



(12) 发明专利申请

(10) 申请公布号 CN 117058713 A

(43) 申请公布日 2023. 11. 14

(21) 申请号 202310992087.4
(22) 申请日 2023.08.08
(71) 申请人 东南大学
地址 211189 江苏省南京市江宁区东南大学路2号
(72) 发明人 王帅 梅洛瑜 李文炜
(74) 专利代理机构 南京众联专利代理有限公司
32206
专利代理师 叶倩
(51) Int. Cl.
G06V 40/10 (2022.01)
G01C 21/16 (2006.01)
G01C 21/00 (2006.01)
G01C 21/20 (2006.01)
G01S 19/47 (2010.01)

G01S 13/88 (2006.01)
G01S 7/02 (2006.01)
G01S 13/06 (2006.01)
G01S 13/66 (2006.01)
G06V 10/77 (2022.01)
G06V 10/80 (2022.01)
G06V 10/34 (2022.01)
G06T 7/73 (2017.01)
G06V 10/82 (2022.01)
G06N 3/0464 (2023.01)
G06N 3/0442 (2023.01)
G06N 3/0455 (2023.01)
G06N 3/084 (2023.01)

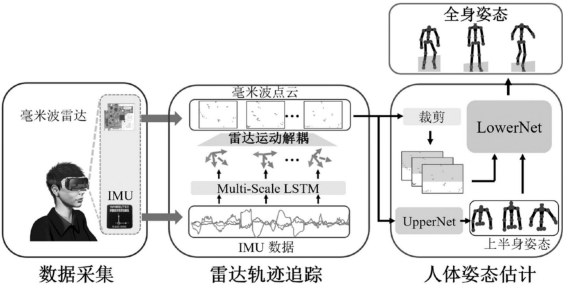
权利要求书3页 说明书8页 附图3页

(54) 发明名称

基于头戴式毫米波雷达的第一人称姿态估计方法及系统

(57) 摘要

本发明公开了一种基于头戴式毫米波雷达的第一人称姿态估计方法及系统,首先通过与雷达集成的IMU中获得雷达的运动数据,将运动数据送入神经网络学习后,获得雷达的相对6D轨迹;再将毫米波点云从雷达坐标系全部转换为H坐标系,解耦雷达运动的影响,通过神经网络学习上半身准确的关节位置;最后利用上半身和下半身之间的关联,通过神经网络分别对上半身骨架和下半身点云进行特征提取,并进行特征融合,学习人体运动先验,从而得到人体下半身准确的关节位置,最终完成全身的人体姿态估计。本方法利用人体运动先验成实现人体下半身的姿态估计,克服了下半身信息缺失的挑战,实现了在各种环境下进行稳健的人体姿态估计并避免隐私泄露,更加的精准安全。



1. 基于头戴式毫米波雷达的第一人称姿态估计方法,其特征在于,包括如下步骤:

S1, 基于头部运动特征的雷达轨迹追踪: 通过与雷达集成的IMU中获得雷达的运动数据, 将运动数据送入神经网络学习后, 获得雷达相对于H点的6D轨迹; 所述H点为人体锁骨的中间位置;

S2, 基于雷达运动解耦的上半身姿态估计: 将毫米波点云从雷达坐标系全部转换为H坐标系, 解耦雷达运动的影响, 通过神经网络学习上半身准确的关节位置;

S3, 基于人体运动先验的下半身姿态估计: 利用上半身和下半身之间的关联, 通过神经网络分别对上半身骨架和下半身点云进行特征提取, 并进行特征融合, 学习人体运动先验, 从而得到人体下半身准确的关节位置。

2. 如权利要求1所述的基于头戴式毫米波雷达的第一人称姿态估计方法, 其特征在于: 所述步骤S1具体包括:

S11: 利用和雷达集成的IMU测量得到雷达的运动数据, 其中第t帧包括3维加速度 α_t 、3维角速度 ω_t 作为输入, 经过一个(6, 512)的全连接层(FC)将其映射为512维的特征向量;

S12: 在20帧特征序列上使用双向两层的长短期记忆神经网络(LSTM)提取微小运动特征, 输出特征维度为1024维;

S13: 经过大小为(1024, 1)的自注意力机制层(self-attention)聚合每个时间步的特征, 得到该20帧序列的特征表示;

S14: 多个聚合后的特征输入同样参数的LSTM提取剧烈运动特征, 输出特征维度同样为1024维;

S15: 通过一个(1024, 9)的全连接层将特征映射到雷达相对于H点的6D运动轨迹, 即六维向量表示的相对朝向 \hat{R} 和三维向量表示的相对位置 \hat{p} , H点表示人体锁骨中间位置;

S16: 使用如下损失函数计算预测出的朝向和位置误差, 计算梯度并进行反向传播, 更新网络参数直至网络收敛:

$$\mathcal{L}_R = \alpha \sum_{t=0}^{T_1} \cos^{-1} \left(\frac{\text{tr}(\hat{R}^t R^{tT}) - 1}{2} \right) + \beta \sum_{t=0}^{T_1} \|\hat{p}^t - p^t\|_2$$

其中 α, β 为权重系数, R^t 为真实的雷达朝向, p^t 为雷达真实的位置, T_1 为点云帧序列长度。

3. 如权利要求2所述的基于头戴式毫米波雷达的第一人称姿态估计方法, 其特征在于: 所述步骤S2具体包括:

S21: 基于步骤S1得到的雷达相对朝向R和位置t, 将毫米波雷达反射人体得到的点云坐标转换到H坐标系, 其中每个点由六维向量表示, 包括xyz三维坐标、距离r、速度v以及强度s;

S22: 将点云输入到共享参数的多层感知机(shared-MLP)中进行特征提取, 其大小为(6, 16, 32, 64), 每一层之后使用批归一化(BN)和ReLU激活函数;

S23: 使用(64, 1)的自注意力机制自适应学习每个点的权重, 对不同点的特征进行聚合, 得到帧级的高维特征表示;

S24: 帧级的特征表示输入三层双向的LSTM中提取时间序列特征, 输出特征为128维;

S25:大小为(128,64,42)的多层感知机(MLP)将时间序列特征映射到上半身关节相对于其父关节的旋转角度 RJ_i ,其中 $i=1,2,\dots,14$, RJ_i 用六维向量表示;

S26:将六维旋转表示使用Schmidt正交化转换为旋转矩阵表示,应用人体正向运动学(forwardkinematics)从关节旋转矩阵得到人体上半身关节位置 \hat{J}_i , \hat{J}_i 用三维向量表示其xyz三维坐标;

S27:使用如下损失函数计算预测出的关节位置误差,计算梯度并进行反向传播,更新参数直至网络收敛;

$$\mathcal{L}_U = \sum_{t=0}^{T_1} \sum_{i=1}^{15} \|\hat{J}_i^t - J_i^t\|_2$$

其中 J_i^t 为真实的上半身关节位置。

4.如权利要求3所述的基于头戴式毫米波雷达的第一人称姿态估计方法,其特征在于:所述上半身关节包括左右手腕、左右手肘、左右肩膀、左右锁骨、头部、颈部、胸部、肚脐、骨盆及左右髋关节,其中颈部作为根关节。

5.如权利要求3所述的基于头戴式毫米波雷达的第一人称姿态估计方法,其特征在于:所述步骤S3具体包括:

S31:将步骤S2得到的人体上半身关节位置送入三层大小分别为(32,64,128)的图卷积神经网络(GCN)进行特征提取,得到上半身姿态特征H;

S32:将毫米波点云以骨盆关节节点的x轴坐标为界限进行裁剪,得到下半身点云,不足64的进行补零,送入大小为(6,32,64,128)的共享参数的多层感知机得到下半身点云特征Z;

S33:将点云特征Z和骨架特征H使用尺寸为(128,128)的互注意力机制及残差连接进行特征融合;

S34:使用(128,1)的自注意力机制聚合为帧级特征;

S35:将帧级特征送入三层双向LSTM中提取时间序列特征;

S36:使用大小为(128,64,36)的多层感知机将时间序列特征映射到六个下半身关节旋转角度,同样用六维向量表示;

S37:将六维旋转表示使用Schmidt正交化转换为旋转矩阵表示,应用正向运动学得到下半身关节位置 \hat{L}_i ;

S38:使用如下损失函数计算预测出的关节位置误差,计算梯度并进行反向传播,更新参数直至网络收敛;

$$\mathcal{L}_L = \sum_{t=0}^{T_1} \sum_{i=1}^8 \|\hat{L}_i^t - L_i^t\|_2$$

其中 L_i^t 为真实的下半身关节位置。

6.如权利要求5所述的基于头戴式毫米波雷达的第一人称姿态估计方法,其特征在于:所述下半身关节包括左右脚、左右脚腕,左右膝盖和左右髋关节,其中左右髋为两个根关节位置。

7. 如权利要求5所述的基于头戴式毫米波雷达的第一人称姿态估计方法,其特征在于:所述步骤S26和步骤S37中正向运动学计算如下:

$$J_i = J_{parent(i)} + R J_i (\bar{J}_i - \bar{J}_{parent(i)})$$

其中 $J_{parent(i)}$ 表示在骨架结构中 J_i 的父节点, \bar{J}_i 和 $\bar{J}_{parent(i)}$ 分别表示 J_i 和 $J_{parent(i)}$ 的初始位置。

8. 基于头戴式毫米波雷达的第一人称姿态估计系统,其特征在于:至少包括雷达轨迹追踪模块、上半身姿态估计模块和下半身姿态估计模块,

所述雷达轨迹追踪模块:利用与雷达集成的IMU,基于头部运动特征的雷达轨迹追踪网络,训练得到准确的雷达相对6D轨迹,所述雷达轨迹追踪网络通过不同时间尺度的LSTM及时间步上的自注意力融合机制来捕获人体头部运动规律;

所述上半身姿态估计模块:基于雷达姿态解耦的上半身姿态估计网络,训练得到上半身姿态;所述模块的输入为转换到H坐标系的毫米波点云,通过PointNet中共享参数的MLP对不同的点提取语义特征,随后通过自注意力机制聚合出帧级特征,并用LSTM捕捉不同帧之间点云的联系,最后通过MLP和正向运动学模块将隐藏层特征映射为人体上半身姿态;

所述下半身姿态估计模块:基于人体运动先验,上半身姿态和下半身点云被融合以估计下半身姿态,从而完成全身的人体姿态估计。

9. 如权利要求8所述的基于头戴式毫米波雷达的第一人称姿态估计系统,其特征在于:所述下半身姿态估计模块的输入上半身姿态模块所输出的上半身姿态,及通过骨盆点裁剪的下半身点云;分别采用对应的特征提取网络提取特征后,对稀疏点云在上半身姿态的语义上下文中进行特征融合,得到更准确的下半身特征,再通过聚合和时序特征提取,将隐藏层特征映射为人体下半身姿态,和上半身姿态拼接后得到全身的人体姿态。

基于头戴式毫米波雷达的第一人称姿态估计方法及系统

技术领域

[0001] 本发明属于无线感知技术领域,主要涉及了一种基于头戴式毫米波雷达的第一人称姿态估计方法及系统。

背景技术

[0002] 人体姿态估计旨在重建三维身体动作,在许多应用领域中起着重要作用,如运动分析、人机交互、增强现实(AR)、虚拟现实(VR)和医疗康复。传统的人体姿态是通过安装在场景周围的运动捕捉设备(如摄像头)或大量穿戴在身体上的惯性测量单元(IMU)来捕捉的。然而,固定的运动捕捉设备基础设施受限于有限的感知范围,限制了人体运动的空间,因此无法满足涉及大范围移动的日常活动。另一方面,佩戴在身体上的IMU需要繁琐的设置和复杂的校准操作,给用户带来不便,并且阻碍了正常活动和社交互动。

[0003] 近年来,头戴设备(如VR眼镜和智能头盔)的普及启发了一种第一人称姿态估计的新方向。具体来说,它通过由用户佩戴的单个头戴设备来估计姿势,为用户提供了移动性和便利性。例如,Apple Vision Pro配备了多个摄像头,用于替代传统的手柄。然而,基于视觉的方法对光照条件、烟雾、灰尘以及人体外观非常敏感,并经常引起隐私方面的问题。

[0004] 新兴的集成毫米波雷达具有低成本和小型化(约10厘米),因此很容易嵌入到头戴设备(如AR头盔)中。此外,最近的应用展示了雷达在提供各种人体感知能力方面的潜力,同时对恶劣的光照或天气具有鲁棒性,保护隐私,对用户不具侵入性。基于雷达的第一人称姿态估计可以实现广泛的应用,如沉浸式虚拟现实、运动辅助分析、用于急救人员的AR视觉增强以及用于驾驶员的安全行为检测。

[0005] 尽管雷达在各种人体感知任务中取得了最近的成功,但第一人称视角对姿势估计提出了两个独特的挑战。首先,头戴设备是非静止的,因此雷达信号不仅包含姿势的变化,还包括由头部运动引起的随机设备移动。现有最先进的姿态估计方法无法捕捉稳定的时空特征,导致关节定位误差显著(最大可达16厘米)。其次,由于自上而下的视角,雷达信号在下半身受到镜面反射和上肢自遮挡的严重影响,导致雷达仅能感知到极少量的下肢运动,因此准确的估计下半身具有一定的挑战性。

发明内容

[0006] 本发明正是针对现有技术中姿态估计方法误差较大且下半身信息缺失的问题,提供一种基于头戴式毫米波雷达的第一人称姿态估计方法及系统,首先进行基于头部运动特征的雷达轨迹追踪,通过与雷达集成的IMU中获得雷达的运动数据,将运动数据送入神经网络学习后,获得雷达的相对6D轨迹;再将毫米波点云从雷达坐标系全部转换为H坐标系,解耦雷达运动的影响,通过神经网络学习上半身准确的关节位置;最后利用上半身和下半身之间的关联,通过神经网络分别对上半身骨架和下半身点云进行特征提取,并进行特征融合,学习人体运动先验,从而得到人体下半身准确的关节位置,最终完成全身的人体姿态估计。本方法利用人体运动先验成实现人体下半身的姿态估计,克服了下半身信息缺失的挑

战,实现了在各种环境下进行稳健的人体姿态估计并避免隐私泄露,更加的精准安全。

[0007] 为了实现上述目的,本发明采取的技术方案是:基于头戴式毫米波雷达的第一人称姿态估计方法,包括如下步骤:

[0008] S1,基于头部运动特征的雷达轨迹追踪:通过与雷达集成的IMU中获得雷达的运动数据,将运动数据送入神经网络学习后,获得雷达相对于H点的6D轨迹;所述H点为人体锁骨的中间位置;

[0009] S2,基于雷达运动解耦的上半身姿态估计:将毫米波点云从雷达坐标系全部转换为H坐标系,解耦雷达运动的影响,通过神经网络学习上半身准确的关节位置;

[0010] S3,基于人体运动先验的下半身姿态估计:利用上半身和下半身之间的关联,通过神经网络分别对上半身骨架和下半身点云进行特征提取,并进行特征融合,学习人体运动先验,从而得到人体下半身准确的关节位置。

[0011] 作为本发明的一种改进,所述步骤S1具体包括:

[0012] S11:利用和雷达集成的IMU测量得到雷达的运动数据,其中第t帧包括3维加速度 α_t 、3维角速度 ω_t 作为输入,经过一个(6,512)的全连接层(FC)将其映射为512维的特征向量;

[0013] S12:在20帧特征序列上使用双向两层的长短期记忆神经网络(LSTM)提取微小运动特征,输出特征维度为1024维;

[0014] S13:经过大小为(1024,1)的自注意力机制层(self-attention)聚合每个时间步的特征,得到该20帧序列的特征表示;

[0015] S14:多个聚合后的特征输入同样参数的LSTM提取剧烈运动特征,输出特征维度同样为1024维;

[0016] S15:通过一个(1024,9)的全连接层将特征映射到雷达相对于H点的运动轨迹,即六维向量表示的朝向 \hat{R} 和三维向量表示的位置 \hat{p} ,H点表示人体锁骨中间位置;

[0017] S16:使用如下损失函数计算预测出的朝向和位置误差,计算梯度并进行反向传播,更新网络参数直至网络收敛:

$$[0018] \quad \mathcal{L}_R = \alpha \sum_{t=0}^{T_1} \cos^{-1} \left(\frac{\text{tr}(\hat{R}^t R^{tT}) - 1}{2} \right) + \beta \sum_{t=0}^{T_1} \|\hat{p}^t - p^t\|_2$$

[0019] 其中 α, β 为权重系数, R^t 为真实的雷达朝向, p^t 为雷达真实的位置, T_1 为点云帧序列长度。

[0020] 作为本发明的一种改进,所述步骤S2具体包括:

[0021] S21:基于步骤S1得到的雷达相对朝向R和位置p,将毫米波雷达反射人体得到的点云坐标转换到H坐标系,其中每个点由六维向量表示,包括xyz三维坐标、距离r、速度v以及强度s;

[0022] S22:将点云输入到共享参数的多层感知机(shared-MLP)中进行特征提取,其大小为(6,16,32,64),每一层之后使用批归一化(BN)和ReLU激活函数;

[0023] S23:使用(64,1)的自注意力机制自适应学习每个点的权重,对不同点的特征进行聚合,得到帧级的高维特征表示;

[0024] S24:帧级的特征表示输入三层双向的LSTM中提取时间序列特征,输出特征为128维;

[0025] S25:大小为(128,64,42)的多层感知机(MLP)将时间序列特征映射到上半身关节相对于其父关节的旋转角度 RJ_i ,其中 $i=1,2,\dots,14$, RJ_i 用六维向量表示;

[0026] S26:将六维旋转表示使用Schmidt正交化转换为旋转矩阵表示,应用人体正向运动学(forwardkinematics)从关节旋转矩阵得到人体上半身关节位置 \hat{J}_i , \hat{J}_i 用三维向量表示其xyz三维坐标;

[0027] S27:使用如下损失函数计算预测出的关节位置误差,计算梯度并进行反向传播,更新参数直至网络收敛;

$$[0028] \quad \mathcal{L}_U = \sum_{t=0}^{T_1} \sum_{i=1}^{15} \|\hat{J}_i^t - J_i^t\|_2$$

[0029] 其中 J_i^t 为真实的上半身关节位置。

[0030] 作为本发明的另一种改进,所述上半身关节包括左右手腕、左右手肘、左右肩膀、左右锁骨、头部、颈部、胸部、肚脐、骨盆及左右髋关节,其中颈部作为根关节。

[0031] 作为本发明的另一种改进,所述步骤S3具体包括:

[0032] S31:将步骤S2得到的人体上半身关节位置送入三层大小分别为(32,64,128)的图卷积神经网络(GCN)进行特征提取,得到上半身姿态特征H;

[0033] S32:将毫米波点云以骨盆关节节点的x轴坐标为界限进行裁剪,得到下半身点云,不足64的进行补零,送入大小为(6,32,64,128)的共享参数的多层感知机得到下半身点云特征Z;

[0034] S33:将点云特征Z和骨架特征H使用尺寸为(128,128)的互注意力机制及残差连接进行特征融合;

[0035] S34:使用(128,1)的自注意力机制聚合为帧级特征;

[0036] S35:将帧级特征送入三层双向LSTM中提取时间序列特征;

[0037] S36:使用大小为(128,64,36)的多层感知机将时间序列特征映射到六个下半身关节旋转角度,同样用六维向量表示;

[0038] S37:将六维旋转表示使用Schmidt正交化转换为旋转矩阵表示,应用正向运动学得到下半身关节位置 \hat{L}_i ;

[0039] S38:使用如下损失函数计算预测出的关节位置误差,计算梯度并进行反向传播,更新参数直至网络收敛;

$$[0040] \quad \mathcal{L}_L = \sum_{t=0}^{T_1} \sum_{i=1}^8 \|\hat{L}_i^t - L_i^t\|_2$$

[0041] 其中 L_i^t 为真实的下半身关节位置。

[0042] 作为本发明的又一种改进,所述下半身关节包括左右脚、左右脚腕,左右膝盖和左右髋关节,其中左右髋为两个根关节位置。

[0043] 作为本发明的又一种改进,所述步骤S26和步骤S37中正向运动学计算如下:

$$[0044] \quad J_i = J_{parent(i)} + R J_i (\bar{J}_i - \bar{J}_{parent(i)})$$

[0045] 其中 $J_{parent(i)}$ 表示在骨架结构中 J_i 的父节点, \bar{J}_i 和 $\bar{J}_{parent(i)}$ 分别表示 J_i 和 $J_{parent(i)}$ 的初始位置。

[0046] 为了实现上述目的,本发明还采取的技术方案是:基于头戴式毫米波雷达的第一人称姿态估计系统,至少包括雷达轨迹追踪模块、上半身姿态估计模块和下半身姿态估计模块,

[0047] 所述雷达轨迹追踪模块:利用与雷达集成的IMU,基于头部运动特征的雷达轨迹追踪网络,训练得到准确的雷达相对6D轨迹,所述雷达轨迹追踪网络通过不同时间尺度的LSTM及时间步上的自注意力融合机制来捕获人体头部运动规律;

[0048] 所述上半身姿态估计模块:基于雷达姿态解耦的上半身姿态估计网络,训练得到上半身姿态;所述模块的输入为转换到H坐标系的毫米波点云,通过PointNet中共享参数的MLP对不同的点提取语义特征,随后通过自注意力机制聚合出帧级特征,并用LSTM捕捉不同帧之间点云的联系,最后通过MLP和正向运动学模块将隐藏层特征映射为人体上半身姿态;

[0049] 所述下半身姿态估计模块:基于人体运动先验,上半身姿态和下半身点云被融合以估计下半身姿态,从而完成全身的人体姿态估计。

[0050] 作为本发明的另一种改进,所述下半身姿态估计模块的输入上半身姿态模块所输出的上半身姿态,及通过骨盆点裁剪的下半身点云;分别采用对应的特征提取网络提取特征后,对稀疏点云在上半身姿态的语义上下文中进行特征融合,得到更准确的下半身特征,再通过聚合和时序特征提取,将隐藏层特征映射为人体下半身姿态,和上半身姿态拼接后得到全身的人体姿态。

[0051] 与现有技术相比,本发明具有的有益效果:

[0052] (1) 本发明所提出的基于头戴式毫米波雷达的第一人称姿态估计方法能够实现各种环境下,活动范围不限且稳健的人体姿态估计,相比于传统的摄像头依赖光照、暴露隐私、易受恶劣天气影响的限制,本发明方法仅使用低成本的雷达和IMU设备,不依赖光照并且无视烟雾粉尘等恶劣条件,更加易于实现,也避免了隐私问题。

[0053] (2) 本发明所提出的步骤S1基于头部运动特征的雷达轨迹追踪方法,避免了IMU位置测量的漂移问题,得到准确的6D雷达轨迹。

[0054] (3) 本发明所提出的步骤S3基于人体运动先验的下半身姿态估计,能够在下半身信息缺乏的情况下实现准确的下半身姿态估计效果,结果更加精准。

附图说明

[0055] 图1是本发明方法的步骤流程示意图;

[0056] 图2是本发明方法步骤S1中相对于颈关节(H点)的雷达姿态示意图;

[0057] 图3是本发明系统中雷达轨迹追踪模块的工作原理图;

[0058] 图4是本发明系统中上半身姿态估计模块(UpperNet)的工作原理图;

[0059] 图5是本发明系统中下半身姿态估计模块(LowerNet)的工作原理图;

[0060] 图6是本发明测试例中系统实验设备原型图。

具体实施方式

[0061] 下面结合附图和具体实施方式,进一步阐明本发明,应理解下述具体实施方式仅用于说明本发明而并不用于限制本发明的范围。

[0062] 实施例1

[0063] 基于头戴式毫米波雷达的第一人称姿态估计系统,至少包括雷达轨迹追踪模块、上半身姿态估计模块和下半身姿态估计模块,

[0064] 其中,雷达轨迹追踪模块如图3所示,利用与雷达集成的IMU,基于头部运动特征的雷达轨迹追踪网络,训练得到准确的雷达相对6D轨迹。所述雷达相对轨迹为相对于人体颈关节(以H点表示)的轨迹如图2所示。雷达轨迹追踪网络通过不同时间尺度的LSTM及时间步上的自注意力融合机制来捕获人体头部运动规律;

[0065] 所述上半身姿态估计模块如图4所示,基于雷达姿态解耦的上半身姿态估计网络,训练得到上半身姿态。该模块的输入为转换到H坐标系的毫米波点云,通过PointNet中共享参数的MLP对不同的点提取语义特征,随后通过自注意力机制聚合出帧级特征,并用LSTM捕捉不同帧之间点云的联系,最后通过MLP和正向运动学模块将隐藏层特征映射为人体上半身姿态;

[0066] 所述下半身姿态估计模块如图5所示,基于人体运动先验,上半身姿态和下半身点云被融合以估计下半身姿态,从而完成全身的人体姿态估计。该模块的输入包含两部分,一部分为上半身姿态模块所输出的上半身姿态,另一部分为通过骨盆点裁剪的下半身点云,由于两种输入共享用一个坐标空间,因此分别采用对应的特征提取网络提取特征后,对稀疏点云在上半身姿态的语义上下文中进行特征融合,从而得到更准确的下半身特征,随后同样通过聚合和时序特征提取,将隐藏层特征映射为人体下半身姿态,和上半身姿态拼接后得到完整的人体姿态。

[0067] 本系统首先利用与雷达集成的IMU,基于头部运动特征的雷达轨迹追踪网络,训练得到准确的雷达相对6D轨迹;其次基于雷达姿态解耦的上半身姿态估计网络,训练得到较准确的上半身姿态;最后,基于人体运动先验,上半身姿态和下半身点云被融合以估计下半身姿态,从而完成全身的人体姿态估计。本系统使用低成本的雷达和IMU设备,克服了传统的摄像头依赖光照、暴露隐私、易受恶劣天气影响的限制,易于实现,同时克服了下半身信息缺失的问题,姿态估计更加的精准安全。

[0068] 实施例2

[0069] 基于头戴式毫米波雷达的第一人称姿态估计方法,如图1所示,通过雷达轨迹追踪和人体运动先验来实现稳定的第一人称姿态估计,具体包括如下步骤:

[0070] S1,基于头部运动特征的雷达轨迹追踪:通过与雷达集成的IMU中获得雷达的运动数据,将运动数据送入神经网络学习后,获得雷达相对于H点的6D轨迹;

[0071] 其中,H点为人体锁骨的中间位置,如图2所示;用以实现雷达轨迹追踪的神经网络模型如图3所示。 $\mathbf{x}_{imu}^{t_1:t_2}$ 表示在 t_1 到 t_2 时刻内的IMU数据,每一时刻的IMU数据为加速度 \mathbf{a}_t 和角速度 $\mathbf{\omega}_t$ 拼接为的六维向量; δ 和 T_2 分别为选取的短时间区间和长时间区间; T_1 为点云帧时间区间长度; R_t 和 P_t 表示雷达的朝向和位置;所述步骤具体包括:

[0072] S11: \mathbf{x}_{imu} 经过一个(6,512)的全连接层将其映射为512维的特征向量;

[0073] S12:在 δ 帧特征序列上使用双向两层的LSTM提取微小运动特征,输出特征维度为1024维;

[0074] S13:随后经过大小为(1024,1)的自注意力机制层聚合每个时间步的特征,得到该20帧序列的特征表示;

[0075] S14:多个聚合后的特征输入同样参数的LSTM提取剧烈运动特征,输出特征维度同样为1024维;

[0076] S15:通过一个(1024,9)的全连接层将特征映射到雷达相对于H点的运动轨迹,即六维向量表示的朝向 \hat{R} 和三维向量表示的位置 \hat{p} ,H点表示人体锁骨中间位置;

[0077] S16:使用如下损失函数计算预测出的朝向和位置误差,计算梯度并进行反向传播,更新网络参数直至网络收敛:

$$[0078] \quad \mathcal{L}_R = \alpha \sum_{t=0}^{T_1} \cos^{-1} \left(\frac{\text{tr}(\hat{R}^t R^{tT}) - 1}{2} \right) + \beta \sum_{t=0}^{T_1} \|\hat{p}^t - p^t\|_2 \quad .$$

[0079] S2,基于雷达运动解耦的上半身姿态估计:将毫米波点云从雷达坐标系全部转换为H坐标系,解耦雷达运动的影响,通过神经网络学习上半身准确的关节位置;

[0080] 在得到雷达姿态之后,本发明将模型过程中的坐标空间均转换为以H点为根节点的坐标空间,具体来说是将毫米波点云和人体姿态均转换到H坐标系下,继而使用如图4所示的神经网络模型得到上半身姿态,所述步骤具体包括:

[0081] S21:基于步骤S1得到的雷达相对旋转R和平移t,将毫米波雷达反射人体得到的点云坐标转换到H坐标系,其中每个点由六维向量表示,包括xyz三维坐标、距离r、速度v以及强度s;点云每一帧不足128个点的进行补零;

[0082] S22:将点云输入到共享参数的多层感知机中进行特征提取,其大小为(6,16,32,64),每一层之后使用批归一化(BN)和ReLU激活函数;

[0083] S23:之后使用(64,1)的自注意力机制自适应学习每个点的权重,从而对不同点的特征进行聚合,进而得到帧级的高维特征表示;

[0084] S24:帧级的特征表示输入三层双向的LSTM中提取时间序列特征,输出特征为128维;

[0085] S25:大小为(128,64,42)的多层感知机将时间序列特征映射到上半身十四个关节相对于其父关节的旋转角度 RJ_i ,其中 $i=1,2,\dots,14$, RJ_i 用六维向量表示;

[0086] S26:将六维旋转表示使用Schmidt正交化转换为旋转矩阵表示,应用人体正向运动学从关节旋转矩阵得到人体上半身关节位置 \hat{J}_i , \hat{J}_i 用三维向量表示其xyz三维坐标;其中正向运动学计算如下:

$$[0087] \quad J_i = J_{parent(i)} + RJ_i(\bar{J}_i - \bar{J}_{parent(i)})$$

[0088] S27:使用如下损失函数计算预测出的关节位置误差,计算梯度并进行反向传播,更新参数直至网络收敛;

$$[0089] \quad \mathcal{L}_U = \sum_{t=0}^{T_1} \sum_{i=1}^{15} \|\hat{J}_i^t - J_i^t\|_2 \quad .$$

[0090] S3,基于人体运动先验的下半身姿态估计:利用上半身和下半身之间的关联,通过

神经网络分别对上半身骨架和下半身点云进行特征提取,并进行特征融合,学习人体运动先验,从而得到人体下半身准确的关节位置。

[0091] 利用步骤S2和步骤S1提取到的雷达姿态解耦后的上半身姿态和毫米波点云,使用如图5所示的深度神经网络模型估计人体下半身姿态,所述步骤具体包括:

[0092] S31:将步骤S2得到的人体上半身关节位置建图,每个关节作为图的一个节点,肢体作为一条边,送入三层大小分别为(32,64,128)的图卷积神经网络(GCN)进行特征提取,得到上半身姿态特征H;

[0093] S32:将毫米波点云以骨盆关节的x轴坐标为界限进行裁剪,得到下半身点云,不足64的进行补零,送入大小为(6,32,64,128)的共享参数的多层感知机得到下半身点云特征Z;

[0094] S33:将点云特征Z和骨架特征H使用尺寸为(128,128)的互注意力机制及残差连接进行特征融合,其中Z作为查询,H作为键值,具体计算方式如下:

$$[0095] \quad A^t = Z^t + \text{softmax}\left(\frac{Z^t W_Q (H^t W_K)^T}{\sqrt{d'}}\right) H^t W_V$$

[0096] 其中 W_Q, W_K, W_V 是可学习的参数,维度为(128,128), d' 为128;

[0097] S34:使用(128,1)的自注意力机制聚合为帧级特征;

[0098] S35:将帧级特征送入三层双向LSTM中提取时间序列特征;

[0099] S36:使用大小为(128,64,36)的多层感知机将时间序列特征映射到六个下半身关节旋转角度,同样用六维向量表示;

[0100] S37:同步步骤S26将六维旋转表示使用Schmidt正交化转换为旋转矩阵表示,应用正向运动学得到下半身关节位置 \hat{L}_i ;

[0101] S38:使用如下损失函数计算预测出的关节位置误差,计算梯度并进行反向传播,更新参数直至网络收敛。

$$[0102] \quad \mathcal{L}_L = \sum_{t=0}^{T_1} \sum_{i=1}^8 \|\hat{L}_i^t - L_i^t\|_2。$$

[0103] 本发明使用数据驱动的方法利用深度神经网络学习人体的头部运动规律和身体运动规律,克服雷达移动和下半身信息稀疏带来的挑战,从而实现准确的第一人称姿态估计。从概念上讲,本发明是一种提供高移动性、环境鲁棒性和隐私保护的姿态估计设计,更加的精准安全。

[0104] 测试例

[0105] 为了验证本方案提出方法的有效性,在室内环境下进行人体姿态估计实验,展示本方案的性能优势及影响。

[0106] 采取如下配置执行实验:

[0107] 硬件设置:1.毫米波雷达IWR6843ISK-ODS,用于采集毫米波点云数据;

[0108] 2.IMU Wheeltec N100,用于采集IMU数据;

[0109] 3.Azure Kinect v2,用于采集人体真实姿态,训练网络;

[0110] 4.装有棋盘格的3D打印头盔,用于安装和校准雷达和IMU。

[0111] 实验设计:

[0112] 1.实验者要求佩戴如图6所示的实验头盔并执行指定动作;

[0113] 2.动作包括以下十三种日常动作:原地走路;走路;手臂伸展;摇头;点头;转头;原地走路时左右看;原地走路时上下看;挥舞手臂时上下看;挥舞手臂时左右看;弓步;高抬腿;下蹲;

[0114] 3.采用平均关节位置误差(用S表示)和平均旋转角度误差(Q)来衡量姿态估计效果的准确性。其中S为预测关节位置和真实位置的平均欧式距离,Q为预测关节旋转角度和真实关节旋转角度的平均角度差;

[0115] 4.选择mmMesh,P4Transformer,PCB,mmPose-NLP作为对比方法,其中mmMesh为现有最好方法。

[0116] 实验结果如下表所示:

方法	全身		上半身		下半身	
	Q (°)	S (cm)	Q (°)	S (cm)	Q (°)	S (cm)
mm-Pose	10.885	10.734	10.439	8.964	11.923	13.682
P4Transformer	11.804	10.873	11.326	9.052	12.921	13.832
PCB	11.278	11.031	10.796	9.070	12.404	14.214
mmPose-NLP	10.640	10.406	10.453	8.492	11.077	13.493
mmMesh	7.215	7.783	7.127	6.639	7.423	9.529
本系统	4.914	4.346	5.217	4.287	4.207	5.460

[0118] 表中可以看出,其中包括了全身、上半身、下半身的估计准确度对比。本方案达到了在全身姿态估计中平均4.3cm和4.9°的准确度,分别在上半身达到了平均4.2cm,5.2°的准确度,下半身达到了平均5.4cm,4.2°的准确度。相比于现有最好的方法(mmMesh),本系统提升了44.2%的准确度。

[0119] 上述实验或数据可看出,在人体的日常动作中,本发明可以实现平均4.3cm的关节定位误差和4.9°的关节旋转角度误差。

[0120] 综上,本方案提出了一种基于头戴式毫米波雷达的第一人称姿态估计方法及系统,本方法利用人体运动先验实现人体下半身的姿态估计,克服了下半身信息缺失的挑战,实现了在各种环境下进行稳健的人体姿态估计,更加的精准安全,同时相比于传统的摄像头依赖光照、暴露隐私、易受恶劣天气影响的限制,本发明方法仅使用低成本的雷达和IMU设备,不依赖光照并且无视烟雾粉尘等恶劣条件,更加易于实现,也避免了隐私问题。

[0121] 需要说明的是,以上内容仅仅说明了本发明的技术思想,不能以此限定本发明的保护范围,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰均落入本发明权利要求书的保护范围之内。

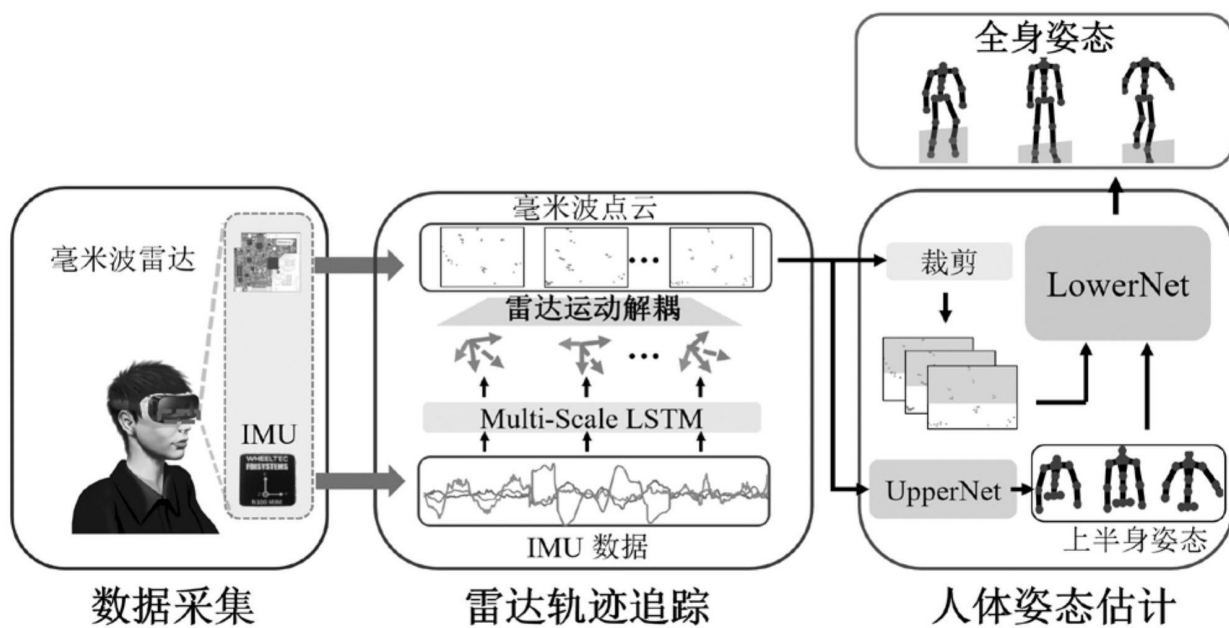


图1

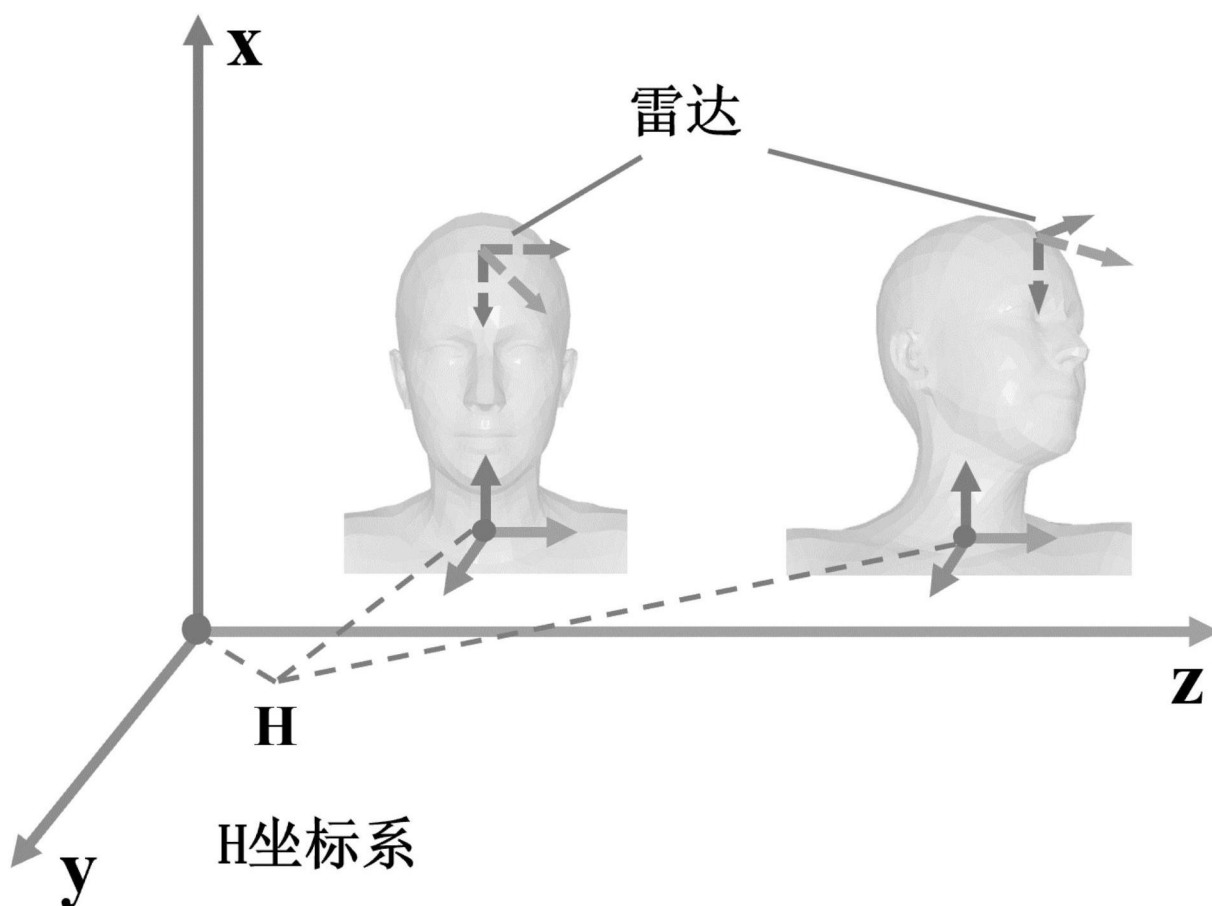


图2

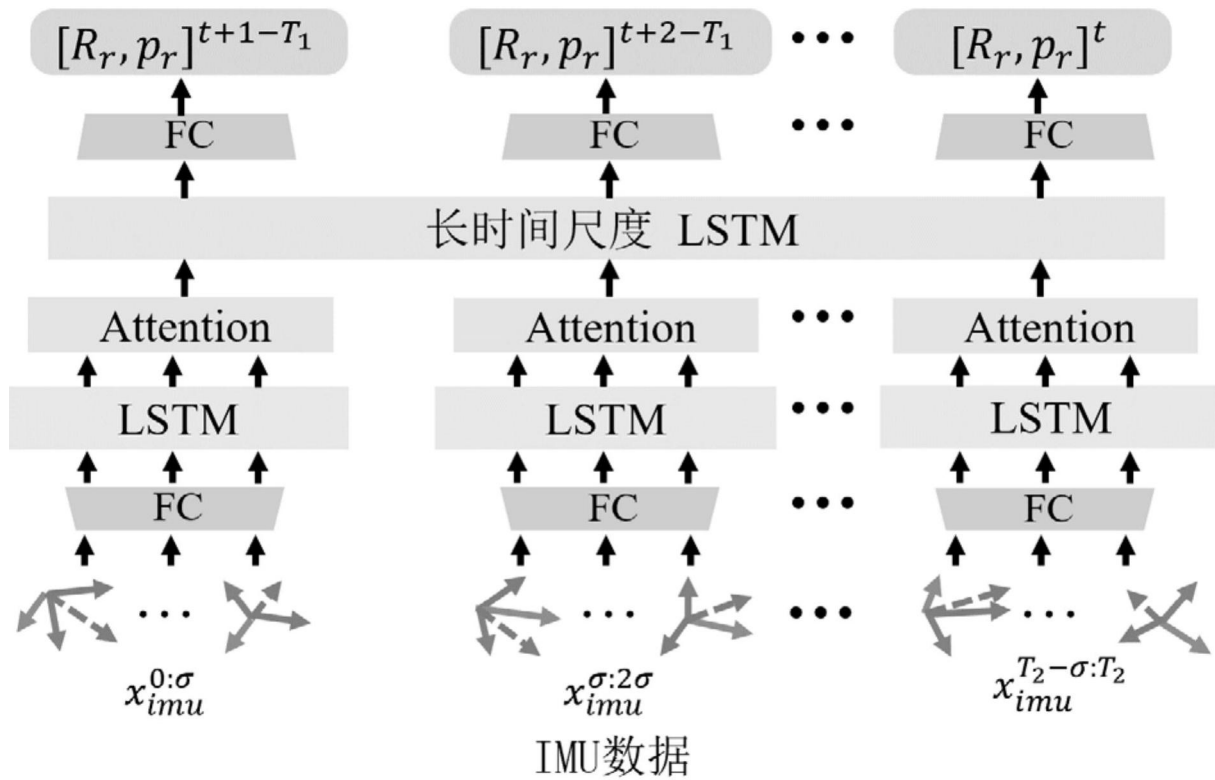


图3

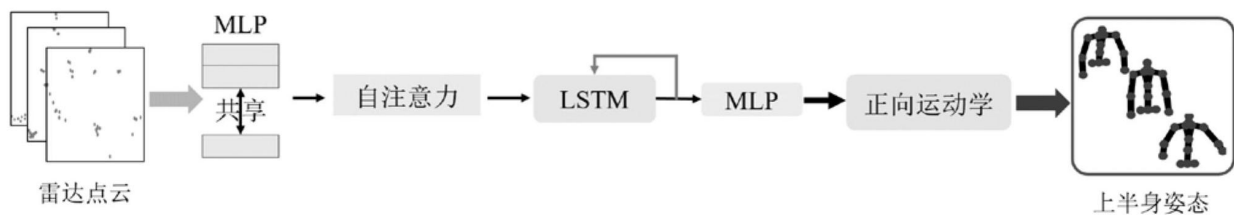


图4

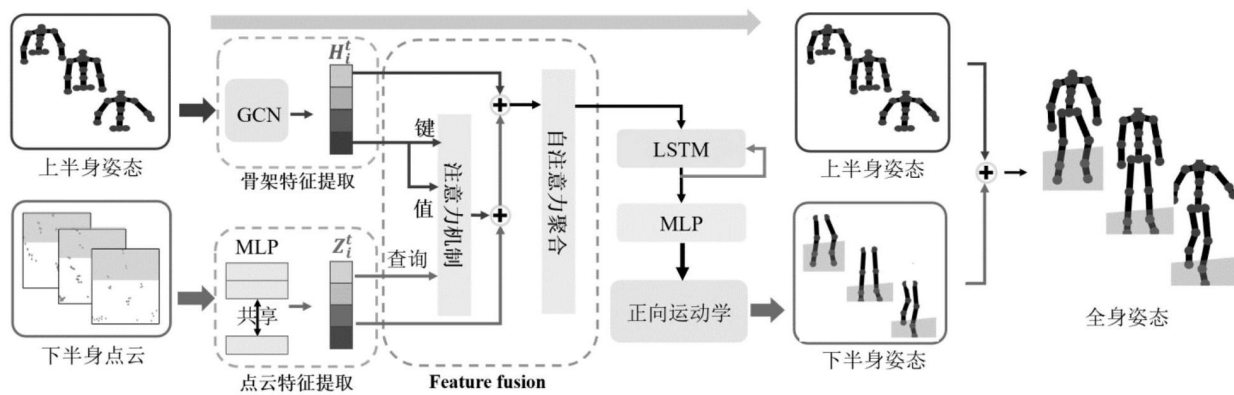


图5

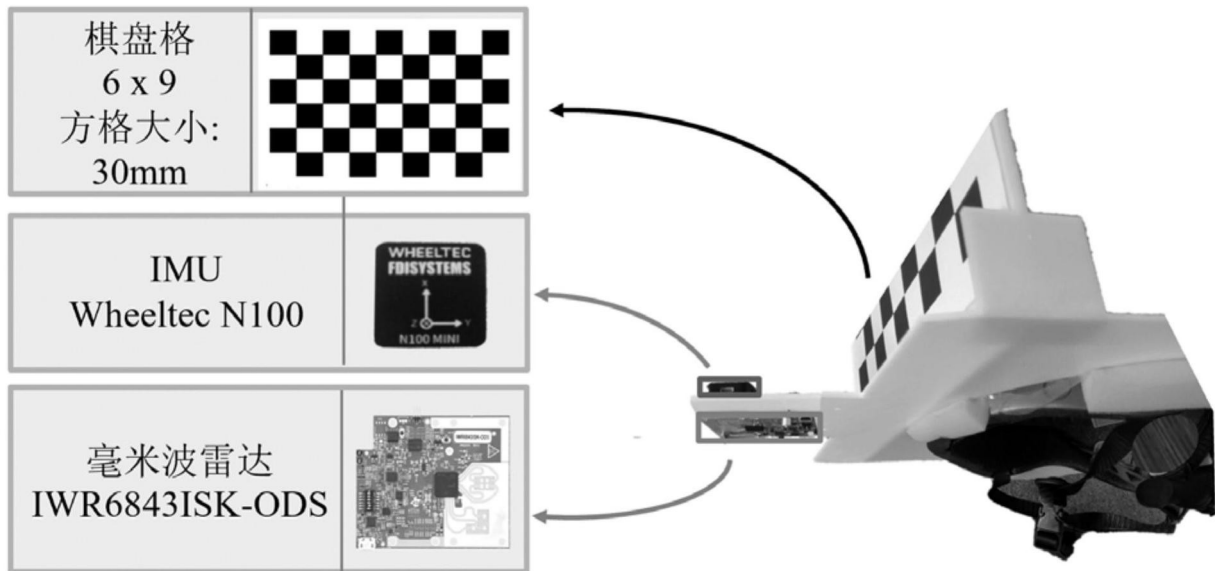


图6