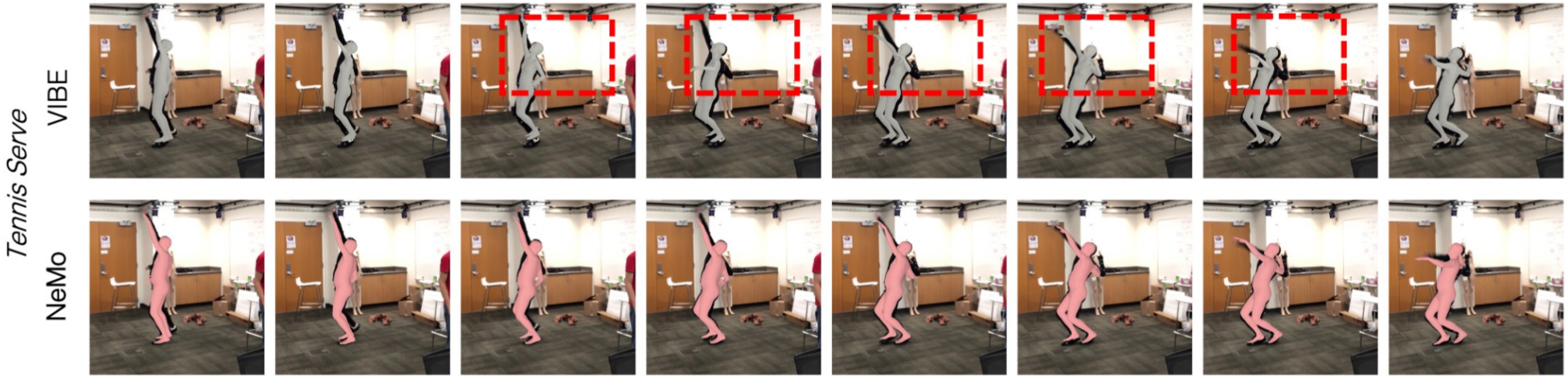
NeMo: 3D Neural Motion Fields from Multiple Video Instances of the Same Action

Kuan-Chieh (Jackson) Wang, Zhenzhen Weng, Maria Xenochristou, Joao Pedro Araujo, Jeffrey Gu,
C. Karen Liu, Serena Yeung

Stanford University

CVPR 2023 Highlight 🔦 (Session: THU-PM-145)

https://sites.google.com/view/nemo-neural-motion-field



JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

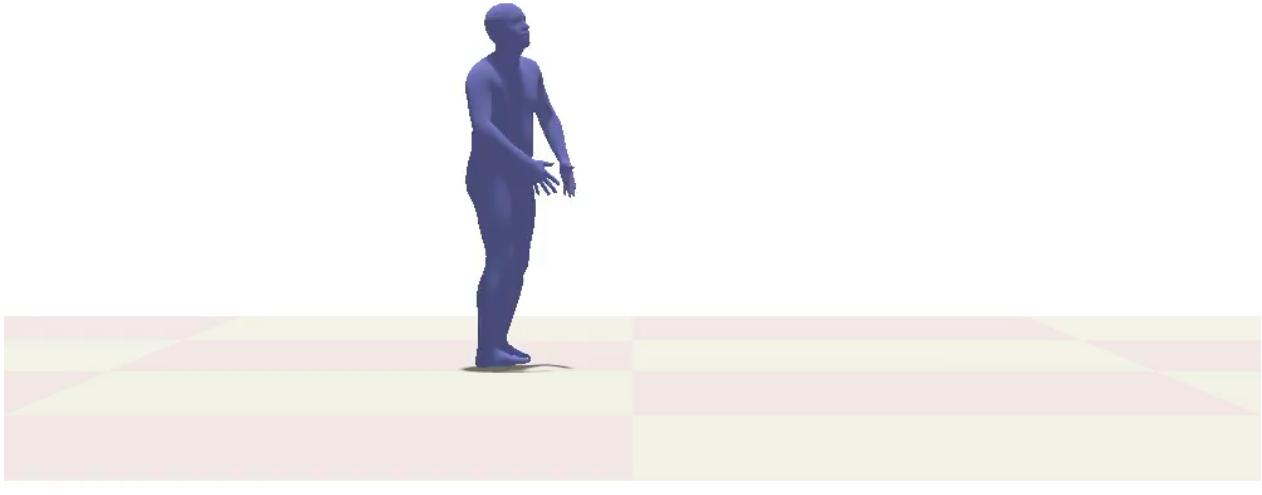
Motivation

"Goal": Get more ***3D human motion*** data.

"Insight": There's an ***abundance of video*** data.

➔ Let's use them as ***sources for human motion data***.

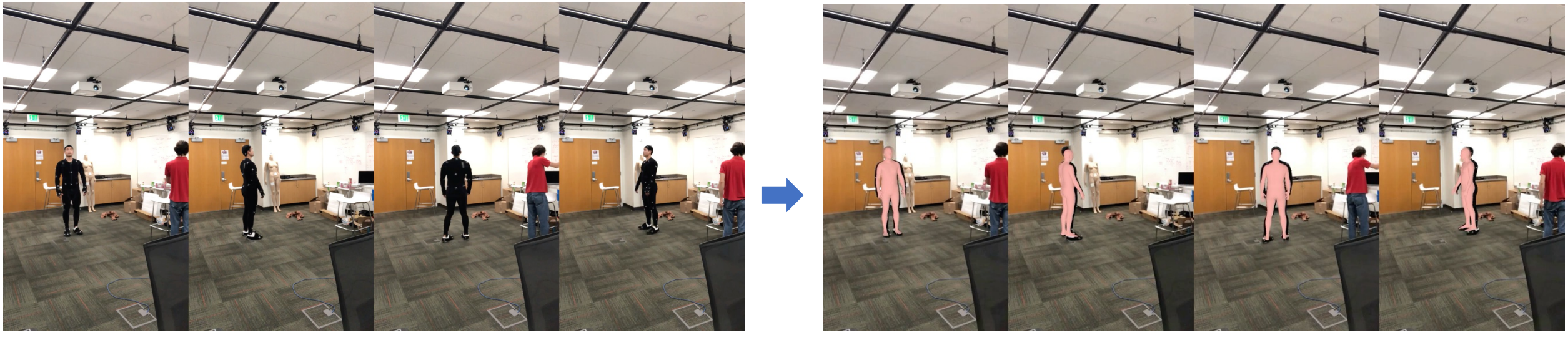
Penn Action Dataset [Zhang et al., 2013 (ICCV)]



Formulation: Multiple Videos of the Same Action

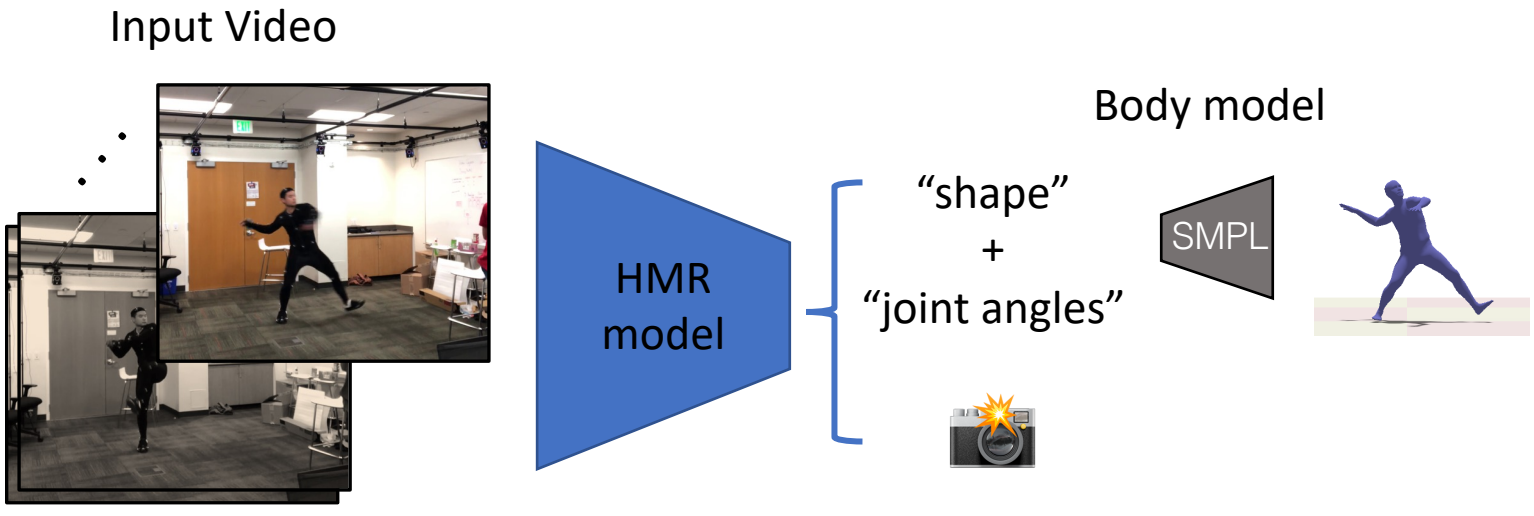
Input: a set of videos (unsynchronized) of the same action.

Output: 3D global motion for each video.



Background: 3D Human Motion Recovery

3D Human Mesh (Motion) Recovery



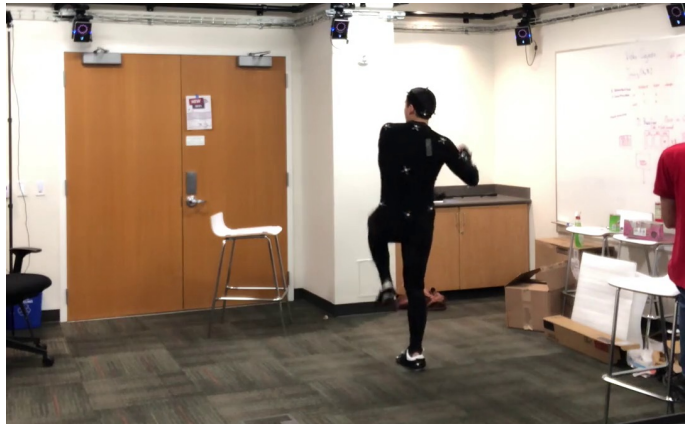
Background: 3D Human Motion Recovery

3D Human Mesh (Motion) Recovery

Input Video

HMR model

"2D to 3D" causes ambiguity



Occluded left arm

A better view

Bridging the Gap between **MoCap** & HMR

Multi-view can resolve ambiguity

H3.6M [Ionescu et al., 2013]

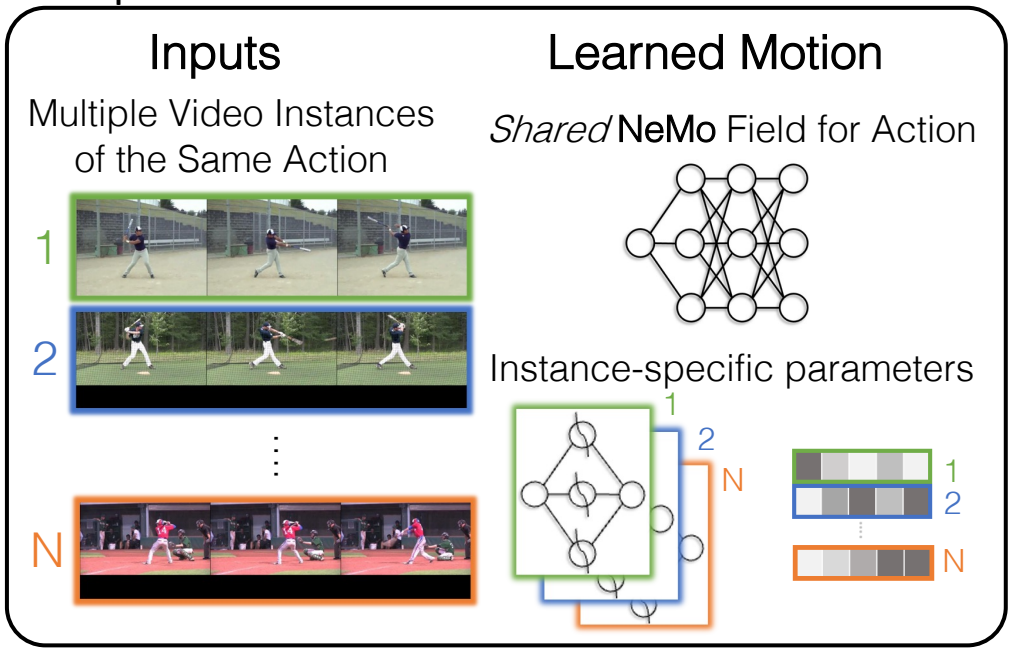
3D data is scarce, and limited in diversity

"Relaxed" **multi-view data**

Penn Action Dataset [Zhang et al., 2013 (ICCV)]

NeMo – Neural Motion Field

Setup

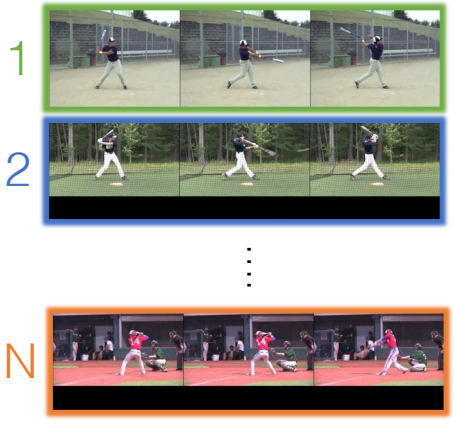


NeMo – Neural Motion Field

Setup

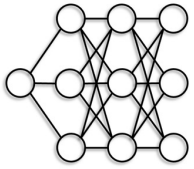
Inputs

Multiple Video Instances of the Same Action

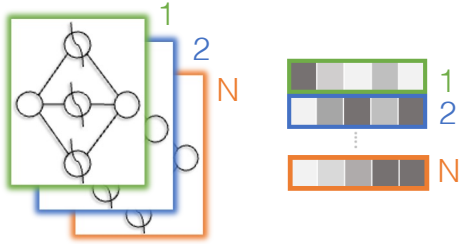


Learned Motion

Shared NeMo Field for Action

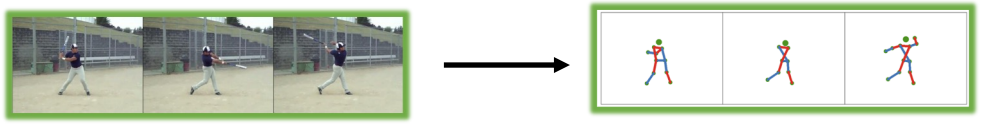


Instance-specific parameters



NeMo optimization for 1 sequence

E.g., optimizing for video 1



NeMo – Neural Motion Field

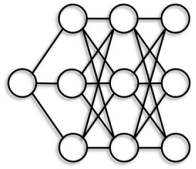
Setup

Inputs

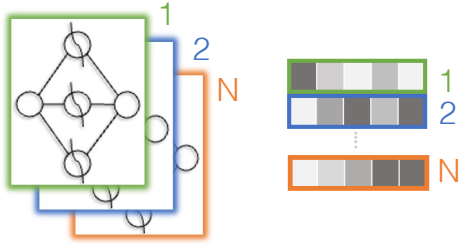
Multiple Video Instances of the Same Action

Learned Motion

Shared NeMo Field for Action



Instance-specific parameters



NeMo optimization for 1 sequence

E.g., optimizing for video 1

Init. 2D/3D Pred.

Reprojection Loss

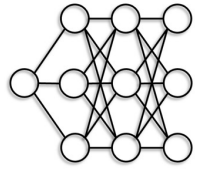
ρ

Learned Camera

R^1, t^1

NeMo

$\phi^1_{1:T(n)}$

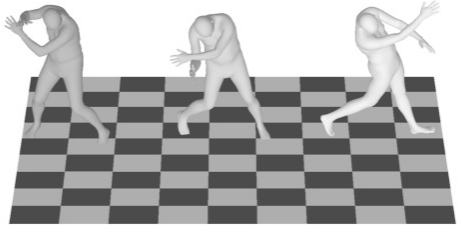


f_{θ}

$x^1_{1:T(n)}$

$\theta^1_{1:T(n)}$

SMPL



NeMo – Neural Motion Field

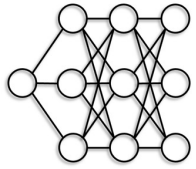
Setup

Inputs

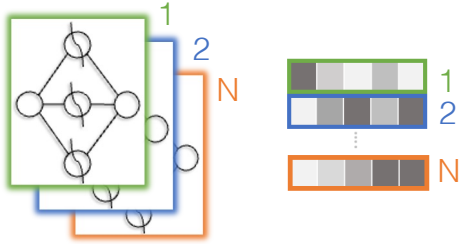
Multiple Video Instances of the Same Action

Learned Motion

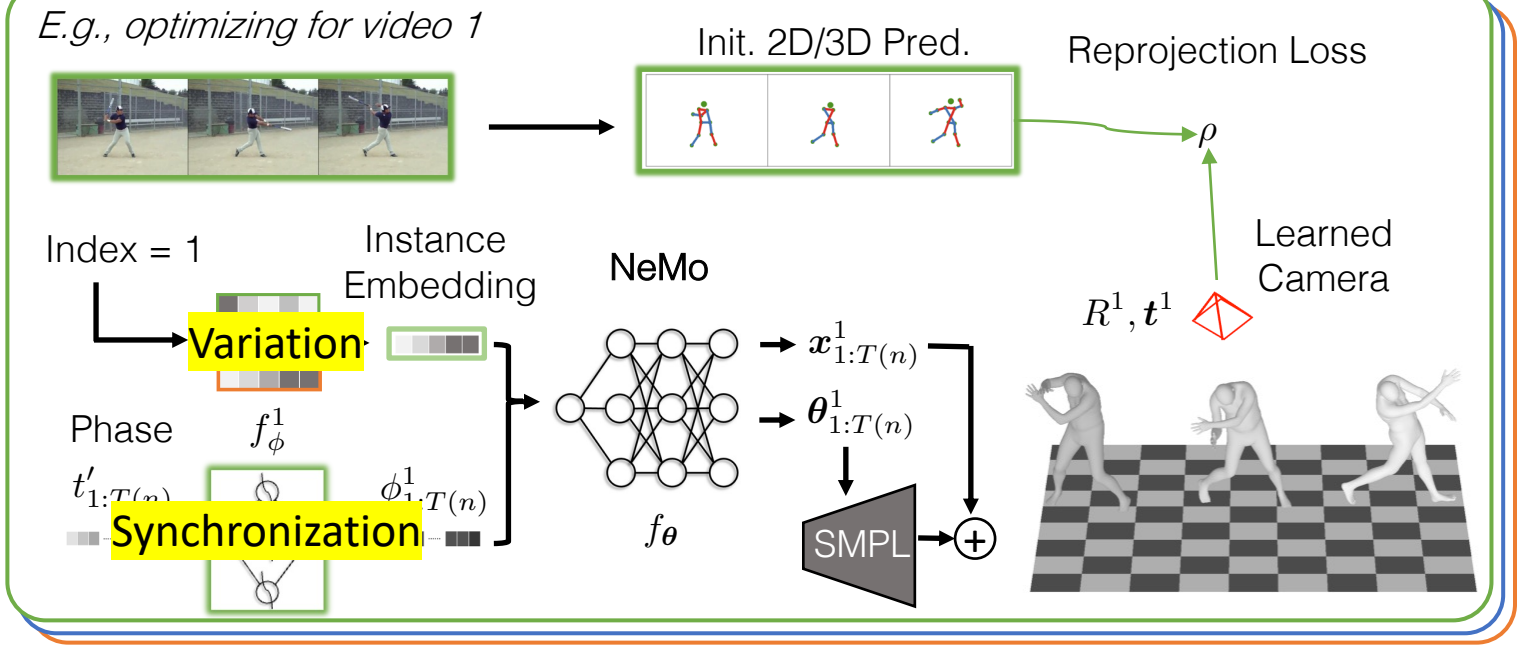
Shared NeMo Field for Action



Instance-specific parameters



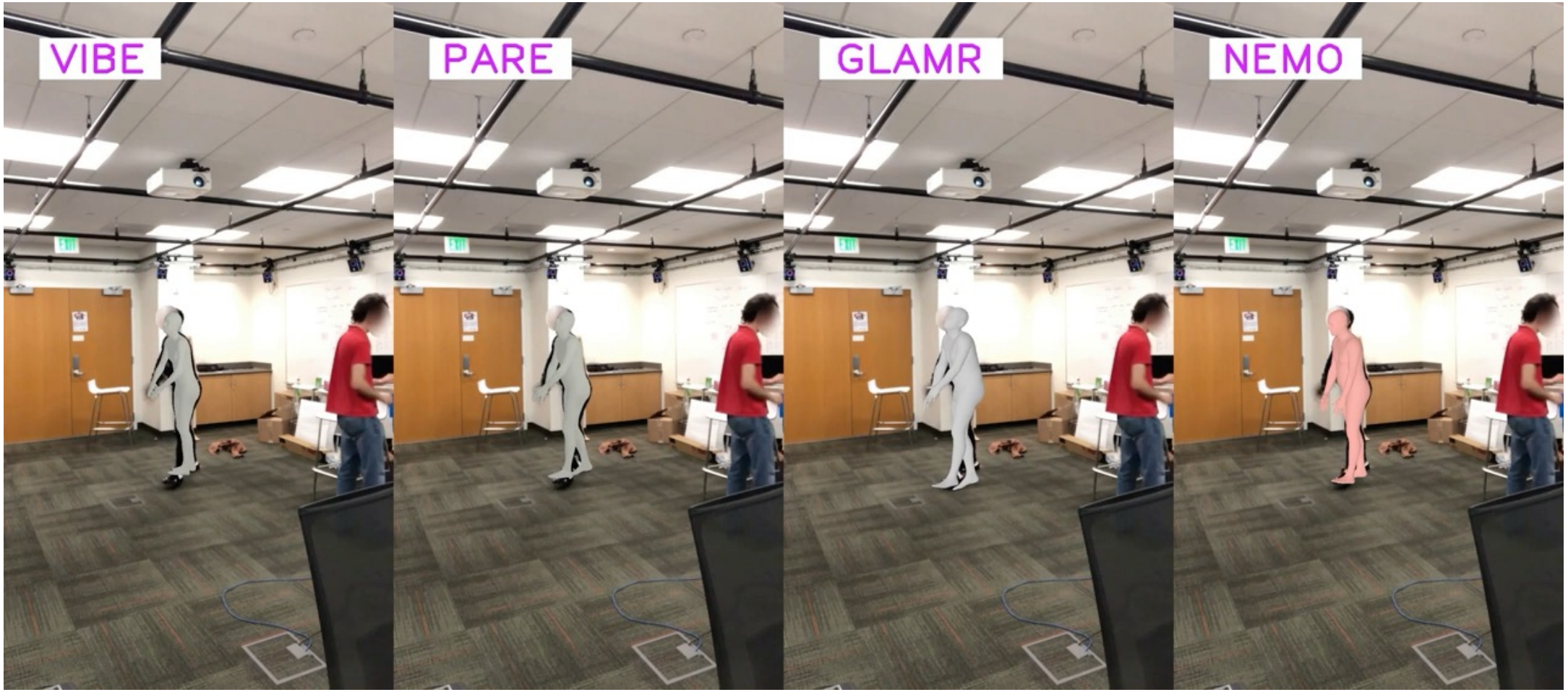
NeMo optimization done jointly for all instances



Results

Validated on our NeMo-MoCap dataset

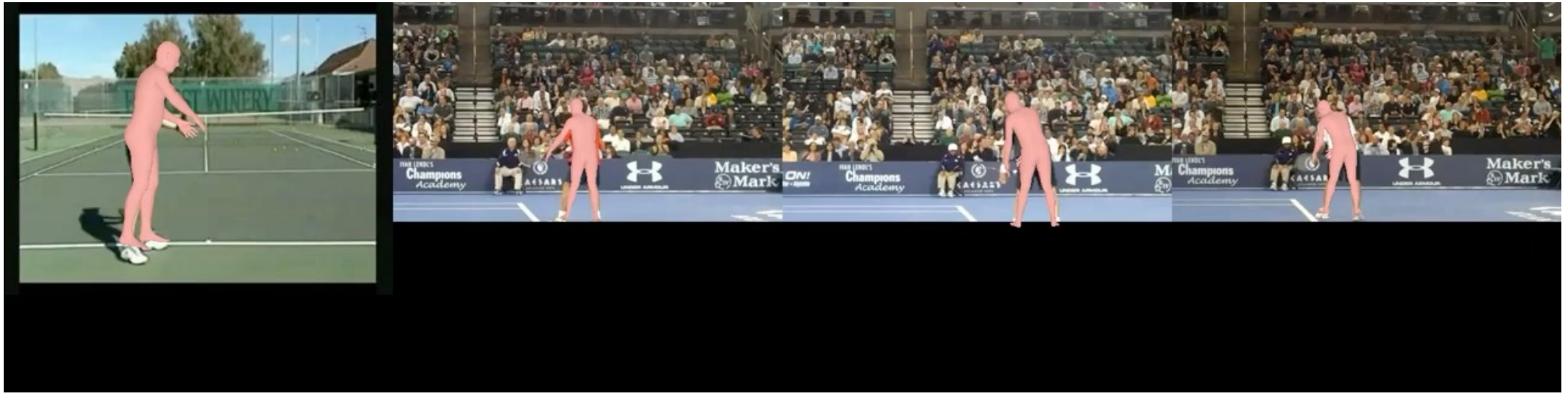
Comparison to other monocular video-based methods



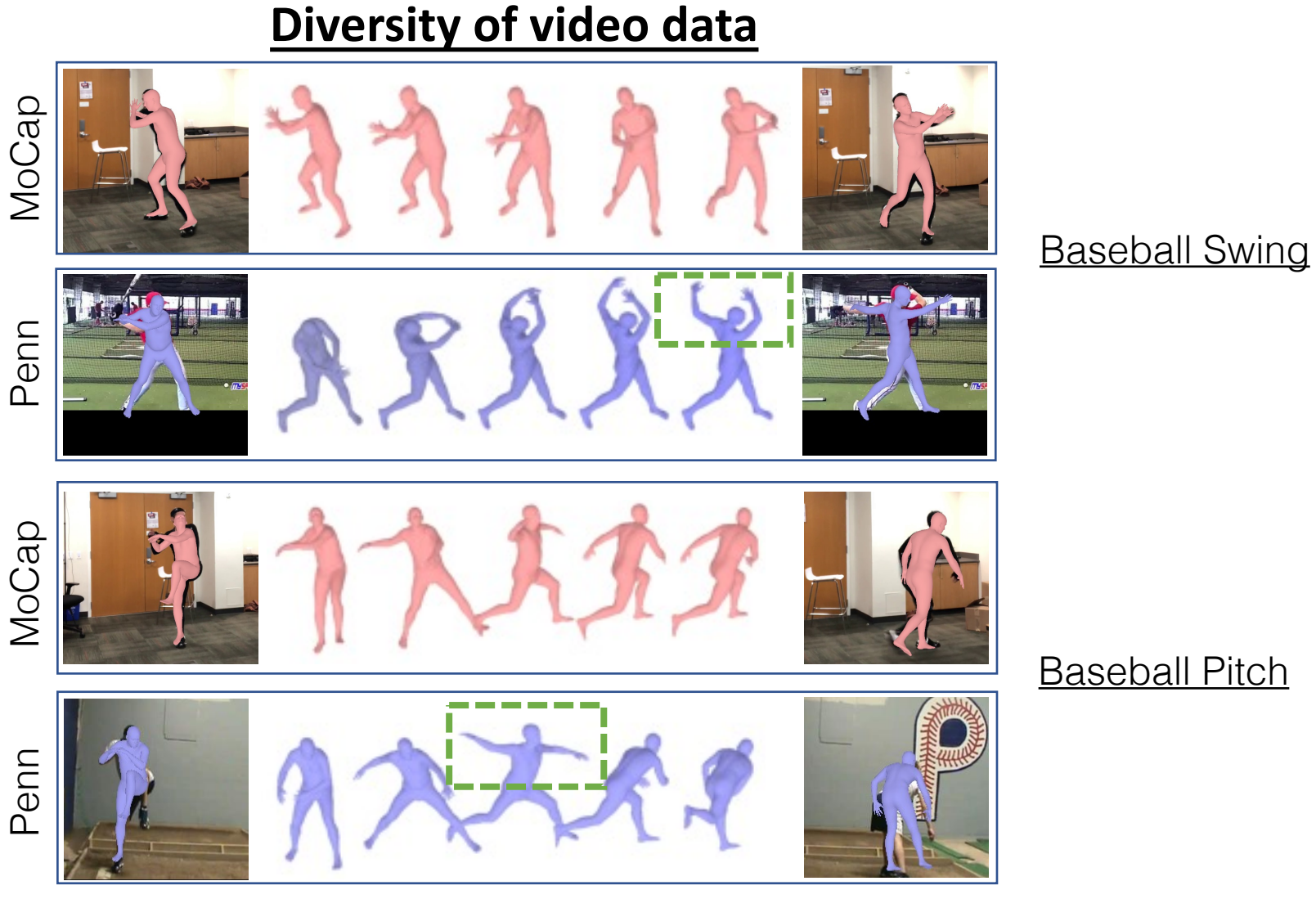
Results

Demo on existing video dataset

Penn Action



Results


Diversity of video data
MoCap
Penn
MoCap
Penn
Baseball Swing
Baseball Pitch

Conclusion

- We propose a method NeMo for 3D human motion reconstruction by ***leveraging information shared across different videos***.
  - *Code and dataset released*

- We hope to facilitate the direction of ***sourcing 3D human data from existing videos***.

Wu Tsai Human Performance Alliance

MARVL Medical AI and ComputeR Vision Lab

The Movement Lab

JUNE 18-22, 2023

CVPR

VANCOUVER, CANADA

Session: Thusday (PM) -145

https://sites.google.com/view/nemo-neural-motion-field

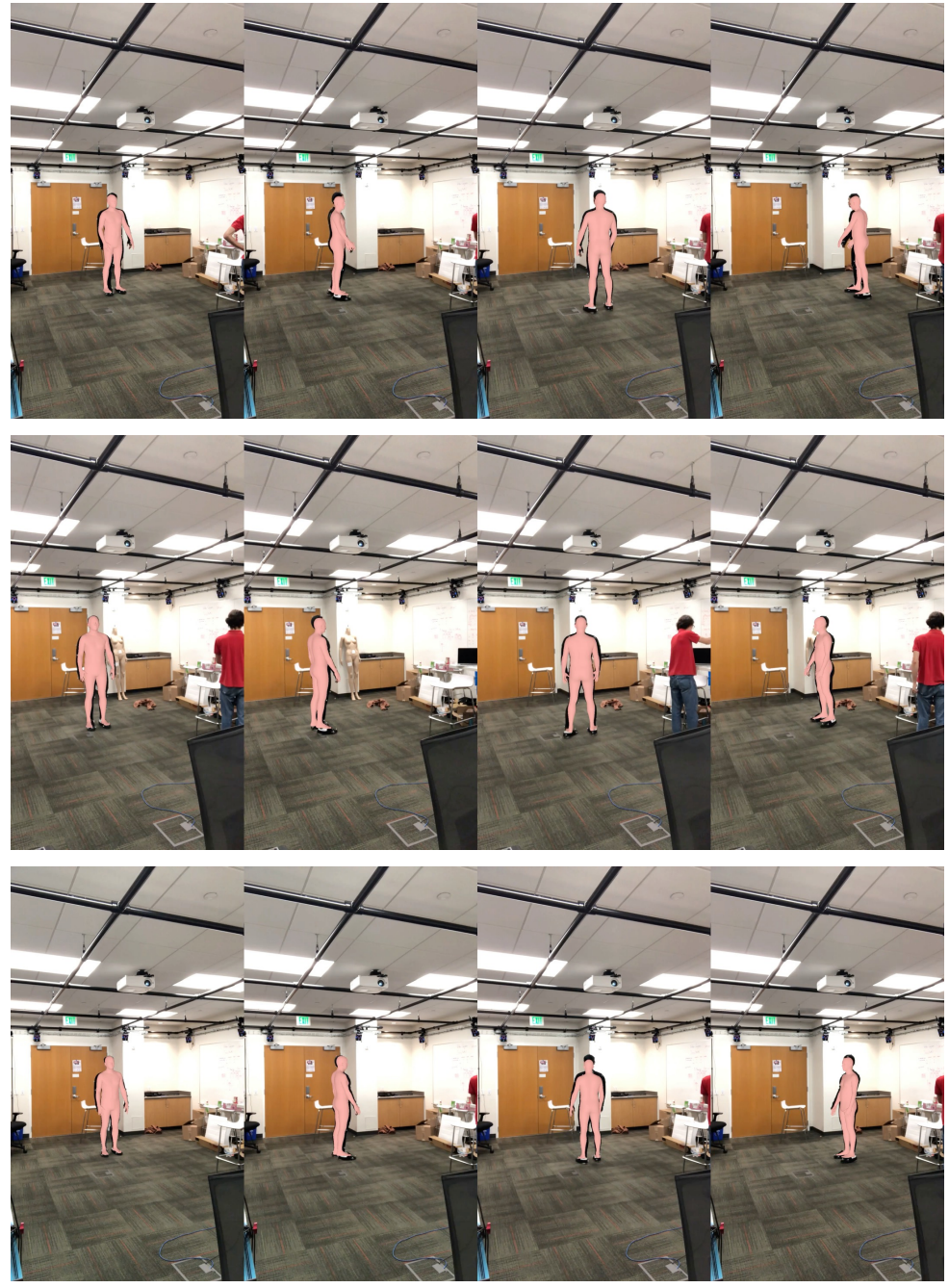
README.md

NeMo [CVPR2023 Highlight]

project page | arxiv 2212.13660

This repo contains the official PyTorch implementation of our paper:

NeMo: 3D Neural Motion Fields from Multiple Video Instances of the Same Action

by Kuan-Chieh (Jackson) Wang, Zhenzhen Weng, Maria Xenochristou, Joao Pedro Araujo, Jeffrey Gu, C. Karen Liu, Serena Yeung

(Project Page | ArXiv | Data)