



NII

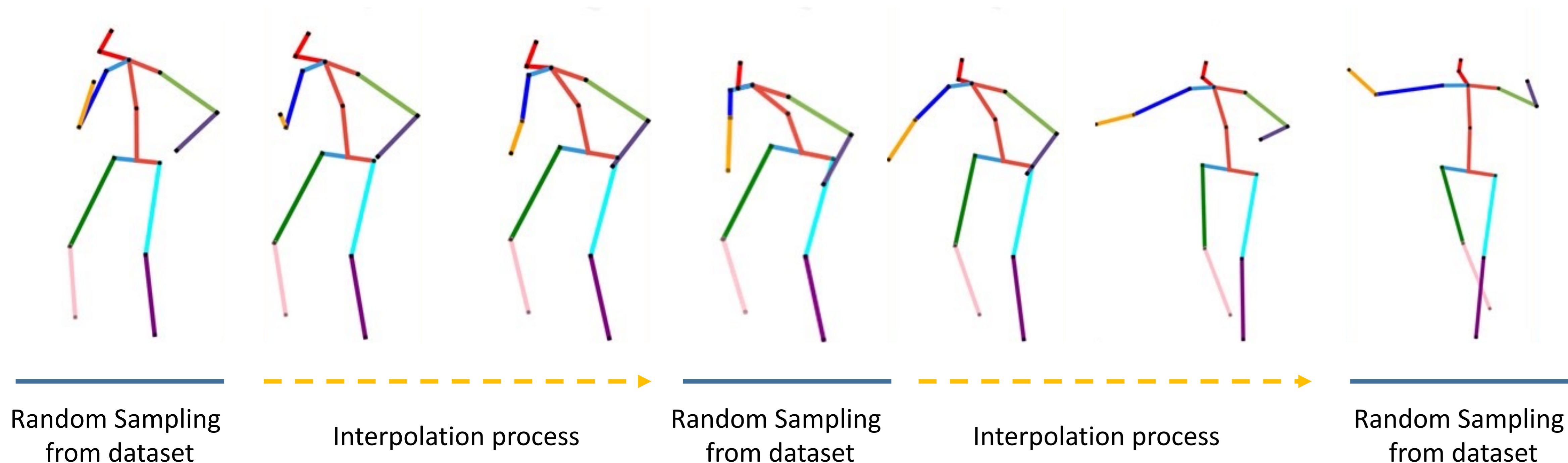
Learning Monocular 3D Human Pose Estimation with Skeletal Interpolation

Ziyi Chen, Akihiro Sugimoto, Shang-Hong Lai

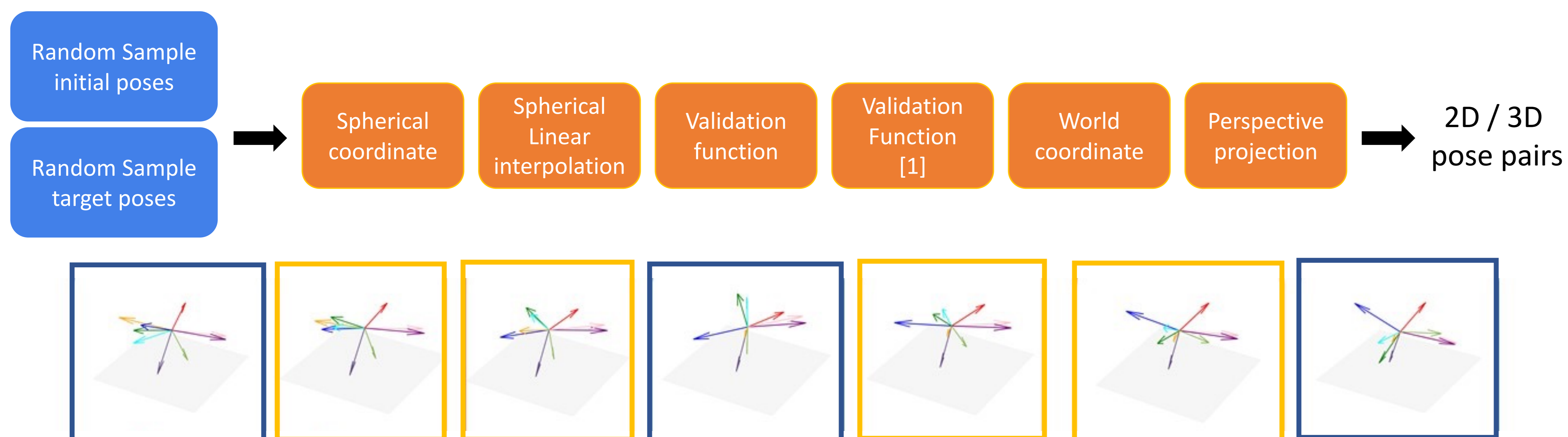
Introduction:

- 3D information is difficult to acquire from single RGB camera
 - One 2D pose in image space corresponds to many varied 3D poses
- 3D Pose are usually captured by Motion Capture System in a highly controlled environment
 - Algorithms developed with these data suffers from poor generalization

Augmentation Process:

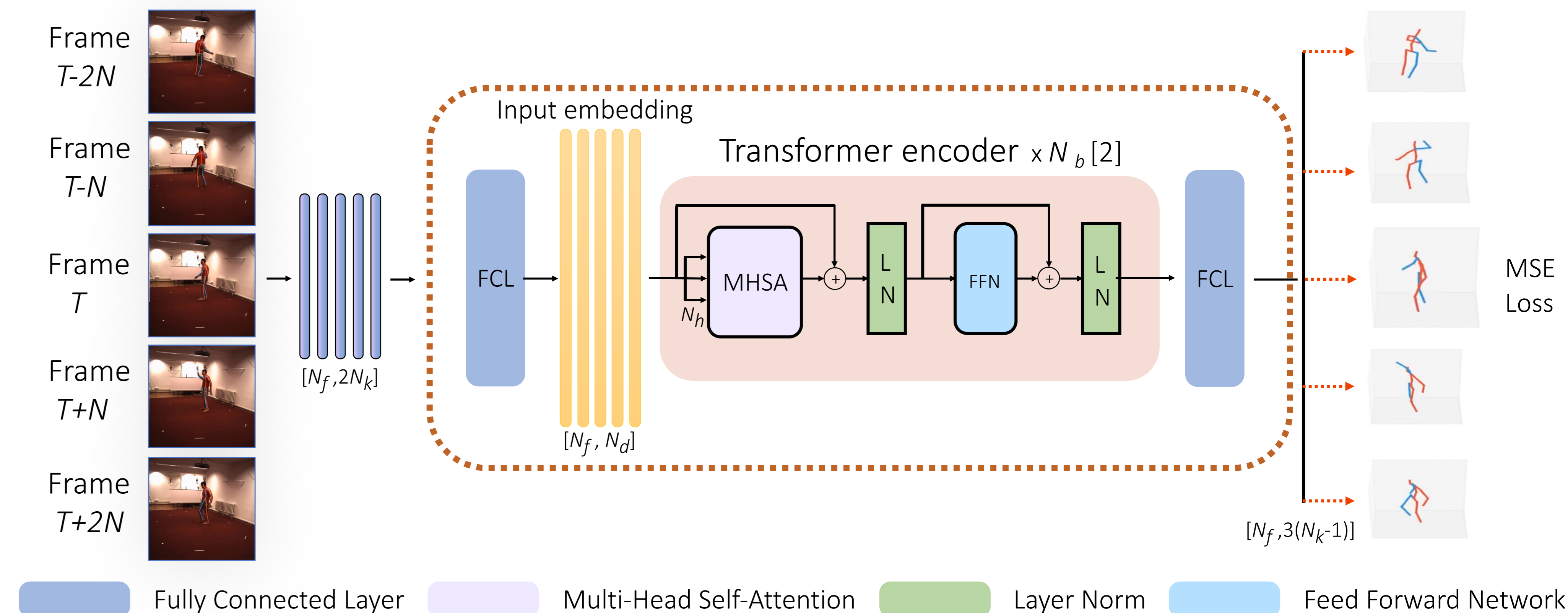


Skeletal Interpolation:



$$\vec{\beta}_{itpl}(T) = \frac{\sin[(1-T)\Omega]}{\sin\Omega} \vec{\beta}_s + \frac{\sin[T\Omega]}{\sin\Omega} \vec{\beta}_t$$

Lifting Network:

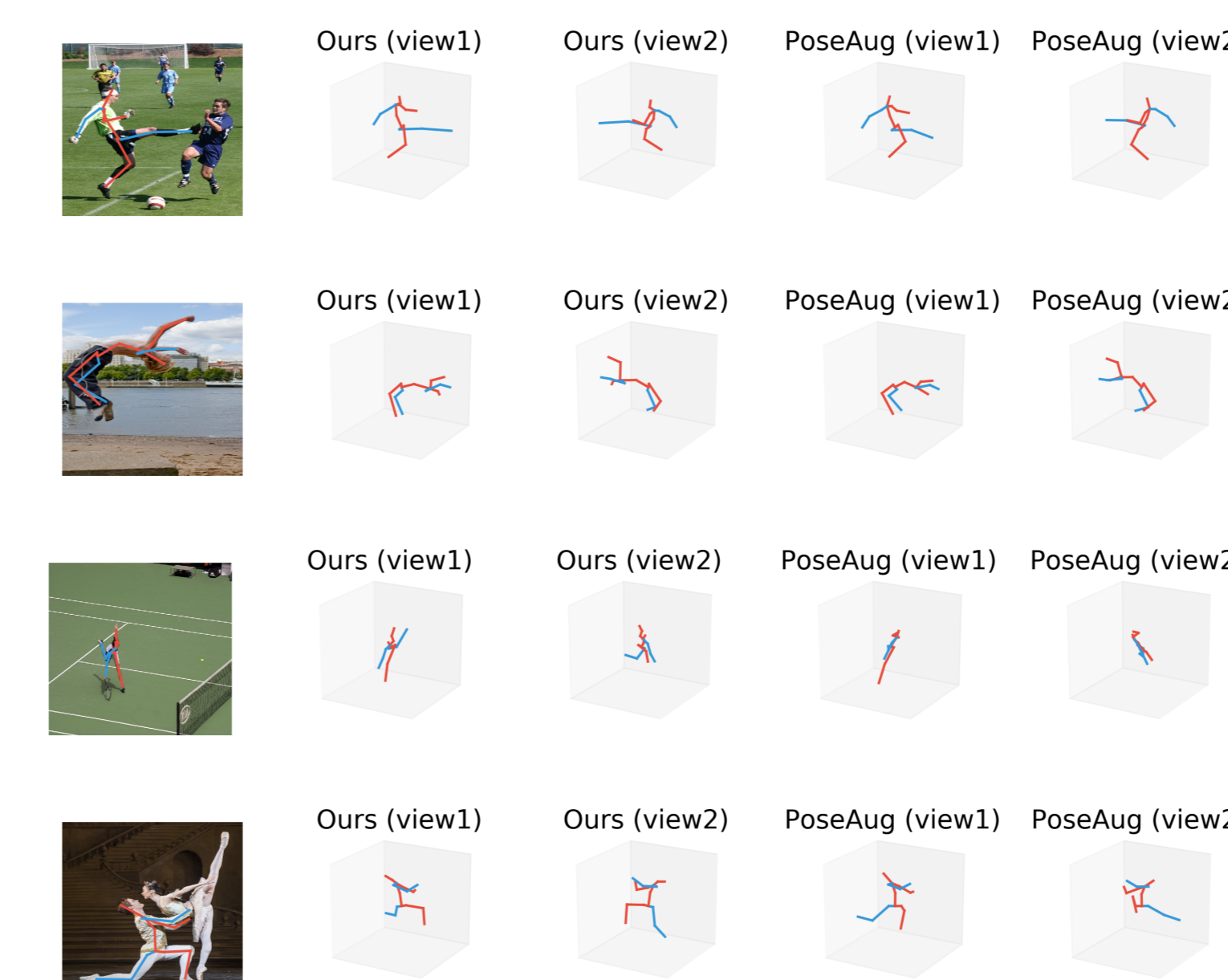


Experiments:

Dim. N_d	Head N_b	Block N_b	Frames N_f	MPJPE	PCK	AUC
256	2	6	5	94.9	81.6	46.6
512	2	6	5	93.4	81.6	48.2
768	2	6	5	95.8	81.3	46.7
512	2	6	5	93.4	81.6	48.2
512	4	6	5	97.2	80.8	46.1
512	8	6	5	94.3	81.8	47.4
512	2	2	5	99.9	78.9	45.2
512	2	4	5	96.5	80.3	46.3
512	2	6	5	93.4	81.6	48.2
512	2	6	3	95.6	80.8	46.6
512	2	6	5	93.4	81.6	48.2
512	2	6	9	97.3	81.1	45.9

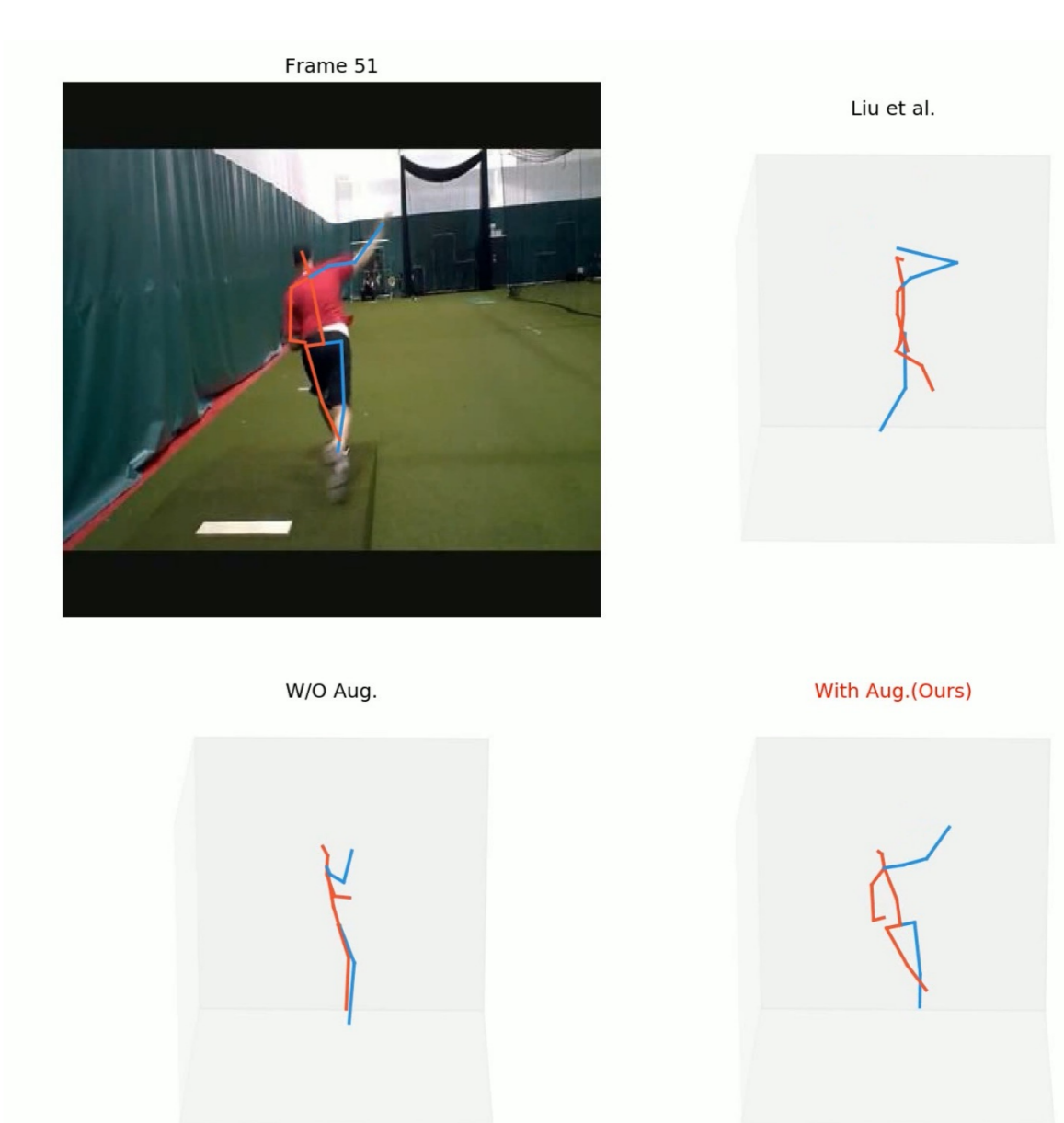
Method	MPJPE	PCK	AUC
Wang <i>et al.</i> (ICCVw'19) [14]	—	71.2	33.8
Habibie <i>et al.</i> (CVPR'19) [15]	127.0	69.6	35.5
Yang <i>et al.</i> (CVPR'18) [16]	—	69.0	32.0
Zeng <i>et al.</i> (ECCV'20) [17]	—	77.6	43.8
Kanazawa <i>et al.</i> (CVPR'18) [18]	113.2	77.1	40.7
Chen <i>et al.</i> (CVPR'19) [19]	—	61.4	29.4
Gong <i>et al.</i> (CVPR'21) [20] †	73.0	88.6	57.3
Ours (with InterAug) *	93.4	81.6	48.2
Li <i>et al.</i> (CVPR'20) [9] *	99.7	81.2	46.1
Ours (without InterAug) *	102.6	79.2	44.4
Ours (with InterAug) *	93.4	81.6	48.2

Ablation study on model architecture



Visualization on U3DPW

Quantitative experiment



Visualization on PennAction

Reference:

- [1] Akhter, Ijaz and Michael J. Black. "Pose-conditioned joint angle limits for 3D human pose reconstruction." *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015): 1446-1455.
- [2] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).