Mohammed VI Polytechnic University (UM6P)
International Artificial Intelligence Center of Morocco (Ai movement)

3D Human Motion Generation

Amal El Fallah Seghrouchni, Naima Otberdout

Key-words: 3D human motion, 3D representations of human body, deep learning for 3D.

1 Research Motivation

Modeling and generating realistic human behavior is a long-standing objective in computer vision. As a matter of fact, the ability to model and reproduce the deformation patterns of the human body in the 3D world has a very high potential impact for important application fields such as film-making, game development, human-computer interaction, robotics, etc. For example, this will help to drive virtual characters capable of mimicking human motions by moving naturally and responding to environmental stimuli.
Nowadays, the emergence of highly-accurate 3D scanning devices makes of the 3D data prominent representations and opens the door to new investigations for human motion generation in 3D. Working on this project will be an added value for developing new approaches that exploit the 3D data and deep generative models to synthesize realistic and diverse 3D and 4D (*i.e.,* 3D+t) human behavior.

2 Problem Statement

Generating and reproducing the human behavior is an open problem that involves several challenges. Indeed, tackling this problem necessitates to come up with suitable representations that capture and model both the shape deformation as well as the dynamic variation in the 3D world. Traditional 3D data representations such as point cloud and meshes come with irregular structure that obstructs the use of deep learning models. In this context, two directions are explored in the recent literature; (1) extending deep learning model to deal with the irregular 3D data structure [1] (2) or investigating more suitable 3D representations that can be learned with standard deep learning models [2].
In this thesis we aim to explore new approaches that tackle the above challenges to respond to the main question of the problem of human motion synthesis that is; How we can generate 3D human body motions that are perceptually realistic while being diverse to cover the wide range of ways in which people carry out the same kind of actions ?
Beyond the studies that tackle the problem of human body motion generation in isolation [3, 4] more recent works [5, 6] begin to explore the influence of the surrounding scenes on human motion synthesis. This will be a promising direction to investigate in this thesis by producing motions of the human body while taking into account the constraints of the environment. This will lead to synthesize more realistic human motions.

3 Research Scope

The aim of this thesis is to explore new deep learning based approaches to model and generate realistic 4D (4D+t) human actions. As a first step, we need to conduct a state-of-the-art review to understand the achieved advance, the existing challenges and the promising directions that can be investigated. Next, a suitable representation of the 3D human motion need to be explored in order to model and capture the dynamic of the 3D shape variation. In addition, more investigations should be conducted to come up with new deep learning architectures that can deal with the non regular structure of the 3D human shape. The last part of this thesis will be interested in generating motion of human body interacting with the environment (*i.e.,* objects and other humans).

4 Admission Criteria

The PhD position is available at Ai movement, the International Center for Artificial Intelligence of Morocco of UM6P. Applicants with excellent academic credentials must be holders of a Master's, an engineering or an equivalent recognized degree with good skills in applied mathematics, in relation to optimization, operations research, and machine learning. The candidate should also be excellent in programming in (Python, Java or C++), should have soft skills, and be fluent in English and French languages. Letters of recommendation are welcome.

References

[1] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[2] Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3d reconstruction in function space. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4460–4470, 2019.

[3] Zhan Xu, , Yang Zhou, Li Yi, and Evangelos Kalogerakis. Morig: Motion-aware rigging of character meshes from point clouds. In *Proc. ACM SIGGRAPH AISA*, 2022.

[4] Chuan Guo, Shihao Zou, Xinxin Zuo, Sen Wang, Wei Ji, Xingyu Li, and Li Cheng. Generating diverse and natural 3d human motions from text. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5152–5161, 2022.

[5] Jingbo Wang, Yu Rong, Jingyuan Liu, Sijie Yan, Dahua Lin, and Bo Dai. Towards diverse and natural scene-aware 3d human motion synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20460–20469, 2022.

[6] Kaifeng Zhao, Shaofei Wang, Yan Zhang, Thabo Beeler, and Siyu Tang. Compositional human-scene interaction synthesis with semantic control. *arXiv preprint arXiv:2207.12824*, 2022.