

基于 BP 神经网络的足球机器人摄像机标定

宗晓萍, 石圣羽, 田 华

ZONG Xiao-ping, SHI Sheng-yu, TIAN Hua

河北大学 电子信息工程学院, 河北 保定 071002

College of Electronic and Informational Engineering, Hebei University, Baoding, Hebei 071002, China

E-mail: shishengyu2002@yahoo.com.cn

ZONG Xiao-ping, SHI Sheng-yu, TIAN Hua. Camera calibration for soccer robots based on BP neural network. Computer Engineering and Applications, 2008, 44(8): 204-206.

Abstract: Accurate camera calibration is required for achieving precise visual measurements. Artificial neural networks are used to learn the relationships between the image information and the 2D information. It neither requires an accurate mathematical model nor needs any prior knowledge about the parameters. The experiment results show that the proposed camera calibration based on neural network has higher precision. It is valid for the camera calibration of the soccer robots.

Key words: BP neural network; soccer robots; camera calibration

摘 要: 摄像机标定是精密视觉测量的基础。利用人工神经网络直接学习图像信息与二维平面信息之间的对应关系, 不需要确定摄像机具体的内部参数和外部参数, 也无需知道有关摄像机模型或参数的先验知识。通过实验表明基于神经网络的标定方法与传统的线性标定方法相比具有较高的标定精度和较强的标定实时性, 适用于足球机器人的摄像机标定。

关键词: BP 神经网络; 足球机器人; 摄像机标定

文章编号: 1002-8331(2008)08-0204-03 **文献标识码:** A **中图分类号:** TP138

1 引言

足球机器人是一个涉及机器视觉、传感器、无线通信、自动控制等多个研究领域的复杂系统, 其中视觉系统是机器人的主要信息来源, 也可以说是整个系统的基础, 而摄像机标定则是视觉系统的重要组成部分。摄像机标定的目的在于建立像平面上的二维坐标与其相应的被测点的三维空间坐标之间的对应关系。目前, 摄像机的标定方法可分为传统标定方法和摄像机自标定方法^[1]。传统的摄像机标定是利用标准参照物与其图像的对应约束关系, 来确定摄像机模型的内外参数。此方法由于其概念清楚、易于理解而被广泛采用, 缺点是标定过程复杂, 特别是在需要考虑镜头畸变时。摄像机的自标定因其不需要标定参照物, 直接利用从图像序列中得到的约束关系来计算摄像机模型的参数, 使实时地、在线地标定摄像机参数成为可能。但是有些摄像机自标定并不能保证解的唯一性和稳定性, 同时, 自标定其精确性和鲁棒性较差。

工程中, 无论应用上述哪一种标定, 首先要选择合适的摄像机模型。线性模型虽然简单, 但不能准确地描述成像几何关系, 影响到标定的精度。当采用非线性模型对摄像机进行标定时, 需要使用非线性优化计算, 但引入过多的非线性参数往往

不仅不能提高精度, 反而引起解的不稳定。事实上, 每一个摄像机的光学特性都不尽相同, 另外, 不同的环境尚有许多不确定因素存在, 没有任何一个摄像机模型可以尽善尽美地描述一个机器视觉系统。

人工神经网络是生物神经网络的一种模拟和近似, 是由一定数量的神经元及各神经元之间的广泛连接所构成的并行、分布式的信息处理网络, 其特点如下: (1) 神经网络主要是根据所提供的样本数据, 通过学习和训练找出输入和输出之间的内在联系, 从而求得问题的解答, 而不是依靠对问题的先验知识和规则, 因而它具有很好的学习能力和自适应能力; (2) 神经网络具有一定的泛化功能, 能够处理那些未经训练过的或含有噪声的不完全的数据, 找出和输出之间的内在联系; (3) 神经网络具有极强的非线性映射能力, 一个三层的前向网络即可逼近任意的非线性连续函数, 因此特别适用于解决高度非线性的函数关系问题。因此, 应用人工神经网络的学习能力、自适应能力以及非线性函数的逼近能力, 建立空间点与图像点的映射关系, 完成摄像机的精确标定。这样即符合人类视觉对信息的处理过程, 又与摄像机的标定机制相近^[2] (通过已知样本数据获取摄像机参数, 并由此来确定未知数据)。

基金项目: 河北省自然科学基金(the Natural Science Foundation of Hebei Province of China under Grant No.A2006000941); 河北省科技厅 2007 科技攻关计划 (No.072135142); 河北省教育厅 2006 科研计划 (No.2006408)。

作者简介: 宗晓萍 (1964-), 女, 博士, 副教授, 主要研究领域为混合动态系统、机器人视觉伺服; 石圣羽 (1983-), 男, 硕士研究生, 主要研究领域为计算机视觉、数据采集及计算机接口技术; 田华 (1972-), 女, 硕士, 实验师, 主要研究领域为光信息处理、嵌入式控制系统。

收稿日期: 2007-07-04 **修回日期:** 2007-10-09

2 基于 BP 神经网络的摄像机标定

2.1 足球机器人视觉系统

足球机器人系统可以分为 4 个子系统: 视觉子系统、决策子系统、无线通信子系统和机器人子系统。其中视觉子系统包括球场上方的摄像头、计算机主机内的图像采集卡和图像识别处理程序。机器人足球比赛的赛场长 1.5 m、宽 1.3 m。比赛场地画有中线、中圈和门区。每队由三个边长不超过 7.5 cm 的立方形的遥控小车(机器人)组成。

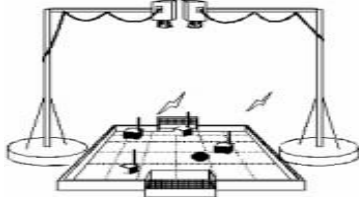


图 1 足球机器人系统

对于足球机器人视觉系统, 其输入为实际的比赛场景图像, 输出为机器人和球的位置、速度等。比赛场景图像由机器人比赛场地中央上方约 2 m 处的摄像头获得, 并由视觉程序实时处理。视觉程序对图像中的颜色块进行识别, 以得到其实际的空间坐标, 并根据几何关系进一步计算得到机器人的位置、速度等其他参数。摄像机标定就是要建立成像面上的二维投影点的坐标与其相应的被测点的二维坐标之间的对应关系。

2.2 BP 神经网络

神经网络的类型很多, 需根据所要完成任务的性质和要求来合适地选择网络的类型。实现映射和函数逼近是前向网络的共同特点。理论证明, 若多层前馈神经网络的隐层节点可以自由设定, 则它可以实现任意精度逼近任何连续函数^[3]。误差反向传播(Back Propagation)网络(简称 BP 网络)是最简单、应用最为广泛的一种前馈网络。

BP 网络由输入层、隐层和输出层节点组成, 其结构如图 2 所示。BP 网络的学习过程是由模式的正向传播和误差反向传播所组成。在正向传播过程中, 输入信息经隐含单元逐层处理并传向输出层。如果输出层不能得到所期望的输出, 则转入反向传播过程, 将实际值与网络输出之间的误差沿原来的联接通路返回, 通过修改各层神经元的联接权重使误差减少, 然后再转入正向传播过程, 如此反复计算, 直至误差小于设定值。

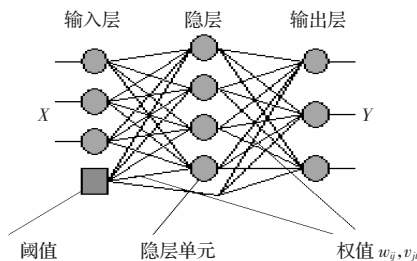


图 2 BP 网络结构

2.3 BP 神经网络的学习算法

常规的 BP 算法本质上为梯度下降法, 其学习速度较慢。另外它是一种局部搜索的优化方法, 但它要解决的问题为求解复杂非线性函数的全局极值, 因此, 算法很有可能陷入局部极值, 使训练失败。基于以上考虑, 为了克服传统 BP 算法的不足, 本文采用了加入动量项或阻尼项的 LM(Levenberg-Marquardt)BP 算法^[4]。经实验表明使用 LM 算法可使网络获得较快的收敛速度。

3 实验设计

3.1 标定样本数据的采集

由于足球机器人仅工作在二维空间, 此时, 深度信息是确定的, 摄像机的标定简化为平面视觉标定问题。本标定实验采用单目 CCD 摄像机及 DH-CG410 图像采集卡, 测量对象为 130 cm×90 cm 的机器人足球场地。标定样本图如图 3 所示。通过对其进行边缘及角点检测, 得到了 320 组有效样本点, 抽出 10 组作为检测样本, 其余为训练样本。



图 3 标定样本图

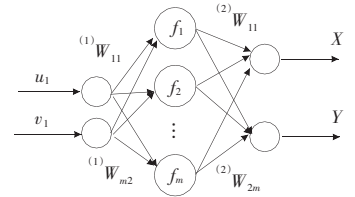


图 4 单目标定的 BP 网络结构

3.2 基于 BP 神经网络的摄像机标定

用于摄像机标定的 BP 网络的结构如图 4 所示。根据摄像机标定的特点, 网络输入为标定样本点的图像像素坐标(u, v), 输出为其对应的世界坐标(X, Y), 即 BP 网络为 2- n -2 结构。 n 为网络的隐层节点数。

目前对 BP 网络隐层及节点数的选择, 尚无理论指导, 应从网络的复杂程度、收敛速度及网络的泛化性等多方面考虑^[4]。在训练过程中, 多次调节 n 值, 当 n 取 12 时, 收敛速度较快并且可获得较好的泛化能力。其网络训练过程如下:

$net=newff(\minmax(p), [2, 12, 2], \{ 'logsig', 'tansig', 'purelin' \}, 'trainlm');$ %建立结构为 2-12-2 的 BP 网络, 其中 p 为输入的训练样本。

$net.trainParam.show=50$

$net.trainParam.epochs=400$

$net.trainParam.goal=0.0001$

$[net, tr]=train(net, p, t)$

在训练到达 103 步时目标函数达到允许误差值 10^{-4} 。网络的训练误差曲线如图 5 所示。应用 10 组检测样本检测其标定的精度, 其结果如表 1 所示。

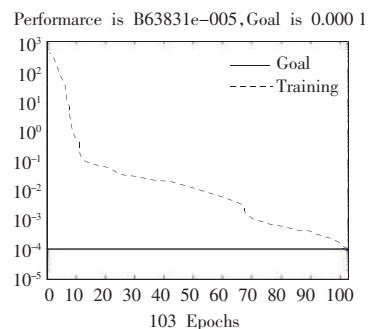


图 5 神经网络的训练误差曲线

\overline{MSE} 为整个测试点集的平均均方根误差, $\overline{MSE} = \frac{MSE_1 + \dots + MSE_{10}}{10}$, 其中测试点的均方根误差为 $MSE = \sqrt{\frac{\Delta x^2 + \Delta y^2}{2}}$ 。

应用基于神经网络的标定方法, 其平均均方根误差为 0.286 2 cm, 相对误差为 0.002 6。对其样本数据加入方差为

表1 基于BP神经网络的摄像机标定精度检测

输入/pixel		期望输出值/cm		网络实际输出值/cm		均方误差/cm
u	v	x_w	y_w	x'_w	y'_w	MSE
125	140	30	15	29.343 3	15.118 3	
125	204	15	15	14.910 0	15.003 3	
189	232	7.5	0	7.816 4	0.245 3	
222	135	30	-7.5	30.440 0	-7.606 9	
255	268	0	-15	-0.224 7	-14.880 7	0.286 2
288	235	7.5	-22.5	7.623 8	-22.651 0	
95	74	45	22.5	44.869 0	22.823 0	
189	333	-15	0	-14.698 4	0.360 1	
253	76	45	-15	-14.860 9	-15.260 4	
255	333	-15	-15	-14.790 5	-15.102 7	

0.02 的高斯白噪声,重复以上实验,得整个测试点集的平均均方根误差为 0.319 8 cm。由于摄像机图像分辨率低,存在量化误差,以及实验条件和样本数据的限制,是造成误差的主要原因。

3.3 直接线性法标定

应用传统的直接线性法标定^[5]:

$$u = \frac{x_w l_0 + y_w l_1 + l_2}{x_w l_3 + y_w l_4 + l_5}, v = \frac{x_w l_6 + y_w l_7 + l_8}{x_w l_3 + y_w l_4 + l_5}$$

这里 (x_w, y_w) 为物体的世界坐标, (u, v) 为相应的像点的图像坐标。不失一般性,可以令 $l_5=1$ 。如果已知 $N(N>3)$ 个标准参照物的世界坐标 (x_w, y_w) 及其对应的图像坐标 (u, v) ,代入上式可得得到一个八元一次方程组,解此方程组即可得到八个参数取值,结果如表2所示。

应用传统的直接线性法标定,其平均均方根误差为 1.245 1 cm,相对误差为 0.011 3。对其样本数据加入方差为 0.02 的高斯白噪声后,其平均均方根误差为 1.312 6 cm。

4 结论

应用BP神经网络描述二维空间中的物点与其对应的图像点之间的复杂的非线性映射关系,不需要预先建立摄像机的成像模型,在包容由于摄像机镜头的畸变等各种因素所造成的复

表2 传统方法摄像机标定检测结果

输入/pixel		期望值/cm		实际计划值/cm		均方误差/cm
u	v	x_w	y_w	x'_w	y'_w	MSE
189	232	7.5	0	8.629 9	-0.455 3	
222	135	30	-7.5	30.231 8	-8.499 3	
255	268	0	-15	-0.469 0	-16.415 0	
125	140	30	15	29.449 9	15.005 6	
125	204	15	15	15.465 4	15.148 3	1.245 1
288	235	7.5	-22.5	6.950 2	-25.911 7	
95	74	45	22.5	43.469 8	21.783 7	
189	333	-15	0	-14.946 5	-0.560 4	
253	76	45	-15	43.149 9	-16.034 8	
255	333	-15	-15	-16.082 0	-17.729 6	

杂的非线性关系的同时,能够有效地避免在复杂的非线性逆运算中可能出现的“病态”现象,且避免了大量的复杂运算,从而具有较好的标定实时性;同时它能够有效地克服光学系统的模型误差、CCD光敏元非均匀等因素所造成的测量误差,从而保证了较高的测量精度;另外,由于神经网络本身具有较强的自学习和自适应能力,使得神经网络标定具有较强的抗噪能力。总之基于神经网络的摄像机标定,为足球机器人的摄像机标定提供了一种新的有效方法。

参考文献:

- [1] 邱茂林,马颂德,李毅.计算机视觉中摄像机标定综述[J].自动化学报,2000,26(1):43-55.
- [2] Zong Xiao-ping, Xu Yan, Hao Lei. Camera calibration based on the radial basis function network with tunable nodes[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006:5708-5712.
- [3] Funahashi K. On the approximate realization of continuous mapping by neural network[J]. Neural Network, 1989, 2:183-192.
- [4] Hagan M T, Demuth H B, Beale M H. 神经网络设计[M]. 戴葵,译.北京:机械工业出版社,2002.
- [5] 邹凤娇,苏显渝,李美菊.基于共面点的摄像机线性标定法[J].光电工程,2005(4):70-74.

(上接174页)

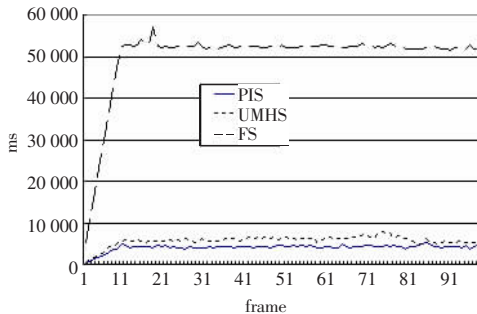


图9 搜索半径为32时的ME时间对比

4 结论

PIS搜索算法将四边形搜索和六边形搜索结合起来,明显地降低了计算量,加快了编码速度,同时图像质量得到了很好

的保存,是一种可行的算法。若在算法中加入类似UMHS算法的早期终止策略,编码效率会进一步提高,这也将是算法进一步完善的主要突破点。

参考文献:

- [1] 李炜,周兵,李波.运动矢量场自适应搜索算法[J].计算机学报,2003(2).
- [2] 杨智勇,郭宝龙,丁贵广.基于预测的方向性菱形快速运动估计算法[J].计算机工程,2005(19).
- [3] Chen Zhi-bo, Zhou Peng, He Yun. Fast integer pel and fractional pel motion estimation for JVT, JVT-F017[S]. Awaji Island JP, 2002.
- [4] Cheung Chun-Ho, Po Lai-Man. Novel cross-diamond-hexagonal search algorithm for fast block motion estimation[J]. IEEE Trans Multimedia, 2005, 7(1).