

Supplementary Material: Neural Rendering and Reenactment of Human Actor Videos

LINGJIE LIU, University of Hong Kong, Max Planck Institute for Informatics

WEIPENG XU, Max Planck Institute for Informatics

MICHAEL ZOLLHÖFER, Stanford University, Max Planck Institute for Informatics

HYEONGWOO KIM, FLORIAN BERNARD, and MARC HABERMANN, Max Planck Institute for Informatics

WENPING WANG, University of Hong Kong

CHRISTIAN THEOBALT, Max Planck Institute for Informatics

ACM Reference Format:

Lingjie Liu, Weipeng Xu, Michael Zollhöfer, Hyeonwoo Kim, Florian Bernard, Marc Habermann, Wenping Wang, and Christian Theobalt. 2019. Supplementary Material: Neural Rendering and Reenactment of Human Actor Videos. 1, 1 (February 2019), 2 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

1 IMPLEMENTATION DETAILS OF CONDITIONING INPUT IMAGES.

We clip the depth values in \mathcal{D} to a depth range of two meters starting from the minimum depth value, and normalize them to the interval $[0, 255]$. The size of \mathbf{X} is $W \times H \times (P \cdot C + B)$, where $W = 256$ and $H = 256$ denote the image width and height, respectively. $P = 6$ is the number of body parts, $C = 4$ denotes the number of channels of the rendered RGBD images, and $B = 3$ denotes the number of channels of the background image (RGB).

2 ADDITIONAL COMPARISONS

In this section, we provide comparisons to the concurrent unpublished work of [Aberman et al. 2018], [Chan et al. 2018] and [Wang et al. 2018]. Note that, neither the code nor the dataset of their work is available. Therefore, we can only provide qualitative comparisons of the results on similar poses. As can be seen from Fig. 2, our method yields clearly much sharper imagery with less artifacts than the methods of [Aberman et al. 2018] and [Wang et al. 2018]. In general, the quality of our results is comparable with or better than that of [Chan et al. 2018]. The results of [Chan et al. 2018] occasionally exhibit strong artifacts such as broken and unnaturally extended arms (see Fig. 1). While our method exhibits occasional failures too, such strong errors are less likely to occur in our method

Authors' addresses: Lingjie Liu, liulingjie0206@gmail.com, University of Hong Kong, Max Planck Institute for Informatics; Weipeng Xu, w Xu@mpi-inf.mpg.de, Max Planck Institute for Informatics; Michael Zollhöfer, zollhoefer@cs.stanford.edu, Stanford University, Max Planck Institute for Informatics; Hyeonwoo Kim, hyeonwoo.kim@mpi-inf.mpg.de; Florian Bernard, fbernard@mpi-inf.mpg.de; Marc Habermann, mhaberma@mpi-inf.mpg.de, Max Planck Institute for Informatics; Wenping Wang, wenping@cs.hku.hk, University of Hong Kong; Christian Theobalt, theobalt@mpi-inf.mpg.de, Max Planck Institute for Informatics.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Association for Computing Machinery.

XXXX-XXXX/2019/2-ART \$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

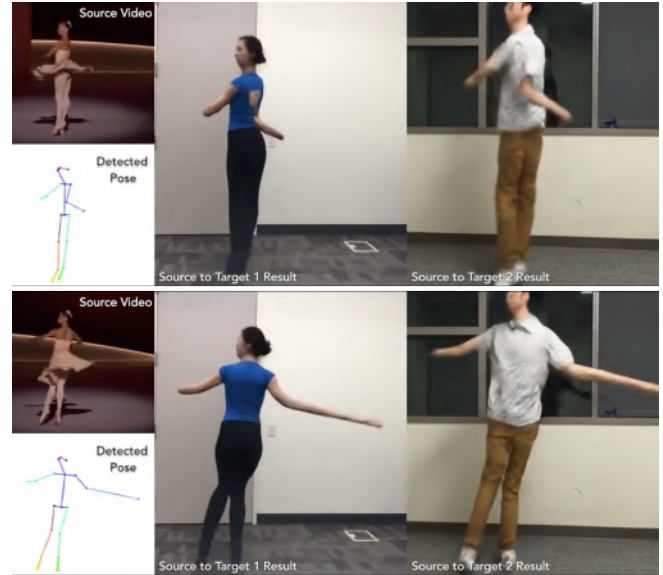


Fig. 1. The results of [Chan et al. 2018] exhibit occasional strong artifacts such as broken and unnaturally extended arms.

which is based on a more sophisticated shape model as conditioning input.

REFERENCES

- Kfir Aberman, Mingyi Shi, Jing Liao, Dani Lischinski, Baoquan Chen, and Daniel Cohen-Or. 2018. Deep Video-Based Performance Cloning. [arXiv:arXiv:1808.06847](https://arxiv.org/abs/1808.06847)
- Caroline Chan, Shiry Ginosar, Tinghui Zhou, and Alexei A. Efros. 2018. Everybody Dance Now. [arXiv:arXiv:1808.07371](https://arxiv.org/abs/1808.07371)
- Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Guilin Liu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. 2018. Video-to-Video Synthesis. [arXiv:arXiv:1808.06601](https://arxiv.org/abs/1808.06601)

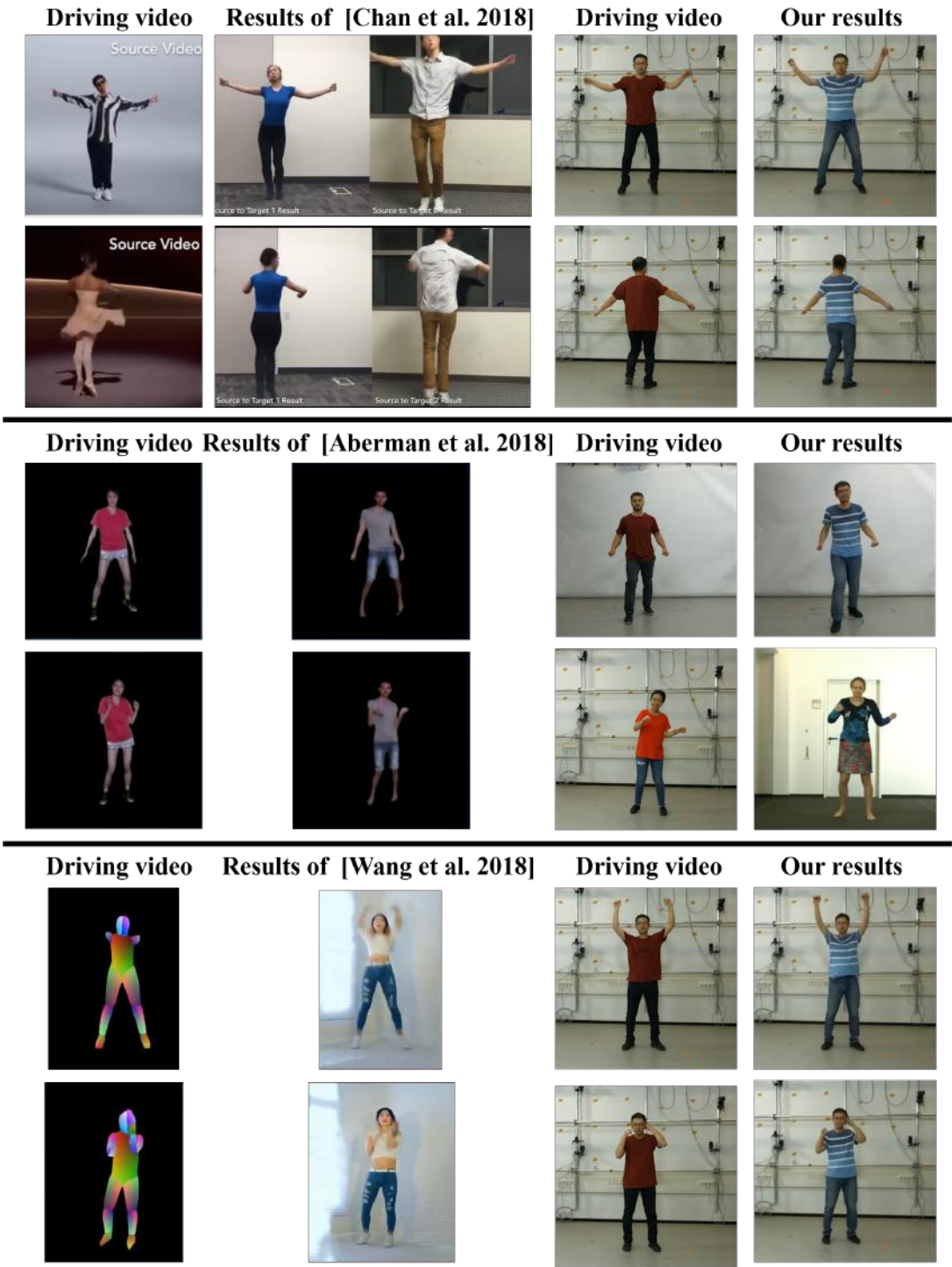


Fig. 2. Qualitative comparison to the concurrent unpublished work of [Aberman et al. 2018], [Chan et al. 2018] and [Wang et al. 2018]: Our method yields clearly much sharper imagery with less artifacts than the methods of [Aberman et al. 2018] and [Wang et al. 2018], and comparable or better results than [Chan et al. 2018].