Supplementary Materials for WACV'24 Paper Titled "3D-Aware Talking-Head Video Motion Transfer"

Haomiao Ni[1] Jiachen Liu[1] Yuan Xue[2] Sharon X. Huang[1]
[1]The Pennsylvania State University, University Park, PA, USA
[2]The Ohio State University, Columbus, OH, USA
[1]{hfn5052, jzl6493, suh972}@psu.edu [2]Yuan.Xue@osumc.edu

1. Information about Attached Videos

We attach 6 MP4 files of example video clips generated by our proposed method in Supp. materials. All the given images are testing (*unseen*) images.

- **canon_rgbd.mp4** shows the synthesized talking-head video clips and the movement of corresponding 3D canonical head using our proposed Head3D under the cross-identity transfer setting. The "RGB of moving 3D head" and "Depth of moving 3D head" are RGB and depth map of the transformed 3D canonical head, which is generated by applying the estimated pose from each driving frame to the canonical head.
- **pose_control.mp4** shows the synthesized examples of pose-controllable novel view synthesis using our proposed Head3D. Each column demonstrates changing of a 3D rotation or translation parameter.
- **same_id.mp4** shows the synthesized talking-head video clips using our proposed Head3D under the self-reconstruction setting.
- **cross_id.mp4** shows the synthesized talking-head video clips using our proposed Head3D under the cross-identity transfer setting.
- **sota.mp4** shows the qualitative comparison of our method Head3D against other state-of-the-art models.
- **256x256.mp4** shows the synthesized talking-head video clips using our proposed Head3D trained at the resolution of 256×256.

2. Positive Negative Societal Impact

Video motion transfer technologies could be misused for unethical purposes, *e.g.*, creating fake news videos of celebrities [2]. Therefore, we will limit the usage of our method to research purposes only. Additionally, we intend to explore fake video detection techniques [1] that can effectively identify fake videos generated using our proposed method.

References

[1] Ashifur Rahman, Md Mazharul Islam, Mohasina Jannat Moon, Tahera Tasnim, Nipo Siddique, Md Shahiduzzaman, and Samsuddin Ahmed. A qualitative survey on deep learning based deep fake video creation and detection method. *Aust. J. Eng. Innov. Technol*, 4(1):13–26, 2022. 1

[2] Sihyun Yu, Jihoon Tack, Sangwoo Mo, Hyunsu Kim, Junho Kim, Jung-Woo Ha, and Jinwoo Shin. Generating videos with dynamics-aware implicit generative adversarial networks. *arXiv preprint arXiv:2202.10571*, 2022. 1