

文章编号: 1001-0920(2010)01-0043-05

一种适用于局部特征的混合视觉伺服方法

谷 雨, 李 平, 韩 波

(浙江大学 a. 工业控制国家重点实验室, b. 工业控制研究所, 杭州 310027)

摘 要: 提出一种基于图像矩的混合视觉伺服方法, 有效解决了因匹配的局部特征改变而引起的雅可比矩阵结构改变及奇异等难点. 该方法首先估计得到单应性矩阵, 然后通过反向映射计算未匹配点的像素坐标, 采用图像矩和单应性矩阵分解得到的姿态分别控制平动和转动自由度, 采用特征的雅可比矩阵不存在奇异情况. 从理论上证明了所采用控制策略的稳定性和对摄像机标定误差的鲁棒性, 仿真实验结果表明该方法适用于局部特征.

关键词: 混合视觉伺服; 局部特征; 图像矩; 稳定性

中图分类号: TP242.6 **文献标识码:** A

Hybrid visual servoing scheme applicable for local features

GU Yu, LI Ping, HAN Bo

(a. National Key Laboratory of Industrial Technology, b. Institute of Industrial Process Control, Zhejiang University, Hangzhou 310027, China. Correspondent: LI Ping, E-mail: pli@ipc.zju.edu.cn)

Abstract: A hybrid visual servoing scheme using image moments is proposed to solve the difficulties of architecture changes and singularities of Jacobian matrix when matched local features change. Homography matrix is obtained, and the pixel coordinates of unmatched points are computed by using back projection. Image moments and poses obtained by homography decomposition are used to control translational and rotational degree of freedom, respectively, where the corresponding Jacobian matrix is not singular. The stability and the robustness to camera calibration errors are analyzed theoretically. The simulation results show that the proposed scheme is applicable for local features.

Key words: Hybrid visual servoing; Local features; Image moments; Stability

1 引 言

机器人视觉伺服控制系统^[1]通过对摄像机获取的图像数据进行分析, 采用一定的控制策略实现定位和抓取等功能, 其在机械零件自动检测、机器人导航和无人机自主着陆等方面具有广泛的应用, 是国内外研究的热点之一^[2].

图像特征选择是机器人视觉伺服系统在复杂环境下能够成功应用的关键因素之一. 目前采用的特征包括点、直线、椭圆和轮廓特征等^[3], 并采用特定标志简化前期图像处理的难度. 由于具有独特性和多量性等优点, 以及在宽基线匹配和目标识别等方面表现出的良好性能, 局部特征^[4]越来越多地应用于机器人视觉伺服控制领域. 文献^[5]首先提出把 SIFT 特征^[6]应用于机器人的视觉伺服控制, 研究了特征点匹配的鲁棒方法, 但没有给出控制策略. Lazar 等^[7]比较了应用于手眼视觉伺服系统的角点

特征和 SIFT 特征的稳定性和鲁棒性. Hoffmann 等^[8]给出了在视觉伺服系统中使用 SIFT 特征的控制策略, 但由于该方法充分利用了 SIFT 特征向量提供的角度信息, 因而不适用于其他局部特征. 文献^[9]提出一种改进的局部特征, 提高了特征提取的实时性, 采用基于图像的视觉伺服方法实现了控制目的.

在视觉伺服控制系统中使用局部特征的难点是摄像机视角变化使匹配的特征点改变, 导致图像雅可比矩阵结构改变, 可能出现奇异或控制信号突变等情况. 本文提出了一种基于图像矩的混合视觉伺服方法, 有效解决了上述难点. 该方法首先通过匹配的特征点估计单应性矩阵, 并通过反向映射计算未匹配点的像素坐标, 然后采用图像矩和单应性矩阵分解获得的姿态信息作为特征分别控制平动和转动自由度, 雅可比矩阵形式不存在奇异情况. 本文从理

收稿日期: 2009-02-11; 修回日期: 2009-05-31.

基金项目: 国家 863 计划项目(2006AA10Z204).

作者简介: 谷雨(1982—), 男, 长春人, 博士, 从事视觉跟踪、视觉伺服等研究; 李平(1954—), 男, 广西北流人, 教授, 博士生导师, 从事空中机器人控制与导航等研究.

论上证明了所采用控制策略的稳定性和对相机标定误差的鲁棒性。

2 传统的混合视觉伺服方法

机器人视觉伺服控制系统的目标是使如下误差函数趋向于零:

$$e(s) = s(t) - s^* \quad (1)$$

其中: s 表示采用的图像特征, s^* 表示在期望位置的特征值。

根据反馈信号的类型,视觉伺服控制系统分为基于位置的视觉伺服(PBVS)和基于图像的视觉伺服(IBVS)^[2]。基于位置的视觉伺服系统误差信号定义在三维世界坐标系中,需要知道目标的三维模型估计相对位姿。由于在图像平面没有控制,不能保证目标一直位于摄像机视野范围内。基于图像的视觉伺服系统误差信号定义在二维图像平面,通过比较当前位置和期望位置对应点的像素坐标误差来实现控制目的。优点是不需要重构位姿,缺点是图像雅可比矩阵存在奇异情况。Malis等^[10]提出了混合视觉伺服方法(HVS),该方法结合了PBVS和IBVS的优点,通过单应性矩阵^[11]分解获得当前位置和期望位置的相对位姿。误差信号一部分定义在图像平面,一部分定义在三维世界坐标系,因此混合视觉伺服方法也称为基于单应性矩阵的视觉伺服方法。

设 $P = [X, Y, Z]^T$ 表示三维世界坐标系中一点,根据透视成像模型得到其在图像平面的像素齐次坐标 $p = [u, v, 1]^T$ 为

$$p = \frac{1}{Z} A P, \quad (2)$$

其中 A 为摄像机内参数矩阵。如图1所示,设在当前位置和期望位置获得的图像平面 I_p 和 I_p^* 中,同一点的像素齐次坐标分别为 p 和 p^* ,则关联两点的单应性矩阵 G 为

$$p = G p^* \quad (3)$$

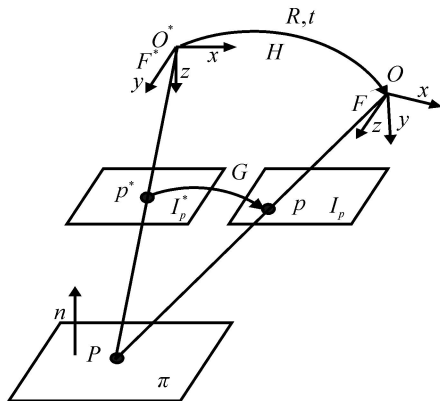


图1 两视角间的单应性矩阵

由式(2)和(3)得到三维平面两点间的单应性矩阵 H 为

$$H = \frac{Z}{Z^*} A^{-1} G A. \quad (4)$$

采用式(5)对 H 进行分解得到 R, t, d, n , 即

$$H = R + \frac{t}{d} n^T. \quad (5)$$

其中: R 为旋转矩阵, t 为平移向量, n 为目标所在平面的法线向量, d 为摄像机和该平面的垂直距离。对 R 进一步分解得到旋转轴 $\phi = [\phi_x, \phi_y, \phi_z]^T$ 和旋转角

文献[10]提出的混合视觉伺服方法采用 $s_w = \phi$ 控制转动自由度,定义扩展的图像坐标

$$m = [x, y, z]^T = [X/Z, Y/Z, \log Z]^T$$

控制平动自由度。采用上述特征实现了平动和转动自由度的部分解耦,雅可比矩阵具有上三角形式,对摄像机标定误差具有一定的鲁棒性。但是该方法需要在线估计 Z , 目前仍是一个亟待解决的问题。

3 基于图像矩的混合视觉伺服方法

采用混合视觉伺服控制方法,需要知道图像平面内 4 对匹配的特征点来确定唯一的 G 。由于局部特征匹配性能好,十分适合在视觉伺服中应用。然而当摄像机视角发生变化时,一些特征点因失去独特性未被检测到,另一些特征点可能位于摄像机视野范围之外,这都将导致特征点匹配关系发生变化,称这种现象为特征改变。在视觉伺服中使用变化的特征仍是一个亟待解决的问题^[12]。

3.1 方法描述

文献[9]在处理特征改变情况时,定义一个对角阵 D ,若期望位置的特征点和当前位置提取的特征点存在匹配,则该处值设为 1,否则为 0。采用基于图像的视觉伺服控制方法的控制率为

$$T_1 = - (DL(s_1))^+ D(s_1 - s_1^*). \quad (6)$$

其中: $s_1 = [u_1, v_1, u_2, v_2, \dots, u_k, v_k]^T$ 为特征点的像素坐标, k 为特征点个数, L 为控制增益矩阵, $L(s_1)$ 为图像雅可比矩阵, $(DL)^+$ 为矩阵 (DL) 的广义逆矩阵。

若匹配的特征点不变,则采用传统的混合视觉伺服能够实现控制目的。当特征改变时,文献[13]通过反向映射计算未匹配特征点的像素坐标。设 $G(i)$ 表示单应性矩阵的第 i 行,则 p^* 在当前位置的像素坐标为

$$u = \frac{G(1) * p^*}{G(3) * p^*}, \quad v = \frac{G(2) * p^*}{G(3) * p^*}. \quad (7)$$

文献[13]采用如下特征控制平动自由度:

$$\bar{m}^T = [\bar{x}, \bar{y}, \bar{z}] = \begin{bmatrix} \mathbf{1}_{k \times 1} & x_i & \mathbf{1}_{k \times 1} & y_i & \mathbf{1}_{k \times 1} & z_i \end{bmatrix}. \quad (8)$$

特征 $s_2 = [\bar{m}^T, \phi^T]^T$ 对应的雅可比矩阵 $L(s_2)$

具体形式见文献[13],采用的控制率为

$$T_2 = - L^+ (s_2) (s_2 - s_2^*), \tag{9}$$

在上述两种伺服方法中,雅可比矩阵包含了三维空间每点相对于摄像机的距离 Z ,一般用期望位置的距离 Z^* 代替.当特征改变时,采用式(6)虽然能够实现连续控制,但图像雅可比矩阵结构改变,可能出现奇异情况,使伺服任务失败.

Chaumette^[14]给出了点和图像区域矩的雅可比矩阵形式,使在视觉伺服中使用图像矩成为可能.文献[15]采用 3 个归一化的矩控制平动自由度,即

$$a_n = Z^* \sqrt{\frac{a}{a^*}}, u_n = a_n u_g, v_n = a_n v_g, \tag{10}$$

$$a = \mu_{20} + \mu_{02}. \tag{11}$$

其中: (u_g, v_g) 为目标重心的像素坐标, μ_{20} 和 μ_{02} 为二阶中心矩.特征 $s_v = [u_n, v_n, a_n]^T$ 对应的雅可比矩阵形式为

$$L(s_v) = [- I \quad L_{vw}], \tag{12}$$

式中 L_{vw} 仅与特征点的像素坐标有关^[15].本文采用如下特征控制机器人在三维空间的运动:

$$s_3 = (s_v^T, s_w^T)^T = (u_n, v_n, a_n, \phi^T)^T. \tag{13}$$

由文献[10]知 $L(s_w) = [0 \quad L_w]$,且 $L_w^{-1} = I$,故特征 s_3 对应的雅可比矩阵形式为

$$L(s_3) = \begin{bmatrix} - I & L_{vw} \\ 0 & L_w \end{bmatrix}. \tag{14}$$

雅可比矩阵 $L(s_3)$ 中不含 Z ,只在特征 s_3 中包含 Z^* ,其可以离线获得.雅可比矩阵具有同传统的混合视觉伺服一样的上三角形,转动自由度和平动自由度实现了部分解耦.当特征点改变时,雅可比矩阵形式不变,通过反向映射使误差信号不会出现突变情况,因而能够实现连续稳定的控制.为使误差函数以指数形式趋向于 0,即

$$\dot{e}(s_3) = - e(s_3), \tag{15}$$

由于满足如下关系式:

$$\dot{s}_3 = L(s_3) T_3, \tag{16}$$

故采用的控制率为

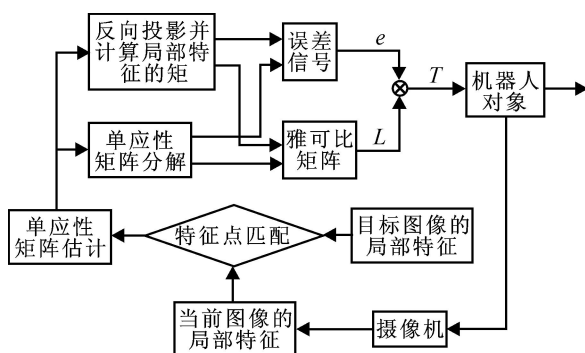


图 2 基于图像矩的混合视觉伺服结构框图

$$T_3 = - L^+ (s_3) e_3 = - \begin{bmatrix} - I & L_{vw} \\ 0 & I \end{bmatrix} e_3. \tag{17}$$

把采用特征和控制率分别如式(13)和(17)所示的伺服方法称为基于图像矩的混合视觉伺服(MHVS),结构框图如图 2 所示.由于在计算控制率时不需要进行矩阵求逆运算,因而非常适合在实际工程中应用.

3.2 稳定性分析

稳定性是控制系统一个重要指标.下面讨论采用如式(17)所示控制率的系统稳定性和对摄像机内参数标定误差的鲁棒性.

由于摄像机内参数标定存在误差,估计得到的误差 e 和真实误差 e 之间满足 $\hat{e} = Ee$,此时雅可比矩阵为 \hat{L} .构造标量函数 $V = 0.5 * e(t)^2$,其对时间的导数为 $\dot{V} = e^T \dot{e} = - e^T \hat{L} \hat{L}^+ Ee$.为使该系统稳定,根据李亚普诺夫稳定性判据,需要满足条件

$$\hat{L} \hat{L}^+ E = \begin{bmatrix} I & L_{vw} - \hat{L}_{vw} \\ 0 & L_w \end{bmatrix} \begin{bmatrix} E_v & 0 \\ 0 & E_w \end{bmatrix} > 0. \tag{18}$$

结合式(1),(16)~(18),整理有

$$\begin{bmatrix} \dot{e}_v \\ \dot{e}_w \end{bmatrix} = \begin{bmatrix} - e_v E_v e_v - w(L_{vw} - \hat{L}_{vw}) E_w e_w \\ - w L_w E_w e_w \end{bmatrix}. \tag{19}$$

式(19)表明,平动和转动自由度实现了部分解耦,故分别讨论其稳定性.转动自由度的局部渐近稳定性证明见文献[10],下面考虑平动自由度的稳定性.当 $e_w(t) \rightarrow 0$ 时, $f(e_w, t) = (L_{vw} - \hat{L}_{vw}) E_w e_w(t)$ 趋向于 0. E_v 是一个只与摄像机内参数有关且特征值具有正实部的 3×3 矩阵,在平衡点 $e_v = 0, e_w = 0$ 附近邻域,考虑如下函数:

$$\dot{e}_v = - e_v E_v e_v + w f(e_w, t), \tag{20}$$

当 $t \rightarrow \infty, e_w \rightarrow 0$ 时,也是局部渐近稳定的.式(19)描述的系统在平衡点(0,0)附近是局部渐近稳定的.

4 仿真实验

采用对极几何工具箱^[16],在 Matlab 下通过仿真验证本文提出的基于图像矩的混合视觉伺服方法的性能.为比较各种视觉伺服控制方法,Gans 等^[17]设计了如下 4 个任务:1)绕光轴 Z 的转动;2)沿光轴 Z 的平动;3)绕 Y 轴的转动;4)一般运动.并采用使任务收敛的迭代次数、最大特征点漂移和最大摄像机漂移等度量,比较各种方法的性能.因改变控制增益矩阵能够控制迭代次数,故本文采用相同的控制增益矩阵和迭代次数,计算任务结束时特征点像素误差和、摄像机位置误差和姿态误差、摄像机路径长度,来衡量式(6)(IBVS), (9)(HVS)和(17)

(MHVS) 三种控制策略的性能.

$$A^* = \begin{bmatrix} 416 & 0 & 160 \\ 0 & 312 & 120 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A = \begin{bmatrix} 417 & 0 & 163 \\ 0 & 313 & 117 \\ 0 & 0 & 1 \end{bmatrix}. \quad (21)$$

仿真实验中, 摄像机三维空间位姿表示为 $[t_x, t_y, t_z, x, y, z]^T$, 其中前 3 项表示摄像机相对于三维世界坐标系原点的平移向量, 后 3 项表示绕坐标轴的转动角度. 局部特征点在垂直于 z 轴且经过点 $[0, 0, 4]^T$ 的平面随机产生, 摄像机期望和实际的内部参数矩阵如式 (21) 所示, 获得的图像分辨率为 320×240 . 迭代次数设置为 300, 控制增益矩阵为 $\text{diag}[0.02, 0.02, 0.02, 0.02, 0.02, 0.02]$. 为检验 3 种伺服方法的抗噪声性能, 将特征点像素坐标添加一个零均值单位方差的高斯噪声. 采用随机方法设置特征点的匹配情况, 但不包含错误匹配情况.

表 1 给出了本文采用的 4 个任务. 表 2 为采用上述 3 种控制策略的度量值比较. 从表 2 可以看出, 3 种方法均成功完成了伺服任务, 任务完成时的位置和姿态误差接近于 0. 由于 IBVS 的目标函数是使对应点的像素坐标误差趋向于 0, 故在后 3 个任务中像素误差和最小, 但由于不能很好地处理摄像机绕光轴转动情况, 此时的像素误差和最大.

表 1 伺服任务描述

任务	初始位置	期望位置
1	$[0, 0, 0, 0, 0, \frac{8}{9}]^T$	
2	$[0, 0, -2, 0, 0, 0]^T$	$[0, 0, 0, 0, 0, 0]^T$
3	$[0, 0, -4, 0, \frac{1}{9}, 0]^T$	
4	$[0, 0, -4, \frac{1}{18}, \frac{1}{18}, \frac{5}{18}]^T$	$[1, 0, 1, -\frac{1}{18}, -\frac{1}{9}, 0]^T$

图 3 为任务 1 中采用 MHVS 方法特征点在图像平面的运动轨迹. 由于采用了混合视觉伺服框架, 特征点在图像平面的轨迹近似为椭圆, 虽然出现了特征点改变情况, 仍能完成伺服任务. 采用 IBVS 方法产生了严重的摄像机回退现象, 摄像机沿轴的最大位移为 8, 路径长度远大于其他两种方法. MHVS 方法在任务 1 中性能最好.

在任务 2 和 3 中, MHVS 在像素、位置和姿态误差表现出了与 IBVS 相同的性能, HVS 的像素误差和最大, 说明其受 z 值影响较大, 而 MHVS 有效消除了这方面的影响, 提高了控制性能.

表 2 三种伺服方法性能比较

任务	伺服方法	像素误差	位置误差	姿态误差	路径长度
1	IBVS	47.48	0.09	2.79	16.40
	HVS	47.18	0.07	2.79	0.62
	MHVS	19.59	0.05	2.79	0.72
2	IBVS	3.63	0.05	0	1.99
	HVS	39.33	0.07	0	2.12
	MHVS	3.81	0.05	0	2.00
3	IBVS	5.76	0.05	0.35	4.05
	HVS	32.08	0.07	0.35	4.28
	MHVS	5.81	0.05	0.35	4.13
4	IBVS	6.47	0.03	1.07	7.36
	HVS	50.78	0.09	1.07	5.91
	MHVS	16.90	0.05	1.07	5.31

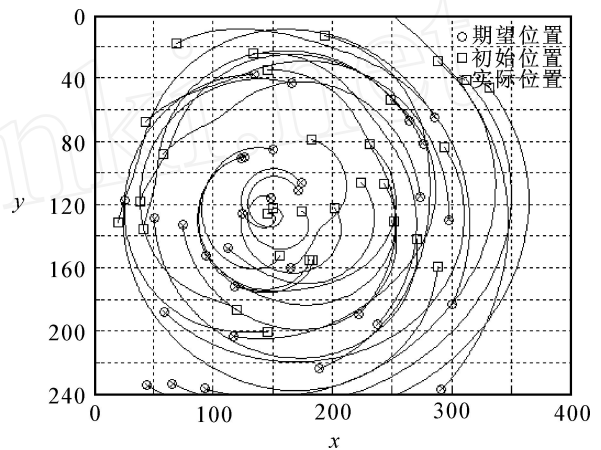


图 3 任务 1, 采用 MHVS 方法特征点在图像平面运动轨迹. 在任务 4 中, 虽然 IBVS 取得了最好的性能, 但从图 4 可以看出, 由于匹配的特征点改变, 使得摄像机的线速度和角速度出现剧烈变化. MHVS 方法采用了反向映射计算未匹配点的像素坐标, 故速度相对平滑, 如图 5 所示.

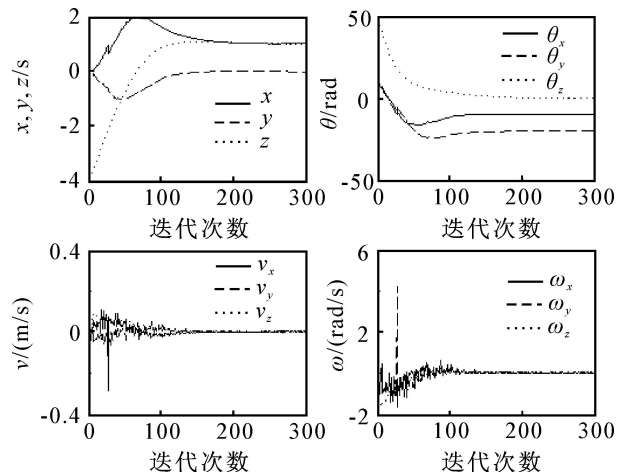


图 4 任务 4, 采用 IBVS 方法的位姿变化和控制输入曲线

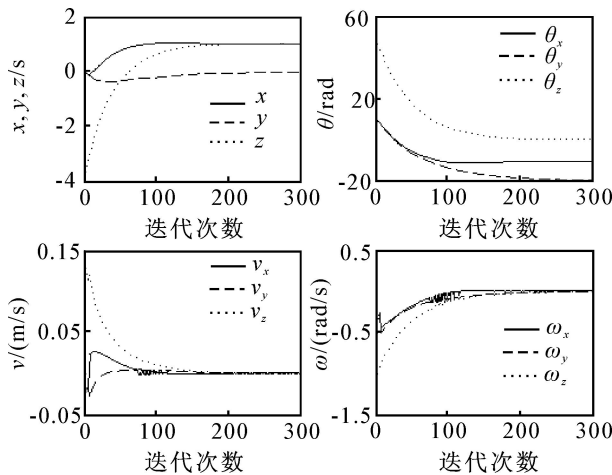


图 5 任务 4,采用 MHVS 方法的位姿变化和控制输入曲线

从上述分析可以看出,基于图像矩的混合视觉伺服方法能够完成各种伺服任务,由于利用像素坐标反向映射有效地消除了特征变化引起的雅可比矩阵结构改变和控制信号跳变情况,适用于局部特征.

5 结 论

本文提出了一种基于图像矩的混合视觉伺服方法,解决了特征改变引起的雅可比矩阵结构改变和控制信号跳变等问题.采用特征的雅可比矩阵不存在奇异情况,计算控制信号不需要矩阵求逆运算,因而非常适合在实际工程中应用.本文从理论上证明了所采用的控制策略的稳定性和对摄像机标定误差的鲁棒性.仿真实验表明该方法提高了控制性能,适用于局部特征.

参考文献(References)

[1] 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 Q J, Lian G Y, Sun Z Q. Survey of robot visual servoing[J]. Control and Decision, 2001, 16(6) : 849-853.)

[2] Hutchinson S, Hager G D, Corke P I. A tutorial on visual servo control[J]. IEEE Trans on Robotics and Automation, 1996, 12(5) : 651-670.

[3] Espiau B, Chaumette F, Rives P. A new approach to visual servoing in robotics[J]. IEEE Trans on Robotics and Automation, 1992, 8(3) : 313-326.

[4] Tuytelaars T, Mikolajczyk K. Local invariant feature detectors: A survey [J]. Foundations and Trends in Computer Graphics and Vision, 2008, 3(3) : 177-280.

[5] Shademan A, Janabi-Sharifi F. Using scale-invariant feature points in visual servoing[C]. Proc of the SPIE Conf on Machine Vision and Its Optomechtronics

Applications. Bellingham : SPIE, 2004 : 63-70.

[6] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. Int J of Computer Vision, 2004, 60(2) : 91-110.

[7] Lazar C, Burlacu A. Performance evaluation of point feature detectors for eye-in-hand visual servoing [C]. Proc of the 5th IEEE Int Conf Industrial Informatics. Piscataway: IEEE, 2007 : 497-502.

[8] Hoffmann F, Nierobisch T, Seyffarth T, et al. Visual Servoing with Moments of SIFT Features[C]. Proc of IEEE Int Conf on Systems, Man and Cybernetics. Piscataway: IEEE, 2006 : 4262-4267.

[9] Thi Thanh Hai T, Marchand E. Real-time keypoints matching: Application to visual servoing [C]. Proc of IEEE Int Conf on Robotics and Automation. Piscataway: IEEE, 2007 : 3787-3792.

[10] Malis E, Chaumette F, Boudet S. 2-1/2-D visual servoing [J]. IEEE Trans on Robotics and Automation, 1999, 15(2) : 238-250.

[11] Hartley R, Zisserman A. Multiple view geometry in computer vision [M]. 2nd ed. London: Cambridge University Press, 2004.

[12] Garcia-Aracil N, Malis E, Aracil-Santonja R, et al. Continuous visual servoing despite the changes of visibility in image features [J]. IEEE Trans on Robotics and Automation, 2005, 21(6) : 1214-1220.

[13] Malis E, Chesi G, Cipolla R. 2-1/2-D visual servoing with respect to planar contours having complex and unknown shapes[J]. Int J of Robotics Research, 2003, 22(10/11) : 841-853.

[14] Chaumette F. Image moments: A general and useful set of features for visual servoing[J]. IEEE Trans on Robotics and Automation, 2004, 20(4) : 713-723.

[15] Tahri O, Chaumette F. Point-based and region-based image moments for visual servoing of planar objects [J]. IEEE Trans on Robotics and Automation, 2005, 21(6) : 1116-1127.

[16] Mariottini G L, Prattichizzo D. EGT for multiple view geometry and visual servoing: Robotics vision with pinhole and panoramic cameras [J]. IEEE Trans on Robotics and Automation, 2005, 12(4) : 26-39.

[17] Gans N R, Hutchinson S A, Corke P I. Performance tests for visual servo control systems, with application to partitioned approaches to visual servo control [J]. Int J of Robotics Research, 2003, 22(10/11) : 955-981.