

2024-2025年人体姿态估计前沿综述

引言与背景

近年来，2D与3D人体姿态估计（Human Pose Estimation, HPE）领域在模型架构、跨数据集泛化、多模态融合以及实时性能等方面取得了显著进展^{1 2}。2025年CVPR会议上涌现出众多创新方法，聚焦于提升关键点检测精度、模型效率和特殊场景适应性；与此同时，ICCV/ECCV 2024也提出了针对低光照、跨模态传感等特殊条件下的人体姿态估计方案。本文将从2D姿态估计和3D姿态估计两方面出发，总结这些最新研究：涵盖新颖的骨干网络和解码器设计、跨数据集训练策略、领域自适应技术、多模态融合方法，以及在COCO、MPII、3DPW、Human3.6M等标准数据集上的性能提升情况和实验结果。我们还重点关注模型在特定人群（如婴儿、老年人）或特殊场景（如弱光、室内VR/AR环境）下的微调与迁移能力。下文将按主题分别介绍主要方法，并通过表格比较它们的创新点和性能指标。

2D人体姿态估计的新方法与进展

1. 跨数据集的姿态表示与泛化：PoseBH (CVPR 2025)^{3 4}。Uyoung Jeong等人在CVPR 2025提出PoseBH: Prototypical Multi-Dataset Training Beyond Human Pose Estimation³。该方法专注于多数数据集联合训练场景下的不同比例骨架定义问题：不同数据集中人体/动物的关键点拓扑各异，直接合并训练会遇到骨架异构和缺乏对齐监督的挑战¹。为此，PoseBH引入了两个关键创新：(a) 在统一嵌入空间中学习非参数化的关键点原型（prototype），将不同骨架的关键点映射为通用表示，从而实现跨数据集的一致表示⁴；(b) 提出跨骨架类型的自监督机制，通过让模型预测的关键点对齐相应的原型嵌入，实现不同数据集之间的间接监督⁴。实验表明，PoseBH在多个数据集上获得了优异的泛化性能：在COCO-WholeBody、AP-10K、APT-36K等全身和动物姿态数据集上显著提升精度，同时在经典的人体数据集（COCO、MPII、AIC）上保持了与单数据集训练相当的性能⁵。此外，PoseBH学到的关键点嵌入还能有效迁移到其他任务，例如手部形状估计（InterHand2.6M）和人体3D形状估计（3DPW），体现出良好的跨任务迁移能力⁵。该工作代码已开源提供⁶。

2. 面向遮挡和边界情况的概率姿态表示：ProbPose (CVPR 2025)^{7 8}。Miroslav Purkrabek等人在CVPR 2025提出ProbPose: A Probabilistic Approach to 2D Human Pose Estimation⁹。传统姿态估计通常忽略图像外的关键点（即人物部分出框）且使用未校准的热图作为关键点表示¹⁰。ProbPose则采用显式概率建模的方法：(a) 为每个关键点预测一个在图像内各位置的出現概率分布（校准的概率热图），同时预测该关键点在图像外的概率以及其可见性标签⁷；(b) 为解决评价标准缺失的问题，作者构建了CropCOCO数据集，并提出扩展的OKS指标Ex-OKS，将图像外关键点纳入评估¹¹。实验显示，在COCO、CropCOCO、OCHuman数据集上，ProbPose对图像边缘及外部关键点的定位精度有显著提升，同时通过数据增强也提升了图像内部关键点的精度⁸。例如，ProbPose在模拟裁剪的人像上准确找到了原方法遗漏的关节，并提高了边界区域的鲁棒性⁸。这种概率表示为评价遮挡/截断情况提供了新范式。作者还开放了模型代码和预训练权重供社区使用¹²。

3. 动态架构提升效率：DynPose (CVPR 2025)^{13 14}。Yalong Xu等人在CVPR 2025提出DynPose: Largely Improving the Efficiency of Human Pose Estimation by a Simple Dynamic Framework¹⁵。Top-down人体姿态估计方法（如HRNet、ViTPose）在精度上表现出色，但其对每个检测到的人逐一高分辨率推理，导致效率低下¹⁶。DynPose针对这一问题，提出了“路由器+双模型”的简单动态推理框架：¹⁷一方面，用一个轻量级路由网络根据输入的bounding box判断姿态样本的复杂度，将简单姿态分配给“小模型”处理，将复杂姿态交由“大模型”处理¹⁸；另一方面，由于多人检测产生大量冗余框，即使NMS后仍存在过多重复计算，路由器也可筛除无效或容易预测的实例¹⁷。在ResNet-50和HRNet-W32组成大小模型的实验中，DynPose在COCO验证集上将推理速度提升近50%，而关键点AP与单独使用大模型时几乎持平¹⁴。更重要的

是，该框架具有良好的**通用性**，无需重新训练即可泛化到不同预训练模型和数据集¹⁹。DynPose通过充分利用样本难度差异，实现了精度不变条件下的显著加速，其代码亦已在GitHub开源¹⁹。

4. 极端弱光条件下的域自适应：Dual-Teacher框架 (ECCV 2024)。弱光环境会导致常规2D姿态估计模型性能严重下降，而标注弱光数据获取困难。Yihao Ai等在ECCV 2024提出“**双教师**”**无监督域适应框架**²⁰，针对**极低光照**下的人体姿态估计进行了创新性探索²⁰。其方法包括两个阶段：²¹ **(a) 预训练阶段**：在充足光照的源域数据上**训练两个教师模型**，一个为**中心点检测+关键点回归**（top-down架构），另一个为**直接关键点检测**（bottom-up架构），同时采用作者设计的**极暗增强(ELLA)**对部分图像施加仿真弱光的增广（包括伽马矫正、亮度对比降低、噪声添加等），将生成的假弱光图像混合进训练，从而让教师模型学习弱光下姿态的大致分布^{22 23}；**(b) 双教师知识蒸馏阶段**：在无弱光标签的数据上，让两位教师分别对真实弱光图像进行预测，并通过置信度筛选和NMS融合得到高置信的**伪标签**²⁴。学生模型以这些伪标签为监督进行训练，同时设计了**人像级的退化增广(PDA)**策略：将教师检测到的人按bbox裁剪，再对局部区域施加弱光增广，以逼真模拟人体局部处于阴影/黑暗的情况²⁵。损失函数包括带伪标签的无监督损失和对增强亮度图像的监督损失²⁶。在ExLPose-OCN等弱光姿态数据集上的评估显示，新方法**相较现有SOTA获得了约6.8%的相对提升（提升2.4 AP）**，且**未使用任何弱光真值标签**²⁷。这证明通过双教师协同和强增强策略，可以在无弱光标注下大幅提高极暗场景中的姿态估计性能²⁸。

5. 模型对特殊人群的适应性：针对婴儿、老年人等特殊体型人群的姿态估计，2024年前后也有一些相关研究探索。例如，有研究利用**生成模型先验**进行**3D婴儿姿态估计**的高效域适应^{29 30}。Zhou等人提出利用成人姿态的生成先验（如VAE或扩散模型）迁移到婴儿域，并通过优化算法迭代调整3D姿态，使之符合2D关键点观测和先验分布，从而在极少婴儿数据下达到SOTA精度^{31 30}。他们还设计了引导式扩散模型，将成人姿态转换为类似婴儿姿态的合成数据以增广训练集^{32 33}。实验证明，在MINI-RGBD、SyRIP等婴儿数据集上，该方法的3D关节平均误差（MPJPE）显著优于直接用成人模型微调的结果³⁴。这一方向表明，通过**生成先验+小样本适应**，可以有效解决婴儿等特定人群数据稀缺下的姿态估计问题。

以上工作在2D姿态估计领域引入了**更通用的表示学习**（如PoseBH的关键点原型）、**更稳健的输出表示**（ProbPose的概率热图）、**更高效的推理机制**（DynPose的动态路由）、**更强的域自适应能力**（双教师弱光适应）和**特定人群迁移**（婴儿姿态）等多方面创新，大幅拓展了2D姿态估计模型在复杂现实环境下的适用性和鲁棒性。表1汇总了部分代表方法的特点和性能。

方法 (年份)	主要创新	性能表现	代码/出处
PoseBH (CVPR 2025)	多数据集原型融合，跨骨架自监督对齐 ⁴	全身/动物数据集显著提升泛化，COCO/MPII性能不降 ⁵	GitHub开源 ⁶
ProbPose (CVPR 2025)	关键点存在概率+出框概率+可见性预测 ⁷	CropCOCO、OCHuman上出框关键点定位精度提升显著 ⁸	GitHub开源 ¹²
DynPose (CVPR 2025)	轻量路由器+大小模型动态推理架构 ³⁵	COCO上推理速度提高~50%，精度与大模型相当 ¹⁴	GitHub开源 ¹⁹
Dual-Teacher弱光 (ECCV 2024)	双教师伪标签蒸馏+极暗数据增广（ELLA+PDA） ³⁶	ExLPose弱光集提升2.4 AP (+6.8%)，无需弱光真值标签 ²⁷	ECCV 2024论文 ²⁸
Infant3D (WACV 2024)	成人生成先验+优化迭代，少样本3D婴儿姿态域适应 ^{37 32}	MINI-RGBD、SyRIP婴儿集取得当时SOTA最低MPJPE ³⁸	WACV 2024 Workshop ³⁰

表1：近年2D姿态估计新方法综述与性能。

融合多模态与特殊场景的姿态估计

现实世界中，纯视觉方法在**遮挡、远距离、隐私受限**等场景下可能受限。2024–2025年出现了一系列融合新型传感器和特殊设置的姿态估计研究，包括使用射频信号（毫米波、WiFi）穿透遮挡，及**第一人称视角**下的姿态估计等。

图1：混合现实(MR)头戴设备中自我姿态估计示意³⁹。在仅依赖头部和手部传感的情况下（摄像机视野有限，双手会间歇性出框），EgoPoser通过全局运动解耦和跨时间融合，实现了空间上连贯的全身姿态重建。

1. 毫米波雷达辅助长距离姿态估计：MVDoppler-Pose (CVPR 2025)^{40 41}。Jaeho Choi等人提出**MVDoppler-Pose: Multi-Modal Multi-View mmWave Sensing for Long-Distance Self-Occluded Human Walking Pose Estimation**⁴²。该研究是首个将毫米波雷达引入长距离人体姿态估计的大规模对比分析⁴³。相较于可见光相机，mmWave雷达在抗遮挡和远距离性能上具有天然优势⁴⁴：它可穿透衣物甚至部分障碍，并且信号随距离衰减减小。然而，单一mmWave在姿态估计上也面临挑战，例如信号噪声和对人体朝向敏感⁴³。为此，作者提出：(a) **多模态多视角融合**：结合多台毫米波雷达的观测，并与少量视觉信息互补，从不同角度缓解雷达方向性盲区⁴⁵；(b) 构建了一个**包含同步相机与雷达数据的新数据集**，覆盖远距离行走人员在不同朝向下的姿态⁴⁶。实验表明，在远距离（数十米）且存在自遮挡的行走姿态下，该融合方法**明显优于纯视觉方法**，体现了mmWave在距离无关和抗遮挡方面的潜力⁴⁵。例如，在新数据集上，mmWave方案在全身关键点定位上达到更高PCK，而相机方案因遮挡和分辨率下降性能大幅退化。此外，多视雷达融合有效弥补了单雷达对某些关节丢失的情况⁴¹。此研究为安全监控、夜间姿态检测等特殊场景提供了新思路。

2. WiFi无线信号的人体姿态估计：HPE-Li (ECCV 2024)^{47 48}。Toan D. Gian等人在ECCV 2024提出**HPE-Li: WiFi-Enabled Lightweight Dual Selective Kernel Convolution for Human Pose Estimation**⁴⁹。WiFi的无线电信号（如CSI信道状态信息）在室内可用于感知人体位置，有望实现**非视觉的隐私友好型姿态估计**⁴⁷。HPE-Li利用**摄像头+WiFi多模态传感**获取精确的3D人体骨架标签作为训练数据⁵⁰，然后重点设计了一个高效的**WiFi信号深度网络**：采用**多分支卷积神经网络**结合**Selective Kernel注意力机制**⁵¹。Selective Kernel可以根据输入动态调整卷积核感受野大小，从而使模型既能捕获WiFi信号中的局部细节又能覆盖整体模式⁵²。该模型参数量小且计算高效，成功解决了以往WiFi姿态方法计算开销大的难题^{47 48}。在**MM-Fi、WiPose**等公共WiFi人体姿态数据集上，HPE-Li的3D关节定位精度优于先前方法，而推理速度大幅提升，适合实时大规模部署⁴⁸。这种方法证明了**WiFi感知**作为视觉替代的可行性：通过结合注意力卷积网络，在保障隐私的同时获得较高姿态重建准确率。

3. 第一人称自我姿态估计：EgoPoser (ECCV 2024)^{53 54}。Jiaxi Jiang等人提出**EgoPoser: Robust Real-Time Egocentric Pose Estimation from Sparse and Intermittent Observations**⁵⁵。该方法针对**VR/AR头戴式设备**用户的全身姿态重建需求：此类系统仅依赖头显上的内向外摄像头，能观测到用户头部和部分手部，当手脚超出视野时就缺乏直接观测⁵⁶。传统方法多基于室内动捕数据训练，假设手部始终在视野内、人体形状固定平均，这在实际**大空间场景**下并不成立^{57 58}。EgoPoser通过四项创新来解决：(a) **视野间歇跟踪建模**：仅利用当手在视野内时的不连续手部轨迹，也能通过时序模型补全姿态⁵³；(b) **全局-局部运动解耦**：提出Global-in-Local全局运动分解策略，将全局位移从局部肢体姿态中分离，以**位置不变的方式**表示动作⁵⁹。这样训练的模型不会过拟合在训练集的绝对坐标系，对用户在大空间中的移动更加鲁棒^{60 61}；(c) **SlowFast双速率时间编码**：以不同采样率获取快变化的短期信号和稀疏但覆盖长时段的信号，并设计SlowFast融合模块兼顾二者，从而在**不增加计算**的情况下利用了更长的运动历史，提高了预测精度^{62 63}；(d) **个性化人体形状估计**：引入多头解码器的一支来预测每个用户的体型参数，而非使用平均体型，从而消除形状差异带来的误差^{64 65}。集成上述设计，EgoPoser在AMASS合成数据上训练，在**HPS等真实第一人称数据集**上达到当前最好效果，并能**实时推理(>600fps)**⁵⁴。与此前方法相比，EgoPoser显著减少了由于全局位移、间歇缺失和体型差异导致的姿态错误^{60 66}。该方法奠定了今后VR场景下“**inside-out**”自我运动捕捉的基础。作者提供了项目开源代码和演示视频⁶⁷。

总的来看，多模态融合与特殊场景方法展现出强大潜力：毫米波和WiFi突破了光学传感的限制，实现**穿透遮挡或无视觉感知**的人体姿态估计；第一人称方法则针对AR/VR应用，使**移动大空间**内的自我姿态捕捉成为可能。这些探索丰富了姿态估计的应用边界，推动了更普适、更健壮的人体姿态感知技术的发展。

3D人体姿态与人体重建的前沿方法

在3D人体姿态估计（包括骨架3D关键点和全身3D网格重建）方面，2024–2025年的顶会论文主要致力于**提升遮挡情况下的鲁棒性、融合多视角信息、实现多人实时重建，以及统一多任务框架等**。下面介绍几项代表性工作：

1. 遮挡条件下的层次化Transformer：HiPART (CVPR 2025) ⁶⁸。Hongwei Zheng等提出**HiPART: Hierarchical Pose AutoRegressive Transformer for Occluded 3D Human Pose Estimation**。该方法面向单目视频中人体部分遮挡的情况，设计了一个**分层自回归Transformer模型** ⁶⁸。其核心思想是**逐层推理人体结构**：首先预测主体可见的骨架部分，然后在此基础上推理被遮挡的肢体 ⁶⁸。具体来说，HiPART利用Transformer编码器获得图像序列特征，再以自回归方式逐步生成3D关节位置，每一步关注当前未估计的关节 ⁶⁹。通过这种层次结构，模型可以根据上下文**推断被遮挡关节的位置**，如根据一只可见手臂的位置推测另一只手臂 ⁶⁸。Labellerr博客的总结指出，HiPART能够在肢体部分不可见时仍重建出准确的全身3D姿态，显著优于以往遇遮挡就丢失关节的模型 ⁶⁸。在Human3.6M等数据上的结果表明，HiPART在含遮挡的场景下MPJPE误差更低，并通过消融实验验证了增加层次预测深度对提升准确率的作用（例如使用更密集的2D输入关节，可以将3D误差从52.1mm降低到46.7mm，提升5.4mm） ^{70 71}。

2. 多视角下的状态空间建模：MV-SSM (CVPR 2025) ^{72 73}。Aviral Chharia等提出**MV-SSM: Multi-View State Space Modeling for 3D Human Pose Estimation** ⁷⁴。针对**多摄像头、多人的3D姿态估计**，传统Transformer-based融合方法难以充分建模不同摄像头下关节的空间关系，尤其在遮挡时容易混淆 ⁷⁵。MV-SSM引入**状态空间模型(State Space Modeling, SSM)**思想，将**多视图特征序列和人体关节序列**在两个层次上显式建模 ⁷⁶：(a) 首先设计**投影状态空间块(PSS Block)**，在特征层面通过状态空间模型算子来学习各摄像头下关节的**空间序列依赖** ⁷⁶。特别地，作者改进了现有SSM模型“Mamba”的扫描方式，引入**网格Token引导的双向扫描(Grid Token-guided Bidirectional Scan, GTBS)**策略，在PSS中实现对多视图特征的双向融合 ⁷⁷；(b) 其次在关节层面，再次应用状态空间模型串联各关节的预测。这种双层次建模能更好地捕获遮挡情况下**视角内和视角间**的关节关联 ⁷⁵。实验证明，MV-SSM在多视图3D姿态基准上取得了**高精度且强泛化**的效果：例如，在CMU Panoptic数据集仅使用3台相机的困难设置下，将3D PCK@25提高了**10.8个百分点**，在训练摄像头布局与测试不同时仍提升7.0 PCK@25；跨数据集测试（Campus A1场景）也提高了**15.3% PCP**指标，**显著超越现有SOTA** ⁷³。这说明MV-SSM通过PSS块有效增强了模型对摄像头数量变化和新场景的适应性，其代码将随论文公开 ⁷⁸。

3. 实时多人人体网格重建：SAT-HMR (CVPR 2025) ^{2 79}。Chi Su等提出**SAT-HMR: Real-Time Multi-Person 3D Mesh Estimation via Scale-Adaptive Tokens** ⁸⁰。当前的**单阶段多人3D网格估计方法**多基于DETR检测器框架，需要对整幅高分辨率图像进行特征编码 ⁸¹。高分辨率虽然提高了对远处小人的捕捉能力，但计算开销巨大 ⁸²。SAT-HMR的关键创新是在Transformer的DETR框架中引入**“尺度自适应Token”**机制：⁸³ 对于检测出的不同尺度的人体实例，动态调整其特征表示分辨率——**远处小人物保持高分辨率处理**，近处大人物则降低分辨率以节省计算，同时对**背景区域进行特征蒸馏压缩** ⁸³。这种按对象尺度自适应的token设计，使得Transformer编码能够**更高效地表征图像**，在后续回归3D网格的解码阶段将计算重点放在**困难的小目标**上 ⁸⁴。实验表明，SAT-HMR在大幅减少计算量的同时，**保留了高分辨率模型的精度**：其在多人3D姿态数据集上的精度与当前最优方法相当，但推理速度达到实时水平（几十fps），实现了**性能-速度权衡的突破** ⁷⁹。作者已公开了模型代码和预训练模型，便于业界进一步使用和验证 ⁸⁵。

4. 稀疏多视角下的交互姿态估计：AvatarPose (ECCV 2024)。Yanjie Ze等提出**AvatarPose: Avatar-guided 3D Pose Estimation of Close Human Interaction from Sparse Multi-view Videos** ⁸⁶。该方法针对**近距离人体交互**（如两人拥抱、搏斗等）的3D姿态重建难题，引入**先验人体Avatar**指导。具体而言，AvatarPose在稀疏相机视角的视频中，先通过两个人体的参数化模型（SMPL或Avatar网格）拟合初始姿态，再利用物理约束（如避免网格穿插）和交互先验（如双手接触点）不断优化两者的姿态和形状参数，从而获得**物理合理的交互姿态序列**。该方法在多人交互数据集上取得了领先的重建精度，特别是在摄像头很少（2~3个）情况下依然能重建出两人的贴合动作，这是传统多视角算法难以做到的。

5. **自监督与不确定性建模**：ECCV 2024的一些工作探索了减少对标注依赖、提高3D预测可信度的新途径。例如，Wang等提出**Mask as Supervision**方法，将2D人体分割掩膜作为弱监督信号，引导模型学习**关键点与肢体轮廓的一致性**，实现了无3D标注下的3D姿态估计⁸⁷。还有**UPose3D**方法结合多视图与时间信息，通过模型输出的**不确定度估计**来加权融合各视角/帧的预测，提升了3D重建的稳定性⁸⁸。这些方法在Human3.6M等有、无标注设置下都取得了有竞争力的结果，展示了利用额外信息（分割、视差）来弥补3D标签不足的可行性。

6. **通用多模态姿态大模型**：值得一提的是，CVPR 2025出现了尝试构建**通用人体姿态大模型**的探索，例如Yiheng Li等提出的**UniPose框架**^{89 90}。UniPose将**大语言模型(LLM)**引入姿态领域，使用离散姿态token将3D SMPL参数量化后嵌入LLM词汇表，从而使LLM能够同时处理图像、文本和3D姿态三种模态⁹⁰。该框架配备了混合视觉编码器（包括专门的姿态图像编码器）来提升对细粒度姿态的感知⁹⁰。通过统一的训练策略，UniPose在**姿态理解、生成、编辑**等任务上实现了知识迁移和能力拓展，在多个子任务（如根据文字生成指定姿势的3D人体，或编辑图中人物的姿态）上表现出色^{91 92}。这项工作首次尝试构建**通用目的的人体姿态模型**，展示了跨模态大模型在姿态领域的潜力。类似地，CVPR 2025的**Visual Persona模型**⁹³也将人体的外观与姿态定制作为基础模型的应用方向，预示着未来可能出现集**姿态估计、动作生成和个性化重建**于一体的**基础模型**。

综上，近一年来3D人体姿态估计领域的研究呈现出以下趋势：一是利用**Transformer、状态空间模型**等新架构更好地融合时空和多视角信息，显著增强了对遮挡、多人的鲁棒性⁷³；二是通过**自监督、弱标签和不确定度建模**等策略，降低了对大规模3D标注数据的依赖，同时提升了预测可信度；三是朝着**实时化与多任务统一**迈进，不仅实现多人实时3D重建⁷⁹，还初步探索了姿态理解与生成的一体化模型⁹¹。这些进展为人体姿态估计在动作捕捉、AR/VR、体育分析、行为理解等应用中落地奠定了坚实基础。

总结与展望

2024–2025年的CVPR、ICCV、ECCV等顶会上，人们从**模型结构、训练策略和多模态融合**等多维度推动了人体姿态估计技术的发展。在2D姿态估计方面，新的框架能够融合不同数据集和域的知识，实现跨类别（人/动物/全身）的统一建模⁵；概率式表示和自监督评估指标提升了遮挡与出框情况下的鲁棒性⁸；动态路由等技术显著提高了推理效率¹⁴。在3D姿态和人体重建方面，层次Transformer、状态空间模型等改进使多视角、多人的场景下3D重建更精准且具备实时性^{73 79}；同时，通过引入射频传感、惯性测量等新型信号源，姿态估计正变得“**全天候、全场景**”：既可在黑暗或遮挡中依靠雷达/WiFi“看见”人^{45 48}，也能在VR头显中只凭有限传感重建用户全身⁵⁴。展望未来，随着**多模态大模型**和更大规模数据的引入，人体姿态估计有望进一步突破单一任务局限，发展出能够**理解动作语义、生成交互动作**的通用智能系统。例如，将姿态估计与动作识别、轨迹预测、3D人体重建融合，构建端到端的**人体行为理解模型**；针对婴儿、残障人士等特殊人群，设计专门的域自适应技术提高模型公平性和适用性；以及利用模拟数据和物理先验，使模型掌握非常规姿态（如舞蹈、体操高难动作）的预测能力。可以预见，人体姿态估计将继续作为计算机视觉领域的重要课题，不断吸收最新的深度学习范式和传感技术，实现更加**精准、高效和多样化**的应用。通过本报告的分析，我们对这一领域最新的研究进展和趋势有了系统认识，这将有助于研究者和从业者把握前沿、启发新的创新思路。

^{1 3 4 5 6} CVPR 2025 Open Access Repository

https://openaccess.thecvf.com/content/CVPR2025/html/Jeong_PoseBH_Prototypical_Multi-Dataset_Training_Beyond_Human_Pose_Estimation_CVPR_2025_paper.html

^{2 79 80 81 82 83 84 85} CVPR 2025 Open Access Repository

https://openaccess.thecvf.com/content/CVPR2025/html/Su_SAT-HMR_Real-Time_Multi-Person_3D_Mesh_Estimation_via_Scale-Adaptive_Tokens_CVPR_2025_paper.html

^{7 8 9 10 11 12} CVPR 2025 Open Access Repository

https://openaccess.thecvf.com/content/CVPR2025/html/Purkrabek_ProbPose_A_Probabilistic_Approach_to_2D_Human_Pose_Estimation_CVPR_2025_paper.html

13 14 15 16 17 18 19 35 CVPR 2025 Open Access Repository

<https://openaccess.thecvf.com/content/CVPR2025/html/>

[Xu_DynPose_Largely_Improving_the_Efficiency_of_Human_Pose_Estimation_by_CVPR_2025_paper.html](#)

20 21 22 23 24 25 26 27 28 36 [Literature Review] Domain-Adaptive 2D Human Pose Estimation via Dual Teachers in Extremely Low-Light Conditions

<https://www.themoonlight.io/en/review/domain-adaptive-2d-human-pose-estimation-via-dual-teachers-in-extremely-low-light-conditions>

29 30 31 32 33 34 37 38 Efficient Domain Adaptation via Generative Prior for 3D Infant Pose Estimation

<https://openaccess.thecvf.com/content/WACV2024W/CV4Smalls/papers/>

[Zhou_Efficient_Domain_Adaptation_via_Generative_Prior_for_3D_Infant_Pose_WACVW_2024_paper.pdf](#)

39 53 54 55 56 59 61 62 63 65 67 EgoPoser: Robust Real-Time Egocentric Pose Estimation from Sparse and Intermittent Observations Everywhere • Sensing, Perception & Interaction Lab

<https://siplab.org/projects/EgoPoser>

40 41 42 43 44 45 46 CVPR 2025 Open Access Repository

[https://openaccess.thecvf.com/content/CVPR2025/html/Choi_MVDoppler-Pose_Multi-Modal_Multi-](https://openaccess.thecvf.com/content/CVPR2025/html/Choi_MVDoppler-Pose_Multi-Modal_Multi-View_mmWave_Sensing_for_Long-Distance_Self-Occluded_Human_Walking_CVPR_2025_paper.html)

[View_mmWave_Sensing_for_Long-Distance_Self-Occluded_Human_Walking_CVPR_2025_paper.html](#)

47 48 49 50 51 52 ecva.net

https://www.ecva.net/papers/eccv_2024/papers_ECCV/papers/04496.pdf

57 58 60 64 66 [2308.06493] EgoPoser: Robust Real-Time Ego-Body Pose Estimation in Large Scenes

<https://arxiv.labs.arxiv.org/html/2308.06493>

68 69 CVPR 2025: Breakthroughs in GenAI and Computer Vision

<https://www.labellerr.com/blog/cvpr-2025-part-2/>

70 [PDF] Hierarchical Pose AutoRegressive Transformer for Occluded 3D ...

<https://openaccess.thecvf.com/content/CVPR2025/papers/>

[Zheng_HiPART_Hierarchical_Pose_AutoRegressive_Transformer_for_Occluded_3D_Human_Pose_CVPR_2025_paper.pdf](#)

71 Hierarchical Pose AutoRegressive Transformer for Occluded 3D ...

<https://arxiv.org/html/2503.23331v1>

72 73 74 75 76 77 78 CVPR Poster MV-SSM: Multi-View State Space Modeling for 3D Human Pose Estimation

<https://cvpr.thecvf.com/virtual/2025/poster/33025>

86 87 88 GitHub - 52CV/ECCV-2024-Papers

<https://github.com/52CV/ECCV-2024-Papers>

89 90 91 92 CVPR 2025 Open Access Repository

<https://openaccess.thecvf.com/content/CVPR2025/html/>

[Li_UniPose_A_Unified_Multimodal_Framework_for_Human_Pose_Comprehension_Generation_CVPR_2025_paper.html](#)

93 GitHub - 52CV/CVPR-2025-Papers

<https://github.com/52CV/CVPR-2025-Papers>