



Offre n°2022-05175

PhD Position F/M Deep-based semantic representation of avatars for virtual reality (Inria/InterDigital Ys.ai project)

Le descriptif de l'offre ci-dessous est en Anglais

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

A propos du centre ou de la direction fonctionnelle

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.

Contexte et atouts du poste

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 and facial animations with semantic-based approaches for the animation of avatars in a context of multi-user immersive telepresence.

Mission confiée

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.

So far, the core of InterDigital's technology is focused on the human face and already enables to extract facial parameters from an input video stream (head pose and facial expressions). These parameters are then encoded and streamed to a video AI-decoder capable of reconstructing a full and complete image. On its side, Inria is investigating new paradigms of animation for multi-user virtual reality experiences, and evaluating the impact the resulting animation quality can have in terms on the users' perception and behavior.

To advance future videoconference and metaverse applications, the main goal of this PhD is to explore novel approaches including both full body and facial elements, in particular by extending the current state of the art to enable full body encoding and decoding for multi-user immersive experiences and the evaluation of the quality of experience.

Principales activités

Leveraging deep learning methods to propose compact representations for avatar animations.

Realistic approaches for controlling the motions of virtual characters in interactive applications have recently emerged thanks to the use of Deep Learning. These recent advances are summarized in [Mourot et al. 2021], and include Phase-Functioned Neural Networks models [Holden et al. 2017], mixture-of-experts-based networks [Zhang et al. 2018, Starke et al. 2019, 2020], etc. However, such approaches have been hardly applied to the context of avatar control. Furthermore, in the context of massive multi-user experiences **hierarchical representations** would also be required. Such applications will require to provide versatile animation systems that can adapt to various devices, potentially from little tracking information (e.g., commercial systems are rarely able to fully capture the user motion, as they only track hand and head motions). Simultaneously, these systems also need to account for potential hardware limitations, such as tracking errors (e.g., noise, tracking losses), as well as limitations that can influence the amount of data to be transmitted (e.g., bandwidth, anonymity). Exploring these challenges will therefore require to propose novel methods based on recent deep learning approaches, tailored for the

specific case of avatar animations.

Ensuring plausible and realistic avatar animations when the semantic data stream is incomplete or corrupted. Controlling avatars' movements typically rely on simple animation techniques, e.g., Inverse Kinematics using the head and hand positions (3-point IK), sometimes including feet (5-point IK) and additional pelvis information (6-point IK). However, such simple animation techniques lead to visual artefacts that can be detrimental for realism and virtual embodiment, such as the well-known elbow or knee orientation problems rising from the ambiguity coming from the limited number of tracked joints. A few recent approaches are going in this direction, either by proposing upper-body VR-tailored IK approaches based on heuristics (i.e., not learned) [Parger et al. 2018] or by relying on deep learning models to predict lower-body poses from head, hands, and pelvis positions [Yang et al. 2021], but are still a long way from being able to generate high-quality motions for avatars in VR, with approaches designed with virtual embodiment in mind. As for previous work on faces, our goal is therefore to provide a unified approach providing several levels of editability, controllability and interactivity of the semantic content from partial information.

Evaluating generative avatar animation methods in a multi-user immersive context. With the development of Virtual Reality applications, avatars have become a major feature for improving the user experience, impacting both user performances [Rybarczyk et al. 2014] and their appreciations of these experiences [Yee and Bailenson 2007]. However, several factors typically impact how users accept their avatars as being their virtual representation in the virtual experience, which is often evaluated through the sense of embodiment [Kilteni et al. 2012]. Amongst these factors, several elements have already been identified as being particularly important to elicit a strong sense of embodiment, in particular the degree of realism of its appearance and animation controls [Argelaguet et al. 2016, Fribourg et al. 2020, Gorisse et al. 2017]. The last part of the project will therefore evaluate the performance of generative approaches for facial and body avatar animations in multi-user immersive applications, and the effect of the factors/parameters influencing their reconstruction on the client application on the user experience. Some of these questions relate to: What is the minimum information that needs to be available to represent a user in a shared application? Should some features be prioritized to others, e.g., facial features vs. body features? What are the novel representations that should be proposed to account for such a context? How can such representations provide an appropriate trade-off between realism and the volume of data required to be transferred to display and animate these avatars. What is the effect of displaying different levels of realism on different parts of the avatar (e.g., realistic appearance vs. low quality animations, or realistic facial animations vs. static hair or body).

References

- Holden, T. Komura, J. Saito. Phase-functioned neural networks for character control. *ACM Trans. Graph.* 36, 4, 2017.
- Mouro, L. Hoyet, F. Le Clerc, F. Schnitzler, P. Hellier. A Survey on Deep Learning for Skeleton-Based Human Animation. *Computer Graphics Forum*, 2021.
- Parger, J. Mueller, D. Schmalstieg, M. Steinberger. Human Upper-Body Inverse Kinematics for Increased Embodiment in Consumer-Grade Virtual Reality. *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, VRST '18*, 2018.
- Starke, H. Zhang, T. Komura, J. Saito. Neural State Machine for Character-Scene Interactions. In: *ACM Trans. Graph.* 38.6, 2019.
- Starke, Y. Zhao, T. Komura, Kazi Z. Local Motion Phases for Learning Multi-Contact Character Movements. In: *ACM Trans. Graph.* 39.4, 2020.
- Yang, D. Kim, S.H. Lee. LoBSTr: Real-time Lower-body Pose Prediction from Sparse Upper-body Tracking Signals. *Computer Graphics Forum* 40.2, pp. 265–275, 2021.
- Zhang, S. Starke, T. Komura, J. Saito. Mode-adaptive neural networks for quadruped motion control. *ACM Trans. Graph.* 37, 4, 2018.
- Argelaguet, L. Hoyet, M. Trico, A. Lécuyer (2016). "The role of interaction in virtual embodiment: Effects of the virtual hand representation". In: *2016 IEEE Virtual Reality (VR)*, pp. 3–10.
- Fribourg, F. Argelaguet, A. Lécuyer, L. Hoyet (2020). "Avatar and Sense of Embodiment: Studying the Relative Preference Between Appearance, Control and Point of View". In: *IEEE Transactions on Visualization and Computer Graphics* 26.5, pp. 2062–2072.
- Gorisse, O. Christmann, E. Armand Amato, S. Richir (2017). "First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and 110 Performance Analysis of Embodied Users". In: *Frontiers in Robotics and AI* 4, p. 33.
- Kilteni, R. Groten, M. Slater (2012). "The Sense of Embodiment in Virtual Reality". In: *Presence* 21.4, pp. 373–387.
- Rybarczyk, T. Coelho, T. Cardoso, R. F. de Oliveira (2014). "Effect of avatars and viewpoints on performance in virtual world: efficiency vs. telepresence". In: *EAI Endorsed Transactions on Creative Technologies* 1.1.
- Yee, J. Bailenson (2007). "The Proteus Effect: The Effect of Transformed Self-Representation on Behavior". In: *Human Communication Research* 33.3, pp. 271–290.

Compétences

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++.
- Evaluation methods and controlled users studies.
- Computer graphics and physical simulation.

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

Avantages

- 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

Rémunération

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

Informations générales

- **Thème/Domaine** : Interaction et visualisation
Plateformes expérimentales logiciel (BAP E)
- **Ville** : Rennes
- **Centre Inria** : [Centre Inria de l'Université de Rennes](#)
- **Date de prise de fonction souhaitée** : 2023-01-01
- **Durée de contrat** : 3 ans
- **Date limite pour postuler** : 2022-10-31

Contacts

- **Équipe Inria** : [MIMETIC](#)
- **Directeur de thèse** :
Hoyet Ludovic / ludovic.hoyet@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

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

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.