Incorporating Test-Time Optimization into Training with Dual Networks for Human Mesh Recovery

Yongwei Nie [1], Mingxian Fan [1], Chengjiang Long [2], Qing Zhang [3], Jian Zhu [4], Xuemiao Xu [1*]

[1] South China University of Technology, [2] Meta Reality Lab, [3] Sun Yat-sen University, [4] Guangdong University of Technology

Introduction

Motivation

Q1

There is a gap between the training and testing optimizations objectives.

Q2

Pretrained model is not specially tailored for the test-time optimization.

Contribution

- We propose a novel dual-network HMR framework with test-time optimization involved into the training procedure, which improves the effectiveness of the test-time optimizations.
- We ensure the test-time objectives identical to the training objectives, further facilitating the joint-training of the test-time and training-time optimizations.
- Extensive experiments validate that our results outperform those of previous approaches both quantitatively and qualitatively.

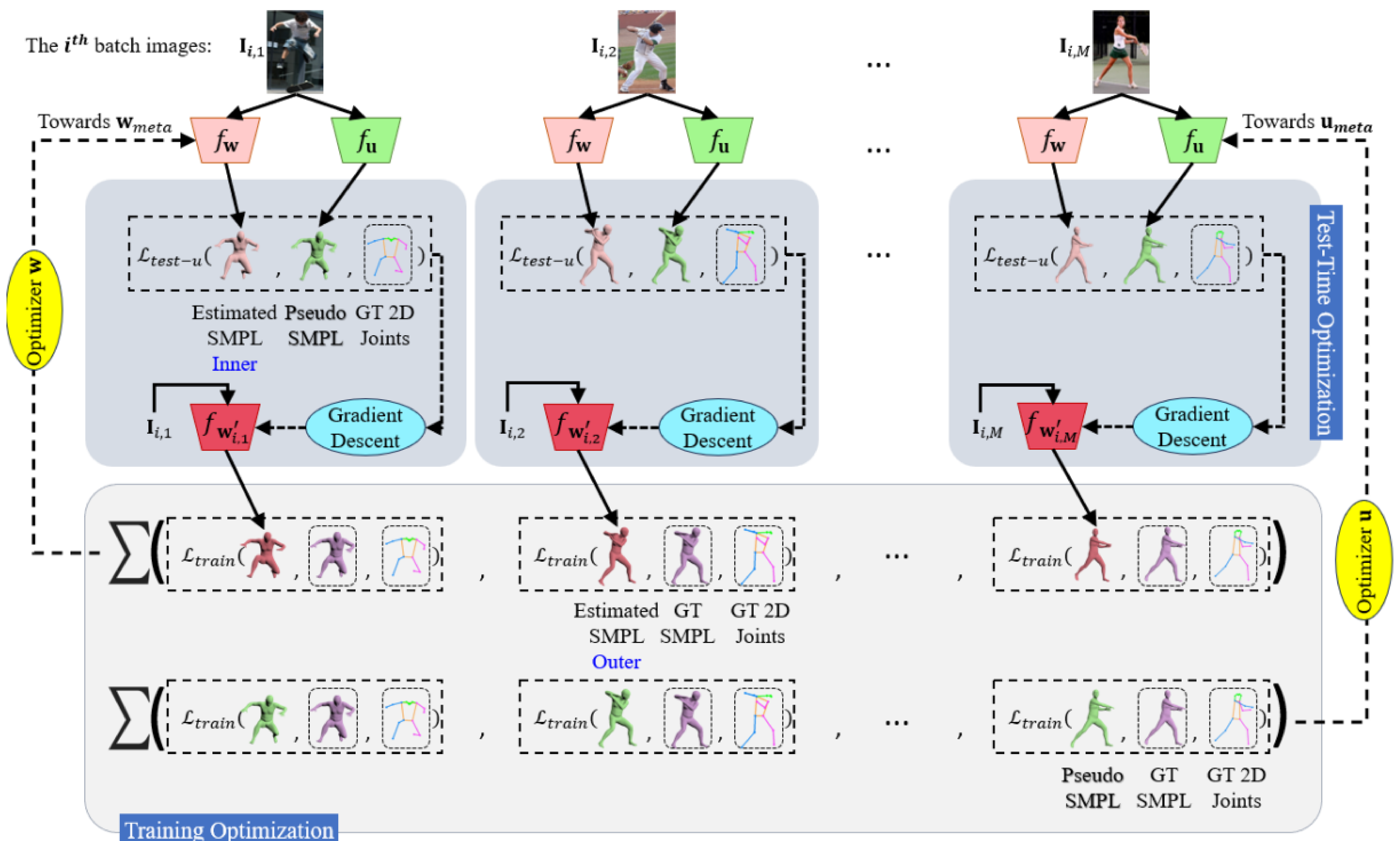
More Information

Github

Paper

Method

Overall Pipeline



Step1

Auxiliary network $f_{\mathbf{u}}$ produces pseudo labels and unifies the training and test-time optimization objectives:

$$\hat{\Theta}^u_{i,j} = f_{\mathbf{u}}(\mathbf{I}_{i,j}) \qquad \mathbf{w}'_{i,j} = \mathbf{w} - \alpha \nabla_{\mathbf{w}} \mathcal{L}_{test-u}(f_{\mathbf{w}}(\mathbf{I}_{i,j}), \hat{\Theta}^u_{i,j}, \hat{\mathbf{J}}_{i,j})$$

Step2

Inspired by optimization-based meta-learning, we perform a step of test-time optimization before conducting training optimization. This results in a meta-model, whose meta-parameters are friendly to test-time optimization.

$$\mathbf{w}_{meta}, \mathbf{u}_{meta} = \arg\min_{\mathbf{w},\mathbf{u}} \sum_{i=1}^{B}\sum_{j=1}^{M} (\mathcal{L}^1_{train}(f_{\mathbf{w}'_{i,j}}(\mathbf{I}_{i,j}), \hat{\Theta}_{i,j}, \hat{\mathbf{J}}_{i,j}) + \mathcal{L}^2_{train}(f_{\mathbf{u}}(\mathbf{I}_{i,j}), \hat{\Theta}_{i,j}, \hat{\mathbf{J}}_{i,j}))$$

Step3

Armed with meta-learning and dual networks, we perform several test-time optimization steps starting from the meta-parameters, ultimately achieving higher HMR accuracy compared to starting from a standard pretrained regression model.

Experiments

Quantitative Comparison

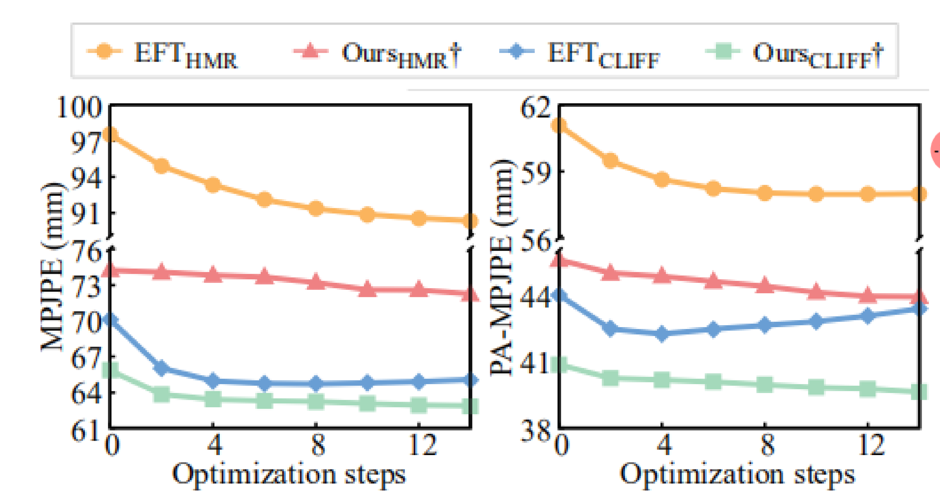
Method	3DPW			Human3.6M	
	MPJPE↓	PA-MPJPE↓	PVE↓	MPJPE↓	PA-MPJPE↓
CLIFF [31]'22	69.0	43.0	81.2	47.1	32.7
LearnSample [58]'22	70.5	43.3	82.7	45.9	33.5
ProPose [11]'23	68.3	40.6	79.4	45.7	**29.1**
POTTER [65]'23	75.0	44.8	87.4	56.5	35.1
DeFormer [59]'23	72.9	44.3	82.6	44.8	31.6
EFT [20]'21	85.1	52.2	98.7	63.2	43.8
NIKI [29]'23	71.3	40.6	86.6	-	-
ReFit [55]'23	65.8	41.0	-	48.4	32.2
PLIKS [47]'23	66.9	42.8	82.6	49.3	34.7
$\text{Ours}_{\text{CLIFF}}\dagger$(OpenPose 2D)	62.9	39.7	80.1	43.9	30.3
$\text{Ours}_{\text{CLIFF}}*$(RSN 2D)	**62.4**	**39.5**	**78.1**	**42.0**	**29.1**

Qualitative Comparison

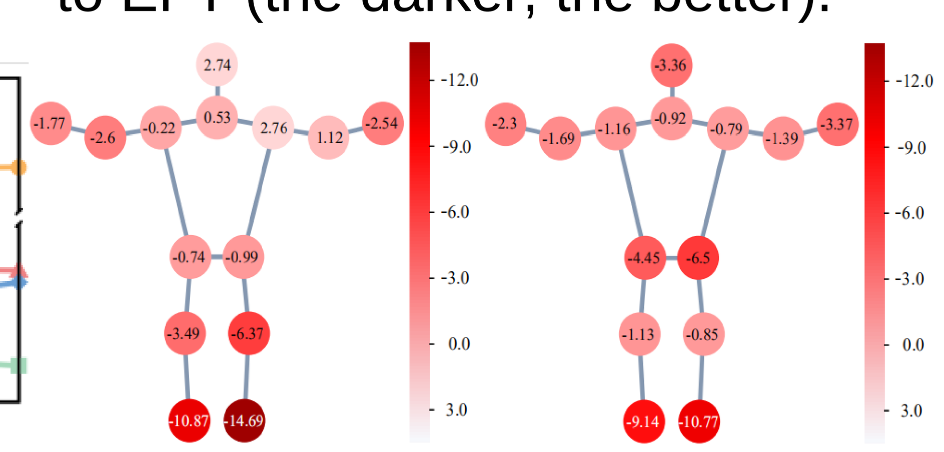
Input CLIFF Ours† Ours∗ Novel view Input Refit Ours† Ours∗ Novel view

Discussion

Influence of optimization steps



Per-joint error analysis compared to EFT (the darker, the better).



Acknowledgements

The work was supported by Guangdong Provincial Natural Science Foundation for Outstanding Youth Team Project (No. 2024B1515040010), the Fundamental Research Funds for the Central Universities (No. 2024ZYGXZR021), etc.