arXiv:2601.01360v1 [cs.CV] 4 Jan 2026

服装惯性去噪器：通过宽松穿戴惯性测量单元去噪器实现精准动作捕捉

Jiawei Fang[1*] Ruonan Zheng[1*] Xiaoxia Gao[1*] Shifan Jiang[1] Anjun Chen[1†] Qi Ye[2] Shihui Guo[1]厦门大学 [2]浙江大学

摘要

可穿戴惯性动作捕捉（*MoCap*）具有便携、无须担忧遮挡且保护隐私等优点，是基于相机的系统的一种替代方案，但其精度依赖于紧密贴附的传感器，而这对日常使用而言既不方便也不舒适。将惯性测量单元嵌入宽松服装是理想的替代方案，但传感器与身体之间的位移会引入严重、结构化且随位置变化的误差，导致标准惯性处理流程失效。我们提出***GID***（服装惯性去噪器），这是一种轻量级、即插即用的 *Transformer*，可将宽松穿戴动作捕捉分解为三个阶段：（*i*）位置特定去噪；（*ii*）自适应跨穿戴融合；（*iii*）通用姿态预测。*GID*采用位置感知专家架构，由共享的时空骨干网络建模全局动作，同时由逐惯性测量单元专家头专门处理局部服装动力学，并通过轻量级融合模块确保跨部位一致性。这种归纳偏置能够实现稳定训练，并从有限的成对宽松与紧身惯性测量单元数据中有效学习。我们还推出***GarMoCap***，这是一个结合了公开数据与新采集数据的数据集，涵盖多样的用户、动作和服装。实验表明，*GID*仅通过单用户训练即可实现精准的实时去噪，并能泛化至未见过的用户、动作和服装类型；将其作为即插即用模块使用时，还能持续提升最先进的惯性动作捕捉方法的性能。

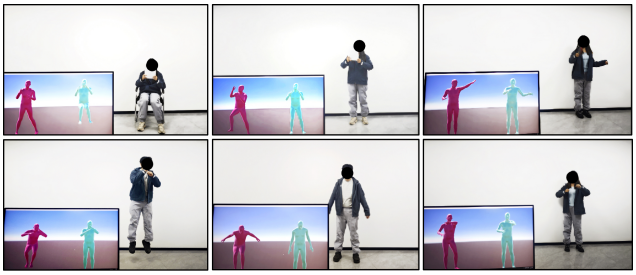

图 1. **服装惯性去噪器（GIP）能够对实时宽松穿着状态下的惯性测量单元数据进行去噪，从而实现鲁棒的全身动作捕捉。**

1. 引言

人体姿态估计是一项基础任务，在医疗保健、体育运动和沉浸式交互领域有着广泛应用 [54]。作为传统视觉方法的替代方案，采用惯性测量单元（IMU）的可穿戴动作捕捉，在便携性、隐私友好性以及抵御遮挡或极端光照方面具有显著优势 [10]。近期研究表明，仅使用少量（3–6 个）惯性测量单元即可实现令人瞩目的全身跟踪 [13, 41, 46–48, 58]。然而，这些系统通常要求将惯性测量单元紧密固定在身体上，以确保惯性读数稳定；这可能会让人感到拘束和不适，不适合日常或长期使用，例如人体数据采集、慢性病监测以及日常以身体为中心的交互。理想的解决方案是将惯性测量单元直接集成到日常宽松服装中，从而大幅减轻用户的负担。

然而，这种宽松服装设置带来了一个关键挑战：服装与身体之间的次级运动（例如滑移、摆动）会严重污染惯性测量单元数据。处理此类数据偏移的一种常见策略是域随机化 [37],，即使用经过随机扰动增强的数据训练模型，以提高鲁棒性。尽管如此，我们认为这种方法从根本上并不适用于该问题。服装动力学所引起的噪声具有内在的结构性，而非独立同分布，其特点是高维、非平稳且存在身体部位相关变化。为捕捉这种复杂分布而扩展增强空间，会导致假设复杂度激增、优化不稳定并增加数据需求，但最终仍无法覆盖现实世界中的结构化偏移。

因此，我们不直接学习从受污染的惯性信号到人体姿态的映射，而是提出一种因子化学习框架，将服装引起的污染与姿态估计解耦。具体而言，我们将

*共同贡献。
†通讯作者

整体映射 $f_{\mathbf{IMU}_{\text{loose}}\to\boldsymbol{\theta}}$ 分解为两个子函数：

$$f_{\mathbf{IMU}_{loose}\to\boldsymbol{\theta}} = f_{\mathbf{IMU}_{loose}\to\mathbf{IMU}_{tight}} \circ f_{\mathbf{IMU}_{tight}\to\boldsymbol{\theta}}. \quad (1)$$

这里，$f_{\mathbf{IMU}tight\to\boldsymbol{\theta}}$ 表示一种成熟且数据丰富的姿态估计器，它使用紧密佩戴的惯性测量单元信号进行训练；而$f_{\mathbf{IMU}loose\to\mathbf{IMU}_{tight}}$ 则是我们提出的服装惯性去噪器（*GID*），它将松散佩戴惯性测量单元测量值映射到对应的紧密佩戴数据（图1）。这种因子化处理将服装引起的干扰明确地与下游姿态预测任务分离开来，在成对松散佩戴与紧密佩戴数据有限时，显著降低了有效学习的复杂度。因此，它提高了数据效率和优化稳定性。然而在实际应用中，将 $f_{\mathbf{IMU}loose\to\mathbf{IMU}tight}$ 直接作为单个端到端网络进行训练并不奏效。松散佩戴惯性测量单元污染具有高度异质性，且取决于佩戴位置：不同身体部位上的传感器（例如手腕与躯干）表现出不同的服装—身体动力学。若用单一整体式去噪器对其进行建模，这些异质模式会引发相互冲突的优化过程，导致学习不稳定（图2）。

为应对这一挑战，GID 采用了基于位置的专家架构。共享的时空 Transformer 主干网络用于捕捉全局运动规律，而轻量级的逐惯性测量单元专家头则专门处理各自传感器位置相关的局部物理规律。小型融合模块进一步保证不同身体部位间的一致性，并消除专家间差异。这种层次化设计融入了恰当的归纳偏置，使模型能够利用极少的配对数据进行可靠学习，同时有效泛化至不同用户和服装配置。

为在真实的设计变化条件下评估我们的方法，我们推出了 GarMoCap，这是一个全面的服装动作捕捉数据集，涵盖 4 种不同的服装设计、43 名用户和 10 种表达性动作。实验结果表明，使用单名用户穿着单件服装的数据训练的 GID，能够有效泛化至未见过的用户和服装设计。作为即插即用模块，GID 无需对现有姿态估计模型进行任何架构修改，便能持续降低角度和位置误差，并提升其鲁棒性。我们的工作贡献总结如下：

• 我们将宽松穿戴惯性动作捕捉重新表述为一项结构化去噪任务，并提出一种因子化框架，将服装引起的伪影与姿态预测分离，从而能够直接复用现有的紧身穿戴惯性测量单元模型。

•我们提出服装惯性去噪器（GID），这是一种轻量级的位置特定去噪模块，用于建模由服装引起的、因身体部位而异的运动伪影。通过为每个惯性测量单元通道分配独立的去噪器，GID 缓解了不同传感器位置之间学习动态的冲突，并提高了训练稳定性。

• 我们提出了**GarMoCap**，这是一个全新的数据集，涵盖多样化的用户、动作和服装设计，专门用于研究真实世界服装—身体交互以及宽松穿戴传感器数据损坏问题。

• 我们证明，GID 具有很高的数据效率——仅使用单件服装和单个用户进行训练，却能够泛化到未见过的用户和服装类型；同时，它还可作为即插即用模块，显著提升现有惯性动作捕捉流程的准确性和鲁棒性。

2. 相关工作

2.1. 动作捕捉

2.1.1. 基于视觉的动作捕捉

光学方法分为基于标记点 [3, 12, 31] 和无标记点 [33] 两类。Vicon 等商业系统 [38] 能够提供高保真度，但需要昂贵的基础设施和严格受控的摄影棚环境，限制了其在日常生活中的可及性。在无标记点方法中，多相机系统 [5, 22, 40, 51] 使用经过校准并同步的摄像机阵列，通过三角测量和体积推理显著减少遮挡并提高精度，但仍需要仔细校准、密集的硬件配置和大量计算资源，因此在许多场景中部署时受到空间和成本限制 [25]。相比之下，单相机 RGB/RGB-D 流程 [9, 20, 35, 36, 42, 43, 45, 52, 53]因简单且易于扩展而日益受到青睐，但仍易受遮挡、光照和服装外观的影响，且通常假设摄像机固定不动——这降低了其在户外或高度动态场景中的稳健性。这些限制仍然阻碍着纯视觉动作捕捉在真实环境中的广泛部署。

2.1.2. 基于惯性的动作捕捉

惯性测量单元（惯性测量单元）通过直接测量肢段方向和线性加速度，为基于视觉的动作捕捉提供了一种便携且不受光照影响的替代方案。Noitom [29] 和 Xsens [27] 等商业系统通过将密集的惯性测量单元阵列紧密绑缚在多个身体部位上，实现了很高的精度，但这类设置具有侵入性，会限制自然运动，且部署成本高昂。这些局限促使研究广泛转向稀疏惯性测量单元配置（通常为 3–6 个 Sensors），旨在平衡实用性与保真度。

早期研究证明了稀疏惯性测量单元姿态重建的可行性：SIP [39]采用离线优化，而 DIP [13] 通过双向循环神经网络实现了近实时性能，但主要关注局部姿态。TransPose [46] 通过引入多阶段网络并对足部与地面的交互进行建模，迈出了关键一步，从而恢复全局平移并减少漂移。随后，

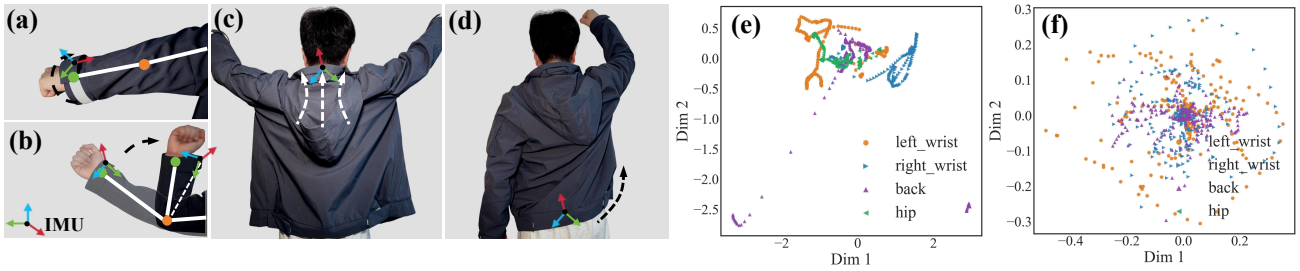

图 2。**动机**。（a–d）宽松穿戴服装会对惯性测量单元信号引入与位置相关的干扰，包括织物滑动、碰撞和服装变形。（e）旋转信号的主成分分析显示，紧身穿戴的惯性测量单元形成了清晰且具有身体部位特异性的流形，而（f）加速度信号的主成分分析表明，在宽松穿戴条件下，信号分布大幅扩散且结构坍缩。这些观察结果凸显了 GID 模块从宽松穿戴测量中恢复紧身穿戴信号特征的必要性。

后续研究沿两个方向取得进展：强化物理一致性，以及设计更优的网络架构。PIP [47] 引入了基于物理的正则化，PNP [48]则通过显式建模非惯性参考系中的虚拟力，进一步提升了鲁棒性。在架构方面，TIP [17] 将 Transformer 引入稀疏惯性测量单元姿态估计，而 ASIP [41] 则将 Transformer 主干网络与对抗训练相结合，以增强物理合理性。近年来，系统的易用性也成为一项重要关注点。TIC [58] 提出了一种动态身体佩戴校准策略，无需进行静态 T字姿势校准，从而显著提升了长期稳定性和用户便利性。与此同时，多模态融合也成为增强鲁棒性的另一研究方向，包括融合超宽带（UWB） [2, 6], 电磁（EM）场 [19] 以改进骨骼追踪和全局位移估计。其他研究，尤其是增强现实/虚拟现实场景中的研究，则着重于根据稀疏输入推断全身姿态，例如仅依据上半身，或仅依据头部和双手 [7, 16, 32, 44]。

尽管稀疏惯性测量单元系统的创新缓解了密集配置带来的一些不适，但其对紧密佩戴绑带的根本依赖依然存在。这种牢固固定的要求严重影响了用户的长期舒适度，给日常长时间使用场景中的持续穿戴性和实用性带来了挑战。

2.2. 基于服装的动作捕捉

将动作捕捉（MoCap）融入日常服装，为实现长期、便携且保护隐私的追踪提供了一条有前景的途径 [8]。现有的基于服装的方法大致可按服装的贴合方式分类。紧身穿戴系统能够保持传感器与身体之间的固定对齐，从而实现稳定、准确的测量。典型示例包括密集惯性测量单元阵列 [8], 、柔性关节追踪传感器 [4, 18], 、姿态识别织物 [21], ，以及结合稀疏深度摄像头与光学标记的混合系统 [3]。SuDA [18] 进一步提出了一种基于支撑的领域自适应策略，用于解决穿戴过程中服装变形导致的信号不稳定问题。

虽然贴身服装能够提供较高的数据准确性，但宽松穿戴选项更符合日常舒适性需求。该领域的研究主要集中于惯性测量单元和电容传感器。关于电容传感器，Zhou 等人 [56]将多通道传感器嵌入夹克中，并采用深度回归器估计上半身坐标。SeamPose [49]将传感器隐蔽地放置在服装接缝内部，由此形成了 SeamFit [50], ——一种可清洗、具有多种尺码的 T恤系统，能够推断未被直接覆盖的关节运动。在惯性测量单元领域，Zhou 等人 [24] 研究了使用多个惯性测量单元，将宽松穿戴数据转换为等效的紧身穿戴数据。LIP [57] 在一件宽松夹克中嵌入了 4 个稀疏惯性测量单元，并引入次级运动自编码器（SeMo-AE），在学习到的潜在空间中将服装–皮肤次级运动建模为加性高斯噪声；同时结合时间连贯性方案，从紧身穿戴数据中合成多样化的宽松穿戴惯性测量单元信号，以增强训练。其他研究则将服装搭载惯性测量单元的用途限制在活动识别 [15, 26, 34]。近期研究开始探索将服装作为多模态传感器融合平台，以获得协同效益。例如，FIP [55]在一件宽松版日常夹克中融合了 4 个惯性测量单元和 2 个肘部弯曲传感器，并引入位移潜在扩散模型，以合成并抵抗惯性测量单元的实时位移，从而实现稳健的实时基于服装的动作捕捉。最近，Ilic 等人 [14] 引入了 GaIP，这是一种服装感知扩散框架，通过合成宽松穿戴信号，并以服装参数作为条件来处理传感器–服装动力学，从而利用稀疏、松散附着的惯性测量单元估计全身姿态。

然而，LIP 和 FIP 等最先进的稀疏惯性测量单元系统在算法上高度依赖数据驱动的噪声建模，这既耗时，又需要大量训练数据来覆盖所有可能的噪声条件，而这几乎是不可能的。为了解决这些局限，我们提

出了首个基于稀疏惯性测量单元（n=6）的全身宽松穿戴动作捕捉系统。我们的方法无需依赖大规模数据合成，只需少量真实数据即可训练出轻量级、即插即用的去噪器，从而提高数据效率和训练稳定性。

3. 方法

3.1. 问题定义

令 $\boldsymbol{\theta} \in \mathbb{R}^{J\times 3}$ 表示 J 身体关节的三维旋转角。服装内惯性测量单元传感器网络由 M个惯性测量单元组成，分别位于$\mathbf{s}_m$，$m = 1, 2, 3, \ldots M$。每个惯性测量单元 m 都会采集其对应位置的加速度$\mathbf{a}_m \in \mathbb{R}^3$ 和旋转 $\mathcal{R}_m \in \mathbb{R}^3$。本文旨在利用稀疏惯性传感器开发一种基于服装的动作捕捉系统，以重建宽松穿戴状态下的惯性测量单元读数 $\mathbf{IMU}_{loose}$ 与底层关节旋转 $\boldsymbol{\theta}$之间的映射函数 f：

$$\boldsymbol{\theta} = f(\mathbf{IMU}_{loose}) \tag{2}$$

然而，直接估计 f 具有挑战性。宽松穿戴状态下的惯性测量单元会因织物与身体之间的位移而产生次级运动，这些运动表现为叠加在理想紧身穿戴信号上的复杂干扰。这些干扰显著扩大了映射函数 f （方程2），从而使直接优化更加复杂。受域随机化 [37],的启发，一种简单的方法是在紧身穿戴信号中添加噪声，合成多种宽松穿戴状态下的惯性测量单元数据。随后，使用这些增强数据训练模型 f，期望其能够泛化至真实世界的**惯性测量单元**$_{\text{loose}}$条件：

$$f_{\mathbf{IMU}_{loose}\rightarrow\boldsymbol{\theta}} = f_{\mathbf{IMU}_{tight}\rightarrow\mathbf{IMU}_{loose}} \circ f_{\mathbf{IMU}_{loose}\rightarrow\boldsymbol{\theta}} \tag{3}$$

然而，对于服装运动问题而言，这种策略在根本上存在错误设定。宽松服装造成的扰动并非独立同分布噪声，而是由惯性、动量和材料属性等物理特性引发的结构化物理扰动。噪声注入方法忽略了这种内在结构，试图通过蛮力搜索庞大的参数空间来实现鲁棒性。这导致了一个固有的两难局面：狭窄的噪声分布无法覆盖真实世界的动力学，而宽泛的分布则会使训练不稳定，并学到一种计算效率低下且不符合物理规律的映射。

3.2. 服装惯性去噪器

因子化参数空间。为克服这些局限，我们采用了不同的视角：我们不再随机化噪声，而是显式地学习一个去噪函数$f_{\mathbf{IMU}_{loose}\rightarrow\text{惯性测量单元}_{tight}}$，将宽松穿戴时的惯性测量单元读数映射为对应的紧身穿戴数据。我们提出这一方法的关键洞见是，f 的参数空间是可因子化的，因此可分解为两个具有可解释性的子函数，分别对应信号去噪和姿态重建：

$$f_{\mathbf{IMU}_{loose}\rightarrow\boldsymbol{\theta}} = f_{\mathbf{IMU}_{loose}\rightarrow\mathbf{IMU}_{tight}} \circ f_{\mathbf{IMU}_{tight}\rightarrow\boldsymbol{\theta}} \tag{4}$$

这种分解策略具有优势。它能够约束假设空间，并让模型从有限的配对数据中更高效地学习。由于映射 $f_{\mathbf{IMU}_{tight}\rightarrow\boldsymbol{\theta}}$ 包含复杂的身体动力学和非线性运动学依赖关系，因此占据了参数空间中大得多的部分。相比之下，去噪映射（$f_{\mathbf{IMU}_{loose}\rightarrow\text{惯性测量单元}_{tight}}$）的复杂度相对较低，因为 **惯性测量单元**$_{loose}$ 和**惯性测量单元**$_{tight}$ 具有相似的时空结构。任务并非从头开始学习姿态，而是学习这种结构化物理扰动的逆过程。这种因式分解使我们能够相应地解耦问题：利用稀缺的配对数据学习扰动模型，同时借助成熟且数据丰富的估计器完成姿态重建任务。

位置特定去噪。然而，直接学习去除$f_{\mathbf{IMU}_{loose}\rightarrow\text{惯性测量单元}_{tight}}$ 在式4 中的去噪函数极具挑战性。主要难点在于宽松佩戴造成的失真具有位置依赖（图2）。安装在不同身体部位（例如手腕、背部和腰部）的惯性测量单元会受到不同关节运动学以及传感器—身体—服装耦合动力学的影响，而这些因素又受到固定几何结构、局部松弛度和服装—皮肤接触模式等因素的共同作用。因此，所得信号统计特性会随位置系统性地变化。使用单一共享模型来表示这类异质分布，会引入相互冲突的优化梯度，从而难以学习有效且具有良好泛化能力的映射。

为了解决这一问题，GID采用位置特定去噪范式，为每个传感器分配一个位置特定去噪器（专家模块），以刻画其独特的物理行为。这些专家运行于不同的时空Transformer骨干网络之上，用于捕捉全局运动规律；与此同时，轻量级融合模块则确保跨部位一致性。该设计融入了适当的归纳偏置，增强了对异质运动伪影的鲁棒性，并保持了参数效率，使式4 中的复合映射既具有可处理性，又具备良好的泛化能力。

自适应跨穿戴融合。虽然每个位置特定的去噪器都会对 $\hat{}$**惯性测量单元**$_{tight}$ 生成精细估计，但由于独立处理，其输出仍可能存在残余不一致。为了减少

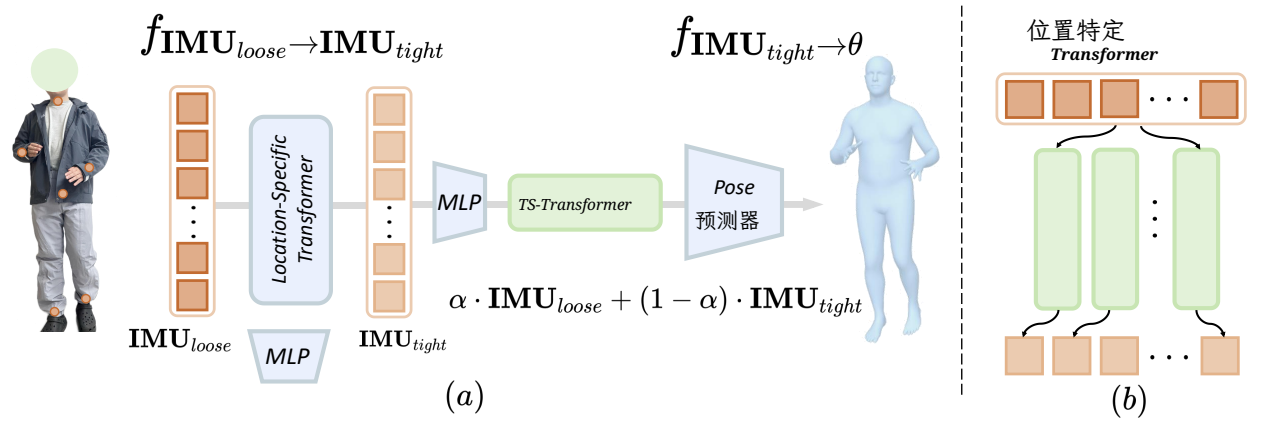

图 3. **服装惯性去噪器流程。** (a) 我们将 f 分解为两个子函数，分别用于信号去噪和姿态重建。(b) 每个传感器读数 **惯性测量单元**$_{loose}$ 都会分配给一个基于时空（TS）Transformer 的位置特定去噪器。

这些差异并利用紧身与宽松穿戴域的互补线索，我们提出一种早期融合策略。具体而言，我们构造可学习凸组合 $\alpha \cdot$**^惯性测量单元**$_{tight}$ $+(1-\alpha)$ **·惯性测量单元**$_{loose}$，其中 $\alpha \in [0,1]$ 是通过 Sigmoid 参数化实现的可训练权重。这种形式使模型能够自适应地平衡紧身穿戴信号的稳定性与宽松穿戴动态的真实性，从而捕捉高频和低频运动模式。随后，融合表示会经过最终的时空 Transformer 进行一致性细化，从而在不增加推理开销的情况下实现更平滑的姿态重建。

姿态预测器。 随后，将滤波后的紧身等效信号输入姿态预测器$f_{\mathbf{IMU}_{\text{tight}}\rightarrow\boldsymbol{\theta}}$，由其将惯性序列映射为相应的人体姿态参数 $\boldsymbol{\theta}$。鉴于有大量紧身穿着条件下的惯性测量单元与动作配对数据，该阶段可以在完全监督的设置下训练，从而精确恢复人体运动学。得益于模块化且轻量化的设计，我们的流程不受 $f_{\mathbf{IMU}_{\text{tight}}\rightarrow\boldsymbol{\theta}}$ 具体实现方式的限制，并且无需修改结构即可无缝集成最先进的惯性测量单元到姿态架构（例如 PIP、PNP）。这种解耦设计将位置感知去噪器与下游姿态预测器分离，有助于灵活适配、稳定收敛，并能利用有限的配对数据进行高效学习。

4. GarMoCap 数据集

为全面评估我们的方法，我们构建了 Gar-MoCap，这是一个复合数据集集合，将现有的公开上半身数据集（15 名用户）与我们新采集的



图 4. **我们完整的基于服装的全身动作捕捉系统原型。** 宽松穿着的夹克和裤子中共嵌入了六个惯性测量单元传感器和两块中央电路板，用于采集数据。

数据（28 名用户）以及全身动作捕捉服装相结合（见图4）。该集合涵盖多种服装类型（$n=4$）、身体区域（$n=43$），同时包含上半身和全身录制数据。

4.1. 硬件

图4 展示了我们基于日常运动服装版型制作的全身动作捕捉原型服装。该服装的尺码为 XL，目标合体尺寸为 50 厘米（肩宽）、 108 厘米（胸围）、 18 厘米（腰围）和 25 厘米（袖围），适用于身高 160–185 厘米、体重 50–85 千克的穿戴者。该系统由一件夹克和一条长裤组成，每件服装均集成了四个 Xsens MTI-3 惯性测量单元 [28]。上装中的传感器分别位于左前臂、右前臂、背部（与数据传输主板共置）和腰线。下装中的传感器分别位于左腰和右腰（集成下部主板）。如技术型外套所示，每个传感器均连接至

其对应的主板。柔性电缆通过热压方式嵌入织物中，从而确保线路平整、耐用且不易察觉。两块主板均以 40 Hz 的频率采集所连接的惯性测量单元数据，并通过蓝牙将测量结果无线传输至主机计算机。系统由一块 500 mAh 的锂电池供电。更多信息请参阅我们的附录。

4.2. 数据集采集

我们使用服装动作捕捉系统采集了两个新数据集，$\mathcal{D}_{GID}^{upper}$ 和 $\mathcal{D}_{GID}^{full}$。随后，我们将它们与两个现有的公开服装动作捕捉数据集 $\mathcal{D}_{FIP}^{upper}$ 和 $\mathcal{D}_{LIP}^{upper}$ 整合，构成了我们的复合数据集 GarMoCap。

- $\mathcal{D}_{GID}^{upper}$：我们采集了**GID-Upper**数据集，其中包含二十名参与者（10 名男性、10 名女性）穿着配有四个惯性测量单元的宽松夹克。参与者依次完成五种上半身动作（例如跑步、拳击、热身），每种动作持续一分钟，总计约 100 分钟的数据。真实值通过 Perception Neuron 3 惯性系统记录 [30]。该数据集专门用于评估跨用户和跨动作的泛化能力。

- $\mathcal{D}_{GID}^{full}$：我们进一步采集了**GID-Full** 数据集，其中包含八名参与者完成十种全身动作的记录（总计约 80 分钟）。这些数据使用我们配备了 6 个惯性测量单元的全身服装套装（夹克和裤子）采集。值得注意的是，真实值通过由 8 台摄像机同步组成的 RGB 阵列获取，因为宽松服装遮挡光学标记时，传统光学标记无法使用。摄像机覆盖 4m×2.5m 的捕捉空间，可实现不受限制的运动。我们使用 EasyMocap [1] 框架对 SMPL [23] 参数（形状 β、关节旋转、平移）进行三维重建，重建误差低于 20 mm。该多摄像机设置有效缓解了遮挡问题，即使在动态的宽松穿着场景中，也能确保可靠地生成真实值。

- $\mathcal{D}_{FIP}^{upper}$：我们利用 **FIP-Upper** 数据集 [55], 评估惯性测量单元服装，该数据集收集了真实世界记录。数据集包含 10 名体型各异的参与者，他们执行了 10 种预定义上半身动作（例如 行走、跑步和拳击）。数据集共包含 371,122 帧，以 60 fps 采集（约 103 分钟）；同时还提供了由 Perception Neuron 3 系统使用 11 个紧身佩戴的惯性测量单元采集的真实姿态。

- $\mathcal{D}_{LIP}^{upper}$：我们还纳入了专为评估宽松穿着的惯性测量单元服装而设计的 **LIP-upper**数据集 [57]。该数据集包括五名参与者，他们在两种条件（拉链闭合和拉链打开）下穿着一件配有四个惯性测量单元的夹克，以模拟不同程度的宽松状态。该数据集包含 212,496 帧，采样率为 30 fps（约 118 分钟），涵盖预定义动作和自由形式动作。真实值同样由 Perception Neuron 3 系统采集，该系统配备十一紧身惯性测量单元。

这些数据集共同构成了一个统一的评估语料库，涵盖多样的服装设计、穿着条件（拉链闭合、拉链打开）以及身体覆盖范围（上半身和全身）。该数据集为研究可穿戴动作捕捉中的宽松–紧身域转换、传感器干扰去噪和跨服装泛化提供了全面基础。有关 GarMoCap 的详细信息，请参阅我们的补充材料。

5. 实验

5.1. 实验设置

指标 我们使用平均绝对误差（MAE）损失来衡量所提方法的去噪性能。遵循 TransPose [46], ，我们使用以下四项指标衡量姿态估计的准确性：

- 角度误差，用于衡量所有身体关节的平均全局旋转误差，单位为度。
- 位置误差，用于衡量所有估计关节的平均欧氏距离误差，单位为厘米。
- 网格误差，用于衡量姿态化 SMPL 网格的顶点误差。
- 抖动误差，用于衡量全局空间中所有身体关节的平均加加速度（加速度的时间导数），反映运动的平滑度 [11]。

训练详情 所有实验均在一台配备 Intel U9-285K CPU 和 NVIDIA RTX 5080 GPU 的工作站上进行。实现基于 PyTorch 2.8.0 和 CUDA 12.9。更多训练配置和网络规格详见补充材料。

5.2. 相同服装设计上的性能

全身动作捕捉。我们首先在全身动作捕捉任务上评估 GID，使用 $\mathcal{D}_{GID}^{full}$ 中的一名受试者进行训练，其余受试者用于测试。与上半身追踪相比，全身动作捕捉的挑战显著更大，因为其参数空间维度更高，而且强烈的肢体间依赖关系会使服装引起的噪声在各关节间传播。尽管存在这些困难，GID 在所有协议上均取得了显著提升，尤其是在 PIP 和 ASIP 上，这表明我们的去噪器能够有效去除服装运动，并在复杂的全身动力学条件下恢复具有物理意义的惯性信号（表 1）。

一个有趣的观察是，将经过 GID 去噪的惯性测量单元信号应用于 PNP，并不会进一步提升其性能。我们认为，这源于表征不匹配：我们的去噪器在根节点相对惯性测量单元空间中运行，而 PNP 依赖于世界坐标系下具有物理结构的惯性线索。将根节点相对的去噪信号注入

表 1. **采用和不采用 GID 的最先进方法在全身性能上的比较。**数值越低，性能越好。

方法	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)
PIP [47]	**16.58**/24.27	**8.95**/10.85	**10.79**/13.01	**0.37**/0.40
PNP [48]	**23.21**/23.21	**10.95**/11.94	**13.28**/14.05	1.03/**0.86**
ASIP [41]	**15.68**/26.09	**8.49**/11.49	**10.45**/14.02	**0.03**/0.07

表 2. **在上半身数据集评估中，采用和未采用 GID 的最先进动作捕捉方法对比。**

方法	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)
PIP	**38.24**/38.42	16.08/**14.63**	18.84/**17.97**	**0.34**/0.57
PNP	**20.39**/22.38	**8.80**/9.22	**11.36**/13.17	**0.35**/0.63
ASIP	**20.08**/24.18	**9.07**/10.13	**11.07**/13.69	**0.13**/0.30

世界坐标系预测器可能会扭曲 PNP 所依赖的时间动态。然而，直接在世界坐标系中训练 GID 会将相机对齐和主体特定的全局运动重新引入学习目标，使任务显著变得更加困难，并在实际中导致不稳定优化和较差的泛化。

上半身动作捕捉。 为评估 GID 在服装内部的泛化能力，我们在从全身录制数据中提取的上半身子集上进行评估。由于上衣可以单独穿着，这种设置直接检验了经过 GID 训练的信号在全身场景之外是否仍然有效。我们使用一名随机选取的用户进行训练 $\mathcal{D}_{GID}^{full}$，并在其余八名用户上测试。如表2所示，尽管仅使用了来自单个受试者的成对的宽松与紧身惯性测量单元数据，我们的去噪器仍能可靠地抑制宽松惯性测量单元伪影，并显著提高姿态估计精度。这些提升与全身实验结果一致，进一步表明 GID 能够有效泛化到不同用户和可穿戴设备配置。

5.3. 泛化至不同服装设计

为进一步评估 GID 的跨服装泛化能力，我们使用一位用户在 $\mathcal{D}_{GID}^{full}$ 上的数据进行训练，并在 $\mathcal{D}_{LIP}^{upper}$、$\mathcal{D}_{FIP}^{upper}$ 和 $\mathcal{D}_{GID}^{upper}$ 上评估模型。这些数据集总共涵盖三种不同的服装设计以及 35 位用户。如表 3所示，尽管仅使用一位用户的数据进行训练，GID 仍表现出稳健的泛化能力，在不同材质、图案和宽松程度的服装上保持较高的去噪精度，从而实现可靠的动作捕捉性能。这一发现与我们的核心假设高度一致：GID 表征能够捕捉服装不变运动特征。

5.4. 与最先进的基于服装的动作捕捉方法比较

我们使用其公开代码库中的官方检查点，将 GID 与两种最先进的基于服装的动作捕捉方法 FIP 和 LIP 进行比较。两种基线方法在 $\mathcal{D}_{LIP}^{upper}$、$\mathcal{D}_{FIP}^{upper}$ 和 $\mathcal{D}_{GID}^{upper}$ 上均表现出色（表 4）。值得注意的是，尽管 GID 仅使用一位参与者穿着一种服装时采集的数据进行训练（38,286 帧；11 分钟），其精度仍可媲美甚至超过这两种方法；相比之下，FIP/LIP 依赖超过 700 万个模拟帧（1,967 分钟），凸显了 GID 在数据效率和训练成本方面的明显优势。此外，GID 支持全身动作捕捉，而 FIP/LIP 仅针对上半身；因此，我们在上半身数据集上报告定量比较结果。

5.5. 消融实验

如表 5所示，完整的 GID 模型在全局角度、位置、网格和抖动误差方面均取得了最低值（22.24°、8.62 cm、10.92 cm、0。11 m/s^3），同时具备最佳的惯性测量单元平均绝对误差（0.130）。将双阶段设计简化为单一的端到端映射会明显恶化所有姿态指标，表明直接从松散惯性测量单元回归姿态会使优化更加困难且准确性更低。使用共享去噪器而非位置特定专家会轻微降低姿态精度并残留更多传感器噪声，说明逐惯性测量单元专家能够更好地捕捉异质动态。去除融合器的变体会产生更大的角度和网格误差以及噪声最多的惯性测量单元信号，凸显了跨惯性测量单元融合对于实现一致去噪的重要性。总体而言，这些结果表明，双阶段因式分解、位置特定去噪和融合模块对于 GID 实现最佳姿态与去噪性能都是必不可少的。

6. 实时演示

我们开发了一套基于 Python 和 Unity 的实时姿态估计与可视化系统。用户像平常一样穿上服装后，系统只需进行简短的 T 姿势校准（约 5 秒）即可完成初始化。系统启动后，会实时传输 IMU测量数据，按照 TransPose [45], 进行校准和归一化处理，并将处理后的信号输入我们的姿态估计模型。随后，重建动作会立即渲染，以供实时可视化。在实时演示中，系统展现了宽松衣物的舒适性与自然穿着体验和动作捕捉精确度之间的有效平衡。即使在快速奔跑和跳跃等动态活动中，系统仍能保持稳定，展现出强大的鲁棒性，适用于广泛的真实场景。补充材料中提供了演示视频。

表3。**跨 $\mathcal{D}_{LIP}^{upper}$ 、 $\mathcal{D}_{FIP}^{upper}$ 和 $\mathcal{D}_{GID}^{upper}$ 数据集上采用和未采用 GID 的最先进动作捕捉方法的比较**。尽管 GID 仅使用单个用户穿着单件服装的数据进行训练，但它对具有不同服装设计、用户和动作的未见领域仍能有效泛化。

方法	$\mathcal{D}_{LIP}^{upper}$				$\mathcal{D}_{FIP}^{upper}$				$\mathcal{D}_{GID}^{upper}$			
	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)
PIP	77.25/**75.77**	22.98/**21.89**	28.89/**27.51**	0.51/**0.88**	27.26/**22.78**	12.47/**9.40**	14.35/**11.83**	**0.28**/0.41	34.35/**31.00**	15.73/**13.41**	17.50/**15.08**	**0.31**/0.55
PNP	**31.15**/33.33	**11.84**/14.61	**15.34**/18.98	**0.50**/1.05	**20.55**/24.02	**7.17**/9.76	**9.74**/14.12	**0.28**/0.45	**25.15**/25.98	**9.44**/10.61	**12.45**/14.43	**0.35**/0.68
ASIP	**30.81**/32.22	**11.57**/14.40	**14.88**/18.34	**0.22**/0.42	**19.37**/25.32	**6.46**/10.55	**8.28**/13.84	**0.08**/0.17	**23.09**/25.83	**7.97**/9.87	**10.08**/12.79	**0.10**/0.24

表4. **使用四项关键指标对基于服装的动作捕捉方法进行性能比较**。尽管 GID 仅使用单个用户穿着单件服装的数据进行训练，其精度仍可媲美最先进的基于服装的动作捕捉系统。

方法	$\mathcal{D}_{LIP}^{upper}$				$\mathcal{D}_{FIP}^{upper}$				$\mathcal{D}_{GID}^{upper}$			
	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)
LIP [57]	29.84	11.39	14.33	0.86	22.40	8.59	10.93	0.30	23.72	8.65	10.25	0.51
FIP [55]	**28.24**	**10.30**	**13.09**	0.26	22.77	8.14	10.57	0.10	**21.56**	**7.33**	**9.02**	0.16
GID（ASIP）	30.81	11.57	14.88	**0.22**	**19.37**	**6.46**	**8.28**	**0.08**	23.09	7.97	10.08	**0.10**

表 5. **GID 的消融实验**。 移除 FPS（分数化参数空间）后，直接使用单用户松散动作配对数据训练 ASIP模型。移除 LSD（位置特定去噪器）后，以单个时空 Transformer 替代多去噪器设计。移除 ACF（自适应跨穿戴融合）后，各去噪器独立运行，不进行跨穿戴融合。

	角度 (°)	位置 (cm)	网格 (cm)	抖动 (m/s^3)	MAE
无 FPS	25.69	9.56	12.21	0.13	–
无 LSD	23.03	8.91	11.38	0.13	0.133
无 ACF	24.53	9.80	12.50	0.13	0.145
Full	**22.24**	**8.62**	**10.92**	**0.11**	**0.130**

7. 局限性与未来工作

局限性。尽管去噪能有效减轻位移引起的扰动，但去噪后的宽松穿戴信号与理想的贴身穿戴真值之间仍存在明显的性能差距（请参阅补充材料）。这表明当前模型尚未充分捕捉衣物—身体相互作用的复杂非线性动力学。此外，我们目前的服装设计仍有一定相似之处，例如传感器布局和织物布局相近，这可能会限制模型在不同服装配置之间的泛化能力。在扩展服装设计多样性的同时，引入可微物理或服装仿真先验，有望进一步缩小这一差距，并实现对真实世界宽松穿戴感知条件更全面的建模。

未来工作。 我们目前的实验依赖于从单个用户穿着同一件服装时采集的成对松散佩戴与紧密佩戴数据。虽然这一设置展现了出色的数据效率，并证实了 GID 在受控条件下的有效性，但它限制了运动模式、体型和织物类型的多样性。未来，我们计划扩大数据收集规模并开发去噪基础模型。这包括构建一个大型异构数据集，将真实世界记录与基于物理知识的合成数据相结合，以便研究扩展行为并提升跨域泛化能力。在此扩展语料库的基础上，我们计划通过多域学习和轻量级物理先验训练一个统一模型，以弥合松散佩戴与紧密佩戴传感之间的差距。最终，我们的目标是构建一个具备良好泛化能力且以物理原理为基础的服装惯性测量单元去噪基础模型，彻底弥合松散佩戴与紧密佩戴传感之间的差距。

8. 结论

本文介绍了服装惯性去噪器（GID），这是一种轻量级、即插即用的基于 Transformer 的去噪器，能够有效去除宽松穿着服装所采集的惯性测量单元信号中的噪声。我们的核心创新在于，将从宽松穿着的惯性测量单元到关节姿态的映射分解为可学习的去噪阶段和姿态估计阶段。这种分解利用了宽松与紧身穿着时惯性测量单元信号之间的结构相似性，同时由位置特定去噪器捕捉异质的运动引起的噪声。这些设计选择相辅相成，使 GID 在基于服装的动作捕捉中兼具鲁棒性和泛化能力。为全面评估我们的方法并推动未来研究，我们进一步推出了 GarMoCap，这是一个综合数据集，包含不同用户、动作和服装设计下成对的宽松穿戴姿态记录。大量实验表明，即使仅在单件服装上进行训练，GID 也能在不同用户和服装设计之间保持良好的泛化能力，显著减少传感器干扰，并提高下游姿态估计的精度。未来，我们计划训练一种基于物理原理、适用于宽松穿着服装的去噪基础模型。

参考文献

[1] Easymocap——让人体动作捕捉更简单。GitHub，2021。6[2] Rayan Armani、Changlin Qian、Jiaxi Jiang 和 Christian Holz。超惯性姿态估计器：利用稀疏惯性传感器和超宽带测距实现可扩展的动作捕捉与追踪。载于 ACM SIGGRAPH 2024 会议论文集，2024，页码 –。[3] Anargyros Chatzitofis、Dimitrios Zarpalas、Petros Daras 和 Stefanos Kollias。Democap：低成本基于标记的动作捕捉。《国际计算机视觉期刊》，129(12)：3338–3366，2021。2，3[4] Xiaowei Chen、Xiao Jiang、Jiawei Fang、Shihui Guo、Juncong Lin、Minghong Liao、Guoliang Luo 和 Hongbo Fu。Dispad：用于稳健关节运动追踪的织物传感器灵活身体穿戴式位移。《ACM交互、移动、可穿戴与普适技术汇刊》，7(1)：1–27，2023。3[5] Rohan Choudhury、Kris M Kitani 和 L´aszl´o A Jeni。Tempo：高效的多视角姿态估计、追踪与预测。载于IEEE/CVF国际计算机视觉会议论文集，页码14750–14760，2023。2[6] Nathan DeVrio、Vimal Mollyn 和 Chris Harrison。Smartposer：利用超宽带和惯性测量单元数据，通过智能手机和智能手表进行手臂姿态估计。载于第36届ACM用户界面软件与技术年度研讨会论文集，页码1–11，2023。3[7] Yuming Du、Robin Kips、Albert Pumarola、Sebastian Starke、Ali Thabet 和 Artsiom Sanakoyeu。虚拟化身长出双腿：利用扩散模型，根据稀疏追踪输入生成人体平滑动作。载于 IEEE/CVF计算机视觉与模式识别会议论文集，页码481–490，2023。3[8] Maryam Adnan Fadhil 和 Waleed F Shareef。宽松衣物下使用可穿戴传感器进行人体活动追踪：综述。《TELKOMNIKA（电信、计算、电子与控制）》，21(6)：1382–1390，2023。3[9] Runyang Feng、Yixing Gao、Tze Ho Elden Tse、Xueqing Ma 和 Hyung Jin Chang。Diffpose：用于基于视频的人体姿态估计的时空扩散模型。载于 IEEE/CVF国际计算机视觉会议论文集，页码14861–14872，2023。2[10] Alessandro Filippeschi、Norbert Schmitz、Markus Miezal、Gabriele Bleser、Emanuele Ruffaldi 和 Didier Stricker。基于惯性传感器的动作追踪方法综述：聚焦于上肢人体动作。《Sensors》，17(1)：257，2017。6[11]Tamar Flash 和 Neville Hogan。手臂运动的协调：经实验验证的数学模型。《神经科学杂志》，5(7)：1688–1703，1985。6[12] Nima Ghorbani 和 Michael J Black。Soma：自动解决光学标记点动作捕捉问题。载于 IEEE/CVF国际计算机视觉会议论文集，页码11117–11126，2021。2[13] Yinghao Huang、Manuel Kaufmann、Emre Aksan、Michael J Black、Otmar Hilliges 和 Gerard Pons-Moll。深度惯性惯性姿态估计器：从稀疏惯性测量中实时学习重建人体姿态。ACM图形学汇刊 (TOG)，37(6):1–15，2018。1, 2[14] Andela Ilic、Jiaxi Jiang、Paul Streli、Xintong Liu 和 Chris- tian Holz。利用服装感知扩散模型，从宽松服装和稀疏惯性传感器中捕捉人体动作。arXiv预印本 arXiv:2506.15290，2025。3[15] Udeni Jayasinghe、William S Harwin 和 Faustina Hwang。比较衣物安装式传感器与可穿戴传感器在运动分析和活动分类中的应用。Sensors，20(1):82，2019。3[16] Jiaxi Jiang、Paul Streli、Huajian Qiu、Andreas Fender、Larissa Laich、Patrick Snape 和 Christian Holz。AvatarPoser：利用稀疏运动感测进行关节式全身姿态跟踪。载于《欧洲计算机视觉会议》，第443–460页。施普林格，2022。3[17] Yifeng Jiang、Yuting Ye、Deepak Gopinath、Jungdam Won、Alexander W Winkler 和 C Karen Liu。Transformer Inertial Poser：利用稀疏惯性测量单元实时重建人体动作，同时生成地形。载于亚洲计算机图形学会议 2022 论文集，第1–9页，2022。3[18] Fang Jiawei、Chengxu Zuo、Xiaoxia Gao、Xiaowei Chen、Shihui Guo、Yipeng Qin 等。SUDA：利用基于支持的域适应，通过柔性传感器实现仿真到现实的铰链关节跟踪。载于第四十一届国际机器学习会议，2024。3[19] Manuel Kaufmann、Yi Zhao、Chengcheng Tang、Lingling Tao、Christopher Twigg、Jie Song、Robert Wang 和 Otmar Hilliges。EM-Pose：利用稀疏电磁跟踪器进行三维人体姿态估计。载于IEEE/CVF国际计算机视觉会议论文集，第11510–11520页，2021。3[20] Muhammed Kocabas、Chun-Hao P Huang、Otmar Hilliges 和 Michael J Black。PARE：用于三维人体估计的部位注意力回归器。载于IEEE/CVF国际计算机视觉会议论文集，第11127–11137页，2021。2[21] Zhen Liang、Dongquan Zhang、Guanghua Xu、Fangting Xie、Hui Cai、Hao Guo 和 Jingyuan Cheng。智能服装：一种长期可行的全身纺织压力感测系统。《IEEE传感器期刊》，23(19):23294–23302，2023。3[22] Ziwei Liao、Jialiang Zhu、Chunyu Wang、Han Hu 和 Steven L Waslander。用于三维人体姿态估计的多视图几何 Transformer。载于IEEE/CVF计算机视觉与模式识别会议论文集，第708–717页，2024。2[23] Matthew Loper、Naureen Mahmood、Javier Romero、Gerard Pons-Moll 和 Michael J Black。SMPL：蒙皮式多人线性模型。载于《图形学经典论文：拓展边界》第2卷，第851–866页。2023。6[24] Michael Lorenz、Gabriele Bleser、Takayuki Akiyama、Take- hiro Niikura、Didier Stricker 和 Bertram Taetz。利用集成于宽松服装中的惯性传感器，实现面向伪影的捕捉式人体动作捕捉。载于2022年国际机器人与自动化会议 (ICRA)，第1682–1688页。电气电子工程师学会，2022。3[25]Dushyant Mehta、Oleksandr Sotnychenko、Franziska Mueller、Weipeng Xu、Mohamed Elgharib、Pascal Fua,

Hans-Peter Seidel、Helge Rhodin、Gerard Pons-Moll 和 Christian Theobalt。Xnect：利用单个RGB相机实时进行多人三维动作捕捉。《ACM图形学汇刊》(TOG)，39(4):82–1，2020。2[26] Brendan Michael 和 Matthew Howard。利用宽松服装上的可穿戴传感器进行活动识别。PLOS ONE，12(10): e0184642，2017。3[27] Movella。Movella：动作捕捉与传感器技术创新。https://www.movella.com/，2024。访问日期：2024-08-10。2[28] Movella。Xsens传感器模块：功能概览。https://www.movella.com/sensor-modules/xsens-functionality，2025。访问日期：2025-11-12。5[29] Noitom。Noitom：动作捕捉系统与解决方案。https://noitom.com/，2024。访问日期：2024-08-10。2[30] Noitom。Perception Neuron 3：无线动作捕捉系统。https://www.noitom.com/perception- neuron-3，2025。访问日期：2025-11-12。6[31] OptiTrack。OptiTrack - 动作捕捉系统。https: //www.optitrack.com/，2024。访问日期：2024-08- 10。2[32] Jose Luis Ponton、Haoran Yun、Andreas Aristidou、Carlos Andujar 和 Nuria Pelechano。SparsePoser：利用稀疏数据实时重建全身动作。《ACM图形学汇刊》，43(1):1–14，2023。3[33] M Rahul。动作捕捉技术综述。《全球计算机科学与技术期刊》，18 26，2018 1 22 2 ( ): – 。[34] Tianchen Shen、Irene Di Giulio 和 Matthew Howard。利用宽松服装进行人体活动识别的概率模型。Sensors，23(10):4669，2023。3[35] Soyong Shin、Juyong Kim、Eni Halilaj 和 Michael J Black。WHAM：以准确的三维动作重建世界锚定人体。载于IEEE/CVF计算机视觉与模式识别会议论文集，第2070–2080页，2024。2[36] Ke Sun、Bin Xiao、Dong Liu 和 Jingdong Wang。用于人体姿态估计的深度高分辨率表征学习。载于IEEE/CVF计算机视觉与模式识别会议论文集，第5693–5703页，2019。2 – 。[37] Josh Tobin、Rachel Fong、Alex Ray、Jonas Schneider、Woj- ciech Zaremba 和 Pieter Abbeel。通过域随机化，将深度神经网络从仿真迁移到现实世界。载于2017年IEEE/RSJ智能机器人与系统国际会议 (IROS)，第23–30页。电气电子工程师学会，2017。1, 4[38] Vicon。Vicon：动作捕捉系统。https://www. vicon.com/，2024。访问日期：2024-08-10。2[39] Timo Von Marcard、Bodo Rosenhahn、Michael J Black 和 Gerard Pons-Moll。Sparse Inertial Poser：利用稀疏惯性测量单元自动进行三维人体姿态估计。载于《计算机图形论坛》，第349–360页。Wiley在线图书馆，2017。2[40] Size Wu、Sheng Jin、Wentao Liu、Lei Bai、Chen Qian、Dong Liu 和 Wanli Ouyang。基于图的三维多人使用多视图图像进行姿态估计。载于《IEEE/CVF国际计算机视觉会议论文集》，第 11148–11157 页，2021 年。2[41] Yinghao Wu、Lu Yin、Shihui Guo、Yipeng Qin 等。通过序列结构学习与调制实现准确、稳定的惯性姿态估计。载于《神经信息处理系统进展》，37:42468–42493，2024 年。1, 3, 7[42] Yan Xia、Xiaowei Zhou、Etienne Vouga、Qixing Huang 和 Georgios Pavlakos。利用生物力学精确的骨骼重建人体。载于《计算机视觉与模式识别会议论文集》，第 5355–5365 页，2025 年。2[43] Yufei Xu、Jing Zhang、Qiming Zhang 和 Dacheng Tao。Vit-pose：用于人体姿态估计的简易视觉变换器基线。载于《神经信息处理系统进展》，35:38571–38584，2022 年。2[44] Dongseok Yang、Jiho Kang、Lingni Ma、Joseph Greer、Yuting Ye 和 Sung-Hee Lee。DivaTrack：利用加速度增强型三点追踪器捕捉多样化的身体与动作。载于《Computer Graphics Forum》，第 e15057 页。Wiley Online Library，2024 年。3[45] Sen Yang、Zhibin Quan、Mu Nie 和 Wankou Yang。TransPose：通过 Transformer 实现关键点定位。载于《IEEE/CVF国际计算机视觉会议论文集》，第 11802–11812 页，2021 年。2, 7[46] Xinyu Yi 、 Yuxiao Zhou 和 Feng Xu。TransPose：利用六个惯性传感器实时进行三维人体平移与姿态估计。《ACM图形学汇刊》（TOG），40(4):1–13，2021 年。1, 2, 6[47] Xinyu Yi、Yuxiao Zhou、Marc Habermann、Soshi Shimada、Vladislav Golyanik、Christian Theobalt 和 Feng Xu。物理惯性姿态估计器（PIP）：利用稀疏惯性传感器进行具备物理感知能力的实时人体动作跟踪。载于《IEEE/CVF计算机视觉与模式识别会议论文集》，第 13167–13178 页，2022 年。3, 7[48] Xinyu Yi、Yuxiao Zhou 和 Feng Xu。物理非惯性姿态估计器（PNP）：对稀疏惯性人体动作捕捉中的非惯性效应进行建模。载于《ACM SIGGRAPH 2024会议论文集》，第 1–11 页，2024 年。1, 3, 7[49] Tianhong Catherine Yu、Manru Mary Zhang、Peter He、Chi-Jung Lee、Cassidy Cheesman、Saif Mahmud、Ruidong Zhang、Francois Guimbretiere 和 Cheng Zhang。Seam-pose：将衬衫接缝改造为电容传感器，用于上半身姿态跟踪。载于《第 37 届 ACM 用户界面软件与技术年度研讨会论文集》，第 1–13 页，2024 年。3[50] Tianhong Catherine Yu、Manru Mary Zhang、Luis Miguel Malenab、Chi-Jung Lee、Jacky Hao Jiang、Ruidong Zhang、Franc.ois Guimbreti`ere 和 Cheng Zhang。SeamFit：迈向实用型智能服装，实现自动运动记录。《ACM交互、移动、可穿戴与泛在技术论文集》，9(1):1–22，2025 年。3[51] Jianfeng Zhang、Yujun Cai、Shuicheng Yan、Jiashi Feng 等。直接进行多视图多人三维姿态估计。载于《神经信息处理系统进展》，34:13153–13164，2021 年。2

[52] Qitao Zhao、Ce Zheng、Mengyuan Liu 和 Chen Chen。对于三维人体姿态估计而言，结合上下文的单个二维姿态胜过数百个姿态。载于《神经信息处理系统进展》，36:27394–27413，2023 年。2[53] Jianan Zhen、Qi Fang、Jiaming Sun、Wentao Liu、Wei Jiang、Hujun Bao 和 Xiaowei Zhou。SMAP：单次拍摄的多人绝对三维姿态估计。载于《欧洲计算机视觉会议》，第 550–566 页。施普林格，2020 年。2[54] Ce Zheng、Wenhan Wu、Chen Chen、Taojiannan Yang、Sijie Zhu、Ju Shen、Nasser Kehtarnavaz 和 Mubarak Shah。基于深度学习的人体姿态估计：综述。《ACM计算调查》，56(1):1–37，2023 年。1[55] Ruonan Zheng、Jiawei Fang、Yuan Yao、Xiaoxia Gao、Chengxu Zuo、Shihui Guo 和 Yiyue Luo。FIP：通过融合柔性传感器与惯性传感器，赋予日常服装稳健的动作捕捉能力。载于《2025 CHI计算机系统人因会议论文集》，第 1–21 页，2025 年。3, 6, 8[56] Bo Zhou、Daniel Geissler、Marc Faulhaber、Clara Elisabeth Gleiss、Esther Friederike Zahn、Lala Shakti Swarup Ray、David Gamarra、Vitor Fortes Rey、Sungho Suh、Sizhen Bian 等。Mocapose：利用集成于纺织品中的电容传感器，在宽松型智能服装中进行动作捕捉。《ACM交互、移动、可穿戴与泛在技术论文集》，7(1):1–40，2023 年。3[57] Chengxu Zuo、Yiming Wang、Lishuang Zhan、Shihui Guo、Xinyu Yi、Feng Xu 和 Yipeng Qin。Loose Inertial Poser：利用附着惯性测量单元的宽松穿戴式夹克进行动作捕捉。载于《IEEE/CVF计算机视觉与模式识别会议论文集》，第 2209–2219 页，2024 年。3, 6, 8[58] Chengxu Zuo、Jiawei Huang、Xiao Jiang、Yuan Yao、Xiangren Shi、Rui Cao、Xinyu Yi、Feng Xu、Shihui Guo 和 Yipeng Qin。Transformer惯性测量单元校准器：用于惯性动作捕捉的动态人体佩戴式惯性测量单元校准。《ACM图形学汇刊》（TOG），44(4):1–14，2025 年。1, 3