文章编号: 1001-0920(2001)06-0849-05

机 器 人 视 觉 伺 服 综 述

赵清杰, 连广宇, 孙增圻
(清华大学 计算机系, 北京 100084)

摘 要: 系统论述了机器人视觉伺服发展的历史和现状。从不同角度对机器人视觉控制系统进行分类, 重点介绍了基于位置的视觉伺服系统和基于图像的视觉伺服系统。对人工神经网络在机器人视觉伺服方面的应用情况作了介绍。讨论了视觉伺服中图像特征的选择问题。对机器人视觉所涉及的前沿问题进行阐述, 并指出了目前研究中所存在的问题及今后发展方向。

关键词: 机器人; 视觉伺服; 人工神经网络; 特征选择
中图分类号: TP 242.6 文献标识码: A

Survey of Robot Visual Servoing

ZHAO Qing-jie, LIAN Guang-yu, SUN Zeng-qi
(Department of Computer, Tsinghua University, Beijing 100084, China)

Abstract: The history and current research of robot visual servoing are reviewed. Robot visual control systems can be categorized according to different criteria. Based on the feedback representation, vision-based robot control techniques can be classified into position-based control and image-based control, which are discussed in detail. The application of neural networks in vision-based robot systems is introduced. Feature selection in visual servoing is discussed also. Several advanced problems and the direction in the research of robot vision are pointed out.

Key words: robots; visual servoing; artificial neural network; feature selection

1 引 言

制造能够替代人劳动的机器一直是人类的梦想。但目前即使是世界上智能最高的机器人, 对外部环境变化的适应能力也非常有限, 距人们预想的目标还有很大距离, 这极大地影响了机器人的推广使用。其中的一个重要原因就是机器人缺乏像人一样的感知能力。为解决这一问题, 学者们开始为机器人添加各种外部传感器, 其中比较重要的一种就是视觉传感器。

在现代工业自动化生产过程中, 机器视觉正成为一种提高生产效率和保证产品质量的关键技术, 如机械零件的自动检测、智能机器人控制及生产线的自动监控等。在国防和航天等领域, 机器视觉也具有相当重要的应用意义, 如运动目标的自动跟踪与识别、自治战车导航、登月舱的自动着陆以及空间机器人的视觉控制等[1]。但在机器人视觉领域中, 大部分研究仍建立在解析式描述的基础上, 并以满足某些特定任务为目标, 因而在实际中的应用十分有限。

收稿日期: 2000-07-06; 修回日期: 2000-10-17
基金项目: 国家 863 项目(863-2-5-2-2)
作者简介: 赵清杰(1966—), 女, 河南新野人, 博士生, 从事机器人视觉、智能控制技术研究; 孙增圻(1943—), 男, 江苏靖江人, 教授, 博士生导师, 从事智能控制、机器人和神经网络等研究。

如何使机器人具有类似人类视觉的能力, 是一个需要长期研究的问题。

2 机器人视觉伺服系统

2.1 机器人视觉发展历史简介

机器视觉是随着 60 年代末计算机与电子技术的快速发展而出现的。把视觉信息用于机械手定位的研究可以追溯到 70 年代。当时出现了一些实用性的视觉系统, 如应用于集成电路生产、精密电子产品装配、饮料罐装场合的检次等。80 年代后期, 出现了专门的图像处理硬件, 人们开始系统地研究机器人视觉控制系统。到了 90 年代, 随着计算机能力的增强和价格下降, 以及图像处理硬件和 CCD 摄像机的快速发展, 机器视觉系统吸引了众多研究人员的注意。在过去的几年里, 机器人视觉控制无论是在理论上还是在应用方面都有很大进步。

最早基于视觉的机器人系统, 采用的是静态 look and move 形式。即先由视觉系统采集图像并进行相应处理, 然后通过计算估计目标的位置来控制机器人运动。这种操作精度直接与视觉传感器、机械手及控制器的性能有关, 这使得机器人很难跟踪运动物体。到 80 年代, 计算机及图像处理硬件得到发展, 使得视觉信息可用于连续反馈, 于是人们提出了基于视觉的伺服(visual servoing)控制形式[2]。这种方式可以克服模型(包括机器人、视觉系统、环境)中存在的不确定性, 提高视觉定位或跟踪的精度。

2.2 机器人视觉控制系统分类

可以从不同的角度, 如反馈信息类型、控制结构和图像处理时间等方面对 vision-based 机器人控制系统进行分类[3~5]。

从反馈信息类型的角度, 机器人视觉系统可分为基于位置(position-based)的视觉控制和基于图像(image-based)的视觉控制。前者的反馈偏差在 3D Cartesian 空间进行计算, 后者的反馈偏差在 2D 图像平面空间进行计算。这将在后面做重点介绍。

从控制结构的角度, 可分为开环控制系统和闭环控制系统。开环控制的视觉信息只用来确定运动前的目标位姿, 系统不要求昂贵的实时硬件, 但要求事先对摄像机-机器人进行精确标定。闭环控制的视觉信息用作反馈, 这种情况下能抵抗摄像机与机器人的标定误差, 但要求快速视觉处理硬件。

根据视觉处理的时间可将系统分为静态和动态两类。

根据摄像机的安装位置可分为 eye-in-hand 安装方式和其它安装方式。前者在摄像机与机器人末端之间存在固定的位置关系, 后者的摄像机则固定于工作区的某个位置。最近也有人把摄像机安装在机械手的腰部, 即具有一个自由度的主动性[5]。

根据所用摄像机的数目可分为单目、双目和多目等。

根据摄像机观测到的内容可分为 EOL 和 ECL 系统[3]。EOL 系统中摄像机只能观察到目标物体; ECL 系统中摄像机同时可观察到目标物体和机械手末端, 这种情况的摄像机一般固定于工作区, 其优点是控制精度与摄像机和末端之间的标定误差无关, 缺点是执行任务时, 机械手会挡住摄像机视线。

根据是否用视觉信息直接控制关节角, 可分为动态 look-and-move 系统和直接视觉伺服(direct visual servo)系统。前者的视觉信息为机器人关节控制器提供设定点输入, 由内环的控制器控制机械手的运动; 后者用视觉伺服控制器代替机器人控制器, 直接控制机器人关节角。由于目前的视频部分采样速度不是很高, 加上一般机器人都有现成的控制器, 所以多数视觉控制系统都采用双环动态 look-and-move 方式。

此外, 也可根据任务进行分类, 如基于视觉的定位、跟踪或抓取等。

2.3 基于位置的视觉伺服系统

基于位置的方法, 视觉伺服误差定义在三维 Cartesian 空间[5]。视觉或特征信息用来估计机械手末端与目标的相对位姿[6~9]。参见图 1。

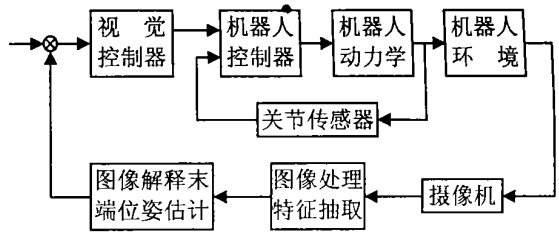

图 1 position-based 的视觉伺服系统

这种方法的主要优点是直接在 Cartesian 空间控制机械手的运动, 另外它把视觉重构问题从机器人控制中分离出来, 这样可以分别对二者进行研究[10]。但这种方法一般需要对视觉系统和机器人进行标定。另外由于要对图像进行解释, 因而增加了计算量。

2.4 基于图像的视觉伺服系统

image-based 视觉伺服系统, 其伺服误差直接定义在图像特征空间[11]。即摄像机观察到的特征信息

直接用于反馈,不需要对三维姿态进行估计。该方案的结构框图如图 2 所示。

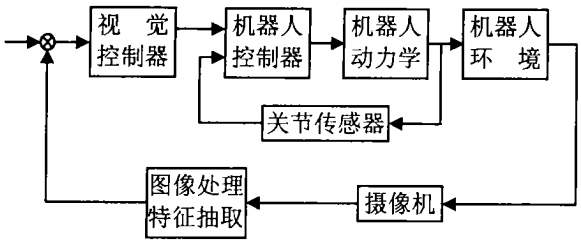

图 2　image-based 的视觉伺服系统

图像雅可比定义如下。

设 $s = [s_1, s_2, \cdots, s_p]^T$ 表示 Cartesian 空间下的末端位姿,$q = [q_1, q_2, \cdots, q_n]^T$ 表示对应的关节坐标,$f = [f_1, f_2, \cdots, f_m]^T$ 表示对应的特征空间坐标;$\dot{s}$ 表示末端的速度矢量(包括平移速度和旋转速度),$\dot{q}$ 表示关节速度矢量,$\dot{f}$ 表示图像特征的变化速度。

关节速度 $\dot{q}$ 与对应的任务空间速度 $\dot{s}$ 之间的关系为 $\dot{s} = J_1\dot{q}$,其中 J_1 称为机器人雅可比矩阵[12],表示向量 s 对向量 q 的导数。

任务空间速度 $\dot{s}$ 与对应的图像特征的变化速度 $\dot{f}$ 之间的关系为 $\dot{f} = J_2\dot{s}$,其中 J_2 称为 Local Jacobian[2],也有人称之为图像雅可比。

图像特征的变化速度 $\dot{f}$ 与对应的关节速度 $\dot{q}$ 之间的关系为 $\dot{f} = J\dot{q}$,其中 $J = J_2J_1$。

多数文献把 J 阵称为图像雅可比矩阵。其它称法还有特征敏感阵、变换阵[11]、B 阵[13]、composite Jacobian[14] 及 visual-motor Jacobian[15] 等。图像雅可比表示了在某个关节角附近图像特征变化与机器人关节角变化之间的关系。

计算图像雅可比的方法有经验法[16,17]、在线估计法[14,15,18] 和学习方法等。经验法可以通过标定或先验模型知识得到;在线估计图像雅可比的方法,可以事先不进行标定,但存在雅可比的初值选择问题;学习方法主要有离线示教和神经网络方法等。

image-based 方法绕过 3D 重构,直接用图像特征控制机器人运动。与 position-based 方法相比,image-based 方法受传感器模型、机器人运动学及标定所带来的误差影响较小,静态定位的精度也比较高。但动态估计雅可比时,需要不断进行更新和求逆,计算时间上也有待于进一步优化。另外,由于系统是耦合的且具有非线性,因此不能保证在整个任务空间都是收敛的。

最近,Malis[19] 提出 2-1/2-D 的视觉伺服方式。该方法将旋转与平移进行解耦,反馈信号一部分采用 3D 笛卡儿空间坐标表示,一部分采用 2D 图像空间坐标表示。这种方法可对 position-based 和 image-based 两种结构进行取长补短,系统的稳定性和收敛域都有所提高。

3　神经网络在视觉伺服中的应用

人工神经网络是生物神经网络的一种模拟和近似。神经网络特别适用于不能用显性公式表示的、具有复杂非线性关系的情况[16],并具有较强的适应和学习功能。如果用人工神经网络来表示机器人视觉控制中的 visual-motor 关系,便可省去标定机器人及视觉系统的复杂计算过程,把复杂的变换过程转化为简单的神经网络映射关系。

Hashimoto[20] 采用 BP 网络学习图像特征变化与机器人关节角变化之间的非线性关系。网络的输入是图像特征的变化,输出是期望的关节角的变化。整个过程分为学习阶段和执行阶段。学习阶段通过随机加入控制量,记录相应图像特征的变化,得到训练样本,离线对网络进行训练。执行时则根据图像特征变化由学习过的神经网络求出期望的关节角变化,然后由机器人控制器控制机器人末端到达目标位置。这种方法的缺点是学习的样本数量要求很大,对不同工件的通用性不强。Wells[21] 分别用 4 点特征、傅立叶描述子和几何矩作为神经网络的输入,对 6DOF 机器人进行定位实验。其与众不同之处在于:采用全局的图像特征,这样可扩大应用范围。但利用全局特征的定位精度比较低。Sun[22] 采用两个神经网:用一个分层 Kohonen 网作为基本网进行全局控制,视觉信号来自两个固定于工作空间的摄像机;另一个调整网采用 BP 网进行局部控制,视觉信号来自安装在末端上的两个摄像机。Stanley[23] 借助于神经网络进行特征抽取和求逆雅可比。

上述方法,多数都是针对具体物体上的具体特征,或只进行了仿真实验。

4　视觉伺服中的特征抽取

视觉伺服的性能依赖于控制回路中所用的图像特征。特征包括几何特征和非几何特征,机械手视觉伺服中常见的是采用几何特征。早期视觉伺服中用到的多是简单的局部几何特征,如点、线、圆圈、矩形、区域面积等以及它们的组合特征[11,13,18,24~27],其中点特征应用最多。局部特征虽然得到了广泛应用,

而且在特征选取恰当的情况下可以实现精确定位, 但当特征超出视域时则很难做出准确的操作。特别是对于真实世界中的物体, 其形状、纹理、遮挡情况、噪声、光照条件等都会影响特征的可见性。所以单独利用局部特征会影响机器人可操作的任务范围。近来有人在视觉控制中利用全局的图像特征, 如特征向量(eigenfeatures)[28], 几何矩(geometric moments)[21], 图像到直线上的投影, 随机变换, Fourier 描述子等[29]。全局特征可以避免局部特征超出视域所带来的问题, 也不需要在参考特征与观察特征之间进行匹配, 适用范围较广, 但定位精度比用局部特征低[21]。总之, 特征的选取没有通用的方法, 必须针对任务、环境、系统的软硬件性能, 在时间、复杂性和系统的稳定性之间进行权衡。

早期的视觉控制机器人, 一般取图像特征的数目与机器人的自由度相同。例如, Weiss 和 Sandersons[2]要求允许的机器人自由度数一定要等于特征数, 这样可以保证图像雅可比是方阵, 同时要求所选的特征是合适的, 以保证图像雅可比非奇异。由于自由度增加会使控制器设计复杂, 所以一般的实验都只采用 2～4 个自由度。文献中把数量比机器人自由度数多的特征集称为冗余特征(redundant feature), 从控制的观点看, 用冗余特征可抑制噪声的影响, 提高视觉伺服的性能[30]。

值得提出的是, 伺服控制用的特征与图像识别用的特征, 其选择指标有一定的差别[31]。从图像识别角度看, 特征应具有鲁棒性和唯一性; 但是从伺服的观点看, 特征又必须对物体姿态的变化具有敏感性, 即如果物体的位置和姿态发生变化, 图像的特征必须变化。另一方面特征必须可控, 即通过一系列的控制行为能够得到所选择的图像特征。

5 总结与展望

由于机器人视觉包括许多自成一体的研究领域, 如图像处理、计算机视觉、运动学、动力学、控制理论及实时计算等, 所以要考虑的问题比较多。其中的视觉部分, 与计算机视觉中的 active vision, shape from stereovision, shape from motion 等有相似之处。但计算机视觉与机器人视觉研究的最终目的不同: 前者主要研究视觉检验, 精度要求高, 速度不是主要问题; 而机器人视觉主要研究在视觉引导下机器人对环境的作用, 有实时性的要求。因此机器人视觉研究存在更多的困难。根据目前情况, 机器人视觉应在以下几个方面进一步加强研究:

1) 图像特征的选择问题。视觉伺服的性能密切依赖于所用的图像特征, 特征的选择不仅要考虑识别的指标, 还要考虑控制指标。从控制的观点看, 用冗余特征可抑制噪声的影响, 提高视觉伺服的性能[30], 但又会给图像处理增加难度。因此如何选择性能最优的特征, 如何处理特征以及如何评价特征, 都是需要进一步研究的问题。针对任务有时可能需要从一套特征切换到另一套, 可以考虑把全局特征与局部特征结合起来。

2) 结合计算机视觉及图像处理的研究成果, 建立机器人视觉系统的专用软件库。在视觉伺服中, 需要进行图像采集、图像处理、特征抽取及由二维信息重构三维信息等, 要处理的数据量较大, 算法复杂多样。如果有这样的软件平台, 在进行视觉伺服任务时, 就可以少走弯路。当然更希望生产出性能价格比较高的相关硬件。

3) 加强系统的动态性能研究。目前的研究多集中于根据图像信息确定期望的机器人运动这一环节上, 而对整个视觉伺服系统的动态性能缺乏研究。

4) 利用智能技术的成果。虽然神经网络在机器人视觉伺服中已得到应用, 但多数都是针对具体物体的具体特征, 或只进行了仿真实验, 还有待于进一步的研究。考虑到人类看到并拿起某个物体时, 事先并没有在数字上准确计算物体的位置, 而是通过不断地观察、判断和推理, 其中包含学习和模糊推理的内容, 由此可以考虑用模糊神经技术解决机器人视觉伺服问题。

5) 利用主动视觉的成果。主动视觉是当今计算机视觉和机器视觉研究领域中的一个热门课题[31]。它强调的是视觉系统与其所处环境之间的交互作用能力, 因而有可能使被动感知方式下一些病态问题转化为良态问题, 使非线性问题转化为线性问题。

6) 多传感器融合问题。视觉传感器具有一定的使用范围, 如能有效地结合其它传感器, 利用它们之间性能互补的优势, 便可以消除不确定性, 取得更加可靠、准确的结果。

参考文献:

[1] 郑南宁. 计算机视觉与模式识别[M]. 北京: 国防工业出版社, 1998.

[2] L E Weiss, A C Sanderson, C P Neuman. Dynamic sensor-based control of robots with visual feedback[J]. IEEE J Robot Automation, 1987, 3(5): 404-417.

[3] S Hutchinson, G Hager, P Corke. A tutorial introduc-

tion to visual servo control [J]. IEEE Trans on Robotics and Automation, 1996, 12(5): 651-670.

[4] S Wijesoma, D Wolfe, R Richards. Eye-to-hand coordination for vision-guided robot control applications[J]. Int J Robotics Research, 1993, 12(1): 65-78.

[5] M C Ngugen. Situation-oriented behavior-based stereo vision to gain robustness adaptation in manipulator control [A]. Proc of SPIE [C]. Massachusetts, 1999. 3837: 90-97.

[6] W J Wilson, C C Hulls, G S Bell. Relative end-effector control using Cartesian position-based visual servoing [J]. IEEE Trans on Robotics and Automation, 1996 12 (5): 684-696.

[7] D B Westmore, W J Wilson. Direct dynamic control of a robot using an end-point mounted camera and Kalman filter position estimation [A]. IEEE Conf on Robotics and Automation[C]. California, 1991. 3: 2376-2384.

[8] J Pauli. Learning to recognize and grasp objects [J]. Machine Learning, 1998, 31(1-3): 239-258.

[9] M Vargas, F R Rubio, A R Malpesa. Pose-estimation and control in a 3D visual servoing system[A]. Proc of the 14th IFAC[C]. Beijing, 1999. 317-323.

[10] G D Hager. A modular system for robust positioning using feedback from stereo vision[J]. IEEE Trans on Robotics and Automation, 1997, 13(4): 582-595.

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

[12] 孙增圻. 机器人智能控制[M]. 北京: 北京教育出版社, 1993.

[13] N P Papanikolopoulos, P K Khosla, T Kanade. Visual tracking of a moving target by a camera mounted on a robot: A combination of vision and control[J]. IEEE Trans on Robotics and Automation, 1993, 9(1): 14-35.

[14] H Sutanto, R Sharma, V Varma. The role of exploratory movement in visual servoing without calibration [J]. Robotics and Autonomous Systems, 1998, 23 (3): 153-169.

[15] M Jagersand. On-line estimation of visual-motor models for robot control and visual simulation[D]. USA: University of Rochester, 1998.

[16] 孙增圻. 智能控制理论与技术[M]. 北京: 清华大学出版社, 1996.

[17] W Jang, Z Bien. Feature-based visual servoing of an eye-in-hand robot with improved tracking performance [A]. IEEE Conf on Robotics and Automation [C]. California, 1991. 2254-2260.

[18] K Hashimoto, T Kimoto, T Ebine et al. Manipulator control with image-based visual servo[A]. IEEE Conf on Robotics and Automation [C]. California, 1991. 2267-2272.

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

[20] H Hashimoto, M Sato, F Harashima. Visual control of robotics manipulator based on neural networks[J]. IEEE Trans on Industrial Electronics, 1992, 39(6): 490-496.

[21] G Wells, C Venaille, C Torras. Promising research vision-based robot positioning using neural networks [J]. Image and Vision Computing, 1996, 14(10): 715-732.

[22] L Sun, CH Doeschner. Visuo-motor coordination of a robot manipulator based on neural networks [A]. IEEE Conf on Robotics and Automation[C]. Belgium, 1998. 1737-1742.

[23] K Stanley, Q M Wu, A Jerbi. Neural network-based vision guided robotics [A]. IEEE Conf on Robotics and Automation[C]. Michigan, 1999. 281-286.

[24] A Castano, S Hutchinson. Visual compliance: Task-directed visual servo control [J]. IEEE Trans on Robotics and Automation, 1994. 10(3): 334-342.

[25] S Skaar, W Brockman, R Hanson. Camera space manipulation [J]. Int J Robotics Research, 1987, 6(4): 20-32.

[26] S B Skaar, W H Brockman, W S Jang. Three-dimensional camera space manipulation [J]. Int J Robotics Research, 1990, 9(4): 22-39.

[27] J Feddema, O Mitchell. Vision- guided servoing with feature-based trajectory generation [J]. IEEE Trans on Robotics and Automation, 1989, 5(5): 691-700.

[28] S K Nayar, S A Nene, H Murase. Subspace methods for robot vision[J]. IEEE Trans on Robotics and Automation, 1996, 12(5): 750-758.

[29] G Wells, C Torras. Selection of image features for robot positioning using mutual information [A]. IEEE Conf on Robotics and Automation [C]. Belgium, 1998. 2819-2826.

[30] K Hashimoto, T Noritsugu. Performance and sensitivity in visual servoing [A]. IEEE Conf on Robotics and Automation[C]. Belgium, 1998. 2321-2326.

[31] H I Christensen, K W Bowyer, H Bunke. Active robot vision[M]. Singapore: World Scientific, 1993.