

文章编号: 1002-0446(2003)05-0461-04

一种基于神经网络的视觉定位系统*

刘艳梅¹, 董再励², 孙茂相¹, 毛艳娥¹

(1. 沈阳工业大学信息科学与工程学院, 辽宁 沈阳 110023;

2. 中国科学院沈阳自动化研究所机器人学开放研究实验室, 辽宁 沈阳 110016)

摘要: 本文提出了一种基于人工神经网络的多线阵相机系统标定与 3D 定位方法, 并应用于基于多线阵相机构成的视觉空间定位系统. 该视觉定位系统利用了线阵相机的快速性与高分辨率的特点, 非平行空间投影面相交定位的基本原理, 实现了这种结构下快速、高精度空间定位. 实验表明, 人工神经网络的定位方法简化了多线阵视觉定位系统标定与定位计算的复杂性, 在定位精度上达到了良好效果. 为机器人位置反馈控制提供了有效的技术途径.

关键词: 立体视觉系统; 相机标定; 位置测量; 人工神经网络

中图分类号: TP24

文献标识码: B

A VISION-LOCATION SYSTEM BASED ON NEURAL NETWORK

LIU Yan mei¹, DONG Zai li², SUN Mao-xiang¹, MAO Yan e¹

(1. School of Information Science and Engineering, Shenyang University of Technology;

2. Robotics Laboratory, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China)

Abstract: A new space location vision system based on one dimensional cameras is presented. High speed and high precision are the features of this position measurement system. In this paper, the measurement methods based on artificial neural networks and the structure of the system are introduced. The experimental results are presented to show the efficiency of this system on position measurement of space targets.

Keywords: stereovision system; camera calibration; position measurement; artificial neural network

1 引言(Introduction)

机器视觉^[1,2]技术是近几十年来发展的一门新兴技术. 机器视觉是研究使机器具有类似于生物视觉部分功能的一门新兴学科. 因此在工业过程质量检测、监控、非接触位置与几何测量、视觉伺服控制、机器人自主行为规划、机器人遥操作、透视 3D 成像、反求工程、遥现与遥视等研究与工业领域有着不可替代作用和巨大的发展前景. 高精度空间定位与动态轨迹测量是机器视觉技术的一个重要应用领域. 近年来, 高精度空间定位与轨迹跟踪技术在先进制造、医疗服务、科学研究等领域的需求日趋广泛. 因此以 CCD 相机为基本传感器的视觉空间定位模式, 如单目、双目、结构光等不同结构组成各种位置测量装置与检测方法, 一直是机器视觉的研究前沿课题.

典型的视觉测量与定位方法主要采用面阵 CCD

相机. CCD 相机的最大特点是可以在线直接获取景物的数字化图象, 信息丰富, 但由于其易受环境光噪声以及图象处理复杂性的影响, 实际应用受到很大限制. 因此, 当视觉系统作为位置检测与反馈装置使用时, 希望能尽量简化信息的复杂性, 以利于高速、准确地提取位置信息. 近年来国外研制的多线阵 CCD 构成位姿测量系统(如医疗辅助导航系统^[9,10]), 就是一种成功应用的范例. 线阵相机可以使用柱面镜头和线阵光电器件, 具有分辨率高、采样速率高等特点, 易于实现快速高精度检测的要求. 因此具有更明显的实用意义. 另一种可以实现快速、高精度视觉位置测量方式是采用 PSD(position sensitive detector)相机. PSD 是 20 世纪 80 年代发展起来的一种新型半导体位置敏感器件. 它通过输出模拟信号, 可以准确地反映出照射在其敏感表面上的光点位置. 与线阵 CCD 类似, 它也具有高分辨率, 高灵敏

度, 以及更高的响应速度, 而且成本低. 这些特点使其在生产 and 科研中具有良好的应用开发前景和实际意义. 因此成为近年来位置精密检测方法 with 装备的研究热点.

视觉位置检测系统的一个重要实现技术就是光电系统的标定问题, 标定效果的好坏和效率, 直接影响视觉测量系统的应用性能, 如精度、环境适应性等. 由于神经网络具有能充分逼近复杂非线性关系的能力等特点, 可以应用于视觉位置检测系统的参数标定.

2 多线阵相机的定位原理 (Location Principle of Linear Cameras)

2.1 使用柱面镜头成像的光学定位机理

使用线阵相机做空间位置测量时, 一般采用柱面镜头采集图像信号, 故成像面上的任一点在镜头曲率方向上对应为一个投影面, 即线阵镜头上的任一投影点 u 都可以定义其相应的一个空间平面. 见图 1.

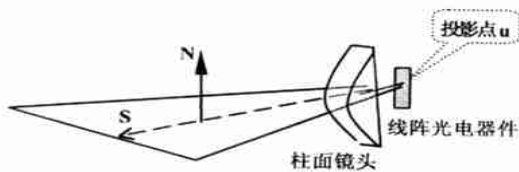


图 1 成像点与投影面在柱面镜头作用下的成像关系

Fig. 1 The imaging relationship between imaging point and projection plan produced by cylindrcal lens

图 1 中, S 表示一个空间平面, 其位置由 S 的法向 N 和点 u 决定, 在平面 S 上的任一点, 其在柱面镜头上的投影点都落在 u 上, 法向 N 和点 u 存在着一一对应关系. 给出 u 值, S 就确定了. 由几何定理可知, 若至少存在空间这样 3 个非平行的平面, 就可以联立计算出其唯一交点 P . 因此, 多线阵相机空间定位原理, 就是通过建立 3 个这样定义在参考系上的平面, 计算其交点, 从而实现点 P 的空间定位. 其一般系统结构见图 2. 由图 2 所示的 3 个线阵相机获取目标点 P 的投影 (其投影位置由相机的输出数据 u 给出), 应用相应的算法可得到目标点的空间位置. 定位系统的标定和位置计算一般采用解析方法, 既通过建立位置关系表达式, 运用光电系统标定模型参数, 完成 3D 定位.

2.2 动量—自适应学习网络模型下的系统参数标定

摄像机标定是计算机视觉中的一个重要内容. 摄像机参数标定过程就是确定摄像机的内部几何与

光电参数和摄像机坐标系与世界坐标系之间的相对位置和方向的过程. 目前的摄像机标定算法一般可分为两种类型: 其一是用最小二乘法求解线性方程. 简单快速, 但没有考虑镜头畸变, 准确性欠佳. 其二是考虑了畸变参数, 引入非线性优化, 但方法较繁, 速度慢, 且初值选择和噪声比较敏感, 而且非线性搜索并不能保证参数收敛到全局最优解. 这两种方法都不能适应快速、高精度视觉检测的需求.

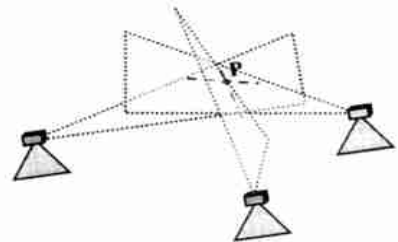


图 2 定位原理图

Fig. 2 The location fundamental theory graphic

本文尝试采用人工神经网络来解决这一问题. 由于网络由相连的非线性单元组成, 因此具有学习非线性成像过程的潜在能力. 这种方法不要求预先建立成像系统的精确的模型结构, 也不需要具体求出摄像机的内部参数和外部参数, 而是利用人工神经网络通过直接学习图像信息与三维世界坐标信息 (x, y, z), 得出二者之间的关系. 实验结果表明所提方法有效可行.

BP 模型是神经网络的重要模型之一, 应用比较广泛, 在分类、预测、故障诊断和参数检测中具有广泛的应用, 本文采用的 BP 网络结构加图 3 所示.

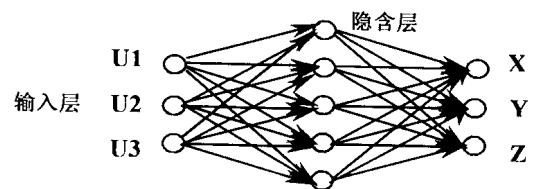


图 3 BP 网络结构

Fig. 3 BP network structure

第一层为输入层, 第二层为隐含层, 第三层为输出层, 隐含层神经元模型采用 S 型激活函数, 输出层采用线性激活函数. 网络输入为目标点在 3 个相机中的图像信息即投影点 u , 输出为目标点的三维世界坐标. 对于输入信号, 首先向前传播到隐节点, 经过作用函数后, 再把隐节点的输出信息传播到输出节点, 最后输出节点通过输出函数给出输出结果. 这个算法的学习过程, 由正向传播和反向传播组成. 在正向传播过程中, 输入信号从输入层经隐单元层逐层

处理,并传向输出层,每一层神经元的状态只影响下一层神经元的状态。如果输出层不能得到期望的输出,则转入反向传播,将误差信号沿原来的连接通路返回,通过修改各层神经元的权值,使得误差信号最小。过程为:1)将网络初始化,即随机设置各层神经元的权值。2)给定网络的输入向量和目标输出。3)计算网络隐层、输出层各单元的输出。4)计算目标值与实际输出的偏差。5)调整各层神经元的权值。6)转至过程3),重新开始学习。直至目标值与实际输出的偏差满足训练精度的要求。

BP 算法理论具有依据可靠、推导过程严谨、精度较高、通用性较好等优点,但是 BP 模型存在几个缺点:1)收敛速度缓慢;2)容易陷入局部极小值;3)难以确定随层节点的数目。针对其存在的缺点出现了许多改进算法,如本文所讨论的动量—自适应学习率 BP 算法即是一种。动量—自适应学习率 BP 算法采用了动量法和学习率自适应调整的策略,从而提高了学习速度并增加了算法的可靠性。动量法降低了网络对于误差曲面局部细节的敏感性,有效地抑制网络陷入局部极小。动量法权值调整算法为

$$w(k+1) = w(k) + a[(1-mc)D(k) + mcD(k-1)] \quad (1)$$

其中 $w(k)$ 为第 k 步时的权值, $D(k) = -dE/dw(k)$ 为第 k 步时的负梯度, a 为学习率, $a > 0$, mc 为动量因子, $0 < mc < 1$, 加入的动量项相当于阻尼项,它减小了学习过程的振荡趋势,从而改善了收敛性。另外,自适应调整学习率的权值调整算法为

$$\begin{aligned} w(k+1) &= w(k) + a(k)D(k) \\ a(k) &= 2ra(k-1) \\ r &= \text{sign}[D(k)D(k-1)] \end{aligned} \quad (2)$$

当连续两次迭代其梯度方向相同时,表示下降太慢,这时可使学习率加倍;当连续两次迭代其梯度方向相反时,表示下降过头,这时可使学习率减半。

将以上两种方法结合起来,就得到动量—自适应学习率 BP 神经网络的权值修正算法

$$w(k+1) = w(k) + a(k)[(1-mc)D(k) + mcD(k-1)] \quad (3)$$

网络法标定过程:首先从实验系统中获得训练神经网络用的样本对,即目标点在 3 个像机中的投影点 u 和目标点的三维世界坐标,由于 S 激活函数的灵敏区有限,因此有必要对样本数据进行归一化处理,应用归一化公式保证输入数据在 $(-1, 1)$ 范

围;确定网络初始权值;应用上述动量—自适应学习率 BP 算法训练网络权值,通过调整网络中可调参数(隐层数、隐层节点数、学习速率、动量因子等特性参数)进行网络的训练,并采用均方根误差(RMS)计算训练精度,该网络最佳的结构隐层选用 5 个神经元。则像机的标定参数隐含在网络权值当中。训练好的网络就可以用于定位测试。

3 系统实现与实验(System realization & experiments)

3.1 系统构成

本文介绍的基于多线阵相机的视觉定位系统(如图 4 所示)由一个安装在三维移动平台上的 LED 红外调制信号源 P, 3 个配置柱面镜头的线阵 PSD 相机相互正交构成测量系统, PSD 信号处理电路, 多路 A/D 数据采集板和数据处理计算机组成。其分辨率主要取决于 A/D 的分辨率。

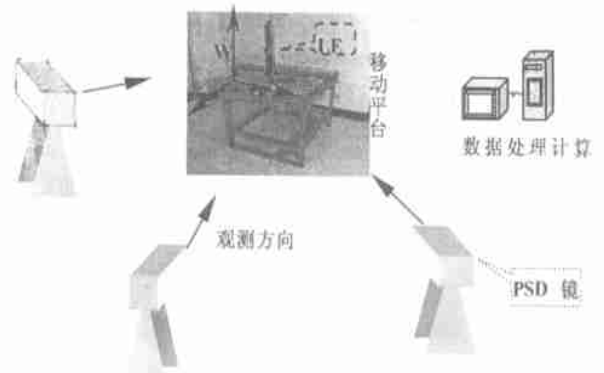


图 4 实验中所用的空间定位

Fig. 4 The space location system applied in experiment

3.2 PSD 的位置测量原理

PSD 是一种新型的光电位置信息探测器,属于分割型器件,也可称为坐标电池。当指示光源照射到光敏面上时,由于平行于结面的横向电场作用,使光生载流子形成向两端电极流动的电流 x_1 和 x_2 , 它们之和等于总电流。当投影点移动时,两个输出电流信号也将随之发生变化,其结构如图 5 所示。

线阵 PSD 上投影点的位置 X 与所产生的输出电流经 x_1 和 x_2 的关系式(4)为。

$$\frac{x_2 - x_1}{x_2 + x_1} = \frac{2x}{L} \quad (4)$$

线阵 PSD 产生的两路连续模拟量,可经信号处理后由式(4)计算出投影点的位置信号 x 即 u 。

由于 PSD 器件的输出量是连续的模拟信号,可以使光电系统达到很高的分辨率。且输出单一,系统

可以具有较高的响应速度. 另一优点是可以对目标光信号进行调制, 从而使信号采集系统可以具有很好的抗干扰能力和环境适应性. 实验系统的线阵 PSD 接收信号光源的波长范围在 320~ 1100nm, 峰值波长为 920nm. 因此本文介绍的实验系统采用波长为 900nm 的 LED 作为观测目标.

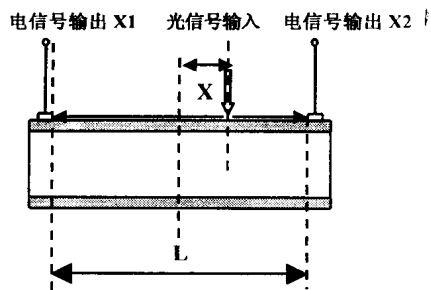


图 5 PSD 结构原理图

Fig. 5 PSD structure graphic

3.3 实验结果

实验在实验室环境下进行, 3 个线阵 PSD 摄像机固定分布如图 4, 距移动平台为 10m~ 15m, 镜头视角为 22°. LED 目标点在空间范围为 80 × 80 × 35cm 三维移动平台(定义为 W)上移动. A/D 分辨率为 4096. 视场空间分辨率为 0.2mm. 步骤如下:

(1) 系统标定: 在 W 中移动 LED 目标, 采集 50 组样本数据对(目标点在 3 个像机中的投影 u 和其在世界坐标系 W 中的三维坐标), 对网络进行权值训练. 并把结果存入计算机.

(2) 定位计算: 目标点在三维移动平台上的随机移动, 得到 3 个摄像机的投影作为网络的输入, 利用训练好的系统, 算出输出即目标点的空间坐标. 对目标点进行静态定位, 实测 36 个点, 定位结果如图 6.

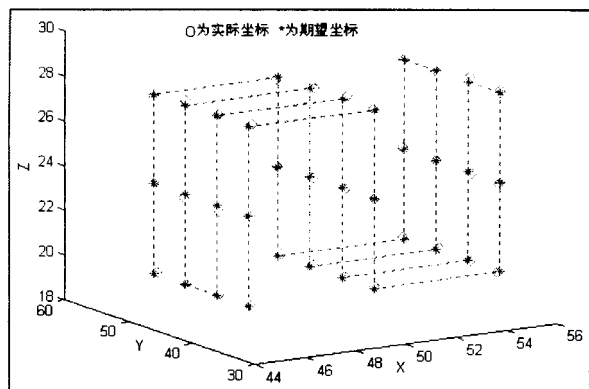


图 6 定位结果

Fig. 6 The location result

实验定位误差(单位 cm): X 方向最大偏差为 0.183, 最小为 0.047, 均方根误差为 0.1338, Y 方向最大偏差为 0.110, 最小为 0.029, 均方根误差为 0.0806, Z 方向最大偏差为 0.112, 最小为 0.013, 均方根误差为 0.08, 合成位置偏差平均值为 0.0982. 图中 ° 为测量值, * 为实际值.

4 结论 (Conclusion)

上述研究表明, 利用多线阵 PSD 像机对目标点进行空间位置测量是可行的. 由于信号采集处理基本是硬件实现, 计算量很小, 因此采样速度可以很高. 由于系统分辨率可以很高, 因而系统的测量精度主要与相机的观测角、信号处理电路的稳定性和线性度有关. 当系统稳定性和线性度可以控制时, 系统的位置测量精度可以满足许多应用领域的高精度位置检测要求. 当观测物体安装有多多个目标点时, 可以通过分时采样技术计算其姿态. 因而可以完成空间目标物体的高精度位姿测量和轨迹跟踪. 应用神经网络技术可以化简视觉定位系统的标定和位置计算, 从而减少了位置检测系统使用复杂性. 利用动量——自适应学习率 BP 算法可提高学习速度并增加了算法的可靠性. 实验验证了此类定位方法与实现技术解决空间定位问题上的有效性、合理性和可实现性.

参考文献 (References)

- [1] 游素亚, 徐光佑. 立体视觉研究的现状与进展[J]. 中国图象图形学报, 1997, 2(1): 17- 23.
- [2] 马颂德, 张正友. 计算机视觉[M]. 科学出版社, 1998.
- [3] Polanski A, Wojciechowski K, Borek A. Stereo calibration by planar grid lines[A]. CAIP Proceedings[C]. Springer Verlag 1995.
- [4] Wei G Q, Ma S D. Implicit and Explicit Camera Calibration: Theory and experiments[J]. IEEE Trans PAMI, 1994, 16: 469- 480.
- [5] 董再励等. 一种基于视觉的移动机器人定位系统[J]. 中国图形图象学报, 2000, 5(8): 688- 692.
- [6] 孙增圻. 智能控制理论与技术[M]. 北京: 清华大学出版社, 1997.
- [7] 赵清杰, 孙增圻. 摄像机神经网络标定技术[J]. 控制与决策, 2002, 17(3).
- [8] 从爽. 面向工具箱的神经网络理论与应用[M]. 合肥: 中国科技大学出版社, 1998.
- [9] 周果宏, 罗述谦. 图像导引的手术导航系统[J]. 世界医疗器械, 1997, 3(3): 27- 31.
- [10] Scot A Tebo, Donald A Leopold, Donglin M Long, S James Zinreich. An optical 3D digitizer for frameless stereotactic surgery[J]. IEEE Computer Graphics and Applications, 1996. 55- 64.

作者简介:

刘艳梅(1974), 女, 博士生. 研究领域: 自动化, 信号处理.
董再励(1952), 男. 研究领域: 模式识别, 传感器系统.