

# 基于混合现实的可移动机械臂操作控制系统

董泽宇, 刘政, 李勇\*, 胡立坤

广西大学电气工程学院, 广西 南宁 530004

**摘要** 为应对特殊条件下不便于实地进行移动式操作的问题,设计了一套基于混合现实的可移动机械臂操作控制系统。该系统分为人机交互、机械驱动、虚拟现实等 3 个模块:人机交互模块通过摄像头识别操作者肢体手势动作并发布相应操作指令;机械驱动模块解析由人机交互模块发布的操作指令后完成相应动作并反馈设备工作状态;虚拟现实模块接收机械驱动模块反馈信息后将设备运行情况在构建好的虚拟场景中还原,实现对设备运行情况的实时监控。在可移动机械臂平台上的测试结果表明,操作者通过所提系统可以实现对可移动机械臂的远程精准控制与实时监控,系统响应速度能够达到 60~100 ms/frame。

**关键词** 机器视觉;混合现实;Unity3D;机械臂;动作捕捉

**中图分类号** TP29;TP39 **文献标志码** A

**DOI:** 10.3788/LOP223235

## Operation Control System for Movable Robotic Arm Based on Mixed Reality

Dong Zeyu, Liu Zheng, Li Yong\*, Hu Likun

School of Electrical Engineering, Guangxi University, Nanning 530004, Guangxi, China

**Abstract** To manage the problem regarding inconvenient mobile operations in the field under special conditions, we have designed a mobile robotic arm operation control system based on mixed reality. The system is categorized into three modules: human-computer interaction, mechanical drive, and virtual reality. The human-computer interaction module delivers corresponding operation instructions after recognizing the body gestures of the operator via camera. These operating instructions are then analyzed by the mechanical drive module that provides the working status of the equipment. Furthermore, this feedback information is received by the virtual reality module, which then restores the operation of the equipment in the built virtual scene to realize its real-time monitoring. The test results on a mobile robot platform show that the operator can achieve remote precise control and real-time monitoring of the mobile robot arm through the proposed system, and the system response speed can reach 60–100 ms/frame.

**Key words** machine vision; mixed reality; Unity3D; robotic arm; motion capture

## 1 引言

2022年5月,国内发生了多起自然灾害以及其他诸如火灾、毒气泄漏等突发事件<sup>[1]</sup>。然而,救援队进入或接近现场进行抢救工作的安全成本与技术难度极大。基于此,本文提出一种基于混合现实技术的远程机械臂操作控制系统。该系统通过搭建混合现实系统框架,辅助使用者对极端环境下工作的灾后救援机器人进行远程操纵与监测。

混合现实技术将可视化、人机交互和自动分析方

法整合为一体,支持用户以视觉的方式感知数据,通过交互方式对数据进行探索<sup>[2]</sup>。传统的人机交互功能依赖可视化交互设备实现可视化交互功能。其中,信息的输出通过二维显示屏实现、指令的输入功能则依赖键盘与鼠标实现。随着混合现实技术的发展,如今出现了更加自然直观的多模态人机交互方式,包含大数据可视化交互、基于声场感知的交互、混合现实实物交互、可穿戴设备交互、人机对话交互,以及其他人机交互方式<sup>[3]</sup>。人机交互的目的在于感知人类的交互意图,当前已有一些用于人机交互的设备。例如:

收稿日期: 2022-12-02; 修回日期: 2023-01-05; 录用日期: 2023-02-07; 网络首发日期: 2023-03-09

基金项目: 国家自然科学基金(61863002)、大学生创新创业训练计划国家级项目(202210593862)

通信作者: \*yongli@gxu.edu.cn

Pooryousef等<sup>[4]</sup>通过对柔性印制电路板进行设计开发,制作出依赖触觉进行交互的VR手套;Microsoft公司开发了能够应用在动作捕捉中的深度摄像头Kinect等<sup>[5]</sup>。然而,这些产品的成本相对较高,且在使用时会受到一定的程序限制。

为此,基于图像的人体关键点检测方法被提出,该方法将人体结构映射为若干关键点的几何拓扑关系,用于人机交互时具有良好的拓展性。目前关于人体关键点检测的研究主要包括基于堆叠漏斗神经网络的stacked hourglass网络模型与基于卷积神经网络和监督学习的open pose网络模型<sup>[6]</sup>。前者使用模块化网络设计,利用resnet思想使用跳级链接辅助采样,实现了先降采样后升采样的中继监督训练。后者则通过caffe框架搭建,通过down-up算法在定位各个关键点后匹配拼接关键点骨架。为应对系统需求,本文通过改进media pipe框架,实现单人3D关键点检测。

目前应用于机械臂的主动操作多为柔顺控制,分为基于机械结构实现的被动柔顺与通过反馈传感设备对机械臂的柔顺控制<sup>[7-8]</sup>。目前主要研究方向为通过自适应阻抗控制技术实现对电流环的动态调整<sup>[9]</sup>。两种控制方式对于操作环境的要求都极高,并不利于遥控

操作机械臂普及。如今新兴的人机交互方式同样需要特定的环境布置传感设备,其应用环境的局限性极大。

因此,为降低应用成本与环境因素限制,使得此类技术能够用于特殊环境下的机械臂操作,采用基于视觉的动作捕捉功能实现人机交互,采用Unity引擎构建虚拟仿真环境,同时通过机器人操作系统(ROS)实现无线数据通讯功能、指令的发布与接收,最终构建了一套基于混合现实的可移动机械臂操作控制系统。系统演示的视频链接为:<https://v.douyin.com/MB1A6ND/>。

## 2 基本原理与系统实现

### 2.1 总体架构

#### 2.1.1 软件系统

图1为基于混合现实的可移动机械臂操作控制系统的软件模块,包括人机交互、机械驱动、虚拟现实等3个功能模块,分别实现指令信号的输入、执行与反馈。不同模块间通过结合基于消息队列遥测传输协议(MQTT)的网络通讯通道与局域网通讯通道<sup>[10-11]</sup>,实现对指令信号的格式化、发布与接收,在保证系统良好稳定性的同时,具有良好的拓展性。

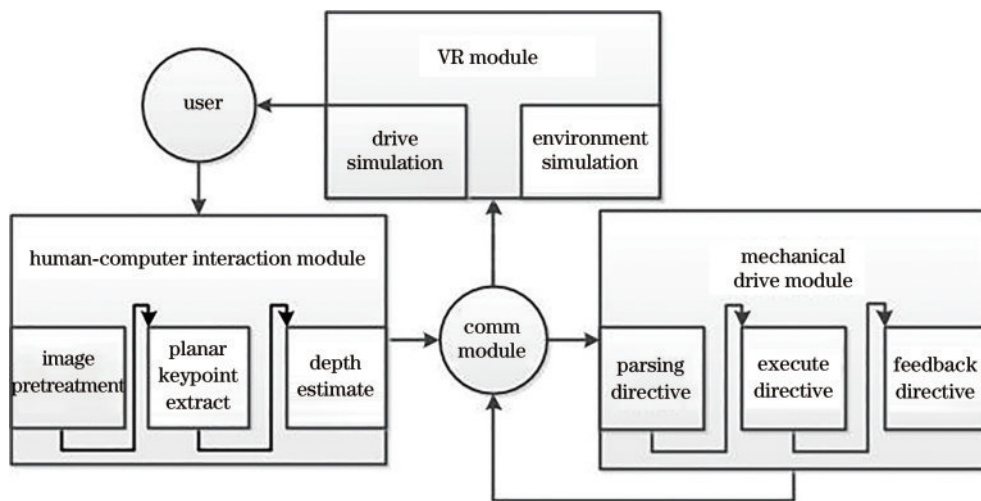


图1 软件系统架构

Fig. 1 Software system architecture

#### 2.1.2 硬件系统

基于混合现实的可移动机械臂操作控制系统硬件构成如图2所示:使用者通过个人计算机中搭载的人机交互模块向机械驱动模块发布指令;机械驱动模块接收指令并执行后,反馈设备的工作情况到虚拟现实模块;使用者通过虚拟现实模块实现对设备的可视化监控。

### 2.2 人机交互模块

肢体动作是一种便捷的人机交互方式。人机交互模块通过摄像头获取包含使用者肢体手势动作的图像数据,通过感知哈希算法滤波后提取使用者的关键点2D数据,最后对关键点进行深度估计,从而实现使用者肢体手势动作识别。

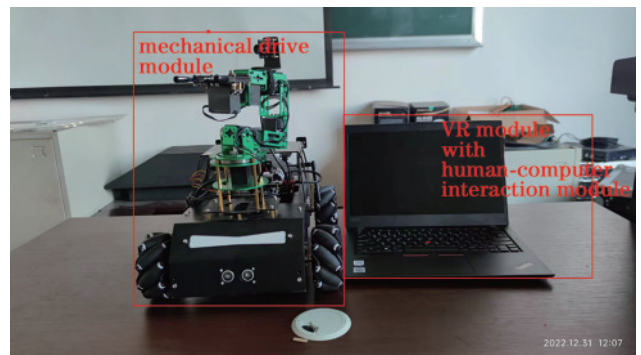


图2 硬件系统架构

Fig. 2 Hardware system architecture

2.2.1 感知哈希滤波

关键点估计对于设备处理器性能要求较高,而对视频数据进行实时目标检测时,摄像头、检测对象以及目标检测算法等多方面因素对于输出结果的精度影响极大。因此,所提系统利用感知哈希算法对输入视频进行预处理,减少关键点检测工作的任务量并减小因使用者抖动等原因产生的误差,提高系统输出指令的有效率与准确性。

感知哈希算法是一种判断图像相似性的一类哈希算法<sup>[12]</sup>。为保证该部分有效提高系统的处理速度,采用差异值哈希算法实现节流器功能,对比历史帧图像与输入帧图像,通过式(1)确定有效动作输出。算法流程如图3所示。

$$d_h = \text{sum} \left\{ \text{abs} \left[ \epsilon(I_0[0:9, 0:9] - I_0[1:10, 1:10]) - \epsilon(I_1[0:9, 0:9] - I_1[1:10, 1:10]) \right] \right\}, \quad (1)$$

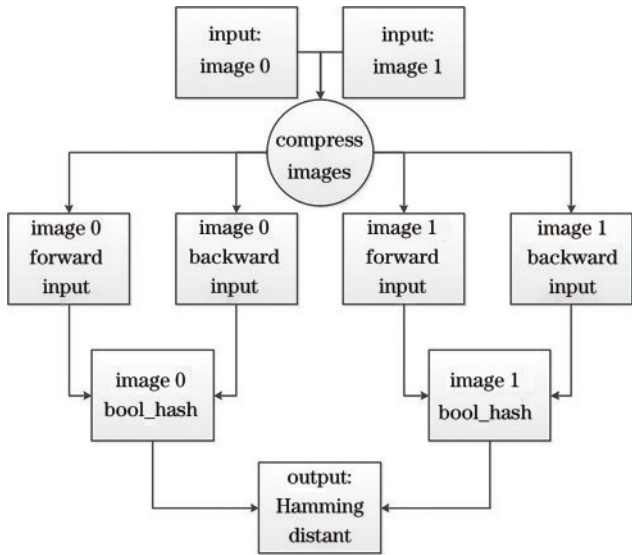


图3 d-Hash算法流程图

Fig. 3 Flowchart of d-Hash algorithm

式中: $I_0$ 为历史帧压缩图像, $I_1$ 为输入帧压缩图像, $d_h$ 为Hamming距离; $\epsilon$ 为阶跃函数; $\text{abs}(\cdot)$ 为求绝对值函数; $\text{sum}(\cdot)$ 为求和函数。

启动摄像头后系统保存第一帧图像作为待处理图像,随后计算摄像头读取图像与待处理图像的Hamming距离。其中,机械臂夹具控制图像差异值最小,且不同使用者多次测试该动作时Hamming距离都不小于5,因此选取Hamming距离为5作为输入图像预处理的阈值。当Hamming距离大于预设阈值5时,采用此帧摄像头读取图像替换待处理图像,并进行后续处理。

2.2.2 基于RGB图像的人体骨骼点坐标估计

所提系统通过media pipe框架<sup>[13]</sup>实现对使用者的关键点坐标提取。关键点坐标提取流程与分布分别如图4和图5所示。

系统在获得待处理图像后,首先定位使用者位置并与环境进行分割处理<sup>[13]</sup>,随后进行人体关键点坐标预测(pose\_landmarker),通过匹配输入图像与人体关

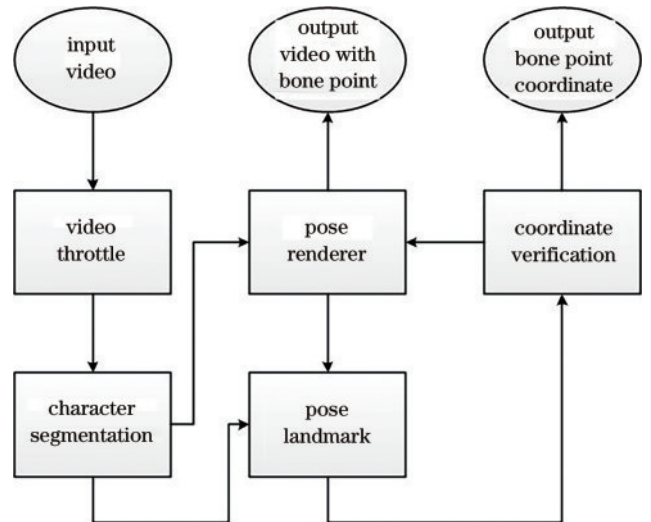
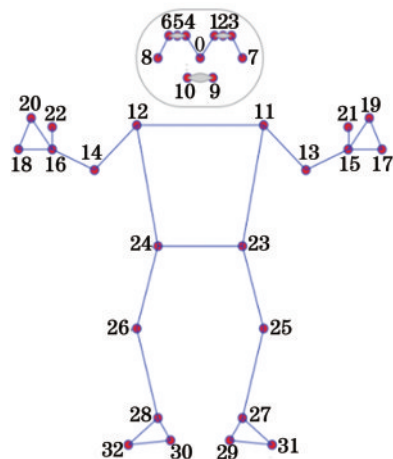


图4 关键点提取流程图

Fig. 4 Flowchart of key point extraction



- 0. nose
- 1. left\_eye\_inner
- 2. left\_eye
- 3. left\_eye\_outer
- 4. right\_eye\_inner
- 5. right\_eye
- 6. right\_eye\_outer
- 7. left\_ear
- 8. right\_ear
- 9. mouth\_left
- 10. mouth\_right
- 11. left\_shoulder
- 12. right\_shoulder
- 13. left\_elbow
- 14. right\_elbow
- 15. left\_wrist
- 16. right\_wrist
- 17. left\_pinky
- 18. right\_pinky
- 19. left\_index
- 20. right\_index
- 21. left\_thumb
- 22. right\_thumb
- 23. left\_hip
- 24. right\_hip
- 25. left\_knee
- 26. right\_knee
- 27. left\_ankle
- 28. right\_ankle
- 29. left\_heel
- 30. right\_heel
- 31. left\_foot\_index
- 32. right\_foot\_index

图5 关键点分布示意图<sup>[13]</sup>

Fig. 5 Schematic of the distribution of key points<sup>[13]</sup>



键点模型(关键点分布如图 5 所示)定位目标关键点在图像中的分布坐标,并通过坐标检测环节判断被遮挡部分关键点置信度。最后由图像渲染环节(pose\_renderer)接收由图像分割处理输出的图像数据流与坐标检测环节校正后输出的关键点分布坐标,在图像数据流中进行标记,最终输出标记了关键点的图像流与关键点坐标数据。

### 2.2.3 关键点深度估计

系统通过人体运动特征结合立体视觉几何原理,实现对关键点的深度估算<sup>[14]</sup>。基于图 6 所示的人体关键点分布,系统将 11、12、23、24 号关键点作为特征关键点,并将特征关键点组成四边形的几何特征建立特征关键点矩阵  $\mathbf{A}$ ,最后通过空间几何原理建立的式(2)~(7)实现对关键点的深度预测,生成深度矩阵  $\mathbf{H}$ 。矩阵  $\mathbf{A}$ 、 $\mathbf{H}$  的表达式分别为

$$\mathbf{A} = \begin{bmatrix} l_{12,11} & l_{11,23} \\ l_{12,24} & l_{24,23} \end{bmatrix}, \quad (2)$$

$$\mathbf{H} = \begin{bmatrix} h_{12} & h_{11} \\ h_{24} & h_{23} \end{bmatrix}, \quad (3)$$

式中: $l_{ij}$ 为关键点  $i$ 、 $j$  间距离; $h_k$ 为  $k$  关键点深度值。

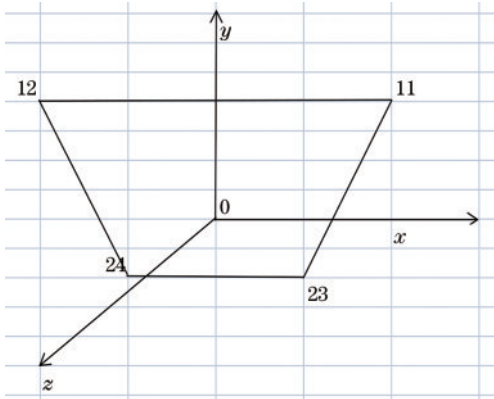


图 6 空间坐标系下人体关键点分布

Fig. 6 The distribution of human body key points in the spatial coordinate system

当使用者所在平面平行于摄像头平面时,以特征关键点所围图形中心为原点、以平行于 11、12 号关键点组成线段所在直线为  $x$  轴、以垂直于该平面且指向摄像头所在平面的直线为  $z$  轴,建立如图 6 所示空间坐标系。建立基准特征关键点矩阵  $\mathbf{A}_0$ ,并保存各关键点间距离,作为基准单位长度  $l_0$ ,此时深度矩阵记作  $\mathbf{H}_0$ 。

$$\mathbf{H}_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \quad (4)$$

系统运行过程中,由矩阵  $\mathbf{A}$  与矩阵  $\mathbf{A}_0$  的变换关系得到基准关键点沿  $z$  轴移动矩阵  $\mathbf{H}_z$ 、沿  $x$  轴旋转矩阵  $\mathbf{H}_x$ 、沿  $y$  轴旋转矩阵  $\mathbf{H}_y$ 。最终通过运动学叠加原理<sup>[15]</sup>,建立基准点深度矩阵  $\mathbf{H}$ 。矩阵  $\mathbf{H}$ 、 $\mathbf{H}_x$ 、 $\mathbf{H}_y$ 、 $\mathbf{H}_z$  的表达式分别为

$$\mathbf{H} = \mathbf{H}_0 + h \times \frac{\sqrt{l_{11,23}} + \sqrt{l_{12,24}}}{2} \cdot \mathbf{H}_x + w \times \frac{\sqrt{l_{11,23}} + \sqrt{l_{12,24}}}{2} \cdot \mathbf{H}_y + \frac{\|\mathbf{A}\|_1}{4} \times \frac{w+h}{2} \cdot \mathbf{H}_z, \quad (5)$$

$$\begin{cases} \mathbf{H}_x = \ln\left(\frac{l_{11,12}}{l_{0,11,12}}\right) \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \ln\left(\frac{l_{23,24}}{l_{0,23,24}}\right) \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \\ \mathbf{H}_y = \ln\left(\frac{l_{12,24}}{l_{0,12,24}}\right) \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \ln\left(\frac{l_{11,23}}{l_{0,11,23}}\right) \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \\ \mathbf{H}_z = \ln\left(\frac{\|\mathbf{A}\|_2}{\|\mathbf{A}_0\|_2}\right) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \end{cases}, \quad (6)$$

式中: $w$ 为输入图像宽度; $h$ 为输入图像高度。

在获取基准关键点的相对深度特征  $\mathbf{H}$  后,以特征关键点为基点对四肢关键点深度进行预测。以左臂为例,系统由大臂实时映射长度  $l_{11,13}$  与基准单位长度  $l_{0,11,13}$  得出大臂与躯干平面夹角  $\alpha_{11}$ ,以此确认大臂的相对深度特征  $h_{11,13}$ ,最终可以得到 13 号关键点深度特征  $h_{13}$ 。

$$h_{13} = h_{11} + l_{0,11,13} \cdot \sin\left[\arccos\left(\frac{l_{11,13}}{l_{0,11,13}}\right)\right]. \quad (7)$$

同理可得其他关键点深度特征,即实现了对所需关键点的深度预测功能。关键点检测输出效果如图 7 所示,其中,图 7(a)为原始输入图像,图 7(b)为图像分割与关键点提取展示图,图 7(c)为关键点坐标展示。

## 2.3 机械驱动模块

机械驱动模块基于 ROS 开发,实现对设备运行情况的实时监测并利用固定在机械臂末端摄像头采集设备的工作环境信息,最终通过通信网络实现对可移动机械臂的远距离操作控制,辅助使用者完成作业。

### 2.3.1 基于 D-H 参数的机械臂驱动控制

所提系统通过机械臂驱动控制模块完成对目标物体的搬运与环境信息的采集工作。系统通过人机交互模块实现对动作的捕捉与判断后将使用者的姿势映射为相应的机械臂驱动控制指令,并对指令执行发布、接收、解析工作,从而实现对机械臂的远程控制。

机械臂的链路由刚体连杆组成,在已知机械臂的各个关节连杆参数的情况下,通过控制机械臂各关节舵机转动角度,即可驱动机械臂运动到指定位置<sup>[15]</sup>。针对本研究的需求,选择改进型 Denavit-Hartenberg (D-H) 法<sup>[16-17]</sup>,为机械臂建立的 D-H 参数表如表 1 所示,旋转矩阵如式(8)所示。

$$\begin{cases} {}^{i-1}T_i = \mathbf{R}_X(\alpha_{i-1}) \times \mathbf{D}_X(a_{i-1}) \times \mathbf{R}_Z(\theta_i) \times \mathbf{D}_Z(d_i) \\ {}^{i-1}T_i = \begin{bmatrix} c\theta_i & -s\theta_i & 0 & \alpha_{i-1} \\ s\theta_i c\alpha_{i-1} & c\theta_i c\alpha_{i-1} & -s\alpha_{i-1} & -s\alpha_{i-1}d_{i-1} \\ s\theta_i s\alpha_{i-1} & c\theta_i s\alpha_{i-1} & c\alpha_{i-1} & c\alpha_{i-1}d_{i-1} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \end{cases} \quad (8)$$

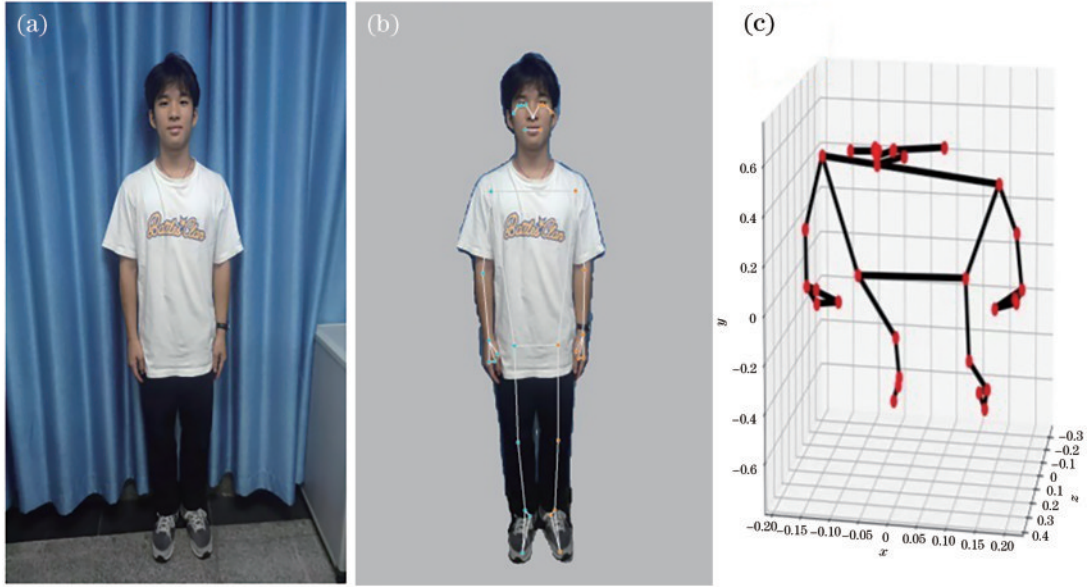


图 7 关键点深度估算效果展示。(a) 原始图像;(b) 关键点提取;(c) 深度估计

Fig. 7 Key point depth estimation effect display. (a) Raw image, (b) key extraction, (c) depth estimation

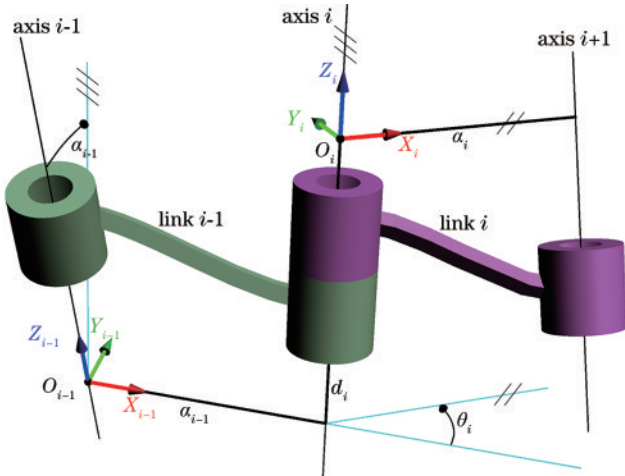


图 8 连杆坐标系

Fig. 8 Connecting rod coordinate system

式中： $a_{i-1}$ 为沿  $X_{i-1}$  轴从  $Z_{i-1}$  移动到  $Z_i$  的距离； $\alpha_{i-1}$  为沿  $X_{i-1}$  轴从  $Z_{i-1}$  旋转到  $Z_i$  的角度； $d_i$  为沿  $Z_i$  轴从  $X_{i-1}$  移动到  $X_i$  的距离； $\theta_i$  为沿  $Z_i$  轴从  $X_{i-1}$  旋转到  $X_i$  的角度。

表 1 机械臂 D-H 参数表

Table 1 Robot arm D-H parameter table

Link	$\theta$	$d$	$a$	$\alpha$	$\cos \alpha$	$\sin \alpha$
1	$\theta_1$	$L_1$	0	$90^\circ$	0	1
2	$\theta_2$	0	$L_2$	0	1	0
3	$\theta_3$	0	$L_3$	0	1	0
4	$\theta_4$	0	$L_4$	$-90^\circ$	0	-1
5	$\theta_5$	0	0	$90^\circ$	0	1
6	$\theta_6$	$L_5$	0	0	1	0

### 2.3.2 基于正运动学的全向轮式机器人运动控制

所提系统选取 Mecanum 轮式结构作为机械臂平台的运动驱动底盘。该部分根据人机交互模块发布的信息实现平面的全方向移动。

Mecanum 轮式机器人的全方位系统结构依赖轮组的速度组合即可实现稳定可靠且控制简单的全方位运动效果<sup>[18]</sup>。然而在复杂地形环境下 Mecanum 轮式机器人的底盘受力情况并非理想状态,实际运动轨迹与预测轨迹存在较大差异<sup>[19]</sup>,因此需要人工辅助控制。所提系统通过人机交互模块实现对 Mecanum 轮式机器人的远程控制<sup>[20]</sup>,以满足在不利于使用者进入的极端条件下对设备进行主动控制的需求。由于轮式机器人在平面上工作,因此,对 Mecanum 轮式机器人建立如图 9 所示坐标系进行运动学分析。

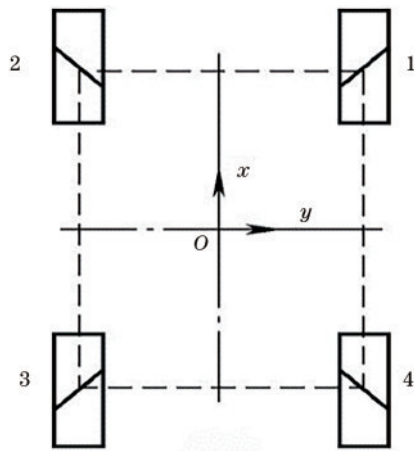


图 9 Mecanum 轮运动学建模

Fig. 9 Kinematic modeling of the Mecanum wheel

对于设备底盘来说,根据式(9)即可在已知小车的运动状态时,实现电机转速精准调节。

$$\begin{bmatrix} v_{1w} \\ v_{2w} \\ v_{3w} \\ v_{4w} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_y \\ v_x \\ \omega(a+b) \end{bmatrix}, \quad (9)$$

式中: $v_x$ 为底盘 $x$ 轴方向速度分量; $v_y$ 为底盘 $y$ 轴方向速度分量; $\omega$ 为底盘 $z$ 轴方向自转角速度; $(a, b)$ 为车的旋转中心到各轮子的距离坐标。

## 2.4 虚拟现实模块

远程操作过程依赖传输图像反馈设备运行状态,这对于服务器的性能要求极高,为改变传统模式下设备运行的“黑箱”状态<sup>[21]</sup>,辅助使用者更好操作设备,系统参考数字孪生技术通过建立多维结构模型实现预测功能<sup>[22]</sup>,利用Unity引擎搭建可移动机械臂的数字仿真模型,通过接收设备反馈数据进行图形化展示。

虚拟现实模块的节点分布如图10所示,建立了car、even两个主要节点。其中,car节点部署了可移动机械臂的数字化模型,even节点则用于搭载工作环境数字化虚拟模型并展示由固定在机械臂上的摄像头采集的工作实际场景。在系统运行状态下,该模块将接收到机械臂的数据反馈,驱动模型运动,给使用者直观形象的设备工况反馈。

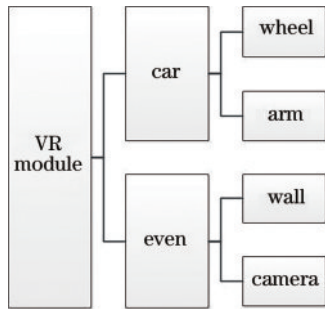


图10 虚拟现实模块节点分布

Fig. 10 Virtual-reality module node distribution

该模块的运动控制脚本挂载在car节点中,负责接收、解析与模拟由移动机械臂平台发布的设备工况数据。even节点中挂载的脚本则执行简单的设备碰撞检测并解析显示安装在机械臂端的摄像头捕捉的画面。场景3D展示效果如图11所示。

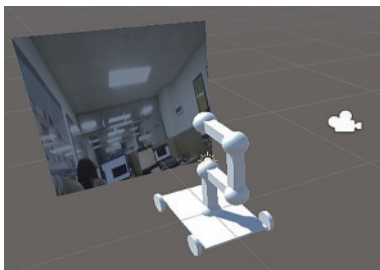


图11 场景3D展示效果

Fig. 11 Scene 3D display effect

## 3 分析与讨论

### 3.1 节流器测试

通过差异值哈希算法对输入图像进行预处理判断。其中,3种感知哈希算法运算效率与动作差异值的结果输出如表2所示。为区分使用者期待发布指令与使用者因抖动等问题造成的抖动误差指令,对3项控制方法分别进行4组测试。经测试,d-Hash处理速度最快,且满足对于手势指令识别的要求,故系统选取d-Hash算法作为人机交互模块节流器核心算法。

表2 感知哈希阈值与速度判断

Table 2 Perception hash threshold and speed judgment

Test image classification	Algorithm type	Speed / (ms/1000 photos)	Hamming distance
Fixture control	a-Hash	86.06	5
	p-Hash	225.15	7
	d-Hash	46.03	5
Flex control	a-Hash	80.05	19
	p-Hash	212.14	31
	d-Hash	45.03	17
Rotary control	a-Hash	79.05	27
	p-Hash	211.14	43
	d-Hash	45.03	24

### 3.2 系统响应速度测试

将系统按模块分别搭建在Windows 8系统(电脑配置为八内核Intel Core i7-3635QM,内存8 GB,CPU主频2.40 GHz)与jetson nano(设备配置为四内核ARMA57,内存4 GB,CPU主频1.43 GHz)上。为实现对系统响应速度的测试,在调取图像后嵌入time.perf\_counter()函数,利用设备性能计数器统计图像处理耗时。通过对5次测试数据的统计,绘制如图12所示的系统响应速度测试折线图。

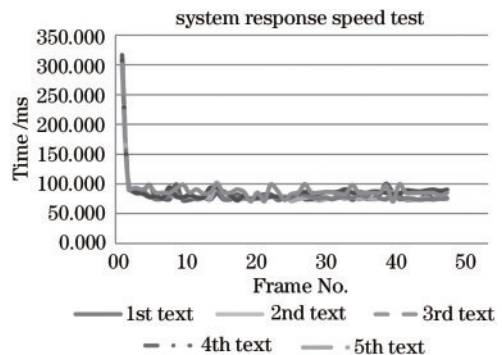


图12 系统响应速度测试

Fig. 12 System response speed test

由图12可知,系统的响应速度稳定在60~100 ms/frame。第一次指令发布过程中,由于系统需要进行服务器连接等工作,因此产生一定的延迟,此后每一次指令的发布耗时都稳定在60~100 ms。



在运行过程中,当使用者的输入图像变化差异值小于阈值时系统处于静默状态,故静默状态系统的响应速度相比于发布指令时的速度有较大差异。因此,加入输入图像预处理环节能降低系统对设备处理性能的占用率,且系统对于设备的性能要求较低。

### 3.3 交互指令测试

选择肢体关键点(图 5)23-11-13、11-13-15、手部关键点 4-5-12 所成夹角进行交互指令测试,运动轨迹如图 13 所示。手臂由图 13(a)运动至图 13(c)时,最终获取的 3 个夹角变化折线图如图 14 所示。

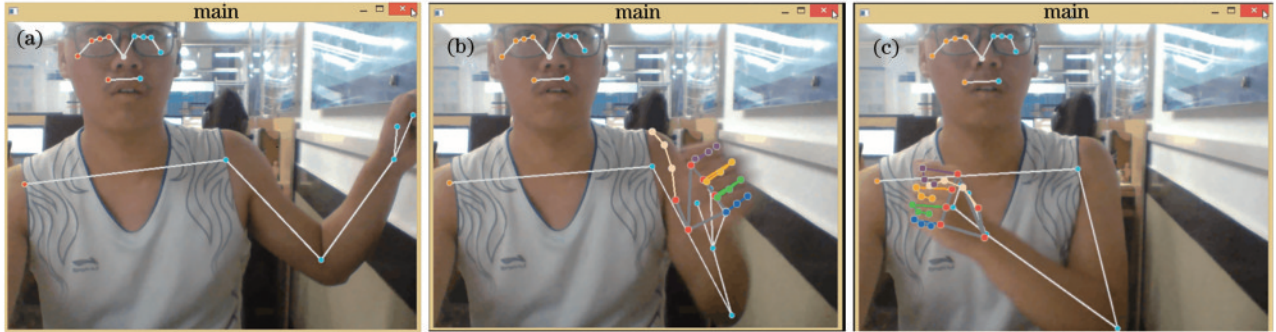


图 13 测试运动轨迹。(a)运动起始端;(b)运动中间时刻;(c)运动终点

Fig.13 Test motion trajectory. (a) Beginning of movement; (b) moment in the middle of movement; (c) end of movement

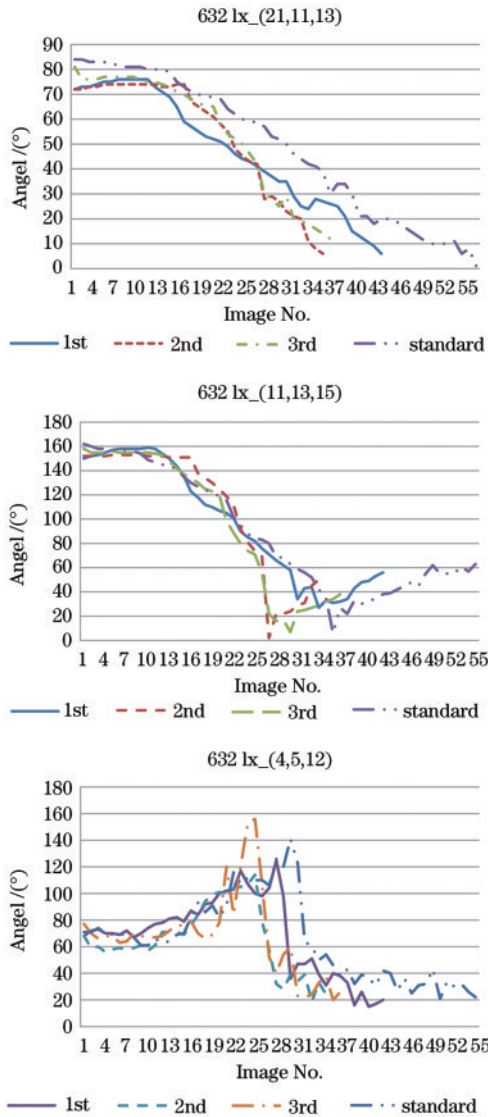


图 14 运动轨迹测试结果

Fig.14 Motion trajectory test results

通过对比室内遮光环境(光照强度 37 lx)、室内日光环境(光照强度 255 lx)、室内灯光环境(光照强度 632 lx)、室外日光环境(光照强度 10661 lx)等 4 种不同光照强度条件,评估所提系统的准确性。经测试,在室内灯光环境下设备的关键点估计准确性最好,但同时增加了因肢体抖动产生的误差。因此为提高系统鲁棒性,所提系统添加节流器实现滤波功能。

在图 14 的测试结果中,(23, 11, 13)项测试平均标准差为 2.49、(11, 13, 15)项测试平均标准差为 3.77、(4, 5, 12)项测试平均标准差为 11.77。本项测试选取 5 人完成该动作,其中,起点位置与终点位置为预先设定位置,故在测试过程中使用者的运动轨迹存在一定波动,但各关键点夹角的变化趋势相似且最终都向目标角度收敛。因此,该方法的稳定性可以满足大部分对于精度要求不高的环境下机械臂的操作工作。

此外,由于手部关键点(4, 5, 12)夹角的变化差异较大,为确保系统运行的稳定性,将夹具的控制函数改为式(10) $\theta_0$ 的阶跃函数。由图 14 所示的实验结果可知,将机械臂夹具的控制阈值设为  $60^\circ$ ,可以保障系统当且仅当手势调整至指定动作后,方可执行夹取与释放指令,有效避免系统运行过程中手势动作存在的波动。

### 3.4 系统运行效果展示

系统运行效果最终展示效果如图 15 所示,图像上方为使用者操作情况、下方为机械臂的运行情况。使用者首先在个人计算机中部署人机交互模块与虚拟现实模块,然后先后启动机械驱动模块与部署在个人计算机的模块,建立两者的通讯通道,随后使用者即可通过肢体手势动作对机械臂平台进行远程控制。

肢体动作与舵机映射关系如式(10)所示:

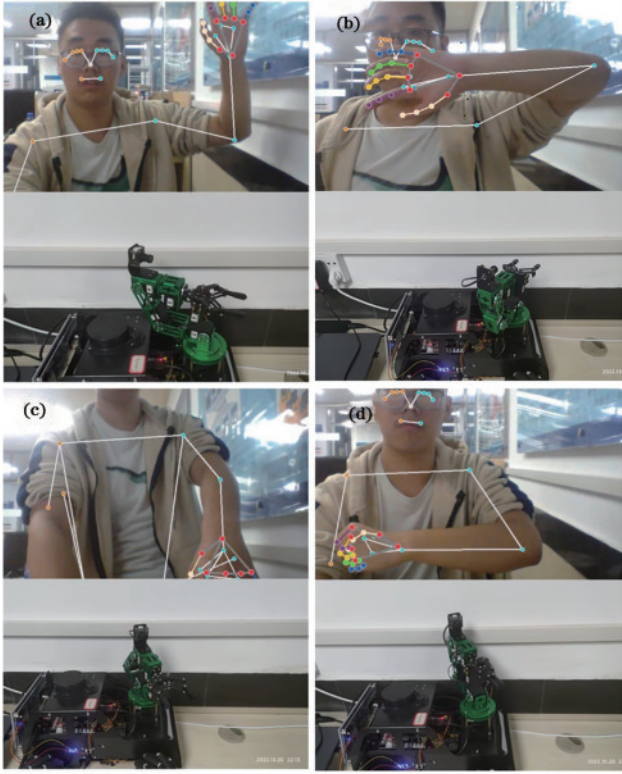


图 15 系统运行效果展示。(a)机械臂初始状态;(b)机械臂转向右侧并收缩;(c)机械臂转向左侧并前伸;(d)机械臂转向左侧并夹紧夹具

Fig.15 System operation effect display. (a) Robot initialization; (b) the arm turns to the right and retracts; (c) the arm turns to the left and extends forward; (d) the arm turns to the left and clamps the gripper

$$\begin{cases} \theta_1 = 0.75 \cdot \theta_{11,13,23} \\ \theta_3 = 0.6 \cdot (90 - \theta_{11,13,15}) \\ \theta_4 = 0.6 \cdot (\theta_{11,13,15} - 90) \\ \theta_6 = 90 \cdot \epsilon(\theta_{4,5,12} - 60) \end{cases} \quad (10)$$

式中： $\theta_1$ 为一号舵机旋转角度； $\theta_3$ 为三号舵机旋转角度； $\theta_4$ 为四号舵机旋转角度； $\theta_6$ 为六号舵机旋转角度； $\theta_{11,13,23}$ 为肢体关键点 11-13-23 夹角； $\theta_{11,13,15}$ 为肢体关键点 11-13-15 夹角； $\theta_{4,5,12}$ 为手部关键点 4-5-12 夹角。

所提系统通过读取使用者右臂及右手的空间状态,并通过式(10)将其映射为机械臂的运动控制指令,最终辅助使用者实现对设备的预期运动控制效果。

由图 15 所示的实验结果可知:设备在运动过程中,由于肢体部分的关键点分布结构相比于手部关键点相互遮挡概率小,故肢体部分关键点分布预测精度相较于手势部分高,更适合作为系统的线性输入变量,并将手势指令通过阶跃函数处理,实现对夹具的二元化控制。此外,相比于未加入感知哈希函数的系统,所提系统对于使用者肢体在使用过程中产生的自然抖动具有良好的滤波效果。

## 4 结 论

混合现实技术是特殊环境或要求下实现使用者远程操控设备的一种重要技术手段,本文提出一种基于混合现实的可移动机械臂操作控制系统。控制系统软件部分包括人机交互、机械驱动、虚拟现实等 3 个主要功能模块。亦搭建了一整套硬件测试平台,测试结果表明,该平台可以实现对使用者肢体手势动作的解析,并将解析后的数据映射为机械控制指令,实现对机械臂的控制,能够解决特殊条件下操作者对设备无法进行实地操作的问题,具有一定的应用价值。测试中,常规状态下平台运行的响应速度能够达到 60~100 ms/frame,且响应速度与硬件设备性能具有较大的相关性,系统具有实现实时处理的前景,为该技术未来应用场景的选择和配套硬件设备的选择提供了参考。

此外,所提系统需要在摄像头捕捉到头部后才会进行下一步的关键点检测功能,这对于个人计算机的摄像头捕捉视角范围提出了一定的要求。而对于未在摄像头视野范围内的肢体部分的预测处理以及虚拟现实反馈模块的完善则是系统下一步发展的方向。

## 参 考 文 献

- [1] 李生才, 安莹. 2022 年 5—6 月国内环境事件[J]. 安全与环境学报, 2022, 22(4): 2292-2296.  
Li S C, An Y. Statistics of environmental events in China during the period from May to June in 2022[J]. Journal of Safety and Environment, 2022, 22(4): 2292-2296.
- [2] 温圳, 李鹏, 王叙萌, 等. 基于混合现实的可视分析综述[J]. 计算机辅助设计与图形学学报, 2022, 34(6): 821-829.  
When Z, Li P, Wang X M, et al. Mixed - reality - based visual analytics[J]. Journal of Computer-Aided Design & Computer Graphics, 2022, 34(6): 821-829.
- [3] 陶建华, 巫英才, 喻纯, 等. 多模态人机交互综述[J]. 中国图象图形学报, 2022, 27(6): 1956-1987.  
Tao J H, Wu Y C, Yu C, et al. A survey on multi-modal human-computer interaction[J]. Journal of Image and Graphics, 2022, 27(6): 1956-1987.
- [4] Pooryousef V, Brown R, Turkey S. Shape recognition and selection in medical-volume visualization with haptic gloves[C]//Proceedings of the 31st Australian Conference on Human-Computer-Interaction, December 2-5, 2019, Fremantle, WA, Australia. New York: ACM Press, 2019: 433-436.
- [5] Cao C X, Shan B, Zhang H Y. Pattern recognition of Wushu routine action decomposition process based on kinect[J]. Mathematical Problems in Engineering, 2022, 2022: 3876487. .
- [6] Newell A, Yang K, Deng J. Stacked hourglass networks for human pose estimation[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in



- computer science. Cham: Springer, 2016, 9912: 483-499.
- [7] 朱晓俊, 王学谦, 马云萱, 等. 绳驱超冗余空间柔性机械臂遥操作系统设计与实验研究[J]. 机器人, 2022, 44(1): 9-18.  
Zhu X J, Wang X Q, Ma Y X, et al. The design and experimental study of a teleoperation system for cable-driven and hyper-redundant space flexible manipulator[J]. Robot, 2022, 44(1): 9-18.
- [8] 高晓科, 孙亮, 张伟军. 高压带电引流作业机器人系统的设计与开发[J]. 机械设计与研究, 2021, 37(5): 10-14, 20.  
Gao X K, Sun L, Zhang W J. Design and development of a high-voltage live drainage robot system[J]. Machine Design & Research, 2021, 37(5): 10-14, 20.
- [9] Zhong B, Guo K Q, Yu H Y, et al. Toward gait symmetry enhancement via a cable-driven exoskeleton powered by series of elastic actuators[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 786-793.
- [10] 孔垂跃, 陈羽, 赵乾名. 基于 MQTT 协议的配电网物联网云边通信映射研究[J]. 电力系统保护与控制, 2021, 49(8): 168-176.  
Kong C Y, Chen Y, Zhao Q M. Research on cloud-side communication mapping of the distribution Internet of Things based on the MQTT protocol[J]. Power System Protection and Control, 2021, 49(8): 168-176.
- [11] 欧军, 吴清秀, 裴云, 等. 基于 socket 的网络通信技术研究[J]. 网络安全技术与应用, 2011(7): 19-21.  
Out J, Wu Q X, Pei Y, et al. Research on network communication based on socket[J]. Network Security Technology & Application, 2011(7): 19-21.
- [12] 许庚林, 冉峰, 邓良, 等. 轻量化神经网络和哈希跟踪算法在嵌入式人脸抓拍系统中的应用[J]. 上海大学学报(自然科学版), 2021, 27(6): 1018-1028.  
Xu G L, Ran F, Deng L, et al. The application of lightweight neural networks and Hash tracking algorithm in an embedded face capture system[J]. Journal of Shanghai University (Natural Science Edition), 2021, 27(6): 1018-1028.
- [13] Bazarevsky V, Grishchenko I, Raveendran K, et al. BlazePose: on-device real-time body pose tracking[EB/OL]. (2020-06-17)[2022-11-12]. <https://doi.org/10.48550/arXiv.2006.10204>.
- [14] 许宇能, 朱西产, 马志雄, 等. 基于单目摄像头的车辆前方道路三维重建[J]. 汽车技术, 2014(2): 48-52.  
Xu Y N, Zhu X C, Ma Z X, et al. Three-dimensional reconstruction of the road ahead of a vehicle based on mono-vision[J]. Automobile Technology, 2014(2): 48-52.
- [15] 任敬轶, 孙汉旭. 一种新颖的笛卡儿空间轨迹规划方法[J]. 机器人, 2002, 24(3): 217-221.  
Ren J Y, Sun H X. A novel method of trajectory planning in Cartesian space[J]. Robot, 2002, 24(3): 217-221.
- [16] When S H, Ma Z Y, When S H, et al. The study of the NAO robot arm based on direct kinematics using the D-H method[C]//2014 UKACC International Conference on Control (CONTROL), July 9-11, 2014, Loughborough, UK. New York: IEEE Press, 2014: 515-518.
- [17] 冷舒, 吴克, 居鹤华. 机械臂运动学建模及解算方法综述[J]. 宇航学报, 2019, 40(11): 1262-1273.  
Leng S, Wu K, Ju H H. Overview of manipulator kinematics modeling and solving method[J]. Journal of Astronautics, 2019, 40(11): 1262-1273.
- [18] 夏长安, 肖雨琴, 岳义, 等. 基于 Mecanum 轮的移动机器人运动建模与仿真[J]. 机械设计与研究, 2022, 38(4): 17-21.  
Xia C A, Xiao Y Q, Yue Y, et al. Movement modeling and simulation of a mobile robot based on the Mecanum wheel[J]. Machine Design & Research, 2022, 38(4): 17-21.
- [19] 杨霖. Mecanum 轮式全向移动平台减振技术的研究[D]. 大连: 大连交通大学, 2016.  
Yang L. The research on vibration reduction technology for Mecanum wheel omnidirectional mobile platform[D]. Dalian: Dalian Jiaotong University, 2016.
- [20] 王一治, 常德功. Mecanum 四轮全方位系统的运动性能分析及结构形式优选[J]. 机械工程学报, 2009, 45(5): 307-310, 316.  
Wang Y Z, Chang D G. Motion performance analysis and layout selection for a motion system with four Mecanum wheels[J]. Journal of Mechanical Engineering, 2009, 45(5): 307-310, 316.
- [21] 陶飞, 刘蔚然, 刘检华, 等. 数字孪生及其应用探索[J]. 计算机集成制造系统, 2018, 24(1): 1-18.  
Tao F, Liu W R, Liu J H, et al. Digital twins and its potential application exploration[J]. Computer Integrated Manufacturing Systems, 2018, 24(1): 1-18.
- [22] Tao F, Zhang M. Digital twin shop-floor: a new shop-floor paradigm toward smart manufacturing[J]. IEEE Access, 2017, 5: 20418-20427.