

一种基于运动基识别的人体关节运动跟踪粒子滤波估计方法

张亚南¹, 朱方², 叶龙¹, 张勤¹

(1.中国传媒大学媒介音视频教育部重点实验室, 北京 100024; 2.中兴通讯股份有限公司南京研发中心, 南京 210004)

摘要: 人体关节运动跟踪是非线性、非高斯系统的运动状态估计问题, 粒子滤波器是实现人体运动跟踪的有效手段。粒子状态的预测与更新是影响粒子滤波器性能的关键, 预测模型反映人体运动规律的程度是决定使用粒子滤波器能否进行关节运动准确跟踪的主要因素之一。本文提出一种基于人体运动模式识别的关节运动跟踪粒子滤波器架构, 在将运动模式定义为运动基的基础上, 利用 $R(2+1)D$ 网络进行运动基类型识别。同时, 根据识别所得到的运动基概率密度分布, 分配每个运动基对应预测模型的粒子个数并进行关节运动状态的先验概率密度分布计算。在粒子状态更新阶段, 选取颜色直方图特征计算粒子适应度, 在对粒子状态进行重采样更新的基础上修正运动基的概率密度分布, 从而达到了基于粒子滤波器的人体运动模式识别与状态跟踪联合实现的目的。实现结果表明, 本文提出的新型粒子滤波器架构在关节跟踪准确性上远优于传统粒子滤波器方案, 与基于深度学习的点对点人体关节运动跟踪方法也具有一定的可比性。

关键词: 运动基; 粒子滤波; 预测模型; 运动跟踪

中图分类号: O422 **文献标识码:** A

A particle filter estimation method for human joint motion tracking based on dyname recognition

ZHANG Yanan¹, ZHU Fang², YE Long¹, ZHANG Qin¹

(1.Key Lab of Media Audio & Video of Ministry of Education, Communication University of China, Beijing 100024, China; 2. Nanjing R&D Center, ZTE Corporation, Nanjing 210004, China)

Abstract: Human joint motion tracking is a non-linear, non-Gaussian system motion state estimation problem, so particle filters have become an effective means to achieve human motion tracking. The prediction and update of the particle state is the key to affecting the performance of the particle filter. The degree to which the prediction model reflects the law of human motion is one of the main factors that determines whether the particle filter can be used to accurately track the joint motion. This paper proposes a joint motion tracking particle filter architecture based on human motion pattern recognition. On the basis of defining the motion pattern as a dyname, the $R(2+1)D$ network is used for dyname type recognition. At the same time, according to the probability density distribution of the dynemes obtained by the recognition, the number of particles of the prediction model corresponding to each dyname is allocated and the prior probability density distribution calculation of the joint motion state is performed. In the particle state update stage, the color histogram feature is selected to calculate the particle fitness, and the probability density distribution of the motion basis is corrected on the basis of re-sampling and updating the particle state, so as to achieve the human body motion pattern recognition and state tracking based on the particle filter The purpose of the joint realization. The implementation

results show that the new particle filter architecture proposed in this paper is far superior to the traditional particle filter scheme in joint tracking accuracy, and it is also comparable to the point-to-point human joint motion tracking method based on deep learning.

Key words: dyneme; particle filter; prediction model; motion tracking

1 引言

人体运动跟踪是视觉监控系统,虚拟现实,医学分析等当前众多领域^{[1][2][3]}中所研究的关键技术之一,但由于跟踪维度高、背景混乱、遮挡现象严重等多方面原因,其效果距离复杂场景中的实际应用还有一定的距离。人体运动跟踪任务是从图像中检测出人物,并对目标人物进行逐帧的位置或关节运动状态的定位与估计。将跟踪问题采用动态系统的贝叶斯估计方法^[7]进行解决是目前常用的人体运动跟踪方法,它的思路是将目标跟踪问题转换成在贝叶斯理论框架下利用状态方程预测目标状态的先验概率密度,并利用最新的测量值对先验概率密度进行修正,不断求解从而得到目标状态的后验概率密度估计。卡尔曼滤波^{[8][9]}是建立在高斯模型上用于解决非平稳随机过程内线性预测问题的方法,但是由于高斯模型是单峰模型,难以对于因遮挡所带来的非线性运动情况进行有效处理,因此适应性更高的粒子滤波算法^{[10][11][21]}逐渐成为实现人体运动贝叶斯跟踪的主要手段。

然而对于粒子滤波算法而言,粒子的数量会随着状态空间的维度增加呈指数增长。由于人体关节运动的高维度属性,因此如何提高粒子采样的有效性成为使用粒子滤波方法进行人体关节运动跟踪过程中所需解决的主要问题^[4]。一种提高粒子采样有效性的方法是在粒子权值赋值后进行粒子状态的优化,Deutscher等人^[12]提出利用模拟退火算法实现高维空间中关节模型的优化。Bray M等人^[13]使用“随机元下降”作为样本优化器来实现高维运动跟踪^[1]。另一种方法是利用人体运动模型来降低状态空间的维度,Ormonet D等人使用PCA算法在低维线性子空间中呈现人类行走,然后估计人和活动之间的自然变化^[14]。Rohr^[15]将问题的维度降低到步行周期的阶段,并使用卡尔曼滤波器进行身体位置的估计。

使用针对特定运动形式的人体关节轨迹预测模

型进行粒子状态转移空间的约束是一种基于先验空间建模的预处理思想。在特定目标的运动建模方面,Ling Guan^{[5][6]}提出了分层时空人体运动模型,将最小运动单元定义为运动基。一个运动序列解析为一组运动片段,单个运动片段即为运动基。Simonyan等人^[16]在2014年首次提出了创造性的双流网络,利用单帧RGB作为输入处理空间维度的信息和使用多帧密度光流场作为输入处理时间维度的信息并融合。为了更好的捕获时序信息,2015年Tran D等人提出了C3D卷积网络,利用三维卷积核处理视频信息^[17]。在此基础上Tran.D与Y.Lecun对时空卷积有了新的认知^{[19][20]},提出R(2+1)D网络即将2D的空间卷积和1D的时间卷积分离开来,使用残差网络进行学习^[18],与C3D对比,R(2+1)D卷积卷积更容易解决达到系统优化的目的。

在本文中,我们提出了一种基于运动基识别的人体关节运动跟踪粒子滤波估计方法。首先,在运动解析的过程中,我们先使用运动基定义运动空间,然后在运动空间中进行运动基的识别,从而得到运动基的概率密度分布,再通过概率密度分布分配粒子的多少。同时在粒子状态更新阶段,选取颜色直方图特征计算粒子适应度,在对粒子状态进行重采样更新的基础上修正运动基的概率密度分布。本文的主要贡献在于提出了一种运动状态不确定条件下的基于预测空间建模的粒子滤波器实现方案,并将其在人体关节运动跟踪应用中进行了应用验证。实现结果表明,本文提出的新型粒子滤波器架构在关节跟踪准确性上远优于传统粒子滤波器方案,并且我们还与现有的基于深度学习的点对点人体关节运动跟踪方法进行了对比,本文所提出方法在遮挡严重的情况下对于关节运动的预测性能对比深度学习方法也有一定的优势。

2 粒子滤波理论架构

2.1 粒子滤波理论基础

贝叶斯滤波是一种利用输入观测值和时域观测

基金项目:国家自然科学基金面上项目(61971383)

作者简介:张亚南(1997-),女(汉族),山东德州人,中国传媒大学硕士研究生, zhangyanan@cuc.edu.cn

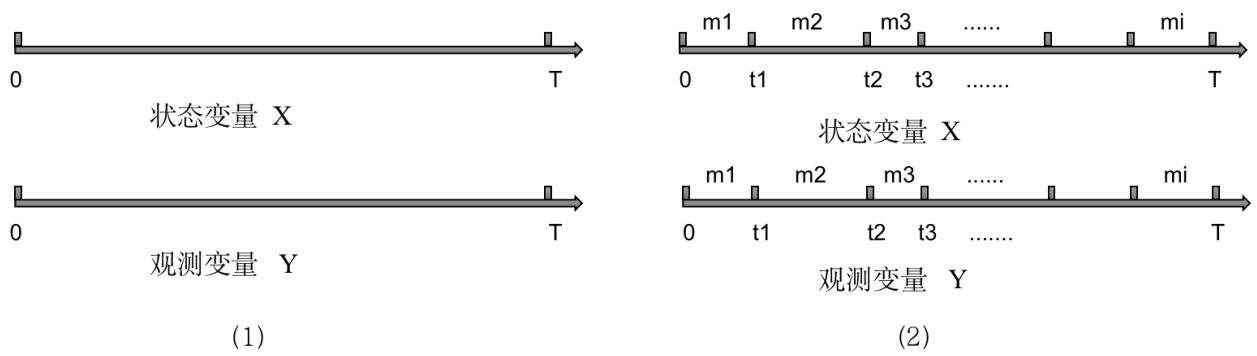


图 1 (1) 状态估计的粒子滤波时间序列 (2) 联合状态估计与模式预测的粒子滤波时间序列

模型递归地估计未知概率密度函数的方法，基于视频的人体运动跟踪的过程通常用动态系统模型来描述，主要包括状态方程和观测方程，如公式(1)(2)所示。

$$x_t = f_t(x_{t-1}, v_t) \quad (1)$$

$$y_t = h_t(x_{t-1}, w_t) \quad (2)$$

其中 x 为系统状态， y 为测量值， f, h 是状态转移函数和测量函数， v, w 为过程噪声和测量噪声，噪声都是独立同分布的。

从贝叶斯理论的观点来看，状态估计问题就是根据之前一系列的已有数据（后验知识）递推的计算出当前状态的可信度。它需要通过预测和更新两个步骤来递推计算，如公式(3)(4)所示。

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \quad (3)$$

$$p(x_t|y_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx \quad (4)$$

贝叶斯后验概率计算公式中需要对积分进行求解，为了解决公式中积分计算较难的问题，采用蒙特卡罗采样计算后验概率。粒子滤波是基于蒙特卡罗思想的贝叶斯滤波算法，基本思想是用一组样本来近似表示系统的后验概率分布，以样本均值代替积分计算，从而获得状态的最小方差估计。假设可以从后验概率中采样到 N 个样本，那么后验概率的计算可表示为：

$$p(x_t|y_{1:t}) = \sum_{i=1}^N \omega_i^t \delta(x_t - x_t^i) \approx \hat{p}(x_t|y_{1:t}) \quad (5)$$

在蒙特卡罗方法中， $f(x) = \delta(x_t - x_t^i)$ 是狄拉克函数， ω 是粒子的归一化权重，粒子权重可以通过以下公式(6)进行更新：

$$\omega_t^i \propto \omega_{t-1}^i \frac{p(y_t|x_t^{(i)})p(x_t^{(i)}|x_{t-1}^{(i)})}{q(x_{t-1}^{(i-1)}|x_t^{(i)}, y_t)} \quad (6)$$

在粒子滤波器的执行过程中，随着时间的推移，会出现粒子权值退化现象。此时一般采用粒子重采样的方法复制高权重的粒子进行解决。粒子重采样的操作使得算法向粒子权重高的地方收敛，进而得到目标状态的最优估计。

2.2 基于预测模式识别的粒子滤波框架

在粒子滤波算法中，状态方程的状态转移函数是唯一的(图 1(1)所示)，但在现实情况下，基于时间序列的运动形式不只有一种，并且状态转移函数也是不唯一的，这就需要算法有运动模型自适应调整的能力。因此我们提出了一种基于运动模式识别的粒子滤波状态估计方法(图 1(2)所示)，即从原始的求解粒子滤波算法后验概率 $p(x_t|y_{1:t})$ 演变成求解 $p(x_t, m_t|y_{1:t})$ 。对于人体运动跟踪问题来说，我们需要根据前几帧的位置来提取特征以此匹配当前运动最合适的模式，然后再对粒子状态进行估计。这种方法是先识别运动模式，再通过匹配最大似然化的模式进行状态的估计。方法如公式(7)表示，其中 t 属于时间序列的某个时间段。

$$p(x_t, m_i|y_{1:t}) = \frac{p(m_i|y_t)p(y_t|x_t)p(x_t|y_{t-1}, m_i^T)}{p(y_t|y_{1:t-1})} \quad (7)$$

其中 $p(y_t|y_{1:t-1}) = \int p(y_t|x_t)p(x_t|y_{1:t-1}, m_i^T)dx$ 为归一化常数，等式中 $p(y_t|x_t)$ 为似然函数，由观测方程决定。

3 粒子滤波理论架构

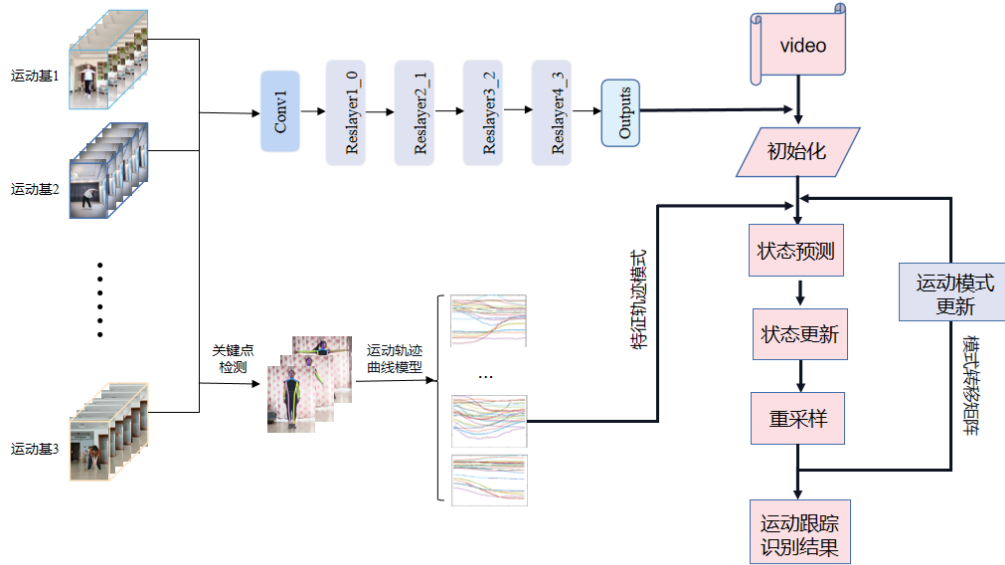


图 2 人体运动跟踪与识别的整体框架

3.1 总体架构

根据 2.2 中提出的基于预测模式识别的粒子滤波理论，本文提出一种基于运动基识别的人体关节运动跟踪粒子滤波估计方法。人体的关节结构具有多维度特性，单对粒子进行预测估计会产生粒子分散和消失现象，因此在粒子状态预测阶段，我们先对人体运动模式进行识别，运动模式被定义为运动基，通过 $R(2+1)D$ 网络识别运动基以获得预测模型中初始粒子个数的概率分布。再通过提取运动基的关节轨迹模式特征，对过程中的粒子进行行为空间的有效约束。在粒子状态更新阶段，我们采用颜色直方图特征对预测过程中的粒子进行修正，并对权值较小的粒子进行重采样。最后根据模式转移矩阵和轨迹特征样本的似然度来更新下一阶段的运动类型和运动状态估计。如图 2 所示，利用运动模式识别对粒子滤波进行估计来实现人体运动状态的跟踪。

3.2 粒子状态预测模块

预测模型是影响关节运动准确跟踪的主要因素之一。在粒子滤波预测阶段，我们需要对运动模式进行识别来获得粒子初始分布状态，再利用提取的模式特征信息对行为空间进行有效的约束。

3.2.1 基于运动基模型的预测模式空间构建

对于人体运动而言，需要对 x 、 y 平面方向的最大最小值来定义运动基开始和结束的人体动作。尽管人体能够完成的运动数量是无限的，但实现任何方向的运动方法是有限的。为简单起见，只考虑发

生在 x, y 平面方向上的运动，如表 1 所示，根据权重大小分为四种类别，可划分成 10 类运动基。

表 1 运动基类型

类别	运动基
身体旋转	全身旋转，上身旋转
质心运动	下蹲，站立，跳跃，膝盖弯曲
重心转移	腿部运动
静态姿势	上肢、下肢、前屈运动

3.2.2 运动基识别

在进行粒子滤波算法的预测和更新过程之前需要先对粒子进行初始化采样。由于人体关节具有多个自由度并且每个自由度的关节位置差异明显，如果同传统粒子滤波算法一样直接进行初始化，则会造成粒子失踪等现象。因此需要根据识别每个运动基的概率来为它们分配粒子。我们采用 $R(2+1)D$ 网络来实现运动基的识别， $R(2+1)D$ 网络在三维卷积网络的基础上增加了残差模块，可以更好地提取时空域的特征。如图 3 所示， $R(2+1)D$ 网络通过初始化五层残差卷积并根据卷积层规格设置每层中的块数，首层只包含一个卷积层，通过在空间轴上执行 2D 卷积到中间子空间，然后在时间轴上执行 1D 卷积来输出；第二层输出和第一层输出大小相同不进行下采

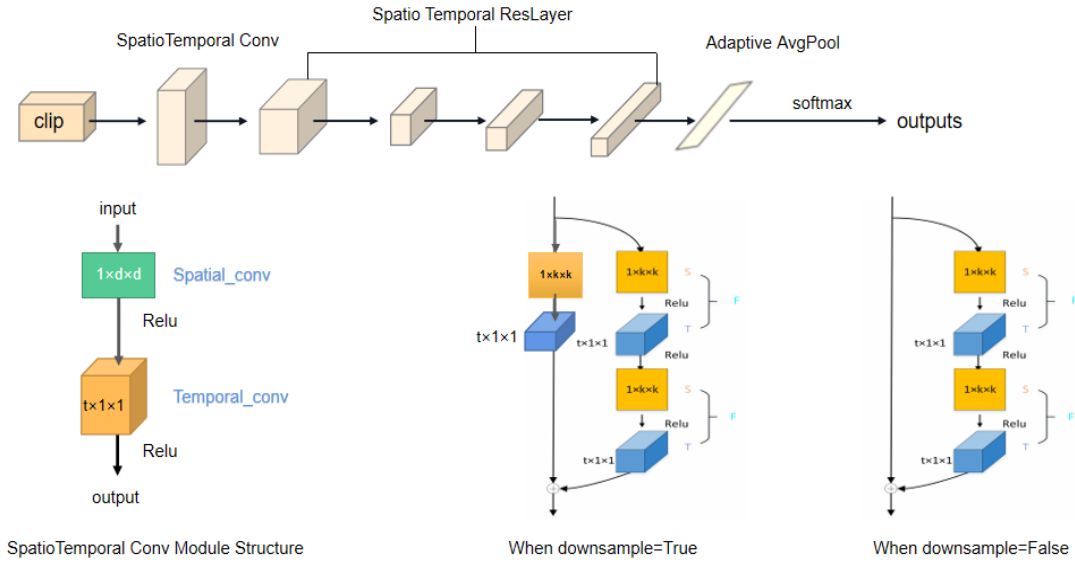


图 3 R(2+1)D网络结构

样，右侧分支直接相加；最后三层在每一层的第一个block进行下采样，对输入层进行图片的长宽压缩，这条主线部分增加了一次卷积包括空间卷积加时间卷积。并最后通过执行全局平均池化层来为批次中的每个元素生成512维向量，来形成整个残差网络特征提取器。

3.2.3 关节运动轨迹模型

(1) 人体骨架模型

人体约有几十到几百个关节，假设用所有关节点的位置信息来建立人体骨架模型，那将会给建立模型带来比较大的复杂度和计算量。所以根据人体骨架拓扑结构以及各关节在人体生理结构中的重要性，本文挑选了14个部位的关节来表示人体骨架模型，旨在简单、准确地表示人体。动作实例就是由这14个骨骼关节点的坐标随时间变化的序列组成，如图4所示。

(2) 关节运动位移信息

每个人的身高、胖瘦、形体各不相同，以至于在表征运动时的初始位置也不尽相同，不同的人表征相同动作会有一些的差异性，故我们选用帧间关节点相对位移近似消除差异性来拟合关节的运动轨迹。

定义第 i 个关节在相邻时刻的帧间位移向量：

$$\Delta s_i(t) = \{\Delta x_i(t), \Delta y_i(t)\} \quad (8)$$

第 i 个关节在模式 m 内的帧间位移表示为：

$$\Delta s_i = \{\Delta s_i(1), \Delta s_i(2), \dots, \Delta s_i(t)\} \quad (9)$$
 所以人体在单一模式内的帧间位移表示如下，其中 $N=14$ ：

$$\Delta S = \{\Delta s_1, \Delta s_2, \dots, \Delta s_N\} \quad (10)$$

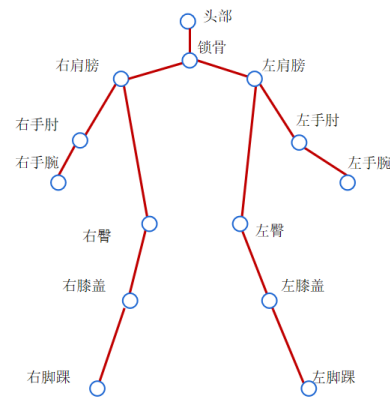


图 4 二维骨架运动模型

(3) 运动轨迹模型

在进行不同类型运动的时候，人体的骨架姿势与各关节位置都具有明显的差异性。但在进行相同类型运动的时候，不同的人的骨架姿势及关节位置的变化是类似的。因而可以通过寻找每种运动类型的骨架关节运动变化和运动轨迹的规律，来对每类运动建立运动模型。关节点的运动轨迹是由该关节点不同时刻的位置向量构成的序列，因此我们采用均值归一化的位移向量序列描述关节点的运动轨

迹:



图 5 运动基序列

$$Tr_i = \frac{(\Delta s_i(t), \dots, \Delta s_i(t+l-1))}{(\|\Delta s_i(t)\| + \dots + \|\Delta s_i(t+L-1)\|)} \quad (11)$$

根据得到的位移向量模板去拟合运动轨迹曲线模型，模式M共含有28条关节曲线，如公式所示。Tri代表关节点i的运动轨迹，I代表自由度的数量，M=1,2,...,10。

$$M = [Tr_1, Tr_2, \dots, Tr_i, \dots, Tr_I] \quad (12)$$

通过上述定义的运动基模型以及对其进行识别获取每类运动基的概率，并分配对应预测模型的粒子个数，进行关节运动状态的先验概率密度分布计算，然后根据每类运动基的运动模型对人体进行行为空间的约束，使粒子在预测阶段可以更好的实现粒子准确性表达。

3.3 粒子状态更新模块

我们利用目标的颜色直方图特征表示观测值，特征是由HSV三个颜色通道组成的110维子块直方图，量化bin为16:16:16。在更新过程中通过提取各个粒子的颜色直方图特征，引入巴氏距离来计算特征直方图分布与参考模型之间的相似性，以此来计算粒子的权重信息以及与上一帧样本的相似度。如果相似度大于0.7，直接用估计的目标位置作为下一帧的初始目标，否则结合上一帧的目标位置以及速度估计下一帧的初始位置。巴氏距离计算公式如下:

$$\rho(p, q) = \sum_{k=1}^m \sqrt{p(k)q(k)} \quad (13)$$

$$d(p, q) = \sqrt{1 - \rho(p, q)} \quad (14)$$

其中 ρ 为两直方图的Bhattacharyya系数； p 为候选目标直方图分布； q 为模板直方图分布； d 为两直

方图的Bhattacharyya距离，其值越小，表明两直方图的相似度越高；反之，两直方图相似度越低。

4 人体运动跟踪实验

4.1 数据准备

为了简化设计，假设人体运动技能序列是有限的，针对上述的运动基概念，我们构建了一个新的数据集来完成跟踪与识别实验。具体来说，我们采集了40个人在简单背景场景下的广播体操视频，视频正好包含了10类运动基，通过剪辑将视频分为若干个运动基视频，最后一共得到了4022个视频，每个视频的时间在2s左右，运动基序列如图5所示。

4.2 运动跟踪过程及实验结果

在跟踪之前需要对粒子初始化，初始化在运动跟踪过程中起着重要作用,它可以利用更多的先验信息去进行跟踪。首先需要分配预测模型中的粒子个数，我们对数据集进行分类，通过环境配置、数据预处理等步骤之后将其输入到网络中进行训练，进而得到识别结果的概率，图6是识别每类运动基的平均准确率。

我们根据已知的前几帧的数据和初始化的粒子可以匹配到有最大似然度的特征样本，因此就可以根据预测模型里的运动轨迹模型得到此刻运动基的模式特征轨迹，由此对运动基的行为空间进行约束，使粒子能够在可见的范围内采样。人体模型的大小和初始化姿势可以固定为先验信息，图7是具体运动基的运动轨迹模型。

粒子在跟踪过程中会因权重方差越来越大导致退化现象，因此我们采用SIR方法对粒子进行重采样，首先判断是否需要重采样，若需要则再判断粒子权

值是否低于重采样的阈值，重采样过程中总的粒子数保持不变。我们对所建立的运动基数据库中的测试视频进行运动跟踪实验，跟踪结果如图8和图9所示。

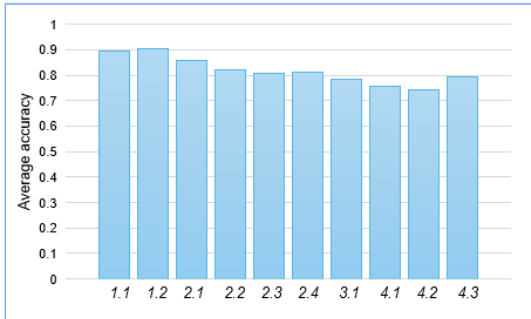


图 6 R(2+1)D 网络识别的 10 类运动基准确率



图 8 不同的人在同一运动序列下的关节跟踪

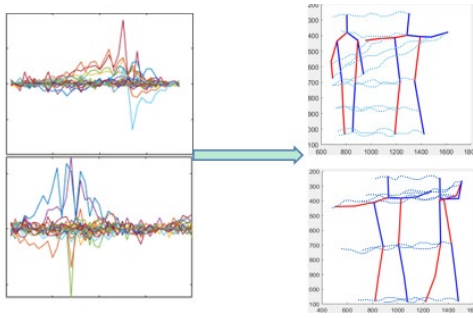


图 7 根据位移矢量模板拟合合成运动轨迹曲线模型

4.3 运动结果评估

为了更加准确地体现改进算法的跟踪结果，本文采用人体运动跟踪平均误差率对算法性能进行对比分析。利用相同的实验平台环境，粒子数为100。图10是将改进的粒子滤波器与传统粒子滤波器进行对比，可以看出随着时间的增加传统粒子滤波器的误差越来越大，且不能较好的适应运动模型的变化。图11是将我们的方法和深度学习的方法进行比较，在转变运动模式时深度学习表现出了较大的误差，即粒子的转变率较低。因此我们的方法在跟踪准确度和性能上是高且稳定的，并且从跟踪效果来看我们的预测模型克服了误差积累的弱点，能够较好的适应运动变化。

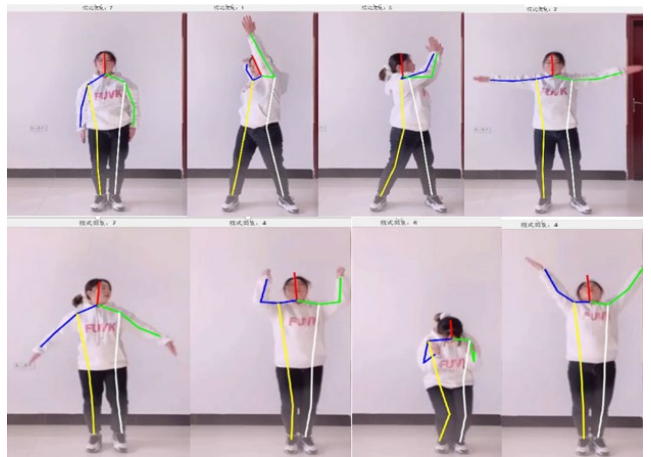


图 9 同一个人不同运动序列下的关节跟踪

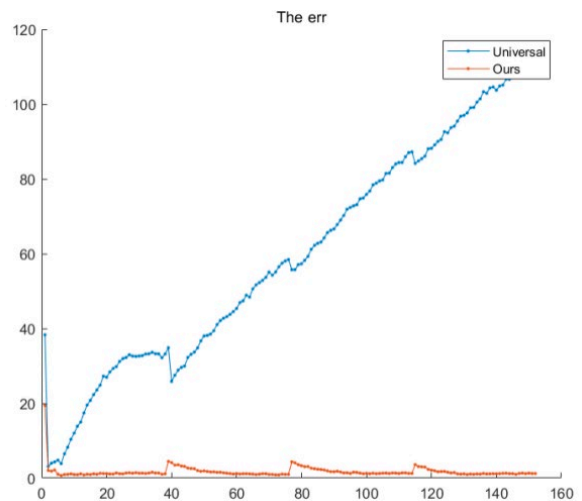


图 10 提出的方法与传统粒子滤波器的平均误差比较

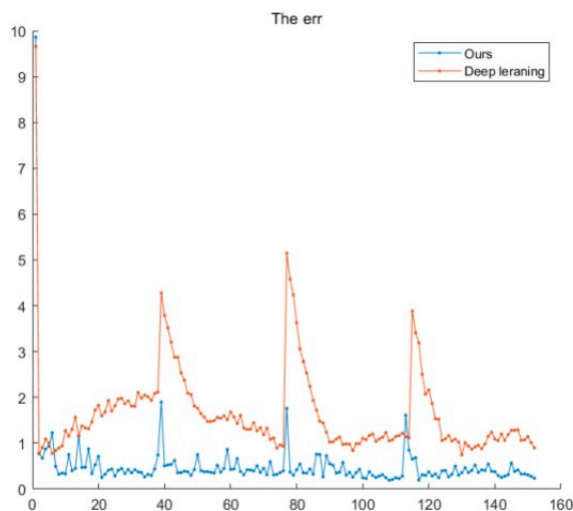


图 11 提出的方法与深度学习点对点跟踪方法的平均误差比较

5 结论

本文提出一种基于强约束预测模型的关节运动跟踪粒子滤波架构,使用运动基定义行为运动空间,然后在运动空间中进行运动基的识别,通过对预测模型中的运动基识别得到先验概率分布来约束粒子的采样,并在预测阶段利用模式轨迹特征约束人体运动,使其达到有效跟踪。本文提出的方法能够鲁棒的适应运动变化以及人体遮挡等跟踪问题,并且在跟踪性能和识别效果方面优于传统的粒子滤波算法以及点对点跟踪的深度学习方法。但是本文的算法适用于单人动作,多人交互运动时模型之间会存在影响,需要进一步改进。

参考文献

- [1] Lin S Y, Chang I C. 3D human motion tracking using progressive particle filter[C]. ISVC 2008: Advances in Visual Computing, 2008 :833-842.
- [2] Baldominos A, Saez Y, Pozo CG. An approach to physical rehabilitation using State-of-the-art virtual reality and motion tracking technologies[J]. Procedia Computer Science, 2015, 64:10-16.
- [3] Win S, Thein TLL. Real-time human motion detection, tracking and activity recognition with skeletal model[C]. IEEE Conference on Computer Applications(ICCA), Yangon, Myanmar ,2020.
- [4] Liu Z, Lee D, Sepp W. Particle filter based monocular human tracking with a 3d cardboard model and a novel deterministic resampling strategy[C]. IEEE/RISJ International Conference on Intelligent Robots and Systems(IROS), San Francisco, USA, 2011.
- [5] Green R D, Guan L. Quantifying and recognizing human movement patterns from monocular video images-part I: a new framework for modeling human motion[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2004, 14(2):179-190.
- [6] Green R D, Ling G. Quantifying and recognizing human movement patterns from monocular video images-part II: applications to biometrics[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2004, 14(2):191-198.
- [7] Alayrac J B, Sivic J, Laptev I, et al. Joint Discovery of Object States and Manipulation Actions[C]. IEEE International Conference on Computer Vision (ICCV), 2017.
- [8] Chui C K, Chen G. Sequential and Square-Root Algorithms[A].Kalman Filtering: With Real-Time Applications [M]. Springer Berlin Heidelberg, 2008: 101-113.
- [9] Song Y, Grizzle J W. The Extended Kalman Filter as a Local Asymptotic Observer for Nonlinear Discrete-Time Systems[C].American Control Conference, Chicago, USA, 1992.
- [10] Carpenter J, Clifford P, Fearnhead R. Improved particle filter for nonlinear problems[J]. IEE Proceedings Radar Sonar & Navigation, 1999, 146(1): 2-7.
- [11] Trees H LV, Bell KL. Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking[M]. Wiley-IEEE Press, 2007. ; Arulampalam MS, Maskell S, Gordon N, et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Transactions on Signal Processing, 2002,50(2):174 - 188
- [12] Deutscher J, Blake A, Reid I. Articulated body motion capture by annealed particle filtering[C].Proceedings IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Hilton Head, SC, USA, 2000.
- [13] Bray M, Koller-Meiera E, Gool LV. Smart particle filtering for high-dimensional tracking[J]. Computer Vision and Image Understanding, 2007, 106(1):116-129..
- [14] Ormoneit D, Black M J , Hastie T, et al. Representing

- cyclic human motion using functional analysis[J]. *Image & Vision Computing*, 2005, 23(14):1264-1276.
- [15] Rohr K. Human movement analysis based on explicit motion models[A]. Shah M, Jain R . *Motion-based recognition* [C]. Netherlands :Springer, 1997: 171-198.
- [16] Simonyan K, Zisserman A. Two-Stream Convolutional Networks for Action Recognition in Videos[DB/OL]. <https://arxiv.org/abs/1406.2199>, 2014.
- [17] Tran D, Bourdev L, Fergus R, et al. Learning Spatiotemporal Features with 3D Convolutional Networks[C]. *International Conference on Computer Vision*, 2015:4489-4497.
- [18] Du T, Wang H, Torresani L, et al. A closer look at spatiotemporal convolutions for action recognition[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018:6450-6459.
- [19] He K, Zhang X, Ren S , et al. Deep residual learning for image recognition[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [20] Sun L, Jia K, Yeung D Y, et al. Human action recognition using factorized spatio-temporal convolutional networks[C].*IEEE International Conference on Computer Vision*, 2015:4597-4605.
- [21] Zhao X, Lyu L, Zhang J, et al. An image-constrained particle filter for 3D human motion tracking[J]. *IEEE Access*, 2019,7:10294 - 10307.