Supplementary Document for "MoCap-Solver: A Neural Solver for Optical Motion Capture Data"

In this document, we give details of the network architectures of MoCap-Solver and present more comparison results for the real MoCap dataset.

The dimensions of the input and outputs of all components of MoCap-Solver are shown in Table 1. The skeleton structure we use is consistent with the SMPL model, which has 24 joints. As the topology of our skeleton is topologically consistent with the atom skeleton in [Aberman et al. 2020], the dimensions of the latent codes of the motion encoder and the template skeleton are consistent with those in [Aberman et al. 2020].

Table 1: Network architectures of various components of MoCap-Solver.

Module	Comp.	Input	Output
Marker Configuration	E_C	$X_C : (56, 24, 3)$	$l_C : (1024)$
	D_C	$l_C : (1024)$	$Y_C : (56, 24, 3)$
Template Skeleton	E_T	$X_T : (24, 3)$	$l_T : (168)$
	D_T	$l_T : (168)$	$Y_T : (24, 3)$
Motion	E_M	$X_M : (99, 64)$	$l_M : (1792)$
	D_M	$l_M : (1792)$	$Y_M : (99, 64)$
MoCap-Solver	E_{MS}	$X : (64, 56, 3)$	$l_C : (1024)$
			$l_T : (168)$
			$l_M : (1792)$
Marker Reliability	F_M	$X_t : (56, 3)$	$P : (8)$

More comparisons of prediction results with [Holden 2018] for the real MoCap dataset are showed in Figure 1.

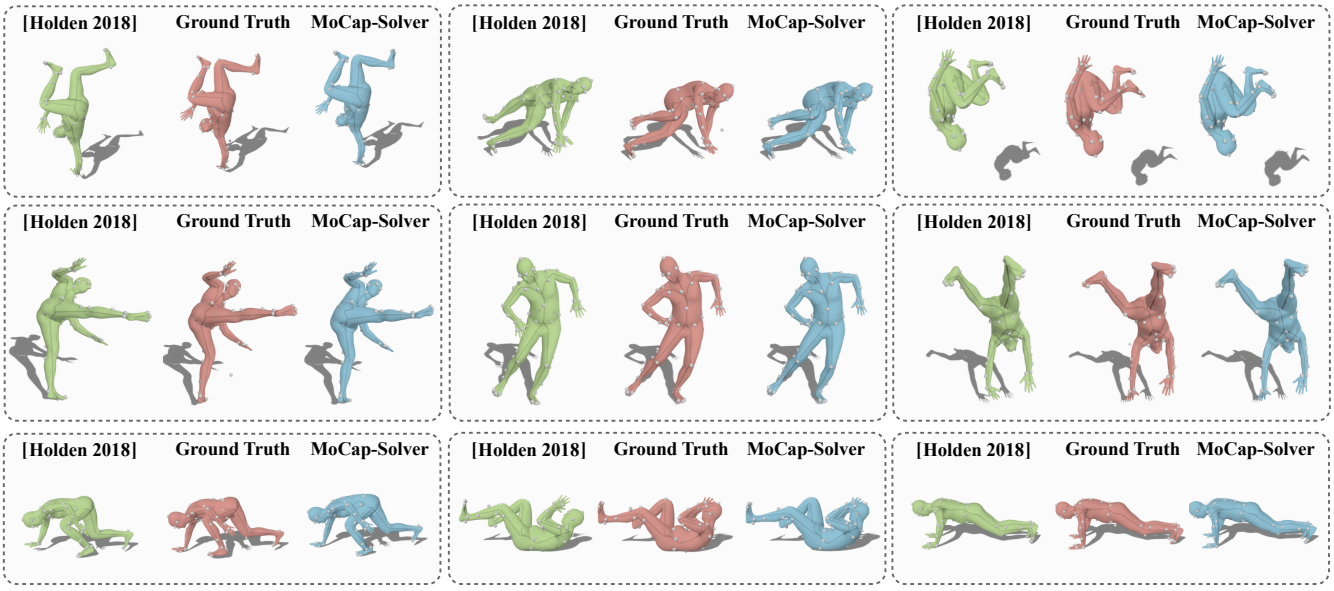

Figure 1: Comparison of prediction results for the real MoCap dataset. Green: [Holden 2018]. Blue: MoCap-Solver. Red: ground-truth. Orientations in our results are more accurate than Holden's for head, hand, arm, leg and feet.