文章编号: 1001-0920(2019)12-2517-10 DOI: 10.13195/j.kzyjc.2018.0342

一种基于Dyna-Q学习的旋翼无人机视觉伺服智能控制方法

史豪斌†, 徐 梦, 刘珈妤, 李继超

(西北工业大学 计算机学院，西安 710072)

摘 要: 基于图像的视觉伺服机器人控制方法通过机器人的视觉获取图像信息,然后形成基于图像信息的闭环反馈来控制机器人的合理运动. 经典视觉伺服的伺服增益的选取在大多数条件下是人工赋值的,故存在鲁棒性差、收敛速度慢等问题. 针对该问题,提出一种基于Dyna-Q的旋翼无人机视觉伺服智能控制方法调节伺服增益以提高其自适应性. 首先,使用基于费尔曼链码的图像特征提取算法提取目标特征点;然后,使用基于图像的视觉伺服形成特征误差的闭环控制;其次,针对旋翼无人机强耦合欠驱动的动力学特性提出一种解耦的视觉伺服控制模型;最后, 建立使用 Dyna-Q 学习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伺服增益. Dyna-Q学习在经典的Q学习的基础上通过建立环境模型来存储经验,环境模型产生的虚拟样本可以作为学习样本来进行值函数的迭代. 实验结果表明,所提出的方法相比于传统控制方法PID控制以及经典的基于图像视觉伺服方法具有收敛速度快、稳定性高的优势.

关键词: 视觉伺服；Dyna-Q学习；增益调节；旋翼无人机；费尔曼连码；强化学习

中图分类号: TP273 **文献标志码:** A

A visual servo intelligent control method for rotor UAV based on Dyna-Q learning

SHI Hao-bin†, XU Meng, LIU Jia-yu, LI Ji-chao

(School of Computer Science，Northwestern Polytechnical University，Xi'an 710072，China)

Abstract: The image-based visual servo control method of robots obtains the image information through the robot's vision and then forms the closed-loop feedback based on the image information to control the robot's reasonable movement. However, due to the problem of poor robustness and slow convergence, the selection of servo gain for classical visual servoing is artificial assignment under most conditions. Therefore, an intelligent servo control method based on Dyna-Q learning is proposed to adjust the servo gain to improve its adaptability. Firstly, this method uses the image feature extraction algorithm based on Felman chain code to extract the target feature point, then uses the image-based visual servoing to form the closed-loop control of the characteristic error. Then, this paper presents a decoupling visual servoing control model for the dynamic characteristics of rotor UAV's strong coupling underactuated. Finally, a reinforcement learning model using Dyna-Q learning to adjust the servo gain is established, through which the rotor UAV can choose the servo gain independently. The Dyna-Q learning method learns to store experience on the basis of classical Q-Learning by setting up an environment model, and the virtual samples generated by the environment model can be used as learning samples to iterate the value function. The experimental results show that the proposed method is faster and more stable than the classical PID control and classical image based visual servo methods.

Keywords: visual servo；Dyna-Q learning；gain adjustment；rotor UAV；Felman chain code；reinforcement learning

0 引 言

旋翼无人机(MRUAV),又称为旋翼无人飞行器,是一种具有3个及以上旋翼轴的无人飞行器,搭载小型高精度摄像头. 旋翼无人机作为一种集欠驱动、半自主、强耦合、非线性等动力学特性于一身的飞行器,被广泛应用于军事、民用等领域. 近年来,由于旋翼无人机相关技术发展迅速,对于旋翼无人机智能控制方法的研究已成为众多学者的研究热点[1-4]. 文献[5]

收稿日期: 2018-03-22；**修回日期:** 2018-08-16.

基金项目: 航空科学基金项目(2016ZC53022)；国家重点研发计划项目(SQ2017YFGX060091)；西北工业大学研究生种子基金项目(ZZ2018169).

责任编委: 程龙.

†通讯作者. E-mail: shihaobin@nwpu.edu.cn.

通过六通道PID控制器实现了旋翼无人机在风扰情况下的悬停控制;文献[6]提出了一种基于图像的视觉PID控制方法控制无人机的运动,但是PID控制算法精度不高,在处理非线性不确定系统时抗干扰能力弱;文献[7]采用基于位置的视觉伺服(PBVS)实现了旋翼无人机的控制,基于位置的视觉伺服的误差定义在三维笛卡尔空间,因此对初始条件、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[8]提出了一种基于视觉的旋翼无人机垂直起降算法,利用基于图像的视觉伺服(IBVS)[9]在二维图像空间中跟踪平台,将生成的速度指令作为自适应滑模控制器的输入;文献[10]提出了一种无标定的基于图像的视觉伺服无人机控制方法,并将该方法与经典的基于图像的视觉伺服控制方法进行对比,验证了所提出方法的效果. 基于图像的视觉伺服的伺服误差直接定义在二维图像平面,不需要对三维位姿进行估计,通过计算出当前目标的位置与期望中的位置误差来形成闭环控制. 但是基于图像的视觉伺服控制方法对于伺服增益的选取是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,因此基于图像的视觉伺服控制方法往往不能够在复杂的非线性环境中实现精确控制[11]. 针对上述旋翼无人机飞行控制方法的不足和经典的基于图像的视觉伺服控制方法收敛速度慢、稳定性差的问题,考虑到旋翼无人机强耦合、欠驱动的动力学特性,本文通过图像特征提取算法提取出特征点,并提出一种解耦的旋翼无人机视觉伺服控制方法,使用该解耦的基于图像的视觉伺服控制方法实现基于误差的闭环控制,并利用Dyna-Q学习[12]调节伺服增益.

1 基于视觉伺服的旋翼无人机控制模型

1.1 视觉模型

假设$P=(X_p,Y_p,Z_p)$为空间中的一点,$P_i=(x,y,z)^{\mathrm{T}}$为P点在图像平面的投影,图1展示了旋翼无人机视觉模型.

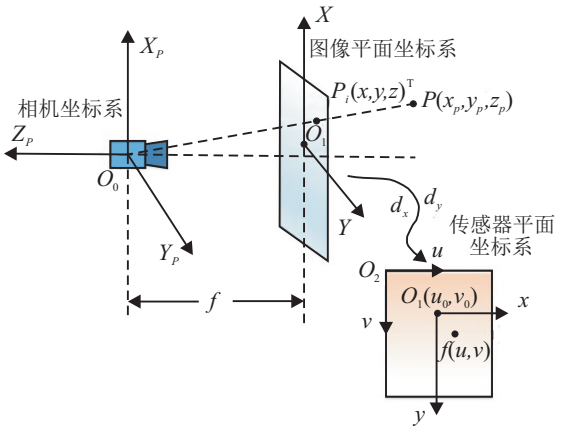


图1 旋翼无人机视觉模型

根据小孔成像的原理可以得到

$$\begin{bmatrix}x\\y\\z\\1\end{bmatrix}=\frac{1}{K_s}\begin{bmatrix}f&0&0&0\\0&f&0&0\\0&0&1&0\\0&0&0&K_s\end{bmatrix}\begin{bmatrix}X_p\\Y_p\\Z_p\\1\end{bmatrix}. \tag{1}$$

其中:K_s为常数,记为比例因子;f为焦距,由于相机坐标系的中心到图像平面坐标系中心的距离为焦距,$z=f$.

视觉传感器采集的图像使用二元函数,存储在计算机中,记为$f(u,v)$.其中:(u,v)为图像平面的坐标,$f(u,v)$为在该点处的像素值.图像平面坐标系中一点(u,v)与视觉传感器平面的坐标系中一点(x,y)的关系为

$$\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}\frac{1}{d_x}&0&u_q\\0&\frac{1}{d_y}&v_q\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}. \tag{2}$$

其中:记视觉传感器平面坐标的原点O_1在图像平面的坐标为(u_q,v_q);d_x,d_y为从图像平面到视觉传感器平面的放缩比例.

1.2 旋翼无人机动力学模型

为了便于描述旋翼无人机飞行时的姿态信息和位置信息,需要分别定义机体坐标系和惯性坐标系,并以此来分析旋翼无人机的动力学模型[13]. 图2展示了旋翼无人机的动力学模型.

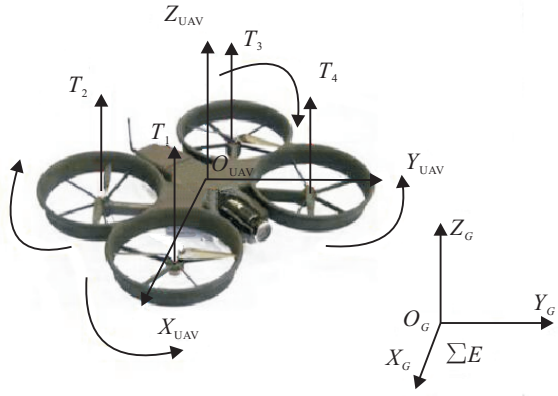


图2 旋翼无人机动力学模型

使用X_{UAV}-Y_{UAV}-Z_{UAV}表示机体坐标系,原点O_{UAV}位于机体质心,X_{UAV}轴与Y_{UAV}轴相互垂直并且构成平面$O_{\mathrm{UAV}}X_{\mathrm{UAV}}Y_{\mathrm{UAV}}$,$Z_{\mathrm{UAV}}$轴始终垂直于$O_{\mathrm{UAV}}X_{\mathrm{UAV}}Y_{\mathrm{UAV}}$平面向上.

使用X_G-Y_G-Z_G表示惯性坐标系. 原点O_G位于地平面上任意选定的某固定点,X_G轴与X_{UAV}轴平行,Y_G轴与Y_{UAV}轴平行,Z_G轴沿铅垂线方向向上.

假设旋翼无人机的质心坐标在惯性坐标系中的

位置坐标是$\zeta = x, y, z$,根据牛顿运动定律,旋翼无人机的动力学模型为

$$\begin{cases} \dot{\xi} = \boldsymbol{R} \cdot \boldsymbol{V}_u, \\ m \cdot \boldsymbol{V}_u = -m\dot{\boldsymbol{\omega}} \times \boldsymbol{V}_u + \boldsymbol{F}, \\ \dot{\boldsymbol{R}} = \boldsymbol{R} \cdot \mathrm{sk}(\boldsymbol{\omega}), \\ \boldsymbol{I} \cdot \boldsymbol{\omega} = -\boldsymbol{\omega} \times \boldsymbol{I}\boldsymbol{\omega} + \boldsymbol{\Gamma}. \end{cases} \tag{3}$$

其中: $\boldsymbol{V}_u$为线速度; $\boldsymbol{R}$为正交旋转矩阵,使用$\boldsymbol{R}$可以从机体坐标系转移到惯性坐标系; $\boldsymbol{I}$为旋翼无人机分别绕机体坐标系中$X_{\mathrm{UAV}}, Y_{\mathrm{UAV}}, Z_{\mathrm{UAV}}$三轴的转动惯量; $\mathrm{sk}(\boldsymbol{\omega})$为斜对称矩阵, $\mathrm{sk}(\boldsymbol{\omega}) = \boldsymbol{\omega} \times \boldsymbol{V}_u$; $\boldsymbol{\Gamma}$为合外力矩; $\boldsymbol{F}$为合外力, $\boldsymbol{F}$可以按下式计算:

$$\boldsymbol{F} = mg\boldsymbol{R}^{\mathrm{T}}\boldsymbol{e}_z - T\boldsymbol{e}_z, \tag{4}$$

$\boldsymbol{e}_z$为惯性坐标系中沿Z_G轴方向的单位向量, $T = F_1 + F_2 + F_3 + F_4$为参数, g为重力加速度.

通过改变旋翼无人机推进器转子的旋转速度,使飞行器产生升力,引起运动. 因此,通过改变4个推进器的转动速度,可以控制飞行器上下运动. 如果相反地控制第2和第4推进器的旋转速度,则会引起翻滚运动;如果相反地控制第1和第3推进器的旋转速度,则会引起俯仰运动;通过共同控制第1～第4推进器的旋转速度,可使飞行器产生偏航运动. 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,其映射公式如下

$$[\omega_1, \omega_2, \omega_3, \omega_4]^{\mathrm{T}} = \boldsymbol{J}_{\boldsymbol{a}}[a_x^v, a_y^v, a_z^v, a_x^\omega, a_y^\omega, a_z^\omega]^{\mathrm{T}}. \tag{5}$$

其中: $\omega_1, \omega_2, \omega_3, \omega_4$为对应的推进器转子的转动速度; $a_x^v, a_y^v, a_z^v, a_x^\omega, a_y^\omega, a_z^\omega$为六自由度的加速度; $\boldsymbol{J}_{\boldsymbol{a}}$是推进器转动速度到加速度的映射矩阵.

1.3　动力学扩展

假定旋翼无人机传入的控制参数只有4个维度,且旋翼无人机保持水平. 但实际中为了实现旋翼无人机在x或y方向运动,不能通过视觉伺服计算来获得滚转角θ_x或者俯仰角θ_y,而是通过旋翼无人机内部控制对v_x和v_y进行运动学解析来计算得出对应的θ_x和θ_y,属于底层控制. 下面介绍该方法的具体过程.

通过动力学原理可以计算出x、y方向的线加速度a_x^v, a_y^v,具体计算公式为

$$\begin{cases} a_x^v = \mathrm{d}v_x/\mathrm{d}t, \\ a_y^v = \mathrm{d}v_y/\mathrm{d}t. \end{cases} \tag{6}$$

以x方向为例,当给定一个预定的加速度a_x^v时,旋翼无人机的运动状态可以简化为图3.

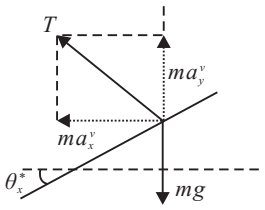


图 3　x方向运动受力分析

在图3中, θ_x^*为旋翼无人机在x方向的期望滚转角,则可得

$$\tan\theta_x^* = \frac{ma_x^v}{mg},\ \theta_x = \arctan\frac{a_x^v}{g}. \tag{7}$$

通过旋翼无人机自身携带的陀螺仪可以获取当前x方向上的实际滚转角θ_x^c. x方向的角速度ω_x可以计算出来,计算公式为$\omega_x = (\theta_x{}^* - \theta_x{}^c)/\Delta t$,其中$\Delta t$为设定的运动时间. 同理可以计算出$y$方向的角速度$\omega_y$,计算公式为

$$\begin{cases} \omega_y = \arctan\dfrac{a_y^v}{g}, \\ \omega_y = (\theta_y^* - \theta_y^c)/\Delta t. \end{cases} \tag{8}$$

其中: θ_y^*为旋翼无人机在y方向上的期望俯仰角, θ_y^c为陀螺仪获取的y方向的当前角度.

通过这样的计算就可以获得x和y方向的角速度,进而将速度扩展为$v = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)^{\mathrm{T}}$.

1.4　解耦的基于图像的视觉伺服

$\boldsymbol{e}$代表特征误差向量,可以表示为$\boldsymbol{e} = \boldsymbol{f}_c - \boldsymbol{f}_*$, $\boldsymbol{f}_c$表示特征点在图像平面中当前的坐标向量(u_c, v_c), $\boldsymbol{f}_*$为期望的特征点的坐标向量(u_*, v_*). 根据动力学关系可以得到特征误差随时间的变化率与角速度、线速度之间的函数关系,即

$$\dot{\boldsymbol{e}} = \frac{\mathrm{d}(\boldsymbol{f}_c - \boldsymbol{f}_*)}{\mathrm{d}t} = \boldsymbol{J}\begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{\omega} \end{bmatrix}. \tag{9}$$

其中: 如果只有一个特征点,则图像雅可比矩阵为$\boldsymbol{J} \subseteq R^{2\times 6}$,如果特征点的个数为$N$,则图像雅可比矩阵为$\boldsymbol{J} \subseteq R^{2N\times 6}$; $\boldsymbol{v} = (v_x, v_y, v_z)^{\mathrm{T}}$为旋翼无人机的线速度向量; $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)^{\mathrm{T}}$为旋翼无人机的俯仰角速度、滚转角速度、偏航角速度的向量. 为了保证特征误差呈指数解耦下降,设定$\dot{\boldsymbol{e}} = -\lambda_{v,\omega}\boldsymbol{e}$, $\lambda_{v,\omega}$为伺服增益,同时为了使各个维度变化都相同以保证系统的稳定性,设定$\lambda_{v,\omega} > 0$,代入式(9)可得

$$\begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{\omega} \end{bmatrix} = -\lambda_{v,\omega}\widehat{\boldsymbol{J}}\boldsymbol{e}, \tag{10}$$

其中$\widehat{\boldsymbol{J}}$为矩阵$\boldsymbol{J}$的伪逆. 如果$\boldsymbol{J}$为方阵,则矩阵的伪逆就为矩阵的逆,记为$\boldsymbol{J}^{-1}$;如果矩阵为行列不同的矩阵,则$\widehat{\boldsymbol{J}} = (\boldsymbol{J}^{\mathrm{T}} \cdot \boldsymbol{J})^{-1} \cdot \boldsymbol{J}^{\mathrm{T}}$[10].

考虑到旋翼无人机的动力学特性,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[14]. 在一个低速的状态下,一个微小的姿态倾斜角度对视觉的影响可以忽略,不会因视觉计算产生俯仰角和滚转角,故可以忽略滚转角和俯仰角的影响,只考虑线速度和偏航角[15]. 在去除俯仰角和滚转角之后,当只有一个特征点时,图像雅可比矩阵可以写为$\boldsymbol{J} \subseteq R^{2\times4}$.

根据运动学公式可得(X_p, Y_p, Z_p)随时间的变化率与线速度、角速度的关系如下:

$$\begin{bmatrix}\dot{\boldsymbol{X}}_{\boldsymbol{p}}\\ \dot{\boldsymbol{Y}}_{\boldsymbol{p}}\\ \dot{\boldsymbol{Z}}_{\boldsymbol{p}}\end{bmatrix} = \begin{bmatrix}-1 & 0 & 0 & Y_p\\ 0 & -1 & 0 & -X_p\\ 0 & 0 & -1 & 0\end{bmatrix}\begin{bmatrix}v_x\\ v_y\\ v_z\\ \omega_z\end{bmatrix}. \tag{11}$$

结合式(1)~(11)计算可以得到

$$\boldsymbol{J} = \begin{bmatrix}-\dfrac{f}{z\cdot d_x} & 0 & \dfrac{x}{z\cdot d_x} & \dfrac{y}{d_x}\\ 0 & -\dfrac{f}{z\cdot d_y} & \dfrac{y}{z\cdot d_y} & -\dfrac{y}{d_x}\end{bmatrix}. \tag{12}$$

若特征点的数目为N,则特征点坐标集为$\{f_i|i=1,2,\cdots,N\}$,故特征点误差为

$$\boldsymbol{e} = (\boldsymbol{f}_1 - \boldsymbol{f}_1^*, \boldsymbol{f}_2 - \boldsymbol{f}_2^*, \cdots, \boldsymbol{f}_N - \boldsymbol{f}_N^*)^{\mathrm{T}} \subseteq R^{2N\times1}. \tag{13}$$

图像雅可比矩阵为$\boldsymbol{J} = (\boldsymbol{J}_1, \boldsymbol{J}_2, \cdots, \boldsymbol{J}_N)^{\mathrm{T}} \subseteq R^{2N\times4}$. 对线速度和角速度分别设置独立的伺服增益$\dot{\boldsymbol{e}} = \boldsymbol{J}\cdot(v,\omega)^{\mathrm{T}}$转换,即

$$\begin{cases}\boldsymbol{v} = -\lambda_v\cdot\widehat{\boldsymbol{J}}_v\cdot\boldsymbol{e},\\ \boldsymbol{\omega} = -\lambda_\omega\cdot\widehat{\boldsymbol{J}}_\omega\cdot\boldsymbol{e}.\end{cases} \tag{14}$$

其中:λ_v和λ_ω分别为线速度和角速度的伺服增益,$\widehat{\boldsymbol{J}}_v \subseteq R^{3\times2N}$为$\widehat{\boldsymbol{J}} \subseteq R^{4\times2N}$的前3行组成的子矩阵,$\widehat{\boldsymbol{J}}_\omega \subseteq R^{1\times2N}$为$\widehat{\boldsymbol{J}} \subseteq R^{4\times2N}$的第4行组成的子矩阵.

分别对线速度和角速度设置伺服增益可以实现线速度增益和角速度增益的单独调节. 文中1.4节给出了旋翼无人机动力学解耦思路,即只调节线速度和偏航角. 在实际应用中,滚转角和俯仰角特别小但是不为0,利用运动学原理,通过$\boldsymbol{v} = (v_x, v_y, v_z, \omega_z)^{\mathrm{T}}$来计算出预计的滚转角和俯仰角,然后使用运动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$\boldsymbol{v} = (v_x, v_y, v_z, \omega_z)^{\mathrm{T}}$扩展为$\boldsymbol{v} = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)^{\mathrm{T}}$再进行计算,通过这种方法可以实现对线速度和角速度的动力学解耦,动力学扩展的方法在1.3节进行了阐述.

本文到此建立旋翼无人机的解耦视觉伺服模型,具体包括两部分:旋翼无人机动力学解耦以及设置单独的线速度增益和角速度增益. 已经有学者使用了李雅普诺夫稳定性理论说明了解耦视觉伺服的稳定性[10,14].

2 基于Dyna-Q学习的视觉伺服控制方法

2.1 基于Dyna-Q学习的视觉伺服控制方法框架

旋翼无人机通过底部视觉传感器采集图像信息,通过基于费尔曼链码的图像轮廓特征提取算法提取出目标的轮廓特征,之后根据轮廓提取出目标的中心特征点,文献[1]对此进行了详细的解释. 图4展示了基于Dyna-Q学习的视觉伺服控制方法的基本框架,在提取出目标的特征点之后使用解耦的基于图像的视觉伺服实现旋翼无人机的闭环反馈控制,使用Dyna-Q学习调节伺服增益.

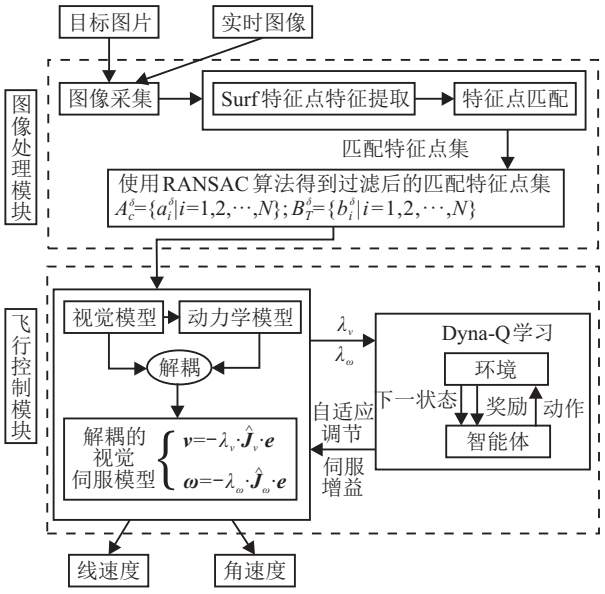


图4 使用Dyna-Q学习的视觉伺服控制方法

2.2 基于费尔曼链码的图像特征提取算法

使用基于轮廓的图像特征提取算法,提取出的目标的轮廓使用费尔曼链码[16]来描述,使用费尔曼链码的轮廓像素记为$A = \{P_i|i = 1,2,\cdots,N\}$,对目标进行轮廓提取得到的链码表示为$C = \{C_i|i = 1,2,\cdots,N\}$,其中费尔曼链码的方向数为$K$. 在使用这种方法之前,首先需要建立标准轮廓库,将获取到的目标的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为$D = \{D_i|i = 1,2,\cdots,N\}$,首先使用一阶差分对费尔曼链码进行旋转归一化,即

$$d_i = \begin{cases}C_i - C_N,\ i = 1;\\ C_i - C_{i-1} + 1,\ i > 1,\ C_i - C_{i-1} + 1 > 0;\\ K + C_i - C_{i-1} + 1,\ i > 1,\ C_i - C_{i-1} + 1 \leqslant 0.\end{cases} \tag{15}$$

在进行旋转归一化后计算提取出的轮廓与标准轮廓之间的lenvenshtein距离[17],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. 在识别出目标轮廓之后,设定第i个目标的所识别的轮廓像素集合为$\boldsymbol{A}_i$,$\boldsymbol{P}_j^i(x_j^i, y_j^i)$为$\boldsymbol{A}_i$中的某一个点坐标,$n_i$为$\boldsymbol{A}_i$中轮廓像素点个数,设定$\boldsymbol{f}_i(x_i, y_i)$为第$i$个目标的中心特征点,则中心特征点为边缘轮廓像素的平均值,中心特征点就是本文所需提取的目标特征点,中心特征点计算公式为

$$\boldsymbol{f}_i = \frac{\sum_{j=1}^{n^i} \boldsymbol{P}_j^i}{n_i},\ \boldsymbol{P}_j^i \subseteq \boldsymbol{A}_i. \tag{16}$$

2.3　Q 学 习

强化学习[18]是一种试错型的学习算法,通过智能体不断与环境进行交互而获得经验来进行决策. 强化学习最重要的两个特征是试错和延迟奖励,强化学习的过程是:智能体对环境执行某种动作,改变环境的状态并获得环境给予的奖励从而强化或是减弱智能体选择某动作的趋势,反复执行这一过程,智能体最终能够获得完成相应目标的最优动作.

Q学习[18]是目前较为有效的无模型的强化学习方法之一,其学习规则如下:

$$Q(s_t, a_t) = (1-\alpha)Q(s_t, a_t) + \alpha(r + \gamma \max_{a \subseteq A} Q(s_{t+1}, a)). \tag{17}$$

其中:$\alpha \subseteq [0,1]$为学习率,$\gamma \subseteq (0,1)$为折扣因子,A为动作的集合,r为奖励函数. Q学习的稳定性从确定性马尔科夫过程和非确定性马尔科夫过程两个方面来证明[19].

基于Dyna-Q学习的伺服增益迭代算法不同于传统的Q学习算法,Dyna-Q学习在经典的Q学习的基础上引入规划来辅助学习. 智能体与环境进行交互获得真实样本,直接强化学习过程使用真实样本更新值函数,而在Dyna学习框架中,该样本用于学习更新值函数和更新虚拟环境模型,规划是指虚拟环境模型产生的虚拟样本被用于更新值函数的过程. Dyna-Q学习有机结合了学习过程和规划过程,更好地利用了已有的学习经验,可以获得在动态环境下更快的学习效果[20-22].

2.4　使用Dyna-Q学习的伺服增益的调节

2.4.1　状态空间的划分

状态空间的划分是建立强化学习模型的关键部分,通过本文提出的状态划分方法可以得到离散化的状态空间. 下面介绍本文使用的状态划分算法:

Step 1: 对于640×360像素的成像平面,将图像沿X轴划分为n_x段,沿Y划分为n_y段,其中n_x, n_y满足$640\%n_x = 0$ and $360\%n_y = 0$,状态空间的大小为$K = \sum_{i=1}^{n_x/2} \left(\frac{n_y}{2} - i + 1\right)$.

Step 2: 假设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即图5中的坐标原点,记为$O(r_a, c_a)$,设定$\hat{\boldsymbol{S}}$为像素坐标距离,计算公式为$\hat{\boldsymbol{S}} = \mathrm{Sgn}(r_c - r_a)^2 + \mathrm{Sgn}(c_c - c_a)^2$,其中Sgn()函数为向上取整函数. 本文设定如果$\mathrm{Sgn}(a) = 0$,则$\mathrm{Sgn}(a) = \mathrm{Sgn}(a) + 1$. 如图5所示,像素坐标距离在实线圆周上表示状态1,在点线圆周上处于状态4. 如果依次选取$R = \{(r_c, c_c)|c = 1, 2, \cdots, K\}$,计算对应的$\hat{\boldsymbol{S}}$,则$\hat{\boldsymbol{S}}$可以构成一个状态空间,大小为$K$.

Step 3: 状态空间记为$\boldsymbol{S} = \{s_i|i = 1, 2, \cdots, K\}$,在某种情形下,在计算出像素坐标距离之后按照以下策略确定当前状态s,如果$s_i < \hat{\boldsymbol{S}} \leqslant s_{i+1}$,则$s = s_{i+1}$,否则$s = s_i$,由此可以计算出在某一情形下的状态.

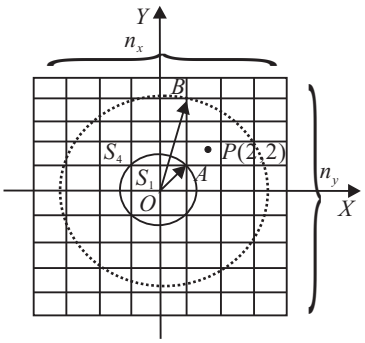


图 5　状态划分

2.4.2　动作空间的划分

由分析可知,如果选择伺服增益的值作为动作,则由于伺服增益的选择范围很广,可能导致动作空间过大,进而产生维数灾难. 因此针对两种不同伺服增益,分别初始化两个伺服增益,然后设定一个有限大小的离散动作集合,通过在动作集合中选择动作进行增益的调整,由于离散动作空间和离散状态空间大小是固定值,可以避免维数灾难.

为了实现伺服增益的调整,选择伺服增益的差值作为动作,选择初始值λ^*作为伺服增益的初始值. 设动作集的大小为$2n^a+1$,则动作集构成了一个等差数列,设定公差为d^a,则动作集$A = \{a_i|i = 1, 2, \cdots, 2n^a\}$,动作集展开为$\{-n^a d^a, -(n^a-1)d^a, \cdots, -d^a, 0, d^a, 2d^a, (n^a-1)d^a, n^a d^a\}$,伺服增益调整公式为

$$\lambda_{t+1}^a - \lambda_t^a = \Delta\lambda_t^a = a_t. \tag{18}$$

由式(18)可知,线速度和角速度方向上的伺服增益的

调整公式为

$$\begin{cases}\lambda_{t+1}^{a_v}=\lambda_t^{a_v}+a_t^v,\\ \lambda_{t+1}^{a_\omega}=\lambda_t^{a_\omega}+a_t^\omega.\end{cases}\tag{19}$$

其中: $\lambda_t^{a_v}$ 为调整之前的线速度伺服增益, $\lambda_{t+1}^{a_v}$ 为选择动作 a_t^v 后调整得到的线速度伺服增益, $\lambda_t^{a_\omega}$ 为调整之前的偏航角的伺服增益,在选择动作 a_t^ω 之后伺服增益调整为 $\lambda_{t+1}^{a_\omega}$.

2.4.3 奖励函数

奖励函数分为3个部分:到达期望目标、追踪目标丢失和其他情况. 如果每一维特征误差 $|f_i^*-f_i|<\delta, i=1,2,\cdots,N$, δ 为阈值,则认为旋翼无人机已经到达了目标位置,可以给予最高奖励.

若旋翼无人机拍摄的实时图像通过特征提取后得到的特征点相比于目标图像的特征点有缺失,则认为无人机已经开始丢失目标,此时回报值为负值. 其他情况依据旋翼无人机距离目标的远近来给予奖励. 因此,使用函数解析式描述奖励函数如下所示:

$$r=\begin{cases}100,\ \text{feature points reach target position};\\ -100,\ \text{missing target feature points};\\ -100\Big(\displaystyle\sum_{i=1}^{N}|f_i^*-f_i|/N\sqrt{\text{row}^2+\text{col}^2}\Big),\ \text{others}.\end{cases}\tag{20}$$

其中row, col分别为图像平面的长度和宽度.

2.4.4 基于Dyna-Q学习的伺服增益迭代算法

Dyna-Q学习在获取真实经验过程中建立的环境模型记为 $\text{model}(s,a)$, $\text{model}(s,a)$ 使用4元组 (s_t, a_t, s_{t+1}, r), Q 表示为 $Q(s,a)$. 根据Dyna-Q学习原理可知,在一次迭代过程中选择动作 a 之后,得到环境的一次奖励并改变了环境的状态. 在线速度与偏航角两种状态空间划分伺服增益,本文设定这两种状态空间相同,都是按照2.4.1节介绍的状态空间划分方法进行划分,故 $\boldsymbol{S}_v=\boldsymbol{S}_\omega=\boldsymbol{S}$.

由式(20)分析得出线速度和角速度的Q学习中 Q 值的迭代公式为

$$\begin{cases}Q_v(s_t^v,a_t^v)=(1-\alpha)Q_v(s_t^v,a_t^v)+\\ \qquad\alpha(r+\gamma\max Q_v(s_{t+1}^v,a_t^v)),\\ Q_\omega(s_t^\omega,a_t^\omega)=(1-\alpha)Q_\omega(s_t^\omega,a_t^\omega)+\alpha(r+\\ \qquad\gamma\max Q_\omega(s_{t+1}^\omega,a_t^\omega)).\end{cases}\tag{21}$$

其中: $Q_v(s_t^v,a_t^v)$, $Q_\omega(s_t^\omega,a_t^\omega)$ 分别为状态 s_t^v 和状态 s_t^ω 下采取动作 a_t^v 和动作 a_t^ω 所对应的 Q 值, r 为奖励值, γ 为折扣因子. 使用Dyna-Q学习的迭代算法的具体步骤如表1所示.

$$\begin{cases}\pi(s_t^v)=\arg\max Q_v(s_v^t,a_t^v),\\ \pi(s_t^\omega)=\arg\max Q_\omega(s_\omega^t,a_t^\omega).\end{cases}\tag{22}$$

表1 基于Dyna-Q学习的伺服增益迭代算法

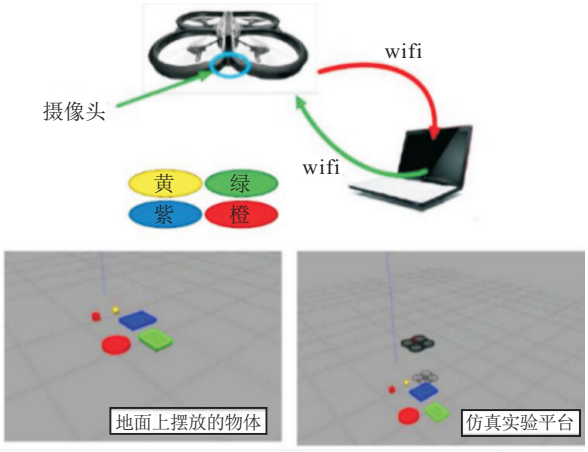
步骤	具体操作
输入	初始化 $Q_v(s_v^t,a_t^v)$, $Q_\omega(s_\omega^t,a_t^\omega)$, $\text{model}(s_v^t,a_t^v)$, $\text{model}(s_\omega^t,a_t^\omega)$,状态 $s_v^t\in\boldsymbol{S}_v$, $s_\omega^t\in\boldsymbol{S}_\omega$, 动作 $a_t^v\in\boldsymbol{A}_v$, $a_t^\omega\in\boldsymbol{A}_\omega$, Q 表中元素的数值全部初始化为0,4元组的模型元素初始化为0.
过程	1)当前的状态为 s_v^t,产生两个随机数分别记为 rand_1 和 rand_2,如果 $\text{rand}_1<\varepsilon$,则随机选择一个动作 a_t^v,否则按照式(22)选择动作;同理如果 $\text{rand}_2<\varepsilon$,则随机选择动作 a_t^ω,否则按式(22)选择动作. 2)采取动作 a_t^v 和 a_t^ω 后,分别得到下一状态 s_v^{t+1} 和 s_ω^{t+1},奖励 r. 3)根据式(21)更新 Q_v 和 Q_ω. 4)分别用4元组 $(s_v^t,a_t^v,s_v^{t+1},r)$ 更新 $\text{model}(s_v^t,a_t^v)$,用4元组 $(s_\omega^t,a_t^\omega,s_\omega^{t+1},r)$ 更新 $\text{model}(s_\omega^t,a_t^\omega)$. 5)重复以下步骤 P 次. a)在 $\text{model}(s_v^t,a_t^v)$ 中随机选择状态-动作对 (s_v^t,a_t^v),在 $\text{model}(s_\omega^t,a_t^\omega)$ 中随机选择状态-动作对 (s_ω^t,a_t^ω). b)将 (s_v^t,a_t^v) 用于 $\text{model}(s_v^t,a_t^v)$,得到 s_v^{t+1} 和 r,将 (s_ω^t,a_t^ω) 用于 $\text{model}(s_\omega^t,a_t^\omega)$,得到 s_ω^{t+1} 和 r. c)根据式(21)更新 Q_v 和 Q_ω. 6)回到步骤2),重复执行 k 次.
输出	$Q_v(s_v^t,a_t^v)$, $Q_\omega(s_\omega^t,a_t^\omega)$ 和4元组 $\text{model}(s_v^t,a_t^v)$, $\text{model}(s_\omega^t,a_t^\omega)$.

3 实验与分析

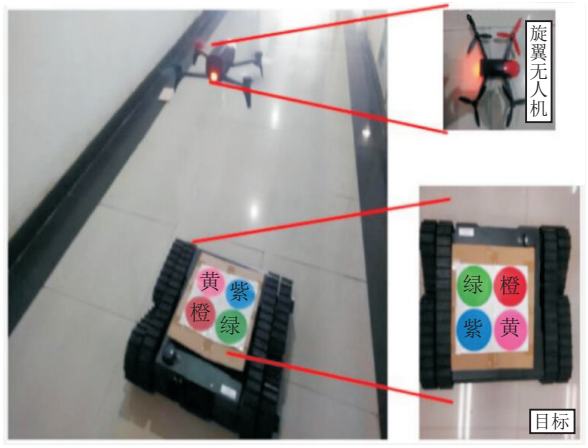
3.1 实验概述

本次实验在机器人操作系统[23](robot operating system, ROS)下进行, Gazebo7.5仿真平台提供了一个较为理想的物理仿真环境,使得模型在建立过程中可以避免因考虑太多因素而导致模型过于复杂[24],旋翼无人机使用底部视觉采集图像,即使用旋翼无人机的下视摄像头进行图像采集. 旋翼无人机通过无线传输将图像数据传输到计算机上进行处理,图6展示了仿真平台下和真实环境中的旋翼无人机的飞行实

验环境.



(a) 仿真实验环境



(b) 实物实验环境

图 6　旋翼无人机仿真和实物实验环境

表2展示了实验参数配置,由于旋翼无人机使用底部视觉进行环境感知,基于图像的视觉伺服需要获取全部的目标特征,调节旋翼无人机起飞并离地面一定距离后再开始进行视觉伺服控制.

表2　实验参数配置

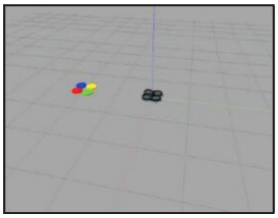
参数	值	描述
λ	0.4	IBVS 伺服增益
λ_v	0.4	DQ-IBVS 线速度方向初始伺服增益
λ_ω	0.45	DQ-IBVS 角速度方向初始伺服增益
n_a^v	7	DQ-IBVS 线速度动作空间大小
n_a^ω	7	DQ-IBVS 角速度动作空间大小
α	0.5	学习率
γ	0.6	折扣因子
$T_{\max}$	1 800	最大时间步
$T_{\min}$	0	最小时间步
ω	120	旋翼无人机初始偏航角
n_x	32	成像平面 X 轴分段数
n_y	18	成像平面 Y 轴分段数

本次实验分别在仿真平台和真实环境下对PID控制方法(PID)[6]、基于图像的视觉伺服[10]的控制方法(IBVS)、使用Dyna-Q学习调节增益的视觉伺服的控制方法(DQ-IBVS)进行实验. 仿真实验和实物场景下的实验结果均表明,本文提出的使用Dyna-Q学习调节增益的视觉伺服具有快速收敛性.

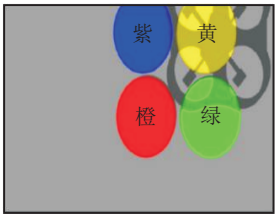
3.2　特征点检测对比

旋翼无人机通过底部视觉获取目标图像,使用基于费尔曼链码的图像特征提取算法进行图像特征提取,使用3种不同的方法分别进行测试对比,设定时间片为$\Delta t = 0.04$. 设置旋翼无人机不同的摆放位置和朝向,在不同的情形下进行测试,对比实验结果.

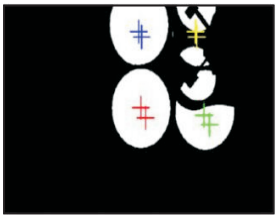
如图7所示的情景中,4种不同颜色的圆的半径为0.05 m,旋翼无人机与惯性坐标系的ΣEx轴的夹角为120°.



(a) 初始旋翼无人机摆放的位置



(b) 最终图像位置



(c) 目标特征点检测

图 7　旋翼无人机特征检测测试情形分布

为了验证3种不同方法的实验效果,分别在仿真平台和真实场景下进行测试,旋翼无人机在真实场景中的摆放与仿真平台下一致,将目标放在轮式机器人上,控制轮式机器人运动,旋翼无人机会对目标进行动态追踪,同时,仿真环境与实物环境下的目标形状、大小均相同. 特征点的识别结果如图8所示,图8展示了3种方法分别在仿真平台和真实场景中的特征点轨迹. 单次实验的偶然性较强,为此本文重复实验50次. 图8展示了3种不同的方法得到的特征轨迹图,特征点使用的是2.2节介绍的图像特征提取算法提取出的目标中心特征点,在仿真场景中使用该图像特征

提取算法分别提取出4个不同颜色的圆形目标的中心特征点,绘制4个特征点的运动轨迹,在实物场景下进行同样的实验.

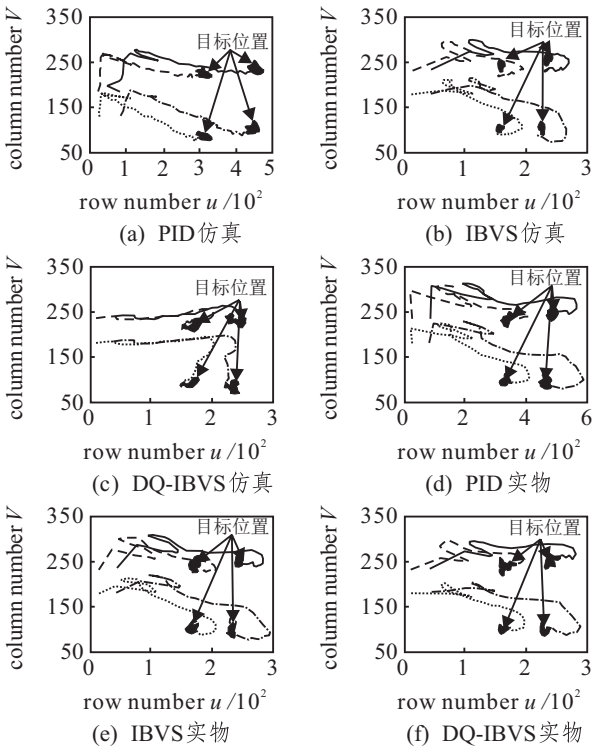


图8 3种不同方法下仿真和实物测试的特征点变化轨迹

在图8中,实线代表绿色,虚线代表紫色,点线代表黄色,点划线代表橙色. 从左到右依次为PID控制、IBVS控制、DQ-IBVS控制方法得到的特征轨迹图,上方的图示为仿真平台下的测试,下方为实际场景下的实验. 由实验结果可以看出,相对于仿真平台,实际场景中的环境更为复杂,在实际运行中,旋翼无人机波动较仿真平台多,但最终都可以收敛到目标位置处上下波动. 实际场景下的实验较仿真平台下实验有一些波动,但是波动范围在像素之间,属于合理的范围,其中,使用IBVS和PID控制会出现在目标位置处拐弯的现象,这是由于PID控制和IBVS控制方法固定增益所导致的各个方向不同步(某个维度已经达到指定的位置,其他维度还没有到达). 使用DQ-IBVS方法在仿真平台下的测试特征轨迹曲线平滑,但是在实际场景中效果不理想,旋翼无人机易受实际环境的影响,但是使用该目标中心特征点提取方法可以较为有效地将目标简化为单一特征点,为后面的实验做准备.

3.3 三维飞行轨迹曲线

设定旋翼无人机起始位置和目标位置如表3所示. 图9展示了使用PID、IBVS、DQ-IBVS三种不同方法获得的仿真三维飞行轨迹图.

表3 旋翼无人机初始与目标位置参数设置

参数	值	描述
X_{init}	5.5000	初始位置横坐标
Y_{init}	0.6000	初始位置纵坐标
Z_{init}	0.0000	初始位置竖坐标
ω_{init}	120	初始位置偏航角
X_{fin}	5.707	目标位置横坐标
Y_{fin}	−0.210	目标位置纵坐标
Z_{fin}	0.825	目标位置竖坐标
ω_{fin}	0	目标位置偏航角

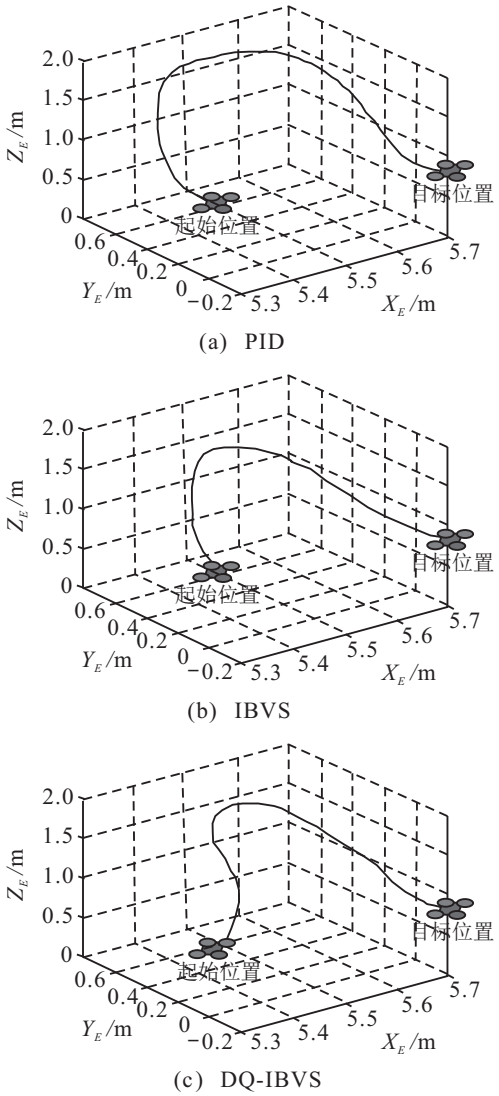


图9 三维飞行轨迹

由图9可以看出,使用PID控制方法、基于图像的视觉伺服控制方法IBVS、基于Dyna-Q学习的视觉伺服控制方法DQ-IBVS三种不同的方法控制旋翼无人机运动,最终都可以使得旋翼无人机从相同的起点运动到相同的终点,但是飞行轨迹却不相同. 同时,为了保证旋翼无人机能够看到目标的全部特征,设定旋翼无人机起飞后距离地面1.30 m后再进行飞行控制. 可以发现,旋翼无人机初始朝向与惯性坐标系轴

具有一定的夹角,会使得旋翼无人机先上升一定的距离后才会下降.

3.4　误差收敛曲线

仿真平台下轨迹误差收敛曲线和实际场景下轨迹特征误差曲线如图10和图11所示.

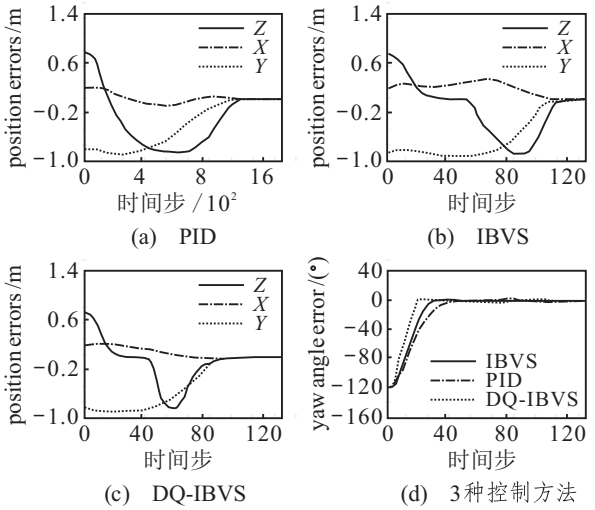


图 10　仿真平台下轨迹误差收敛曲线

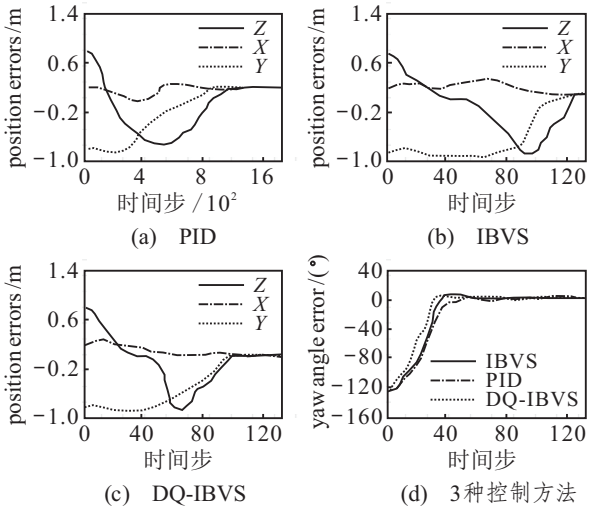


图 11　实际场景下轨迹特征误差曲线

由图10和图11可以看出,旋翼无人机在实际场景中波动较仿真平台多,但是总体上可以实现收敛并稳定. 本文在旋翼无人机动力学建模阶段设定了旋翼无人机的滚转角和俯仰角为0°,通过旋翼无人机的位姿调节并按照1.3节的方法进行动力学扩展. 动力学扩展之后,在实际实验中测得滚转角和俯仰角的波动范围在 $-5^\circ \sim 5^\circ$ 之间,十分微小,说明了本文对模型进行简化,即不通过视觉伺服获取旋翼无人机的俯仰角和滚转角的合理性.

通过图10可以看出,在仿真实验中,经典的PID控制方法在1 200～1 300时间步之间实现位置误差收敛,基于图像的视觉伺服控制方法在110～120个时间步之间位置误差会收敛至0附近,并且最终几乎没有误差波动,DQ-IBVS方法在80～100之间实现了位置误差收敛,DQ-IBVS方法的收敛效果要优于IBVS方法和PID控制方法. 在实际场景中的测试由于受到环境影响,旋翼无人机在飞行过程中稳定性较差,因此会出现误差波动,实验效果与仿真平台下的实验具有一定的差距,但是总体效果与仿真实验一致. 如图11所示,在实际场景中,使用PID控制方法位置误差收敛时间在1 200～1 600时间步之间,并且在收敛之后误差波动较大,IBVS方法在120～140时间步之间实现位置误差收敛,DQ-IBVS方法在100时间步的时候基本实现了位置误差收敛. 观察图10和图11可以发现,固定增益的控制方法使得旋翼无人机的收敛效果不高,甚至会在某些维度上发生波动,基于此,本文认为增益应该是一个变量,在不同的状态下,增益的值不同.考虑到强化学习是一种十分有效的学习算法,是一种智能体可以通过自学习获得经验而进行最优决策的学习算法[25],本文使用一种强化学习的方法进行增益的自调节,即Dyna-Q学习.

设定Dyna-Q学习在一个学习回合中所花费的最大时间单位为400个时间步,在旋翼无人机摆放位置位于可行范围内的条件下(即每个回合中旋翼无人机在保证起飞1.30 m之后能够看到目标的全部特征的条件下)随机摆放,5 000个回合为一次训练. 限定:如果400个时间步之后旋翼无人机仍然没有到达目标位置,则强制回到起点进行下一回合;如果旋翼无人机在飞行过程中失去了目标的特征点,则本回合结束,进行下一回合;如果旋翼无人机在飞行过程中,保持一定时间相对于目标位置的距离在5像素以内,认为到达了目标位置,则结束本回合,进行下一回合;在每回合结束之后,根据动作选择策略式(19)进行增益的迭代. 通过图10和图11可以发现,使用Dyna-Q学习调节增益的视觉伺服控制方法较其余两种方法的收敛时间更少,在仿真和实际条件下均在100时间步以内可以实现收敛,并且误差波动最小,稳定性较高. 这也验证了本文提出的Dyna-Q学习调节增益的方法较另外两种方法具有更稳定的运行效果和更快的收敛速度.

4　结　论

针对旋翼无人机的智能控制问题,本文通过总结分析PID控制方法和经典的基于图像的视觉伺服控制方法的不足,提出了使用Dyna-Q学习算法自适应调节视觉伺服增益的智能控制方法,并结合一种旋翼无人机动力学解耦的方法,为旋翼无人机的智能控制提出了一种稳定性更强、收敛更快的控制方法. 由实验结果可知,该方法相对于PID控制、IBVS控制算法的收敛速度更快、稳定性更高,使得旋翼无人机可以

通过自学习调节伺服增益,有效提升旋翼无人机的智能控制效果.

参考文献(References)

[1] Shi H, Li X, Hwang K S, et al. Decoupled visual servoing with fuzzy Q-learning[J]. IEEE Transactions on Industrial Informatics, 2016, 14(1): 241-252.

[2] Zhou Y, Dong X, Lu G, et al. Time-varying formation control for unmanned aerial vehicles with switching interaction topologies[C]. International Conference on Unmanned Aircraft Systems. Orlando: IEEE, 2014: 26-36.

[3] Hafez A T, Marasco A J, Givigi S N, et al. Solving multi-UAV dynamic encirclement via model predictive control[J]. IEEE Transactions on Control Systems Technology, 2015, 23(6): 2251-2265.

[4] Jabbari H, Oriolo G, Bolandi H. An adaptive scheme for image-based visual servoing of an underactuated UAV[J]. International Journal of Robotics & Automation, 2014, 29(29): 92-104.

[5] 屈耀红, 邢哲文, 袁冬莉, 等. 基于悬停旋翼无人机位置姿态信息的风场估计方法研究[J]. 西北工业大学学报, 2016, 34(4): 684-690.
(Qu Y H, Xing Z W, Yuan D L, et al. Wind field estimation based on position and attitude information of quadrotor in hover[J]. Journal of Northwestern Polytechnical University, 2016, 34(4): 684-690.)

[6] Watanabe K, Yoshihata Y, Iwatani Y, et al. Image-based visual PID control of a micro helicopter using a stationary camera[C]. Sice Annual Conference 2007. Takamatsu: IEEE, 2007: 3001-3006.

[7] Ryuta Ozawa, François Chaumette. Dynamic visual servoing with image moments for an unmanned aerial vehicle using a virtual spring approach[J]. Advanced Robotics, 2013, 27(9): 683-696.

[8] Lee D, Ryan T, Kim H J. Autonomous landing of a VTOL UAV on a moving platform using image-based visual servoing[C]. IEEE International Conference on Robotics and Automation. Saint Paul: IEEE, 2012: 971-976.

[9] Janabi-Sharifi F, Deng L, Wilson W J. Comparison of basic visual servoing methods[J]. IEEE/ASME Transactions on Mechatronics, 2011, 16(5): 967-983.

[10] Àngel Santamaria-Navarro, Andrade-Cetto J. Uncalibrated image-based visual servoing[C]. IEEE International Conference on Robotics and Automation. Wuhan: IEEE, 2013: 5247-5252.

[11] Siradjuddin I, Behera L, Mcginnity T M, et al. Image-based visual servoing of a 7-DOF robot manipulator using an adaptive distributed fuzzy PD controller[J]. IEEE/ASME Transactions on Mechatronics, 2014, 19(2): 512-523.

[12] Hwang K S, Jiang W C, Chen Y J, et al. Model learning for multistep backward prediction in Dyna-Q learning[J]. IEEE Transactions on Systems, Man, & Cybernetics Systems, 2017, 48(9): 1470-1481.

[13] 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.

[14] Chaumette F, Hutchinson S. Visual servo control, part I: Basic approaches[J]. IEEE Robotics & Automation Magazine, 2007, 13(4): 82-90.

[15] Gomez-Balderas J E, Flores G, Lozano R. Tracking a ground moving target with a quadrotor using switching control[J]. Journal of Intelligent & Robotic Systems, 2013, 70(1/2/3/4): 65-78.

[16] Lee D, Kim S J. Modified chain-code-based object recognition[J]. Electronics Letters, 2015, 51(24): 1996-1997.

[17] Faes J, Gillis J, Gillis S. Phonemic accuracy development in children with cochlear implants up to five years of age by using Levenshtein distance[J]. Journal of Communication Disorders, 2016, 59: 40-48.

[18] Sugimoto T, Gouko M. Acquisition of hovering by actual UAV using reinforcement learning[C]. International Conference on Information Science and Control Engineering. Beijing: IEEE, 2016: 148-152.

[19] Watkins C J C H, Dayan P. Technical note: Q-learning[J]. Machine Learning, 1992, 8(3/4): 279-292.

[20] Hwang K S, Jiang W C, Chen Y J. Model learning and knowledge sharing for a multiagent system with Dyna-Q learning[J]. IEEE Transactions Cybernetics, 2015, 45(5): 964-976.

[21] Sutton R S, Barto A G. Reinforcement learning: An introduction[J]. IEEE Transactions on Neural Networks, 1998, 9(5): 1054.

[22] Sutton R S. Integrated architectures for learning, planning, and reacting based on approximating dynamic programming[J]. Machine Learning Proceedings, 1990, 2(4): 216-224.

[23] Scholl P M, Majoub B E, Santini S, et al. Connecting wireless sensor networks to the robot operating system[J]. Procedia Computer Science, 2013, 19: 1121-1128.

[24] Yao W, Dai W, Xiao J, et al. A simulation system based on ROS and gazebo for robocup middle size league[C]. IEEE International Conference on Robotics and Biomimetics. Zhuhai: IEEE, 2016: 54-59.

[25] Hwang K S, Li C W, Jiang W C. Adaptive exploration strategies for reinforcement learning[C]. International Conference on System Science and Engineering. Ho Chi Minh City: IEEE, 2017: 16-19.

作者简介

史豪斌(1978−), 男, 副教授, 博士, 从事智能决策与控制、机器学习及其应用等研究, E-mail: shihaobin@nwpu.edu.cn;

徐梦(1993−), 男, 硕士生, 从事智能决策与控制、机器学习及其应用的研究, E-mail: menghsu@mail.nwpu.edu.cn;

刘珈妤(1993−), 女, 硕士生, 从事智能决策与控制、机器学习及其应用的研究, E-mail: 942227597@qq.com;

李继超(1994−), 男, 硕士生, 从事智能决策与控制、机器学习及其应用的研究, E-mail: 948998893@qq.com.

(责任编辑：闫 妍)