中国科学: 技术科学 2026 年 第 56 卷 第 6 期: 1434 ~ 1446
SCIENTIA SINICA Technologica techcn.scichina.com
《中国科学》杂志社 SCIENCE CHINA PRESS
论 文 人形机器人专题

基于视觉重定向与卡尔曼滤波的航天员助手机器人遥操作控制方法研究

王思越[1,2†], 杨妍[1,2†], 刘金国[1*], 邵哲铭[3]

1. 中国科学院沈阳自动化研究所机器人与智能系统全国重点实验室, 沈阳 110016
2. 中国科学院大学, 北京 100049
3. 辽宁工业大学, 锦州 121001
† 同等贡献
* E-mail: liujinguo@sia.cn

收稿日期: 2025-09-09; 接受日期: 2026-03-04; 网络版发表日期: 2026-06-10

摘要 随着航天员助手机器人在微重力环境中执行舱内任务的需求不断增加, 灵巧手的高精度遥操作控制方法已成为实现复杂作业的关键技术, 直接决定了空间任务的执行精度与效率. 然而, 在视觉遥操作控制过程中, 手部关键点检测易受舱内动态环境干扰, 产生抖动与短时异常, 若将其直接映射至灵巧手, 将导致动作稳定性下降和响应延迟增加, 严重制约遥操作控制的可靠性与精准度. 针对上述问题, 本文提出了一种基于视觉运动重定向与卡尔曼滤波的遥操作控制方法. 首先, 依托视觉感知模块, 精准提取人手21个关键点的三维坐标, 并结合人手与灵巧手的结构特征, 构建多自由度运动映射模型, 实现人手多关节动作向灵巧手弯曲与旋转角度序列的转化; 其次, 对灵巧手关节角度与角速度进行联合状态估计, 同时融合指数加权移动平均算法抑制关键点数据跳变, 再通过卡尔曼滤波的状态预测功能补偿遥操作延迟, 最终生成实时与稳定的灵巧手控制指令. 实验结果表明, 该方法能够显著减小关节抖动幅度, 提升动作轨迹的平滑性与系统响应速度, 并在不确定性干扰条件下表现出良好的稳定性与适应性, 验证了其在复杂空间任务中的应用潜力.

关键词 航天员助手机器人, 遥操作控制, 灵巧手, 姿态检测, 重定向, 卡尔曼滤波

1 引言

近年来, 随着载人航天任务进入长期化与常态化阶段, 航天员助手机器人在空间站、深空探测器等任务中承担着日益重要的角色[1~3]. 特别是在舱内微重力环境中, 漂浮式助手机器人依靠多向推进器实现自由飞行与姿态调整, 可协助航天员完成设备检测、物资搬运、实验监测等多类型任务[4~6]. 相比固定基座或轨道式机械臂, 舱内助手机器人具备机动灵活、部署成本低、适应性强等显著优势, 能够在狭小且结构复杂的舱内环境中快速重定位, 并可根据任务需要搭载视觉、力觉等多模态传感器与灵巧手、操作工具等末端执行器, 形成开放、可重构的多用途作业平台[7]. 然而, 这类机器人要在动态、受限且任务多样的舱内环境中

引用格式: 王思越, 杨妍, 刘金国, 等. 基于视觉重定向与卡尔曼滤波的航天员助手机器人遥操作控制方法研究. 中国科学: 技术科学, 2026, 56: 1434–1446
Wang S Y, Yang Y, Liu J G, et al. A study on teleoperation control methods for astronaut assistant robots based on visual redirection and Kalman filtering (in Chinese). Sci Sin Tech, 2026, 56: 1434–1446, doi: 10.1360/SST-2025-0256

© 2026 《中国科学》杂志社

实现稳定、精准、高效运行，对其自主环境感知、实时定位控制和高可靠性人机协同能力提出了更高要求[8~12].

随着空间站建设与在轨任务的持续推进，空间机械臂及灵巧手的关键技术研究不断取得进展. 相关研究针对天宫二号机械手开展了在轨实验，验证了抓取、操作和交互等多类任务能力，为复杂舱内外作业提供了重要工程经验[13]. 同时，系统梳理了空间站核心舱机械臂的关键技术，结果表明视觉感知与高精度控制仍是制约因素[14]. 在机构与控制层面，绳簇混合驱动的空间机械臂及其“力-位-型”融合控制方法展现了提升灵巧操作性能的新思路[15]. 空间作业智能操控技术的系统性研究与展望进一步强调了自主感知与远程交互的协同发展方向[16]. 然而，这些成果更多集中于舱外大尺度机械臂及地面验证平台[17]，对于舱内漂浮式助手机器人在微重力环境下如何实现高效稳定的灵巧手遥操作控制方法仍缺乏系统性探索，这也成为制约其实用化的重要因素.

与此同时，国内外在视觉驱动的灵巧手遥操作方法上也开展了大量探索. 已有研究提出腱驱动空间五指灵巧手的控制系统，验证了高自由度手部机构在复杂任务中的可行性[18]. Dexpilot与AnyTeleop等系统通过单目或深度视觉，实现了跨平台的手臂协同遥操作，展现了视觉重定向与动作映射技术在提升通用性与自然度方面的潜力[19,20].

同时, 基于red green blue-depth (RGB-D)的手势检测框架被提出以改善手部关键点检测的实时性与精度，为灵巧手映射提供了可靠数据支撑[21]. 与基于数据手套或机械外骨骼的传统遥操作方式相比，视觉驱动方法能减轻操作者负担，提高交互自然度，但也引入关键点抖动与遮挡问题. 现有部分研究在协作机器人中采用视觉粗定位结合力控实现复杂抓取[22,23]，但在微重力舱内，末端力控和自主策略易受不确定性干扰，系统稳定性难以保证. 本文提出的视觉运动重定向与卡尔曼滤波框架，在保证遥操作精度和实时性的基础上，实现高可靠的舱内操作. 这些研究共同揭示了在复杂空间舱内环境下，视觉遥操作方法需要同时解决实时性、鲁棒性与交互自然度的问题[24~26]，本文采用一种视觉重定向与卡尔曼滤波相结合的策略实现航天员助手机器人的遥操作控制.

针对上述问题，本文面向空间舱内助手机器人的灵巧手遥操作任务，提出了一种融合视觉重定向与卡尔曼滤波预测的联合控制方法. 本文研究聚焦于操作者人手视觉输入的三维姿态识别、关节级状态预测与运动重定向，在舱内漂浮式助手机器人遥操作场景下优化视觉噪声处理与延迟补偿，实现高精度、低抖动的关节控制. 本文未展开灵巧手空间位姿控制，其成熟的伺服系统可保证稳定性. 该方法通过视觉重定向机制抑制因参考系变化引起的位姿映射误差，并利用卡尔曼滤波器对关节角序列进行平滑与预测，以有效抑制信号抖动、增强系统在干扰下的稳定性与鲁棒性, 同时提升动作执行的连续性与自然度.

为验证所提方法的有效性，本研究基于中国科学院沈阳自动化研究所研制的舱内助手机器人平台AAR-2 (astronaut assistant robot-2)开展实验验证. 该平台采用12个涵道风扇推进实现六自由度悬浮运动，集成RH56DF3五指灵巧手与Intel RealSense D455视觉传感器，并依托大型气浮台模拟微重力操作环境，能够真实复现空间舱内动态漂浮操作场景. 实验结果表明，该方法在典型抓取、精细操作与人机交互任务中显著提升了任务完成率、控制精度及操作流畅性，为未来空间舱内智能协作机器人的发展提供了关键技术支持与系统设计参考[27~29].

本文主要创新点如下：在传统卡尔曼滤波框架基础上，针对舱内漂浮式助手机器人的遥操作视觉场景提出创新性改进，通过联合状态建模、视觉关键点级联平滑以及基于预测的延迟补偿，实现高噪声、异步输入环境下的关节级平滑控制. 该方法在轨迹连续性、响应速度与操作稳定性方面优于传统滤波方法，为微重力环境下高精度遥操作提供可靠技术方案.

2　手部三维姿态识别方法

本节所述的手部三维姿态识别针对操作者人手，用作遥操作输入信号，机器人灵巧手的关节角与位姿由第3节运动重定向与状态预测算法生成，并通过现有伺服系统执行，保证动作的连续性与稳定性. 灵巧手空间位置与姿态控制未在本文展开，其运动能力及稳定性由航天员助手机器人AAR-2及六自由度驱动系统保障.

作为遥操作的输入基础，视觉感知是灵巧手控制的关键环节，其核心任务是在空间舱内动态环境中实

时、准确地获取人手21个关键点的三维坐标[30]，为后续运动重定向提供稳定输入，如图1所示. 本系统的视觉模块通过硬件选型优化、高精度检测算法设计及环境适应性处理, 实现舱内环境下人手运动的可靠感知.

2.1 硬件配置与数据采集

针对空间舱内微重力、光照不均、设备密集的特殊情况, 要求实时捕获操作者手部运动信息, 本系统采用Intel RealSense D455深度相机作为核心视觉传感器. 该设备通过双目红外成像与激光散斑结构光融合技术, 在0.6至6 m范围内生成同步的RGB (red green blue)与深度数据流. 可将相机刚性固定AAR-2机器人末端执行器，按照舱内部署布局，将镜头朝向操作者手部运动区域，通过航空级线缆与机器人主控单元连接, 确保数据传输延迟符合空间试验参数要求[31].

2.2 基于MediaPipe的手部关键点检测算法

2.2.1 MediaPipe算法介绍

本研究采用Google开源的MediaPipe Hands框架作为手部实时姿态感知的核心算法. 该框架基于端到端的轻量化卷积神经网络，可从单目RGB图像中高效检测出21个手部关键点的精确位置. 其网络采用两阶段级联结构: 该流程首先通过BlazePalm检测器快速定位手掌并预测其边界框，随后在此裁剪出的手部区域内回归出归一化的关键点坐标. 这些关键点完整覆盖手腕、指节关节及指尖等解剖学部位，能够充分表达手部的复杂构型[32,33]. 除二维像素坐标外，模型还可输出各关键点的相对深度值. 尽管不具备绝对物理尺度，但该值能够有效表征关键点相对于手腕中心的深度偏差，为三维手部姿态估计提供几何结构先验. 这一特性使其易于与Intel RealSense D455等深度传感器提供的绝对深度信息进行融合，共同构建稳定准确的三维关键点坐标. MediaPipe Hands网络架构如图2所示[34].

2.2.2 手部关键点检测流程

基于MediaPipe Hands框架，结合Intel RealSense D455同步输出RGB图像和深度图，实现21个手部姿态三维关键点的实时检测，处理流程包含三级级联模型: 手掌检测、手部关键点回归与三维坐标重构[35]. 获取手部关键点三维坐标点的具体方案如图3所示，主要

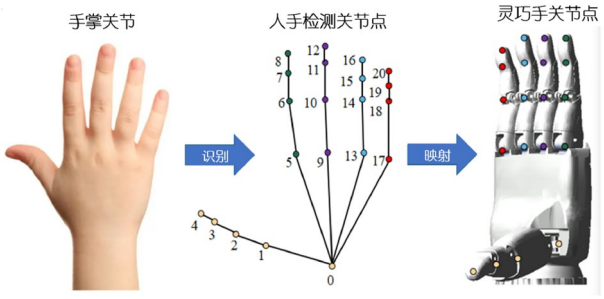


图 1 (网络版彩图)人手关节点与灵巧手映射关系
Figure 1 (Color online) The mapping relationship between the hand-dominant joint points and the dexterous hand.

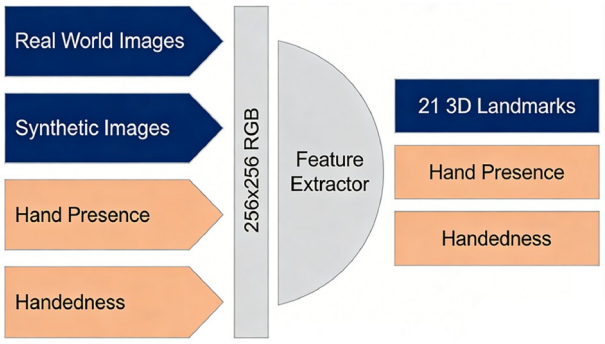


图 2 (网络版彩图) MediaPipe Hands网络架构
Figure 2 (Color online) MediaPipe Hands network architecture.

分为两个部分.

(1) MediaPipe Hands的手部关键点回归

用MediaPipe Hands框架通过两级轻量级神经网络实现RGB图像中手掌定位与关键点回归，该网络以RGB图像为输入, 专注于快速定位手掌区域.

它通过对图像进行多尺度特征提取和分析，能够在复杂背景中高效识别出可能包含手掌的候选区域, 并输出手掌边界框信息. 输出包括手掌中心、5根手指的指根、关节及指尖的21个关键点的归一化二维坐标(u_i, v_i), 以及相对深度$z^{(i)}$, 即点到相机光心的直线距离.

(2) 深度信息处理

对MidiaPipe Hands输出的信息，进一步精细化处理. 它基于手掌区域的局部特征, 通过对特征的深度挖掘和学习, 将MediaPipe输出的归一化二维坐标转换为RGB图像上的像素坐标，利用深度相机从对齐后的深度图中获取物理深度值[36]. 使用Intel RealSense D455的深度传感器进行参数标定，将像素坐标转换为三维

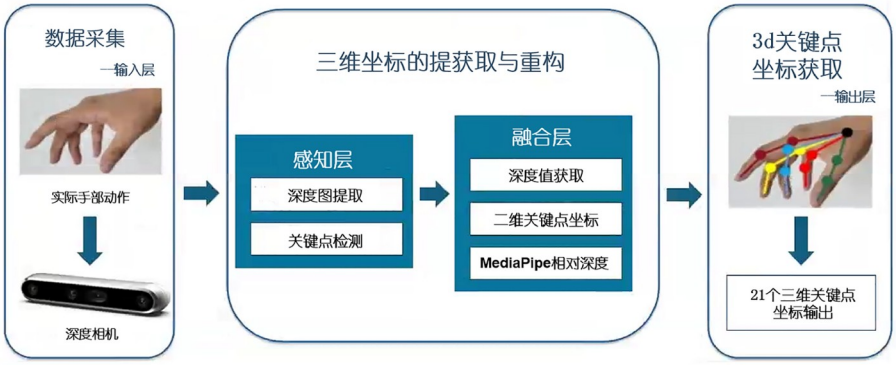


图 3　(网络版彩图)手部关键点检测流程

Figure 3　(Color online) Hand key point detection process.

相机坐标, 给出相机内参矩阵:

$$\mathbf{K}=\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

其中, f_x 与 f_y 分别为x与y方向焦距, c_x与c_y为主点坐标. 将像素坐标结合采样的相机的物理深度值z_d, 转化为三维点(x,y,z), 并结合像素坐标(u,v), 其齐次坐标表示为

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} (u-c_x)z_d/f_x \\ (v-c_y)z_d/f_y \\ z_d \end{bmatrix}, \tag{2}$$

$$\mathbf{p}_c = z_d\mathbf{K}^{-1}[u\ \ v\ \ 1]^{\mathrm{T}}, \tag{3}$$

其中, $\mathbf{p}_c$是手部关键点在相机坐标系下的三维坐标向量.

为提高稳定性(尤其是存在遮挡或深度噪声较大的区域), 我们将MediaPipe Hand输出的相对深度值转换为绝对尺度, 再与相机深度图的观测值进行加权融合, 以此优化最终的深度值. 记手腕关键点索引为0, 有 $\Delta z_d^{(i)}=z_d^{(i)}-z_d^{(0)}$, $\Delta\hat{z}^{(i)}=\hat{z}^{(i)}-\hat{z}^{(0)}$, $z_d^{(i)}$表示第i个手部关键点在相机坐标系下的像素深度估值. $\Delta\hat{z}^{(i)}$表示第i个手部关键点的相对深度, 即该点到相机光心的相对距离. 利用最小二乘法对相对深度缩放因子s进行缩放:

$$s=\frac{\sum_i w_i\Delta z_d^{(i)}\Delta\hat{z}^{(i)}}{\sum_i\left(\Delta\hat{z}^{(i)}\right)^2 w_i}, \tag{4}$$

其中, w_i是加权不同关键点对深度优化的权重, 可根据经验值进行设置, 将优化后的深度值应用于每个关键点, 得到最终的深度值$z^{(i)}$.

$$z^{(i)}=\omega_{\mathrm{depth}}z_d^{(i)}+\left(1-\omega_{\mathrm{depth}}\right)\left(z_d^{(0)}+s\triangle\hat{z}^{(i)}\right) \tag{5}$$

$$\mathbf{a}_c^{(\mathrm{i})}=\left(\frac{z^{(i)}(u-c_x)}{f_x},\frac{z^{(i)}(v-c_y)}{f_y},z^{(i)}\right) \tag{6}$$

其中, ω_{depth}为融合权重, 随深度置信度、局部方差、空洞率递减, 这个权重控制了深度图和相对深度在最终深度值中的影响程度, 确保在噪声或遮挡的情况下更加依赖深度图数据. $\mathbf{a}_c^{(\mathrm{i})}$为最终得到的三维坐标. 该试验通过加权融合深度图和相对深度信息, 提高了在复杂环境中的鲁棒性, 满足遥操作系统在空间舱内动态环境下对低延迟与高连续性的要求, 同时也满足手部关键点定位低复杂度与高精度识别的目的, 试验中手部关键点检测效果图如图4所示.

3　运动重定向与卡尔曼滤波

在机器人视觉遥操作系统中, 如何将图像中识别出的人手关键信息稳定、连续地映射至灵巧手, 是实现高精度遥操作的关键问题. 第2节已介绍, 本系统使用基于MediaPipe的关键点识别算法, 结合RGB-D双目深度相机, 实现对21个手部关键点的实时三维坐标提取. 然而, 深度图像在动态环境下存在遮挡、关键点误检、识别跳变等问题, 直接将原始识别结果映射至灵巧手各自由度会导致输出信号抖动甚至失控. 为解决这一问题, 视觉模块输出的人手关键点三维轨迹经运动重定向映射为灵巧手各自由度目标关节角, 并结

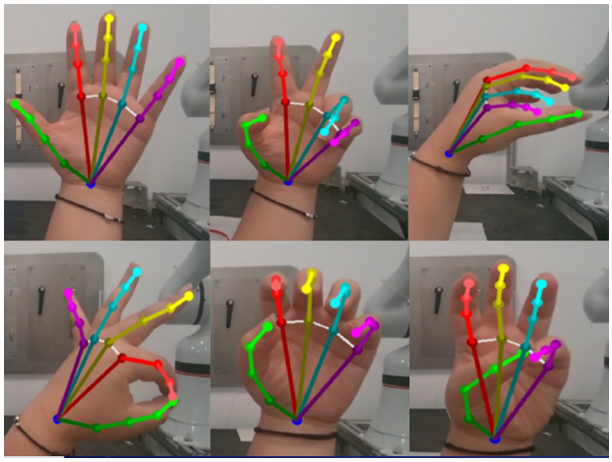


图 4 (网络版彩图)手部关键点检测效果图

Figure 4 (Color online) Image shows the detection effect of hand key points.

合卡尔曼滤波预测实现轨迹平滑与延迟补偿，最终通过现有伺服系统执行. 本文重点关注舱内微重力环境下的视觉噪声抑制、关节动作平滑、延迟补偿及操作精度优化，以保证操作任务板中旋钮旋转、按钮按压和把手拉动任务的高成功率与轨迹连续性，为未来将视觉输入与空间位姿控制结合的研究提供方法基础和实验参考.

针对舱内漂浮式助手机器人的高噪声、延迟及非线性动态特性，本文将“关节角-角速度”构成二维状态向量, 显式建模每个自由度的动态特性, 使状态预测可直接用于延迟补偿和轨迹前瞻. 同时，针对MediaPipe关键点在复杂光照与遮挡下的短时跳变, 具体如下.

(1) 视觉前处理：对原始关键点坐标进行EWMA处理.

(2) 卡尔曼滤波: 对关节角与角速度进行联合滤波.

(3) 滤波后平滑：对输出关节角再次进行EWMA二次平滑.

此外，本文设计了基于预测的两帧延迟补偿策略，并在操作任务板实验中进行验证，显著降低了滞后感，提升了响应速度与轨迹流畅性.

3.1 自由度建模与视觉输入映射

本研究在航天员助手机器人AAR-2平台上实现了从人手到五指灵巧手的运动重定向控制. 视觉识别模块获取人手关键点坐标，包括每根手指的4个关键点: 指根、近端关节、远端关节、指尖以及手掌中心关键点. 设时刻t时第i个关键点在三维空间中的坐标为

$$\mathbf{a}_c^{(i)} = \begin{bmatrix} x_t^{(i)} \\ y_t^{(i)} \\ z_t^{(i)} \end{bmatrix} \in \mathbf{R}^3, \tag{7}$$

其中, $x_t^{(i)}$, $y_t^{(i)}$, $z_t^{(i)}$分别表示深度相机检测到的第i个关键点的空间坐标的水平、垂直和深度方向分量.

本研究所用五指灵巧手具有6个主要自由度: 大拇指包含弯曲和旋转两个自由度，其他四指则各具单自由度的弯曲控制. 每个自由度均由一组直线伺服电机通过腱绳系统(一种仿生肌腱的传动机构)传动，因此灵巧手每根手指的近端关节与指根之间存在固定的角度关系，无法独立控制. 灵巧手每根手指实际仅有3个关节点(指根、中间关节、指尖)，而视觉识别获得的4个关键点中间的两个关节角度需通过固定比例关系映射到单一的弯曲自由度上.

以食指为例，其运动状态通过5个关键点确定，分别为手掌关键点P_0、指根P_1、近端关节P_2、远端关节P_3和指尖P_4. 根据这5个点计算4个骨骼向量:

$$\mathbf{v}_i = \mathbf{a}_t^{(P_i)} - \mathbf{a}_t^{(P_{i-1})},\ i = 1,2,3,4. \tag{8}$$

基于这4个向量, 计算3个连续骨骼向量之间的夹角:

$$\theta_i = \arccos\frac{\mathbf{v}_i\mathbf{v}_{i+1}}{\|\mathbf{v}_i\|\cdot\|\mathbf{v}_i\|},\ i = 1,2,3,4. \tag{9}$$

由于五指灵巧手由腱绳驱动, 5个手指均存在关节角度耦合, 实际控制角度定义为3个夹角的加权和:

$$\theta_t^{(j)} = k_1\theta_1 + k_2\theta_2 + k_3\theta_3, k_1 + k_2 + k_3 = 1, \tag{10}$$

其中, 权重系数k_1, k_2, k_3根据机械结构和人手运动学特性确定，一般满足$k_1 > k_2 > k_3$，表明近端关节对手指整体弯曲的主导作用. 大拇指的弯曲和旋转角度则分别通过骨骼向量的投影及夹角定义得到. 最终构造六维自由度角度向量:

$$\boldsymbol{\theta}_t = \begin{bmatrix} \theta_t^{(\text{thumb_bend})} \\ \theta_t^{(\text{thumb_rot})} \\ \theta_t^{(\text{index})} \\ \theta_t^{(\text{middle})} \\ \theta_t^{(\text{ring})} \\ \theta_t^{(\text{pinky})} \end{bmatrix} \in \mathbf{R}^6. \tag{11}$$

该角度向量作为灵巧手的控制输入，实现了人手与灵巧手在AAR-2平台上的运动同步.

3.2　卡尔曼滤波建模与状态预测

为抑制视觉识别关键点的随机噪声及测量误差，采用离散时间卡尔曼滤波对关键点轨迹进行状态估计[37]. 每个自由度$j\in\{1,...,6\}$的状态向量定义为

$$\mathbf{x}_t^{(j)}=\begin{bmatrix}\theta_t^{(j)}\\ \dot{\theta}_t^{(j)}\end{bmatrix}, \tag{12}$$

其中，$\theta_t^{(j)}$指的是第j个自由度的角度值，$\dot{\theta}_t^{(j)}$表示第j个自由度的角速度值，t表示当前时刻. 状态转移模型假设匀速运动:

$$\mathbf{x}_t^{(j)}=\mathbf{F}\mathbf{x}_{t-1}^{(j)}+\mathbf{w}_t^{(j)},\mathbf{F}=\begin{bmatrix}1 & \Delta t\\ 0 & 1\end{bmatrix}, \tag{13}$$

其中, $\mathbf{F}$为状态转移矩阵, Δt = 1/30 s为采样周期, 对应视觉系统输入帧率30 Hz, $\mathbf{w}_t^{(j)}$为过程噪声, 服从均值为零、方差为Q的高斯分布. 观测模型定义为

$$z_t^{(j)}=\mathbf{H}\mathbf{x}_t^{(j)}+v_t^{(j)},\mathbf{H}=[1\ \ 0], \tag{14}$$

其中，$z_t^{(j)}$为视觉识别系统在时刻t观测到的第j个自由度的角度值, $\mathbf{H}$是观测矩阵, 用于从状态向量中选取出与观测值对应的角度分量. $v_t^{(j)}$为观测噪声，服从均值为零、方差为R的高斯分布. 卡尔曼滤波的预测过程与更新过程如下.

(a) 预测步骤

$$\hat{\mathbf{x}}_{t|t-1}=\mathbf{F}\hat{\mathbf{x}}_{t-1|t-1}, \tag{15}$$

$$\mathbf{P}_{t|t-1}=\mathbf{F}\mathbf{P}_{t-1|t-1}\mathbf{F}^{\mathrm{T}}+\mathbf{Q}, \tag{16}$$

其中，$\hat{\mathbf{x}}_{t|t-1}$是基于$t-1$时刻的状态估计对t时刻状态的预测值，$\mathbf{P}_{t|t-1}$是已知$t-1$状态估计下，对时刻t的预测误差协方差矩阵, $\mathbf{P}_{t-1|t-1}$为时刻$t-1$的误差协方差矩阵.

(b) 更新步骤

$$\mathbf{K}_t=\mathbf{P}_{t|t-1}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{P}_{t|t-1}\mathbf{H}^{\mathrm{T}}+\mathbf{R}\right)^{-1}, \tag{17}$$

$$\hat{\mathbf{x}}_{t|t}=\hat{\mathbf{x}}_{t|t-1}+\mathbf{K}_t\left(z_t^{(j)}-\mathbf{H}\hat{\mathbf{x}}_{t|t-1}\right), \tag{18}$$

$$\mathbf{P}_{t|t}=(\mathbf{I}-\mathbf{K}_t\mathbf{H})\mathbf{P}_{t|t-1}, \tag{19}$$

其中，$\mathbf{K}_t$为卡尔曼增益矩阵，决定预测与观测的权重平衡, $\hat{\mathbf{x}}_{t|t}$为时刻t的后验状态估计, $\mathbf{P}_{t|t}$为后验误差协方差, $\mathbf{I}$为单位矩阵.

为进一步抑制视觉关键点检测过程中因遮挡或环境变化产生的瞬时跳变，对卡尔曼滤波输出的估计角度应用指数加权移动平均滤波进行平滑处理:

$$y_t^{(j)}=\alpha\hat{\theta}_{t|t}^{(j)}+(1-\alpha)y_{t-1}^{(j)},\alpha\in(0,1], \tag{20}$$

其中，$\hat{\theta}_{t|t}^{(j)}$为经卡尔曼滤波后第j个自由度的角度估计, $y_t^{(j)}$为滤波后平滑角度，平滑系数α调节响应速度与平滑效果的权衡. 较大α值保证响应灵敏，较小α值增强动作连续性和平滑度. 本研究通过实验调整选取α = 0.4, 在响应速度与动作平稳性之间达到平衡.

3.3　基于状态预测的延迟补偿策略

考虑到视觉识别、数据处理及机械执行存在一定延迟，直接将时刻t的滤波估计角度$\hat{\theta}_{t|t}^{(j)}$发送给机械手控制器，可能导致控制响应滞后，影响动作的流畅性和准确性. 为此，利用卡尔曼滤波器的状态预测能力，在时刻t预测未来两帧，即时刻t+2的状态，作为控制器的输入角度, 实现延迟补偿. 预测未来两步状态计算:

$$\hat{\mathbf{x}}_{t+2|t}^{(j)}=\mathbf{F}^2\hat{\mathbf{x}}_{t|t}^{(j)}=\begin{bmatrix}y_t^{(j)}+2\Delta t\hat{\dot{\theta}}_t^{(j)}\\ \hat{\dot{\theta}}_t^{(j)}\end{bmatrix}, \tag{21}$$

其中，$y_t^{(j)}$为经过平滑的当前角度估计, $\hat{\dot{\theta}}_t^{(j)}$为当前角速度估计，$\Delta t$为采样周期. 该预测角度用于抵消系统时延，确保机械手动作的及时性和流畅性. 最终控制信号为延迟补偿后预测的角度，基于对卡尔曼滤波结果的平滑处理与状态预测，实现对航天员助手机器人AAR-2灵巧手各自由度的稳定控制.

4　实验验证与性能评估

4.1　实验平台

本研究实验平台基于中国科学院沈阳自动化研究所研制的舱内宇航员助手机器人. AAR-2采用轻量化铝合金主体框架，具备良好的刚度和空间环境适应性.

其移动基座搭载12组高精度涵道风扇推进器，可实现六自由度位姿调整与悬停控制，并在微重力环境下保持运动稳定性．AAR-2机器人本体采用3层功能布局: 上层为通信与配重块；中层构成为控制核心、感知模块和动力系统；下层为锂电电源系统，AAR-2布局如图5所示[38].

末端执行机构为RH56DF3型五指仿人灵巧手，该手具备6个独立自由度与12个耦合关节, 拇指与四指最大抓握力分别达到15和10 N，并集成分布式压力传感与关节位置反馈.

手腕设计支持俯仰与侧摆双自由度转动，可适应不同抓取角度与精细操作需求．视觉感知子系统基于Intel RealSense D455 RGB-D相机, 通过主动红外散斑与立体视觉融合实现1280 × 720分辨率下的深度感知, 在1至3 m范围内深度精度达到±2%．手部姿态检测采用基于MediaPipe Hands的改进关键点识别算法, 可实时输出21个手部关节点的三维坐标信息．为在地面充分模拟空间微重力环境，实验平台配备大型高平整度花岗岩气浮台, 通过喷射高压气体形成气膜, 使AAR-2机器人悬浮于台面之上，可模拟空间微重力条件下的悬浮运动与操作任务．该平台为验证视觉识别关键点、延迟补偿算法和人机协同策略提供了可靠的物理环境.

图6所示为微重力环境模拟气浮台的结构示意图, 该设备以立方体框架为核心承载结构，框架尺寸规格为200 mm × 200 mm × 200 mm，其主体结构体系由支撑底座、3个空气轴承、四根支撑柱及支撑平面共同构成，各组件通过模块化装配形成稳定的力学支撑结构.

其中，支撑底座与3个空气轴承构成承载单元，支撑底座底部与空气轴承顶部采用“球端-球端槽”的配合设计, 利用球面接触的自定心特性, 实现支撑底座与空气轴承之间的自动对中，有效保障了结构装配的同轴度与受力均匀性．当空气轴承的出口压力达到0.3 MPa时, 其底部与大理石桌面之间会形成一层厚度为5~10 μm的气膜，该气膜可通过气体润滑效应抵消系统自身重力，使整个气浮台系统具备无摩擦自由流动特性, 从而模拟微重力环境下的运动状态.

该气浮台的气动系统采用双回路设计，即入口回路与出口回路相互独立且协同工作，共同完成高压氮气的储存、调节与输送，为气膜生成提供稳定的气源保障．该系统配置的单体储气气瓶，其额定容积为0.22 L, 标称工作压力为10 MPa, 可满足气浮台单次实验的气源需求. 当气瓶内氮气压力达到额定值(即完成充气过程)后，需关闭手动阀以切断气源通路，系统配备的压力表可实时采集并显示气瓶内的压力值，为充气过程的终止判断及后续压力监测提供量化依据．将上述各部件按照图7结构搭建实验平台，实际实验平台如图8所示.

4.2 实验与评估

(1) 常见物体抓取实验

为了验证本文方法在复杂太空环境下的抓取稳定性与操作适应性, 开展了常见物体的抓取实验. 实验对

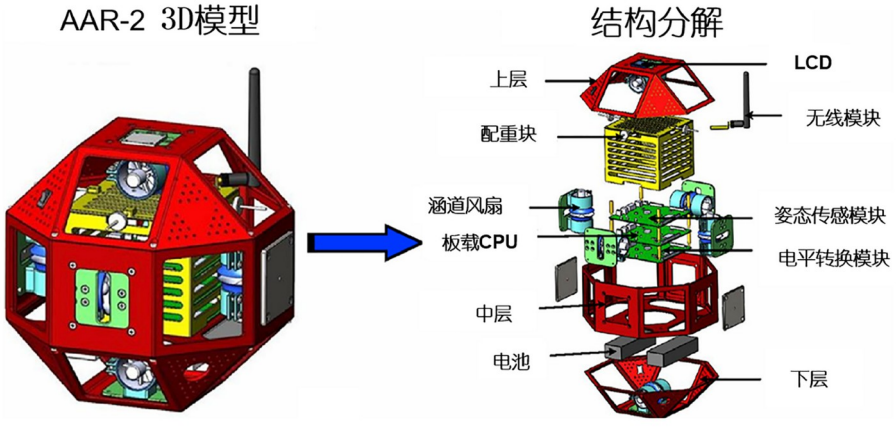


图 5 (网络版彩图) AAR-2结构组成与分解

Figure 5 (Color online) The composition and disassembly of AAR-2.

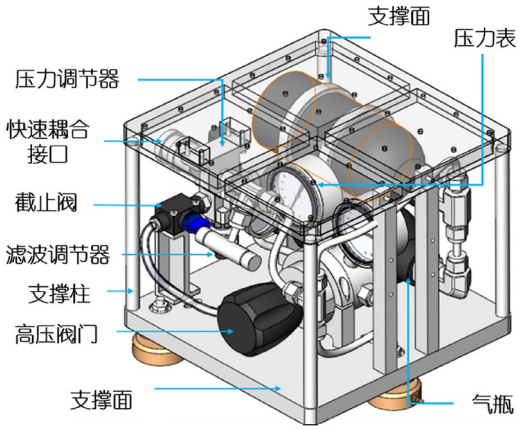


图 6　(网络版彩图)气浮台的内部结构

Figure 6　(Color online) The internal structure of the air flotation platform.

象包括: 地球仪、太空碎片模拟体、3D打印零件、贴片元件、笔以及剪刀等. 上述物体在几何形状、尺寸大小、质量分布和表面摩擦特性方面均存在明显差异, 能够全面考察灵巧手在不同抓取情境下的稳定性与精确度. 在整个实验过程中, 灵巧手关节动作平滑、无显著抖动, 指尖与目标物体的接触位置准确. 抓取实验图示见图9, 具体实验数据对比如表1所示.

实验结果表明: 在简易抓捕任务(T1~T3)中, 系统成功率达到100%, 平均完成时间均在2 s左右, 表明该方法在处理常规、低精度要求的抓取任务时具有极高的可靠性和效率. 在精细操作任务(T4~T6)中, 尽管面对更复杂的操作对象和更高精度要求, 系统仍保持80%以上的成功率, 完成时间的增加反映了任务本身复杂度的提升, 系统仍能保持可接受的性能水平, 证明了其在精细任务中的适用性.

(2) 操作任务板实验

为进一步验证所提方法在空间舱内复杂交互任务中的适用性、精度与稳定性, 本研究设计了操作任务板实验. 任务板集成旋钮、按钮和把手3类常见交互装置, 模拟空间舱内部件调节与交互操作场景. 实验中, 操作者通过人手动作遥操作使灵巧手对任务板上的各类部件执行3种典型操作: 旋转指定角度的旋钮、按压按钮以及拉动把手, 实验效果实现如图10所示(操作目标在虚线框中标注). 每种任务重复10次以统计性能, 得到操作任务板实验性能对比如表2所示.

结果显示, 旋钮操作中, 系统对连续旋转动作的跟踪非常平滑, 无卡顿或跳变, 最终角度误差小, 这充分体现了卡尔曼滤波预测模型在补偿系统延迟、保证运动连续性方面的卓越效果. 按钮操作的成功率最高, 且

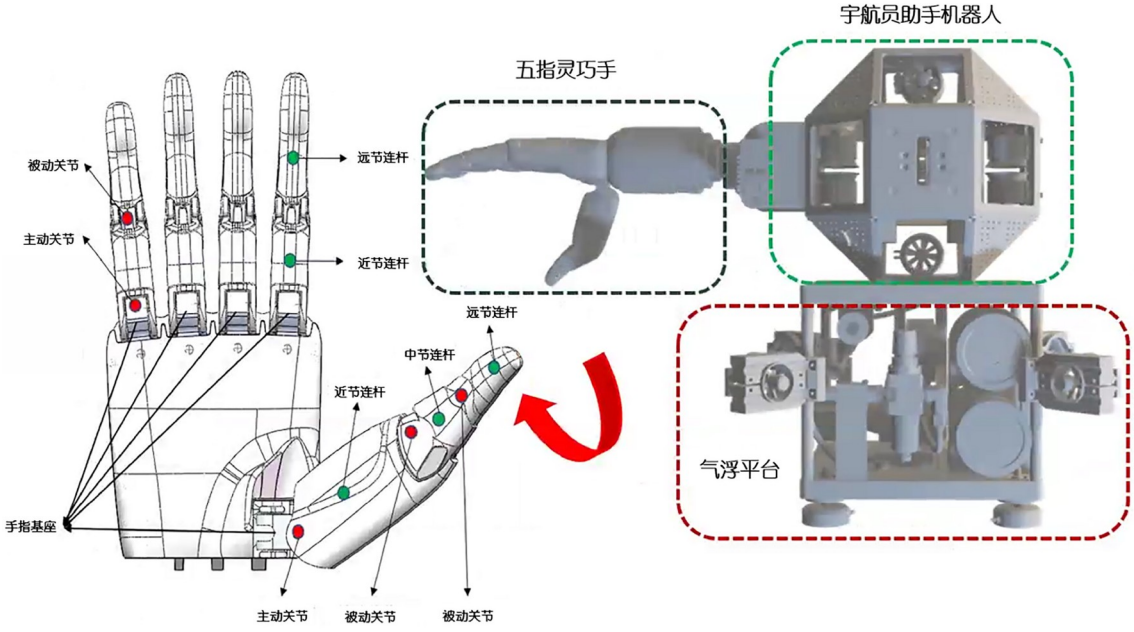


图 7　(网络版彩图)实验平台结构设计图

Figure 7　(Color online) Experimental platform structure design drawing.

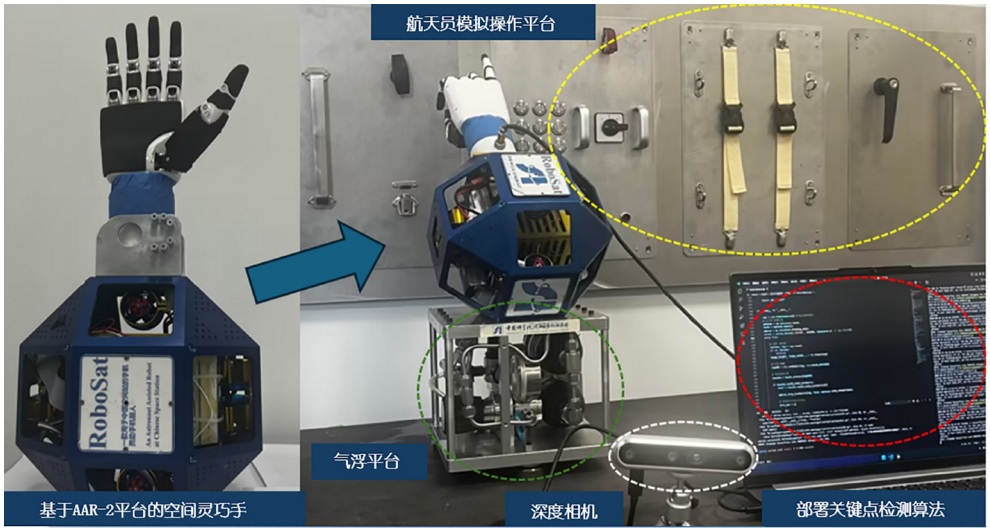


图 8 (网络版彩图) 航天员助手机器人遥操作实验平台

Figure 8 (Color online) Astronaut assistance robot for conducting remote operation experimental platform.

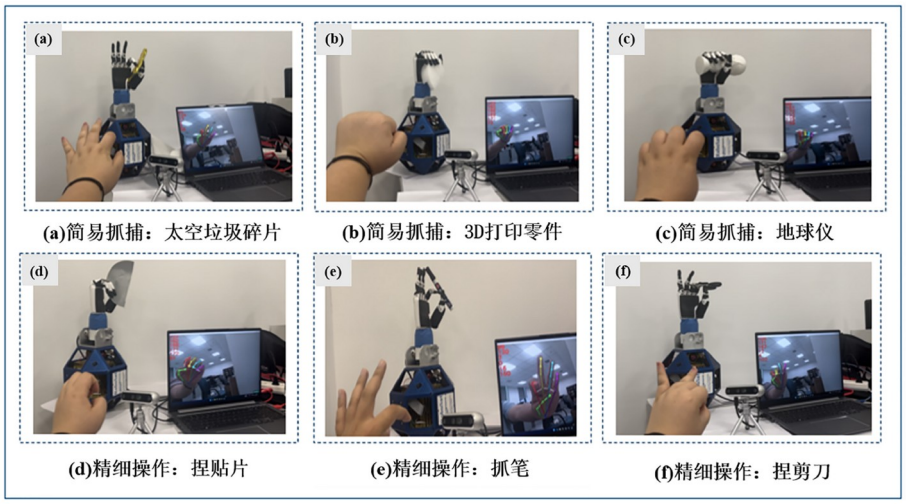


图 9 (网络版彩图)地面遥操作常见物体抓取实验

Figure 9 (Color online) Common object grasping experiments in ground remote operation.

平均用时最短，表明该方法在短行程、高精度定位任务中响应迅速且可靠. 把手拉动任务体现了系统在抵抗微小扰动、维持稳定力控方面的能力，误差主要源于把手机构的微小晃动被系统敏感捕获，但最终的位移精度依然在空间舱内宇航员操作精度范围内.

综上所述，经过视觉运动重定向和卡尔曼滤波处理后的控制信号能够保证灵巧手在任务执行过程中轨迹平稳，指尖与目标部件接触精确、操作顺畅且无明显滞后. 特别是在旋钮连续旋转和按钮反复按压等任务中，系统保持了较高的响应速度和稳定性，充分体现了延迟补偿与轨迹平滑的效果. 本文通过视觉输入处理、关节级状态预测及延迟补偿实现了舱内漂浮式助手机器人的高精度、低抖动、平滑遥操作. 尽管未直接控制空间位姿，该方法在旋钮操作、按钮按压及把手操作中均保持高成功率和轨迹连续性，验证了其工程适用性与创新价值.

表 1　常见物品抓捕操作试验数据对比

Table 1　Comparison of test data for common items' capture operation

任务编号	操作内容	手指数	操作对象	成功率 (%)	平均完成时间 (s)
T1	简易抓捕	二指	太空垃圾	100	2.18
T2	简易抓捕	四指	3D打印零件	100	1.93
T3	简易抓捕	五指	地球仪	100	1.80
T4	精细操作	二指	铁片	80	3.37
T5	精细操作	二指	笔	80	4.16
T6	精细操作	三指	剪刀	90	4.53

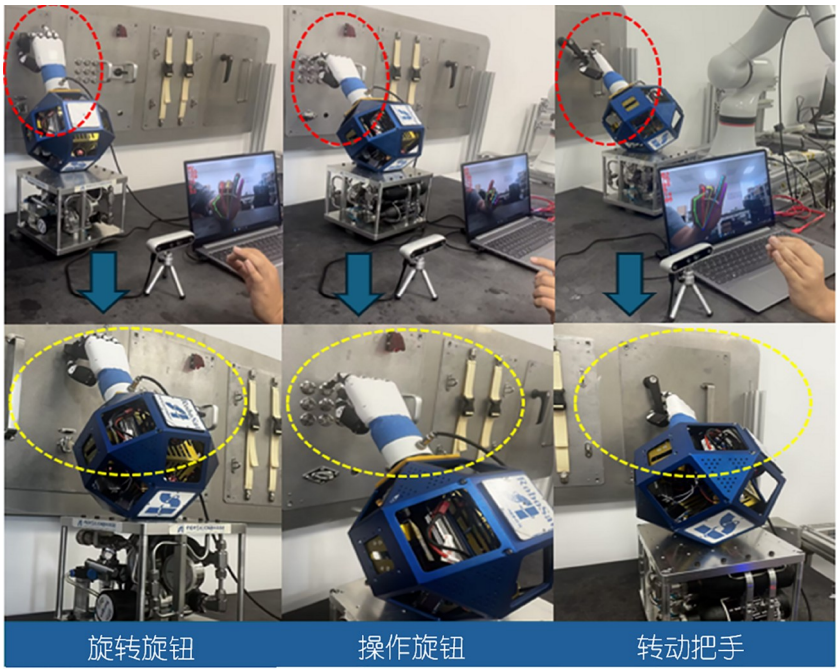


图 10　(网络版彩图)航天员助手操作任务板实验

Figure 10　(Color online) Astronaut assistant operation task board experiment.

表 2　航天员助手操作任务板实验操作性能对比

Table 2　Comparison of operational performance of the astronaut assistant in conducting the task board experiment

任务类型	操作描述	评估指标	成功率 (%)	平均操作时长 (s)	平均误差位置 (°)
旋钮操作	顺时针旋转90°	角度误差	100	3.81	3.2
	逆时针旋转180°	角度误差	90	6.33	5.8
按钮按压	按压至触发深度	触发成功率	100	1.54	–
把手拉动	顺时针旋转90°	角度误差	90	4.61	6.1
	逆时针旋转180°	角度误差	80	7.17	8.7

5 总结与讨论

本文围绕航天员助手机器人灵巧手在微重力环境下的高精度遥操作问题，提出了一种融合视觉运动重定向与卡尔曼滤波的控制方法，有效解决了因视觉感知噪声、信号抖动和系统延迟导致的控制不稳定问题. 通过在AAR-2机器人平台上进行系统实验, 验证了该方法在多种操作任务中的有效性、稳定性和优越性. 实验结果表明, 本文方法在常见物体抓取任务中成功率高达100%(简易抓捕)和80%以上(精细操作), 平均完成时间短, 表现出极高的操作效率和可靠性; 在操作任务板交互任务中，系统在旋钮旋转、按钮按压和把手拉动等操作中均保持高成功率与低误差，尤其在连续操作中表现出良好的平滑性和实时性.

本文通过联合状态建模、视觉关键点级联平滑及基于预测的延迟补偿，充分利用视觉输入特性与机械动态特性，实现舱内漂浮式助手机器人的高精度、低抖动、平滑实时遥操作. 实验结果表明，该方法在旋钮操作、按钮按压及把手操作中均保持高成功率与轨迹连续性, 验证了其工程适用性与创新性.

本方法的核心优势在于: (1) 视觉部分采用了深度信息加权融合优化策略，有效克服了遮挡、光照变化和噪声干扰; (2) 通过卡尔曼滤波对关节角度与角速度进行联合状态估计，有效抑制了视觉输入中的高频噪声和异常跳变; (3) 利用状态预测机制补偿系统延迟，显著增强了操作的实时性和流畅性; 整体算法在保证精度的同时具备较低计算复杂度，可扩展至视觉结合力控的远程自主或半自主操作场景，为微重力环境下高精度遥操作提供可靠技术方案.

视觉结合力控实现远程自主或半自主操作是航天员助手机器人发展的重要方向. 协作机器人领域已有大量工作采用视觉粗定位+力/阻抗控制策略，实现柔顺抓取与接触保持. 在微重力舱内，由于动力学不确定性、视角遮挡和计算资源限制，实现完全自主仍存在挑战. 本文提出的人手-灵巧手运动重定向与状态预测框架具有良好扩展性: 可在遥操作控制外层增加局部自主策略，如抓取过程中的夹持力自适应调节、滑移检测及接触保持, 为未来视觉+力控融合提供基础.

本文为航天员助手机器人在轨服务、舱内操作与人机协作任务提供了一项可靠的技术方案. 未来工作将进一步优化视觉算法在极端环境下的适应性，并探索与强化学习等智能方法的结合，以提升系统的自主决策与泛化.

参考文献

1 Ding X L, Gao H B, Huang P F, et al. Booming space robot technology and applications (in Chinese). Robot, 2022, 44: 1 [丁希仑, 高海波, 黄攀峰, 等. 蓬勃发展的空间机器人技术与应用. 机器人, 2022, 44: 1]

2 Xue Z H, Liu J G. Review of space manipulator control technologies (in Chinese). Robot, 2022, 44: 107–128 [薛智慧, 刘金国. 空间机械臂操控技术研究综述. 机器人, 2022, 44: 107–128]

3 Wang Y J, Sun F C, Liu H P, et al. Control scheme and Kinematics analysis for space teleoperation based on 6-DOF decoupled manipulator (in Chinese). Inf Control, 2012, 41: 38–43 [王裕基, 孙富春, 刘华平, 等. 基于6-DOF解耦操作臂的空间遥操作控制及运动学分析. 信息与控制, 2012, 41: 38–43]

4 Wang X H. Space on-orbit service technology and current development status and trends (in Chinese). Satell Netw, 2016: 70–76 [王晓海. 空间在轨服务技术及发展现状与趋势. 卫星与网络, 2016: 70–76]

5 Amaya-Mejía L M, Ghita M, Dentler J, et al. Visual servoing for robotic on-orbit servicing: A survey. In: Proceedings of 2024 International Conference on Space Robotics (iSpaRo). Piscataway: IEEE, 2024. 178–185

6 Yang G, Ji J, Liu Y, et al. A novel visual servo method for space manipulators by planning jerk. Adv Space Res, 2023, 71: 2591–2608

7 Liu H, Liu Y, Jiang L. Space robot and its teleoperation (in Chinese). Harbin: Harbin Institute of Technology Press, 2012 [刘宏, 刘宇, 姜立. 空间机器人及其遥操作. 哈尔滨: 哈尔滨工业大学出版社, 2012]

8 Ding X L, Chen Y T, Wang C C, et al. Research status and prospect of space robot operation technology (in Chinese). Chin J Aeronaut, 2025, 46: 294–322 [丁希仑, 陈一同, 王成才, 等. 空间机器人操作技术研究现状与展望. 航空学报, 2025, 46: 294–322]

9 Liang B, Wang X Q, Chen Z. Space robots teleoperation: System and Control (in Chinese). Beijing: People's Posts and Telecommunications Press, 2020 [梁斌, 王学谦, 陈章. 空间机器人遥操作系统及控制. 北京: 人民邮电出版社, 2020]

10 Ghasemi A, Li P, Xie W F. Adaptive switch image-based visual servoing for industrial robots. Int J Control Autom Syst, 2020, 18: 1324–1334

11 Ding W, Zhang F, Cui L, et al. Visual image processing and Image stitching method for flapping-wing flying robot based on clustering analysis (in Chinese). Inf Control, 2024, 53: 416–432 [丁伟, 张峰, 崔龙, 等. 基于聚类分析的扑翼飞行机器人视觉图像处理和图像拼接方法. 信息与控制, 2024, 53: 416–432]

12 Zhu Y, Gao Q, Shi H, et al. Gestures recognition based on multimodal fusion by using 3D CNNs. J Intell Fuzzy Syst, 2024, 46: 1647–1661

13 Liu H, Li Z Q, Liu Y W, et al. Key technologies of TianGong-2 robotic hand and its on-orbit experiments (in Chinese). Sci Sin Tech, 2018, 48: 1313–1320 [刘宏, 李志奇, 刘伊威, 等. 天宫二号机械手关键技术及在轨试验. 中国科学: 技术科学, 2018, 48: 1313–1320]

14 Hu C W, Gao S, Xiong M H, et al. Key technologies of the China space station core module manipulator (in Chinese). Sci Sin Tech, 2022, 52: 1299–1331 [胡成威, 高升, 熊明华, 等. 空间站核心舱机械臂关键技术. 中国科学: 技术科学, 2022, 52: 1299–1331]

15 Liang B, Huang X X, Liu T L, et al. Design and force-position-model fusion control of a hybrid active and passive cable cluster-driven hyper-redundant space manipulator (in Chinese). Sci Sin Tech, 2024, 54: 2319–2336 [梁斌, 黄学祥, 刘天亮, 等. 绳簇主被动混合驱动空间机械臂设计及其力-位-型融合控制. 中国科学: 技术科学, 2024, 54: 2319–2336]

16 Zhang Z, Liu C K, Wang M M, et al. Development and prospects of space intelligent operation (in Chinese). Sci Sin Tech, 2024, 54: 289–303 [张哲, 刘传凯, 王明明, 等. 空间作业智能操控技术研究与展望. 中国科学: 技术科学, 2024, 54: 289–303]

17 Wang X Q, Liang B, Xu W F, et al. Ground-based verification technology for teleoperation of space robot (in Chinese). Robot, 2009, 31: 8–14 [王学谦, 梁斌, 徐文福, 等. 空间机器人遥操作地面验证技术研究. 机器人, 2009, 31: 8–14]

18 Han D, Nie H, Chen J B, et al. The control system of the tendon-driven space five-fingered dexterous hand (in Chinese). Robot, 2016, 38: 129–134 [韩冬, 聂宏, 陈金宝, 等. 腱驱动空间五指灵巧手控制系统研究. 机器人, 2016, 38: 129–134]

19 Ajoudani A, Zanchettin A M, Ivaldi S, et al. Progress and prospects of the human–robot collaboration. Auton Robot, 2018, 42: 957–975

20 Handa A, Van Wyk K, Yang W, et al. Dexpilot: Vision-based teleoperation of dexterous robotic hand-arm system. In: Proceedings of 2020 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2020. 9164–9170

21 Gao Q, Ju Z, Chen Y, et al. An efficient RGB-D hand gesture detection framework for dexterous robot hand-arm teleoperation system. IEEE Trans Hum-Mach Syst, 2022, 53: 13–23

22 Zhang R, Wang Z, Zhang Y. Astronaut visual tracking of flying assistant robot in space station based on deep learning and probabilistic model. Int J Aerospace Eng, 2018, 2018: 1–17

23 Eisenberg T, Reichert G, Christe R, et al. Assistant in Space. Cham: Springer, 2024

24 Zhang L X, Xiao G Z, Wang D Z, et al. Review and prospects of orbit-to-surface teleoperation (in Chinese). Sci Sin Tech, 2020, 50: 716–728 [张立宪, 肖广洲, 王东哲, 等. 在轨对星球表面遥操作技术现状与展望. 中国科学: 技术科学, 2020, 50: 716–728]

25 Tong Y, Liu J, Zhou H, et al. Adaptive tracking control of robotic manipulators with unknown kinematics and uncertain dynamics. IEEE Trans Automat Sci Eng, 2024, 21: 5252–5269

26 Wen H, Liu X, Yang Y. Position and force control in teleoperation systems with time-varying delays using dynamic event-triggered communication (in Chinese). Inf Control, 2025, 54: 110–122 [文辉, 刘霞, 杨勇. 基于动态事件触发通信的时变时延遥操作系统的位置和力控制. 信息与控制, 2025, 54: 110–122]

27 Daraviña G C, Valencia J L, Holguin G A, et al. Visual servoing and Kalman filter applied to parallel manipulator 3-RRR. Electronics, 2024, 13: 2703

28 Tao Q. Design and control of an indoor miniature autonomous blimp. Dissertation for Doctoral Degree. Atlanta: Georgia Institute of Technology, 2020

29 Meng G, Han L L, Zhang C F. Research progress and technical challenges of space robot (in Chinese). Chin J Aeronaut, 2021, 42: 523963 [孟光, 韩亮亮, 张崇峰. 空间机器人研究进展及技术挑战. 航空学报, 2021, 42: 523963]

30 Liu H. Compliant grabbing of dexterous hand in complex scenarios for a visual-guided collaborative robot (in Chinese). Dissertation for Master's Degree. Harbin: Harbin Institute of Technology, 2022 [刘海超. 基于视觉引导的协作机器人灵巧手复杂场景下的柔顺抓取. 硕士学位论文. 哈尔滨: 哈尔滨工业大学, 2022]

31 Chou W S, Meng C, Chen J X, et al. Robot assisted teleoperation of space scientific experiment (in Chinese). Chin Space Sci Technol, 2003, 6: 10–16 [丑武胜, 孟偲, 陈建新, 等. 空间科学实验机器人辅助遥操作系统. 中国空间科学技术, 2003, 6: 10–16]

32 Rodríguez-Martínez E A, Flores-Fuentes W, Achakir F, et al. Vision-based navigation and perception for autonomous robots: Sensors, SLAM, control strategies, and cross-domain applications—A review. Eng, 2025, 6: 153

33 Lugaresi C, Tang J, Nash H, et al. Mediapipe: A framework for perceiving and processing reality. In: Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019
34 Qi J, Ma L, Cui Z, et al. Computer vision-based hand gesture recognition for human-robot interaction: A review. Complex Intell Syst, 2024, 10: 1581–1606
35 Wameed M, M. Alkamachi A, Erçelebi E. Tracked robot control with hand gesture based on mediapipe. Al-Khwarizmi Eng J, 2023, 19: 56–71
36 Deng H, Birdal T, Ilic S. 3D local features for direct pairwise registration. In: Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019. 3244–3253
37 Khodarahmi M, Maihami V. A review on Kalman filter models. Arch Computat Methods Eng, 2023, 30: 727–747
38 Gao Q, Liu J, Tian T, et al. Free-flying dynamics and control of an astronaut assistant robot based on fuzzy sliding mode algorithm. Acta Astronaut, 2017, 138: 462–474

A study on teleoperation control methods for astronaut assistant robots based on visual redirection and Kalman filtering

WANG SiYue[1,2], YANG Yan[1,2], LIU JinGuo[1] & SHAO ZheMing[3]

[1] *State Key Laboratory of Robotics and Intelligent Systems, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China*
[2] *University of Chinese Academy of Sciences, Beijing 110049, China*
[3] *Liaoning University of Technology, Jinzhou 121001, China*

With the growing demand for astronaut-assistant robots to perform intravehicular tasks in microgravity environments, high-precision teleoperation control of dexterous hands has become a key enabling technology for executing complex operations. This technology directly determines the accuracy and efficiency of space missions. However, during vision-based teleoperation, hand keypoint detection is highly susceptible to disturbances from dynamic cabin environments. This often results in jitter and short-term anomalies in the detection output. Directly mapping such noisy visual data to a dexterous hand degrades motion stability and increases response latency, severely limiting the reliability and precision of teleoperation control. To address these challenges, this paper proposes a teleoperation control method based on visual motion retargeting and Kalman filtering. First, a vision perception module is employed to accurately extract the 3D coordinates of 21 human hand keypoints. By incorporating the structural characteristics of both the human hand and the dexterous hand, a multi-degree-of-freedom motion mapping model is constructed. This model transforms human hand joint motions into corresponding sequences of bending and rotational angles for the dexterous hand. Subsequently, the joint angles and angular velocities of the dexterous hand are jointly estimated as system states. An exponentially weighted moving average (EWMA) algorithm is integrated to suppress abrupt variations in the keypoint data. Meanwhile, the state prediction capability of the Kalman filter is utilized to compensate for teleoperation latency. This process ultimately generates real-time and stable control commands for the dexterous hand. Experimental results demonstrate that the proposed method significantly reduces joint jitter and improves both motion trajectory smoothness and system response speed. It also exhibits strong stability and adaptability under uncertain disturbances. These findings validate its potential for application in complex space tasks.

astronaut assistant robot, teleoperation control, dexterous hand, pose detection, retargeting, Kalman filtering

doi: 10.1360/SST-2025-0256