JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Skinned Motion Retargeting with Residual Perception of Motion Semantics & Geometry

Jiaxu Zhang[1], Junwu Weng[2], Di Kang[2], Fang Zhao[2], Shaoli Huang[2], Xuefei Zhe[2], Linchao Bao[2], Ying Shan[2], Jue Wang[2] and Zhigang Tu[1*]

[1]Wuhan University [2]Tencent AI Lab

Problem Statement

Motion retargeting

SAN (TOG 2020), Kfir Aberman, et al.

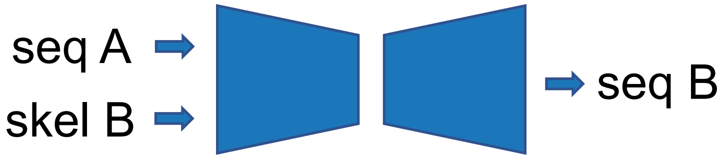
Mapping the motion of a source character to a target character without losing plausibility.

Traditional methods：

- Optimization with kinematic constraints
- Motion-specific and hand-designed constraints
- Post-processing and hand-tuning steps

Deep neural network methods：

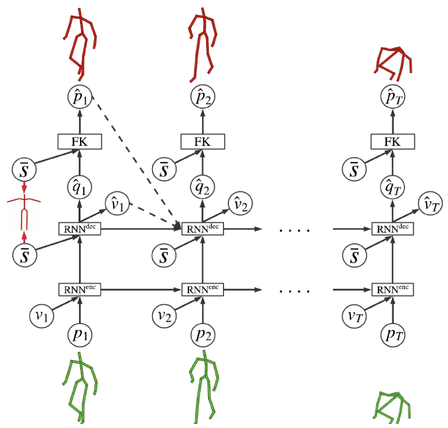
- Data-driven
- End-to-end solution



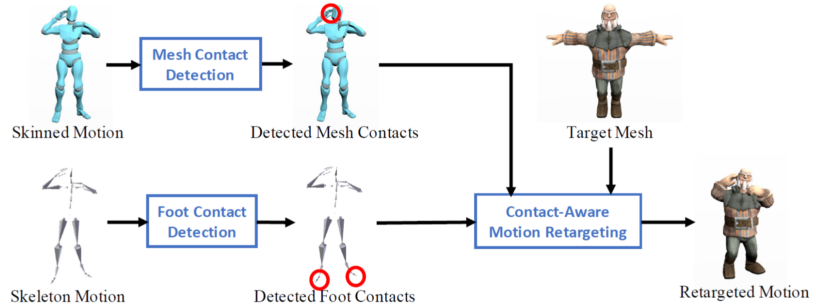
Related Work

Previous learning-based methods

- **Full mapping and without consideration of geometry:**
  - NKN (CVPR 2018), Ruben Villegas, et al.
  - PMnet (BMVC 2019), Lim Jongin, et al.
  - SAN (TOG 2020), Kfir Aberman, et al.
  - ItMRnet (C&G 2022), Shujie Li, et al.



- **Post-processing for geometry preserving:**
  - Contact-aware (CVPR 2021), Ruben Villegas, et al.

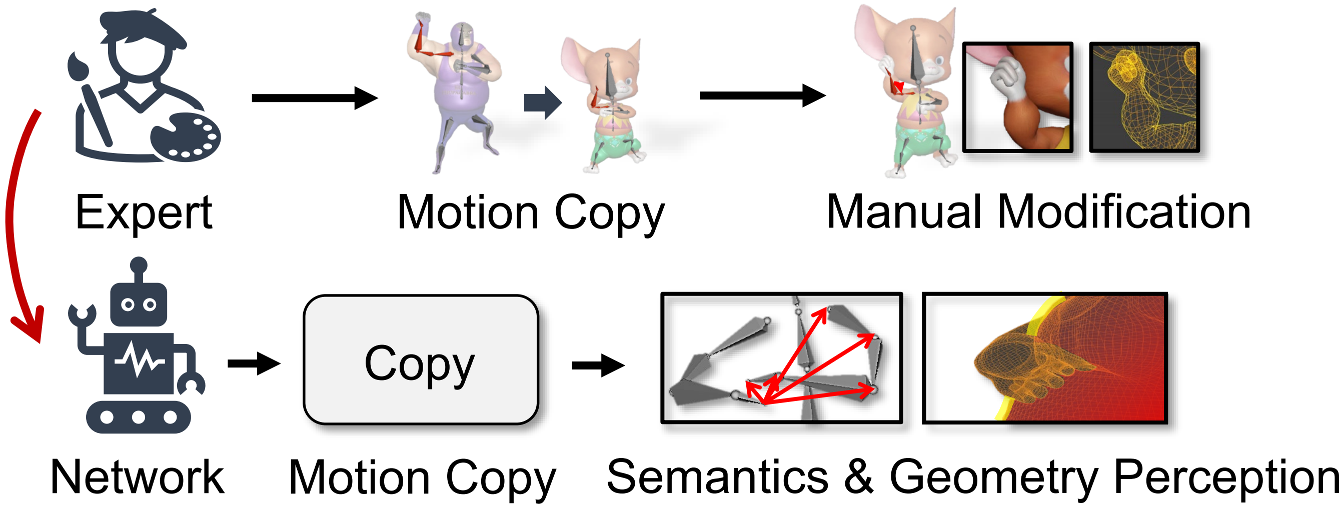


Our goal:

- Enable the network to perceive both the semantics of the motion and the geometry of the character during the inference process without requiring post-processing.

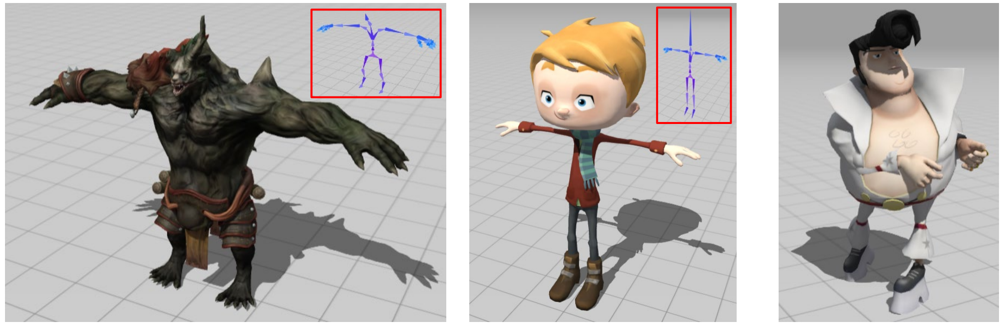
Motivation & Challenges

Motivation



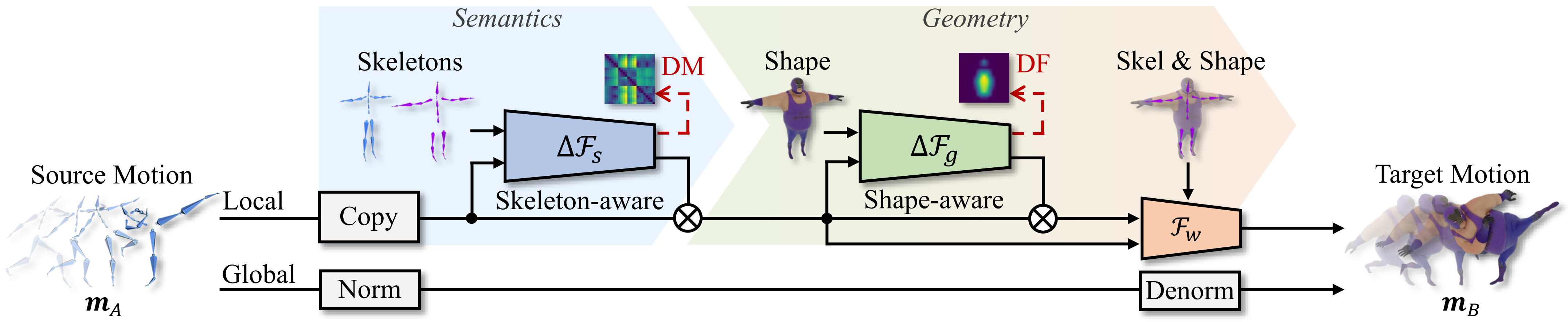
- We observe that artists usually copy the motion of the source character, and then manually modify it to preserve motion semantics and avoid translation artifacts, e.g., interpenetration, during motion reuse in new characters.

Challenges



- Lack of paired motion data from different characters.
- Differences in bone lengths and proportions.
- Various of shape geometries.
- Motion should be plausible and realistic.

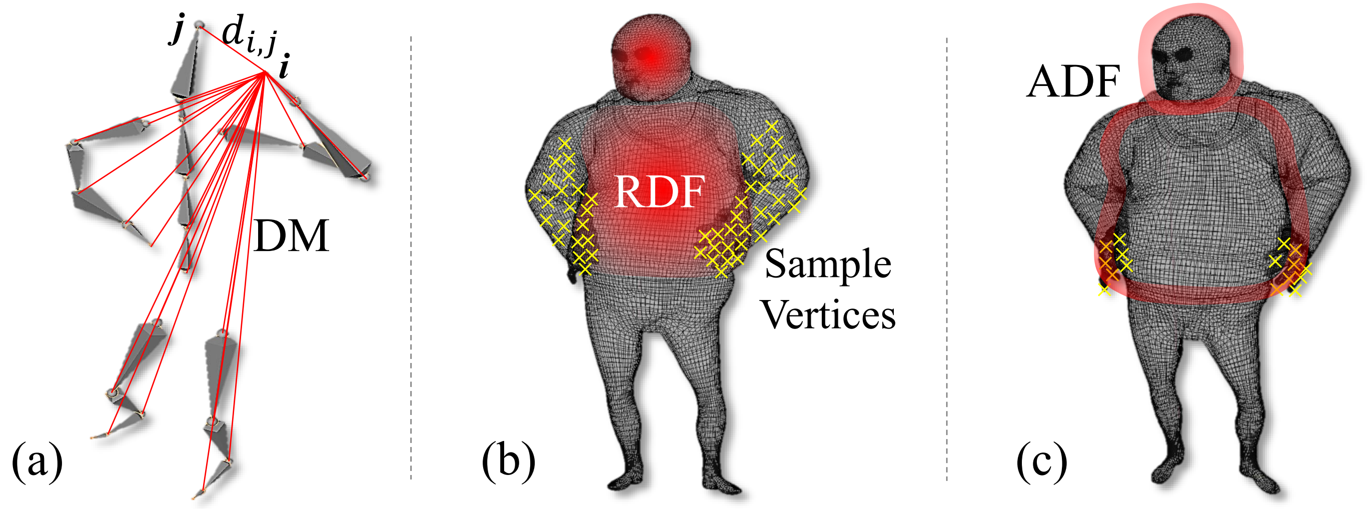
Residual RETargeting network (R^2ET)



Key designs

- A *residual network structure* for neural motion retargeting
- A *skeleton-aware modification module* for motion semantics perceiving.
- A *shape-aware modification module* for shape geometry perceiving.
- A *balancing gate* for to make a trade-off between two modifications.
- *Distance-based measurements for* motion semantics and geometry learning method.

Residual RETargeting network (R²ET)



Distance-based measurements:

- Normalized Distance Matrix
- Repulsive Distance Field
- Attractive Distance Field

Loss Functions:

- Semantics Similarity Loss

$$\mathcal{L}_{sem} = \left\| \eta\left(\frac{\boldsymbol{D}_A}{h_A}\right) - \eta\left(\frac{\boldsymbol{D}_B}{h_B}\right) \right\|_2^2,$$

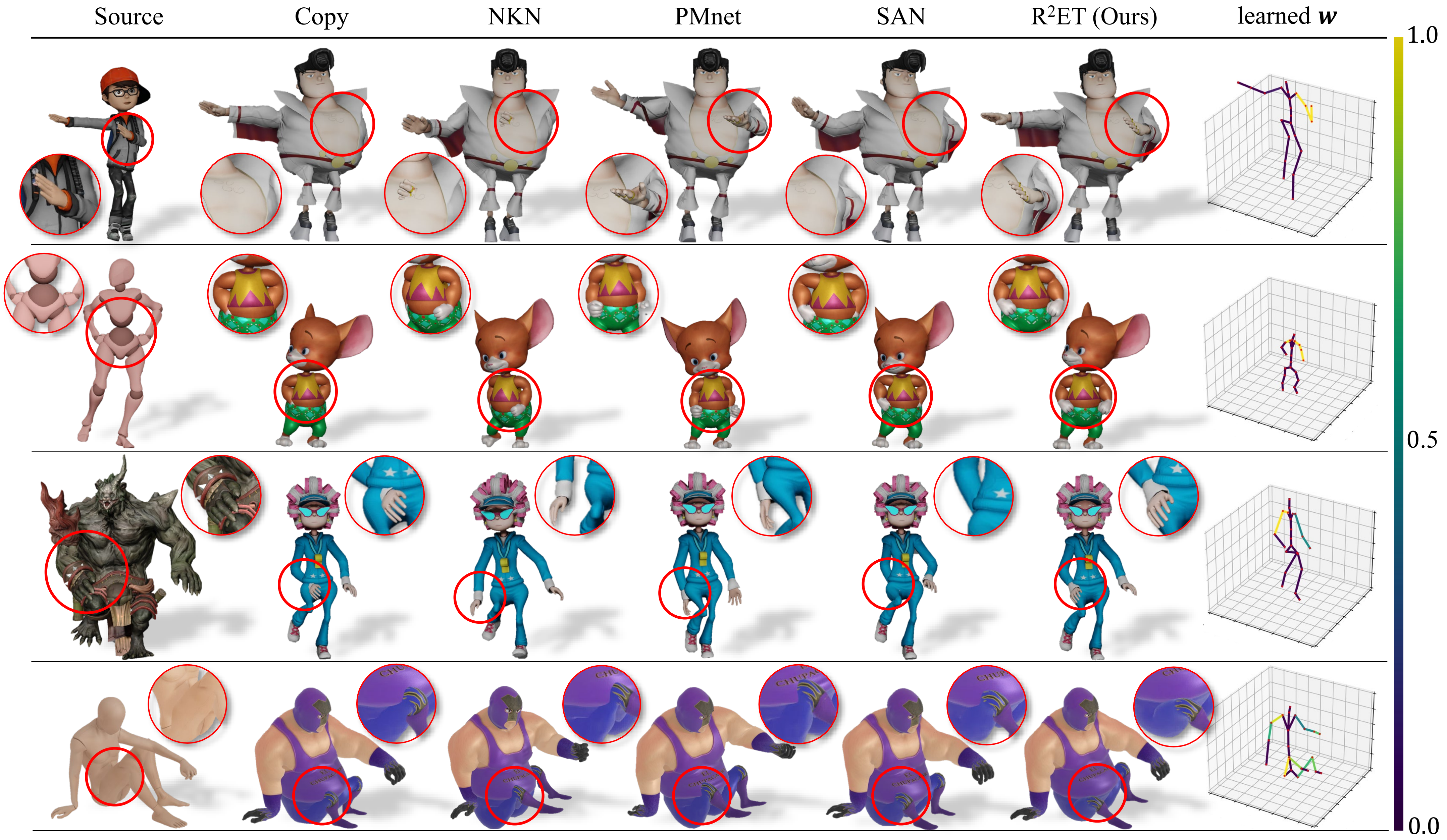
- Repulsive Loss

$$\mathcal{L}_{rep} = \frac{1}{N_l} \sum_{e \in E_l} \psi_R(e),$$

- Attractive Loss

$$\mathcal{L}_{att} = \frac{1}{N_h} \sum_{e \in E_h} \psi_A(e),$$

Results



Results

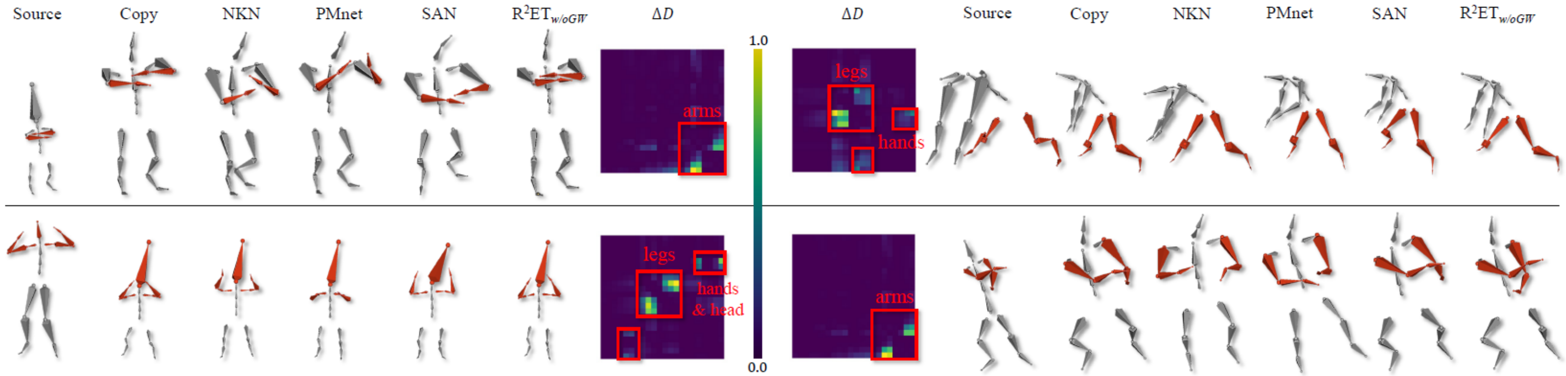
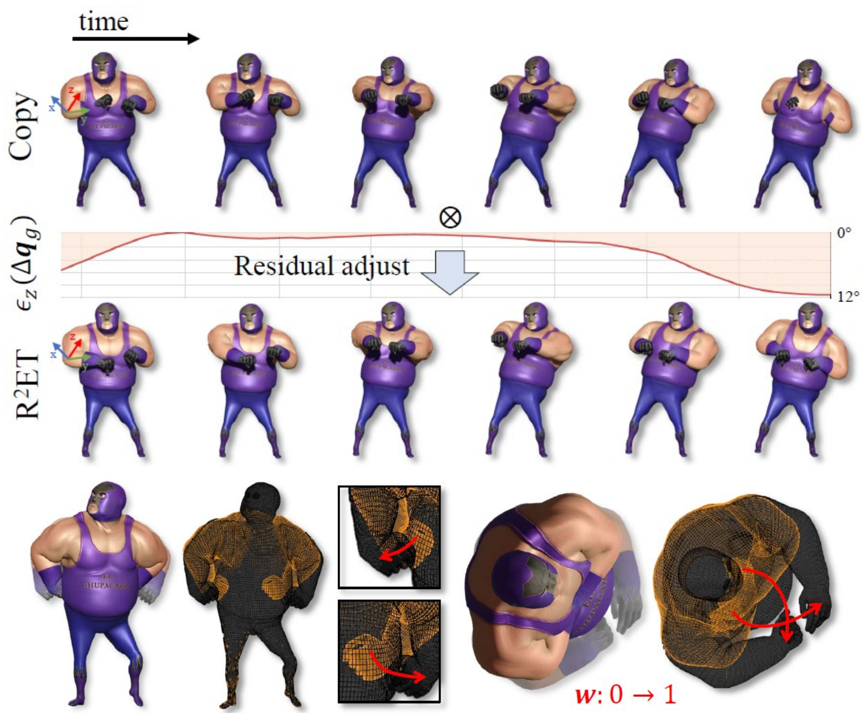


Figure 5. Qualitative results of skeletal motion retargeting. ΔD indicates the DM difference between the motion copy and our result.



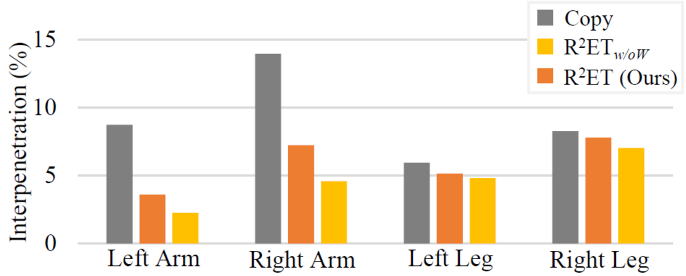


Figure 9. Comparison of the interp. rates of different limbs.

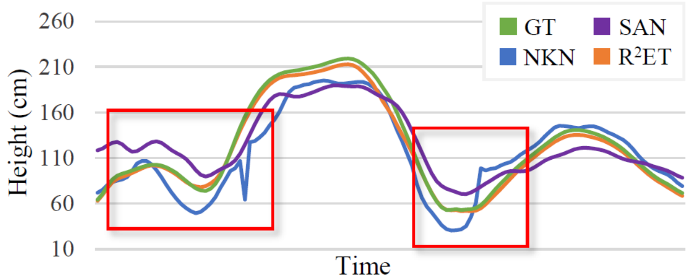


Figure 10. The change of the left-hand end-effector's height of a retargeted motion on time domain.

Table 1. Comparison with the state-of-the-arts. MSE^{lc} is the local MSE. R^{2}ET$_{w/oGW}$ is the model with the skeleton-aware module only. R^{2}ET$_{w/oW}$ is the model without the balancing gate. Copy† is the motion copy without the global motion normalization.

Methods	Inp.	MSE$_\downarrow$	MSE$^{lc}_\downarrow$	Pen.$^{\%}_\downarrow$	Con.$^{cm}_\downarrow$
GT	-	-	-	9.02	4.92
NKN [25]	Pos.	2.298	0.575	8.96	4.42
PMnet [16]		0.806	0.281	7.11	14.7
Copy	Rot.	0.267	0.060	9.23	4.95
Copy†		3.087	0.060	9.23	4.95
SAN [1]		0.321	0.118	8.91	4.86
PMnet*		0.374	0.120	9.03	5.24
R^{2}ET$_{w/oGW}$	Rot.	**0.297**	**0.094**	9.09	4.93
R^{2}ET$_{w/oW}$		0.378	0.178	**4.68**	5.31
R^{2}ET (Ours)		0.318	0.116	5.94	**3.57**

Table 2. Ranking results of the user study. We invite 100 users to compare our retargeting results to that of the recent methods from three aspects, i.e., overall quality (Q), semantics preservation (S), and motion details (D).

Methods	Skeletal Motion			Skinned Motion		
	Q$_\downarrow$	S$_\downarrow$	D$_\downarrow$	Q$_\downarrow$	S$_\downarrow$	D$_\downarrow$
Copy	1.88	1.83	1.84	1.84	1.84	1.93
NKN [25]	3.37	3.45	3.40	3.44	3.44	3.42
PMnet [16]	3.06	3.06	3.06	3.10	3.07	3.00
R^{2}ET (Ours)	**1.69**	**1.67**	**1.70**	**1.63**	**1.65**	**1.64**

Thanks for watching!

More details