情形: 到达期望目标、追踪目标丢失以及其他情况。

1) 如果每一维特征误差 $|f_i^* - f_i| < \delta$, $i=1, 2, \dots, N$, δ 为阈值, 那么就认为旋翼无人机已经到达了目标位置, 给予最高奖励。

2) 如果旋翼无人机所拍摄的实时图像提取出的特征点相比于目标图像的特征点有缺失的话, 那么就认为旋翼无人机已经开始丢失目标, 则奖励值为负值。

3) 其他情况依据旋翼无人机距离目标的远近来给予奖励。

奖励的计算公式为:

$$r = \begin{cases} 100 & \text{到达目标位置} \\ -100 & \text{丢失目标} \\ -100 \left(\sum_{i=1}^N |f_i^* - f_i| / N \sqrt{\text{row}^2 + \text{col}^2} \right) & \text{其他情况} \end{cases} \quad (17)$$

式中, row、col 分别代表图像平面的长度与宽度。

2.4 Sarsa学习算法

使用Sarsa学习的伺服增益迭代算法的执行过程如下:

算法1 Sarsa学习的增益调节算法

Definition

rand := Random number

α := Learning rate

ε := Threshold

a^{t+1} := Next action

a^t := Current action

r := Reward

s^{t+1} := Next state

s^t := Current state

γ := Discount factor

t := Time steps

Initialization

each value of Q matrix \leftarrow Zero;

$\alpha \leftarrow 0.5, \gamma \leftarrow 0.6,$

Repeat (for each step)

Choose an initial state s_0 ;

$t \leftarrow 0$;

Repeat (for each step of episode)

Generating random number: rand;

If rand $< \varepsilon$:

Choose action a^t at random.

Else:

Choose action a^t as the policy:

$\pi(s^t) = \arg \max Q(s^t, a^t)$

Take action a^t in state s^t ;

Observe s^{t+1}, r ;

Take action a^{t+1} form s^{t+1} as the ε -greedy

$Q(s^t, a^t) = Q(s^t, a^t) + \alpha[r + \gamma Q(s^{t+1}, a^{t+1}) - Q(s^t, a^t)]$

$s^t \leftarrow s^{t+1}$;

$a^t \leftarrow a^{t+1}$;

$t++$;

Until s is terminal.

Until Q matrix is convergence.

3 实验及结果分析

实验在机器人操作系统下(robot operating system, ROS)下进行, 仿真实验在Gazebo7.5下进行。旋翼无人机通过底部视觉采集图像, 采集的图像通过无线传输到计算机进行处理。除了仿真实验之外, 本文同时设计了实物场景下的实验。通过设计PID追踪算法、基于图像的视觉伺服追踪算法和基于Sarsa算法的视觉伺服(不同学习率)这3种不同的追踪算法的对比实验来验证本文提出的追踪算法的效果。实验对比角度为: 1) 追踪不同速度的相同目标时的丢失目标概率; 2) 同一目标在不同运动速度下的追踪曲线。实验参数设置如表1所示。

表1 实验参数设置

参数	值	描述
λ	0.5	伺服增益初始值
n	7	动作空间大小
α	0.5	学习率
γ	0.6	折扣因子
T_{\max}	1 800	最大时间步
T_{\min}	0	最小时间步
$\omega/(^\circ)$	120	旋翼无人机初始偏航角
n_x	10	成像平面沿X轴划分段数
n_y	10	成像平面沿Y轴划分段数

在仿真实验中, 在仿真平台下放置轮式机器人, 编写Python脚本使用键盘控制轮式机器人的运动, 设定回合目标运动的路径相同, 每回合目标采取不同的运动速度, 在每回合下测试不同的追踪算法的追踪效果。轮式机器人的运动学参数使用线速度与转角速度两个属性来描述, 转角速度以顺时针方向为正方向, 编写Python脚本控制线速度与角速度的大小。

3.1 不同线速度目标动态追踪对比实验

为了保证在视觉伺服的过程中, 旋翼无人机可以看到目标的全部特征, 设定旋翼无人机起飞飞行至离目标1 m的正上方之后才开始进行视觉伺服控制实验。基于强化学习的增益调整规则为: 在每个回合中, 旋翼无人机在保证起飞至目标上方1 m处可以看到目标的全部特征, 同时摆放位置随机。5 000回合为一次训练, 设定规则: 1) 如果旋翼无人机在追踪过程中目标丢失, 则结束本回合进入下一回合; 2) 如果旋翼无人机保持在全程追踪过程目标始终出现在底部视觉传感器视野以内, 表示追踪成功, 结束本回合, 进入下一回合; 3) 每个回合结束后,

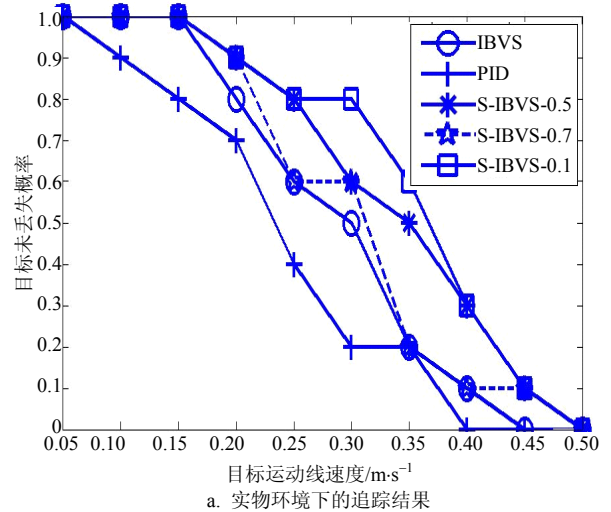
根据式(16)更新伺服增益。在实际场景的动态追踪测试中,将目标放在轮式机器人上,控制轮式机器人的运动,测试分别在PID控制、视觉伺服(IBVS)和使用Sarsa算法调节增益(S-IBVS)这3种追踪算法的情形下旋翼无人机的追踪效果,仿真条件与实物条件下的实验场景如图5所示。通过反复实验得出轮式机器人角速度在小于1.1 rad/s时,旋翼无人机采用3种不同的方法均可以较好的追踪目标。因此可以控制在轮式机器人的角速度1.1 rad/s的条件下,测试在不同线速度的追踪情况,测试结果如图6所示,其中图6a展示了仿真实验结果,图6b展示了实物实验结果。



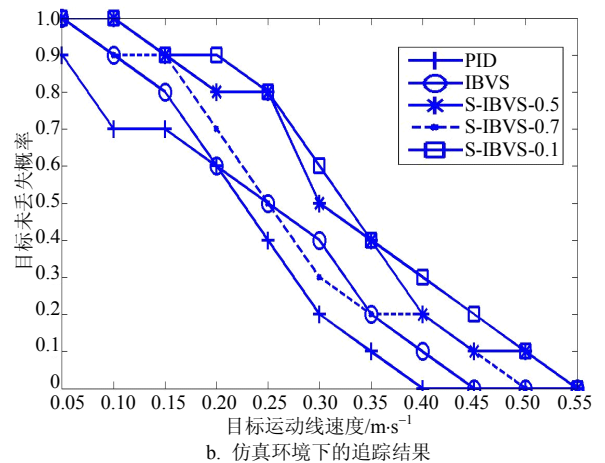
图5 仿真与实物场景下的追踪实验

学习率是影响强化学习性能的一个关键因素。为了对比强化学习的效果,本文以强化学习的学习率为变量,设置Sarsa学习的学习率分别为0.1、0.5、0.7进行对比实验。实验分别设置目标不同的线速度运动,线速度的范围是0.05~0.5 m/s。旋翼无人机起飞追踪,目标的每一次线速度追踪测试反复进行10次,统计目标未丢失概率(目标未丢失次数/总实验次

数)。由图6~图7可以看出,仿真场景提供了一种较为理想的实验环境,但是在实际场景下的实验往往更为复杂,在实物场景中旋翼无人机实验效果也相对仿真实验效果有异,但是总体的变化情况相同,说明实验的合理性。通过分析实验结果可知,PID控制与视觉伺服控制在目标速度逐渐增大的条件下未丢失目标概率快速减少,对应曲线下降斜率增大,这是这两种方法的稳定性不高导致的。而本文提出的方法(学习率为0.5)在前两种方法均完全丢失目标之后仍然有一定的概率能够追踪到目标。对比不同学习率的强化学习方法,学习率为0.1时旋翼无人机丢失目标的概率最低,学习率为0.7时旋翼无人机丢失目标的概率最高,随着学习率的增加,智能体接近目标的速度越快,但是对经验的利用就会越少。



a. 实物环境下的追踪结果



b. 仿真环境下的追踪结果

图6 目标不同线速度追踪实验对比图

3.2 不同角速度目标动态追踪对比实验

控制旋翼无人机与轮式机器人同方向前进,控制轮式机器人转弯时的线速度为0.05 m/s,调节轮式机器人以不同的角速度转向时旋翼无人机在仿真场

景图5a与实物场景图5b下的对比实验如图7所示。与线速度实验方法类似, 为了避免实验的偶然性, 每一次不同的轮式机器人的角速度实验反复进行10次, 轮式机器人的角速度从1.1 rad/s~2.0 rad/s逐渐增加。从图中可以看出, 随着轮式机器人转向角速度的不断增加, 导致旋翼无人机在追踪目标过程中丢失目标的概率(丢失目标次数/总实验次数)逐渐增大, PID控制方法在1.6 rad/s时已经完全丢失目标, 视觉伺服控制方法在2.0 rad/s之前会完全丢失目标。本文提出的方法(学习率为0.5)在2.0 rad/s之后完全丢失目标, 由此可以说明本文所提出的方法(学习率为0.5)有着较好的收敛性。

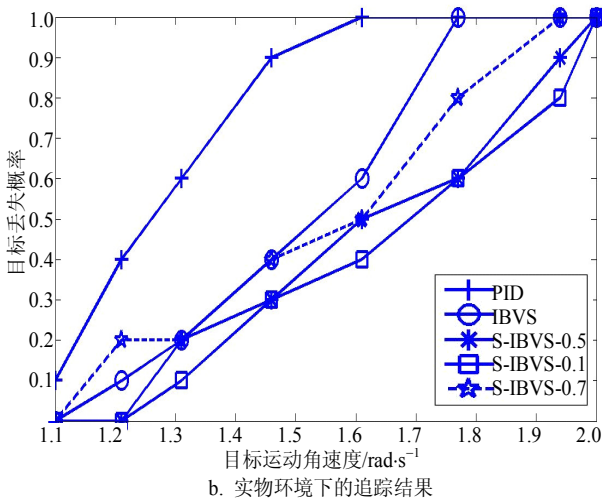
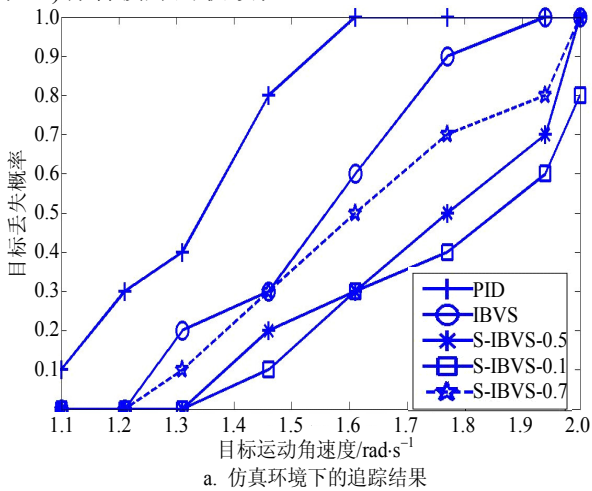


图7 目标不同角速度追踪实验对比图

4 结束语

针对在旋翼无人机追踪场景下的PID控制与基于图像视觉伺服控制的不足, 本文结合强化学习与视觉伺服控制提出了一种基于Sarsa学习的旋翼无人机视觉伺服智能追踪方法。本文建立了基于强化学习的旋翼无人机视觉伺服增益调节模型, 通过对旋

翼无人机在不同的场景下进行训练, 使得旋翼无人机通过训练能够选择相对合理的伺服增益值。本文设计了PID控制、视觉伺服控制和使用Sarsa学习算法调节伺服增益的旋翼无人机目标追踪实验, 实验结果说明了本文所提出方法具有更好的追踪效果。

参 考 文 献

- [1] CHOI S, KIM S, KIM H J. Inverse reinforcement learning control for trajectory tracking of a multirotor UAV[J]. International Journal of Control Automation & Systems, 2017, 15(4): 1826-1834.
- [2] LIU Y, MONTENBRUCK J M, ZELAZO D, et al. A distributed control approach to formation balancing and maneuvering of multiple multirotor UAVs[J]. IEEE Transactions on Robotics, 2018, 34(4): 870-882.
- [3] SI W, SHE H, WANG Z. Fuzzy PID controller for UAV tracking moving target[C]//Control and Decision Conference. Chongqing, China: IEEE, 2017: 3023-3027.
- [4] TEULIERE C, MARCHAND E, ECK L. 3-D Model-based tracking for UAV indoor localization[J]. IEEE Transactions on Cybernetics, 2015, 45(5): 869-879.
- [5] ZHENG D, WANG H, CHEN W, et al. Planning and tracking in image space for image-based visual servoing of a quadrotor[J]. IEEE Transactions on Industrial Electronics, 2018, 65(4): 3376-3385.
- [6] WANG Y, ZHANG G L, LANG H, et al. A modified image-based visual servo controller with hybrid camera configuration for robust robotic grasping[J]. Robotics & Autonomous Systems, 2014, 62(10): 1398-1407.
- [7] CHAUMETTE F, HUTCHINSON S. Visual servo control, part I: Basic approaches[J]. IEEE Robotics & Automation Magazine, 2009, 13(4): 82-90.
- [8] CHO S, LEE D, SHIM D H. Image-based visual servoing framework for a multirotor UAV using sampling-based path planning[C]//AIAA Guidance, Navigation, & Control Conference. Kissimmee, Florida: AIAA, 2015: 1-15.
- [9] SERRA P, CUNHA R, HAMEL T, et al. Landing of a quadrotor on a moving target using dynamic image-based visual servo control[J]. IEEE Transactions on Robotics, 2016, 32(6): 1524-1535.
- [10] LEE D, KIM S J. Modified chain-code-based object recognition[J]. Electronics Letters, 2015, 51(24): 1996-1997.
- [11] ZHENG N, MA Q, JIN M, et al. Abdominal-waving control of tethered bumblebees based on Sarsa with transformed reward[J]. IEEE Transactions on Cybernetics, 2018, 49(8): 3064-3073.
- [12] LI J, CHAI T, LEWIS F, et al. Off-policy Q-learning: Set-point design for optimizing dual-rate rougher flotation operational processes[J]. IEEE Transactions on Industrial Electronics, 2017, 65(5): 4092-4102.