

# 基于多模态感知的机械臂遥操作系统设计与实现

冯莹盈, 贾丹平\*, 赵璐, 刘振宇

沈阳工业大学信息科学与工程学院, 辽宁 沈阳

收稿日期: 2026年3月22日; 录用日期: 2026年5月17日; 发布日期: 2026年5月28日

## 摘要

针对传统手部动作感知系统依赖单一传感模态、穿戴舒适性差等问题, 设计了一种轻量化、无线化的多模态感知系统, 并将其应用于机械臂遥操作。该系统由集成弯曲与压力传感器的柔性数据手套及八通道前臂肌电手环构成, 通过双蓝牙架构实现运动学与生理学数据的亚毫秒级同步采集。在遥操作控制中, 利用手环内置惯性测量单元(IMU)解算前臂姿态, 控制机械臂空间位姿; 基于数据手套手势识别驱动灵巧手完成抓取动作。实验结果表明, 该系统在抓取任务中成功率达100%, 验证了其在精细操作与实时交互中的可靠性与有效性。为自然人机交互提供了一种轻量化、高集成度的技术方案。

## 关键词

多模态感知, 数据手套, 肌电信号, 机械臂遥操作

# Design and Implementation of a Robotic Arm Teleoperation System Based on Multimodal Perception

Yingying Feng, Danping Jia\*, Lu Zhao, Zhenyu Liu

School of Information Science Engineering, Shenyang University of Technology, Shenyang Liaoning

Received: March 22, 2026; accepted: May 17, 2026; published: May 28, 2026

## Abstract

To address the limitations of single-modal hand motion perception systems, this paper designs a lightweight, wireless multimodal sensing system and applies it to robotic arm teleoperation. The

\*通讯作者。

文章引用: 冯莹盈, 贾丹平, 赵璐, 刘振宇. 基于多模态感知的机械臂遥操作系统设计与实现[J]. 人工智能与机器人研究, 2026, 15(3): 867-875. DOI: 10.12677/airr.2026.153080

system consists of a flexible data glove integrated with bend and pressure sensors, and an eight-channel forearm EMG wristband. A dual-Bluetooth architecture enables sub-millisecond synchronous acquisition of kinematic and physiological data. In the teleoperation framework, the wristband's IMU estimates forearm posture to control the robotic arm's spatial position, while gesture recognition from the data glove drives the dexterous hand for grasping. Experiments show a 100% success rate in grasping tasks and a response latency of about 120 ms in dynamic interactions, validating the system's reliability and real-time performance. This work offers a lightweight, highly integrated solution for natural human-robot interaction.

## Keywords

Multimodal Perception, Data Glove, SEMG, Robotic Arm Teleoperation

Copyright © 2026 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

近年来,随着自然人机交互[1]技术的发展,基于可穿戴传感器[2]的手部动作意图识别[3]在机器人遥操作等领域展现出巨大的应用潜力。然而,现有感知系统多依赖单一传感模态,存在意图预判困难或易受干扰等局限,且传统的多模态融合设备[4]往往面临硬件体积庞大、线缆繁琐及穿戴舒适性差等工程化痛点,难以满足真实场景下的高机动性需求。针对上述问题,本文设计并实现了一种轻量化、无线化的多模态手部动作感知系统。该系统创新性地将集成多路弯曲与压力传感器的数据手套与多通道前臂肌电手环[5]相结合,通过定制化的高集成采集电路与双蓝牙通信架构,实现了运动学与生理学特征的优势互补以及数据的低延迟同步传输。在此基础上,本文进一步构建了“手环导航-手套操作”的仿生控制框架,并在机械臂[6]精细抓握与高动态人机交互实验中,充分验证了该系统在复杂任务执行中的高可靠性与工程落地价值。

## 2. 系统总体架构

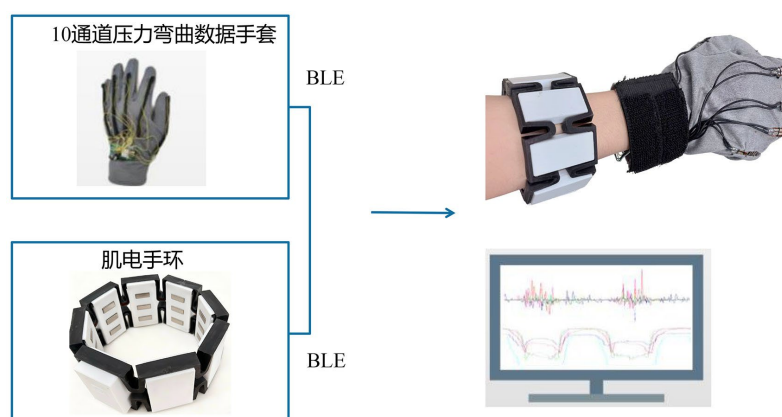


Figure 1. Multimodal sensing system

图 1. 多模态感知系统

面向机器人遥操作的高机动性需求,设计了一套轻量化、无线化的多模态感知系统如图 1 所示。系

统由捕捉手指运动学特征的“多传感数据手套”与获取前臂生理学信息的“多通道肌电(sEMG)手环”构成。系统数据流向包含信号感知、调理与传输三层：前端柔性传感器与肌电电极实时捕获手部形态与肌肉电生理变化；原始微弱信号经高集成电路完成放大与滤波去噪调理；随后经微控制器(MCU)模数转换，依托双蓝牙节点将多源异构数据低延迟同步发送至上位机。该无缆化架构为后端的特征融合与意图识别奠定了硬件基础。

### 2.1. 柔性数据手套设计

柔性数据手套以轻量化、高透气性的麂皮材质为载体，整只手套经过精密剪裁与缝制，重量控制在40克以内，实现了近乎无感的佩戴体验，在保证机械强度的同时，兼顾了柔韧性、透气性与长期穿戴舒适性，可满足长时间连续使用场景下的用户需求。传感布局严格遵循“背侧运动感知-掌侧交互感知”的仿生学原理，构建了覆盖运动学与动力学的完整信息采集体系，为后续手势识别提供了丰富的数据基础。在手套背侧，沿五指掌指关节(MCP)至远端指间关节(DIP)的解剖轴线，精确贴合5路柔性弯曲传感器(Flex Sensor 2.2)。该传感器基于可变电阻原理，当手指屈曲时其电阻值随弯曲角度增大而相应增加，通过标定实验建立了角度-电阻的精确映射关系，弯曲寿命可达百万次以上，能够适应长期高频使用场景。实际应用中，传感器贴合于手指表面，使其与手指屈伸活动同步弯曲，引起电阻值连续变化，从而实现对手指屈伸运动轨迹的连续、高精度量化监测。在掌侧五指指腹区域，嵌入5路微型薄膜压力传感器(FSR RP-C10-LT)，以网格化形式实时捕捉抓握、捏取等操作中的接触压力分布。该传感器基于力敏电阻原理，敏感区域外径10mm，厚度仅0.4mm，具备20g至2kg的宽范围压力检测能力与低于10ms的快速响应特性，重复测量精度达 $\pm 2\%$ 。当手指与物体接触时，传感器电阻值随压力增大而非线性递减，经信号采集电路转换为电压信号后，可精确反映接触力的大小与分布变化。该设计在保障信号采集可靠性的同时，最大程度减少了对自然触觉感知的干扰，为手势交互中的力度控制与接触状态识别提供了关键数据支撑，数据手套设计如图2所示。



Figure 2. Data glove  
图2. 数据手套

### 2.2. 表面肌电手环设计

为实现对手部动作神经驱动信息的同步采集，系统集成自主研发的多通道表面肌电(sEMG)手环。该手环以轻量化、低功耗为设计原则，采用柔性可伸展软体材料作为基底，兼顾长期佩戴的舒适性与运动形变适应性，并支持便捷清洁维护。硬件架构以nRF52840低功耗蓝牙芯片为核心处理单元，集成了8

通道医疗级银电极与前端预处理电路。电极贴合于前臂桡侧腕屈肌、指浅屈肌等关键肌群，拾取肌肉活动产生的微弱电生理信号，原始信号幅值通常为微伏级，经前置放大、带通滤波(20~500 Hz)及电压抬升等关键调理环节后，有效抑制工频噪声、运动伪影及肌颤干扰，信噪比显著提升。芯片内置 12 位模数转换器(ADC)完成数字化采样，支持 1000 Hz 与 500 Hz 双采样率可选。在无线传输方面，手环采用优化的批量传输策略，平均每秒发送约 62 个数据包，每个数据包完整封装 16 个连续采样时刻的 8 通道原始肌电信号，通过低功耗蓝牙(BLE)模块将缓存数据实时上传至上位机。该设计在有限无线带宽下实现了多通道肌电信号的高精度采集与可靠传输。肌电手环实物如图 3 所示及信号预处理流程如图 4 所示。



Figure 3. SEMG wristband  
图 3. 肌电手环

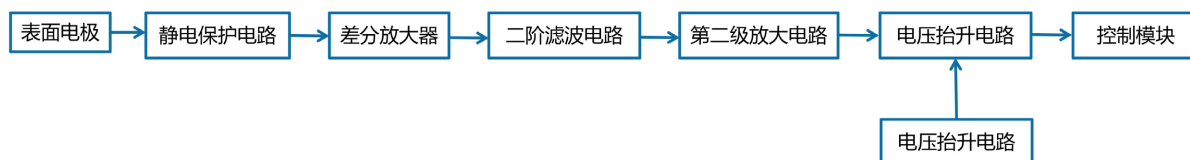


Figure 4. Preprocessing circuit process circuit  
图 4. 预处理电路流程电路

### 2.3. 多模态数据同步与无线通信

为实现运动学与生理学特征的亚毫秒级时空对齐，系统采用双蓝牙并行通信架构。数据手套端基于 ESP32-WROOM-32D 模组[7]构建嵌入式采集系统，其下位机工作流程如图 5 所示。系统上电后完成串口通信、蓝牙协议栈初始化及广播参数配置，随后持续检测蓝牙连接状态。当客户端连接建立后，系统进入指令响应循环，通过命令特征值(UUID:19b10005)接收上位机控制指令。接收到 0x01 启动采集指令后，系统进入高速数据采集与缓存阶段：以 1000 Hz 采样频率轮询采集 10 路传感器信号(5 路弯曲 + 5 路压力)，每次完整采集 10 路数据后存入循环缓冲区，累积完成 25 轮采集(总计 250 个数据点)后缓冲区达到满载状态，随后经数据特征值(UUID:19b10004)以 GATT 通知方式批量发送至上位机，实现约 1000 Hz 的有效采样率。发送完成后缓冲区清零，系统继续等待下一次采集指令。

肌电手环端基于 nRF52840 芯片，同样以 1000 Hz 采样频率工作，采用批量传输策略，平均每秒发送约 62 个数据包，每个数据包完整封装 16 个连续采样时刻的 8 通道原始肌电信号。系统通过硬件同步脉冲触发与软件高精度时间戳校正协同机制，结合动态时钟漂移补偿算法，确保了两类异构数据在时间维度上的亚毫秒级严格对齐。该双节点无线通信架构彻底摆脱了传统有线束缚，在保证数据实时性与同步精度的同时，显著提升了穿戴设备的移动操作自由度与场景适应性。

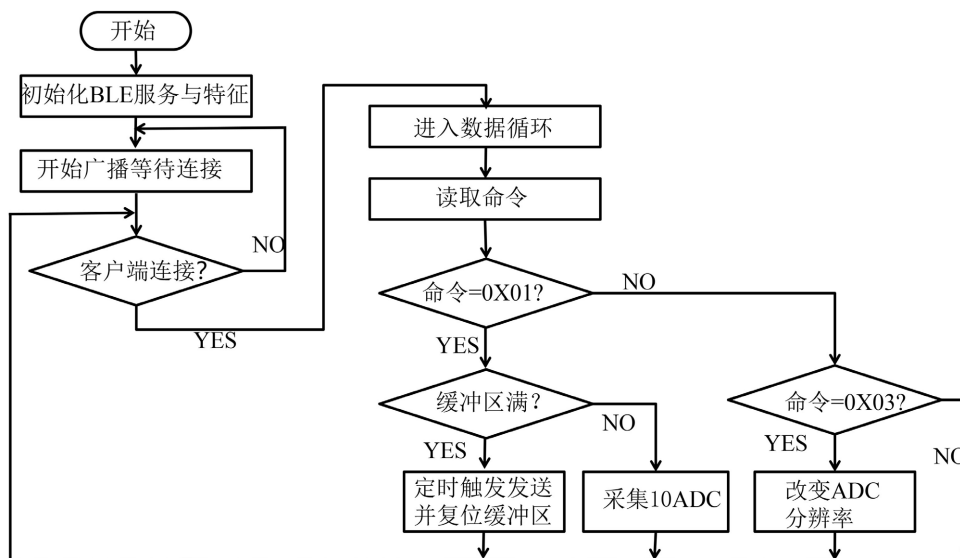


Figure 5. Flow chart of lower computer acquisition

图 5. 下位机采集流程图

### 3. 基于多模态感知机械臂遥操作

#### 3.1. 遥操作控制策略

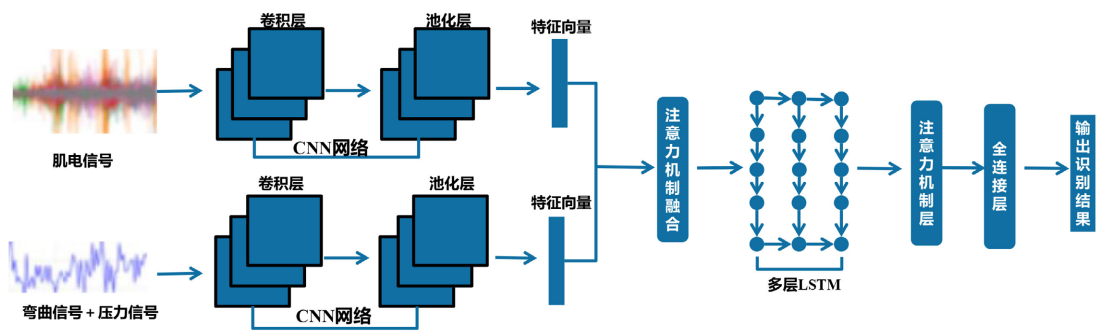
为实现对机械臂的自然、直观控制，本文构建了“手环导航 - 手套操作”的仿生控制框架，将人体上肢运动与手部动作分别映射为机械臂的空间位姿与灵巧手抓取动作，形成粗粒度导航与细粒度操作的分层控制策略。

导航控制(粗粒度)：利用肌电手环内置的惯性测量单元(IMU)实时采集前臂在三维空间中的姿态变化(俯仰、偏航、横滚)，通过惯性导航解算获得手臂运动的方向与速度信息。系统将手臂姿态角线性映射为机械臂末端执行器的平移速度指令，实现机械臂在三维空间中的自然导航。该控制方式无需外部视觉或定位系统，为用户提供了无约束、高自由度的操作体验。

操作控制(细粒度)：基于数据手套采集的 5 路弯曲与 5 路压力信号，通过多模态融合模型实时识别用户手势。系统将识别结果映射为灵巧手的抓取动作指令：手势“five”(五指张开)对应灵巧手完全张开，手势“fist”(握拳)对应灵巧手完全闭合，手势“two”(剪刀手)对应灵巧手做出剪刀姿态。该映射策略使得用户能够通过自然手势完成对目标物体的抓取、搬运与释放。

#### 3.2. 手势识别算法模型设计

为了克服决策级融合中各模态信号关联性挖掘不足及信息利用过于松散的问题，本研究转向特征级融合范式，构建了卷积 - 注意力 - LSTM (CNN-Attention-LSTM)混合架构，如图 6 所示。该模型采用双流并行输入结构，旨在实现异构数据在特征层面的深层关联与动态协同。其中，“组合运动流”合并了 5 通道弯曲与 5 通道压力信号，维度设定为(30, 10)，用于表征手部空间形态与力学交互的内在同步性；“生理特征流”则输入 8 通道表面肌电(sEMG)信号，维度为(30, 8)，侧重于反映神经肌肉驱动层面的电生理活动。通过卷积层(1D-CNN)对双流信号进行局部时空特征提取，并利用通道注意力机制(Channel Attention)根据不同手势上下文对各模态贡献度进行动态权重分配，该架构能有效弥补单一模态在表征复杂手势意图时的信息缺失，从而显著提升系统对跨模态动态关联的解析能力。

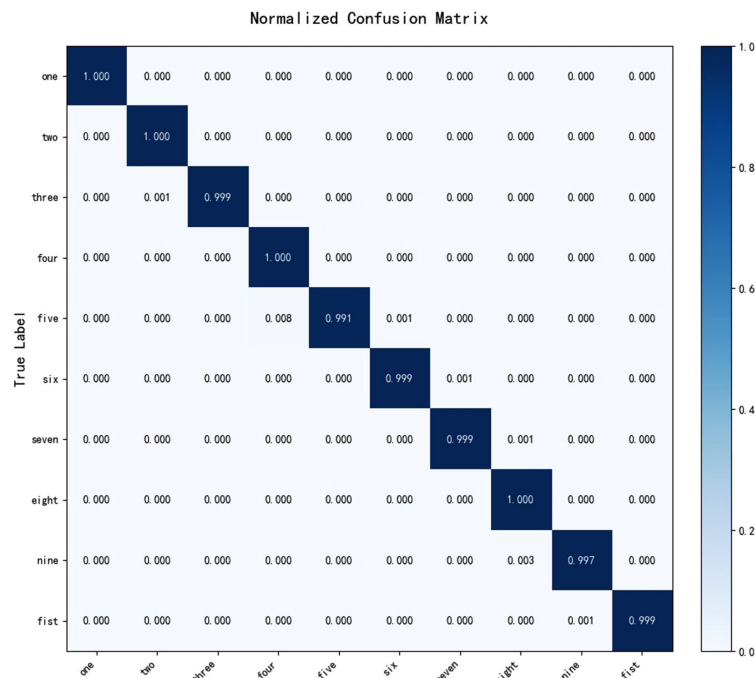


**Figure 6.** Dual-Stream CNN-LSTM network with attention-based fusion  
**图 6.** 基于注意力融合的双流 CNN-LSTM 网络

本研究构建的多模态融合网络采用双流并行架构，输入层分别为 8 通道肌电信号(30, 8)与 10 通道弯曲 - 压力组合流(30, 10)。特征提取层选用 64 个核大小为 5 的一维卷积核，配合 ReLU 激活、批量归一化及步长为 2 的最大池化层，实现时域特征的 4 倍降采样。随后，通过通道注意力机制对拼接后的 128 维融合特征进行动态权重分配，并输入至三层堆叠的 LSTM 网络(每层 128 个隐藏单元，Dropout 率为 0.2)进行时序建模。最终经由时序注意力机制生成的 128 维上下文向量，通过三层全连接网络(神经元数分别为 256、128、64)与 Softmax 层实现 10 类手势分类。模型训练采用 Adam 优化器，学习率设为 0.001，批量大小为 32，数据集按 4:1 比例随机划分。

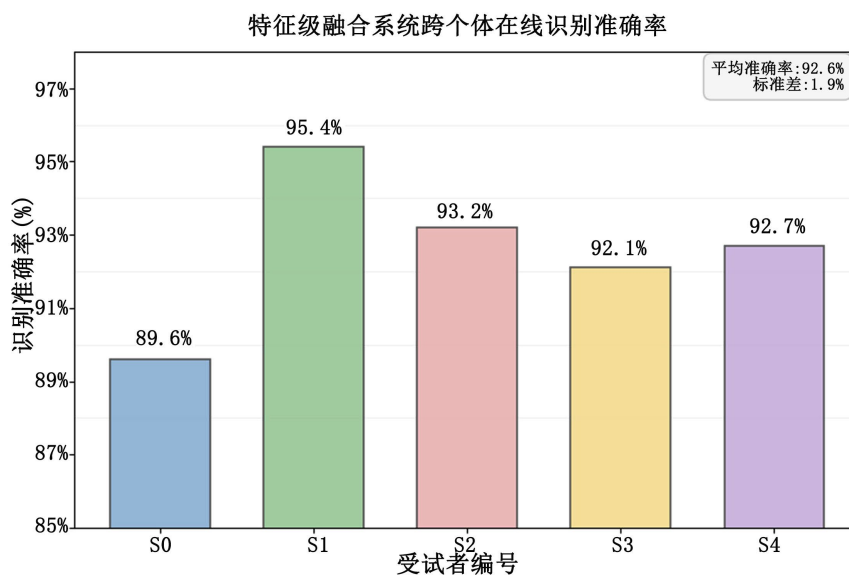
### 3.3. 手势识别实验验证与结果分析

为评估卷积 - 注意力 - LSTM 模型的性能，本文在包含 10 类手势的标准数据集上进行了离线训练与测试。实验结果显示，该特征级融合模型在测试集上的平均识别准确率达到 99.55%，混淆矩阵如图 7 所示。



**Figure 7.** Confusion matrix  
**图 7.** 混淆矩阵

为进一步验证系统的工程泛化能力,开展了面向 5 名新用户的跨个体在线实验,结果如图 8 所示。在不引入任何针对新用户的校准或微调策略的情况下,系统实现了“开箱即用”的高效识别。实验统计显示,系统平均在线识别准确率达 92.6%,最高个体准确率可达 95.4%。相比于传统的决策级融合系统,充分证明了通道注意力机制在自适应调节模态权重、削弱个体差异干扰方面的有效性。



**Figure 8.** Distribution of online recognition accuracy across subjects

**图 8.** 跨个体在线识别准确率分布

### 3.4. 遥操作实验平台搭建及系统验证

实验平台硬件由三部分组成: (1) 多模态感知端: 集成弯曲与压力传感的数据手套, 以及集成了表面肌电传感器与惯性测量单元的前臂手环; (2) 计算与控制端: 一台高性能工控机, 运行多模态融合识别模型, 并生成控制指令; (3) 执行端: 一台六自由度机械臂及其末端搭载的五指灵巧手。系统通过双蓝牙将多模态数据实时上传至上位机, 上位机完成手势识别与姿态解算后, 通过网络协议将控制指令下发至机械臂控制器, 如图 9 所示。



**Figure 9.** Robotic end-effector: the 6-DOF manipulator and five-fingered dexterous hand

**图 9.** 机器人执行端: 六自由度机械臂与五指灵巧手

### 3.5. 抓取搬运任务验证

为验证系统在精细操作任务中的有效性，设计了“人传物-机搬运”的完整闭环任务：

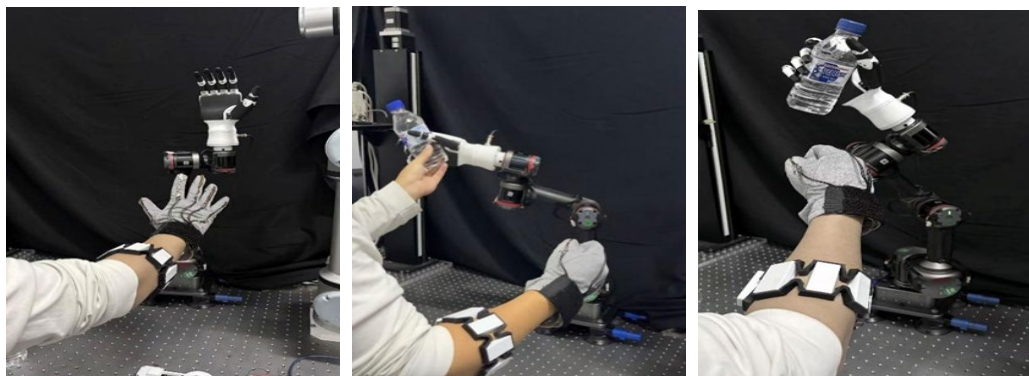
(1) 准备与装载阶段：操作者通过前臂运动控制机械臂，将灵巧手导航至固定准备位置，并做出“five”手势使灵巧手保持张开状态。随后，操作者用另一只自由手将一个 350 ml 的矿泉水瓶放置于灵巧手掌心。

(2) 抓握与确认阶段：装载完成后，操作者对数据手套做出“fist”手势，系统识别后控制灵巧手闭合，稳定抓取矿泉水瓶。

(3) 搬运阶段：操作者通过手臂运动控制机械臂，将矿泉水瓶平稳搬运至 50 cm 外的指定目标区域上方。

(4) 释放阶段：到达目标位置后，操作者做出“five”手势，系统识别后控制灵巧手张开，将水瓶释放至目标区域内。

实验共进行 10 次重复测试，任务成功率达到 100%，平均单次任务完成时间约 45 秒，验证了该系统在精细操作任务中的可靠性与实用性实验如图 10 所示。



**Figure 10.** Key collaborative steps in teleoperated grasping. (a) Gesture control preparation; (b) Manual object transfer; (c) Gesture-triggered stable grasping of the dexterous hand

**图 10.** 遥操作抓取任务中人机协作关键步骤。(a) 手势控制准备；(b) 人工物体传递；(c) 手势触发灵巧手稳定抓取

## 4. 总结

本文设计并实现了一种轻量化、无线化的多模态感知系统，并将其成功应用于机械臂遥操作场景。该系统由集成弯曲与压力传感器的柔性数据手套及八通道前臂肌电手环构成，通过双蓝牙并行通信架构，实现了运动学与生理学数据的亚毫秒级同步采集与无线传输。在此基础上，构建了“手环导航-手套操作”的仿生控制框架，利用手环内置惯性测量单元(IMU)解算前臂姿态，控制机械臂空间位姿；同时基于数据手套的手势识别结果驱动灵巧手完成抓取动作。实验结果表明，该系统在抓取搬运任务中成功率达 100%，验证了其在精细操作与实时交互中的可靠性与有效性。本文工作为自然人机交互与机器人遥操作提供了一种轻量化、高集成度的技术方案，现将其与现有的遥操作感知系统进行对比，如表 1 所示。

**Table 1.** Comparison of technical schemes for different teleoperation perception systems

**表 1.** 不同遥操作感知系统技术方案对比

参考文献	遥操作感知模式	硬件形态	无线方案
Fang B 等[6]	IMU + 磁力计	重型手套	蓝牙 2.0
J. Li 等[8]	sEMG + 视觉	穿戴 + 外部相机	WiFi/有线
Yu, Z 等[9]	弯曲 + 光学捕捉	>300 g	ZigBee/光捕
本文系统	Semg + 弯曲 + 压力	轻量柔性穿戴	双蓝牙并行

通过与上述典型研究的横向对比可以发现,本系统在保障感知深度的前提下,显著提升了遥操作过程中的佩戴舒适度与环境鲁棒性。相比于文献[6]等受电磁干扰限制的磁力计方案,本文的“生理-运动-力学”全模态感知不受物理场环境影响,可靠性更高;相较于文献[8][9]等依赖外部相机或高成本光学捕捉设备的方案,本系统彻底摆脱了视觉遮挡与操作空间的限制,实现了真正的“无缆化”与“全场景适用”。此外,得益于双蓝牙并行架构与特征级融合算法,系统在保持约120ms低延迟响应的同时,实现了极高识别精度。综上所述,本系统在轻量化、高机动性及意图解析精度之间取得了优异的平衡,为复杂任务下的机器人自然遥操作提供了具备工程落地价值的技术路径。

## 基金项目

(1) 国家重点研发项目(2022YFF1202502): 多模态运动信息高效感知与意图精准识别技术;

(2) 辽宁省科技创新团队领军人才项目(LJ222410142060): 基于多模态多传感器融合导航的复合机器人系统研究。

## 参考文献

- [1] Shushtari, M., Foellmer, J. and Arami, A. (2024) Human-Exoskeleton Interaction Portrait. *Journal of NeuroEngineering and Rehabilitation*, **21**, Article No. 152. <https://doi.org/10.1186/s12984-024-01447-1>
- [2] Singh, R., Mozaffari, S., Akhshik, M., Ahamed, M.J., Rondeau-Gagné, S. and Alirezaee, S. (2023) Human-Robot Interaction Using Learning from Demonstrations and a Wearable Glove with Multiple Sensors. *Sensors*, **23**, Article 9780. <https://doi.org/10.3390/s23249780>
- [3] Zhao, S., Cai, H., Li, W., Liu, Y. and Liu, C. (2021) Hand Gesture Recognition on a Resource-Limited Interactive Wristband. *Sensors*, **21**, Article 5713. <https://doi.org/10.3390/s21175713>
- [4] Khaleghi, B., Khamis, A., Karray, F.O. and Razavi, S.N. (2013) Multisensor Data Fusion: A Review of the State-of-the-art. *Information Fusion*, **14**, 28-44. <https://doi.org/10.1016/j.inffus.2011.08.001>
- [5] Shen, C., Pei, Z., Chen, W., Wang, J., Zhang, J. and Chen, Z. (2022) Toward Generalization of sEMG-Based Pattern Recognition: A Novel Feature Extraction for Gesture Recognition. *IEEE Transactions on Instrumentation and Measurement*, **71**, 1-12. <https://doi.org/10.1109/tim.2022.3141163>
- [6] Fang, B., Sun, F., Liu, H. and Guo, D. (2017) A Novel Data Glove Using Inertial and Magnetic Sensors for Motion Capture and Robotic Arm-Hand Teleoperation. *Industrial Robot: An International Journal*, **44**, 155-165. <https://doi.org/10.1108/ir-07-2016-0179>
- [7] 刘昶江, 林媛媛, 刘梦杰. 基于ESP32芯片的智能物联电源设计[J]. 电子设计工程, 2024, 32(11): 141-145, 150.
- [8] Li, J., et al. (2022) A Multi-Modal Gesture Recognition Architecture Using sEMG and Skeletal Information. *Journal of Intelligent & Fuzzy Systems*, **43**, 643-656.
- [9] Yu, Z., Lu, C., Zhang, Y. and Jing, L. (2024) Gesture-controlled Robotic Arm for Agricultural Harvesting Using a Data Glove with Bending Sensor and Optitrack Systems. *Micromachines*, **15**, Article 918. <https://doi.org/10.3390/mi15070918>