**人体动作捕捉算法综述**

# （一）概述

人体动作捕捉算法（简称动捕算法）是人体动作捕捉系统（简称动捕系统）的重要组成部分。动捕算法是一种工程算法，动捕系统的使用范围、使用环境和硬件设备决定了算法的主要特性，动捕算法与动捕系统的关系非常紧密。动捕算法的基本通用特性是：将人体简化为一个多体系统，一般为铰接的多刚体系统；以测量或模式等方式获得输入数据；通过描述多体系统的运动来模拟人体的动作，作为输出结果。动捕算法会涉及到多种科学算法，如滤波算法、智能算法，具体包括：EKF算法，AHRS算法，IK算法等（待进一步归纳）。

# （二）动捕算法的发展

现代动捕系统的发展最早起源于20世纪70年代末的动画制作行业，这也可以看做是动捕算法的现代起源。近几年动捕算法的研究越来越广泛。

# （三）动捕算法的当前研究机构

1.光学动捕解决方案：

* VICON：新一代Shogun解决方案[1]

2.惯性动捕解决方案：

* Xsens：运动分析[2-3]

3.动捕技术与应用：

* Disney Research：混合动捕技术，动捕在动画制作中的应用[4]
* IKinema：全身IK技术，基于VIVIE追踪器的全身动捕[5]
* University of Kaisersalutern：传感器融合技术，惯性动捕误差模型，自动校准[6-9]
* TU Berlin：即穿即用的惯性动捕技术，运动分析[10-12]
* 中科院计算所：人体运动仿真，光学动捕技术[13-16]

4.基于图像的动捕：

* University of Toronto：基于多目图像和物理模型的动捕[17-18]
* CMU：单相机图像人体特征点识别[19-20]
* TAMU/中科院计算所：单深度相机全身动捕[21-22]
* MPI Informatik: 单彩色相机、双鱼眼相机、单帧图像三维姿态估计[23-26]
* 清华：单彩色相机三维姿态估计[27-28]
* 浙江大学动画与感知课题组 http://zju-capg.org/

5.运动合成：

* University of Edinburgh：基于机器学习的运动合成技术[29-30]
* 中科院计算所：人体运动仿真，个性化运动合成[31]
* UBC：基于物理模型的运动仿真与合成技术[32-33]

# （四）动捕算法的当前研究热点

1.低成本/快捷动捕方案

* VICON Shogun - 快捷光学动捕方案：在线快速校准
* Disney Research - 基于少量传感器的混合动捕
* IKinema - 使用6或8个VIVE追踪器的全身动捕
* Xsens - 探索只用5个IMU的动捕
* University of Kaisersalutern - 自动校准

2.基于物理模型的动捕和运动合成

* Disney Research - 三维扫描建模，逆动力学，接触计算，基于物理模型的姿态滤波
* UBC - 基于物理仿真的运动合成技术
* 中科院计算所 - 基于逆动力学和运动数据库的实时接触力计算与运动合成

3.基于动作数据库和机器学习方法的动捕和运动合成方法

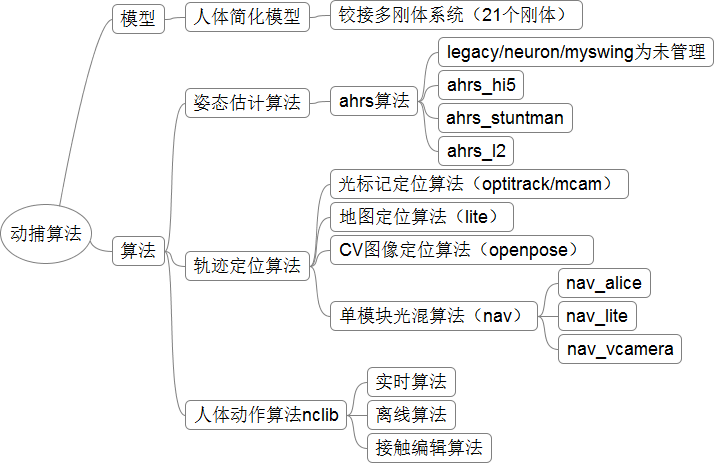
* Disney Research - 运动数据库作为先验动作知识参与姿态优化计算
* University of Edinburgh - 基于运动数据库、机器学习的运动合成技术
* MPI Informatik - 基于深度学习的单帧图像三维姿态估计

4.基于图像的动作捕捉

* CMU - 单相机图像人体特征点识别OpenPose
* TAMU/中科院计算所 - 单深度相机全身动捕，使用Kinect，基于先验运动数据库
* MPI Informatik - 单彩色相机动作捕捉

# （五）我司的动捕算法

我司的动捕算法可以总结为一个模型（人体简化模型），三个部分（姿态估计算法，轨迹定位算法，人体动作算法）。



（1）人体简化模型

将人体简化为一个多体系统，一般为铰接的多刚体系统。我司将模型主要简化为21个刚体骨骼，另外有时也会用到57个、23个或167个刚体骨骼模型。

（2）姿态估计算法

姿态估计算法指通过测量设备的测量数据计算出测量设备的姿态的算法，目前主要包括ahrs算法。公司最初的姿态估计算法，即用于legacy、neuron和myswing的姿态估计算法没有纳入ahrs算法体系。目前ahrs算法体系包括ahrs\_hi5、ahrs\_stuntman（L2无线）和ahrs\_l2（L2有线）三个。

（3）轨迹定位算法

轨迹定位算法指通过测量设备的测量数据计算出测量设备的位置的算法，目前包括光标记定位算法（optitrack/mcam），地图定位算法（lite），CV图像定位算法（openpose），单模块光混算法（nav）。

光标记定位算法（optitrack），目前还只能说说是一种方法，因为现阶段是直接使用了光学设备厂商提供的结果。自主的算法正在MCAM项目研发中。

地图定位算法（lite），在alicelite项目中研发的inside-out算法，已经可以在项目中应用，目前在整理归纳中。

CV图像定位算法（openpose），在开源项目openpose算法中发现，该算法经过改造可以实现轨迹定位的目的，该算法正在研究中。

单模块光混算法（nav），将基本的定位数据（例如光点坐标）输入到惯性测量单元，可以得到更高精度的定位结果。该算法同时也可以输出单模块的姿态。该算法已经在alice，lite等项目中应用。其算法体系包括nav\_alice，nav\_lite，nav\_vcamera。

（4）人体动作算法

人体动作算法指通过姿态估计算法和轨迹定位算法计算后，提供了测量单元的姿态和位置，将这些输入计算为人体的多体系统的运动。根据输入数据的不同和输入数据实时性等要求，可以分为实时算法、离线算法、接触编辑算法。实时算法是人体动作算法的基础，其算法主要处理的问题包括：

1）没有相应的输入数据，例如惯性设备的丢包会导致没有姿态数据，只有惯性设备会导致没有位置数据，光学设备的丢光会导致没有位置数据。数据带宽（？60k，90k，180k都会有不同的算法，目前是70k左右）。

2）相应的测量输入数据不正确，这里主要指出现了明显噪声，例如光学数据在大多时是准确的，可是突然出现了短时噪声。

3）相应的测量输入数据不精确，例如目前惯性设备的姿态、加速度和角速度的精度都不太高。

4）测量数据向人体数据转换的问题，例如设备穿戴后，在运动过程中设备发生的相对骨骼的运动，又如人体尺寸测量不准确。

5）人体运动的固有约束，包括基本物理约束，环境约束，及关节活动度约束，例如不会有高频抖动，不会出现跳变，脚部接触地不能穿地，走路时脚部接触地不能滑动，不能反关节等。



# （六）动捕算法的下一步工作方案

我司动捕算法发展的总体策略为成熟一代、研发一代、规划一代。三代的动捕算法分别为惯性动捕算法、光混动捕算法、智能动捕算法，下面对三代算法的工作给以介绍。

（1）完善惯性动捕算法（第一代算法）

惯性动捕算法是公司的传统优势算法，在过去的一年多已经将惯性动捕算法完全进行了梳理和重构。算法的各部分已经模块化，每一个模块也实现了方法思路、公式推导、代码实现的一致性。在未来可以逐步提升各模块的效果和性能，例如EKF算法、接触判断算法、接触修正算法、IK算法、坐标系一致性等等。这些功能的优化可以结合光混动捕算法的发展一并考虑。

（2）发展光混动捕算法（第二代算法）

光混动捕系统是公司重点发展的新一代动捕系统，光混动捕算法也是目前最重要的工作。光学设备的引入后可以提供准确的位置，解决惯性动捕系统对轨迹定位的不足。经过8个月的努力，已经完成了具有一定实用价值的光混动捕算法，但同时也发现了许多问题，需要在未来发展完善。

目前光混动捕算法发现的主要问题包括：

1）正在研发多种光学定位算法、以降低成本。

2）匹配错误的光学数据会导致明显错误。

3）丢光导致的光惯切换，进而影响了动作的连续性。

4）光学数据的高精度与惯性数据的低精度的不协调、引发多种问题。

针对以上这些问题，计划的主要处理方法如下：

1）研发新的定位算法：optitrack/mcam，lite，openpose等。

2）研发错误数据的识别、剔除和修复的方法。

3）优化光惯数据融合策略，改善光惯融合及切换时动作的连续性。

4）基于光学数据对惯性数据建立误差修正模型，用于改善切换不连续或者精度不协调问题。

5）改进校准方法，改善精度不协调问题和测量不准确的问题。

6）离线算法：接触状态修正，光点轨迹定位算法等。

（3）智能动捕算法（暂定，第三代动捕算法）

随着动作数据库的逐渐丰富和深度学习算法的发展，基于动作数据库的动捕算法和基于机器学习的动捕算法，已经有了越来越广泛的应用。这种使用少量测试数据和大量积累数据的方法，在游戏、VR等领域的应用已经崭露头角，同时动作数据库的积累也是公司的宝贵资产，并可以应用到各个领域中去。

对于智能动捕算法，目前还处于规划中，其主要任务是研究目前的国际的先进方法，并规划动捕数据的积累方法。

# 附. 参考文献

[1] VICON Shogun: <https://www.vicon.com/products/software/shogun>

[2] T. Cloete “Benchmark inertial measurement for gait analysis” Thesis, Stellenbosch University, 2009

[3] F. Wouda, etc. “Estimation of full-body poses using only five inertial sensors: an eager or lazy learning approach?” Sensors 2016, 16, 2138

[4] S. Andrew, etc. “Real-time physics-based motion capture with sparse sensors,” Conference for Visual Media Production (CVMP) 2016, Disney Research Project website: <https://www.disneyresearch.com/publication/real-time-physics-based-motion-capture-with-sparse-sensors/>

[5] IKinema Orion: <https://ikinema.com/orion>

[6] wearHealth website: <http://agw.cs.uni-kl.de/en/research/>

[7] B. Taetz, etc. “Towards Self-Calibrating Inertial Body Motion Capture,” 2016 19th International Conference on [Information Fusion (FUSION)](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=7518993)

[[8]](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=7518993) G. Bleser, etc. “A generic approach to inertial tracking of arbitrary kinematic chains,” 8th International Conference on Body Area Networks, 2013

[9] M. Miezal, etc. “On Inertial Body Tracking in the Presence of Model Calibration Errors,” Sensors 2016, 16, 1132

[10] Prof. Thomas Seel’s website: <http://www.control.tu-berlin.de/User:Thomas_Seel>

[11] T. Seel, etc. “IMU-based joint angle measurement,” Sensors 2014, 14, 6891-6909

[12] P. Muller, etc. “Alignment-free, self-calibrating elbow angles measurement using inertia sensors,” IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 21, NO. 2, MARCH 2017

[13] 中科院计算所人体运动仿真组：<http://humanmotion.ict.ac.cn/index.html>

[14] S. Xia, etc. “A survey on human performance capture and animation,” JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 2017

[15] S. Xia, etc. “Toward accurate realtime marker labeling for live optical motion capture,” Computer Graphics International 2017

[16] X. Lv, etc. “Data-driven inverse dynamics for human motion, ” SIGGRAPH Asia 2016

[17] Prof. David Fleet’s website: <http://www.cs.toronto.edu/~fleet/research/projects.html>

[18] D. Fleet, “Motion model for people tracking,” Guide to Visual Analysis of Humans: Looking at People, 2011

[19] Z. Cao, etc. “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” CVPR 2017

[20] OpenPose website: <https://github.com/CMU-Perceptual-Computing-Lab/openpose>

[21] X. Wei, etc. “Accurate Realtime Full-body Motion Capture Using a Single Depth Camera,” ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH Asia 2012

[22] 苏乐等, ”基于局部姿态先验的深度图像3D人体运动捕获方法,” 软件学报，2016

[23] MPI Informatik’s website: <http://gvv.mpi-inf.mpg.de/index.html>

[24] D. Mehta, etc. “VNect: Real-time 3D human pose estimation with a single RGB camera,” ACM Transactions on Graphics ([SIGGRAPH 2017](http://s2017.siggraph.org/)),

project website: <http://gvv.mpi-inf.mpg.de/projects/VNect/>

[25] H. Rhodin, etc. “EgoCap: Egocentric Marker-less Motion Capture with Two Fisheye Cameras,” SIGGRAPH Asia 2016

[26] F. Bogo, etc. “Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image,” ECCV 2016

[27] Prof. Yebin Liu’s website: <http://www.liuyebin.com/>

[28] Y. Wang, etc. “Outdoor Markerless Motion Capture with Sparse Handheld Video Cameras,” [IEEE Trans Vis Comput Graph.](https://www.ncbi.nlm.nih.gov/pubmed/28422685) 2017

[29] D. Holden, etc. “Phase-Functioned Neural Networks for Character Control,” SIGGRAPH 2017, Website: <http://theorangeduck.com/page/phase-functioned-neural-networks-character-control>

[30] D. Holden, etc. “A deep learning framework for character motion synthesis and editing,” SIGGRAPH 16

[31] S. Xia, etc. “Realtime style transfer for unlabeled heterogeneous human motion,” ACM Transactions on Graphics (SIGGRAPH), 2015

[32] Prof. Michiel van de Panne’s website: <http://www.cs.ubc.ca/~van/>

[33] L. Liu, etc. “Guided learning of control graphs for physics-based characters,” Transactions on Graphics (SIGGRAPH 2016),

Project website: <http://www.cs.ubc.ca/~van/papers/2016-TOG-controlGraphs/index.html>