Motion Capture via Video

Final Project UPDATE! by Joseph Aguilar

The Problem: Motion Capture

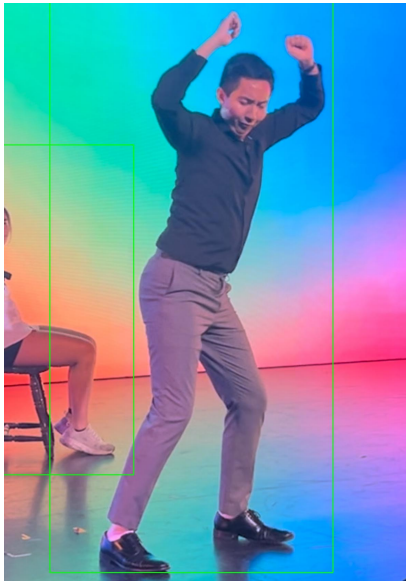
- Relatively unexplored in the context of live theater.
- Requires expensive, high-quality suits and many sensor cameras for high-quality results.
- How to reconcile this?

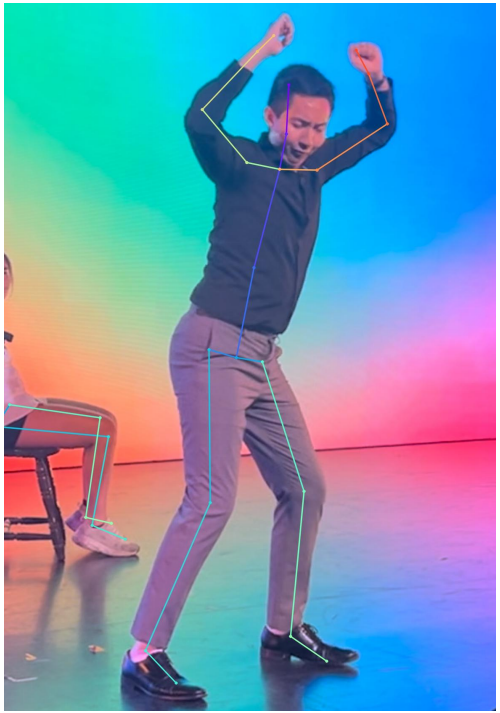
My INITIAL Proposed Approach

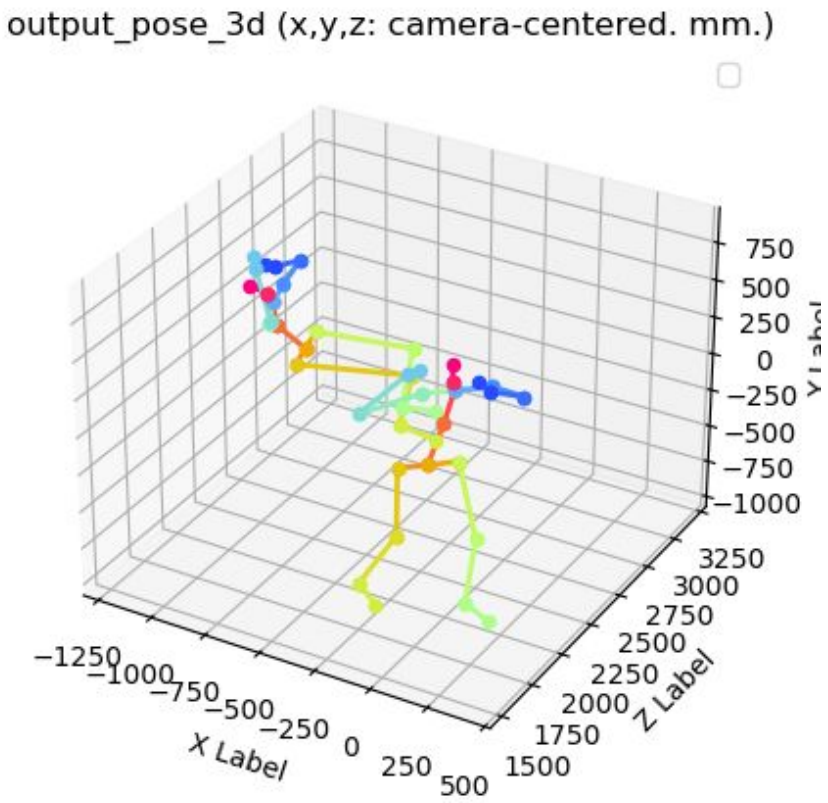
- A convolutional neural network that takes image input and outputs 3D vector coordinates of a person (like motion capture!)
- Trained on data from Human3.6M
- A continuation of the work done in these papers [1] [2] [3].
    - Reworked for use on video, focusing:
        - Speed
        - Algorithms for efficiency

How the Plan has Changed

1. Switched from making my own model to working with pre-trained snapshots because...
2. Went from working from 1 CNN... to 3 CNNs...
3. A CNN for...
   a. Pose matching
   b. Root detecting
   c. Bounding box tracking
4. Developed a framework for custom input, using 3 CNNs.







The Progress So Far!

Where to Go from Here

1. SPEED!
    a. GPUs/CUDA
    b. Frame sampling

2. REALISM!
    a. “Tweening”
    b. Frame sampling

3. Make cooler video?

```
✓ 73m 21.1s

460
>>> Using GPU: 0
>>> Using GPU: 0
```

Thank you!