

眼在手上机器人手眼无标定三维视觉跟踪¹⁾

潘且鲁 苏剑波 席裕庚

(上海交通大学自动化研究所 上海 200030)

(E-mail: panqielu@263.net)

摘要 在未标定手眼关系及摄像机模型的情况下,建立了机器人三维视觉跟踪问题非线性视觉映射关系模型,并据此设计了基于人工神经网络的视觉跟踪控制器.仿真结果表明该算法能完全消除稳态跟踪误差,具有很强的环境适应性和容错能力,算法简单,易于实时实现.

关键词 视觉跟踪,神经网络,无标定,手眼协调

中图分类号 TP24

EYE-IN-HAND 3D ROBOTIC VISUAL TRACKING WITHOUT CALIBRATION

PAN Qie-Lu SU Jian-Bo XI Yu-Geng

(Institute of Automation, Shanghai Jiaotong University, Shanghai 200030)

(E-mail: panqielu@263.net)

Abstract In this paper, without explicit external and internal calibration, a nonlinear visual mapping model for 3D robotic visual tracking problem is proposed and a new visual tracking controller based on artificial neural network is designed. Simulation results show that this method can drive the static tracking error to zero quickly and keep good robustness and adaptability at the same time. Additionally, the algorithm is very easy to be implemented with low computational complexity.

Key words Visual tracking, neural network, calibration free, hand-eye coordination

1 引言

近年来,在智能机器人领域,手眼协调的研究受到了日益广泛的重视^[1].传统的手眼系统(如 look then move 系统)采用基于模型的控制方法,本质上属于开环静态控制,机械臂定位精度与模型标定准确度有关,误差累积现象严重,难以工程实现^[2].近年来,无标定方法引起各国学者的广泛注意^[3~5],其核心是在不精确或未知手眼关系条件下,利用视觉反馈误差规划机器人运动,从而完成视觉伺服、跟踪任务.由于采用了闭环反馈控制策略,从根本上消

1) 国家自然科学基金(69875010)和上海市高校青年科学基金(98QN18)资助

除了模型准确度对机械臂定位精度的影响. 问题的关键在于, 如何将视觉空间误差信息映射到机械臂所在的三维运动空间中去. 文[6,7]利用图像雅可比矩阵描述该映射关系, 以完成手爪定位任务, 但图像雅可比矩阵方法无法应用于目标和摄像机同时运动的场合(眼在手上的场合). 为克服这一困难, 文[8,9]将 ARMAX 模型应用于视觉跟踪, 取得了较好的效果, 但需要对手眼关系加以标定, 工程实现困难. 此外, 由于受雅可比矩阵收敛速度及参数估计精度的限制, 以上方法的控制性能难以进一步提高. 文[10]利用人工神经网络(ANN)来逼近逆图像雅可比矩阵, 完成视觉定位任务, 但由于逆图像雅可比矩阵与许多变量有关(如目标与手爪间的相对位置等), 该方法仅在局部范围内有效.

为克服以上方法的缺点, 必须尽可能充分地利用模型信息, 将其与 ANN 紧密结合起来. 换句话说, ANN 结构的选取并不是随意的, 而必须受到有关模型的制约和指导. 基于这一思想, 本文在手眼关系及摄像机模型完全未知的情况下, 建立了一种新的视觉跟踪问题非线性输入输出映射关系模型, 并据此设计了基于 ANN 的视觉跟踪控制器, 以完成三维目标跟踪任务.

2 问题描述

机器人视觉跟踪问题可定义为移动机器人手爪, 使得运动目标在图像特征空间中的投影始终固定在期望位置. 为防止图像特征

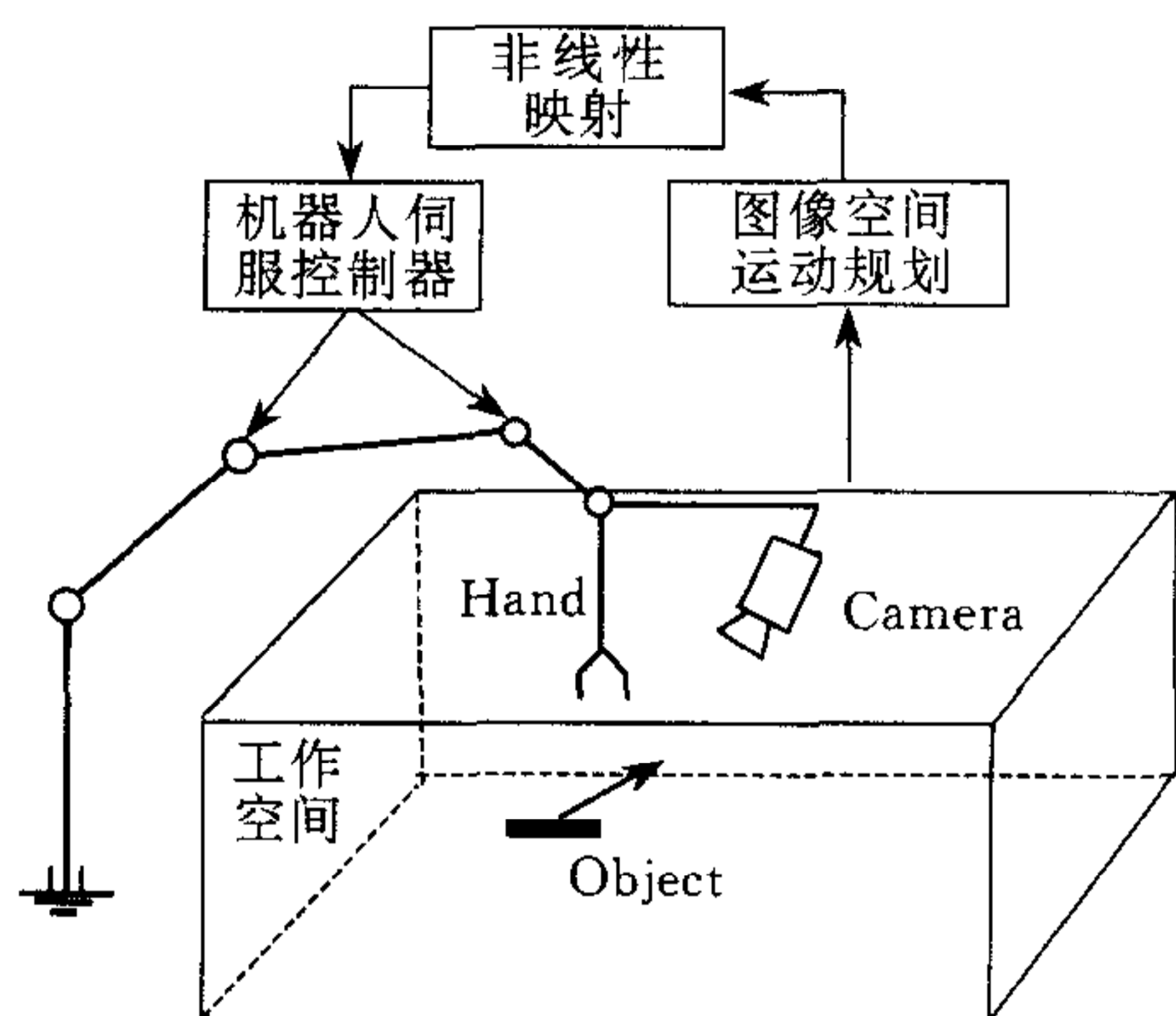


图 1 手眼系统结构示意图

被遮盖, 本文采用眼在手上构造方式, 如图 1 所示. 摄像机固定于手爪之上, 手眼关系和摄像机模型完全未知. 杆形目标在三维工作空间内自由平动, 视觉控制器在图像特征空间中进行运动规划, 并将规划结果映射到机器人运动空间中去, 作为运动指令下达给机器人伺服控制器, 从而驱动手爪进行平动跟踪, 直至追上目标并完成抓取任务. 为此, 需要建立图像特征空间与机器人运动空间之间的映射关系模型.

3 视觉映射关系模型

定义图像特征空间为 $\Omega = \text{span}\{x, y, z\}$, 其中 x, y 分别为杆形目标在图像平面上投影的中心点坐标, z 为图像平面上目标特征尺度, 此处为像的长度. 下面求取从机器人运动空间到图像特征空间的正映射模型. 设 k 时刻图像特征的瞬时速度(变化率)为 $[u(k), v(k), w(k)]^T$, 加速度为 $[\Delta u(k), \Delta v(k), \Delta w(k)]^T$. 由文献[8]可知

$$\begin{bmatrix} u(k) \\ v(k) \\ w(k) \end{bmatrix} = \begin{bmatrix} u_o(k) \\ v_o(k) \\ w_o(k) \end{bmatrix} + \begin{bmatrix} u_c(k) \\ v_c(k) \\ w_c(k) \end{bmatrix} \quad (1)$$

其中 $[u_o, v_o, w_o]^T$ 表示由于目标运动产生的速度分量, $[u_c, v_c, w_c]^T$ 表示因摄像机平移产生

的速度分量. 另外, 当帧率足够高时, $[u(k), v(k), w(k)]^T$ 可由下式近似求取^[8]

$$\begin{bmatrix} u(k) \\ v(k) \\ w(k) \end{bmatrix} = \frac{1}{T} \begin{bmatrix} x(k) \\ y(k) \\ z(k) \end{bmatrix} - \begin{bmatrix} x(k-1) \\ y(k-1) \\ z(k-1) \end{bmatrix} \quad (2)$$

其中 T 为视频采样周期, 为便于分析问题, 不失一般性地可令 $T=1$.

图像特征加速度 $[\Delta u, \Delta v, \Delta w]^T$ 是由目标和摄像机间的相对运动决定的. 在基坐标系中, k 时刻目标与摄像机之间的相对运动总可由以下三个物理量加以描述, 即目标瞬时加速度 $\mathbf{a}_o(k)$ 、摄像机与目标间的瞬时相对速度 $\mathbf{s}_{co}(k)$ 、摄像机瞬时加速度 $\mathbf{a}_c(k)$. 故可将目标与摄像机间的相对运动分解成: 相对速度为 $\mathbf{s}_{co}(k)$ 的匀速运动 + 目标速度为零、加速度为 $\mathbf{a}_o(k)$ 的加速运动 + 摄像机速度为零、加速度为 $\mathbf{a}_c(k)$ 的加速运动. 因此图像特征加速度 $[\Delta u, \Delta v, \Delta w]^T$ 也可分解为与之相对应的三部分, 即

$$\begin{bmatrix} \Delta u(k) \\ \Delta v(k) \\ \Delta w(k) \end{bmatrix} = \begin{bmatrix} \Delta u_o(k) \\ \Delta v_o(k) \\ \Delta w_o(k) \end{bmatrix} + \begin{bmatrix} \Delta u_{co}(k) \\ \Delta v_{co}(k) \\ \Delta w_{co}(k) \end{bmatrix} + \begin{bmatrix} \Delta u_c(k) \\ \Delta v_c(k) \\ \Delta w_c(k) \end{bmatrix} \quad (3)$$

上式中 $[\Delta u_o, \Delta v_o, \Delta w_o]^T$ 是由目标加速度 $\mathbf{a}_o(k)$ 引起的图像特征的加速度分量, 本文假设 $\mathbf{a}_o(k) \equiv 0$, 则此项为零; $[\Delta u_{co}, \Delta v_{co}, \Delta w_{co}]^T$ 表示相对速度 $\mathbf{s}_{co}(k)$ 经非线性透视投影引起的图像特征加速度分量, 其物理意义在于尽管目标与摄像机之间以相对速度 $\mathbf{s}_{co}(k)$ 作匀速运动, 但经过非线性透视投影后, 图像空间中投影点的运动将不再是匀速的, 从而产生了该加速度分量. 由于 $\mathbf{s}_{co}(k)$ 可由图像特征空间中投影点的坐标 $[x(k), y(k), z(k)]^T$ 和速度 $[u(k), v(k), w(k)]^T$ 唯一确定, 故 $[\Delta u_{co}, \Delta v_{co}, \Delta w_{co}]^T$ 可表示为

$$\begin{bmatrix} \Delta u_{co}(k) \\ \Delta v_{co}(k) \\ \Delta w_{co}(k) \end{bmatrix} = f_{co} \left(\begin{bmatrix} x(k) \\ y(k) \\ z(k) \end{bmatrix}, \begin{bmatrix} u(k) \\ v(k) \\ w(k) \end{bmatrix} \right) \quad (4)$$

式(3)中 $[\Delta u_c, \Delta v_c, \Delta w_c]^T$ 是摄像机加速度 $\mathbf{a}_c(k)$ 引起的图像特征加速度分量. 由于手爪作平动跟踪, 故摄像机加速度 $\mathbf{a}_c(k)$ 与手爪平移加速度 $[\Delta T_{hX}, \Delta T_{hY}, \Delta T_{hZ}]^T$ 完全相同, 因此 $[\Delta u_c, \Delta v_c, \Delta w_c]^T$ 可表示为

$$\begin{bmatrix} \Delta u_c(k) \\ \Delta v_c(k) \\ \Delta w_c(k) \end{bmatrix} = f_c \left(\begin{bmatrix} x(k) \\ y(k) \\ z(k) \end{bmatrix}, \begin{bmatrix} \Delta T_{hX}(k) \\ \Delta T_{hY}(k) \\ \Delta T_{hZ}(k) \end{bmatrix} \right) \quad (5)$$

其中手爪加速度 $[\Delta T_{hX}, \Delta T_{hY}, \Delta T_{hZ}]^T$ 可由手爪位置坐标 $[X_h(k), Y_h(k), Z_h(k)]^T$ 经二阶差分求取. 另外, 由于从机器人运动空间到图像特征空间的映射具有非线性, 故式(4)和(5)中 $f_{co}(*, *)$ 和 $f_c(*, *)$ 均为非线性函数, 且均与图像空间位置坐标 $[x(k), y(k), z(k)]^T$ 有关. 将式(4)和(5)代入式(3)可得正映射模型为

$$\begin{bmatrix} \Delta u(k) \\ \Delta v(k) \\ \Delta w(k) \end{bmatrix} = f' \left(\begin{bmatrix} x(k) \\ y(k) \\ z(k) \end{bmatrix}, \begin{bmatrix} u(k) \\ v(k) \\ w(k) \end{bmatrix}, \begin{bmatrix} \Delta T_{hX}(k) \\ \Delta T_{hY}(k) \\ \Delta T_{hZ}(k) \end{bmatrix} \right) \quad (6)$$

由于该模型描述的是一种光学映射关系, 系统响应无时间延迟或时间展宽, 故可将其输入、输出关系置换, 得到从图像特征空间到机器人运动空间的逆映射模型

$$\begin{bmatrix} \Delta T_{hX}(k) \\ \Delta T_{hY}(k) \\ \Delta T_{hZ}(k) \end{bmatrix} = f \left(\begin{bmatrix} x(k) \\ y(k) \\ z(k) \end{bmatrix}, \begin{bmatrix} u(k) \\ v(k) \\ w(k) \end{bmatrix}, \begin{bmatrix} \Delta u(k) \\ \Delta v(k) \\ \Delta w(k) \end{bmatrix} \right) \quad (7)$$

4 控制策略

为完成目标跟踪任务,必须把控制策略和视觉信息紧密结合起来.由于手眼关系未知及透视投影具有严重非线性,传统的基于自适应参数估计的控制方法难以利用.为此,我们根据式(7)所示的非线性映射关系构造了一个人工神经网络(ANN),将运动规划得到的控制指令映射到机器人运动空间中去,从而避免参数辩识的困难.控制系统结构如图 2 所示.

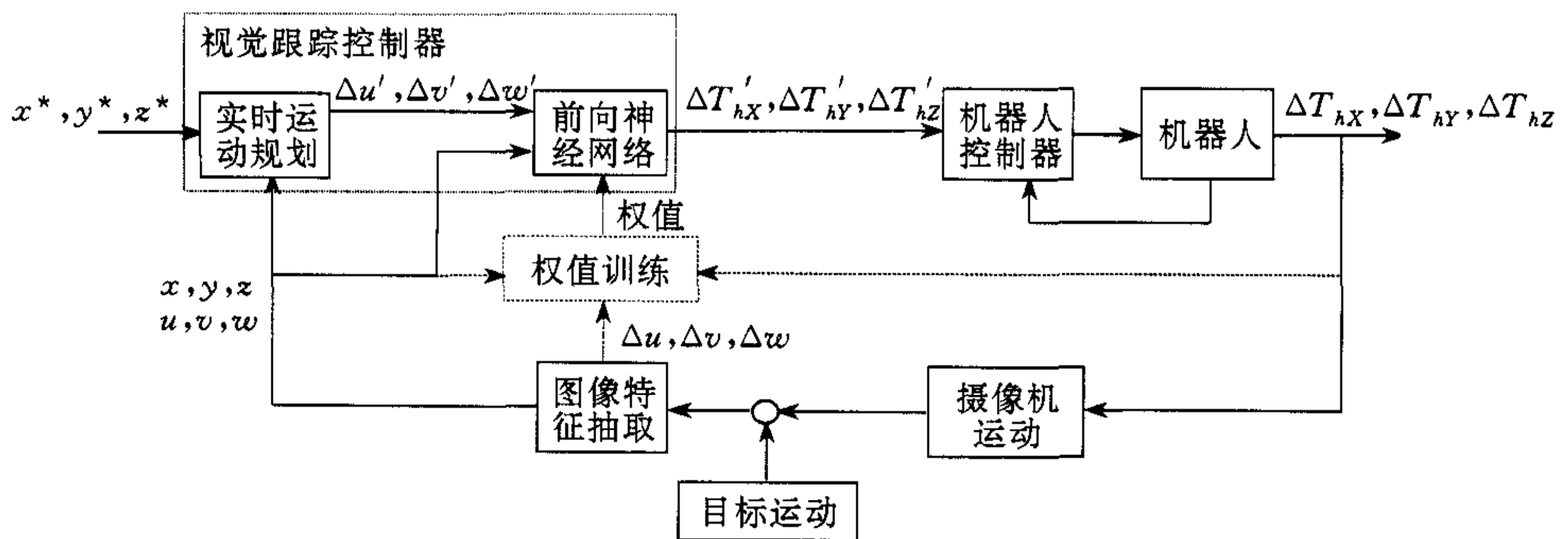


图 2 机器人3D 视觉跟踪控制结构图

4.1 实时运动规划

运动规划在三维图像特征空间中进行,其目的是根据图像特征序列(包括目标在图像空间中投影点的位置 $[x(k), y(k), z(k)]^T$ 、速度 $[u(k), v(k), w(k)]^T$ 及其期望位置 $[x^*, y^*, z^*]^T$),来实时地规划未来时刻目标投影点的运动状态(包括位置 $[x'(k+1), y'(k+1), z'(k+1)]^T$ 、速度 $[u'(k+1), v'(k+1), w'(k+1)]^T$ 、加速度 $[\Delta u'(k+1), \Delta v'(k+1), \Delta w'(k+1)]^T$),从而保证手爪快速、无静差地跟踪运动目标,克服因目标运动速度变化、视觉映射模型不准确所引起的跟踪误差.为描述方便,将速度、加速度表示为矢量形式 $s(k) = [u(k), v(k), w(k)]^T$, $s'(k) = [u'(k), v'(k), w'(k)]^T$, $\Delta s'(k) = [\Delta u'(k), \Delta v'(k), \Delta w'(k)]^T$. 设 k 时刻图像特征空间中目标投影点与期望位置 $[x^*, y^*, z^*]^T$ 间的距离为 $d(k)$,为使手爪尽快靠近目标, $s'(k+1)$ 的方向应始终指向 $[x^*, y^*, z^*]^T$,如图 3 所示,其幅值 $|s'(k+1)|$ 按以下规则变化

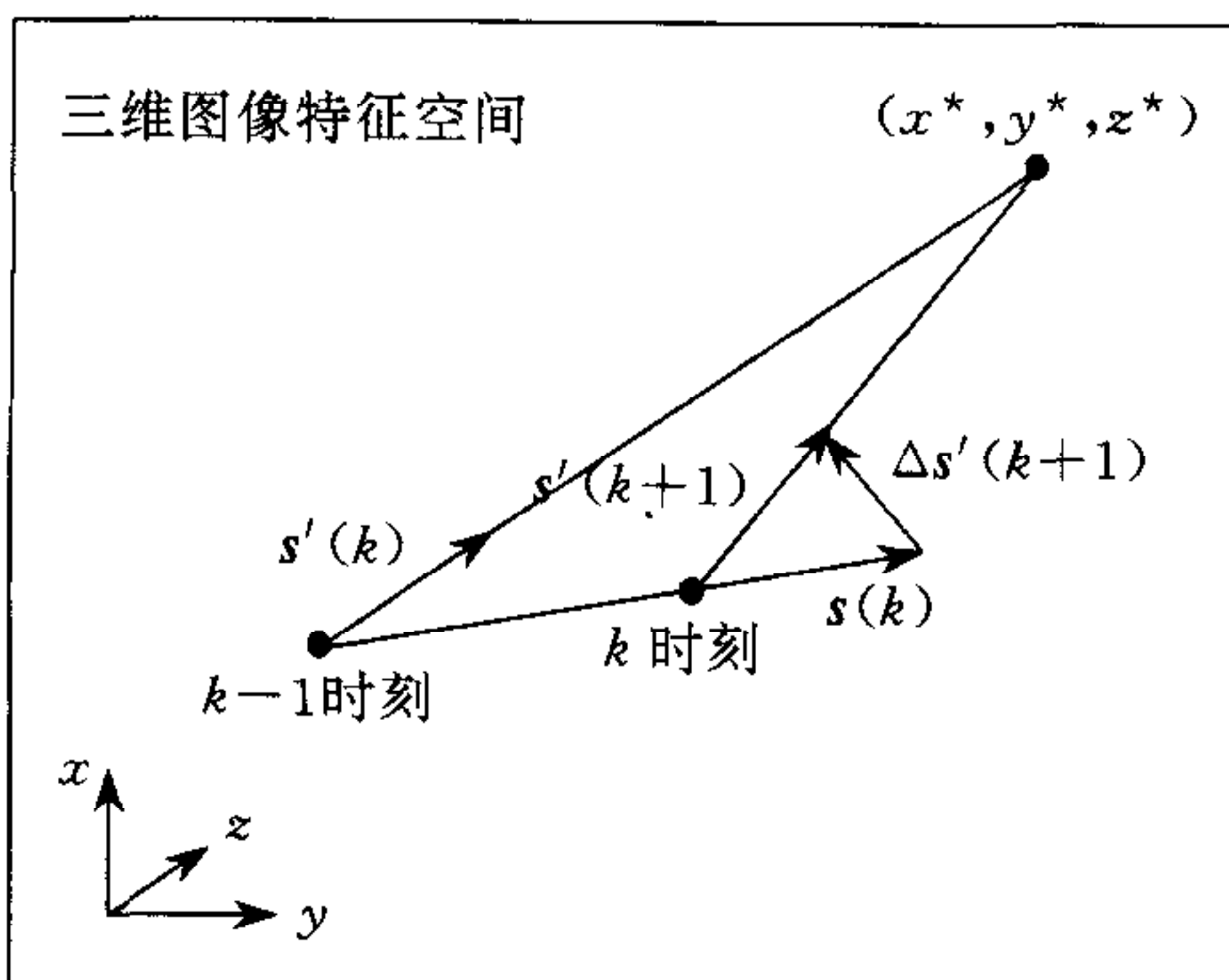


图 3 运动规划示意图

1) 当 $d(k) \geq d_1$ 时,逐渐增加 $|s'(k+1)|$,直到手爪达到允许的最大速度运行;

2) 当 $d_1 > d(k) \geq d_2$ 时,逐渐减小 $|s'(k+1)|$,直到手爪减速至一较低的速度运行;

3) 当 $d(k) < d_2$ 时,根据图像特征空间中当前位置与期望位置的偏差,通过 PI 控制律计算出

下一时刻目标运动速度大小,从而做到无差跟踪,计算公式为

$$\begin{bmatrix} u'(k+1) \\ v'(k+1) \\ w'(k+1) \end{bmatrix} = \begin{bmatrix} u(k) \\ v(k) \\ w(k) \end{bmatrix} + c_1 \begin{bmatrix} x(k) - x^* \\ y(k) - y^* \\ z(k) - z^* \end{bmatrix} - c_2 \begin{bmatrix} x(k-1) - x^* \\ y(k-1) - y^* \\ z(k-1) - z^* \end{bmatrix} \quad (8)$$

以上跟踪策略中,阈值 $d_1 > d_2 > 0$, c_1 和 c_2 分别是 PI 控制器的比例、积分系数. 运动规划器的输出是目标特征加速度规划值 $[\Delta u'(k+1), \Delta v'(k+1), \Delta w'(k+1)]^T$, 计算公式为

$$\begin{bmatrix} \Delta u'(k+1) \\ \Delta v'(k+1) \\ \Delta w'(k+1) \end{bmatrix} = \begin{bmatrix} u'(k+1) \\ v'(k+1) \\ w'(k+1) \end{bmatrix} - \begin{bmatrix} u(k) \\ v(k) \\ w(k) \end{bmatrix} \quad (9)$$

4.2 神经网络映射器

神经网络的作用是将图像特征空间中运动规划结果映射到机器人所在的三维笛卡尔空间中去,作为运动指令下达给机器人伺服控制器. 根据式(7)所建立的逆映射关系模型,构建 9 输入、3 输出三层前向网络,其学习律采用带动量项加速因子的反播(BP)算法.

神经网络权值采用离线训练方式获得,并用于在线控制. 训练样本按以下方式获得:让目标在工作空间内以不同的起始点和初速度作匀速直线运动,其轨迹尽可能布满整个工作范围(仿真中取为 $X_0 \in [-0.2, 0.2]$, $Y_0 \in [-0.2, 0.2]$, $Z_0 \in [-0.2, 0.2]$. 一种特殊情况是让目标速度为零,也即目标静止,这更易于工程实现). 手爪作三维随机平动,亦即其加速度为随机变量. 记录图像特征空间中目标投影的位置 $[x(k), y(k), z(k)]^T$ 、速度 $[u(k), v(k), w(k)]^T$ 、加速度 $[\Delta u(k), \Delta v(k), \Delta w(k)]^T$ 序列和机器人空间中手爪加速度数据 $[\Delta T_{hx}, \Delta T_{hy}, \Delta T_{hz}]^T$ 作为训练样本(手爪加速度 $[\Delta T_{hx}, \Delta T_{hy}, \Delta T_{hz}]^T$ 由其位置坐标 $[X_h(k), Y_h(k), Z_h(k)]^T$ 经二阶差分求取). 值得指出的是,在训练样本采集过程中,尽管图像量化噪声等因素将导致单个样本的求取存在一定的误差,但由于神经网络是由大量的样本训练得到,类似于求取平均值的过程,因此仍可以比较准确地逼近逆视觉映射模型.

ANN 用于控制时,除输入当前时刻图像特征空间中的位置 $[x(k), y(k), z(k)]^T$ 、速度 $[u(k), v(k), w(k)]^T$ 以外,还要输入由运动规划得到的图像特征加速度 $[\Delta u'(k+1), \Delta v'(k+1), \Delta w'(k+1)]^T$. ANN 的输出就是下达给机器人伺服控制器的加速度指令 $[\Delta T'_{hx}(k+1), \Delta T'_{hy}(k+1), \Delta T'_{hz}(k+1)]^T$.

5 仿真结果

利用 Matlab5.1 软件包神经网络工具箱对以上控制方案进行了仿真. 图 4 和 5 分别为手爪跟踪匀速运动和螺旋线运动目标的坐标变化曲线. ANN 权值由 3 061 个训练样本通过 2 万次迭代得到,ANN 隐层节点数为 40. 摄像机坐标系相对与手爪坐标系的偏转角为 30 度,俯仰角为 20 度. 控制器主要参数为 $c_1 = 40$, $c_2 = 0.5$, $d_1 = 0.003$, $d_2 = 0.001$. 由图可见,稳态位置误差为零,跟踪效果极佳. 另外,由于神经网络具有良好的泛化能力,即使目标点在工作平面上的运动轨迹超出训练范围,手爪仍能有效地跟踪(见图 4). 图 6 为采用同样权值条件下,手眼关系发生变化后的跟踪情况(偏转角为 0 度,俯仰角为 -20 度). 由图可见,尽管未针对变化后的手眼关系重新训练神经网络,仍取得了良好的控制效果,说明了基于 ANN 的控制器具有极强的环境适应能力. 图 7 是在 ANN 训练不充分情况下的跟踪曲线. 训练样

本数仅为 638 个,迭代次数为 1 000,隐节点数为 16. 考虑到摄像机图像分辨率受限,为模拟由此带来的量化噪声,该仿真中在目标投影点位置坐标中加入了高斯白噪声,其均方差为 0.36 个像素宽度. 由图 7 可见,尽管此处采用的 ANN 模型极其粗糙,仍然能对目标进行有效跟踪.

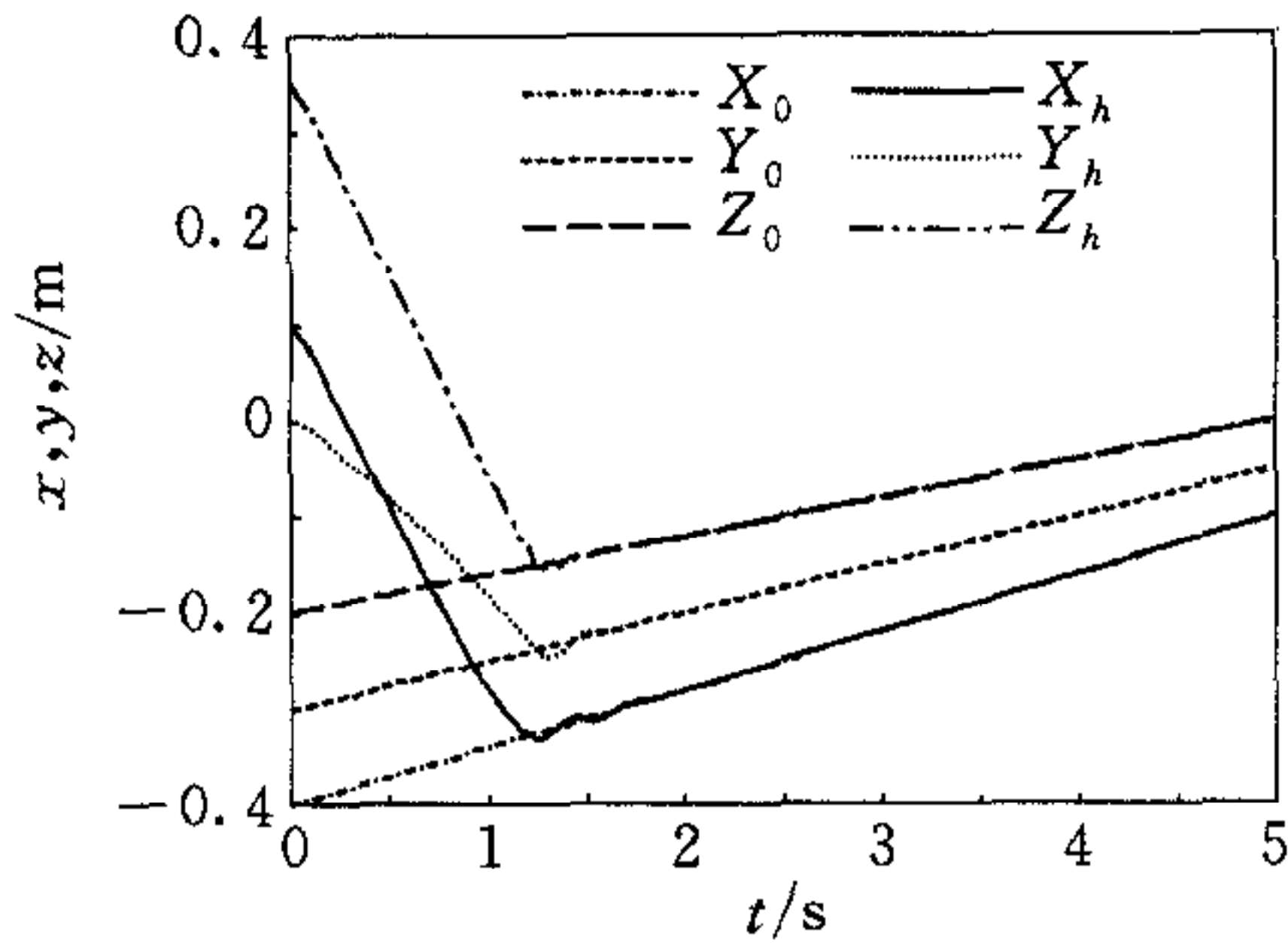


图 4 跟踪匀速运动目标

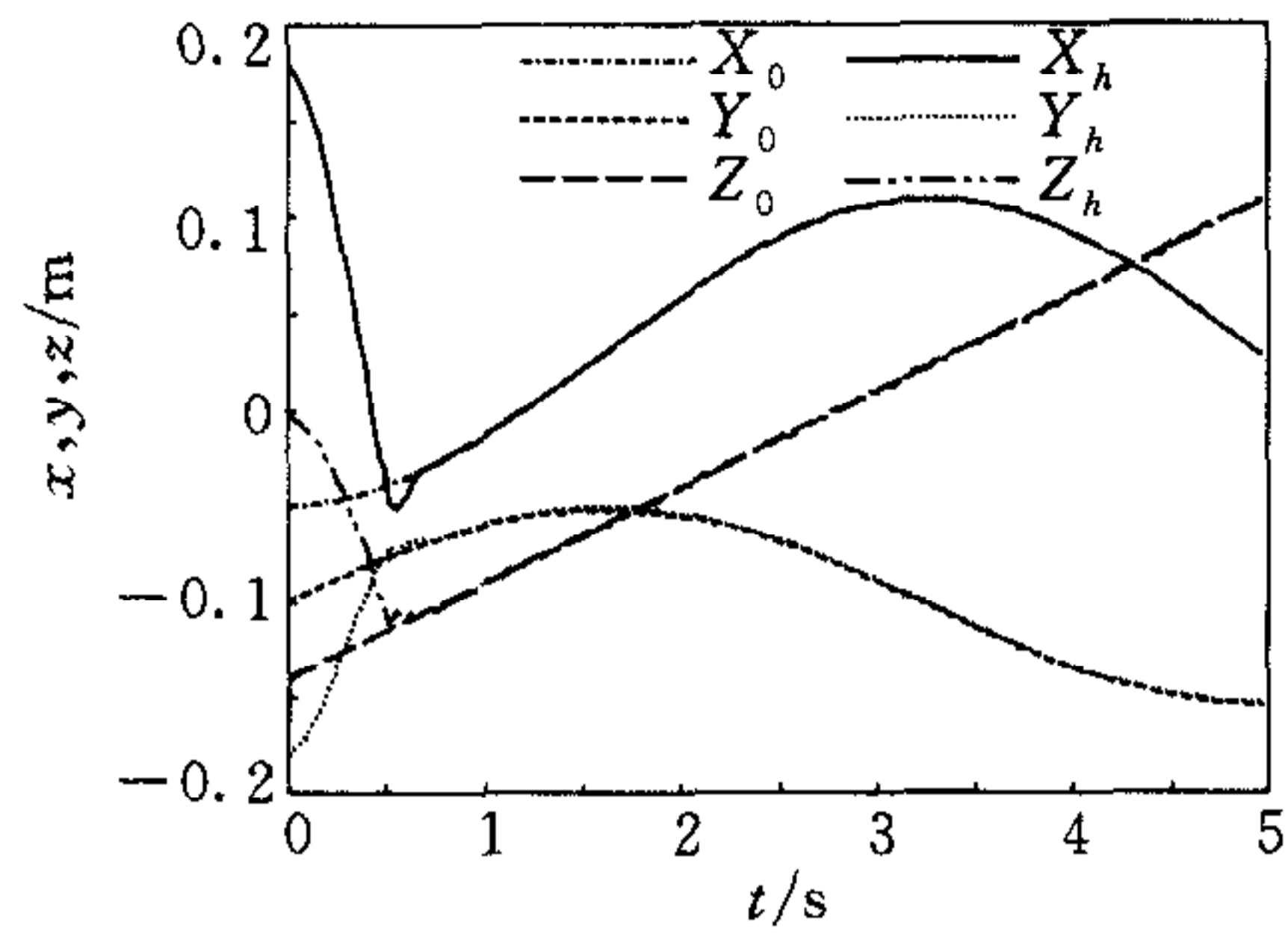


图 5 跟踪螺旋线运动目标

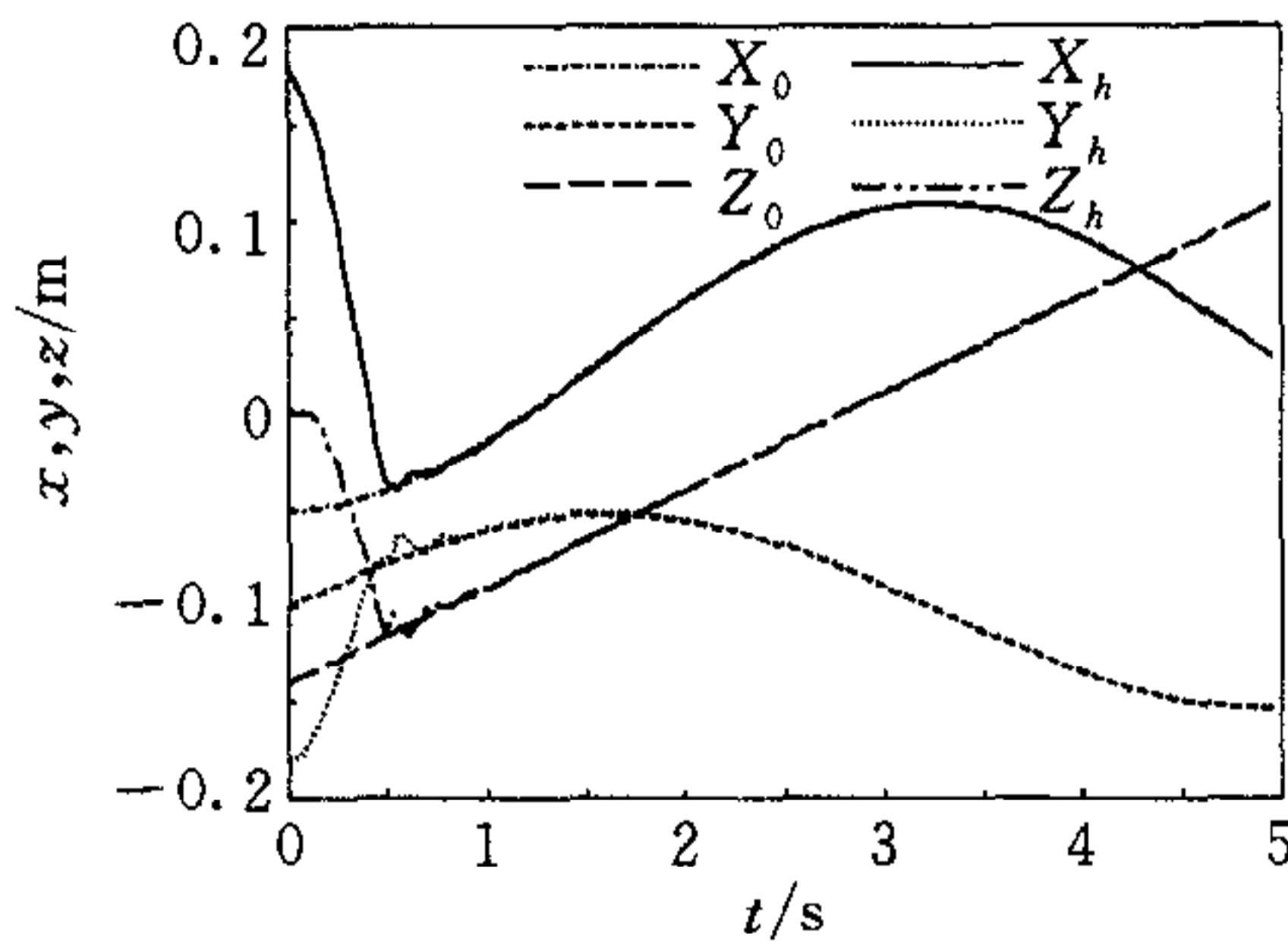


图 6 手眼关系变化时跟踪曲线

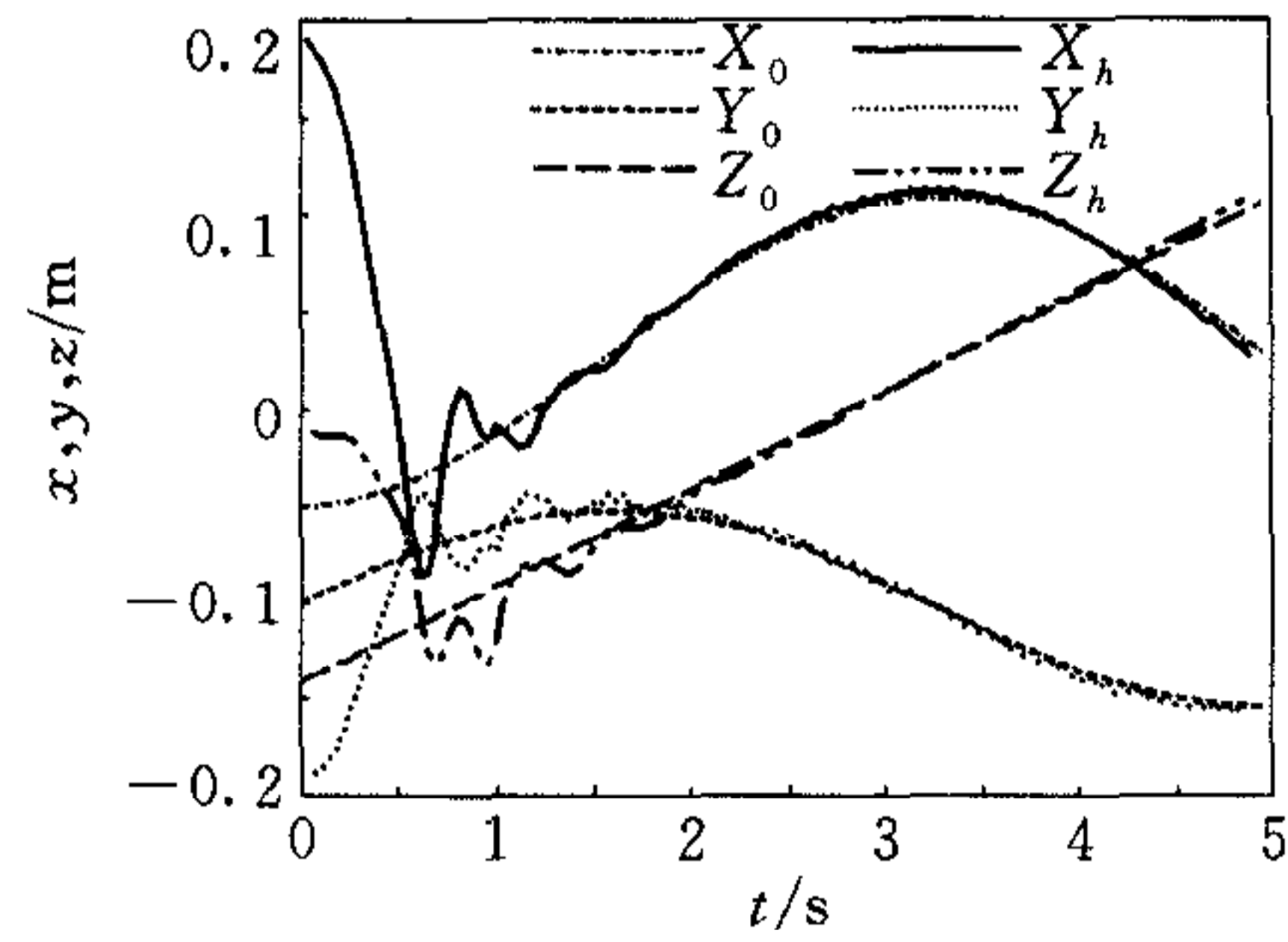


图 7 训练不充分时跟踪曲线

6 结束语

本文在手眼关系及摄像机模型完全未知的情况下,针对眼在手上机器人三维视觉跟踪问题提出了一种新的非线性视觉映射关系模型,并在此基础上设计了基于人工神经网络的跟踪控制方案,避免了对大量未知参数进行辨识或标定,算法简单,易于工程实现. 仿真结果表明,该方案跟踪效果好,具有极强的环境适应性和容错能力,在整个工作空间内均能适用.

参 考 文 献

- 1 Hutchinson S, Hager D, Corke P. A tutorial on visual servo control. *IEEE Trans. Robot. Autom.*, 1996, **RA-12**(5):651~670
- 2 Sanderson A C, Weiss L E, Neuman C P. Dynamic sensor-based control of robots with visual feedback. *IEEE Trans. Robot. Autom.*, 1987, **RA-3**(5):404~417
- 3 Yoshimi B H, Allen P K. Alignment using an uncalibrated camera system. *IEEE Trans. Robot. Autom.*, 1995, **RA-11**(4):516~521
- 4 苏剑波,李耀通. 手眼关系的实时估计及平面机器人运动跟踪. 见:第二届全球华人智能控制与智能自动化大会论文集(CWC-ICIA'97). 西安:西安交通大学出版社,1997. 73~76
- 5 Hespanha J, Dodds Z, Hager G D *et al.* What can be done with an uncalibrated stereo systems? In: Proc. 1998

IEEE Inter. Conf. on Robot. Autom., Belgium; Leuven, 1998. 1366~1372

6 Sutanto H, Sharma R, Varma V. Image based autodocking without calibration. In: Proc. 1997 IEEE Inter. Conf. on Robot. Autom., New Mexico; Albuquerque, 1997. 974~979

7 Scheering C, Kersting B. Uncalibrated hand-eye coordination with a redundant camera system. In: Proc. 1998 IEEE Inter. Conf. on Robot. Autom., Belgium; Leuven, 1998. 2953~2958

8 Papanikolopoulos N P, Khosla P K. Adaptive robotic visual tracking: theory and experiments. *IEEE Trans. Robot. Autom.*, 1993, **RA-38**(3):429~445

9 Smith C E, Brandt S A, Papanikolopoulos N P. Eye-in-hand robotic tasks in uncalibrated environments. *IEEE Trans. Robot. Autom.*, 1997, **RA-13**(6):903~914

10 Hashimoto H, Kubota T, Sato M *et al.* Visual control of robotic manipulator based on neural networks. *IEEE Trans. Industrial Electronics*, 1992, **IE-39**(6):490~496

潘且鲁 1990年至2000年先后就读于上海交大电力学院、电子信息学院,获得学士、硕士和博士学位.研究领域为机器人视觉、数字信号处理、移动通信等.

苏剑波 分别于1989年、1992年、1995年在上海交大、中科院自动化所和东南大学获得学士、硕士和博士学位,1997年博士后出站,现任上海交大自动化系教授.研究领域为机器人视觉与控制.

席裕庚 1984年于德国慕尼黑工业大学获博士学位,现为上海交通大学教授,博士生导师.研究方向为预测控制理论、电信网络优化规划、机器人路径规划与控制等.

中国自动化学会 2002 年一般专题学术活动计划

项目名称	主要内容	时间	地点	联系人
第 21 届中国控制会议	线性系统;最优控制;鲁棒控制;预测控制;随机控制;自适应控制;模型降阶;优化方法;电力系统;交通系统;社会经济系统;生物环境系统等	3 季度	杭州	北京中关村中国科学院系统科学研究所 刘智敏 电话:62532161 邮编:100080
全国(第七届)自动化应用技术学术交流会	交流和推广自动化新技术、新理论及其在工业应用中的科技成果和论文	9 月	新疆	同上
全国(第八届)工业控制系统应用学术交流会	交流各种工业控制系统应用、研制、开发的新成果、新技术、新方法	7 月	山西五台山	山西太原理工大学山西省自动化学会 张忠怀 电话:(0351)6041057 邮编:030024
第十一届全国电气自动化与电控系统学术年会	直流传动技术、交流传动技术、应用电力电子技术的变流器、计算机与 PLC 控制技术、伺服电机步进电机控制系统、控制理论在工程中的应用、工业应用等	3 季度	西安	天津市河东区津塘路 174 号电气自动化专业委员会秘书处 电话:(022)25962354 邮编:300180 E-mail: caaea@public. tpt. tj. cn
全国第五届 Java 技术及应用学术会议	国内外 Java 技术及应用发展综述, Java 技术、平台开发工具及产品开发及应用; Java 技术在企业信息化、电子商务及远程教学、医疗的求解方案; Java 技术及产品在工业、交通、金融等领域与行业应用经验等	3 季度	待定	信息产业部 6 所(北京 927 信箱) 熊强、罗安 电话:82923953 82923958 邮编:100083
第十三届中国过程控制年会	过程建模、仿真与辨识技术、先进控制、优化控制、模式识别与图像处理、动态系统故障诊断与容错技术、DCS 系统、流程工业 CIMS 等	11 月	澳门	广州华南理工大学 朱学锋 电话:(020)87111580 邮编:510641 E-mail: xfzhu@scut. edu. cn