

硕 士 研 究 生 读 书 报 告



题目Capturing Detailed Deformations of Moving Human Bodies

作者姓名 李秋惠

作者学号 22151233

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 2021 年 月 日

Capturing Detailed Deformations of Moving Human Bodies

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: Qilei Li

By

Qiuhui Li

Zhejiang University, P.R. China

2021

# 摘要

文章提出了一个基于新型运动捕捉套装和神经网络的人体动作捕捉方法。新型动作捕捉套装的图案由黑白棋盘和白色棋盘格角上的两个字母代码构成。通过多摄像头系统采集2D图像，由一系列神经网络处理图像以定位角点并识别上面的字母代码，由此捕获到演员身体上1000多个角点并自动为每个点指定唯一的标签。点的位置和唯一标签仅从单个2D输入图像推断，使得捕获的点轨迹包含输入图像的所有细节，包括呼吸、肌肉收缩和肌肉变形引起的运动。系统对身体的伸展和自我遮挡具有鲁棒性，而且仅依赖于标准RGB或单色传感器、全被动照明和被动防护服，易于复制、部署和使用。文章最后展示了对各种人体姿势的高度精确捕捉，包括挑战性动作，如瑜伽、体操或在地面上滚动。

**关键词**：人体动画 运动捕捉

# Abstract

This paper presents a human motion capture method based on a new motion capture suit and neural network. The pattern of the new motion capture suit is composed of a black-and-white chessboard and two letter codes on white points (corners). The 2D image is collected through the multi camera system, and a series of neural networks process the image to locate the corner points and identify the above letter code, so as to capture more than 1000 corner points on the actor's body and automatically assign a unique label to each point. The position and unique label of the point are inferred only from a single 2D input image, so that the captured point trajectory contains all the details of the input image, including the movement caused by respiration, muscle contraction and muscle deformation. The system is robust to body stretching and self shielding, and only depends on standard RGB or monochrome sensors, all passive lighting and passive protective clothing, which is easy to copy, deploy and use. Finally, the article shows the highly accurate capture of various human postures, including challenging movements, such as yoga, gymnastics or rolling on the ground.

**Keywords**: human animation, motion capture, skin deformation

# 1. 引言

在大多数现实世界的图像中，人体会被衣服遮挡，使得精确的身体动作捕捉变得困难当需要精确测量时，传统的工作采用附在动作捕捉服上或粘在皮肤上的反光标记或者在皮肤上绘制彩色图案来进行标记，存在一定的局限性，不能捕捉因呼吸而导致的肌肉变形或运动的细节。

为了捕捉具有高细节的运动物体，Bogo等人提出了DFAUST方法：首先从几何角度将模板人体模型注册到3D扫描，然后使用皮肤上的彩色图案通过光流获得高精度的时间对应。DFAUST方法依赖于初始几何配准和时间跟踪，这可能会受到误差累积的影响，并可能在更具挑战性的姿势或快速运动中导致不正确的局部极小值。另外，DFAUST数据集虽然包含各种高度详细的人体动画，但仅限于直立站立类型的运动。

文章提出了一个基于新型运动捕捉套装和神经网络的人体动作捕捉方法。新型动作捕捉套装的图案由黑白棋盘和白色棋盘格角上的两个字母代码构成，能够精确定位身体表面的点，并允许通过识别代码直接标记角落，这是DFAUST方法做不到的。文章还展示了更广泛的各种运动的捕捉，包括体操练习、瑜伽姿势或在地面上滚动。

# 2. 内容介绍

## 2.1 新型运动捕捉套装

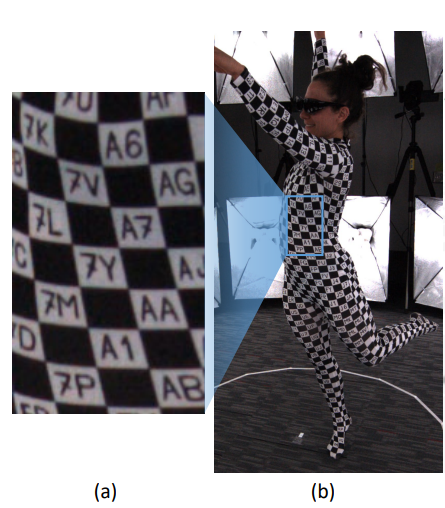


图2-1 新型运动捕捉套装

如图2-1所示，新型的运动捕捉套装带有特殊的基准标记，由棋盘形的角组成，用于精确的局部定位，两个字母代码用于唯一的标记。即使衣服有明显的拉伸，棋盘格状的角仍然是棋盘格状的，因此新型的运动捕捉套装对显著拉伸以及投影失真具有鲁棒性。

## 2.2 数据采集

通过多摄像机系统中各个摄像机之间的对应关系来建立角点标签建立了时间和空间上的对应关系，由此将二维角点位置三角化为三维点。多摄像机设置包含16个标准（RGB）摄像机，这些摄像机围绕捕获体积排列成一个圆圈。每台摄像机以30 FPS的速度以原始格式拍摄4000×2160张图像。围绕着捕获体积，放置了32个软盒，它们产生均匀的漫射光。通过在其中一个摄像机上的传统的校准棋盘来校准摄像机，使用光束调整进一步细化相机参数和世界坐标中的3D棋盘格角位置。

然而，由于自遮挡和摄像机数量有限，三维重建（三角化）点不可避免地会缺失观测，如图2-2所示。为了填补（插值）这些缺失的观测值，首先拟合STAR模型，然后捕获的健美操式动作序列中的点轨迹，并为每个演员细化它，然后使用这种改进的身体模型来插值静止姿势中缺失的角点，从而生成没有任何孔的最终网格。

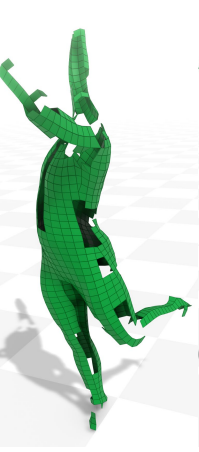


图2-2 观测缺失

## 2.3 图像处理

文章将三个卷积神经网络与几何算法相结合进行图像处理。卷积神经网络（CNNs）的第一个是一个角点检测器，它定位输入图像（4000×2160）中的所有角点。为了纠正白色方块内部两个字母代码的失真，将四个角元组连接成四边形（quads），并应用单应变换来纠正拉伸和投影失真。RejectorNet用于质量控制和处理清晰、垂直的字母代码。其余的代码被传递给RecogNet，RecogNet读取两个字母代码中的字符。

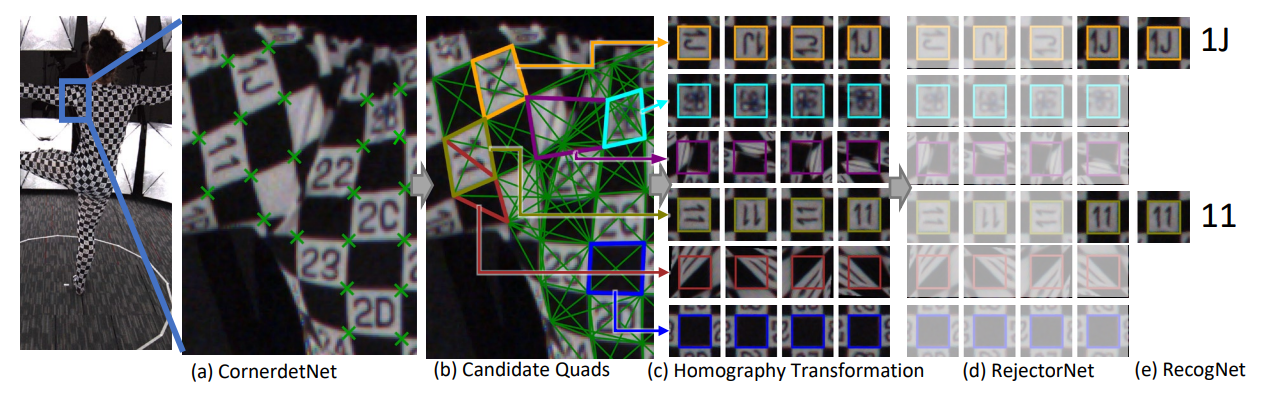


图2-3 图像处理模块结构

# 3. 创新点及应用

（1）文章提出了一种新的方法来测量每一帧的三维标记位置，并自动推断相应的标记标签。这是在没有任何人体形状或运动学先验知识的情况下实现的，也就是说，文章的数据是“原始测量”，不受任何类型的建模或归纳偏见的影响。

（2）文章介绍了一种新型的基准标记和运动捕捉套装，它仅使用局部图像块实现标记定位和唯一标记。由于不使用时间跟踪，该套装对标记的遮挡具有鲁棒性，并且还可以并行处理，因为每个帧都可以独立处理。

（3）文章中的所有结果都是通过使用16台商品RGB摄像机和被动照明的实验多摄像机系统获得的。与需要大量的硬件投资和工程专业知识的系统相比，该系统很容易从廉价的现成部件构建，易于复制、部署和使用。

（4）文章提出的方法可用于对各种人体姿势的高度精确捕捉，包括挑战性动作，如瑜伽、体操或在地面上滚动。

# 4. 局限性

（1）必须穿一套特殊的动作捕捉服。

（2）如果对裸露的皮肤进行成像，当前的相机分辨率（4000×2160）无法从皮肤获得足够的细节。

（3）关注点在与身体的运动捕捉，忽略了脸部和手部的运动捕捉。

（4）文章的数据处理是离线的，将来可以创建一个实时的版本。

参考文献

[1]. Chen H, Park H, Macit K, et al. Capturing Detailed Deformations of Moving Human Bodies[J]. ACM Transactions on Graphics (TOG), 2021, 40(4): 1-18.