

基于改进 Harris 角点提取的摄像机神经网络标定技术^{*}

崔岸 袁智 王丹 张海鹏

(吉林大学汽车动态模拟国家重点实验室, 长春 130025)

【摘要】 以前馈型 BP 神经网络进行双目立体视觉系统的摄像机标定研究, 基于 Harris 角点提取, 提出了增加约束的改进方法, 从而提高网络训练样本集的数据精度; 探讨了神经网络的欠泛化、过泛化问题, 综合运用归一化、提前终止等多种策略, 进一步提高网络泛化能力, 并与经典标定方法进行对比。试验结果表明, 该方法能够获得较高的摄像机标定精度。

关键词: 摄像机标定 神经网络 Harris 角点 泛化

中图分类号: TP391.4 **文献标识码:** A

Neural Network Technique on Camera Calibration Based on Improved Harris Corner Extraction Method

Cui An Yuan Zhi Wang Dan Zhang Haipeng

(State Key Laboratory of Automotive Dynamic Simulation, Jilin University, Changchun 130025, China)

Abstract

Camera calibration in a binocular stereo vision system was studied based on BP neural network techniques. A neural network was built to investigate the relationship between the image coordinates and the space coordinates. To improve the accuracy of training data, a corner extraction algorithm based on Harris algorithm was modified by increasing constraints. The generation ability of neural networks was discussed, and was further improved by synthesizing many strategies such as the regularization and stopped training strategies. At last, compared with the traditional calibration method, the test result shows that this method is available and can get a higher precision for binocular camera calibration.

Key words Camera calibration, Neural network, Harris corner, Generality

引言

摄像机标定是计算机视觉测量中的重要环节, 它建立起二维图像坐标到三维空间坐标的映射关系。由于图像的成像过程受到镜头畸变和光介质等因素的影响, 使得映射呈现一种非线性关系。传统的摄像机标定技术多采用针孔成像模型, 线性标定或两步法等, 但线性标定精度较差, 两步法只考虑径向畸变, 当切向畸变较大时不实用, 因而如何有效合理确定非线性畸变校正模型的参数一直是研究的难点。由于神经网络技术具有很强的非线性逼近能

力, 可以直接学习图像信息和三维信息之间的关系, 不需确定摄像机具体的内部参数和外部参数, 也不需要了解有关模型和参数的先验知识, 而是将所有非线性因素都包含在网络当中, 作为一种隐式的立体标定技术而显示出它的优越性。

建立性能优越的神经网络需要高精度的训练样本数据, 一些学者提出了很好的角点提取算法^[1~3], 但是提取出的结果有时并不令人满意。本文鉴于原有 Harris 角点提取算法存在多义性或漏检的缺欠, 对其进行进一步约束改进, 从而既消除了多义性又使角点提取精度显著提高。同时, 神经网络

的泛化能力是评价其性能的一个重要指标^[4],提高泛化能力一直是研究者关注的问题,但总体上说仍然是一个没有完全解决的问题。本文结合多种策略并进行对比分析使其达到一个较好的泛化效果。

1 标定数据的准备

本文以双目立体视觉为系统模型,用左右两个摄像机分别拍摄图片,摄像机型号为 DH-HV1302UM-T 的 CMOS,分辨率为 1 280 × 1 024,镜头为 12.5 ~ 75 mm F1.8 的 CCTV lens。

以 BP 网络建立神经网络时,为了获得神经网络所需的样本数据集,采用 20 × 20 棋盘格式立体标靶。对角点的空间坐标进行预先设置来作为输出数据样本集,利用 Harris 角点提取算法得到角点的二维坐标作为输入数据样本集,如图 1 所示。

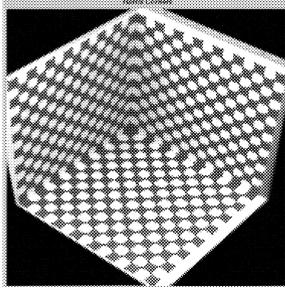


图 1 Harris 角点提取
Fig.1 Harris corner extraction

由于打印机分辨率及照明的原因使得 Harris 检测出的角点效果不佳,有些角点出现两个图像对应的情况,如图 2 为图 1 的局部放大图。为了获得准确单一的图像角点坐标,本文对 Harris 角点提取结果进行了改进。

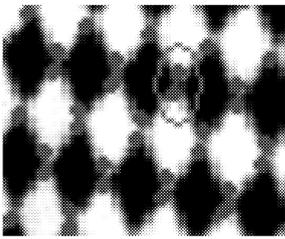


图 2 Harris 角点提取局部放大图
Fig.2 Details of Harris corner extraction

2 Harris 角点提取算法及其改进

2.1 Harris 角点提取算法

角点是反映图像信息的重要特征,如图 3 中的 *a* 所表示的部分,在水平、垂直方向的灰度变化都很大。图 3 中的 *b* 表示边缘,其仅在水平或仅在垂直方向的灰度变化量大;图 3 中的 *c* 所表示的平坦区域,其水平或垂直灰度变化量都较小。

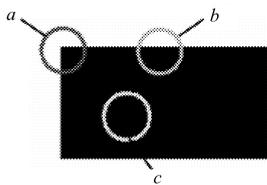


图 3 图像信息特征图
Fig.3 Image characteristics

Harris 角点提取算法步骤^[2]:

(1)利用水平、竖直差分算子对图像每个像素进行滤波以求得水平与竖直的灰度梯度 I_x 和 I_y ,进而得到矩阵

$$m = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (1)$$

(2)对 m 中 4 个元素进行高斯滤波,得到新的 m 。离散二维零均值高斯函数为

$$G = \exp \left\{ -\frac{x^2 + y^2}{2\sigma^2} \right\} \quad (2)$$

(3)利用 m 计算对应于每个像素的角点量为

$$cim = \frac{I_x^2 I_y^2 - (I_x I_y)^2}{I_x^2 + I_y^2} \quad (3)$$

(4)在角点量矩阵 cim 中,同时满足 cim 大于阈值和 cim 是某邻域内的局部极大值的点被认为是角点。

2.2 改进 Harris 角点提取

针对上述算法,改进的 Harris 角点提取方法增加了第(5)步:

步骤(5)对步骤(4)返回的角点坐标进行检验,如果 $|u_i - u_j| \leq \epsilon, |v_i - v_j| \leq \epsilon$ 则把它们合并成一个点, ϵ 为合并阈值,它的选取与图像的拍摄角度有关。通常情况下 ϵ 的最大值应小于相邻棋盘点所含像素最小值的一半, ϵ 最小值应大于歧义点纵横坐标差值的最大者,本文选取 $\epsilon = 10$ 。进行平均插值运算

$$\begin{cases} u_m = \frac{u_i + u_j}{2} \\ v_n = \frac{v_i + v_j}{2} \end{cases} \quad (4)$$

式中 u_i, u_j ——步骤(4)提取出的角点行坐标
 v_i, v_j ——步骤(4)提取出的角点纵坐标
 u_m, v_n ——改进后角点的新坐标

改进后的 Harris 角点提取效果如图 4 所示。

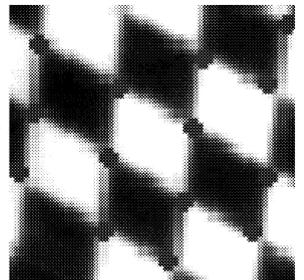


图 4 改进后 Harris 角点提取局部放大图
Fig.4 Details of modified Harris corner extraction

3 摄像机神经网络标定设计

在标定图像与标靶中共选取了 588 个坐标点对,为了使神经网络能够合理、顺利地建立,将其中 294 对(分布均匀,下同)作为训练样本数据, 147

对作为确认样本数据,另 147 对作为神经网络的测试样本数据。

3.1 初步建立的神经网络标定

本文采用 3 层 BP 网络对摄像机进行标定^[5],

直接运用获取的角点坐标和空间坐标进行神经网络的创建与训练。由于样本数较多,网络采用批处理的训练方式。选取隐层的传递函数为 logsig 函数,输出层为 purelin 函数,性能指标均方误差 MSE 设

为 0.05,由输出结果维数可知输出层神经元个数为 3,分别取隐层神经元个数为 5、10、15、20、25、30、35 对 BP 及其改进型网络进行训练,其训练均方误差、次数及时间如表 1 所示。

表 1 网络训练结果

Tab.1 Training result of direct-coordinate network

神经元数	traingd			traingdm			traingdx			trainrp			trainscg			trainlm			
	均方误差	迭代次数	时间/s	均方误差	迭代次数	时间/s	均方误差	迭代次数	时间/s	均方误差	迭代次数	时间/s	均方误差	迭代次数	时间/s	均方误差	迭代次数	时间/s	
5	9 221.38	130	1.687	661.035	527	4.406	-	-	-	5 592.61	5 000	23.782	5861.84	3 701	25.563	53.796	6	200	13.859
10	7 537.58	767	7.123	7 070.04	796	6.875	-	-	-	2 444.79	3 701	20.672	5 805.31	3 101	29.468	4.527	3	801	58.969
15	6 285.62	2 978	27.156	6 135.4	3 338	30.609	-	-	-	3 515.66	3 901	27.469	4 583.10	3 901	48.063	4.723	5	701	60.031
20	6 399.52	2 571	25.484	6 529.2	1 040	8.359	-	-	-	2 343.77	3 401	26.656	5 287.46	2 201	31.860	218.207	0	501	41.219
25	6 501.01	396	4.922	5 900.48	541	4.812	-	-	-	2 309.28	3 701	34.172	6 107.00	2 401	40.375	4.747	6	501	58.951
30	6 002.10	5 000	68.704	5 405.25	3 544	36.797	-	-	-	2 385.54	3 001	31.531	6 746.79	1 601	31.906	82.841	0	401	64.438
35	6 117.70	2 455	35.843	6 490.32	1 843	18.625	-	-	-	2 880.57	2 401	42.562	6 725.62	1 801	40.516	21.138	9	301	58.406

从表 1 中可以看出,当隐层神经元个数为 10~15,训练函数为 trainlm 时,迭代次数最少,最后的性能参数 MSE 较小,同时也可以看出随着隐层神经元个数的增多,其每次迭代所需要的时间增加,所以选隐层神经元个数为 10 较好;当训练函数为 traingdx 时性能参数 MSE 出现了一些波动,因而在表中没有将其结果列出;其余训练结果 MSE 都十分大。总体来看,包括训练函数 trainlm 在内的所有网络训练结果都不能达到理想的性能参数,不能很好地记住样本,属于欠泛化问题,而在有些问题中,能记住样本但是不能逼近非样本数据,出现了过泛化问题^[6]。

3.2 摄像机神经网络标定

基于网络泛化能力的考虑,初始建立的神经网络中,样本数据采用了均匀分布的的选取;神经网络初始参数的设置参考了先验知识,但结果依然不理想。为了进一步提高神经网络的泛化能力,针对所

建立的网络模型,采取了如下改进措施。

(1)在建立神经网络之前,先对数据进行归一化处理,使其数据在 0~1 之间,以控制网络的有效复杂程度,加快收敛速度。具体步骤为:首先,找出样本数据集 P 的最大值 P_{\max} 、最小值 P_{\min} ;然后对样本数据集 P 进行归一化处理,其算法公式为: $P'_i = (P_i - P_{\min}) / (P_{\max} - P_{\min})$,其中 i 为样本序列号。

(2)在网络训练过程中。一般来说学习误差会随着迭代次数的增加而降低,当出现性能函数长时间变化范围很小时,基于泛化精度及训练时间的考虑,采用提前终止策略。

隐层的传递函数设为 tansig,输出层的传递函数为 logsig,性能指标 MSE 设为 1×10^{-5} 。取隐层神经元个数为 5、8、10、15、18、20、25 对 BP 及其改进算法进行训练,其训练均方误差、次数及时间如表 2 所示。

表 2 改进的网络训练结果

Tab.2 Training result of indirect-coordinate network

神经元数	traingd			traingdm			traingdx			trainrp			trainscg			trainlm		
	均方误差	迭代次数	时间/s	均方误差	迭代次数	时间/s												
5	0.004 67	5 000	22.547	0.004 24	5 000	25.547	0.003 08	4 151	20.641	0.002 95	3 901	19.984	0.002 87	2 201	19.696	0.000 41	1 451	26.828
8	0.003 28	5 000	30.875	0.003 17	4 451	26.203	0.002 09	3 151	17.610	0.002 60	2 651	16.703	0.000 42	3 751	37.157	0.000 26	1 101	30.984
10	0.001 83	5 000	31.203	0.002 54	4 501	28.266	0.001 79	3 151	21.796	0.001 57	3 351	21.796	0.000 77	3 851	43.140	0.000 22	901	32.375
15	0.001 84	5 000	35.578	0.001 88	3 601	27.609	0.000 99	5 000	39.375	0.001 48	2 901	22.532	0.000 29	2 701	36.078	1.00×10^{-5}	3 186	173.270
18	0.001 22	5 000	41.734	0.002 75	3 001	26.234	0.001 37	3 701	30.703	0.001 25	3 351	28.975	0.000 31	2 701	42.032	1.00×10^{-5}	4 527	292.890
20	0.001 29	5 000	45.297	0.001 69	3 201	29.985	0.001 30	3 901	33.297	0.001 09	3 601	31.515	0.000 27	3 151	52.203	1.00×10^{-5}	682	54.109
25	0.001 40	5 000	48.485	0.002 81	3 201	32.391	0.001 63	2 901	28.063	0.000 94	2 801	28.562	0.000 28	2 001	39.625	1.00×10^{-5}	366	40.640

从表 2 中可以看出采用归一化后的网络训练次数大幅减少,所需时间大量缩短,所能达到的性能参数 MSE 也较小,其中当隐层神经元个数为 25,训练函数为 trainlm 时,效果最佳,其网络训练过程如图 5 所示。

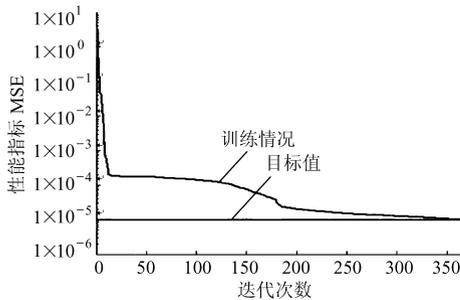


图 5 $n = 25$ 时的 LMBP 网络训练图

Fig.5 Training map of LMBP network when $n = 25$

选用该网络,并保存权阈值对确认样本数据进行试验,经反归一化后得出结果显示坐标误差小于 1 mm,比直接运用样本数据坐标建立的网络有了非常大的改善。故而在样本空间较大的情况下采取归一化和提前终止相结合的方法能够很好地改善神经网络的泛化能力和训练所耗时间。

4 摄像机标定试验

为了进一步验证该研究方法的可行性,本文还

将其标定结果与经典的摄像机标定方法^[7-8]作了对比,这里选用自由平面标定法作为试验参照法,运用左右摄像机分别拍摄 20 幅图像对,如图 6 所示(只列出 2 对)。

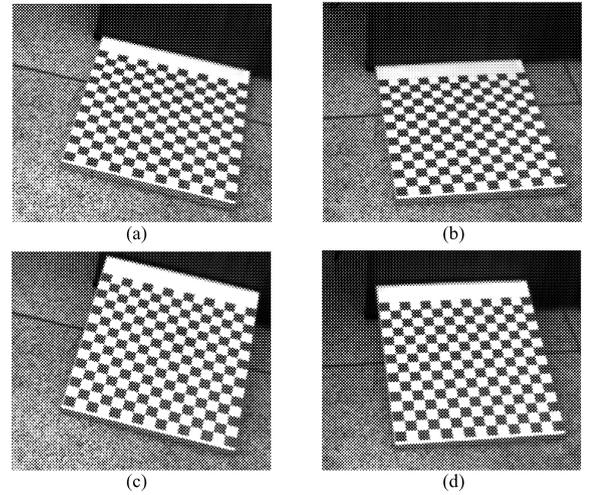


图 6 平面标定法图像对

Fig.6 Image pairs of free-plane calibration

- (a) 左(第 1 对) (b) 右(第 1 对)
- (c) 左(第 2 对) (d) 右(第 2 对)

经过标定优化计算后,分别得出 2 个摄像机的内外部参数,再对它们进行融合,得到的最后参数如表 3 所示。

表 3 自由平面标定法内外部参数结果

Tab.3 Result of interior and exterior parameters using free-plane calibration

	左摄像机	右摄像机
焦距/mm	$[2\ 437.493\ 19\ 2\ 438.082\ 83] \pm [2.374\ 63\ 2.405\ 16]$	$[2\ 433.256\ 05\ 2\ 432.687\ 81] \pm [2.4610\ 3\ 2.555\ 41]$
图像中心	$[175.725\ 05\ 223.188\ 85] \pm [5.199\ 31\ 4.199\ 20]$	$[211.148\ 80\ 276.411\ 72] \pm [5.471\ 95\ 4.534\ 63]$
歪斜系数	$[0.000\ 00] \pm [0.000\ 00]$	$[0.000\ 00] \pm [0.000\ 00]$
扭曲系数	$[-0.257\ 83\ 0.744\ 42\ -0.001\ 82\ 0.002\ 82\ 0.000\ 00] \pm [0.012\ 63\ 0.109\ 39\ 0.000\ 32\ 0.000\ 70\ 0.000\ 00]$	$[-0.224\ 43\ 0.706\ 47\ -0.000\ 04\ 0.000\ 99\ 0.000\ 00] \pm [0.012\ 66\ 0.133\ 03\ 0.000\ 33\ 0.000\ 70\ 0.000\ 00]$
旋转矢量	$[0.004\ 22\ -0.218\ 22\ -0.285\ 44] \pm [0.002\ 35\ 0.003\ 02\ 0.000\ 33]$	
平移矢量	$[267.668\ 52\ 0.627\ 09\ 24.769\ 20] \pm [0.197\ 52\ 0.124\ 22\ 1.469\ 19]$	

运用本文所建立的神经网络模型对测试样本数据进行测试,2 种方法的结果列于表 4(只列出其中的 10 组数据)。从上述对比结果看,经典的自由平面标定法 x 坐标平均偏差为 2.43 mm, y 坐标平均偏差为 0.95 mm, z 坐标平均偏差为 1.55 mm,而采用本文研究的神经网络标定技术, x 坐标平均偏差为 0.19 mm, y 坐标平均偏差为 0.16 mm, z 坐标平均偏差为 0.26 mm,很显然达到了更高的精度。

5 结束语

神经网络技术更多的是作为分类器使用的,然而同样具有很强的非线性逼近能力。本文结合双目立体视觉测量方法,利用前馈型 BP 网络进行了

表 4 2 种标定结果对比

Tab.4 Comparison of two calibration results mm

实际坐标 (x, y)	图像坐标 (x, y)	标定方法(x, y, z)	
		神经网络	自由平面
260,0,280	92,154.5	259.85,0.09,280.01	257.39,-1.07,284.12
160,0,280	264,113.5	159.67,-0.04,279.67	155.59,-0.78,283.38
240,0,240	123.5,208.5	239.93,-0.15,240.31	237.11,-1.45,242.88
120,0,240	322,158.5	119.77,0.04,239.98	115.84,-0.33,243.01
180,40,0	224,552	179.74,39.83,-0.05	179.85,39.55,-0.10
80,220,0	530,708	79.84,219.85,0.18	80.09,220.22,0.04
220,260,0	351,842	220.13,260.11,0.17	219.43,260.04,-0.23
0,100,220	599.5,243	-0.34,99.84,220.05	-5.04,101.90,220.54
0,20,160	504.5,247	-0.84,19.35,158.68	-3.62,21.92,160.99
0,260,80	716,614.5	0.18,260.04,80.16	-0.79,261.31,79.75

摄像机标定技术研究。通过试验与分析可知,基于改进的Harris角点提取方法提高网络训练样本集的数据精度,综合采用归一化、提前终止等多种策略,合理选择隐层神经元个数和网络训练函数,从而改

善网络的泛化能力,可以显著提高摄像机的标定精度,并能大大降低传统方法操作的复杂程度,减少坐标间转换造成的累计误差。本文研究方法可应用于高精度的三维测量工作。

参 考 文 献

- 1 Yang Haojun, Zhang Guilin. Feature correlation tracking algorithm using SUSAN operator [J]. *Infrared and Laser Engineering*, 2000, 29(4): 34~37.
- 2 Harris C G, Stephens M J. A combined corner and edge detector [C] // *Proceedings Fourth Alvey Vision Conference*, Manchester, 1988: 147~151.
- 3 刘阳成,朱枫.一种新的棋盘格图像角点检测算法[J]. *中国图像图形学报*, 2006, 11(5): 656~660.
Liu Yangcheng, Zhu Feng. A new kind of arithmetic of chessboard corner detection [J]. *Journal of Image and Graphics*, 2006, 11(5): 656~660. (in Chinese)
- 4 Martin T H, Howard B D, Mark B. Neural network design [M]. Beijing: China Machine Press, CITIC Publishing House, 2008.
- 5 Mark B L, Cihan H D, Mahesh V. The use of feed forward neural networks for machine vision calibration [J]. *International Journal of Production Economics*, 1999, 60~61: 479~489.
- 6 周开利,康耀红.神经网络及其MATLAB仿真程序设计[M].北京:清华大学出版社,2005.
- 7 Tsai R. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses [J]. *IEEE Journal of Robotics & Automation*, 1987, 3(4): 323~344.
- 8 张健新,段发阶,叶声华.“两步法”求取双目视觉传感器中摄像机位置关系[J]. *光电工程*, 1998, 125(5): 38~42.
- 9 崔岸,刘业峰,袁智,等.摄像机参数的三线性插值误差补偿标定方法[J]. *农业机械学报*, 2008, 39(10): 187~190.
Cui An, Liu Yefeng, Yuan Zhi, et al. Camera calibration method of tri-linear interpolation error compensation [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2008, 39(10): 187~190. (in Chinese)

(上接第 174 页)

- 5 Henten E J, Tuijl B J, Hoogaker G J, et al. An autonomous robot for de-leafing cucumber plants grown in a high-wire cultivation system [C] // *Biosystems Engineering*, 2006, 94(3): 317~323.
- 6 Xu Huirong, Zhu Shengpan, Ying Yibin, et al. Application of multispectral reflectance for early detection of tomato disease [C] // *Proc. of SPIE 2006 Optics for Natural Resources, Agriculture, and Foods*, 2006, 6381: 63810R.1~63810R.8.
- 7 毛罕平,张红涛.储粮害虫图像识别的研究进展及展望[J]. *农业机械学报*, 2008, 39(4): 175~179, 186.
Mao Hanping, Zhang Hongtao. Research progress and prospect for image recognition of stored-grain pests [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2008, 39(4): 175~179, 186. (in Chinese)
- 8 Tsai W. Moment-preserving thresholding: a new approach [J]. *CVGIP*, 1985, 29(3): 377~393.
- 9 罗诗途,罗飞路,张■,等.基于梯度调整的矩不变自动阈值图像分割算法[J]. *计算机应用*, 2004(6): 11~13.
- 10 Arlma S, Kondo N, Shibaxo Y. Studies on cucumber harvesting robot: II [J]. *Journal of the Japanese Society of Agriculture Machinery*, 1994, 56(6): 69~76.

(上接第 213 页)

- 3 陈敏华.机器人操作手运动能力指数的应用[J]. *机床与液压*, 2003(2): 102~104.
Chen Minhua. Applications of moving capability index of robotic manipulator [J]. *Machine Tool & Hydraulics*, 2003(2): 102~104. (in Chinese)
- 4 理查德·摩雷,李泽湘,夏恩卡·萨里特里.机器人操作的数学导论[M].北京:机械工业出版社,1998.
- 5 林义忠,黄玉美,程祥.串联机器人诱导运动的分析和研究[J]. *中国机械工程*, 2004, 15(15): 1 323~1 326.
Lin Yizhong, Huang Yumei, Cheng Xiang. Analysis and research of inductive motion in serial robots [J]. *China Mechanical Engineering*, 2004, 15(15): 1 323~1 326. (in Chinese)
- 6 Goritov A N, Korikov A M. Optimality in robot design and control [J]. *Automation and Remote Control*, 2001, 62(7): 1 097~1 103.
- 7 Gosselin C M, Angeles J. A globe performance index for the kinematic optimization of robotic manipulators [J]. *ASME J. Mech. Des.*, 1991, 113(3): 220~226.