南京航空航天大學
Nanjing University of Aeronautics and Astronautics

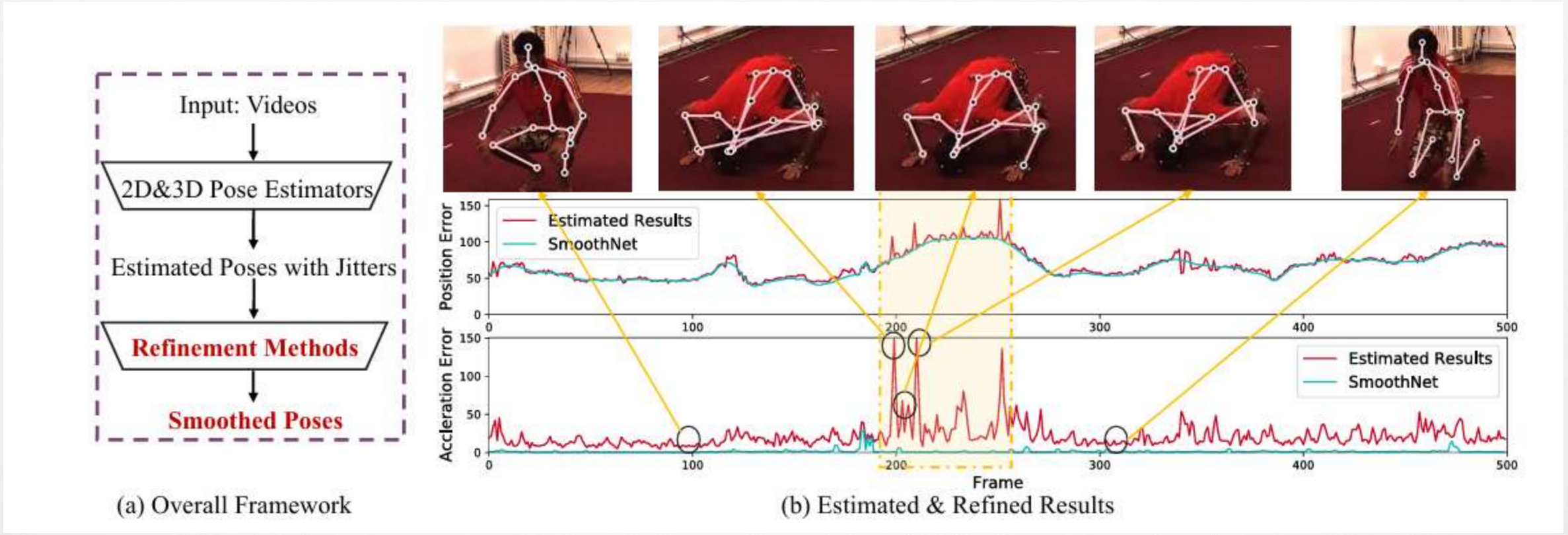
SynSP: Synergy of Smoothness and Precision in Pose Sequences Refinement

Tao Wang[1], Lei Jin[1], Zheng Wang[2], Jianshu Li[3], Liang Li[4], Fang Zhao[5], Yu Cheng[6], Li Yuan[7], Li Zhou[7], Junliang Xing[8], Jian Zhao[9,10]

Beijing University of Posts and Telecommunication[1],Wuhan University[2], Ant Group[3], Institute of computing technology Chinese Academ of Sciences[4], Nanjing University[5], National University of Singapore[6], Peking University[7],Tsinghua University[8], China Telecom Institute of AI[9], Northwestern Polytechnical Universit[10]

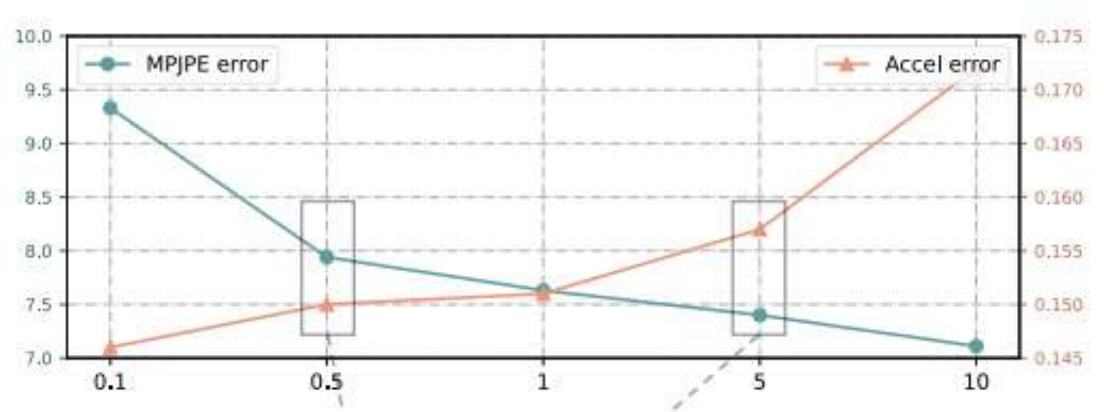
CVPR2024

Introduction

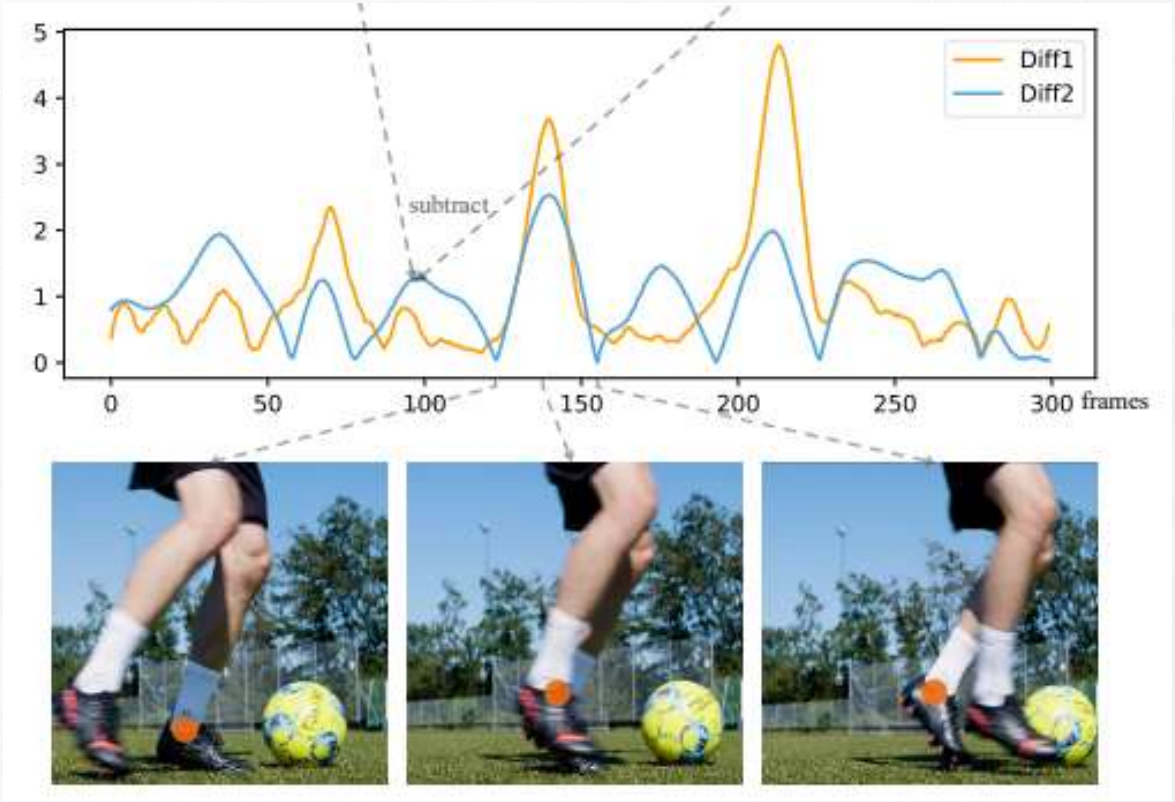


Pose Sequences Refinement

Introduction

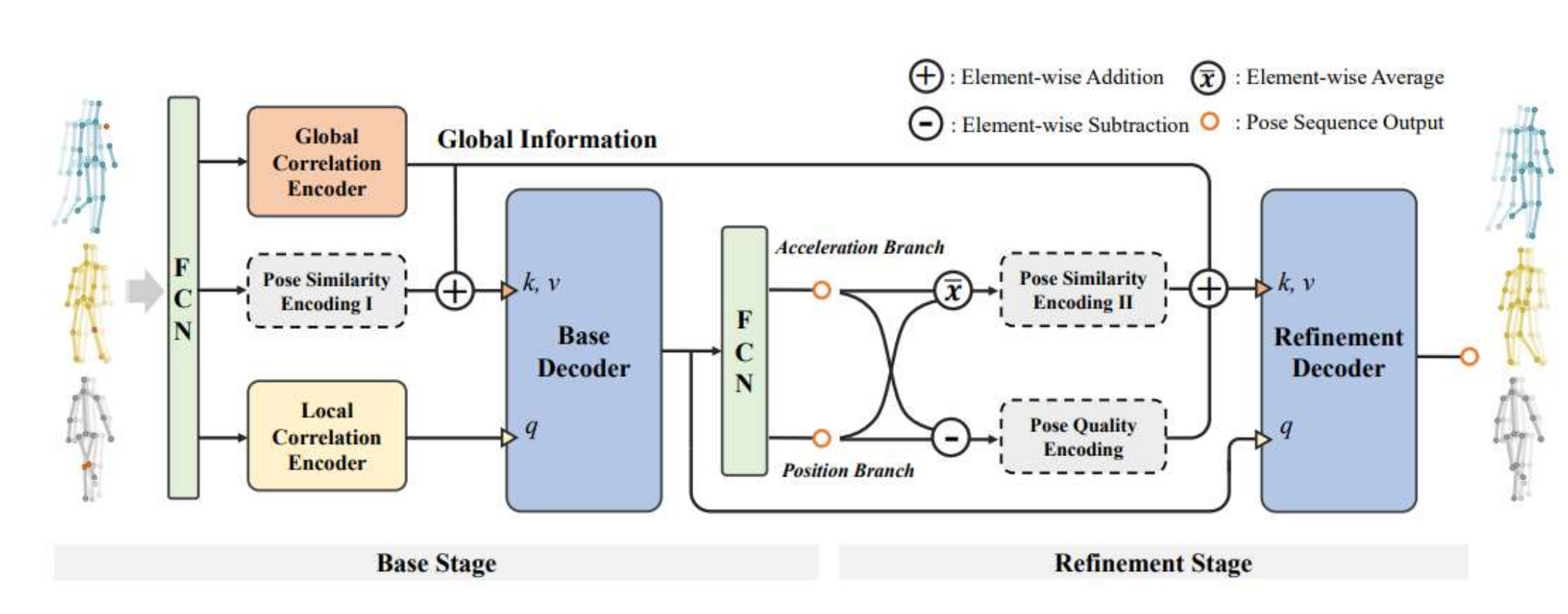


MPJPE (Precision) and Acceleration error (Smoothness) exhibit a conflicting relationship



Diff1: the difference between the predicted sequence and the gound-truth sequence
Diff2: the difference between two output sequences from training SynSP with a loss ratio of 0.5 and 5, respectively.

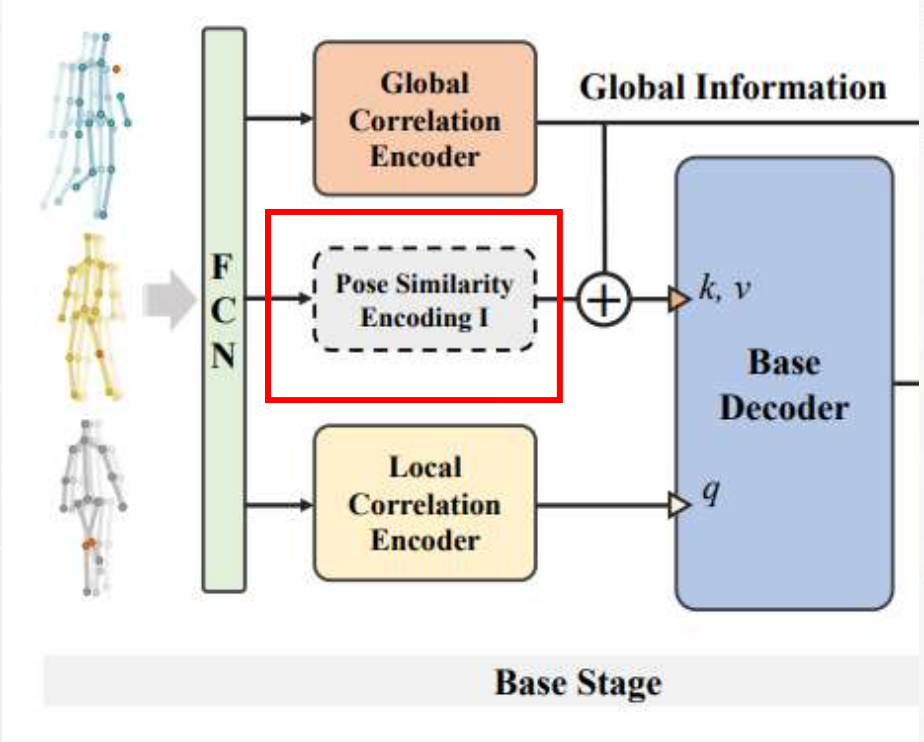
Method



Overview of SynSP

Method

Pose Similarity Encoding:

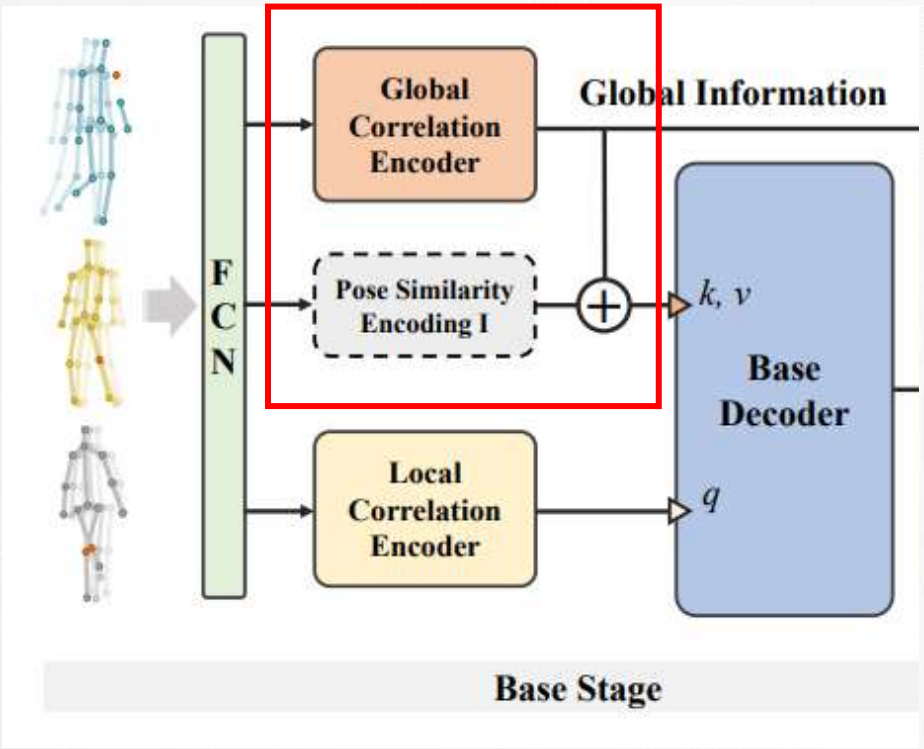


$$P^{rel}_{(v,t,j)} = \hat{P}_{(v,t,j)} - P_{(v,t,c)},$$

$$PSE_{(v_1,v_2,t)} = exp(-\sum_{j=1}^{J}(P^{rel}_{(v_1,t,j)} - P^{rel}_{(v_2,t,j)})^2)$$

Method

Global Correlation Encoder:



$$E_{(\boldsymbol{v},\boldsymbol{t},\boldsymbol{e})} = FCN\ (P_{(\boldsymbol{v},\boldsymbol{t},\boldsymbol{j})})$$

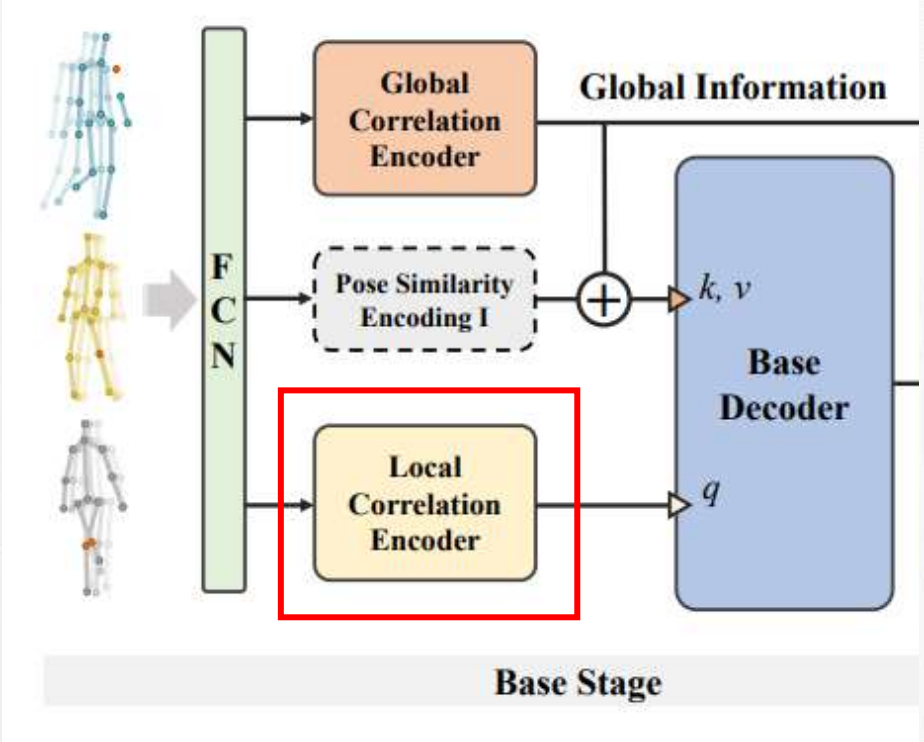
$$E^{g}_{((v,t),e)} = Global\ Encoder\ (E_{((\boldsymbol{v},\boldsymbol{t}),\boldsymbol{e})})$$

$$E^{g}_{((v_2,t),e)} + PSE_{(v_1,v_2,t)} = E^{g}_{(1,(v_2,t),e)} + PSE_{(v_1,(v_2,t),e)}$$

$$= E^{g}_{(v_1,(v_2,t),e)}$$

Method

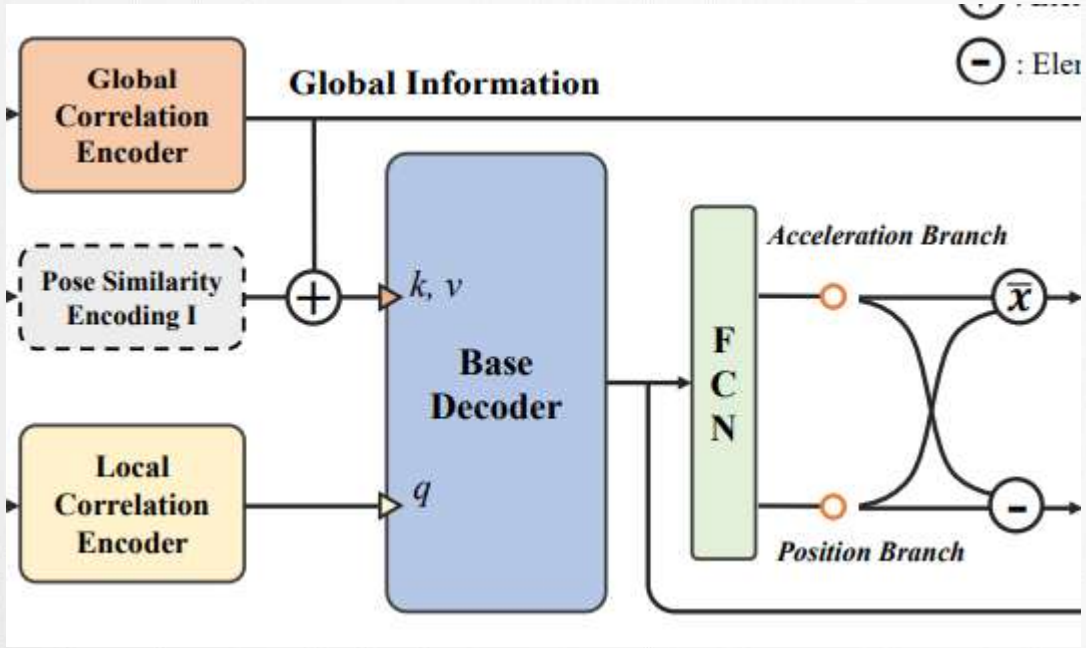
Local Correlation Encoder:



$$E^{l}_{(v_1,t,e)} = Local\ Encoder\ (E_{(v_1,t,e)})$$

Method

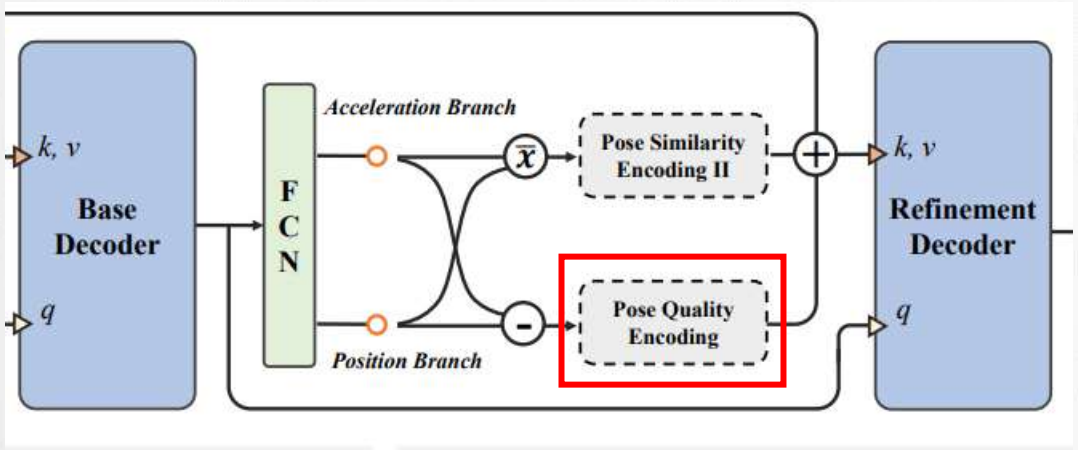
Base Decoder:



$$K, V = E^{g}_{(v_1,(v_2,t),e)},\ Q^{0} = E^{l}_{(v_1,t,e)},$$
$$Q^{d_{th}}_{(v_1,t,e)} = \mathbf{DLayer}^{d_{th}}(Q^{(d-1)_{th}}, K, V), d \in \{1,2,3\},$$
$$P^{pos}_{(v_1,t,j)} = \mathrm{FCN}(\mathbf{DLayer}^{4\mathrm{th}}(\mathrm{Q}^{3\mathrm{th}}, \mathrm{K}, \mathrm{V})),$$
$$P^{acc}_{(v_1,t,j)} = \mathrm{FCN}(\mathbf{DLayer}^{5\mathrm{th}}(\mathrm{Q}^{3\mathrm{th}}, \mathrm{K}, \mathrm{V})),$$

Method

Pose Quality Encoding:



$$P^{dis}_{(v_1,v_2,t,j)} = |P^{acc}_{(v_1,t,j)} - P^{pos}_{(v_2,t,j)}|,$$

$$PQE'_{(v_1,v_2,t)} = \frac{1}{J}\sum_{j=1}^{J} exp(-\frac{P^{dis}_{(v_1,v_2,t,j)}}{\max\limits_{j\in J} P^{dis}_{(v_1,v_2,t,j)} + \sigma}),$$

$$PQE_{(v_1,v_2,t)} = \begin{cases} PQE'_{(v_1,v_2,t)}, & v_1 = v_2; \\ 0, & v_1 \neq v_2. \end{cases}$$

Experiments

Method	WS	Human3.6M / 2D			Human3.6M / 3D			AIST++ / SMPL		
		MPJPE↓	PA-MPJPE↓	Accel↓	MPJPE↓	PA-MPJPE↓	Accel↓	MPJPE↓	PA-MPJPE↓	Accel↓
Input	N/A	9.42	7.64	1.54	54.55	42.2	19.17	107.72	74.40	33.20
One-Euro [3]	1	10.69	7.98	0.34	55.20	42.73	3.80	108.97	75.27	14.70
Gaussian1d [40]	32	9.37	7.56	0.51	53.67	41.60	2.43	104.84	72.18	10.05
Savitzky-Golay [30]	32	9.35	7.55	0.17	53.48	41.19	1.34	104.58	72.30	6.07
SmoothNet [42]	32	9.25	7.57	**0.15**	52.72	40.92	1.03	103.00	71.19	**5.72**
SynSP	8	8.13	6.09	**0.15**	51.36	40.13	1.02	**84.63**	**59.02**	6.08
SynSP*	8	**7.62**	**5.64**	**0.15**	**41.78**	**33.32**	**0.98**	-	-	-

* indicates that SynSP is trained and tested in multi-view scenes.

Experiments

Dataset	Method	WS	MPJPE↓	Accel↓
3DPW	MAED [36]	16	79.00	-
	MPS-Net [38]	16	84.30	-
	TCMR [6]	16	86.46	6.75
	TCMR+SmoothNet [42]	16+32	86.50	6.00
	TCMR+SynSP	16+8	86.10	**5.90**
	PARE [18]	1	79.00	25.60
	PARE+SmoothNet [42]	1+32	78.10	5.91
	PARE+SynSP	1+8	**76.20**	6.16
AIST++	VIBE [17]	16	106.9	31.60
	VIBE+SmoothNet [42]	16+32	97.47	**4.15**
	VIBE+SynSP	16+8	**77.00**	4.32
H36M	TransFusion* [24]	1	25.52	-
	PPT* [25]	1	25.16	11.60
	PPT*+SmoothNet [42]	1+32	23.97	1.31
	PPT*+SynSP	1+8	**21.51**	**1.10**

* indicates multi-view inputs.

Experiments

Ablation Study

Table 5. Ablation study on PSE I, RS (Refinement Stage), PSE II, and PQE modules. * denotes directly using the output from base stage as input of refinement stage to re-optimize these sequnence.

PSE I	RS	PSE II	PQE	MPJPE	Accel
				51.2	1.09
√				48.7	1.07
	√*			51.4	1.08
	√			49.6	1.07
√	√			46.7	1.05
√	√	√		46.4	1.04
√	√		√	42.5	0.99
√	√	√	√	41.8	0.98

Thank you for watching!