Master 2 Internship subject:

Deep Learning Frameworks for Generative Models of 4D Human

Hosting institute

ICube Laboratory (Le laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie : The Engineering science, computer science and imaging laboratory) at the University of Strasbourg is a leading research center in Computer Science, with more than 300 permanent researchers, with the recently opened AI graduate school supported by the French government.

Work place and salary

The internship will take place in CAID (Computer-Aided Intervention & Design, 2021–) research team located at the hospital site of the laboratory, 10 min walking distance to the city center of Strasbourg, which is a UNESCO world heritage site.

Salary: 500€/month approximately for a duration of 5-6 months.

Staring date

December 2020 – February 2021 (adjustable).

Supervisors

director: Hyewon Seo (ICube, Univ. Strasbourg), co-supervisor: Cédric Bobenrieth (ICube, ECAM)

Context

Recently, there is a huge interest in applying deep learning techniques for synthesizing novel data from the learned model. It is true also for the human shape and motion data, for which several deep learning approaches have been proposed. Examples include a feedforward neural network that maps high level control parameters to the low level human motion over a manifold space found by a convolutional autoencoder [HSK16], CNN-based architecture combined with deep correlated 2D features for full shape recovery from image silhouettes [DJOZ+17], auto-conditioned recurrent neural networks to synthesize arbitrary motions with highly complex styles [ZLXH +18], RNNs (recurrent neural networks) trained for time-series prediction on shape and pose change [FLFM15, SL20, ZS21] during animation, Phase-Functioned neural network which takes the geometry of the scene into account to produce character motion along a user-defined path [HKS17], networks that can produce a distribution of next-state predictions in the context of character motion generation [HHSY+17, LZCV20], among others.

In this internship, we will focus on generative models of new types of data, 4D human, i.e. 3D human shape data under motions. The challenging problem of high spatiotemporal dimension of data, physical/environmental constraints, and user-defined controls will be addressed, along with the architectures of deep neural networks that can handle long sequences without an accumulation of errors.

Objectives

The objective of this internship is to develop deep-learning frameworks for the generation of realistic and controllable 4D human models. Given the user-controllable goal (task, style, constraints, etc), the trained network

should be able to generate the desired model in real-time. There are several ways to approach the problem, depending on the representation of dataset, the choice of the network architecture, and the types of goals and the way they are specified/controlled by the user. As for the network architecture, we will focus on the combinations of RNN and variational autoencoder, allowing a stochastic prediction of shape- and pose-sequences in a latent space. Several preprocessing of datasets from different sources may be required, in order to homogenize them into a uniform representation for the training. Different data representations and network hyperparameters will be experimented, to obtain the best results. Evaluation and comparison of the performance to the state-of-the-art methods is strongly recommended, whenever applicable.

Candidate profile

— Master student in Computer Science or in (Applied) Mathematics

— Solid programming skills in deep learning platforms: Tensorflow/Pytorch

— Background in geometric modeling and statistics

— Good communication skills

Application

Send your CV and your academic transcripts (Bachelor and Master courses) to seo@unistra.fr.

References

[DJOZ+17] Dibra, Endri & Jain, Himanshu & Oztireli, Cengiz & Ziegler, Remo & Gross, Markus. (2017). Human Shape from Silhouettes Using Generative HKS Descriptors and Cross-Modal Neural Networks. 5504-5514. 10.1109/CVPR.2017.584.

[FLFM15] Fragkiadaki, Katerina & Levine, Sergey & Felsen, Panna & Malik, Jitendra. (2015). Recurrent Network Models for Human Dynamics. 4346-4354. 10.1109/ICCV.2015.494.

[HHSY+17] Habibie, Ikhsanul & Holden, Daniel & Schwarz, Jonathan & Yearsley, Joe & Komura, Taku. (2017). A Recurrent Variational Autoencoder for Human Motion Synthesis.

[HKS17] Holden, Daniel & Komura, Taku & Saito, Jun. (2017). Phase-functioned neural networks for character control. ACM Transactions on Graphics. 36. 1-13. 10.1145/3072959.3073663.

[HSK16] Daniel Holden, Jun Saito, and Taku Komura. 2016. A deep learning framework for character motion synthesis and editing. ACM Trans. Graph. 35, 4, Article 138 (July 2016), 11 pages.

[LZCV20] Hung Yu Ling, Fabio Zinno, George Cheng, and Michiel Van De Panne. 2020. Character controllers using motion VAEs. ACM Trans. Graph. 39, 4, Article 40 (July 2020), 12 pages.

[SL20] Hyewon Seo and Guoliang Luo, Generating 3D Facial Expressions with Neural Networks, to appear as a book chapter in “AI Techniques for Scene Recognition and Modelling”, Springer, 2020.

[ZLXH+18] Y Zhou, Z Li, S Xiao, C He, Z Huang, H Li, Auto-conditioned recurrent networks for extended complex human motion synthesis, International Conference on Learning Representations 2018.

[ZS21] Kaifeng Zou and Hyewon Seo, DSNet: Dynamic skin deformation prediction by Recurrent Neural Network, under submission.