文章编号:1674-2869(2024)05-0574-05

基于LSTM网络的机器人异空间手眼标定方法

乐恒韬[1,2],赵康康[1,2],吴松林[1,2],付中涛*[1,2],陈绪兵[1,2]
1. 武汉工程大学机电工程学院,湖北 武汉 430205;
2. 智能焊接装备与软件工程技术湖北省研究中心(武汉工程大学),湖北 武汉 430205

摘 要:针对现有机器人操作空间与相机视野异空间的情况,基于长短期记忆(LSTM)网络,提出了一种新颖的机器人异空间手眼标定方法。首先依次提取并记录标定板上各圆心的像素坐标,然后利用传送带将标定板平移送至机器人工作空间内,并记录机器人末端顶针的位姿信息。其次利用LSTM网络数据训练获得手眼映射关系。最后使用采集的36组真实数据作为验证集来验证预测精度。结果表明,该方法训练的模型所预测的机器人基坐标系坐标平均平移误差仅为0.69 mm,并且针对随机分布于传送带所有工作空间中的验证集数据,平移误差波动值均小于1 mm,有效验证了该方法的鲁棒性和有效性。相较于经典平面标定方法,本文所提出的方法有效工作空间大,标定精度高,并且可以有效补偿相机镜头畸变、深度值变化等因素所带来的误差。

关键词:机器人;手眼标定;LSTM网络;异空间
中图分类号:TP241 **文献标识码**:A **DOI**:10.19843/j.cnki.CN42-1779/TQ.202401027

Robot hand-eye calibration in different spaces based on LSTM network

YUE Hengtao[1,2], *ZHAO Kangkang*[1,2], *WU Songlin*[1,2], *FU Zhongtao*[*1,2], *CHEN Xubing*[1,2]
1. School of Mechanical and Electrical Engineering, Wuhan Institute of Technology, Wuhan 430205, China;
2. Hubei Research Center of Intelligent Welding Equipment and Software Engineering Technology (Wuhan Institute of Technology), Wuhan 430205, China

Abstract: We present a novel robot hand-eye calibration method in an alien space based on the long short-term memory (LSTM) network, addressing the mismatch between the current robot operating space and camera field of view. Firstly, pixel coordinates of each circle center on the calibration board were extracted and recorded sequentially. Subsequently, the calibration board was translated into the robot workspace using a conveyor belt, and the pose information of the robot end effector was recorded. The LSTM network was then employed for data training to establish the hand-eye mapping relationship. Finally, 36 sets of real data were collected as a validation set to assess prediction accuracy. The results indicate that the model trained by this method predicts with an average translation error of only 0.69 mm in the robot base coordinate system. Moreover, for the validation dataset distributed randomly across all workspaces on the conveyor belt, the fluctuation value of translation error is less than 1 mm, effectively affirming the robustness and effectiveness of this method. In comparison to classical planar calibration methods, the proposed method operates efficiently in larger workspaces, provides high calibration accuracy, and can effectively compensate for errors arising from factors such as camera lens distortion and depth value changes.

Keywords: robot; eye-hand calibration; LSTM network; different spaces

收稿日期:2024-01-31
基金项目:湖北省技术创新重点研发计划(2023BAB071);武汉东湖新技术开发区"揭榜挂帅"项目(2024KJB352)
作者简介:乐恒韬,硕士研究生。 Email:yhtengineer@foxmail.com
***通信作者**:付中涛,博士,副教授。 Email:zhongtao.fu@wit.edu.cn
引文格式:乐恒韬,赵康康,吴松林,等. 基于LSTM网络的机器人异空间手眼标定方法[J]. 武汉工程大学学报,2024,46(5):574-578.

视觉传感器为机器人操作任务（如抓取、装配、焊接）提供了各种环境感知信息[1]。通常，视觉传感器固定在机器人末端，或者独立地固定在机器人外部从而构建机器人手眼系统，即分为“眼在手上”和“眼在手外”两种形式[2]。机器人手眼关系的精确程度对于基于视觉引导的机器人操作任务至关重要。因此，国内外学者针对机器人手眼关系的求解方法进行了大量研究。

机器人手眼标定问题实质上是求解矩阵方程 $\boldsymbol{AX}=\boldsymbol{XB}$[3]，其中 $\boldsymbol{A}$ 为机械臂末端的坐标变换关系，$\boldsymbol{B}$ 为相机的坐标变换关系，$\boldsymbol{X}$ 为机械臂末端到相机之间的坐标变换关系。袁泉等[4]利用棋盘格标定法对相机进行标定，利用特征点匹配求解相机参数。对于相机视野模糊的情况，章秀华等[5]提出了一种动目标多视点基于标定的去模糊方法。Wang等[6]建立了惩罚函数的目标方程来更新旋转矩阵，使其单位正交化。然而，旋转矩阵的正交化可能造成精度退化，尤其当求解不准确时。王一凡等[7]提出了一种分别求出四元数实部和对偶部的方法求解矩阵方程。这类方法可以通过优化算法更好地处理旋转部分的噪声和误差，但是计算复杂度较高，需要较长的计算时间。Liu等[8]使用Kronecker积、奇异值分解（singular value decomposition，SVD）算法求解标定方程以获得初始解，最后通过迭代算法进行优化，但是如果没有合适的停止条件会导致过拟合。此外，还有一些数学方法被提出来解决该问题，比如对偶四元数[9]、Clifford代数[10]、旋量理论[11]等。以上方法大致可以分为求解解析解以及求解迭代解两种，这两种方法计算成本较大，不适合快速在线求解。Chen等[12]提出了基于对偶非线性优化和样本筛选的标定方法，但是该方法并没有考虑标定时来自机器人的位姿误差。Chen等[13]使用目标配准误差模型（target registration error，TRE）进行手眼校准，消除了机器人位姿误差的影响，但是没有考虑来自相机模型的噪声和干扰，影响了标定结果的准确性。

长短期记忆（long short-term memory，LSTM）网络在处理序列数据方面表现出优良的性能[14]。Bilal等[15]提出了基于视觉的位姿估计系统，并使用LSTM网络和稀疏回归方法提高其准确性。Kuo等[16]提出了一种传感器融合方案，借助惯性传感器对机器人进行静态标定和动态补偿。但是以上方法都默认相机和机器人处在同一工作空间内，且标定结果的作用空间较小。因此，本文利用LSTM网络，通过引入空间偏移量来处理异空间机器人手眼标定问题，将机器人和相机转换为同一工作空间内，再通过LSTM网络对像素坐标-实际坐标对应数据点进行学习，最后得到能精准反映机器人手眼非线性映射关系的预测模型。相比于经典的平面标定方法，不仅大幅度提高了标定精度，并且克服了经典平面标定方法工作空间小，且仅在标定板所在平面区域内精度较高的缺陷。

1　机器人异空间手眼标定的问题描述

如图1所示，在机器人的工作任务中，其工作空间与相机的视野有时不在同一范围内，本文采用LSTM网络来确定机器人基坐标系$\{B\}$与相机像素平面坐标系$\{C\}$之间的映射关系。

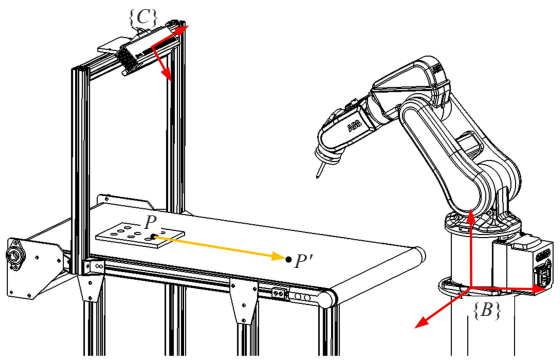

图1　机器人异空间的手眼标定示意图

Fig. 1　Schematic of robot hand-eye calibration in different space

设标定板上的某一角点 P 通过相机识别到的像素坐标为 $P_C=(u,v)$，然后由传送带带动标定板运动一段距离 d 到点 P'，点 P' 位于机器人的工作空间内，用机器人末端顶尖接触点 P'，记录此时点 P' 在机器人基坐标系$\{B\}$的坐标为 $P_B=(x,y,z)$，则点 P 在坐标系$\{B\}$和$\{C\}$的映射关系 f 可表示为：

$$P_C=f\left(P_B\right) \tag{1}$$

1.1　LSTM网络的求解过程

为了确定式（1）的映射关系，本文采用LSTM网络模型，其网络结构如图2所示。

整个LSTM网络模型学习过程的前向传递包含了遗忘门 $\boldsymbol{f}_t$、输入门 $\boldsymbol{i}_t$ 和输出门 $\boldsymbol{o}_t$。其中遗忘门 $\boldsymbol{f}_t$ 用来控制 $t-1$ 时刻的记忆单元状态 $\boldsymbol{C}_{t-1}$ 是否被保留，表达式为：

$$\boldsymbol{f}_t=\sigma\left(\boldsymbol{W}_f\cdot\left[\boldsymbol{h}_{t-1}\quad\boldsymbol{x}_t\right]+\boldsymbol{b}_f\right) \tag{2}$$

式中，$\boldsymbol{x}_t$ 为 t 时刻的输入向量，$\boldsymbol{h}_{t-1}$ 为 $t-1$ 时刻的隐藏状态向量，σ 为sigmoid激活函数，$\boldsymbol{W}_f$ 和 $\boldsymbol{b}_f$ 分别为遗忘门的权重矩阵与偏置向量。

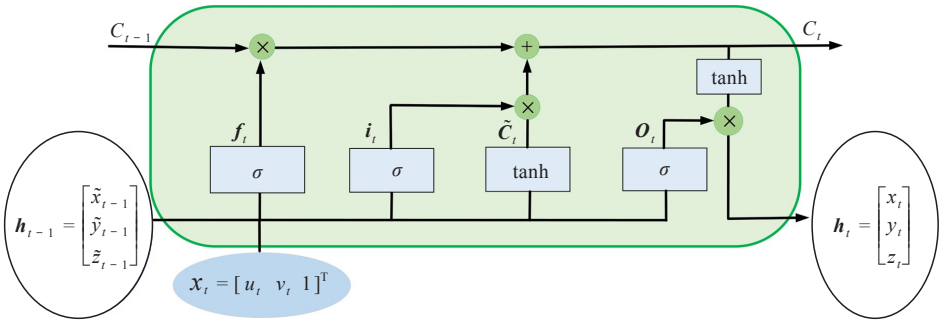

图2 LSTM网络的结构示意图

Fig. 2 Schematic diagram of the LSTM network structure

$\boldsymbol{i}_t$用于控制t时刻的输入向量$\boldsymbol{x}_t$是否被更新到t时刻的记忆单元更新值$\tilde{\boldsymbol{C}}_t$，分别表示为：

$$\boldsymbol{i}_t=\sigma(\boldsymbol{W}_i\cdot[\boldsymbol{h}_{t-1}\quad \boldsymbol{x}_t]+\boldsymbol{b}_i) \tag{3}$$

$$\tilde{\boldsymbol{C}}_t=\tanh(\boldsymbol{W}_c\cdot[\boldsymbol{h}_{t-1}\quad \boldsymbol{x}_t]+\boldsymbol{b}_c) \tag{4}$$

式中，$\boldsymbol{W}_i$和$\boldsymbol{b}_i$为输入门的权重矩阵与偏置向量，$\boldsymbol{W}_c$和$\boldsymbol{b}_c$为记忆单元更新值的权重矩阵与偏置向量。

t时刻的记忆单元状态$\boldsymbol{C}_t$通过遗忘门$\boldsymbol{f}_t$和输入门$\boldsymbol{i}_t$控制$t-1$时刻的记忆单元状态$\boldsymbol{C}_{t-1}$以及更新值$\tilde{\boldsymbol{C}}_t$得到：

$$\boldsymbol{C}_t=\boldsymbol{f}_t\cdot\boldsymbol{C}_{t-1}+\boldsymbol{i}_t\cdot\tilde{\boldsymbol{C}}_t \tag{5}$$

$\boldsymbol{o}_t$基于$\boldsymbol{C}_t$更新t时刻的隐藏状态向量$\boldsymbol{h}_t$，表达式分别为：

$$\boldsymbol{o}_t=\sigma(\boldsymbol{W}_o\cdot[\boldsymbol{h}_{t-1}\quad \boldsymbol{x}_t]+\boldsymbol{b}_o) \tag{6}$$

$$\boldsymbol{h}_t=\boldsymbol{o}_t\cdot\tanh(\boldsymbol{C}_t) \tag{7}$$

式中，$\boldsymbol{W}_o$和$\boldsymbol{b}_o$为输出门的权重矩阵与偏置向量。

1.2 标定点的像素坐标提取

为了更准确地获得标定点P的像素坐标P_C，本文采用便于算法检测的圆形标定板。选用跨平台计算机视觉库(open source computer vision library，OpenCV)的Hough圆检测函数HoughCircles来检测标定板的圆孔，进而得到圆心的像素坐标。Hough圆检测通过Canny边缘检测算子找到潜在的圆心。在圆心基础上，根据边缘的非零像素的灰度值确定圆孔半径，具体流程如图3所示。其中区域分割采用cv2::getRectSubPix函数；高斯滤波采用cv2::GaussianBlur函数；霍夫圆检测采用cv2::HoughCircles函数。

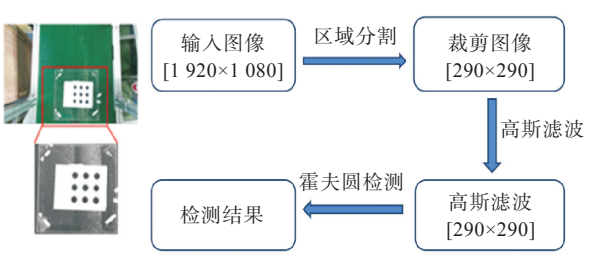

图3 圆心像素坐标检测流程

Fig. 3 Detection of circle center pixel coordinates

2 验证与讨论

为了验证所提方法的有效性，开展了机器人异空间手眼标定的实验并进行了分析。如图4所示，实验装置包括UR5e机器人、KinectV2相机、标定板、顶针和传送带。标定板上9个圆的直径均为20 mm。相机安装在传送带上方，标定板放置在传送带上。为了降低光照带来的影响，实验时在传送带上方使用灯条来做固定光源。标定程序运行在Ubuntu I7-10750电脑上。整个标定实验的流程为：

步骤1：在KinectV2相机视野内放置标定板，使用KinectV2相机拍摄当前标定板的图像，依次提取并记录标定板上各圆心的像素坐标；

步骤2：开启传送带，将标定板平移送至机器人工作空间内；

步骤3：拖动UR5e机器人，使其末端顶针以S型轨迹依次接触每个圆的圆心，按照轨迹顺序依次记录各圆心的像素坐标，同时记录对应的机器人末端顶针的位姿信息。为保证顶针和圆心的正确对应，标定时使用相机算法实时提取出标定圆的圆心，通过程序校准加人为校准，尽可能降低对

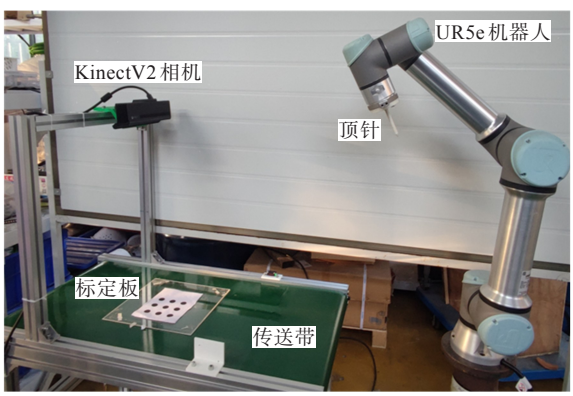

图4 实验装置

Fig. 4 Experimental setup

应关系所带来的误差。重复上述过程,直至采集完36组数据。

设像素坐标为标定板圆心坐标,单位为pixel。像素坐标输入LSTM网络模型解算后得到的输出为模型预测坐标,单位为mm。机器人顶尖顶住标定板圆心时的顶尖末端坐标为机器人实际坐标,单位为mm。采集的36组真实数据作为验证集来验证训练模型的预测精度,数据记录见表1。

表1　标定点的预测坐标与实际坐标

Tab. 1　Predicted and actual coordinates of calibrated points

序列号	像素坐标 / mm	模型预测坐标 / mm	机器人实际坐标 / mm
点1	(570,366)	(802.26,−587.49)	(802.81,−586.84)
点2	(498,362)	(774.07,−614.73)	(773.12,−613.95)
…	…	…	…
点35	(504,222)	(716.60,−552.04)	(715.13,−551.27)
点36	(432,220)	(685.16,−579.91)	(686.31,−579.25)

使用平移误差E来评估LSTM网络的预测精度,表达式为:

$$E = \sqrt{(x_{\mathrm{P}} - x_{\mathrm{A}})^2 + (y_{\mathrm{P}} - y_{\mathrm{A}})^2} \tag{8}$$

式中:x_{P}和y_{P}表示预测值的坐标值;x_{A}和y_{A}表示真实值的坐标值。

LSTM网络模型的训练基于高层神经网络(advanced neural network, ANN)的深度学习框架Keras实现。共采集450组数据,其中训练集和测试集的比例为8:2。数据归一化函数选择最小最大值标准化函数MinMaxScaler,激活函数选择ReLU,全连接层选择Dense,优化器选择Adam,训练完整遍历次数epochs设置为200,训练参数个数batch_size与日志显示verbose均设置为1。训练流程如图5所示。

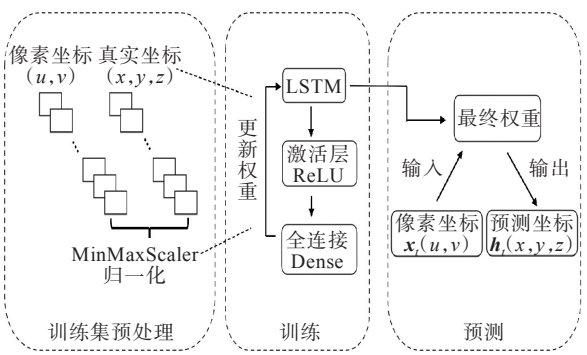

图5　LSTM网络训练流程图

Fig. 5　Training pipeline of LSTM network

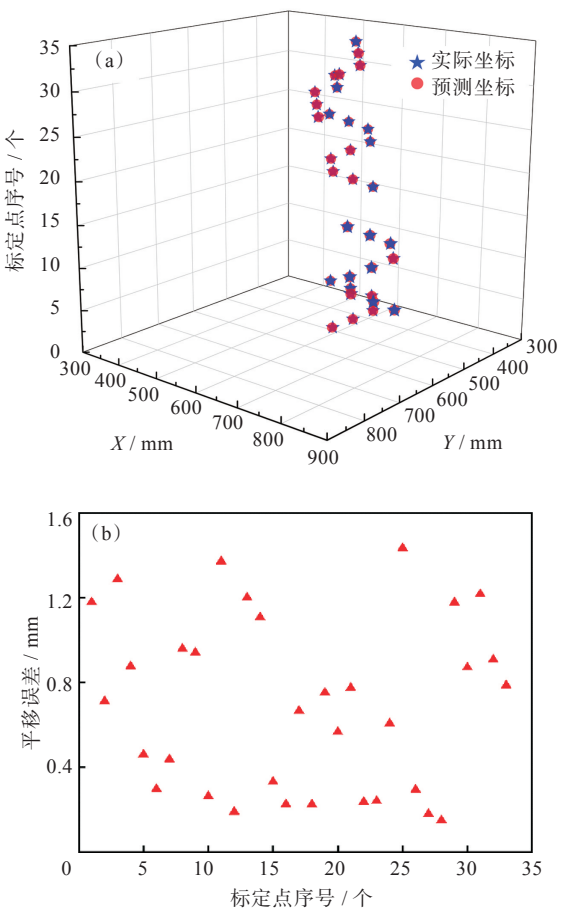

图6　拟合效果图(a);预测坐标平移误差(b)

Fig. 6　Fitting diagram(a);Predicted coordinates of the translational errors(b)

LSTM网络模型的评价指标设定为预测坐标的平移误差。将36组真实数据组成的验证集作为输入,预测结果如图6(a)所示。

如图6(b)所示,该网络模型验证集输出的预测坐标的平均平移误差仅为0.69 mm。验证集数据随机分布于传送带的所有工作空间中,模型预测结果显示平移误差波动值均小于1 mm,证明了该方法的有效性。

3　结　论

针对机器人异空间平面标定问题,创新性地提出了一种基于LSTM网络的机器人手眼关系标定方法。通过引入空间偏移量,利用LSTM网络求解出像素坐标系和机器人坐标系的映射关系。

通过标定实验,得到如下结论:

(1)通过引入移动距离偏移量,解决了相机空间和机器人工作空间不匹配的问题。在传送带稳定工作的情况下,LSTM网络预测的坐标准确度较高。对于工作不稳定的传送带,由于扰动形式多样,如果扰动均匀分布在标定数据集里,在训练过程中也可以很好地被补偿。

(2)基于LSTM网络的机器人异空间手眼关系标定方法不仅可以精准匹配像素坐标和机器人工作平面坐标,并且平移误差波动小,证明了该方法的泛化性。相比于经典的九点标定算法,该方法可以更好地补偿相机镜头畸变、深度值变化等误差,所适用的工作空间更大。

参考文献

[1] 熊有伦,李文龙,陈文斌,等.机器人学:建模、控制与视觉[M].2版.武汉:华中科技大学出版社,2020.

[2] 郭源,张爱军.基于3D相机的拆垛机器人手眼标定方法研究[J].组合机床与自动化加工技术,2024(1):1-4.

[3] SHIU Y C,AHMAD S.Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form *AX*=*XB*[J].IEEE Transactions on Robotics and Automation,1989,5(1):16-29.

[4] 袁泉,邹冲,闵锋.双目视觉在类人机器人测距中的应用[J].武汉工程大学学报,2017,39(2):193-198.

[5] 章秀华,洪汉玉,华夏,等.动目标多视点基于标定的去模糊方法[J].武汉工程大学学报,2014,36(12):56-62.

[6] WANG W,LIU F,YUN C.Calibration method of robot base frame using unit quaternion form [J].Precision Engineering,2015,41:47-54.

[7] 王一凡,段锁林,高仁洲,等.基于对偶四元数的机器人手眼标定算法研究[J].机电工程,2019,36(7):755-760,765.

[8] LIU Z Y,LIU X,DUAN G F,et al.Precise hand-eye calibration method based on spatial distance and epipolar constraints[J].Robotics and Autonomous Systems,2021,145:103868.

[9] FU Z T,PAN J B,SPYRAKOS-PAPASTAVRIDIS E,et al.A dual quaternion-based approach for coordinate calibration of dual robots in collaborative motion[J].IEEE Robotics and Automation Letters,2020,5(3):4086-4093.

[10] BAYRO-CORROCHANO E,DANIILIDIS K,SOMMER G.Motor algebra for 3D kinematics:the case of the hand-eye calibration[J].Journal of Mathematical Imaging and Vision,2000,13:79-100.

[11] ZHAO Z J,LIU Y C.A hand-eye calibration algorithm based on screw motions[J].Robotica,2009,27(2):217-223.

[12] CHEN L,ZHONG G W,WAN Z H,et al.A novel binocular vision-robot hand-eye calibration method using dual nonlinear optimization and sample screening[J].Mechatronics,2023,96:103083.

[13] CHEN E C S,MORGAN I,JAYARATHNE U,et al.Hand-eye calibration using a target registration error model[J].Healthcare Technology Letters,2017,4(5):157-162.

[14] GRAVES A.Supervised sequence labelling with recurrent neural networks[M].Berlin:Springer,2012.

[15] BILAL D K,UNEL M,TUNC L T,et al.Development of a vision based pose estimation system for robotic machining and improving its accuracy using LSTM neural networks and sparse regression[J].Robotics and Computer-Integrated Manufacturing,2022,74:102262.

[16] KUO Y L,HSIEH C H.Static calibration and dynamic compensation of the SCORBOT robot using sensor fusion and LSTM networks[J].Journal of the Chinese Institute of Engineers,2023,46(8):881-894.

本文编辑:霍淑利