



Offer #2022-05186

PhD Position F/M Deep interactive control of virtual character's motion based on separating identity, motion and style (Inria/InterDigital Ys.ai project)

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : PhD Position

About the research centre or Inria department

The Inria Rennes - Bretagne Atlantique Centre is one of Inria's eight centres and has more than thirty research teams. The Inria Center is a major and recognized player in the field of digital sciences. It is at the heart of a rich R&D and innovation ecosystem: highly innovative PMEs, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institute, etc.

Context

Inria and InterDigital recently launched the Nemo.ai lab dedicated to research on Artificial Intelligence (AI) for the e-society. Within this collaborative framework, we recently initiated the Ys.ai project which focuses on representation formats for digital avatars and their behavior in a digital and responsive environment, and are looking for several PhDs and post-docs to work on the user representation within the future metaverse.

This PhD position will focus on exploring, proposing and evaluating novel solutions to represent both body shape and movements in a compact latent representation. This representation aims at simplifying the adaptation of the shape (identity) of a user, or/and his motion, and/or the style of both his shape and motion (such as transferring the user's moving shape to a fictional character with different properties and style).

Assignment

For its current and future standard video and immersive activities, Interdigital is aiming at providing semantic-based data solutions for videoconference and Metaverse applications. The goal is to stream data enabling the editability, controllability and interactivity of the content, while keeping the data throughput low to enable the use of existing and coming networks.

Character animation has a large set of potential applications, in videogames, movie industry, sports, rehabilitation, ergonomics, training, etc. To capture a 3D human shape in motion, two options are available at the moment: either a direct acquisition of the surface mesh thanks to a calibrated multi-camera setup, or a skinning technique from the animated skeleton. However, being able to capture and reproduce the expressivity of a human motion is difficult, as it involves several subtle parameters, some of them being lost when modeling human performance as joint angles: angles, contacts on the body surface, velocity profiles, accelerations, distance or coordination between body parts, etc. With the development of robust body shape reconstruction based on cheap sensors, such as RGB cameras or depth sensors, directly manipulating the visible surface of the character has become a very active field of research.

With the growing interest in persistent shared virtual worlds, such as the MetaVerse immersive social network, specific problems for character animation are raised. Indeed, in these environments, users are represented by avatars with different shapes and morphologies. Compared to the face, which has been studied for decades, there is no semantic controller for the body mesh, where one could easily change the motion type and style. The character animation platform should consequently be able to adapt the motion of the user to his/her specific shape (retargeting problem), or adapt the identity of the avatar so that the user is recognizable by his/her friends, or change the style of the motion to convey a given emotion or adapt to the expected behavior of the avatar. For example, a Hulk avatar is expected to move with a specific style, but should also mimic the characteristics of the user. Finally, the distribution of these avatar models over the network is a practical challenge due to the potential scale of the shared virtual worlds. Therefore, learning a representation that allows for efficient transmission and dynamic editing has a high practical impact.

Main activities

Motion retargeting has been explored a long time ago, by satisfying hand-tuned constraints that the

character has to be preserved, and use inverse kinematics to adapt the joint angles accordingly [5]. Other works introduced the idea of Interaction Mesh to replace hand-tuned geometric constraints by automatically preserving distances between body joints [6]. However these methods do not take the body surface into account, whereas it conveys a lot of relevant information about the motion. Recent works suggested that transferring the pose from a source to a target character was an ill-defined problem [2]. Alternatively, they proposed to transfer the shape of the target character to the source character in its desired pose, while preserving his identity.

Other recent works using deep learning aimed at separating the identity and shape, especially in the RGB video space [1]. Other works obtained impressive results to change the style of a picture, to make a human seem older, change head orientation, or the color [4]. However this work is very difficult to transfer from 2D RGB videos to 3D shapes.

Identity transfer for mesh-based motion.

Recent works suggested that transferring the pose from a source to a target character was an ill-defined problem [3, 2]. Alternatively, they proposed to transfer the shape of the target character to the source character in its desired pose while preserving his identity, either using optimization [3] or deep learning approaches [2]. Unlike previous works such as adapting joint angle [5] or preserving body joints distance [6], transferring the shape does take body surface into account, leading to the preservation of body contact. Unfortunately, these approaches are currently limited to static poses and therefore not directly suitable for character animation and style transfer in the Metaverse. A possible first research direction is to extend these state-of-the-art contact preserving retargeting methods to motion instead of isolated poses. Tackling the dynamics of the animation will require identifying dynamic user characteristics, extracting them from references and transferring identity while generating a temporally coherent animated mesh.

Efficient latent representation of identity, motion and style.

Metaverse is likely to involve cooperation between many entities. Hence, a huge volume of different environments, animations, characters, user identity and style must be shared and exchanged between entities. Creating or learning efficient and consistent representations of these items is necessary. There are two main deep learning based approaches to compress 3D data: implicit neural representations such as [7] and autoencoders such as [8, 9]. Unfortunately, it is usually hard to manipulate the encoding learned by these models to edit or transfer identity, motion or style. A solution could be to leverage and/or adapt recent advances in GAN inversion, where an encoder is trained to yield a latent space that can easily be manipulated [11]. Another possible direction is to enable acquisition of consistent embeddings on heterogeneous devices. One possible approach is to train neural networks with varying complexity [12].

Retargeting for other animations

Retargeting through identity transfer could be applied to other types of animation, such as skeleton or multi-body based animations. This would open up the possibility to use these animation techniques for Metaverse rendering. Recently, parametric controllers for multi-body that enables body shape variation have been proposed [10]. However, this approach currently lacks any mean to transfer identity. It could benefit from techniques similar to [3, 2]

The PhD will be co-supervised by:

- Franck Multon, Inria Rennes, MimeTIC team
- Pierre Hellier, InterDigital Rennes
- Adnane Boukhayma, Inria Rennes, MimeTIC team
- François Schnitzler, InterDigital Rennes

For more information, please contact: Franck Multon (fmulton@irisa.fr) or Pierre Hellier (Pierre.Hellier@InterDigital.com)

References

- [1] Kfir Aberman, Yijia Weng, Dani Lischinski, Daniel Cohen-Or, and Bao-quan Chen. Unpaired motion style transfer from video to animation. *ACM Trans. Graph.*, 39(4), jul 2020.
- [2] Jean Basset, Adnane Boukhayma, Stefanie Wuhler, Franck Multon, and Edmond Boyer. Neural human deformation transfer. In *2021 International Conference on 3D Vision (3DV)*, pages 545–554, 2021.
- [3] Jean Basset, Stefanie Wuhler, Edmond Boyer, and Franck Multon. Contact preserving shape transfer: Retargeting motion from one shape to another. *Computers & Graphics*, 89:11–23, 2020.
- [4] Amit H. Bermano, Rinon Gal, Yuval Alaluf, Ron Mokady, Yotam Nitzan, Omer Tov, Or Patashnik, and Daniel Cohen-Or. State-of-the-art in the architecture, methods and applications of stylegan, 2022.
- [5] Michael Gleicher. Retargeting motion to new characters. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques, SIGGRAPH '98*, pages 33–42, New York, NY, USA, July 1998. Association for Computing Machinery.
- [6] Edmond S. L. Ho, Taku Komura, and Chiew-Lan Tai. Spatial relationship preserving character motion adaptation. *ACM Transactions on Graphics*, 29(4):33:1–33:8, July 2010.

[7] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In European conference on computer vision, pages 405–421. Springer, 2020.

[8] Maurice Quach, Giuseppe Valenzise, and Frederic Dufaux. Learning convolutional transforms for lossy point cloud geometry compression. In 2019 IEEE international conference on image processing (ICIP), pages 4320–4324. IEEE, 2019.

[9] Danhang Tang, Saurabh Singh, Philip A. Chou, Christian Hane, Mingsong Dou, Sean Fanello, Jonathan Taylor, Philip Davidson, Onur G. Guleryuz, Yinda Zhang, Shahram Izadi, Andrea Tagliasacchi, Sofien Bouaziz, and Cem Keskin. Deep implicit volume compression.

In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.

[10] Jungdam Won and Jehee Lee. Learning body shape variation in physics-based characters. ACM Trans. Graph, 38(6):207:1–207:12, 2019.

[11] Xu Yao, Alasdair Newson, Yann Gousseau, and Pierre Hellier. Feature-Style Encoder for Style-Based GAN Inversion. Technical report, February 2022. ADS Bibcode: 2022arXiv220202183Y Type: article.

[12] Jiahui Yu, Linjie Yang, Ning Xu, Jianchao Yang, and Thomas Huang. Slimmable neural networks. In International Conference on Learning Representations, 2019.

Skills

The candidate must have MsC in computer sciences, with a focus either on machine learning, computer graphics or on virtual reality. In addition, the candidate should be comfortable with as much following items as possible:

- Deep learning
- Development of 3D/VR applications (e.g. Unity3D) in C# or C++.
- Character simulation and animation
- Geometry processing
- Evaluation methods and controlled users studies.
- Computer graphics and physical simulation.

The candidate must have good communication skills, and be fluent in English.

Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Possibility of teleworking (90 days per year) and flexible organization of working hours
- Partial payment of insurance costs

Remuneration

Monthly gross salary amounting to 1982 euros for the first and second years and 2085 euros for the third year

General Information

- **Theme/Domain** : Interaction and visualization
Software Experimental platforms (BAP E)
- **Town/city** : Rennes
- **Inria Center** : [Centre Inria de l'Université de Rennes](#)
- **Starting date** : 2023-01-01
- **Duration of contract** : 3 years
- **Deadline to apply** : 2023-03-31

Contacts

- **Inria Team** : [MIMETIC](#)
- **PhD Supervisor** :
Multon Franck / Franck.Multon@irisa.fr

About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

The keys to success

The candidate should like working in a team, should be curious and propose ideas and solutions. Creativity, is also an interesting key to success in such a research project.

Warning : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

Instruction to apply

Please submit online : your resume, cover letter and letters of recommendation eventually

Defence Security :

This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy :

As part of its diversity policy, all Inria positions are accessible to people with disabilities.