

基于混合视觉-物理模型的类人臂机器人控制策略可行性分析与实施计划

1. 引言

• 1.1. 问题陈述

实现由人类操作员手臂自然运动驱动的机器人手臂的实时、直观且精确的控制是一项重大挑战。这种控制方式需要在响应性方面具有低延迟, 在精确操作方面具有高精度, 并且对传感器噪声和环境变化具有鲁棒性。现有的控制方法往往难以平衡这些需求, 它们要么完全依赖于快速但可能不精确的视觉方法, 要么依赖于精确但可能较慢或需要预先标定的基于物理模型的方法。

• 1.2. 拟议混合策略概述

本报告旨在评估一种分阶段的混合控制策略, 旨在结合不同方法的优势:

- **阶段 1 (初始化):** 基于低延迟的2.5D视觉手部/臂部姿态估计(例如, 使用 MediaPipe Hands¹ 等系统) 直接驱动控制。此阶段的核心目标是即时响应性。
- **阶段 2 (融合/精化):** 逐步整合在线估计/优化的用户手臂物理模型(运动学)。利用传感器融合技术(例如, 扩展卡尔曼滤波器 EKF²) 结合2.5D视觉数据和演化中的物理模型的预测, 在模型收敛的同时提高精度和平滑度。
- **阶段 3 (基于物理模型的控制):** 过渡到主要或完全基于收敛后的高保真物理手臂模型的控制, 以实现精确和稳定的操作。

• 1.3. 潜在优势

这种混合方法旨在提供与直接视觉控制相当的初始低延迟, 随着物理模型的精化逐步提高精度和平滑度, 并最终实现对短暂视觉错误具有鲁棒性的高精度控制。

• 1.4. 关键挑战

主要挑战包括: 确保融合和过渡阶段的稳定性; 管理实时性能所需的计算负载(涵盖视觉、估计、融合、控制等所有模块); 实现物理模型参数的快速、准确的在线估计; 处理传感器噪声以及视觉估计和物理估计之间可能存在的差异。人类运动的高速度, 特别是在乒乓球等任务中⁴, 凸显了对低延迟的严格要求。

2. 可行性分析与技术现状回顾

• 2.1. 混合视觉-运动学控制的技术可行性

将视觉感知与运动学/动力学模型相结合是机器人学中用于状态估计和控制的一种常见且有效的策略²。视觉反馈提供外部观测, 而运动学模型则强制施加物理约束, 并在视觉遮挡或噪声期间进行预测。

所提出的分阶段方法(初始视觉 -> 融合 -> 物理模型)在概念上是合理的, 旨在弥补每种单一模态的局限性。初始阶段利用了现代姿态估计器的高速度¹⁰, 而后续阶段则引入了物理真实性和鲁棒性¹⁴。

该方法的核心可行性取决于融合机制(参见第3.3节) 平滑融合可能存在分歧的估计值的能力, 以及在线参数估计(参见第2.3节) 能否足够快速和准确地收敛以支持最终阶

段。视觉估计(如MediaPipe 1)虽然提供度量单位的3D坐标,但在没有深度传感器的情况下存在固有的精度限制。物理模型在开始时是未标定的。融合机制必须处理这种初始差异,并随着物理模型置信度(参数协方差减小)的增加而调整其权重。成功与否取决于所选滤波器(如EKF)能否正确建模两个来源的不确定性,以及在线估计算法(如RLS 19)能否在典型的任务持续时间内收敛。

- **2.2. 延迟、精度、平滑度与稳定性分析**

- **延迟:** 阶段1的延迟受限于2.5D姿态估计器的延迟(例如, MediaPipe流水线在Pixel 6 GPU/CPU上约为12-17ms¹², RTMPose-m在GPU上为2.3ms¹⁰, ReJSHand在2080Ti上为13.9ms¹³)加上控制信号生成时间。阶段2增加了融合计算(EKF更新²¹)。阶段3取决于物理模型计算(例如, IK求解)和控制律执行。实时控制通常需要远高于30Hz的循环频率,对于动态任务可能需要超过100Hz⁴。强化学习(RL)方法的策略推理也存在延迟约束²⁴。较新的扩散模型旨在实现单步、低延迟的轨迹生成(例如, FRMD报告约为17ms²⁸)。
- **精度:** 阶段1的精度取决于2.5D估计器(MediaPipe的3D世界坐标精度限制在¹⁸中有提及)。阶段2旨在通过结合运动学约束来提高精度³⁰。阶段3的精度取决于收敛后的物理模型的保真度和标定质量³³。
- **平滑度:** 直接视觉控制(阶段1)可能因噪声估计而产生抖动³⁵。滤波(如卡尔曼滤波²¹)或结合时间信息(例如, PoseFormer³⁶)可以改善平滑度。物理模型本身就能提供更平滑的轨迹。融合(阶段2)应能平滑过渡。扩散模型最初也面临平滑度挑战,但像FRMD这样的新方法旨在改善这一点²⁸。
- **稳定性:** 模型切换/融合是稳定性的关键点。必须避免控制信号的突变。EKF框架通过传播不确定性(协方差)来提供帮助²。自适应控制技术¹⁹具有相关性。确保物理模型保持稳定且参数正确收敛至关重要。一个根本性的权衡存在:阶段1优先考虑低延迟,牺牲部分精度/平滑度。阶段3优先考虑精度/平滑度,可能以初始延迟(模型计算)为代价。阶段2是桥梁,其有效性取决于在实时循环预算内管理融合和参数估计的计算成本,同时确保稳定的融合。每个控制周期的总延迟预算(例如, 100Hz对应10ms)必须容纳感知、状态估计(融合)、参数更新和控制律计算。如果组合成本超过预算,性能会下降⁴。因此,仔细选择轻量级算法(例如,高效的姿态估计器¹⁰,优化的EKF²¹)和高效的实现(C++⁴³)至关重要。过渡逻辑(参见第3.4节)的设计必须防止因视觉估计和物理模型预测之间的控制权限突然转移而导致的不稳定性。

- **2.3. 在线运动学参数估计技术回顾**

- **目标:** 在线估计用户手臂的参数,如臂段长度、关节限制,甚至惯性参数。
- **方法:**
 - **递归最小二乘法 (RLS):** 适用于估计线性模型(或线性化模型)中的参数。已用于估计人臂阻抗参数¹⁹。需要系统动态的充分激励。
 - **基于滤波器的方法 (例如, EKF/UKF):** 可以将参数包含在EKF/UKF的状态向量中,并与姿态/速度一起估计¹⁴。参数的过程模型通常是随机游走或常数,观测模型将传感器读数(视觉姿态、IMU)与基于当前参数估计的预测相关联。已

用于传感器到段的标定¹⁴和机器人运动学标定⁴⁸。

- 基于优化的方法: 批处理或滑动窗口优化可以通过最小化模型预测与传感器数据之间的误差来估计参数¹⁴。因子图提供了一个模块化的框架¹⁷。
 - 数据要求: 通常需要多样化的运动, 以充分激励运动链, 使参数可观测。可能需要特定的初始标定运动¹⁴。视觉跟踪¹⁹提供了必要的观测数据。在线估计是可行的, 但收敛速度和精度在很大程度上取决于操作过程中观测到的运动数据的质量和丰富性。可能需要在启动时进行初始标定运动, 以引导该过程, 然后再完全依赖于任务相关的运动。用户查询意味着立即从2.5D控制开始。如果物理模型参数最初未知, 融合阶段需要处理大的参数不确定性。集成到EKF中的RLS或基于滤波器的方法似乎最适合在线更新。参数的可观测性(例如, 在没有肘部弯曲的情况下区分上臂和前臂长度)取决于观测到的姿态。系统设计必须考虑是否需要在启动时执行一个简短的、专门的标定运动序列, 以确保合理快速的收敛。
- 2.4. 传感器融合方法回顾 (EKF/KF 用于视觉 + 运动学/动力学)
 - 相关性: 直接适用于阶段2, 融合2.5D视觉姿态估计(观测)与来自运动学模型的预测(过程模型)。
 - 卡尔曼滤波器 (KF): 最适用于具有高斯噪声的线性系统。由于非线性运动学和观测模型, 在此处直接适用性较差⁴¹。
 - 扩展卡尔曼滤波器 (EKF): 广泛用于非线性系统, 如机器人/人体姿态估计²。在每一步使用雅可比矩阵线性化过程和观测模型。如果非线性很强, 可能会受到线性化误差和一致性问题的影响。实现通常使用李群表示法处理方向(SO(3))以避免奇异性问题³⁰。
 - 无迹卡尔曼滤波器 (UKF): EKF的一种替代方法, 通过使用确定性采样(sigma点)来捕捉通过非线性变换的均值和协方差, 从而避免了分析雅可比矩阵。对于高度非线性的系统, 通常比EKF更准确, 但计算成本更高³。
 - 粒子滤波器 (PF): 可以处理任意非线性和非高斯噪声, 但计算成本非常高, 可能不适用于这种实时应用⁵⁵。
 - 因子图: 为批处理或滑动窗口优化提供了一个灵活的框架, 自然地处理异步多传感器融合¹⁷。可以提供比滤波器更高的精度, 但通常会引入延迟(平滑), 除非在移动水平估计(MHE)设置中使用⁶³。
 - 关键考虑因素: 状态表示(关节角度、速度、参数)、过程模型(运动学方程、噪声假设)、观测模型(将状态映射到2.5D视觉关键点、噪声特性)、适当的协方差管理。EKF代表了解决此类实时融合问题最实用且被广泛采用的方法, 平衡了性能和计算成本。仔细的实现, 可能使用李群表示法处理方向和鲁棒的噪声建模, 是成功的关键。系统涉及非线性运动学(将关节角度转换为笛卡尔关键点)和潜在的非线性观测模型(如果将3D模型映射到2D图像平面, 则为透视投影, 尽管2.5D估计器通常直接提供3D世界坐标¹)。EKF是标准选择²。如果EKF在精度方面遇到困难, 可以考虑UKF³, 但需要根据延迟预算评估增加的计算量。因子图⁴⁸功能强大, 但除非实现为MHE⁶³, 否则可能太慢, 而MHE会增加复杂性。因此, EKF是推荐的起点。
- 2.5. 混合/自适应控制策略回顾

- 概念: 根据变化的条件或性能反馈调整其控制律或模型的系统。这与本项目相关, 因为控制器在不同模式(视觉、融合、物理)之间转换, 并适应在线估计的模型。
- 示例: 切换控制、增益调度¹⁹、模型参考自适应控制(MRAC)、自校正调节器。模糊PID控制器与RL(PPO)相结合已用于适应数字孪生系统中的误差⁶⁴。结合视觉和力反馈也是研究方向⁸。
- 相关性: 拟议的系统本身就使用了一种切换/混合控制的形式。融合阶段充当过渡, 最终阶段依赖于适应(在线估计)的模型。过渡期间的稳定性至关重要。需要确保平滑切换和适应的技术。基于二次规划的安全滤波器可以确保满足约束(如可见性⁷), 同时调整控制输入。系统设计需要明确的机制来实现平滑过渡, 并可能根据当前阶段和对物理模型估计的置信度(例如, 参数协方差)自适应地调整控制增益。简单地在满足阈值时切换控制律可能会导致不稳定。融合阶段自然地提供了一种混合, 但加权机制(参见第3.3.iv节)需要仔细设计。随着物理模型参数收敛(协方差减小), EKF内物理模型预测的权重/信任度应增加, 从而实现平滑的交接。控制增益也可能需要根据当前的控制源(视觉 vs. 融合 vs. 物理)进行调度, 以保持所需的性能和稳定性。

3. 拟议的混合控制框架

● 3.1. 概念架构与分阶段方法

- 图示: 高层框图展示: 相机 -> 2.5D姿态估计器 -> 融合模块/控制选择器 -> 机器人控制器 -> 机器人手臂。并行路径: 2.5D姿态估计器 -> 在线参数估计器 -> 物理模型 -> 融合模块/控制选择器。
- 阶段 1 (初始化): 从2.5D姿态估计(例如, 手部位置/方向)通过机器人手臂的逆运动学(IK)直接映射到机器人末端执行器的目标姿态。同时, 2.5D数据馈送给参数估计器。低延迟是首要任务。
- 阶段 2 (融合): EKF融合2.5D姿态估计(测量)与来自演化中的物理手臂模型的预测(过程模型)。EKF状态输出(滤波后的关节角度/速度)驱动机器人控制器(例如, 通过机器人IK或关节指令)。参数估计继续进行。平衡延迟/精度。
- 阶段 3 (基于物理模型的控制): 控制依赖于收敛后的物理手臂模型。这可能涉及直接使用估计的人体关节角度(如果自由度匹配)或使用模型进行预测控制或约束强制执行。高精度/平滑度优先。视觉输入仍可用于监控或偶尔的漂移校正。

● 3.2. 详细控制流程图与数据流水线

- 图示: 详细流程图展示每个时间步的数据流:
 1. 图像采集 (相机)。
 2. 2.5D姿态估计 (例如, MediaPipe -> 3D手部关节点 multi_hand_world_landmarks¹)。
 3. 参数估计更新 (使用2.5D姿态和先前状态)。
 4. EKF预测步骤 (使用先前状态和带有更新参数的运动学模型)。
 5. EKF观测 (使用当前2.5D姿态)。
 6. EKF更新步骤 (计算融合状态: 关节角度 θ_{fused} , 角速度 ω_{fused} , 协方差

Pfused)。

7. 置信度/权重计算 (基于 Pfused, 预测误差等)。
 8. 控制模式选择/混合 (基于置信度/阶段)。
 9. 目标生成 (例如, 目标关节角度 θ_{target} 基于所选模式: $\theta_{2.5D}$, θ_{fused} , 或 θ_{physics})。
 10. 机器人控制指令计算 (例如, 用于末端执行器姿态的机器人IK, 或直接关节指令)。
 11. 指令执行 (机器人手臂)。
- 同步: 处理传感器 (相机) 和控制循环频率之间潜在的时序不匹配问题。EKF在一定程度上自然地处理异步更新⁴³。
- **3.3. 数学公式推导 (基于EKF的融合示例)**
 - **(a) 状态向量定义 (x):**
 - $x = T$
 - θ : 人体手臂关节角度向量 (例如, 肩、肘、腕共7自由度)。
 - ω : 人体手臂关节角速度向量。
 - p : 缓慢变化的运动学参数向量 (例如, 上臂长度、前臂长度)。包含其不确定性。
 - 关联的协方差矩阵 P 。
 - **(b) 融合机制 (EKF):**
 - **i. 过程模型 ($x_k = f(x_{k-1}, u_{k-1}) + w_k$):**
 - 基于当前状态和控制输入 (本例中无控制输入, 假设由人驱动运动) 预测下一时刻状态。
 - $\theta_k = \theta_{k-1} + \omega_{k-1} \Delta t$ (角度的简单匀速预测)
 - $\omega_k = \omega_{k-1}$ (假设匀速, 或将加速度建模为噪声)
 - $p_k = p_{k-1}$ (参数假设为常数或缓慢漂移 - 由过程噪声建模)
 - w : 过程噪声 (高斯分布, 零均值, 协方差 Q), 用于解释模型不准确性和未建模的动态 (例如, 加速度)。
 - 需要计算 f 相对于 x 的雅可比矩阵 F 用于协方差预测:
 $P_{k|k-1} = F P_{k-1|k-1} F^T + Q$ 。
 - **ii. 观测模型 ($z_k = h(x_k) + v_k$):**
 - 将预测状态 x_k 与期望的视觉观测 z_k 相关联。
 - z_k : 观测向量, 例如, 来自2.5D估计器的关键手/臂关节节点的3D坐标 (例如, MediaPipe¹ 或 RTMW3D⁶⁵ 的手腕、肘、肩世界坐标)。
 - $h(x_k)$: 正向运动学函数。根据估计的关节角度 θ_k 和运动学参数 p_k 计算相机或世界坐标系中期望的3D关节节点位置。
 - v : 观测噪声 (高斯分布, 零均值, 协方差 R), 代表2.5D姿态估计器的不确定性/噪声。 R 可以根据估计器的置信度得分进行调整¹。
 - 需要计算 h 相对于 x 的雅可比矩阵 H 用于卡尔曼增益计算。
 - **iii. 状态更新与协方差更新方程: 标准EKF方程⁴¹:**
 - 卡尔曼增益: $K_k = P_{k|k-1} H^T (H P_{k|k-1} H^T + R_k)^{-1}$

- 状态更新: $x_k|k = x_k|k-1 + K_k(z_k - h(x_k|k-1))$
- 协方差更新: $P_k|k = (I - K_k H_k) P_k|k-1$
- (c) 置信度加权策略:
 - 隐式: EKF根据预测状态协方差 $P_k|k-1$ 和测量噪声协方差 R_k 的相对大小, 固有地对信息进行加权。随着物理模型参数 p 收敛, 其在 P 中对应的条目将减小, 自动相对于噪声较大的视觉测量给予模型预测更大的权重。
 - 显式: 可以使用一个额外的加权因子 α (范围0到1) 来混合直接视觉控制 u_{vis} 和基于EKF的控制 u_{ekf} : $u_{final} = \alpha u_{ekf} + (1-\alpha) u_{vis}$ 。
 - 权重更新逻辑: α 可以随时间增加, 或者与 P 中参数子协方差矩阵的迹/行列式相关联, 或者基于新息的一致性 (z_k 和 $h(x_k|k-1)$ 之间的差异)。如果新息很大, 可能表明发散或模型拟合不佳, 此时应减小 α 。
- (d) 过渡逻辑与触发条件 (阶段 2 -> 阶段 3):
 - 触发: 当物理模型被认为足够可靠时进行过渡。
 - 条件:
 - 参数收敛: P 中参数 p 的协方差低于预定义阈值。
 - 预测精度: EKF预测 $h(x_k|k-1)$ 与视觉测量 z_k 之间的均方误差在一个时间窗口内持续保持较低水平。
 - 基于时间: 在阶段2中经过固定持续时间后过渡(更简单, 但鲁棒性较差)。
 - 过渡: 逐渐将控制权限完全转移到物理模型(例如, 如果使用显式混合, 则设置 $\alpha=1$, 或者简单地使用 $x_k|k$ 来驱动控制)。EKF提供了一种原则性的方法来融合数据和管理不确定性。关键的设计选择涉及准确地建模过程噪声 Q (状态在步骤之间可能改变多少) 和观测噪声 R (视觉估计的可靠性如何), 并可能根据视觉估计器的置信度调整 R 。过渡逻辑需要对过早切换具有鲁棒性。EKF是标准选择²。状态必须包括角度、速度和参数¹⁴。过程模型是运动学+噪声²¹。观测模型是正向运动学, 将状态映射到视觉关节点²¹。噪声矩阵 Q 和 R 是关键调整参数。 R 可以根据姿态估计器的置信度来确定¹。基于参数协方差收敛的过渡似乎最为鲁棒, 表明模型已经学习了用户的特定手臂尺寸。

4. 系统架构与组件选型

● 4.1. 2.5D姿态估计模块:

- 候选方案:
 - **MediaPipe Hands/Pose:**¹ 在移动设备/CPU/GPU上具有实时性能。提供21个3D手部关节点(米制世界坐标, 原点在手部中心¹)或33个3D身体关节点⁷¹。延迟: 约12-17ms(完整流水线, Pixel 6¹²), 每个组件约1-2ms(Snapdragon¹¹)。在没有深度信息的情况下, 精度有限¹⁸。易于实现。
 - **RTMPose/RTMW/RTMW3D:**⁷⁶ 在基准测试中达到SOTA精度(COCO-Wholebody⁶⁵)。实时速度(RTMPose-m在CPU上90+ FPS, GPU上430+ FPS¹⁰)。提供2D/3D全身关键点。RTMW3D专门针对3D⁷⁶。需要MMPose框架。

- **ReJSHand:** ¹³ 专注于实时 (72 FPS on 2080Ti ¹³) 手部姿态和网格重建。使用 FastViT 骨干网络 ⁸⁵。声称在速度和精度之间取得了良好平衡 ¹³。
- **OpenPose:** ⁷⁶ 早期的实时多人姿态估计系统。可以输出身体、手部、面部关键点 ⁸⁸。在其他研究中被用作组件 ⁷⁰。在现代硬件上的性能可能低于较新的方法 (⁸⁹ 指出资源占用高)。
- **FrankMocap:** ⁹⁰ 专注于通过回归和集成进行全身 (SMPL-X) 估计。报告速度约 9.5 FPS ⁹³，对于此处的直接实时控制可能太慢。
- **其他单目3D方法:** 基于事件的方法 (用于低光/高速) ⁹⁴，基于Transformer的方法 ¹⁶，直接回归方法 ³⁵。

○ 对比分析 (表):

- 表 4.1: 实时手部/臂部姿态估计器比较

方法 (Method)	输出 (Output) (关节点/网 格/参数, 坐 标系)	报告的延 迟/FPS (硬 件)	精度基准 (Accurac y Benchma rk) (例如, PA-MPJ E, COCO AP)	优点 (Strength s)	缺点/限制 (Weaknes ses/Limit ations)	实现难度 (Impleme ntation Effort)
MediaPipe Hands	21个3D关 节点 (世界 坐标, 米) ¹	~12-17ms (Pixel 6) ¹² , ~1-2ms/组 件 (Snapdrag on) ¹¹	临床应用 中空间精 度有限 ¹⁸	易用, 跨平 台, 低延迟, 直接输出 3D世界坐 标	纯RGB精 度相对较 低, 对遮挡 敏感	低 (Low)
RTMW3D-I	133个3D关 节点 (模型/ 世界坐标?)	延迟未明 确报告, 但 RTMPose- I (2D) 130+ FPS (CPU) ¹⁰	H3WB MPJPE ~56mm ⁷⁷ , COCO-WB AP高 ⁶⁵	高精度, 针 对3D全身 姿态	依赖 MMPose框 架, 硬件要 求可能较 高	中 (Medium)
ReJSHand	3D关节点 + 网格 ⁸⁴	72 FPS (NVIDIA 2080Ti) ¹³	PA-MPJPE 6.3mm, PA-MPVPE 6.4mm ¹³	速度与精度 平衡良好, 输出网格	依赖特定 骨干网络 (FastViT) ⁸⁵ , 相对较 新	中 (Medium)

OpenPose	2D/3D 关节点 (身体, 手, 脸) ⁸⁸	早期报告实时, 但资源消耗大 ⁸⁹	COCO AP (2D)	成熟, 多部位检测	性能可能不如新方法, 资源消耗大	中 (Medium)
FrankMocap	SMPL-X 参数 (全身) ⁹¹	~9.5 FPS (GeForce RTX 2080) ⁹³	定性效果好	输出参数化模型, 统一全身表示	速度较慢, 可能不适用于直接控制	高 (High)

* **推荐选择:** **MediaPipe Hands** 是一个强有力的初始候选方案, 因为它易于使用、跨平台可用、在各种设备上有记录的低延迟 [11, 12], 并且直接输出米制3D世界坐标 [1], 这简化了EKF观测模型。虽然其绝对精度可能低于SOTA研究模型 [18], 但这对于阶段1来说是可以接受的, 并将在阶段2/3中得到补偿。如果初始精度至关重要且MMPose依赖性可以接受, **RTMW3D** [65, 78] 是一个强有力的替代方案。**ReJSHand** [13] 也具有竞争力。选择涉及易于实现/部署 (MediaPipe) 和潜在更高精度 (RTMW3D) 之间的权衡。鉴于混合策略旨在纠正初始不准确性, 从更易于集成的MediaPipe开始似乎是务实的。MediaPipe文档齐全, 使用广泛 [68, 70, 72, 73], 并提供必要的3D输出 [1]。其已知的精度限制 [18] 正是需要融合/物理模型阶段的原因, 这也验证了混合方法本身的合理性。从MediaPipe开始可以让我们将开发精力集中在创新的融合和参数估计部分。

● 4.2. 物理模型模块:

○ 运动学模型选择:

- **结构:** 代表人体手臂的简化运动学链。通常为7自由度 (DoF): 3 DoF肩部 (球状), 1 DoF肘部 (旋转), 3 DoF腕部 (球状)。可以使用标准的 Denavit-Hartenberg (DH) 参数或指数积 (PoE) 表示法。运动学模型在人体姿态估计和机器人学中是标准的¹⁵。
- **参数:** 需要在线估计的关键参数: 上臂长度、前臂长度。关节限制最初可以假设为固定值或离线估计。

○ 在线参数估计算法:

- **选择:** EKF状态增强 (如第3.3节所述) 或递归最小二乘法 (RLS)¹⁹。
- **理由:** 如果已经使用EKF进行融合, EKF集成是很自然的。如果单独实现, RLS可能更简单, 但需要仔细构建回归问题。两者都需要足够的运动变化。

○ 通过物理模型进行控制实现 (阶段 3):

- **选项 1: 直接关节映射:** 使用估计的人体关节角度 θ_{fused} (来自EKF状态) 作为机器人手臂关节的目标角度 (需要自由度匹配或映射)。最简单的方法。

- 选项 2: 末端执行器控制: 对 θ_{fused} 使用正向运动学得到估计的人手3D姿态(位置/方向)。将其用作机器人末端执行器的目标姿态, 通过机器人的逆运动学 (IK) 求解。对于人与机器人之间的运动学差异更具灵活性。
- 选项 3: 模型预测控制 (MPC): 使用物理模型预测未来的人体运动, 并在一个时间窗口内优化机器人控制动作, 可能结合约束(关节限制、速度限制、碰撞避免)。计算成本更高。
- 推荐: 从选项2(末端执行器控制)开始, 以获得灵活性。物理模型的复杂性应加以平衡。一个带有长度估计的简单7自由度运动学链最初可能就足够了。更复杂的动力学模型会增加计算开销, 并需要估计更多参数(质量、惯性)。目标是实时控制。运动学模型¹⁵的计算成本低于完整的动力学模型。与估计惯性参数相比, 仅估计臂段长度¹⁹简化了在线估计问题。使用估计的人手姿态进行末端执行器姿态控制, 可以实现人与机器人运动学的解耦。
- 4.3. 融合模块:
 - 算法选择: EKF (如第3.3节详述)。
 - 实现细节:
 - 状态表示: 使用最小表示法(例如, 如果需要, 使用四元数表示方向, 尽管关节角度可能足以表示手臂模型)。
 - 雅可比矩阵: 解析推导或使用自动微分计算雅可比矩阵 F 和 H 。数值微分是可能的, 但效率/精度较低。
 - 噪声建模: Q (过程噪声) 需要根据预期的运动动态/平滑度进行调整。 R (观测噪声) 应反映2.5D估计器的不确定性, 可能根据置信度得分进行调整¹。 P 中的初始参数不确定性应设置较高。
 - 库: Eigen⁴³ 用于线性代数。如果使用ROS, 可考虑ROS robot_localization⁵⁷, 或独立的C++实现⁴³。不变EKF⁴⁵ 在李群上的姿态估计方面具有理论优势。EKF的仔细实现和调整对于成功至关重要。数值稳定性、正确的雅可比计算以及适当的噪声协方差选择直接影响性能和收敛性。EKF实现的细节非常重要²¹。使用成熟的库(Eigen⁴³, ROS⁹⁹)可以节省工作量。调整 Q 和 R 通常是经验性的, 但至关重要。根据报告的姿态估计器精度(例如, 来自¹³)设置 R 的初始值, 并根据预期的人体运动平滑度设置 Q 的初始值, 是一个合理的起点。

5. 数据要求与采集策略

- 5.1. 2.5D模块所需数据:
 - 如果使用预训练模型(例如, MediaPipe, RTMPose): 最初可能不需要额外的训练数据。但是, 在特定领域数据(例如, 特定的相机视角、光照、用户外观)上进行微调可能会提高性能。像COCO-Wholebody⁶⁵、UBody⁷⁶、H3WB⁶⁵、FreiHAND³⁵这样的数据集被用于训练SOTA模型。合成数据也很常见¹。
 - 如果从头训练/大量微调: 需要具有多样化图像/视频和精确2D/3D标注的大型数据集³⁷。
- 5.2. 物理模型参数的标定协议:

- 目标: 收集能够稳健估计臂段长度的数据。
- 流程: 设计一个简短的用户手臂特定运动序列, 在启动时执行。
 - 示例: 保持上臂相对静止的情况下缓慢进行肘部屈伸(约90度)(分离前臂长度)。保持肘部伸直的情况下缓慢进行肩部外展/内收(分离总臂长, 有助于确定上臂长度)。包括跨越工作空间不同部分的运动。
 - 在这些运动期间记录视觉姿态估计。
 - 将数据馈送到在线估计器(EKF或RLS), 以比仅依赖任意任务运动更快地初始化/收敛参数。灵感来自¹⁴中的标定姿势。
- 5.3. 融合模块训练所需数据:
 - **EKF:** 在传统意义上不需要训练数据。需要调整噪声参数 Q 和 R 。这可以使用记录的人体运动序列以及相应的视觉姿态估计来完成, 可能需要将EKF输出与地面真值(如果可用, 例如来自光学运动捕捉³²)进行比较。
 - 如果使用基于学习的融合模块: 需要成对的带有噪声的视觉输入序列和地面真值(或高保真估计)的运动学状态序列。这会显著增加复杂性。应专注于调整EKF而不是训练融合模块。标定协议对于引导物理模型估计至关重要。EKF是基于模型的滤波器, 而不是学习模型⁴¹。其性能取决于模型精度和噪声调整。需要数据来调整 Q 和 R , 并验证性能(第6节), 而不是用于训练EKF本身。标定序列(第5.2节)直接解决了对良好初始参数估计的需求。

6. 实验设计与验证方案

- 6.1. 组件级基准测试:
 - **2.5D模块:** 在目标硬件上测量延迟(ms)和精度(例如, 与地面真值相比的平均每关节位置误差 - MPJPE, 或一致性指标)¹⁰。在不同条件下(光照、速度)进行测试。使用像FreiHAND³⁵、H3WB⁶⁵这样的数据集, 或带有地面真值的自定义录制。
 - 参数估计: 使用标定协议(第5.2节)和可能更长的运动序列, 测量估计参数(例如, 臂长)的收敛时间和最终精度。将估计长度与手动测量或参考值进行比较。评估对不同初始化和运动类型的鲁棒性。
- 6.2. 端到端系统评估:
 - **关键性能指标 (KPIs):**
 - 端到端延迟: 从图像捕获到机器人指令执行的时间(ms)。分阶段测量。对响应性至关重要⁴。
 - 轨迹跟踪误差: 人手运动(地面真值或参考)与机器人末端执行器运动之间的差异(例如, 3D位置/方向的均方根误差 - RMSE)。分阶段测量。(灵感来自⁶⁷)。
 - 运动平滑度: 分析机器人末端执行器轨迹的加加速度(加速度的导数)。跨阶段和基准进行比较²⁸。
 - 参数收敛时间: 在阶段2期间, 估计的物理模型参数稳定在容差带内所需的时间。
 - 切换稳定性: 对阶段之间(特别是阶段2 -> 阶段3)过渡期间的平滑度和稳定性进行定性和定量评估。检查不连续性或振荡。

- **任务成功率 (可选):** 如果适用于特定的操作任务 (例如, 拾取和放置), 测量与基准相比的成功率。(灵感来自²⁸)。
- **比较基准:**
 - **基准 1: 纯2.5D控制:** 整个过程由选定的2.5D姿态估计器直接控制机器人 (仅阶段1)。预期: 最低延迟, 可能最低的精度/平滑度。
 - **基准 2: 预标定物理模型控制:** 使用离线标定的、固定的用户手臂物理模型来驱动机器人 (相当于阶段3, 但没有在线估计/融合)。预期: 如果标定准确, 则精度/平滑度良好, 但需要事先标定且无法在线适应。
- **表 6.1: 拟议评估指标与比较基准**

指标 (KPI)	混合策略 - 阶段1	混合策略 - 阶段2	混合策略 - 阶段3	基准 1 (纯 2.5D)	基准 2 (预标定物理模型)
端到端延迟 (ms)	低 (Low)	中 (Medium)	中 (Medium)	低 (Low)	中 (Medium)
轨迹跟踪误差 (RMSE)	高 (High)	中 (Medium)	低 (Low)	高 (High)	低 (Low)
运动平滑度 (Jerk)	低 (Low)	中 (Medium)	高 (High)	低 (Low)	高 (High)
参数收敛时间 (s)	N/A	测量值 (Value)	N/A	N/A	N/A
切换稳定性 (定性/定量)	N/A	N/A	评估 (Evaluate)	N/A	N/A
任务成功率 (%) (可选)	评估 (Evaluate)	评估 (Evaluate)	评估 (Evaluate)	评估 (Evaluate)	评估 (Evaluate)

该表清晰地定义了如何衡量所提出系统的成功。它为查询和文献 [4, 28, 38, 67] 中确定的每个关键性能方面 (延迟、精度、平滑度、收敛性、稳定性) 建立了量化指标。与明确定义的基准进行比较, 提供了背景信息, 并允许客观评估混合策略的优势和权衡。

● 6.3. 鲁棒性测试场景:

- **快速运动:** 用户执行快速手臂运动⁴。评估延迟影响和跟踪精度。

- 遮挡: 暂时遮挡相机视图或特定身体部位(手、肘)。评估滤波器的预测能力和重新检测后的恢复能力³⁶。
- 不同起始姿态: 在用户手臂处于不同配置的情况下初始化系统。
- 光照变化: 在不同环境光照条件下进行测试(与视觉估计器相关⁶⁹)。
- 长时间运行: 评估参数估计和姿态跟踪在长时间内的漂移情况。
- **6.4. 硬件/软件环境规范:**
 - 硬件: 指定相机(型号、分辨率、FPS)、计算平台(CPU、GPU型号, 例如, Nvidia Jetson、台式PC¹⁰)、机器人手臂(型号、控制器接口)。
 - 软件: 操作系统(例如, Ubuntu/ROS⁵⁷)、关键库(OpenCV、姿态估计库, 例如 MediaPipe¹、MMPose⁷⁶、EKF库, 例如Eigen⁴³、robot_localization⁹⁹、运动学库, 例如KDL)、编程语言(Python/C++⁴³)。性能依赖于硬件¹⁰。验证必须在代表目标部署环境的硬件上进行。声明用于基准测试(第6.1, 6.2节)的具体硬件/软件对于可重复性和相关性至关重要。平台的选择(例如, 嵌入式 vs. 台式机)将严重影响可实现的延迟以及计算密集型组件(如UKF或MPC)的可行性。

7. 实施路线图与建议

- **7.1. 分阶段开发计划:**
 - 阶段 **A (组件设置与基准1)**: 实现2.5D姿态估计器(例如, MediaPipe)和直接机器人控制(基准1)。建立基本的通信流水线。
 - 阶段 **B (物理模型与基准2)**: 实现手臂运动学模型和离线标定。实现基于预标定模型的控制(基准2)。
 - 阶段 **C (在线估计与EKF融合)**: 实现(例如, 集成到EKF中的)在线参数估计。实现EKF融合模块(阶段2)。调整噪声参数。
 - 阶段 **D (过渡逻辑与完整系统)**: 实现置信度加权和过渡逻辑(阶段2 -> 阶段3)。集成所有组件。
 - 阶段 **E (测试与验证)**: 执行完整的验证计划(第6节)。根据结果进行迭代和改进。
- **7.2. 风险评估与缓解措施:**
 - 风险 **1: 实时性能**: 组合计算超出预算。缓解: 对每个组件进行性能分析; 优化关键代码(C++); 选择高效算法; 可能将计算卸载到GPU; 如有必要, 降低控制频率。
 - 风险 **2: EKF发散/不稳定**: 调整不当, 强非线性。缓解: 仔细调整噪声(Q, R); 使用鲁棒的EKF实现(例如, 平方根形式, 李群表示法⁴⁵); 监控新息序列; 实施健全性检查; 如果EKF失败, 考虑使用UKF。
 - 风险 **3: 参数收敛缓慢/不准确**: 运动激励不足。缓解: 实施强制性标定运动序列(第5.2节); 分析参数可观测性; 可能固定不太关键的参数。
 - 风险 **4: 过渡不稳定**: 控制突变。缓解: 设计平滑的加权/混合机制(第3.3.iv节); 在过渡逻辑中实施滞后处理。
- **7.3. 算法/参数调整指南:**
 - **EKF噪声协方差 (Q, R)**: 从基于传感器规格/文献的值开始, 然后使用记录的数据

进行经验性调整。R 可以与视觉估计器的置信度相关联。Q 反映了对过程模型的信任度(较低的 Q 假设运动更平滑)。

- 参数估计: 调整学习率(RLS)或EKF中参数的过程噪声。平衡适应速度与稳定性。
- 过渡阈值: 根据观察到的收敛行为和融合模式与切换到基于物理模型的控制之间的期望权衡进行调整。

- **7.4. 未来工作方向:**

- 动力学模型: 结合动力学(质量、惯性)以实现力控制或更准确的预测, 但这会增加复杂性。
- 障碍物规避: 集成环境感知和碰撞检查。
- 高级视觉输入: 使用深度相机(RGB-D)提高3D姿态精度¹⁸ 或探索事件相机以应对高速/低光照⁹⁴。NeRF/3DGS¹⁰⁵ 未来可能用于实时深度/场景理解。
- 基于学习的组件: 如果性能瓶颈持续存在, 探索用学习模型替换EKF或参数估计器, 但要注意数据要求。用于轨迹优化的扩散模型²⁹ 可以增强阶段3。

8. 结论

- **8.1. 研究结果总结:** 拟议的混合控制策略在技术上是可行的, 它利用了姿态估计¹、在线参数估计¹⁴ 和传感器融合² 领域的成熟技术。分阶段方法为平衡低延迟和高精度提供了一条合理的路径。关键挑战在于实时计算、融合稳定性以及确保快速的参数收敛。
- **8.2. 拟议框架概述:** 该系统从直接的2.5D视觉控制开始, 通过一个带有在线运动学参数估计的基于EKF的融合阶段, 过渡到依赖于精化物理模型的最终阶段。报告提供了具体的组件建议(MediaPipe、EKF、7自由度运动学模型)和详细的验证计划。
- **8.3. 结论性意见:** 与单一模态方法相比, 这种混合策略在为手臂控制创建更直观、响应更快、更精确的人机接口方面具有巨大潜力。成功的实施需要仔细的工程设计, 特别是在融合模块和过渡逻辑方面, 以及在目标硬件上进行彻底的验证。所提供的计划为开发和测试提供了一种结构化的方法。

Works cited

1. layout: forward target: <https://developers.google.com/mediapipe> ..., accessed April 20, 2025, <https://mediapipe.readthedocs.io/en/latest/solutions/hands.html>
2. Perception for Humanoid Robots, accessed April 20, 2025, <https://arxiv.org/pdf/2309.15616>
3. The Kinetics Observer: A Tightly Coupled Estimator for Legged Robots - arXiv, accessed April 20, 2025, <https://arxiv.org/html/2406.13267v1>
4. Robotic Table Tennis: A Case Study - Google Sites, accessed April 19, 2025, <https://sites.google.com/view/robotictabletennis>
5. Robotic Table Tennis: A Case Study into a High Speed Learning System, accessed April 19, 2025, <https://roboticsconference.org/2023/program/papers/006/>
6. Robotic Table Tennis: A Case Study into a High Speed Learning System - arXiv, accessed April 19, 2025, <https://arxiv.org/html/2309.03315v2>

7. Enhancing Feature Tracking Reliability for Visual Navigation using Real-Time Safety Filter, accessed April 20, 2025, https://www.researchgate.net/publication/388658267_Enhancing_Feature_Tracking_Reliability_for_Visual_Navigation_using_Real-Time_Safety_Filter
8. arXiv daily: Robotics (cs.RO) - ScienceCast, accessed April 20, 2025, https://sciencecast.org/podcasts/arxiv_daily/robotics
9. (PDF) Wearable Data Synthesis for Portable Biomechanics - ResearchGate, accessed April 20, 2025, https://www.researchgate.net/publication/381483986_Wearable_Data_Synthesis_for_Portable_Biomechanics
10. (PDF) RTMPose: Real-Time Multi-Person Pose Estimation based on ..., accessed April 19, 2025, https://www.researchgate.net/publication/369233598_RTMPose_Real-Time_Multi-Person_Pose_Estimation_based_on_MMPose
11. MediaPipe Hand Detection · Models · Dataloop, accessed April 19, 2025, https://dataloop.ai/library/model/qualcomm_mediapipe-hand-detection/
12. Hand landmarks detection guide | Google AI Edge | Google AI for ..., accessed April 19, 2025, https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker
13. Efficient Real-Time Hand Pose Estimation and Mesh Reconstruction Using Refined Joint and Skeleton Features - arXiv, accessed April 19, 2025, <https://www.arxiv.org/pdf/2503.05995>
14. Towards Self-Calibrating Inertial Body Motion Capture - ResearchGate, accessed April 20, 2025, https://www.researchgate.net/publication/303824157_Towards_Self-Calibrating_Inertial_Body_Motion_Capture
15. Human Pose Estimation - Analytics Vidhya, accessed April 20, 2025, <https://www.analyticsvidhya.com/blog/2022/01/a-comprehensive-guide-on-human-pose-estimation/>
16. wuyenlin/pebrt: A Novel Kinematic Approach to 3D Human Pose Estimation - GitHub, accessed April 20, 2025, <https://github.com/wuyenlin/pebrt>
17. tmcg0/bioslam: IMU-based human skeletal pose estimation in C++11 - GitHub, accessed April 20, 2025, <https://github.com/tmcg0/bioslam>
18. (PDF) Hand tracking for clinical applications: validation of the ..., accessed April 20, 2025, https://www.researchgate.net/publication/372858609_Hand_tracking_for_clinical_applications_validation_of_the_Google_MediaPipe_Hand_GMH_and_the_depth-enhanced_GMH-D_frameworks
19. (PDF) Adaptable force control in robotic rehabilitation - ResearchGate, accessed April 20, 2025, https://www.researchgate.net/publication/4177631_Adaptable_force_control_in_robotic_rehabilitation
20. Human multi-robot physical interaction: a distributed framework - IRIS, accessed April 20, 2025, <https://iris.unicas.it/retrieve/7ccec0c1-6bf7-4a6f-a7e5-b30affcb64b5/MAIN.pdf>

21. Using an Extended Kalman Filter for Rigid Body Pose Estimation - ResearchGate, accessed April 20, 2025, https://www.researchgate.net/publication/277485134_Using_an_Extended_Kalman_Filter_for_Rigid_Body_Pose_Estimation
22. Adaptive extended Kalman filtering for visual motion estimation of 3D objects - Lippiello, V., accessed April 20, 2025, <http://wpage.unina.it/lvillani/ftp/pubblicazioni/cep07.pdf>
23. Cross-View Tracking for Multi-Human 3D Pose Estimation at Over 100 FPS - CVF Open Access, accessed April 19, 2025, https://openaccess.thecvf.com/content_CVPR_2020/papers/Chen_Cross-View_Tracking_for_Multi-Human_3D_Pose_Estimation_at_Over_100_CVPR_2020_paper.pdf
24. HEPPPO: HARDWARE-EFFICIENT PROXIMAL POLICY OPTIMIZATION - MacSphere, accessed April 19, 2025, https://macsphere.mcmaster.ca/bitstream/11375/30655/2/taha_hazem_a_202412_masc.pdf
25. Low-Precision Reinforcement Learning - CS@Cornell, accessed April 19, 2025, https://www.cs.cornell.edu/gomes/pdf/2021_bjorck_arxiv_rl.pdf
26. Embedding Koopman Optimal Control in Robot Policy Learning - GitHub Pages, accessed April 19, 2025, https://koopman-learning.github.io/web/files/hang_embedding.pdf
27. Low-Precision Reinforcement Learning: Running Soft Actor-Critic in Half Precision - CS@Cornell, accessed April 19, 2025, https://www.cs.cornell.edu/gomes/pdf/2021_bjorck_icml_sac.pdf
28. FRMD: Fast Robot Motion Diffusion with Consistency-Distilled Movement Primitives for Smooth Action Generation - arXiv, accessed April 19, 2025, <https://arxiv.org/html/2503.02048v1>
29. arxiv.org, accessed April 19, 2025, <https://arxiv.org/pdf/2503.02048>
30. Multiple moving objects tracking based on random finite sets and Lie groups, accessed April 20, 2025, <https://lamor.fer.hr/images/50028521/Cesic2017.pdf>
31. IMU-based Suit for Strength Exercises: Design, Calibration and Tracking - kluedo, accessed April 20, 2025, https://kluedo.ub.rptu.de/files/6470/PhD_thesis_SarvenazSalehiMourkani.pdf
32. arXiv:2110.04431v1 [cs.CV] 9 Oct 2021, accessed April 20, 2025, <https://arxiv.org/pdf/2110.04431>
33. A Robust and Efficient Ensemble of Diversified Evolutionary Computing Algorithms for Accurate Robot Calibration | Request PDF - ResearchGate, accessed April 20, 2025, https://www.researchgate.net/publication/378094501_A_Robust_and_Efficient_Ensemble_of_Diversified_Evolutionary_Computing_Algorithms_for_Accurate_Robot_Calibration
34. Adaptive Kinematic Mapping Based on Chebyshev Interpolation: Application to Flexure-Jointed Micromanipulator Control | Request PDF - ResearchGate, accessed April 20, 2025, https://www.researchgate.net/publication/338004665_Adaptive_Kinematic_Mapp

[ing_Based_on_Chebyshev_Interpolation_Application_to_Flexure-Jointed_Micromanipulator_Control](#)

35. Direction-Aware Hybrid Representation Learning for 3D Hand Pose and Shape Estimation, accessed April 19, 2025, <https://arxiv.org/html/2504.01298v1>
36. Deep Learning for 3D Human Pose Estimation and Mesh Recovery: A Survey - arXiv, accessed April 19, 2025, <https://arxiv.org/html/2402.18844v1>
37. 3D Human Pose Estimation with Occlusions: Introducing BlendMimic3D Dataset and GCN Refinement - CVF Open Access, accessed April 20, 2025, https://openaccess.thecvf.com/content/CVPR2024W/ABAW/papers/Lino_3D_Human_Pose_Estimation_with_Occlusions_Introducing_BlendMimic3D_Dataset_and_CVPRW_2024_paper.pdf
38. FRMD: Fast Robot Motion Diffusion with Consistency-Distilled Movement Primitives for Smooth Action Generation - arXiv, accessed April 19, 2025, <https://arxiv.org/pdf/2503.02048?>
39. [Literature Review] FRMD: Fast Robot Motion Diffusion with Consistency-Distilled Movement Primitives for Smooth Action Generation - Moonlight, accessed April 19, 2025, <https://www.themoonlight.io/review/frmd-fast-robot-motion-diffusion-with-consistency-distilled-movement-primitives-for-smooth-action-generation>
40. [2503.02048] FRMD: Fast Robot Motion Diffusion with Consistency-Distilled Movement Primitives for Smooth Action Generation - arXiv, accessed April 19, 2025, <https://arxiv.org/abs/2503.02048>
41. PROBABILISTIC ROBOTICS, accessed April 20, 2025, <https://docs.ufpr.br/~danielsantos/ProbabilisticRobotics.pdf>
42. Trajectory Planning and Control of Cooperative Robotic System for Automated Fiber Placement - Spectrum Research Repository, accessed April 20, 2025, https://spectrum.library.concordia.ca/994048/1/Zhu_PhD_F2024.pdf
43. dekucheng/Extended-Kalman-Filter-Sensor-Fusion - GitHub, accessed April 20, 2025, <https://github.com/dekucheng/Extended-Kalman-Filter-Sensor-Fusion>
44. cggos/lidar_radar_fusion_ekf_ukf: Lidar and Radar Fusion with EKF and UKF - GitHub, accessed April 20, 2025, https://github.com/cggos/lidar_radar_fusion_ekf_ukf
45. RossHartley/invariant-ekf: C++ library to implement invariant extended Kalman filtering for aided inertial navigation. - GitHub, accessed April 20, 2025, <https://github.com/RossHartley/invariant-ekf>
46. Human multi-robot interaction: from workspace ... - EleA@Unisa, accessed April 20, 2025, http://elea.unisa.it/jspui/bitstream/10556/6209/8/tesi_di_dottorato_M_Lippi.pdf
47. Real-time inertial kinematic chain tracking with joint position estimation[version 1; peer review: awaiting, accessed April 20, 2025, https://www.dfki.de/fileadmin/user_upload/import/14792_fc2f06b2-4c59-47d6-a433-f2b169b99a9f_16939_-_bertram_taetz_%281%29.pdf
48. (PDF) Enhancing Robotic Precision in Construction: A Modular Factor Graph-Based Framework to Deflection and Backlash Compensation Using High-Accuracy Accelerometers - ResearchGate, accessed April 20, 2025,

- https://www.researchgate.net/publication/388402479_Enhancing_Robotic_Precision_in_Construction_A_Modular_Factor_Graph-Based_Framework_to_Deflection_and_Backlash_Compensation_Using_High-Accuracy_Accelerometers
49. A Factor Graph Approach to Multi-Camera Extrinsic Calibration on Legged Robots - arXiv, accessed April 20, 2025, <https://arxiv.org/pdf/1811.01254>
 50. Graph-based Multi-sensor Fusion for Consistent Localization of Autonomous Construction Robots - ResearchGate, accessed April 20, 2025, https://www.researchgate.net/publication/361969193_Graph-based_Multi-sensor_Fusion_for_Consistent_Localization_of_Autonomous_Construction_Robots
 51. A human hand motion tracking method for robot programming by demonstration - ORBi, accessed April 20, 2025, https://orbi.uliege.be/bitstream/2268/297546/1/Manuscript_RobinPellois_Dec2022.pdf
 52. MULTIPLE MODEL ADAPTIVE EXTENDED KALMAN FILTER FOR THE ROBUST LOCALIZATION OF A MOBILE ROBOT - SciTePress, accessed April 20, 2025, <https://www.scitepress.org/Papers/2007/16337/16337.pdf>
 53. Relative multiplicative extended Kalman filter for observable GPS-denied navigation, accessed April 20, 2025, <https://par.nsf.gov/servlets/purl/10316379>
 54. IMU-BASED KINEMATIC CHAIN POSE ESTIMATION USING EXTENDED KALMAN FILTER: Proceedings of the 19th International Conference on CLAWAR 2016 - ResearchGate, accessed April 20, 2025, https://www.researchgate.net/publication/305872082_IMU-BASED_KINEMATIC_CHAIN_POSE_ESTIMATION_USING_EXTENDED_KALMAN_FILTER_Proceedings_of_the_19th_International_Conference_on_CLAWAR_2016
 55. 6DOF Pose Estimation Using 3D Sensors Technical Report CSE-2010-08, accessed April 20, 2025, <https://www.eecs.yorku.ca/research/techreports/2010/CSE-2010-08.pdf>
 56. ser94mor/sensor-fusion: Filters: KF, EKF, UKF || Process Models: CV, CTRV - GitHub, accessed April 20, 2025, <https://github.com/ser94mor/sensor-fusion>
 57. rulshid/Robot-localization-using-EKF: Practical implementation of robot_localization package - GitHub, accessed April 20, 2025, <https://github.com/rulshid/Robot-localization-using-EKF>
 58. farzingkh/EKF-Localization: Application of robot_pose_ekf ROS package on turtlebot gazebo package for localization - GitHub, accessed April 20, 2025, <https://github.com/farzingkh/EKF-Localization>
 59. mithi/fusion-ekf: An extended Kalman Filter implementation in C++ for fusing lidar and radar sensor measurements. - GitHub, accessed April 20, 2025, <https://github.com/mithi/fusion-ekf>
 60. xiaozhengxu/CompRobo_IMU_Sensor_fusion: This is our final project for Computational Robotics class to incorporate a razor IMU sensor to improve the neato's wheel odometry. - GitHub, accessed April 20, 2025, https://github.com/xiaozhengxu/CompRobo_IMU_Sensor_fusion
 61. Wearable Movement Sensors for Rehabilitation - MDPI, accessed April 20, 2025, https://mdpi-res.com/bookfiles/book/4361/Wearable_Movement_Sensors_for_Rehabilitation_From_Technology_to_Clinical_Practice.pdf?v=1713488522

62. Articulated Pose Estimation via Over-parametrization and Noise Projection
Jonathan David Brookshire, accessed April 20, 2025,
https://rvsn.csail.mit.edu/Pubs/phd_brookshire_2013aug_ArticulatedPoseEstimationviaoverparametrizationandnoiseprojection.pdf
63. High-Accuracy Mobile Manipulation for On-Site Robotic Building Construction - Research Collection, accessed April 20, 2025,
https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/307844/1/tsandy_phd_thesis.pdf
64. (PDF) Digital Twin-Empowered Robotic Arm Control: An Integrated PPO and Fuzzy PID Approach - ResearchGate, accessed April 19, 2025,
https://www.researchgate.net/publication/387922790_Digital_Twin-Empowered_Robotic_Arm_Control_An_Integrated_PPO_and_Fuzzy_PID_Approach
65. RTMW: Real-Time Multi-Person 2D and 3D Whole-body Pose Estimation, accessed April 19, 2025,
<https://paperswithcode.com/paper/rtmw-real-time-multi-person-2d-and-3d-whole>
66. Real-Time Gesture-Based Hand Landmark Detection for Optimized Mobile Photo Capture and Synchronization - MDPI, accessed April 19, 2025,
<https://www.mdpi.com/2079-9292/14/4/704>
67. Quantifying Similarities Between MediaPipe and a Known Standard to Address Issues in Tracking 2D Upper Limb Trajectories: Proof of Concept Study, accessed April 19, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11683656/>
68. 1. Introduction - CEUR-WS.org, accessed April 20, 2025,
<https://ceur-ws.org/Vol-3861/paper2.pdf>
69. ThermoHands: A Benchmark for 3D Hand Pose Estimation from Egocentric Thermal Image, accessed April 20, 2025, <https://arxiv.org/html/2403.09871v1>
70. Coordinate-Corrected and Graph-Convolution-Based Hand Pose Estimation Method - MDPI, accessed April 20, 2025,
<https://www.mdpi.com/1424-8220/24/22/7289>
71. mediapipe/docs/solutions/pose.md at master - GitHub, accessed April 20, 2025,
<https://github.com/google-ai-edge/mediapipe/blob/master/docs/solutions/pose.md>
72. Hi5 : 2D Hand Pose Estimation with Zero Human Annotation - arXiv, accessed April 20, 2025, <https://arxiv.org/html/2406.03599v1>
73. Hand Tracking in 3D Space using MediaPipe and PnP Method for Intuitive Control of Virtual Globe | Request PDF - ResearchGate, accessed April 20, 2025,
https://www.researchgate.net/publication/357260494_Hand_Tracking_in_3D_Space_using_MediaPipe_and_PnP_Method_for_Intuitive_Control_of_Virtual_Globe
74. MediaPipe for Dummies - AssemblyAI, accessed April 20, 2025,
<https://www.assemblyai.com/blog/mediapipe-for-dummies>
75. ThermoHands: A Benchmark for 3D Hand Pose Estimation from, accessed April 20, 2025, https://project-archive.inf.ed.ac.uk/ug4/20244276/ug4_proj.pdf
76. Releases · open-mmlab/mmpose - GitHub, accessed April 19, 2025,
<https://github.com/open-mmlab/mmpose/releases>
77. We've released the highly efficient 2D/3D whole-body pose estimation models

- RTMW/RTMW3D. : r/computervision - Reddit, accessed April 19, 2025, https://www.reddit.com/r/computervision/comments/1e3pjlj/weve_released_the_highly_efficient_2d3d_wholebody/
78. [2407.08634] RTMW: Real-Time Multi-Person 2D and 3D Whole-body Pose Estimation, accessed April 19, 2025, <https://arxiv.org/abs/2407.08634>
 79. RTMPose: Real-Time Multi-Person Pose Estimation based on MMPose | Papers With Code, accessed April 19, 2025, <https://paperswithcode.com/paper/rtmpose-real-time-multi-person-pose>
 80. RTMW: Real-Time Multi-Person 2D and 3D Whole-body Pose Estimation - arXiv, accessed April 19, 2025, <https://arxiv.org/html/2407.08634v1>
 81. 3D Pose Estimation | Papers With Code, accessed April 19, 2025, <https://paperswithcode.com/task/3d-pose-estimation/latest>
 82. RTMPose: Real-Time Multi-Person Pose Estimation based on ..., accessed April 19, 2025, <https://arxiv.org/pdf/2303.07399>
 83. arXiv:1712.03866v1 [cs.CV] 11 Dec 2017, accessed April 19, 2025, <http://arxiv.org/pdf/1712.03866>
 84. ReJSHand: Efficient Real-Time Hand Pose Estimation and Mesh Reconstruction Using Refined Joint and Skeleton Features | Request PDF - ResearchGate, accessed April 19, 2025, https://www.researchgate.net/publication/389714661_ReJSHand_Efficient_Real-Time_Hand_Pose_Estimation_and_Mesh_Reconstruction_Using_Refined_Joint_and_Skeleton_Features
 85. ReJSHand: Efficient Real-Time Hand Pose Estimation and Mesh Reconstruction Using Refined Joint and Skeleton Features - arXiv, accessed April 19, 2025, <https://arxiv.org/html/2503.05995v1>
 86. [2503.05995] ReJSHand: Efficient Real-Time Hand Pose Estimation and Mesh Reconstruction Using Refined Joint and Skeleton Features - arXiv, accessed April 19, 2025, <https://arxiv.org/abs/2503.05995>
 87. Using a single RGB frame for real time 3D hand pose estimation in the wild - ar5iv - arXiv, accessed April 19, 2025, <https://ar5iv.labs.arxiv.org/html/1712.03866>
 88. What is OpenPose? A Guide for Beginners. - Roboflow Blog, accessed April 19, 2025, <https://blog.roboflow.com/what-is-openpose/>
 89. www.scitepress.org, accessed April 19, 2025, <https://www.scitepress.org/Papers/2019/73947/73947.pdf>
 90. FrankMocap: A Monocular 3D Whole-Body Pose Estimation System via Regression and Integration - ResearchGate, accessed April 19, 2025, https://www.researchgate.net/publication/356510588_FrankMocap_A_Monocular_3D_Whole-Body_Pose_Estimation_System_via_Regression_and_Integration
 91. FrankMocap: A Monocular 3D Whole-Body Pose Estimation System via Regression and Integration - CVF Open Access, accessed April 19, 2025, https://openaccess.thecvf.com/content/ICCV2021W/ACVR/papers/Rong_FrankMocap_A_Monocular_3D_Whole-Body_Pose_Estimation_System_via_Regression_ICCVW_2021_paper.pdf
 92. FrankMocap: Fast Monocular 3D Hand and Body Motion Capture by Regression and Integration - Yu Rong, accessed April 19, 2025,

- https://penincillin.github.io/frankmocap_iccvw2021
93. [2008.08324] FrankMocap: Fast Monocular 3D Hand and Body ..., accessed April 19, 2025, <https://arxiv.org/html/2008.08324>
 94. 3D Pose Estimation of Two Interacting Hands from a Monocular Event Camera - OpenReview, accessed April 19, 2025, <https://openreview.net/pdf?id=DExV8hqPOU>
 95. 3D Pose Estimation of Two Interacting Hands from a Monocular Event Camera - arXiv, accessed April 19, 2025, <https://arxiv.org/html/2312.14157v1>
 96. TransPose: Real-time 3D Human Translation and Pose Estimation with Six Inertial Sensors - Xinyu Yi, accessed April 20, 2025, <https://xinyu-yi.github.io/TransPose/files/TransPose.pdf>
 97. Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model - MDPI, accessed April 20, 2025, <https://www.mdpi.com/2076-3417/13/4/2700>
 98. Accepted at 6th Workshop and Competition on Affective Behavior Analysis in-the-wild CVPR 2024 Workshop 3D Human Pose Estimation with Occlusions: Introducing BlendMimic3D Dataset and GCN Refinement - arXiv, accessed April 20, 2025, <https://arxiv.org/html/2404.16136v1>
 99. ROS and Sensor Fusion Tutorial.md at master - GitHub, accessed April 20, 2025, <https://github.com/methylDragon/ros-sensor-fusion-tutorial/blob/master/01%20-%20ROS%20and%20Sensor%20Fusion%20Tutorial.md>
 100. methylDragon/ros-sensor-fusion-tutorial - GitHub, accessed April 20, 2025, <https://github.com/methylDragon/ros-sensor-fusion-tutorial>
 101. Toward Approaches to Scalability in 3D Human Pose Estimation - OpenReview, accessed April 20, 2025, [https://openreview.net/forum?id=xse8QMgnyM&referrer=%5Bthe%20profile%20of%20Seong-Whan%20Lee%5D\(%2Fprofile%3Fid%3D~Seong-Whan_Lee3\)](https://openreview.net/forum?id=xse8QMgnyM&referrer=%5Bthe%20profile%20of%20Seong-Whan%20Lee%5D(%2Fprofile%3Fid%3D~Seong-Whan_Lee3))
 102. Optimizing Autonomous Vehicle Performance Using Improved Proximal Policy Optimization, accessed April 19, 2025, <https://www.mdpi.com/1424-8220/25/6/1941>
 103. 3D Human Pose Estimation Experiments and Analysis - KDnuggets, accessed April 19, 2025, <https://www.kdnuggets.com/2020/08/3d-human-pose-estimation-experiments-analysis.html>
 104. CVPR Poster EventNeRF: Neural Radiance Fields From a Single Colour Event Camera, accessed April 19, 2025, <https://cvpr2023.thecvf.com/virtual/2023/poster/22210>
 105. 3D-Vision-World/awesome-NeRF-and-3DGS-SLAM - GitHub, accessed April 19, 2025, <https://github.com/3D-Vision-World/awesome-NeRF-and-3DGS-SLAM>
 106. Evaluating Human Perception of Novel View Synthesis: Subjective Quality Assessment of Gaussian Splatting and NeRF in Dynamic Scenes - arXiv, accessed April 19, 2025, <https://arxiv.org/html/2501.08072v1>
 107. Dynamic NeRF: A Review - arXiv, accessed April 19, 2025, <https://arxiv.org/html/2405.08609>
 108. Quark: Real-time, High-resolution, and General Neural View Synthesis - arXiv,

accessed April 19, 2025, <https://arxiv.org/html/2411.16680v1>