

Evaluating the Quality of a Synthesized Motion with the Fréchet Motion Distance

PRESENTER: Antoine Maiorca

PROBLEM

- Evaluate** the performance of models generating motion :
- Relies on human subjects
 - Lacks of reproductibility
 - Resources consuming process

Development of objective metric that evaluates **at low cost** how good the synthesized motion are.

METHOD

- 1) Motion to **image** conversion
- 2) Image encoded into a **latent space** with autoencoder model
- 3) **Fréchet distance** between latent spaces of generated and ground truth motion dataset

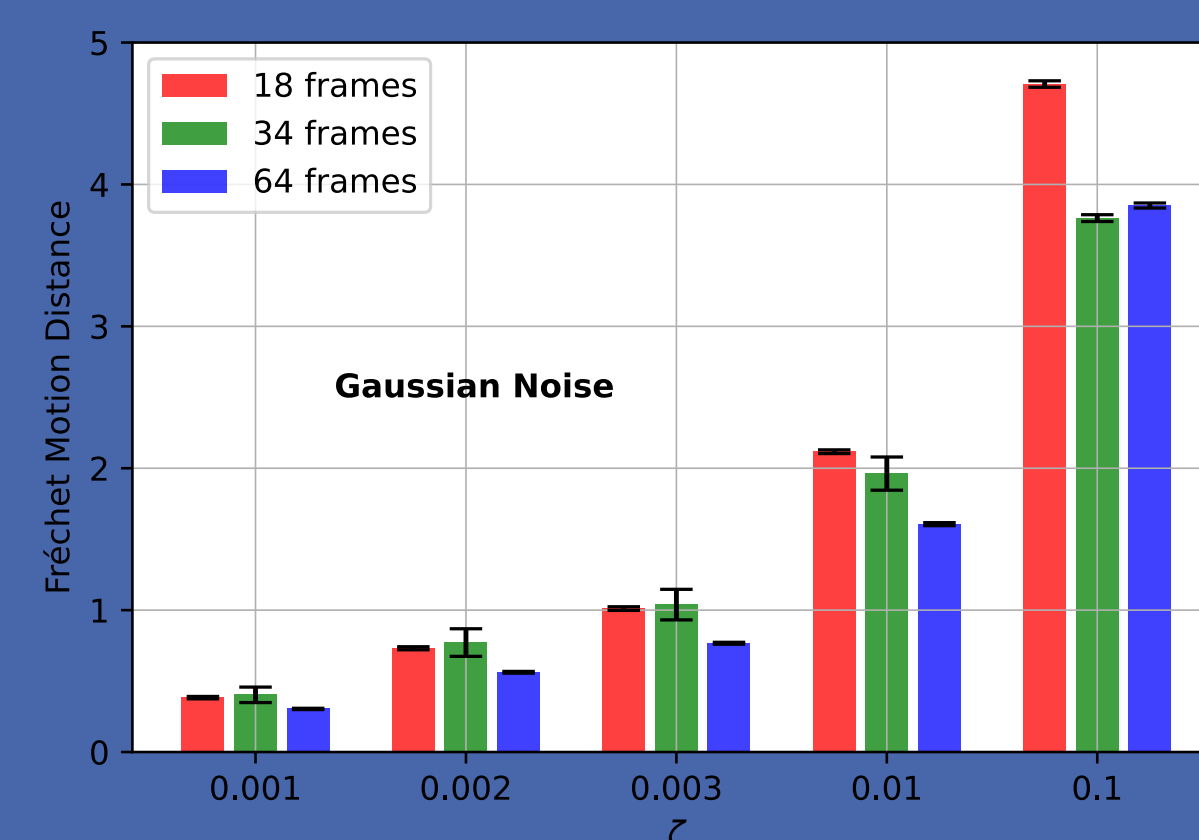
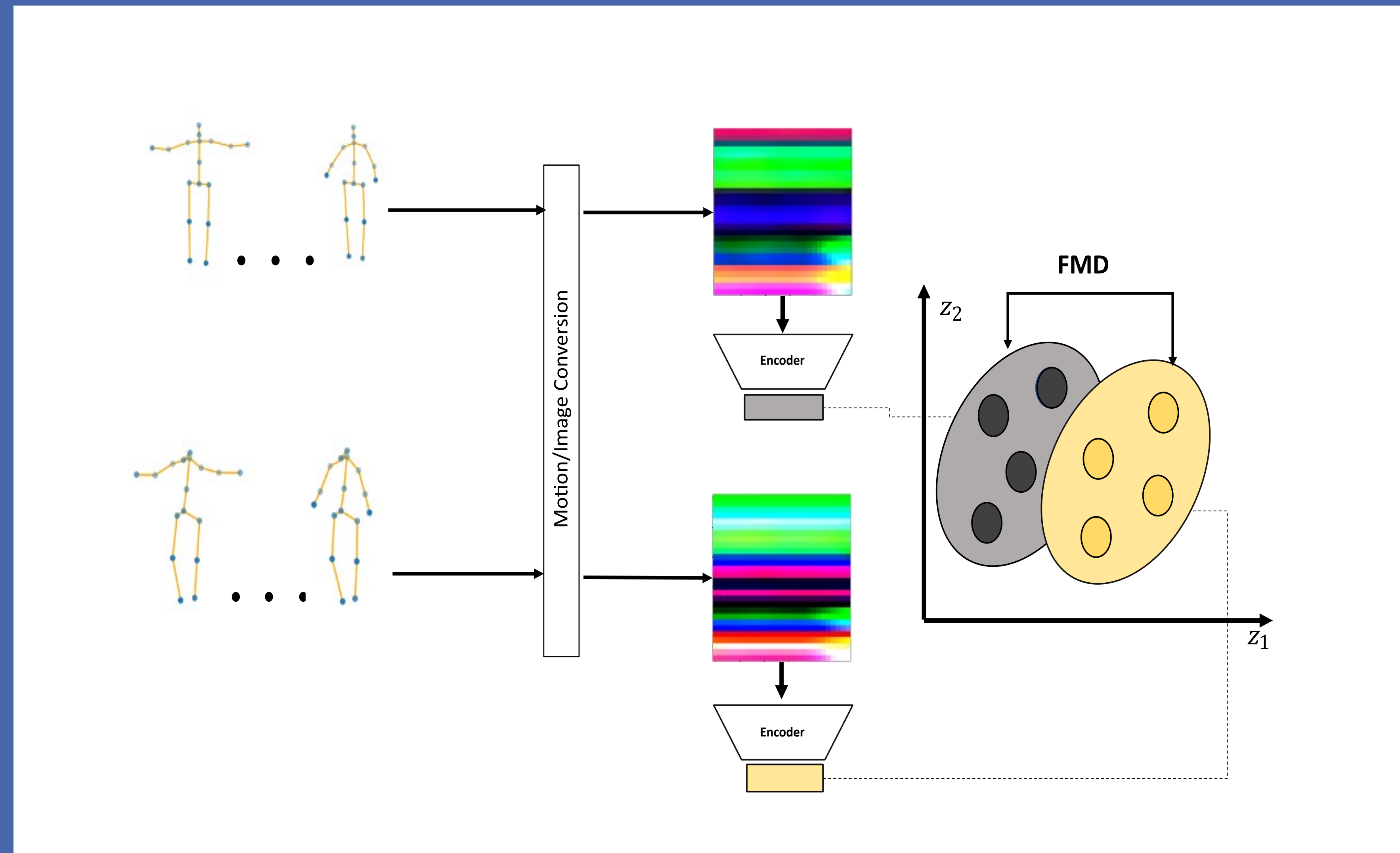
$$FMD = \|\mu_g - \mu_r\|_2^2 + Tr(\Sigma_g + \Sigma_r - 2\sqrt{(\Sigma_r \Sigma_g)})$$

RESULTS

Higher FMD means less qualitative and diverse synthesized motion

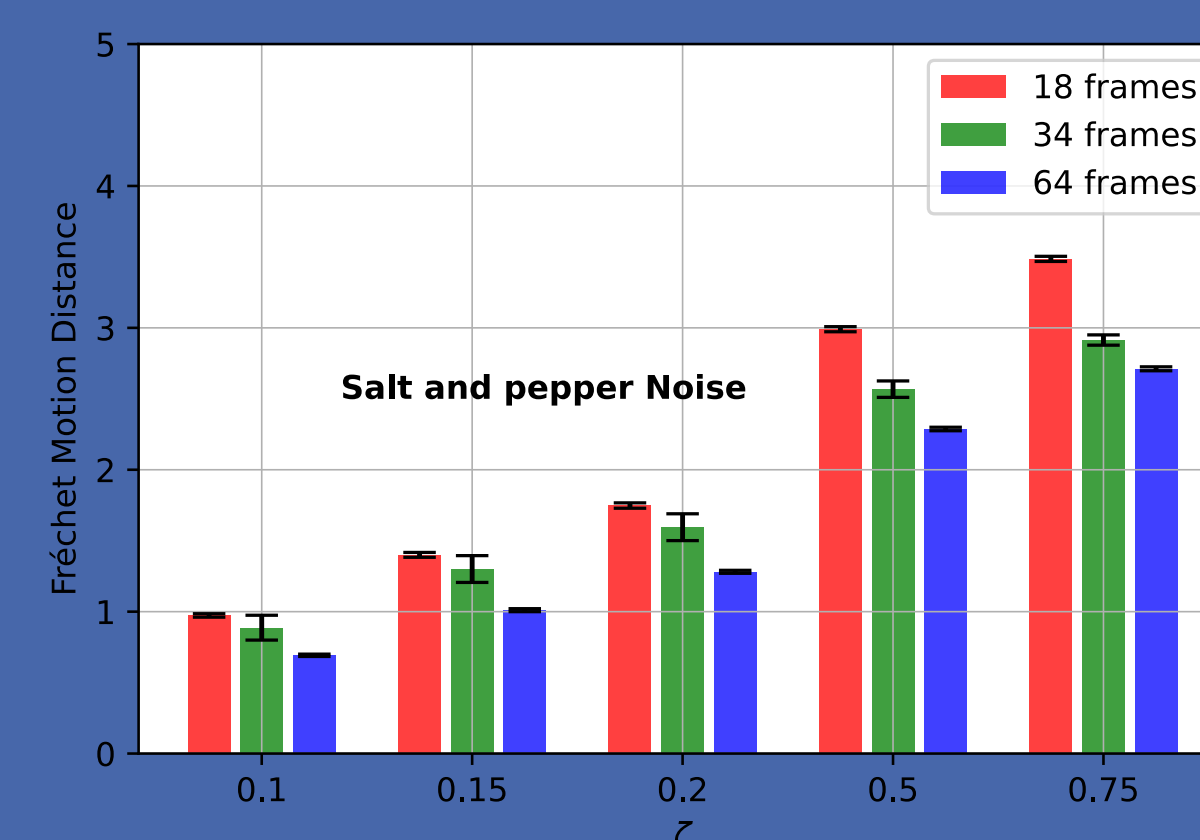
- FMD increases with noise intensity
- FMD trends with noise intensity is kept regardless motion length
- FMD captures less efficiently temporal discontinuities

Fréchet Motion Distance captures motion degradation and evaluates the performance of generative models



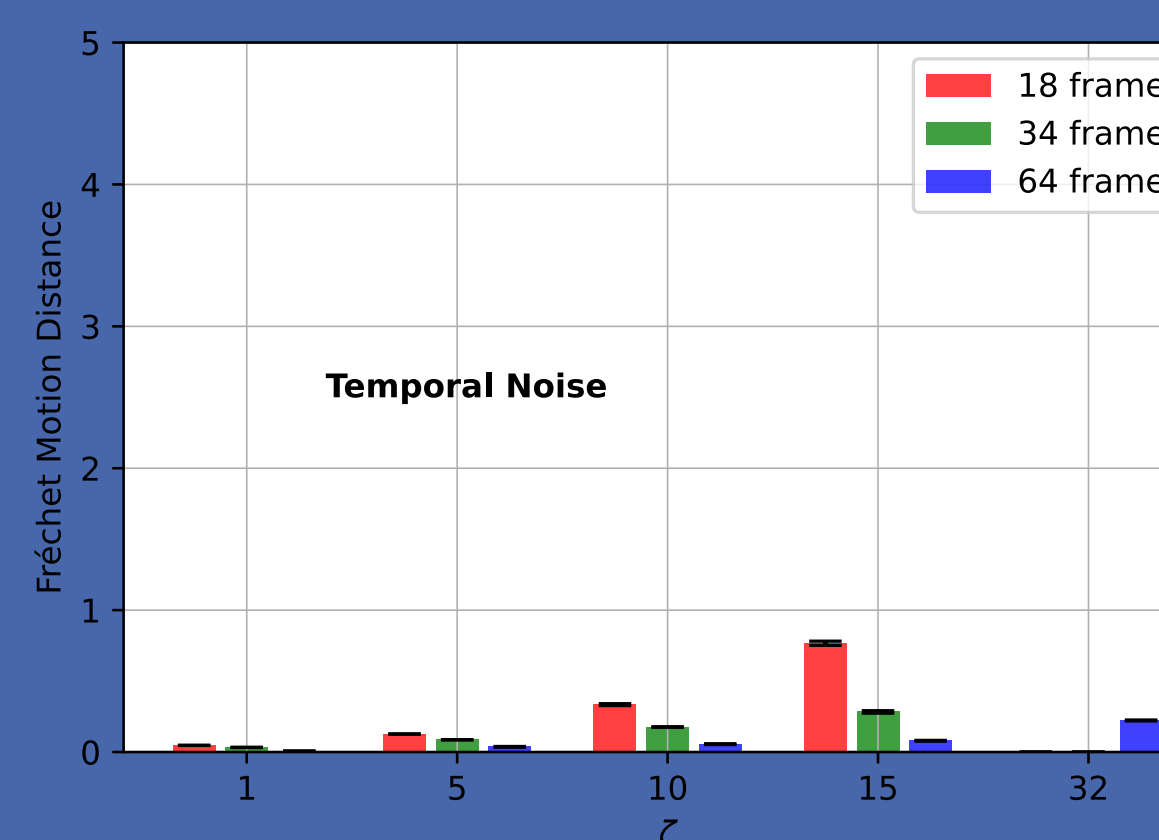
Normal noise added to motion

$$n_j \sim \mathcal{N}(0, \zeta)$$



Impulsive noise added to motion

$$n_j = \begin{cases} 0.2 & \text{if } \zeta/2 < u \leq \zeta \\ -0.2 & \text{if } u \leq \zeta/2 \\ 0 & \text{otherwise} \end{cases} \quad u \sim \mathcal{U}(0, 1)$$



Normal noise added to f motion frames (induces **discontinuity**)

$$n_j^f = \begin{cases} \mathcal{N}(0, 0.003) & \text{if } r \leq f < r + \zeta \\ 0 & \text{otherwise} \end{cases}$$

Benefits :

- **Unsupervised** fine-tuning (requires no labeled data)
- Latent space can be pushed to a **Gaussian distribution** by variational encoding (fits the Gaussian hypothesis of FID)

OUR APPROACH

Fréchet Motion Distance (FMD) measures the distance between the distribution of ground truth and synthesized motion. FMD evaluates the synthesized motion **quality** and **diversity** compared to the ground truth motion dataset.

RELATED WORK

Fréchet Distance has been applied in motion evaluation but no further validation were performed [1].

- [2] investigates the validation process of such metric but:
- Limited to upper body gesture
 - Sensitive to gesture length

REFERENCES

[1] W. Xi, G. Devineau, F. Moutarde, and J. Yang. 2020. Generative Model for Skeletal Human Movements based on conditional DC-GAN applied to pseudo-images. *Algorithms* 13, 12 (2020), 319.

[2] Y.Yoon, B.Cha, J-H. Lee, M. Jang, J. Lee, J.Kim, and G. Lee. 2020. Speech Gesture Generation from the Trimodal Context of Text, Audio, and Speaker Identity. *ACM Transactions on Graphics* 39, 6 (2020), 1–16

CONTACT

Authors

antoine.maiorca@umons.ac.be
youngwoo@etri.re.kr
thierry.dutoit@umons.ac.be

