Towards Practical Single-shot Motion Synthesis

Konstantinos Roditakis, Spyridon Thermos, Nikolaos Zioulis
Moverse

moverseai.github.io/single-shot

CVPR SEATTLE, WA JUNE 17-21, 2024

1. Motivation

- Contemporary generative models require **massive** computational resources and are trained on vast amounts of **annotated data.**
- Lack of interpretability and control poses risks for sensitive and/or protected data memorization and reproduction during generation.
- Existing single sample generation models [1], [2] constitute an promising alternative but are **slow** to train or generate new data.

2. Contributions

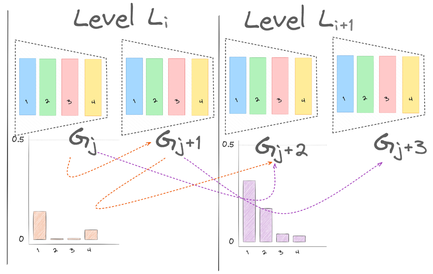
- We show an overall x10 **performance** increase by combining **mini-batch** training and cross-stage **transfer learning**.
- We expand the single sample GAN **application** domain and showcase diverse motion compositing in a **single forwards pass**.

3. Single-shot motion synthesis

- *GANimator* [1], albeit slower to train, offers very fast inference but supports limited applications without re-training.
- *SinMDM* [2] is faster to train, supports more applications with a single train, but suffers from slower iterative inference.

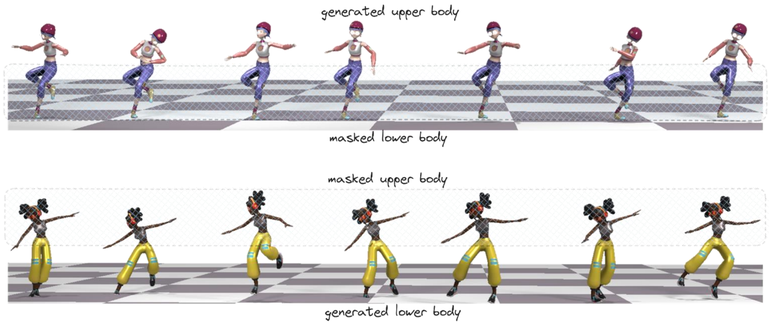
4. Mini-batch training & transfer learning

- Single-sample generative models need to balance training between an adversarial game whose latent space is anchored by a reconstruction objective.
- We prevent **mode collapse** by carefully annealing the weights of the used losses.
- We observe higher **correlation** between specific layers across stages that allow for cross-stage transfer to reduce training time.



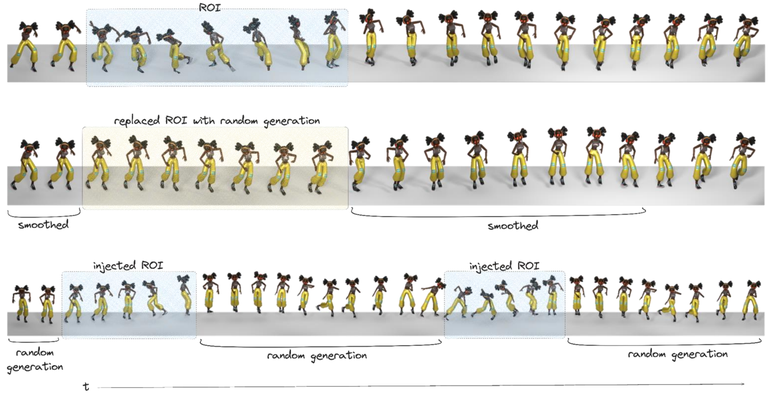
5. Body-part composition

- We focus on applications that do not need re-training,
- Inspired by *SinMDM* [2] we show how a GAN model can be used to for **body-part** motion composition.



6. Full-body composition

- We **compose** new **motions** by mixing mini motion clips of interest.
- The selected mini-clip(s) is ***injected*** at different temporal spots of interest while the rest of the motion is ***inpainted*** by the GAN.



- Even though x6.8 faster to train, performance on the original *GANimator* [1] applications like re-styling and crowd generation are preserved.
- Motion sequences from the **Mixamo** library are used for training and evaluation.
- All experiments are conducted on an Nvidia RTX 3060.

References:

[1] P. Li et al., "GANimator: Neural motion synthesis from a single sequence", ToG, 2022.
[2] S. Raab et al., "Single motion diffusion", ICLR, 2024

Acknowledgement: This work was supported by the Horizon Europe project EMIL-XR [GA101070533].
European Media and Immersion Lab