

机器人手眼协调的图象直接反馈方法*

蒋 平 胡凤轩 陈辉堂 王月娟

(同济大学电气系·上海, 200092)

摘要: 本文从视觉与控制集成观点出发, 提出了一种机械人手眼协调的新方法, 利用手部摄像机于期望抓取(操作)位置获取待抓(加工)物体图象, 以此作为期望图象, 而后可以对任意位置的同类物体进行实时抓取. 首先将期望图象与实时采样图象进行比较生成差图象, 从而依据差图象进行实时伺服反馈, 最终达到消除图象差的目的.

关键词: 机器人视觉; 伺服; 手眼协调; 反馈

1 引言

手眼协调问题是视觉引导下的机器人运动控制的基本问题. 以往基于视觉传感器的机器人操作是根据图象获取目标对象的位姿信息后, 驱动机器人完成作业, 这一过程可以采用视觉伺服方式加以控制^[1~2], 但都离不开图象特征抽取过程(抽取特征点^[1]、估计方位角^[3~4]等), 因此特征抽取质量直接影响控制精度, 例如采取计算图象主轴的方法进行方位角测量时^[3], 当物体无明显主轴时, 算法的数值误差将明显增大; 即使假定这一特征抽取绝对精确, 也无法确保机械手末端的准确定位, 因为这里视觉只作为一种由象平面到运动空间的映射工具, 而实际控制则以这一映射量作为控制目标, 因此其定位精度直接依赖于图象系统的标定精度, 而准确标定是非常困难甚至无法做到的.

本文针对平面物体抓取这一基本过程, 试图直接运用图象误差进行机械手位姿控制, 在抓取某一类物体前, 进行一次期望图象学习, 在实际抓取过程中, 由期望图象与实际待抓物体图象生成差图象, 实时控制摄像机运动以消除这一图象差, 因此整个过程是基于图象误差的实时视觉反馈, 由于直接以整个图象是否一致为控制目标, 因此本方案控制精度较高, 同时由于存在一次实际摄取期望图象的学习过程, 标定误差同样被引入期望图象, 因此本方案对标定误差并不敏感. 基于以上思想, 本文首先提出了基于图象误差的最优速降法, 它具有局部的收敛特性; 尔后我们将它与快速模拟退火算法相结合, 实现全局寻优, 这类似于人类的“尝试”与“摸索”过程.

2 具有局部收敛特性的图象误差控制

考虑到生产线上机器人往往对同一种类部件进行重复操作, 我们首先将一待加工物样本置于工作台上, 手动控制机械手到达理想的加工位置, 此时手部摄影机摄取待加工物图象, 这幅图象可以完全反映待加工物体与机械手末端操作器间的期望位置与姿态关系, 这一过程是学习过程, 可获得期望图象 I_d , 而后可进行实际加工或抓取. 以传送带或其它方式输送至工作区的待加工物, 可能具有任意的位置及方向, 机械手将由期望图象 I_d 及实时采集图象 I_r 生成图象差 $e = I_d \ominus I_r$, 其中 $\ominus$ 为遵循某种定义的图象差, 并根据这一图象差驱动机械手运动, 直至消除该误差. 本文仅讨论平面物体抓取问题.

* 国家自然科学基金资助项目.

本文于 1996 年 1 月 3 日收到. 1996 年 9 月 26 日收到修改稿.

首先我们讨论具有连续灰度情况, 设经学习后获得的期望图象为 $I_d = \{f_d(x, y)\}$, 实时采样图象为 $I_s = \{f_s(x, y)\}$, 其中 $f_d(x, y)$ 、 $f_s(x, y)$ 为图象平面 (x, y) 处象素点灰度, 差图象定义如下:

$$e = I_d \ominus I_s = \{1/2(f_d(x, y) - f_s(x, y))^2\}. \quad (1)$$

可以看出差图象是两幅图象对应点灰度值的某种操作, 它仍是一幅图, 差图象中各点强度反映了两幅图象间的差别大小. 由于摄像机与末端操作器间的位置关系固定, 因此若差图象中各象素值为零, 即 $e = \{0\}$ 时, 表明摄像机与待抓取物体间已达到理想的抓取位置, 可令机械手进行操作.

考虑摄像机系统 R_s , 它具有单位长度焦距 f , 空间一点 P 在 R_s 坐标系中为 $r = (X_s, Y_s, Z_s)$, 则对应于图象平面中的投影点 p 的坐标为 (x, y) , 设机械手运动引起手部摄影机在静止环境中运动, 其平动速度 $T = (T_x, T_y, T_z)^T$ 及转动角度速度 $\omega = (\omega_x, \omega_y, \omega_z)^T$, 则图象平面光流 $u = \dot{x}$, $v = \dot{y}$ 可计算如下^[3]:

$$u = \frac{d}{dt} \left(\frac{X_s}{Z_s} \right) = \frac{\dot{X}_s}{Z_s} - \frac{X_s \dot{Z}_s}{Z_s^2} = \left[x \frac{T_x}{Z_s} - \frac{T_x}{Z_s} \right] + [xy\omega_x - (1 + x^2)\omega_y + y\omega_z], \quad (2)$$

$$v = \frac{d}{dt} \left(\frac{Y_s}{Z_s} \right) = \frac{\dot{Y}_s}{Z_s} - \frac{Y_s \dot{Z}_s}{Z_s^2} = \left[y \frac{T_y}{Z_s} - \frac{T_y}{Z_s} \right] + [(1 + y^2)\omega_x - xy\omega_y - x\omega_z]. \quad (3)$$

当摄像机光轴保持与工作平面垂直且高度不变时, I_s 图象平面光流由机械手运动 T_x, T_y, ω_z 决定. 事实上差图象的灰度变化既可以认为是由于 I_s 运动引起的, 也可以认为是在基坐标系下, I_d 做反方向运动引起的, 由于对 I_d 的处理可以离线完成, 所以我们以 I_d 运动 I_s 静止这种等价形式进行讨论. 此时光流方程为:

$$u' = -u = \frac{T_x}{Z_s} - y\omega_z, \quad (4)$$

$$v' = -v = \frac{T_y}{Z_s} + x\omega_z. \quad (5)$$

所以控制目标为通过机械手运动使 I_d 灰度变化达到 $I_d = I_s$, 首先建立 Lyapunov 函数:

$$S = \iint_e (f_d(x, y) \ominus f_s(x, y)) dx dy = \frac{1}{2} \iint_e (f_d(x, y) - f_s(x, y))^2 dx dy,$$

其中 I_d 运动 I_s 静止, 所以

$$\begin{aligned} \dot{S} &= \iint_e ((f_d(x, y) - f_s(x, y)) \frac{d(f_d(x, y))}{dt}) dx dy \\ &= \iint_e (f_d(x, y) - f_s(x, y)) \left(\frac{\partial(f_d(x, y))}{\partial x} u' + \frac{\partial(f_d(x, y))}{\partial y} v' \right) dx dy. \end{aligned}$$

代入(4), (5)式得

$$\begin{aligned} \dot{S} &= \iint_e (f_d(x, y) - f_s(x, y)) \left[\frac{\partial f_d(x, y)}{\partial x} \cdot \frac{T_x}{Z_s} + \frac{\partial f_d(x, y)}{\partial y} \cdot \frac{T_y}{Z_s} \right. \\ &\quad \left. + \left(\frac{\partial f_d(x, y)}{\partial y} \cdot x - \frac{\partial f_d(x, y)}{\partial x} \cdot y \right) \omega_z \right] dx dy. \end{aligned}$$

由于 $Z_s > 0$, 所以当

$$\begin{cases} T_x = -K_x \iint_e \frac{\partial f_d(x,y)}{\partial x} (f_d(x,y) - f_s(x,y)) dx dy, \\ T_y = -K_y \iint_e \frac{\partial f_d(x,y)}{\partial y} (f_d(x,y) - f_s(x,y)) dx dy, \\ \omega_z = -K_\omega \iint_e \left(\frac{\partial f_d(x,y)}{\partial y} \cdot x - \frac{\partial f_d(x,y)}{\partial x} \cdot y \right) (f_d(x,y) - f_s(x,y)) dx dy \end{cases} \quad (6)$$

时, $\dot{S} \leq 0$, 因此图象差稳定, 但由于(6)式无法保证 $S \neq 0$ 时 $\dot{S}$ 恒为零, 所以图象差不具有大范围渐近稳定特性, 而只能使能量函数 S 达到某个局部最小值, 而具体收敛情况与物体外形模式及初始状态有关。

由控制律(6)可以看出, 它只与期望图象梯度及期望图象与采样图象间的灰度差有关, 而期望图象梯度是可以离线计算的, 为了进一步减少运算量, 对图象进行二值化处理后, 算法具有非常简洁的形式。

设经学习后获得期望图象 $\{f_d(x,y)\}$, 实时采样图象为 $\{f_s(x,y)\}$, 其中 $x, y = 0, \dots, N-1$, 经二值化处理后可得期望二值图象和采样二值图象 $I_d = \{g_d(x,y)\}$, $I_s = \{g_s(x,y)\}$, 针对二值化图象我们给出另一种图象差定义:

$$e = I_d \ominus I_s = \{(g_s(x,y) - g_d(x,y)) \cdot g_s(x,y)\}. \quad (7)$$

上式表示采样图象中目标对象未能被期望图象覆盖部分, 如图 1 阴影部分。

定义差图象面积为新的 Lyapunov 函数如下:

$$S = \sum_{x=0}^N \sum_{y=0}^N (g_d(x,y) \ominus g_s(x,y)) = \sum_{x=0}^N \sum_{y=0}^N [(g_s(x,y) - g_d(x,y)) \cdot g_s(x,y)],$$

与连续灰度情况讨论方式一样, 考虑 I_s 静止 I_d 运动引起灰度变化, 则

$$\begin{aligned} \dot{S} = & - \sum_{x=0}^N \sum_{y=0}^N g_s(x,y) \left[\frac{\partial (g_d(x,y))}{\partial x} \cdot \frac{T_x}{Z_s} + \frac{\partial (g_d(x,y))}{\partial y} \cdot \frac{T_y}{Z_s} \right. \\ & \left. + \left(\frac{\partial (g_d(x,y))}{\partial y} \cdot x - \frac{\partial (g_d(x,y))}{\partial x} \cdot y \right) \cdot \omega_z \right]. \end{aligned}$$

由于 I_d 为二值图象, 我们补充定义二值图象梯度如下, 而不影响稳定性分析:

$$\frac{\partial g(r)}{\partial r} = G(r) \cdot \delta(r - \rho),$$

其中 ρ 为期望图象边界, δ 为高度等于 1 的脉冲函数, $G(r) = [G_x, G_y]^T$, G_x, G_y 为利用 Sobel 算子获得的梯度分量。

当下式满足时:

$$\begin{cases} T_x = K_x \sum_{\rho} g_s(x,y) G_{dx}(x,y), \\ T_y = K_y \sum_{\rho} g_s(x,y) G_{dy}(x,y), \\ \omega_z = K_\omega \sum_{\rho} g_s(x,y) (G_{dy}(x,y)x - G_{dx}(x,y)y). \end{cases} \quad (8)$$

可得 $\dot{S} \leq 0$, 图象差 e 稳定。

由控制律(8)可以看出本控制方案计算量与考虑连续灰度情况相比有明显减少, 图象平

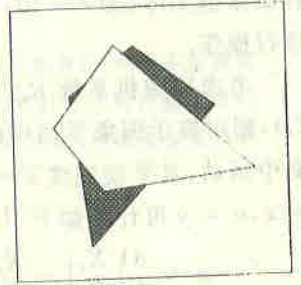


图 1 差图象 e

面的计算转化为期望图象的边界计算,而期望图象的边界抽取和梯度计算是可以离线获得的,因此,完全可以达到实时控制的目的,其实现框图如图 2。

由图 2 可以看出算法良好的实时性,机器人在进行期望图象学习后,离线生成两个数组,其一为边界坐标,其二为对应梯度,在实时控制时,控制器根据第一个数组中的边界坐标,判断采样图象对应点是否为黑色像素,若不是判断下一像素点,若是则进行(8)式累加计算,边界数组查询完毕后,则可输出一个采样周期的控制量(T_x, T_y, ω_z)。

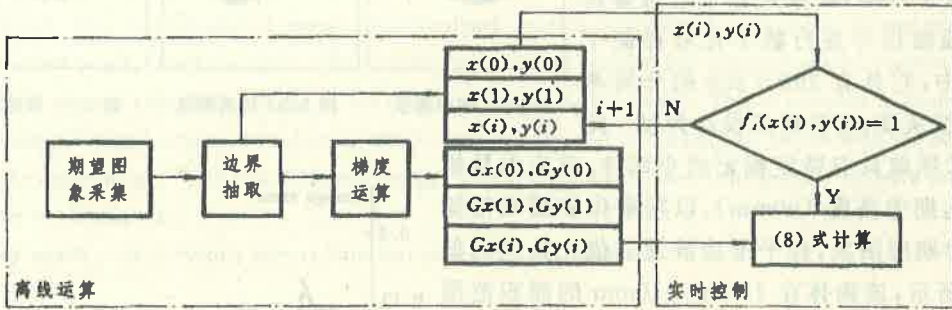


图 2 控制律实现框图

3 基于图象误差的全局控制方法

由以上分析可以发现,控制律无法保证在任意初始偏差下图象差收敛于零,它只能达到一个局部最小,对于平动部分全局只有唯一最小值,但有平坦的两翼,它表示物体尚未重合时的情况;而在旋转控制中这一问题就非常突出,通常其方位与误差面积关系 $\theta \sim S$ 图中,具有唯一零误差最小,但全程有若干个局部最小,所以视初始状态不同,采用(8)式控制可能收敛于任意一个局部最小,因此无法确保收敛于零误差。为了获得具有全局收敛性的图象误差控制,我们在平动部分引入物体中心控制,使大偏差时离开水平段,而转动部分采用有噪声的梯度下降法——模拟退火算法,实现全局寻优。所以平动部分控制律:

$$T = K_1 \sum_{\rho} g_s(x, y) G_d(x, y) + K_2 (\bar{r}_d - \bar{r}_s) \quad (9)$$

其中 $T = [T_x, T_y]^T$, $\bar{r}_d$ 为期望图象中心坐标, $\bar{r}_s$ 为采样图象中心坐标。

对于方位角控制问题,我们将(8)式描述的最优速降法与快速模拟退火算法^[5]相结合,得出了如下具有全局收敛特性和良好的实时性的控制策略:

首先摄像机摄取期望图象 $I_d(0)$ 后,离线生成一系列子图 $I_d(-M), \dots, I_d(M)$, 其中 $I_d(i)$, $i = -M, \dots, M$, 是经过 $I_d(0)$ 以 $\pi i/M$ 为转角增量旋转而得的图象,因此 $I_d(-M), \dots, I_d(M)$ 遍历一周。而 M 值的大小与待抓物体形状有关,当其 $\theta \sim S$ 图中有较多局部极小时,密度应相应增大,以描述出几个典型方位。之后,给定初温 $T(0)$,可以进行实时控制。

- 1) 在均匀分布区间 $[-0.5, 0.5]$ 中取一随机数 y ;
- 2) 采用快速模拟退火算法的柯西型生成函数,随机求取 $\Delta\theta$ 增量;
- 3) 根据 $\Delta\theta = 2k\pi + \Delta\theta'$, 计算 $\Delta\theta'$ 使 $-\pi < \Delta\theta' < \pi$, 并提取与其最接近的旋转图象 $I_d(i)$, 其中 $i = \text{int}(M\Delta\theta/\pi)$;
- 4) 求取以 $I_d(i)$ 为期望图象的差图象面积 $S(i)$;
- 5) 若 $S(i) < S(0)$, 其中 $S(0)$ 为当前差图象面积, 则驱动机械手运动 $\Delta\theta'$;
- 6) 若 $S(i) \geq S(0)$, 则由容忍函数决定是否接受 $\Delta\theta'$ 为新状态增量。按通常模拟退火算法此时应开始新的搜索, 但从实时控制角度考虑, 应在伺服周期内及时输出控制量, 我们令机械

手按(8)式控制律驱动手部摄像机旋转,实现最优速降;

7) $T(t) = T(0)/t$ 降温;重复 1)~7)。

4 实验结果

本次实验在 scorbobot ER-VII 五关节机器人上完成,采用基于 68000 的运动控制器,一台 SXH-II 型 CCD 摄像机安装于机械手末端,图象采集卡将摄像机摄取的视频信号进行数字化后存储于 DRAM 中,它具有 256×256 的分辨率以及 16 级灰度,实验任务设计为以一定方位要求抓取具有特定标记的小部件。首先引导机械手到达期望高度 (90mm),以期望位姿摄取待抓物体获得期望图象,经平滑滤波及二值化处理后如图 3(a) 所示,该物体在 $10\text{mm} \times 10\text{mm}$ 的面积范围内,并且无明显主轴。

我们采用模拟退火方法保证方位角的全局收敛,以 10° 为间隔共存储 35 幅期望子图,模拟退火算法初温 $T(0) = 10$,降温比 $t = 2$,取控制参数为: $K_{1x} = K_{1y} = 0$; $K_{2x} = K_{2y} = 0.1\text{mm/pixel}$; $k_w = 0.01p^\circ/\text{pixel}$ 其中 p 为图象误差面积与物体总面积之比,目的在于改善平衡点附近的控制特性。随机放置待抓物体如图 3(b),最终达到图 3(c) 的期望定位位置,其收敛过程的图象误差如图 4 所示,可以看出经 20 次反馈系统可以达到 1% 的面积误差率以下。

5 结 论

作为一种尝试,本文提出了一种采用视觉误差的图象直接反馈方法,用于机械手定位与抓取中。结果表明,它特别适合应用于难以用全局性描述,又有严格方位要求的机器人定位控制系统中;由于它的闭环控制作用,可以获得较高的定位精度,并与标定误差无关;同时由于期望图象的离线处理,保证了算法良好实时性;实验表明这一方法对于各种复杂物体具有广泛性和一般性,而不象基于特征的控制方法那样,对于不同物体需选取不同特征以便易于物体外形的准确描述。

参 考 文 献

- 1 Papanikolopoulos, N. P., Khosla, P. K. and Kanade, T.. Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision. IEEE Trans. on Robotics and Automation, 1993, 1(9): 14-35
- 2 Wijesoma, S. W., Wolfe, D. F. H. and Richards, R. J.. Eye-to-hand coordination for vision-guided robot control applications. Int. J. of Robotics Research, 1993, 12(1): 64-78
- 3 Horn, B. K. P.. Robot vision. The MIT Press, 1986
- 4 苏剑波,李耀通.一种用于机器人的物体运动参数快速识别方法.自动化学报,1994,20(3):286-290
- 5 斯华龄.电脑人脑化:神经网络——第六代计算机.北京:北京大学出版社,1992

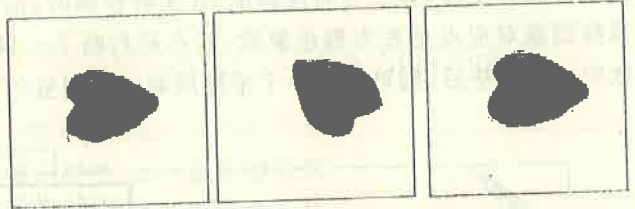


图 3(a) 期望图象

图 3(b) 初始图象

图 3(c) 最终图象

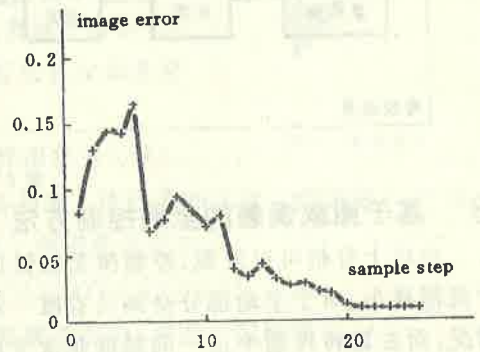


图 4 控制误差

Robot Hand-Eye Coordination with Direct Image Feedback

JIANG Ping, HU Fengxuan, CHEN Huitang and WANG Yuejuan

(Department of Electrical Engineering, Tongji University • Shanghai, 200092, PRC)

Abstract: This paper analyzes robot hand-eye coordination from the view point of integration of vision and control. The new scheme of image direct feedback control is proposed. The manipulating process is as follows: first of all, move the arm to the desired manipulating position and orientation with respect to the object of interest and sample the image of the object by eye-in-hand camera. This is referred to desired image. In the manipulating process, the scheme, which compares the real time sampled. Image with the desired one, forms the error image. With this 2-D error image, The control action steers the gripper to move until the image error is eliminated.

Key words: robot vision; servo; hand-eye coordination; feedback

本文作者简介

蒋 平 1963 年生, 1992 年毕业于西安交大信息与控制工程系, 获工学博士学位. 现为同济大学电气系副研究员. 研究领域为机器人控制与智能机器人.

胡凤轩 1972 年生. 同济大学电气系硕士生. 研究领域为机器人控制与智能机器人.

陈辉堂 1933 年生. 1953 年毕业于交通大学电力系, 现为同济大学电气系教授、博士生导师. 研究领域为机器人控制与智能机器人.

王月娟 1935 年生. 1958 年毕业于交通大学电机系, 现为同济大学电气系教授、博士生导师. 研究领域为机器人控制与智能机器人.