

基于图像矩与神经网络的机器人四自由度视觉伺服

李优新^{1,2}, 毛宗源¹, 田联房¹

(1. 华南理工大学 自动化科学与工程学院, 广东 广州 510640; 2. 广东工业大学 信息工程学院, 广东 广州 510006)

摘要: 针对传统的视觉伺服方法中图像几何特征的标记、提取与匹配过程复杂且通用性差等问题, 本文提出了一种基于图像矩的机器人四自由度(4DOF)视觉伺服方法. 首先建立了眼在手系统中图像矩与机器人位姿之间的非线性增量变换关系, 为利用图像矩进行机器人视觉伺服控制提供了理论基础, 然后在未对摄像机与手眼关系进行标定的情况下, 利用反向传播(BP)神经网络的非线性映射特性设计了基于图像矩的机器人视觉伺服控制方案, 最后用训练好的神经网络进行了视觉伺服跟踪控制. 实验结果表明基于本文算法可实现0.5 mm的位置与0.5°的姿态跟踪精度, 验证了算法的有效性较好的伺服性能.

关键词: 图像矩; 神经网络; 机器人; 视觉伺服

中图分类号: TP242.6 **文献标识码:** A

Visual servoing of 4DOF using image moments and neural network

LI You-xin^{1,2}, MAO Zong-yuan¹, TIAN Lian-fang¹

(1. School of Automation Science and Engineering, South China University of Technology, Guangzhou Guangdong 510640, China;

2. School of Information Engineering, Guangdong University of Technology, Guangzhou Guangdong 510006, China)

Abstract: To avoid the complicated marking, extracting and matching of image features in the traditional visual servoing systems and to improve the universality of the algorithm, a novel visual servoing of 4-degrees of freedom(4DOF) is proposed for an eye-in-hand robot based on image moments and neural network. First, the nonlinear transform relationship between image moments and the robot pose is developed, which provides the theoretical basis for the visual servoing using image moments. Then, a back propagation(BP) neural network is designed to map the transformation from image moments variation to the robot pose displacement with 4DOF without the external and internal parameters calibration for the camera. After this, the proposed control scheme can be applied to the robotic visual servoing. The experiment results show that the tracking error is less than 0.5 mm and 0.5° respectively in position and in orientation. This confirms the validity and satisfactory servoing performance of the proposed method.

Key words: image moments; neural network; robot; visual servoing

1 引言(Introduction)

根据控制方案的不同, 机器人视觉伺服系统主要分为基于位置的视觉伺服系统(PBVS)和基于图像的视觉伺服系统(IBVS), 相比前者而言, 基于图像的视觉伺服系统控制方案由于不需要精确的机器人系统模型与视觉系统模型、对噪声和标定误差 insensitive 以及具有更好的控制精度等优点而得到更广泛应用并成为当前机器人视觉伺服系统的主要研究方向^[1,2].

由于几何特征的可靠提取严重依赖于目标物体的形状与环境, 因此目前的基于图像的视觉控制方案主要从图像中提取点、线、夹角等简单的几何特征来构建图像雅可比矩阵, 并且经常需要借助在场

景中插入的特定的视觉线索^[3]. 即使这样, 几何特征的提取与匹配也需要复杂且耗时的运算, 并且, 不同的目标物体需要采用不同的几何特征, 其选取并没有通用的方法, 甚至同一物体在相对摄像机坐标的位姿发生变化后采用的特征也要相应变化^[4], 因此基于图像几何特征的方法缺乏适用于各种物体对象的通用性, 这些成为影响该方案深入研究与应前景的主要因素. 相对于目前主要采用的图像几何特征, 文献[3,5~7]等采用了基于图像矩特征的视觉伺服方法, 由于应用了图像的全局特征信息, 因此具有更好的鲁棒性与通用性, 但文献[3]借助了点特征加以辅助, 因此是矩特征与几何特征的混合算法, 文献[5~7]推导了基于矩特征算法下的反馈控制律并

验证了其有效性,但都只实现了2自由度或3自由度的视觉伺服任务,并且涉及到复杂的广义逆矩阵的求解,反馈系数矩阵的参数选择并无规律可循、只能依赖经验选取等问题.本文将图像矩特征描述算子与神经网络结合起来,利用BP神经网络的非线性逼近与学习功能来映射图像矩特征与机器人四自由度(4DOF, X, Y, Z 与绕 Z 轴的旋转)位姿变换关系,然后进行机器人的视觉伺服控制,既无需对摄像机与手眼关系进行标定,也避免了图像矩雅可比矩阵的直接求取与逆运算以及传统的基于图像几何特征的视觉伺服系统中所要面对的图像特征的标记、提取与匹配等困难而复杂的问题.

2 图像矩与摄像机位姿增量变换关系(Transformation of image moments and robot pose)

设 $f(x, y)$ 是大小为 $M \times N$ 的数字图像 I_f 中像素点 (x, y) 处的灰度值,则图像 I_f 的 $i + j$ 阶几何矩与中心矩分别定义为

$$m_{ij} = \sum_{x=1}^M \sum_{y=1}^N x^i y^j f(x, y), \quad (1)$$

$$\mu_{ij} = \sum_{x=1}^M \sum_{y=1}^N (x - \bar{x})^i (y - \bar{y})^j f(x, y). \quad (2)$$

其中 i, j 为非负整数,且有

$$\bar{x} = m_{10}/m_{00}, \bar{y} = m_{01}/m_{00}.$$

不同阶次与类型的图像矩特征对应不同的物理解释,如 m_{00} 代表了图像的总“质量”,当摄像机产生垂直于目标平面的运动时, m_{00} 产生变化,对应图像的缩放操作, m_{10}, m_{01} 分别代表图像质心位置,当摄像机产生平行于目标平面的运动时,质心位置随之发生变换, m_{20}, m_{02} 分别是图像绕 x, y 轴的惯量矩; $\mu_{11}, \mu_{20}, \mu_{02}$ 则反映了摄像机与目标之间相对姿态关系^[3].

同时,矩描述子唯一性定理说明:矩特征完全可以反映一幅图像的所有特征,几何矩集合 $\{m_{ij}|i, j = 0, 1, 2, \dots\}$ 与图像 I_f 之间是互相唯一确定的^[8].但对于具体的任务,需要从中选择能提供足够信息的矩描述的子集,如对于视觉伺服控制而言,应选择随摄像机位姿变化而变化的矩的描述子集,矩集合 $\{m_{00}, m_{10}, m_{01}, \mu_{11}, \mu_{20}, \mu_{02}\}$ 的变化可以反映摄像机的位姿的变化^[3].

设 (X_c, Y_c, Z_c) 表示目标在摄像机坐标系 $\{C\}$ 下的坐标, (x, y) 表示其在图像平面上的坐标,摄像机镜头焦距为 f ,根据摄像机成像原理可得

$$x = fK_x \frac{X_c}{Z_c}, y = fK_y \frac{Y_c}{Z_c}. \quad (3)$$

式中 K_x, K_y 是反映 x, y 方向单位距离上像素数量的比例系数.

如果目标静止不动,摄像机在 $\{C\}$ 下的 X, Y, Z 3个方向的微小平移量分别为 $\delta T_x, \delta T_y, \delta T_z$,绕 Z 轴的微小旋转量为 $\delta\theta_z$,则经过简单位姿空间变换及应用式(3)可得目标在图像平面上坐标的相应变化量 $(\delta x, \delta y)$ 近似为

$$\delta x = -\frac{fK_x}{Z_c} \delta T_x + \frac{x\delta T_z}{Z_c - \delta T_z} - \frac{yK_x}{K_y} \sin(-\delta\theta_z), \quad (4)$$

$$\delta y = -\frac{fK_y}{Z_c} \delta T_y + \frac{y\delta T_z}{Z_c - \delta T_z} + \frac{xK_y}{K_x} \sin(-\delta\theta_z). \quad (5)$$

为进一步建立图像矩与摄像机位姿之间的增量变换关系,设图像平面点 (x, y) 处的灰度值在 t 与 $t + \delta t$ 时刻分别为 $f(x, y, t)$ 与 $f(x, y, t + \delta t)$,据矩特征的定义,有

$$\begin{aligned} m_{ij}(t + \delta t) - m_{ij}(t) &= \\ \sum_{x=1}^M \sum_{y=1}^N x^i y^j f(x, y, t + \delta t) - \sum_{x=1}^M \sum_{y=1}^N x^i y^j f(x, y, t), \end{aligned} \quad (6)$$

$$\begin{aligned} \sum_{x=1}^M \sum_{y=1}^N x^i y^j f(x, y, t + \delta t) &= \\ \sum_{x=1}^M \sum_{y=1}^N x^i y^j [f(x, y, t) + f_t(x, y, t)\delta t]. \end{aligned}$$

将上式代入式(6)得

$$\begin{aligned} \delta m_{ij}(t) &= \\ m_{ij}(t + \delta t) - m_{ij}(t) &= \\ \sum_{x=1}^M \sum_{y=1}^N x^i y^j f_t(x, y, t)\delta t. \end{aligned} \quad (7)$$

式中 $f_t(x, y, t)$ 为图像上任一点的灰度变化率,并且根据光流约束方程有

$$f_t(x, y, t) = -f_x(x, y, t)\dot{x} - f_y(x, y, t)\dot{y}. \quad (8)$$

其中: $f_x(x, y, t), f_y(x, y, t), f_t(x, y, t)$ 分别是灰度沿 x, y, t 等3个方向的梯度值.将式(8)代入式(7)可得

$$\begin{aligned} \delta m_{ij}(t) &= \\ - \sum_{x=1}^M \sum_{y=1}^N x^i y^j [f_x(x, y, t)\delta x + \\ f_y(x, y, t)\delta y]. \end{aligned}$$

上式说明了目标在图像平面位置产生微小量改变时,像素点灰度沿 x, y 的梯度值与图像矩特征的变化对应关系,与时间 t 已无直接关系,因此上式可写为

$$\delta m_{ij} = - \sum_{x=1}^M \sum_{y=1}^N x^i y^j [f_x(x, y) \delta x + f_y(x, y) \delta y]. \quad (9)$$

对于目标物理轮廓外的点都具有灰度零的二值化图形而言,有以下关系式^[7]:

$$\sum_{x=1}^M \sum_{y=1}^N x^i y^j f_x(x, y) = \begin{cases} 0, & \text{如果 } i = 0, \\ -i \times m_{(i-1)j}, & \text{如果 } i > 0; \end{cases} \quad (10)$$

$$\sum_{x=1}^M \sum_{y=1}^N x^i y^j f_y(x, y) = \begin{cases} 0, & \text{如果 } j = 0, \\ -j \times m_{i(j-1)}, & \text{如果 } j > 0. \end{cases} \quad (11)$$

将式(4)(5)(10)(11)代入式(9)可得所需各阶图像矩与摄像机位姿之间微量变化关系如下:

$$\delta m_{00} = 2m_{00} \frac{\delta T_Z}{Z_c - \delta T_Z}, \quad (12)$$

$$\delta m_{10} = -m_{00} \frac{fK_x}{Z_c} \delta T_X + 3m_{10} \frac{\delta T_Z}{Z_c - \delta T_Z} - m_{01} \frac{K_x}{K_y} \sin(-\delta \theta_Z), \quad (13)$$

$$\delta m_{01} = -m_{00} \frac{fK_y}{Z_c} \delta T_Y + 3m_{01} \frac{\delta T_Z}{Z_c - \delta T_Z} + m_{10} \frac{K_y}{K_x} \sin(-\delta \theta_Z). \quad (14)$$

根据中心矩与几何矩之间的相互关系^[7],可推导出

$$\delta u_{11} = 4u_{11} \frac{\delta T_Z}{Z_c - \delta T_Z} + \left(\frac{K_y}{K_x} u_{20} - \frac{K_x}{K_y} u_{02} \right) \sin(-\delta \theta_Z), \quad (15)$$

$$\delta u_{20} = 4u_{20} \frac{\delta T_Z}{Z_c - \delta T_Z} - 2u_{11} \frac{K_x}{K_y} \sin(-\delta \theta_Z), \quad (16)$$

$$\delta u_{02} = 4u_{02} \frac{\delta T_Z}{Z_c - \delta T_Z} + 2u_{11} \frac{K_y}{K_x} \sin(-\delta \theta_Z). \quad (17)$$

至此,由式(12)~(17)建立了摄像机位姿增量到图像矩增量之间的直接映射关系,可描述为

$$\delta m = H(\delta P). \quad (18)$$

式中:

$$\delta m = [\delta m_{00}, \delta m_{10}, \delta m_{01}, \delta u_{11}, \delta u_{20}, \delta u_{02}]^T, \\ \delta P = [\delta T_X, \delta T_Y, \delta T_Z, \delta \theta_Z]^T.$$

置换式(18)的输入输出关系,可得到由图像矩增

量到摄像机位姿增量的逆映射模型:

$$\delta P = H'(\delta m). \quad (19)$$

3 神经网络设计与训练(Design and training of neural network)

式(12)~(19)构成了基于图像矩特征进行视觉伺服控制的理论基础,但存在: δm 与 δP 间是非线性函数关系、且随深度 Z_c 而变化,式中包含难以精确标定的摄像机内外参数,要进行非常复杂的广义逆矩阵的求解等一系列问题.因此,传统的控制方法难以利用或取得好的控制性能.为此,本文利用神经网络的非线性逼近能力来映射式(19)的函数关系.

3.1 系统结构设计(Design of system structure)

本文的研究对象是机器人在 X, Y, Z 3维空间的平动与绕 Z 轴的旋转共4个自由度的视觉伺服问题,采用眼在手(eye-in-hand)视觉伺服系统,摄像机固连在机器人末端执行器(简称机器人)上并使其光轴与工具坐标的 Z_T 轴平行,因此摄像机与机器人之间有恒定的位姿变换关系,如图1所示.图中: $\{R\}$ 为基坐标系, $\{T\}$ 为工具坐标系, $\{C\}$ 为摄像机坐标系.

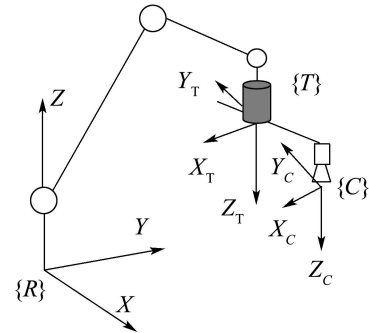


图1 坐标系定义示意图

Fig. 1 Coordinate systems

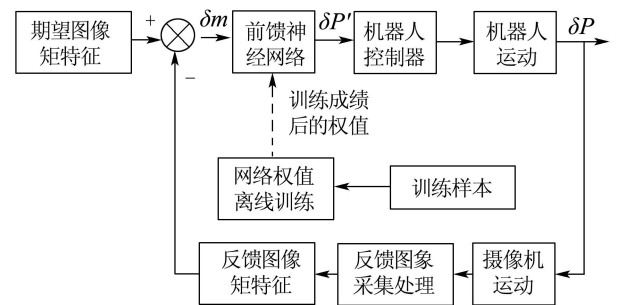


图2 视觉伺服系统结构

Fig. 2 Visual servoing system

网络的作用是将图像空间所计算的目标图像与反馈图像的矩特征误差映射到机器人所在的笛卡尔空间,获得位姿增量送给机器人伺服控制器.根

据输入输出的参数需要与系统特点, 本文采用了6输入、4输出3层前馈网络, 隐层节点数量在后面实验中确定, 学习律采用带动量加速因子的BP算法, 网络权值通过离线训练方式获取然后用于在线控制, 系统结构如图2所示.

3.2 训练样本采集与处理(Samples extracting and processing)

原始训练样本通过如下方式获得: 移动机器人使摄像机垂直于目标工作平面并在以后的实验中维持绕 X 与 Y 轴的旋转角度不变. 先定义一个参考位姿并记录下此时的图像与位姿向量, 分别作为参考图像与参考位姿, 然后在前述的4个自由度上的参考位姿的一定范围内各随机选择 n 个值, 组合成 $n^4 = N$ 个位姿向量, 移动机器人到这些位姿并记录下这些位姿及其对应的图像. 通过二值化进行目标提取等处理后计算这些图像分别与参考图像的图像矩向量差:

$$\delta m^{(i)} = [\delta m_{00}^{(i)}, \delta m_{10}^{(i)}, \delta m_{01}^{(i)}, \delta u_{11}^{(i)}, \delta u_{20}^{(i)}, \delta u_{02}^{(i)}]^T, \\ i = 1, 2, \dots, N.$$

然后按矩的阶次分类进行归一化处理后作为网络训练样本的输入.

同理, 计算出对应 N 个位姿向量与参考位姿向量的差值:

$$\delta P^{(i)} = [\delta T_X^{(i)}, \delta T_Y^{(i)}, \delta T_Z^{(i)}, \delta \theta_Z^{(i)}]^T, \\ i = 1, 2, \dots, N.$$

并按自由度分类进行归一化处理后作为网络训练样本的期望输出.

3.3 神经网络训练(Neural network training)

已知输入层节点数为6, 输出层节点数为4. 算法步骤采用批处理方式, 因此误差性能指标取:

$$E = \frac{1}{2} \sum_{i=1}^N e(i)^T e(i).$$

式中 $e(i)$ 为第 i 个样本的网络输出与期望输出的误差向量.

网络训练包括正向传播与反向传播两个过程, 隐层神经元激活函数采用Sigmoid函数, 输出层神经元激活函数采用线性函数, 网络权值的修正算法采用带动量因子的梯度下降法, 迭代运算的终止条件采用误差终止法.

4 实验结果与分析(Experiment results and analysis)

实验平台由计算机、加拿大CRS PLUS A461 6自由度工业机器人、日本Watec公司CCD摄像头、深圳微视公司V510图象采集卡等组成. 首先按照3.2所述

方法进行了训练样本的采集与处理, 其中: 根据机器人作业区间的情况选定的参考图像所在位姿为: $X = 0 \text{ mm}$, $Y = 330 \text{ mm}$, $Z = 558 \text{ mm}$, $YAW = 0^\circ$, $PITCH = 90^\circ$, $ROLL = -90^\circ$, 并在其后的试验中, 维持 YAW 与 $PITCH$ 不变, 只对其余4个自由度进行视觉伺服控制, 前述的旋转角度 θ_Z 与 $ROLL$ 的关系定义为: $\theta_Z = ROLL + 90^\circ$. 在4个自由度上位姿采集的范围为: $-50 \text{ mm} \leq X \leq 50 \text{ mm}$, $280 \text{ mm} \leq Y \leq 380 \text{ mm}$, $500 \text{ mm} \leq Z \leq 600 \text{ mm}$, $-45^\circ \leq \theta_Z \leq 45^\circ$, 每个自由度上的采样数为 $n = 5$, 共有 $N = 5^4 = 625$ 个学习样本. 然后用自己编写的MATLAB7.0程序离线进行网络训练, 迭代终止条件为误差 $E < 10^{-4}$. 通过多次对比实验, 确定当隐层节点数 $Q = 40$, 学习步长 $\eta = 0.1$, 动量因子 $\alpha = 0.06$ 时, 训练效果与速度较好. 最后, 将离线训练好的网络应用于机器人视觉伺服实时控制系统中, 先移动机器人到参考位姿处采集期望图像如图3(a)所示, 再将机器人移到另外一个位姿后启动视觉伺服自动控制程序, 系统自动采集反馈图像, 完成二值化处理(目标提取)、各阶图像矩计算、与期望图像的图像矩差的计算、通过神经网络映射计算处理对应的机器人位姿控制增量并发送给机器人控制器等一些列流程, 待收到机器人控制器反馈回来的上个控制周期结束的信息后, 重复上述流程直到达到设定条件(如视觉伺服周期数或误差)或人为终止. 本次实验中, 自动控制流程终止的条件采用了视觉伺服周期数, 经过12个视觉伺服周期的控制后反馈图像与期望图像的图像差如图3(b)所示, 实验表明达到了较好的跟踪效果.

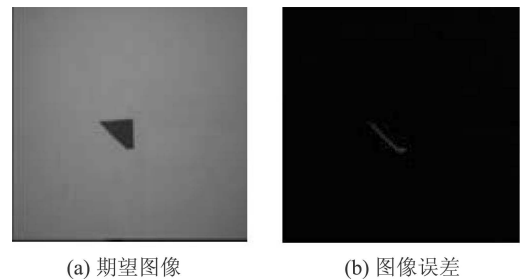
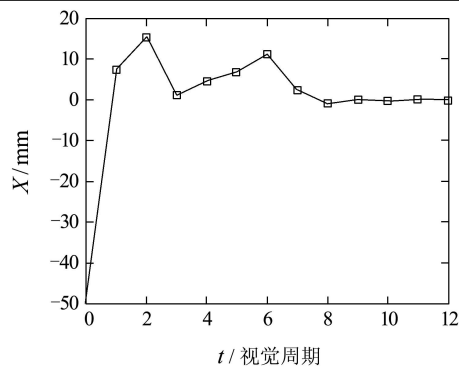


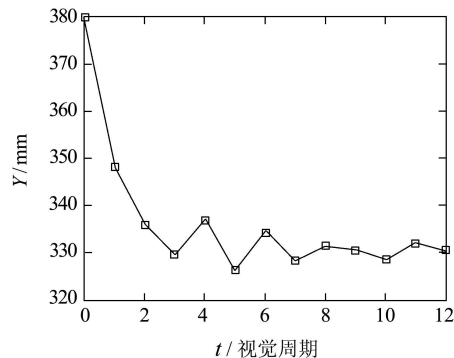
图 3 实验图像

Fig. 3 Experiment images

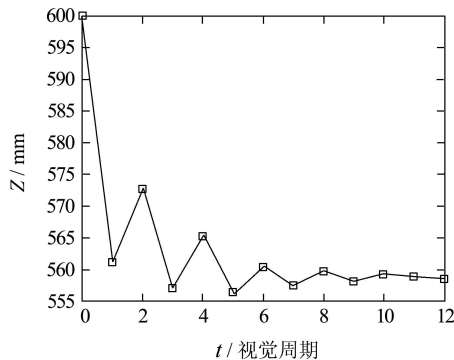
为了辅助证明本文所提视觉控制方法的效果, 系统还记录了与期望图像和各控制周期下反馈图像对应的机器人的位姿信号, 由此生成的4个自由度上的跟踪曲线如图4(a)~4(d)所示. 其中, 在 X , Y , Z 3个方向的定位误差分别在 $\pm 0.4 \text{ mm}$, $\pm 0.4 \text{ mm}$ 与 $\pm 0.5 \text{ mm}$ 之内, 绕 Z 轴的旋转角误差在 ± 0.5 度范围之内, 进一步证明了采用本文算法能获得较好的控制精度.



(a) X轴方向位置跟踪曲线



(b) Y轴方向位置跟踪曲线



(c) Z轴方向位置跟踪曲线

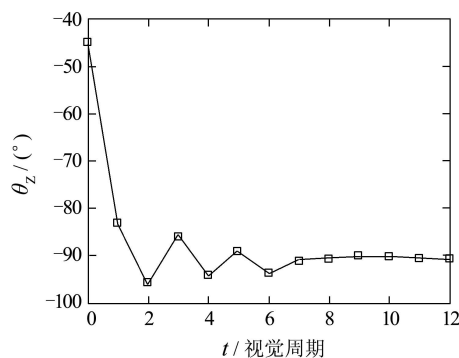
(d) 旋转角 θ_z 的跟踪曲线

图4 位姿跟踪曲线

Fig. 4 Tracking curves of pose

伺服控制的可行性的基础上,利用BP神经网络的非线性映射特性在未进行摄像机与机器人参数标定的条件下设计了基于图像矩的四自由度视觉伺服控制器并进行了实验.从期望图像与反馈图像的跟踪效果以及各自自由度的位姿参数的跟踪曲线验证了算法的有效性.同时,该方法避免了复杂而困难的对大量未知参数进行在线辨识、离线标定以及广义逆矩阵的求解,算法清晰简单、具有通用性与实用价值.实验中在初始反馈位姿与期望位姿的差值并不算小的情况下仍取得了较高的伺服控制精度与较快的收敛速度,说明了本文中基于图像矩与神经网络相结合的视觉闭环控制方案具有良好的泛化能力与鲁棒性.

参考文献(References):

- [1] HUTCHISON S, HAGER G, CORKE P I. A tutorial introduction to visual servoing control[J]. *IEEE Transactions on Robotics and Automation*, 1996, 12(5): 651 – 670.
- [2] ESPIAU B, CHAUMETTE F, RIVES P. A new approach to visual servoing in robotics[J]. *IEEE Transactions on Robotics and Automation*, 1992, 8(3): 313 – 326.
- [3] WELLS G, VENAILLE C, TORRAS C. Vision-based robot robot positioning using neural networks[J]. *Image and Servo Computing*, 1996, 14(10): 715 – 732.
- [4] 赵清杰, 连广宇, 孙增圻. 机器人视觉伺服综述[J]. *控制与决策*, 2001, 16(6): 849 – 853.
(ZHAO Qingjie, LIAN Guangyu, SUN Zengqi. Survey of robot visual servoing[J]. *Control and Decision*, 2001, 16(6): 849 – 853.)
- [5] 林靖, 陈辉堂, 王月娟. 基于图像矩的机器人视觉伺服[J]. *同济大学学报*, 2000, 28(4): 438 – 442.
(LIN Jing, CHEN Huitang, WANG Yuejuan. Visual servoing using image moments[J]. *Journal of Tongji University*, 2000, 28(4): 438 – 442.)
- [6] 杜建军, 王学影, 赵万生. 基于Adept机器人的视觉伺服控制系统[J]. *控制理论与应用*, 2007, 24(4): 565 – 568.
(DU Jianjun, WANG Xueying, ZHAO Wansheng. A visual servoing system based on Adept robot[J]. *Control Theory & Applications*, 2007, 24(4): 565 – 568.)
- [7] 林靖, 陈辉堂, 王月娟. 基于图像矩的运动目标3D平动视觉跟踪[J]. *机器人*, 2000, 22(3): 217 – 223.
(LIN Jing, CHEN Huitang, WANG Yuejuan. Image moments based visual tracking of 3D translational motion[J]. *Robot*, 2000, 22(3): 217 – 223.)
- [8] PROKOP R J, REEVES A P. A survey of moment-based object representation and recognition[J]. *Graphical Models and Image Processing*, 1992, 54(5): 438 – 460.

作者简介:

李优新 (1966—), 男, 博士研究生, 广东工业大学信息工程学院副教授, 主要研究方向为智能控制与机器视觉等, E-mail: Liyx@163.com;

毛宗源 (1936—), 男, 教授, 博士生导师, 主要研究领域为人工智能、智能控制与模糊控制等, E-mail: auzymao@scut.edu.cn;

田联房 (1968—), 男, 教授, 博士生导师, 主要研究方向为智能机器人系统, E-mail: chlftian@yahoo.com.

5 结论(Conclusions)

本文在分析了基于图像矩特征进行机器人视觉