

基于向量平均的6DOF估计值波动抑制方法研究

赵宇晴

西安微电子技术研究所科研生产部, 陕西 西安

收稿日期: 2025年3月31日; 录用日期: 2025年6月4日; 发布日期: 2025年6月12日

摘要

针对DenseFusion模型直接输出值波动导致引导机械臂抓取任务失败问题, 本文提出基于向量平均的6DOF估计值波动抑制方法。通过分析DenseFusion算法输出的6DOF估计结果, 我们发现其波动符合正态分布, 并提出了利用向量平均数对位置和姿态参数进行平滑处理的方法。实验结果表明, 该方法能够有效抑制6DOF估计值的波动, 显著提升了机械臂在不同姿态下的抓取成功率, 从70%以上提升至90%以上。本文的研究为6DOF估计值的波动抑制提供了新的思路, 并为手眼协同系统的实际应用提供了技术支持。

关键词

6DOF, 波动抑制, DenseFusion

Vector Averaging-Based Fluctuation Suppression for 6DOF Estimation

Yuqing Zhao

Research and Production Department, Xi'an Microelectronic Technology Institute, Xi'an Shaanxi

Received: Mar. 31st, 2025; accepted: Jun. 4th, 2025; published: Jun. 12th, 2025

Abstract

To address the failure of robotic arm grasping tasks caused by fluctuations in the direct output values of the DenseFusion model, this paper proposes a vector averaging-based fluctuation suppression method for 6DOF estimation. By analyzing the 6DOF estimation results output by the DenseFusion algorithm, we observe that their fluctuations follow a normal distribution. Accordingly, we

introduce a vector averaging approach to smooth both position and orientation parameters. Experimental results demonstrate that the proposed method effectively suppresses fluctuations in 6DOF estimation, significantly improving the robotic arm's grasping success rate across different poses—from over 70% to above 90%. This study provides a novel approach for mitigating 6DOF estimation fluctuations and offers technical support for the practical application of hand-eye coordination systems.

Keywords

6DOF, Fluctuation Suppression, DenseFusion

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在机器人手眼协同系统中, 6 自由度(6DOF)估计是实现精准操作与环境交互的核心技术之一, 其目标是从传感器数据中恢复目标物体或相机在三维空间中的位置和姿态[1]。尽管近年来基于深度学习与多模态融合的方法在 6DOF 估计中取得了显著进展, 估计值的波动问题仍然是一个亟待解决的挑战。这种波动不仅会导致估计结果的不一致性, 还可能在实际应用中引发严重后果, 例如机器人抓取失败或虚拟现实中的视觉抖动。因此, 如何有效抑制 6DOF 估计值的波动, 成为当前研究的热点问题。研究者们从多个角度提出了解决 6DOF 估计值波动问题的方法。基于深度学习的方法, 如 PoseCNN [2]和 PVNet [3], 通过端到端的训练框架直接预测物体的 6DOF 姿态, 显著提升了估计精度。然而, 这些方法对遮挡和光照变化较为敏感, 且依赖大规模标注数据, 导致在复杂场景中波动较大。为了解决这一问题, GDR-Net [4]等研究提出了几何解耦表示网络, 将旋转和平移分支分离, 并通过几何一致性约束减少误差传播, 有效缓解了波动问题, 但其泛化能力仍受限于训练数据分布。多模态融合方法通过结合 RGB-D 相机、IMU 等多传感器数据, 利用卡尔曼滤波或图像优化技术抑制传感器噪声引起的波动。自监督学习方法通过渲染图像与真实图像的差异优化初始姿态估计, 减少了对标注数据的依赖, 并在复杂光照和纹理变化场景中表现出色, 但其优化过程可能陷入局部最优, 且计算资源消耗较大。

尽管上述方法在抑制 6DOF 估计值波动方面取得了一定进展, 但仍存在实时性、泛化能力和硬件依赖性等方面的局限。本文旨在系统分析 6DOF 估计值波动的成因, 并提出一种基于多模态数据融合与自适应优化的新方法, 以在复杂场景中实现更稳定、更高效的 6DOF 估计。通过实验验证, 我们希望对相关领域的研究提供新的思路和技术支持。

2. 方法

2.1. 6DOF 估计

在手眼协同操作中, 6DOF (六自由度)估计是实现高精度操作的核心技术之一。6DOF 估计能够准确获取目标物体或末端执行器在三维空间中的位置(X、Y、Z 轴)和姿态(俯仰、偏航、滚转), 为机械臂的运动规划提供关键输入[5]。通过视觉传感器(如相机或深度相机)与机械臂的协同, 6DOF 估计可以实时更新目标的状态信息, 确保系统在动态环境中的快速响应和精准操作。例如, 在精密装配、抓取不规则物体或手术机器人操作中, 6DOF 估计的准确性直接决定了任务的成败。此外, 6DOF 估计还能优化手眼标定

过程，提升系统的整体稳定性和鲁棒性，为复杂场景下的自主操作奠定基础。

随着计算机的性能提升，和深度学习方法在各个领域的成功应用，近些年和深度学习相结合的 6DOF 估计的算法也有不少，其也有优缺点，常用的算法对比如表 1 所示。经对比本方案采用 DenseFusion 作为 6DOF 估计的基础算法。

Table 1. Comparison of common 6DOF methods

表 1. 常用 6DOF 方法对比

序号	方法名称	优点	缺点
1	DenseFusion	结合 RGB 和深度信息，精度高，对遮挡和复杂背景鲁棒。	计算复杂度较高。
2	PVNet	基于关键点检测，对遮挡鲁棒，无需精确的 3D 模型。	对关键点定位精度依赖高，对低纹理物体效果有限。
3	CDPN	分离旋转和平移估计，提升精度，对遮挡和光照变化鲁棒。	需要大量标注数据，训练过程复杂 [6]。
4	FFB6D	基于 RGB-D 的端到端网络，精度高，对复杂场景适应性强。	依赖高质量的深度传感器，训练数据需求量大。
5	SO-Pose	基于语义和几何信息融合，鲁棒性强，对遮挡和形变鲁棒。	实现复杂度高，对初始位姿敏感 [7]。
6	Gen6D	基于生成模型，适应性强，对未见过的物体有一定泛化能力。	实现复杂度高，实时性较差 [8]。

2.2. 6DOF 估计结果分析

本文以纹理特征丰富的盒子为实验对象如图 1 所示，采用 DenseFusion 算法直接输出会有很大的波动，估计效果如图 2 所示。

由图 2 可见，DenseFusion 估计的结果引导机械臂进行抓取，将会造成机械手与盒子发生碰撞，因此要修正其误差，才能安全地将目标盒子抓取。输出位姿结果在一定的范围内波动，并且符合一定的正态分布状态。当目标处于不同位置时，输出的位姿参数波动量相差较大，有些位置下的输出结果亦为随机数。因此需要进一步对结果进行处理才能传递给机械臂进行抓取操作。



Figure 1. Test object

图 1. 试验对象

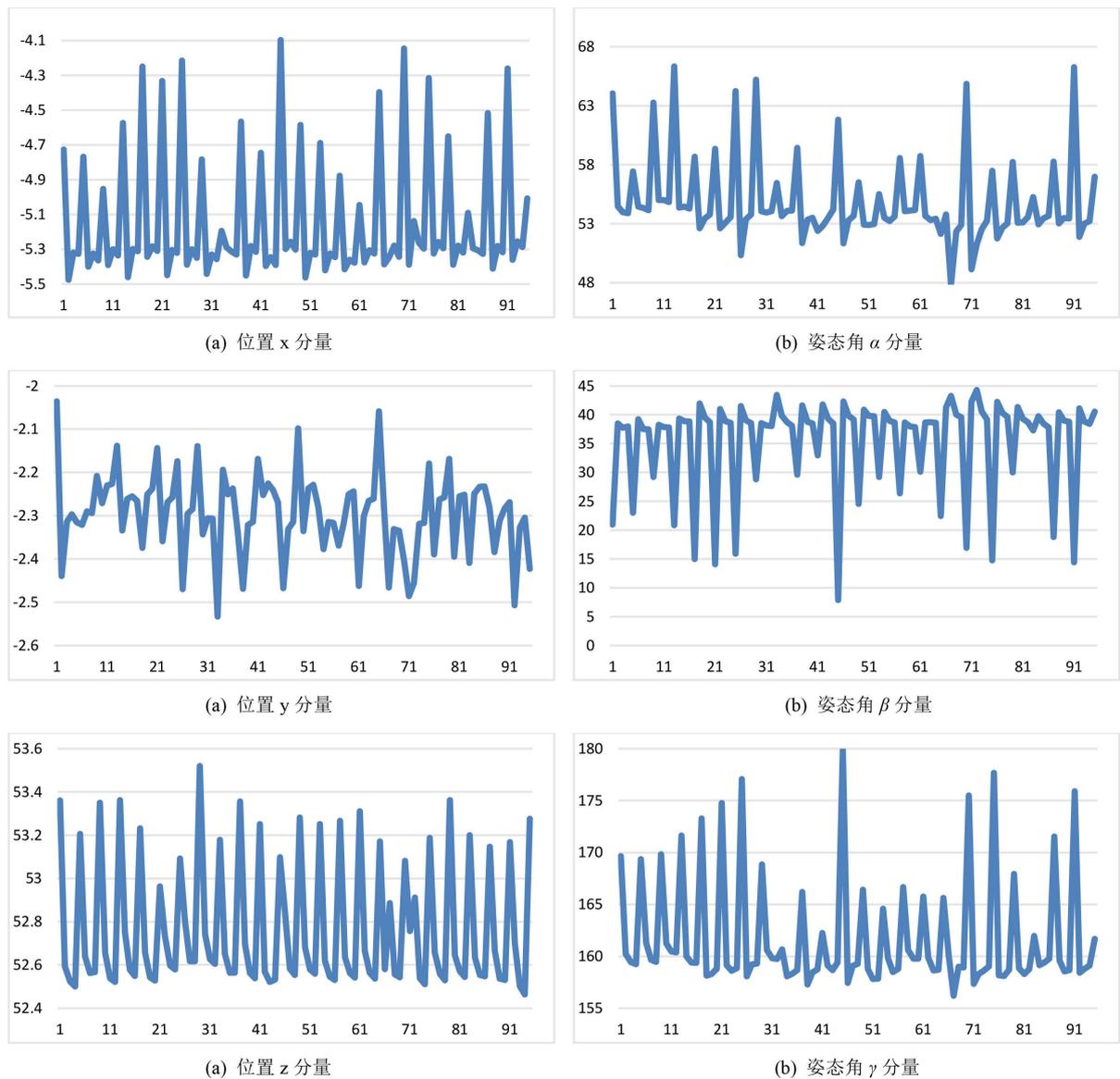


Figure 2. Estimation results of the DenseFusion algorithm

图 2. DenseFusion 算法估计的结果

2.3. 基于向量平均数抑制波动法

由于 DenseFusion 输出的结果符合正态分布,因此可以使用平均数的方法进行处理。由于 DenseFusion 输出结果发生大范围随机现象,可以使用方差进行约束,如连续几帧输出的数据计算的方差大于预定值即可调节识别视角,重新估计目标的 6DOF 参数。

DenseFusion 输出的结果包含 7 个数值($x, y, z, r_x, r_y, r_z, r_w$),分别是坐标的三个参数和姿态的 4 个参数。位置参数(x_m, y_m, z_m)可由连续 n 组估计值进行相加求平均数即可,计算公式如下所示。

$$\begin{cases} x_m = (x_1 + x_2 + x_3 + \dots + x_n) / n \\ y_m = (y_1 + y_2 + y_3 + \dots + y_n) / n \\ z_m = (z_1 + z_2 + z_3 + \dots + z_n) / n \end{cases}$$

姿态参数代表目标物体的基坐标与相机坐标各坐标轴之间的夹角数据，如直接使用连续的四元数求平均数的话，会造成目标基坐标系的坐标轴向量不相互垂直问题，因此本文将四元数 (r_x, r_y, r_z, r_w) 转换成坐标轴转角 (α, β, γ) 再求平均数，求平均数公式如下所示。

$$\begin{cases} \alpha = (\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_n) / n \\ \beta = (\beta_1 + \beta_2 + \beta_3 + \dots + \beta_n) / n \\ \gamma = (\gamma_1 + \gamma_2 + \gamma_3 + \dots + \gamma_n) / n \end{cases}$$

通过以上过程即可抑制 DenseFusion 输出值的波动。以连续 10 组数据为例，以图 1 为姿态估计目标，并采用上述方法抑制 DenseFusion 输出值。为探究方法的可行性，本文在不同光照下进行试验，如图 3 所示，试验结果如图 4~6 所示。



(a) 53.9 lx



(b) 642.3 lx

(c) 1132 lx

Figure 3. Target objects under different lighting conditions**图 3.** 不同光照下的目标

由图 4~6 可见本文方法可以有效抑制 DenseFusion 输出值的波动，从而进一步提升传输各机械臂的位姿精度，进一步提升机械臂抓取目标的安全性。

由图 4 可知，位置抑制后的数据波动在 5 mm 以内，角度抑制后的数据波动在 5° 以内；由图 5 可知，位置抑制后的数据波动在 3 mm 以内，角度抑制后的数据波动在 5° 以内；由图 6 可知，位置抑制后的数据波动在 2 mm 以内，角度抑制后的数据波动在 3° 以内。由以上数据可见由于光照强度的升高，抑制后的数据波动逐渐减小，因此光照环境也是影响位姿估计的因素之一，这是因为由于光照强度的增高，目标的表面在图像中的成像纹理更为明显，导致位姿估计的数据与真实值更接近。有数据显示，抑制后的数据位置在 5 mm 以内波动，角度在 5° 以内波动，可以适应生活中的绝大部分场景。

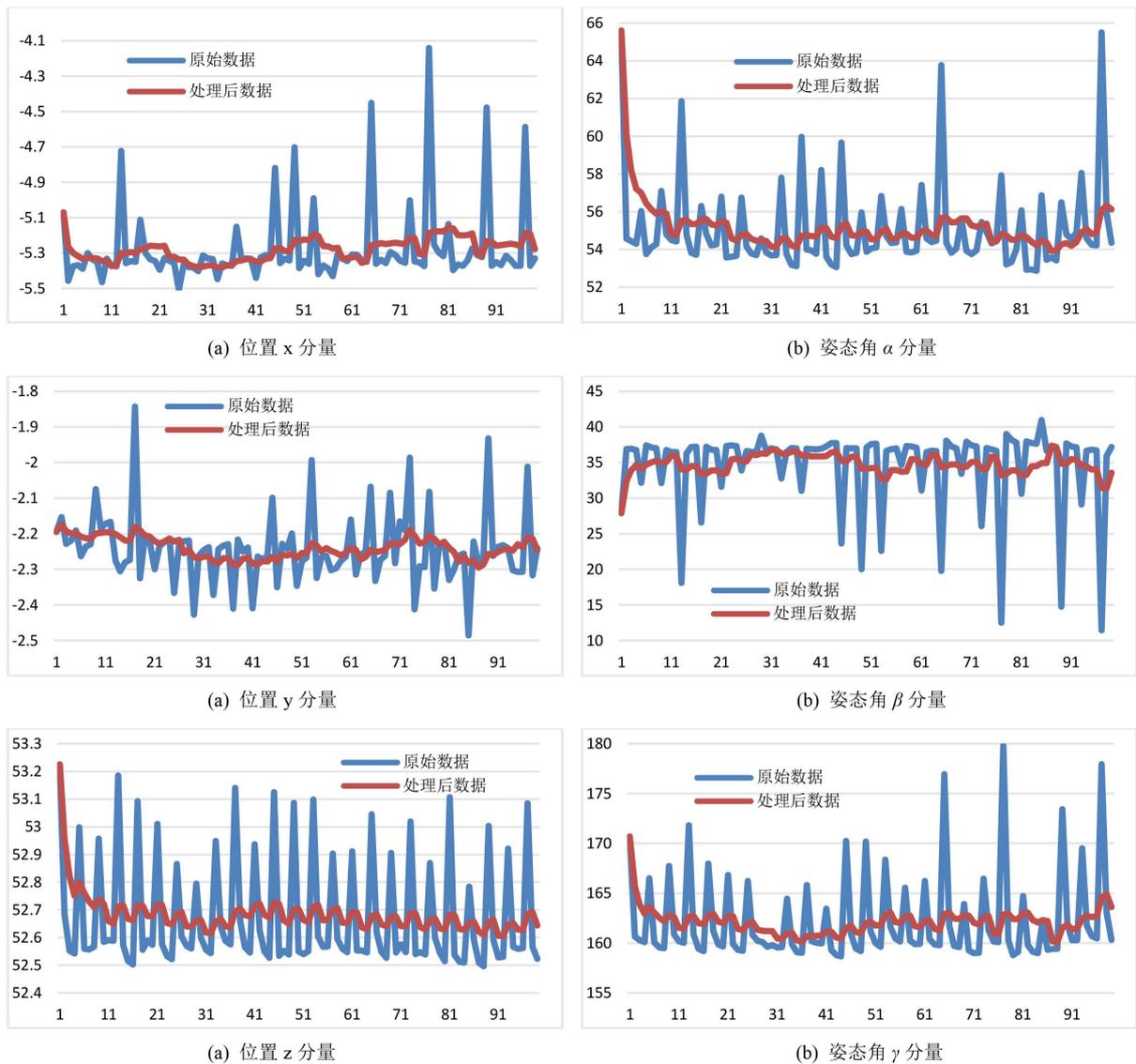


Figure 4. Suppression effect under 53.9 lx illumination
图 4. 53.9 lx 的光照环境下的抑制效果

3. 试验及分析

由于本文研究的 6DOF 估计值的波动抑制方法, 无法直接展现方法的效果, 因此本文设计采用抓取实验验证本文方法在实际应用的效果。本文采用的实验装置如图 7 所示, 包括机械臂、机械手、抓取目标和运算电脑。

其中机械臂的作用是携带机械手前往待抓取目标的位置, 机械手的作用是夹持目标物体, 运算电脑的作用是为图像数据、点云数据等数据的处理, 和 DenseFusion 算法、本文算法和机械臂控制算法等功能模块提供运算平台。

为验证本文方法的有效性, 通过调节目标的位姿和不同光照下进行重复性抓取实验。首先进行不同位置的抓取试验, 每个位姿进行 50 次重复实验, 实验结果如表 2 和表 3 所示。

由表 2 所见在加入本文的方法后目标抓取成功率明显提升, 且在均在抓取成功率均在 90% 以上。

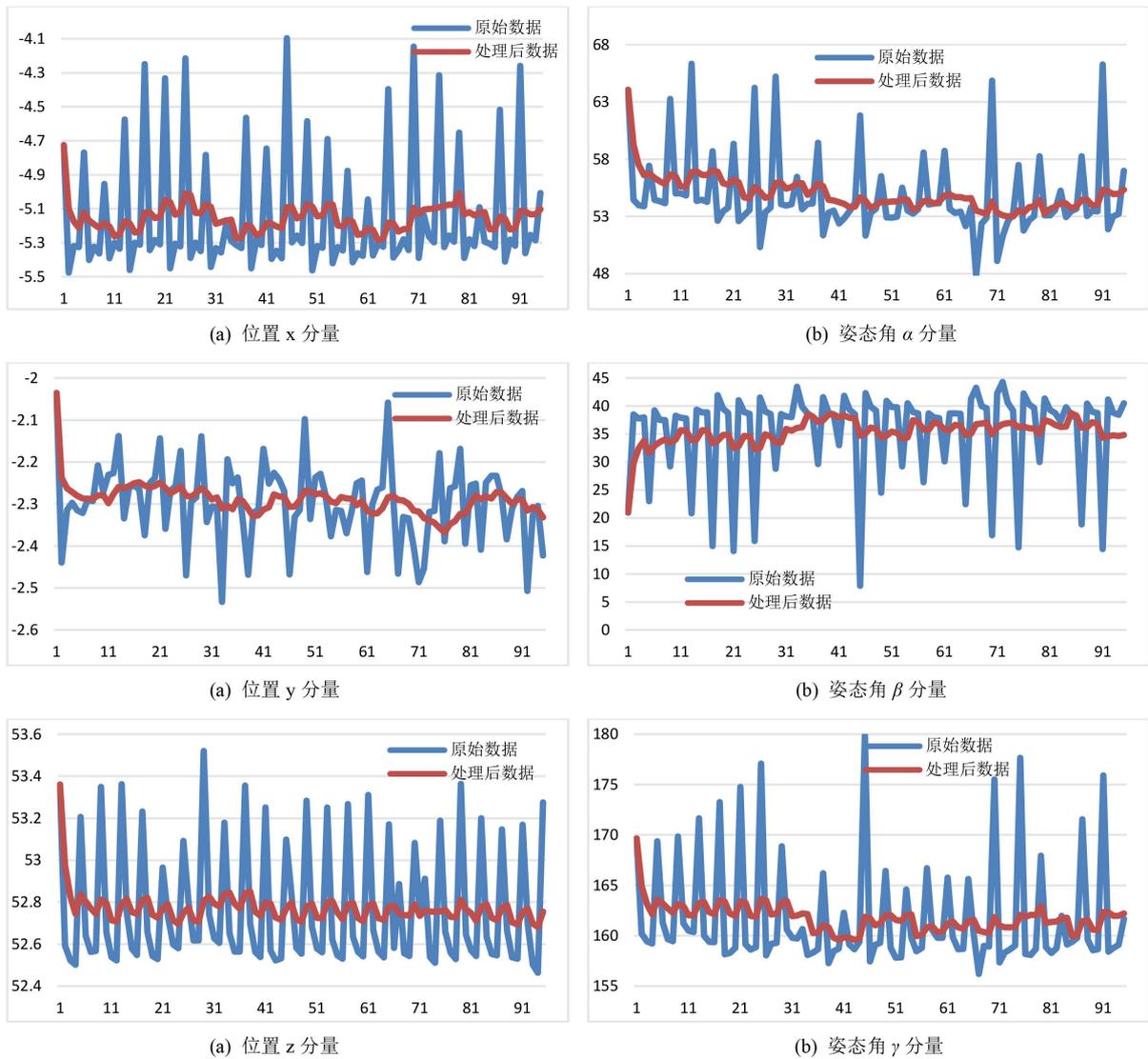


Figure 5. Suppression effect under 642.3 lx illumination
图 5. 642.3 lx 的光照环境下的抑制效果

Table 2. Grasping experiment statistics for different poses
表 2. 不同位姿的抓取实验统计表

姿态角		0°	90°	180°	270°
抓取成功率	未采用本文方法	80%	76%	72%	76%
	采用本文方法	98%	94%	94%	96%

Table 3. Grasping experiment statistics under different illumination conditions
表 3. 不同光照的抓取实验统计表

检测环境	摆放角度					
	0°	90°	180°	270°	360°	任意角
正常光照	98%	92%	94%	94%	96%	92%
强光光照	96%	92%	96%	96%	94%	98%

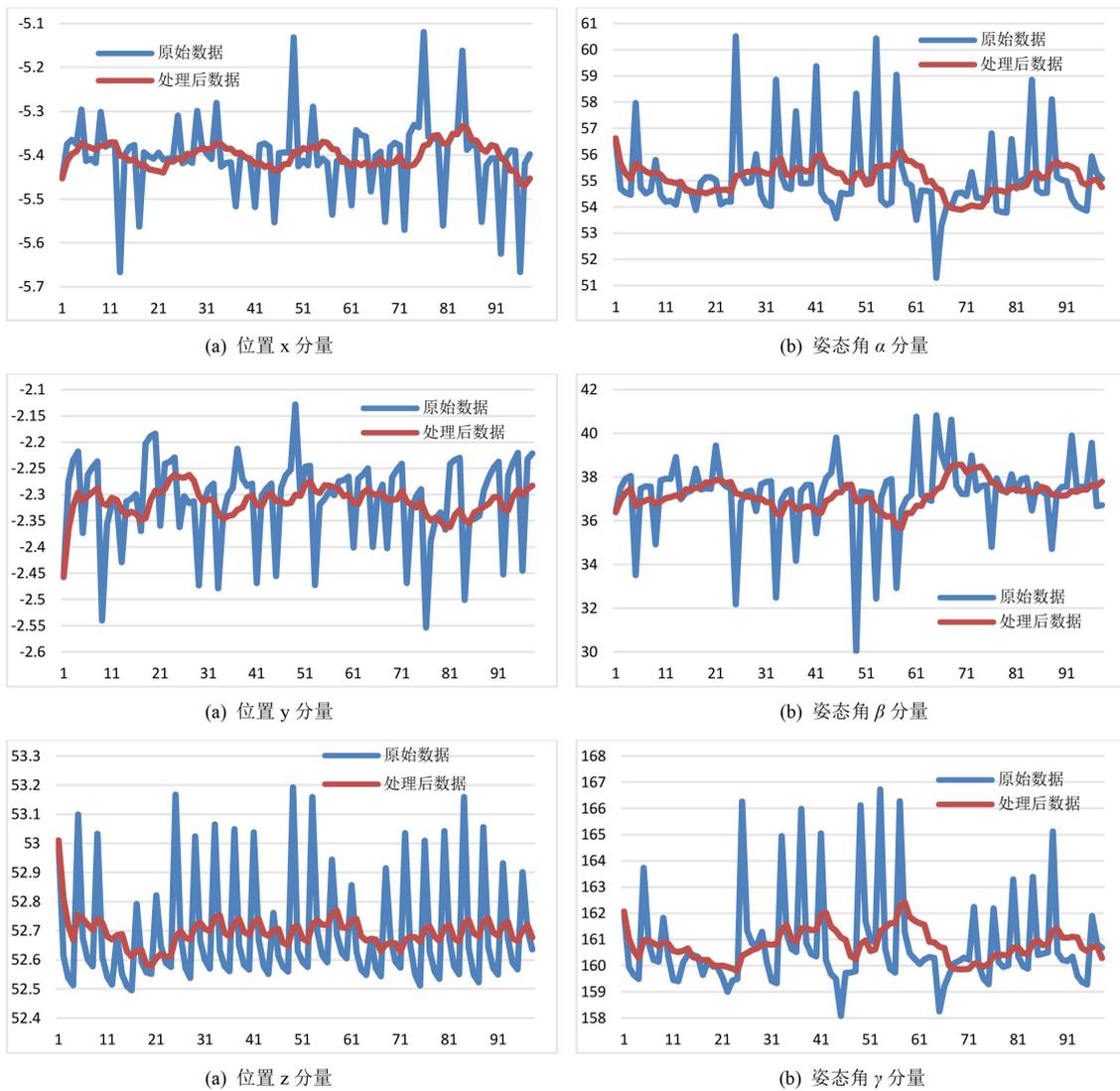


Figure 6. Suppression effect under 1132 lx illumination condition
图 6. 1132 lx 的光照环境下的抑制效果

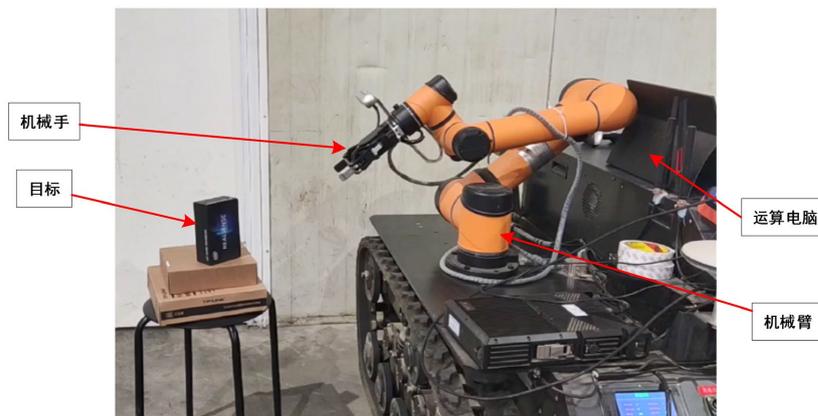


Figure 7. Experimental platform
图 7. 实验平台

由于暗光和正常光照的角度波动相同, 可以通过调节机械手指的初始间距可容许 2 mm 的间距, 在角度方向的误差会由于机械手的宽度影响导致误差容允都不一样, 因此在试验时只进行正常光照(642.3 lx)和强光光照(1132 lx)下的抓取实验, 实验结果如表 3 所示。

由表 3 可知, 正常光照下抓取成功率在 94%, 在强光光照下抓取成功率在 95%, 由此可见位姿估计结果直接影响到支取成功率, 并且抓取成功率与识别精度成正相关。

4. 结论

本文采用基于向量平均数抑制波动法抑制 DenseFusion 的波动, 抑制效果很明显, 并且有效改善由于位姿估计引起的抓取成功率低的问题。本文通过将基于向量平均数抑制波动法融入到手眼协同抓取的任务中, 将抓取成功率 70%多提升到 90%多, 进一步推进基于 6DOF 的手眼协同功能落地。

参考文献

- [1] 刘道鹏. 基于卷积神经网络的机器人抓取目标位姿估计技术研究[D]: [硕士学位论文]. 镇江: 江苏大学, 2023.
- [2] Xiang, Y., Schmidt, T., Narayanan, V. and Fox, D. (2018) PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes. <https://arxiv.org/abs/1711.00199>
- [3] Peng, S., Liu, Y., Huang, Q., Zhou, X. and Bao, H. (2019) PVNet: Pixel-Wise Voting Network for 6DoF Pose Estimation. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 4556-4565. <https://doi.org/10.1109/cvpr.2019.00469>
- [4] Wang, G., Manhardt, F., Tombari, F. and Ji, X. (2021) GDR-Net: Geometry-Guided Direct Regression Network for Monocular 6D Object Pose Estimation. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 16606-16616. <https://doi.org/10.1109/cvpr46437.2021.01634>
- [5] 沈冬春. AI 开放平台技术初探[J]. 中国安防, 2022(5): 41-45.
- [6] 许鑫. 基于三维视觉的轴孔自动化装配方法研究[D]: [硕士学位论文]. 重庆: 重庆大学, 2020.
- [7] 唐梦莹. 基于强化学习的空间双臂智能协同捕获策略研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2019.
- [8] 杨霄鹏, 黄洋, 杨朝阳, 等. 基于混合高斯粒子滤波的 OFDM 频偏估计算法[J]. 探测与控制学报, 2015, 37(6): 71-76.