# Validating Objective Evaluation Metric: Is Fréchet Motion Distance able to Capture Foot Skating Artifacts ?

Antoine Maiorca [1] Youngwoo Yoon [2] Thierry Dutoit [1]

[1]ISIA Lab, University of Mons, Mons, Belgium [2]Electronics and Telecommunications Research Institute, Daejoon, Republic of Korea

## Introduction

**Fréchet Distance ...**

- **Objective evaluation metric** for motion generation is essential in the development of immersive spaces that include animation of avatars.
- **Fréchet Distance** ($FD$) is a popular metric used in generative methods evaluation that assesses jointly the quality and diversity of a generated dataset [2] (see Equation 1).

$$FD = ||\mu_r - \mu_g||_2^2 + Tr(\Sigma_r + \Sigma_g - 2\sqrt{\Sigma_r \Sigma_g}) \quad (1)$$

**... In Motion Generation**

- Validation protocol for **Fréchet Motion Distance** ($FMD$) in the context of motion-generative models evaluation is limited.
- No validation on common motion artifacts produced by deep generative models.

**Aim of this work**

- Identify if the metric proposed in [5] is sensitive enough to measure the intensity of **skating artifacts**.

## Definition: Foot Skating Artifacts

- Foot skating artifacts refer to **unnatural foot sliding when in contact with the ground.** (see Figure 1).
- Mainly appears in regression models trained to minimize the reconstruction error.
- Poorly designed networks suffer from mean pose regression, resulting in unnaturally rigid motion and induced foot skating.

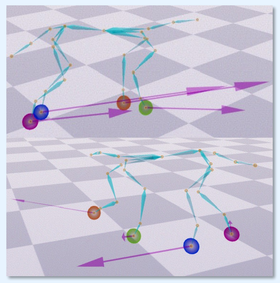

Figure 1. Visualization of foot skating artifacts. The purple arrow represents the foot velocity. The top motion is polluted with a higher degree of skating intensity than the bottom. Scan the QR code to visualize motion samples polluted by foot skating.

## Datasets

- Deep neural animation model to animate quadruped [7].
- The final loss is inversely proportional to $h_{size}$ and the skating is more intense in lower $h_{size}$ models.
- Scan the QR code in Figure 1 to see the output animations.

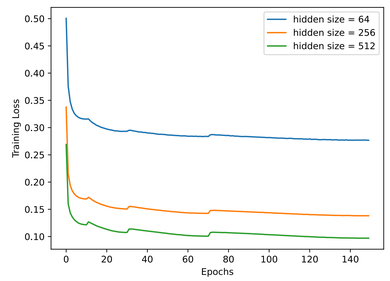


Figure 2. Training curves when reducing the number of parameters of the original model ($h_{size} = 512$). Decreasing the number of parameters of the hidden layer has a negative impact on the resulting motion and increases the skating intensity.

## Method

- **Unsupervised** method to learn motion underlying patterns (no labeled data).
- *Resnet34-based* autoencoder as in [5].
- Since the autoencoder is based on 2D convolutional layers, the motion samples are first **converted into an image representation**.

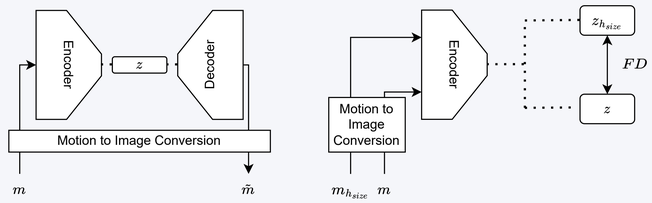


Figure 3. Overview of the method designed in [5]. Top: training procedure. - Bottom: Usage to compute $FD$-based score. The $FD$ is computed by Equation 1 from the statistics of the latent spaces of ground truth and generated motion data.

### Motion to Image Conversion

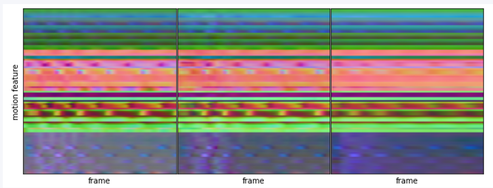


Figure 4. Image representation of motion data. From left to right: motion produced by the model with $h_{size} = 512, 256$ and $64$.

## Results

Table 1 shows the result of these analyzes. The mean and covariance matrix is computed for the whole set of animation samples produced by each deep generative model ($h_{size} = 512, 256$ and $64$). The $FD$ is computed between these statistics and those from the ground truth dataset.

| Generation model size ($h_{size}$) | $FMD \downarrow$ |
|---|---|
| 64 | 89.99 |
| 256 | **83.18** |
| 512 | 101.60 |

Table 1. Fréchet Motion Distance evaluation of motion generated by model with $h_{size} = 64, 256$ and $512$ (lower is better). The FMD in [5] is not proportional to foot skating intensity

This table shows that the $FD$ is **not able to capture the intensity of the foot skating artifacts** in a synthesized motion dataset.

## Perspectives

- More powerful encoding-decoding process involving more recent architectures such as transformers [6].
- Analyzing the latent space structure and keeping the similarity between the original and latent dimension [1].
- In metric analysis, involving human ratings on the motion dataset quality.

## Conclusion

This work presents an analysis of a metric designed to evaluate the quality and diversity of synthesized motion datasets. It highlights one of the limitation of this metric, which is its insensitivity to foot skating artifacts, a common anomaly produced by deep motion-generative models.

## References

[1] Imtiaz Ahmed, Travis Galoppo, Xia Hu, and Yu Ding. Graph regularized autoencoder and its application in unsupervised anomaly detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(8):4110–4124, 2021.

[2] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.

[3] Yong Li, Zihang He, Xiang Ye, Zuguo He, and Kangrong Han. Spatial temporal graph convolutional networks for skeleton-based dynamic hand gesture recognition. *EURASIP Journal on Image and Video Processing*, 2019(1):1–7, 2019.

[4] Antoine Maiorca, Nathan Hubens, Sohaib Laraba, and Thierry Dutoit. Towards lightweight neural animation: Exploration of neural network pruning in mixture of experts-based animation models. *arXiv preprint arXiv:2201.04042*, 2022.

[5] Antoine Maiorca, Youngwoo Yoon, and Thierry Dutoit. Evaluating the quality of a synthesized motion with the fréchet motion distance. In *ACM SIGGRAPH 2022 Posters*, SIGGRAPH '22, New York, NY, USA, 2022. Association for Computing Machinery.

[6] Mathis Petrovich, Michael J Black, and Gül Varol. Action-conditioned 3d human motion synthesis with transformer vae. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10985–10995, 2021.

[7] He Zhang, Sebastian Starke, Taku Komura, and Jun Saito. Mode-adaptive neural networks for quadruped motion control. *ACM Trans. Graph.*, 37(4), jul 2018.

## Contact Authors

**Antoine Maiorca**: antoine.maiorca@umons.ac.be
**Youngwoo Yoon**: youngwoo@etri.re.kr
**Thierry Dutoit**: thierry.dutoit@umons.ac.be

ACM International Conference on Interactive Media Experiences (IMX '23), June 12–15, 2023, Nantes, France