PhD subject:

4D Human Modeling for Real-World Human Cognition by Care-Robots for Aged People
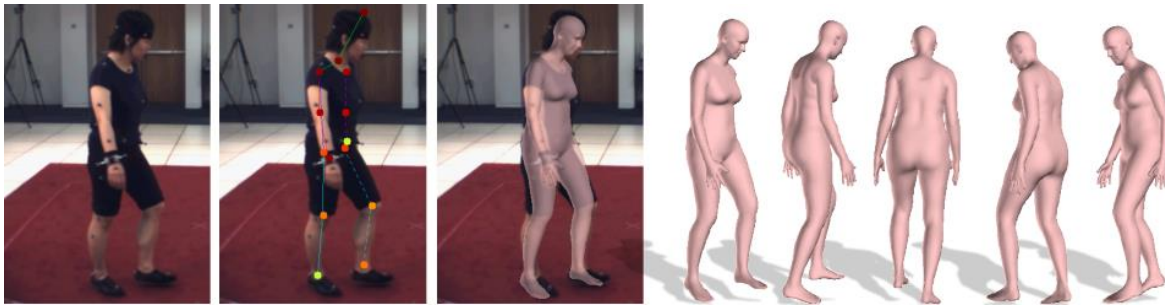

Figure 1: A model-based method can efficiently constrain the solution space for the human action recognition[1].

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 MLMS (Machine Learning, Modeling & Simulation) research team located at the hospital site of the laboratory, 10 min walking distance to the heart of the Strasbourg city center, which is a UNESCO world heritage site. One or two mid-term visiting research at ETRI (Daejeon, South Korea) is envisaged for a collaboration.

Salary: gross salary of 2135€ /month approximately for a duration of 3 years.

Supervisors

– director: Hyewon Seo (ICube, Univ. Strasbourg)

– co-supervisors: Frederic Cordier (Univ. Haut-Alsace), Stephane Cotin (INRIA Strasbourg)

Staring date

January 2022 – October 2022.

Context

Robot vision for human cognition often fails to work well in the real-world situation, despite the disruptive results achieved in Computer Vision and Artificial Intelligence. While most training data have been collected in well-conditioned, easy-to-isolate backgrounds, wild videos from the real-world may contain

---

[1] F. Bogo, A. Kanazawa, C. Lassner, P. Gehler, J. Romero, and M. J. Black. Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. In European Conf. on Computer Vision. Springer International Publishing, 2016.

various environmental conditions such as lighting, background patterns, and, most notoriously, occlusions. The latter becomes the source of recurrent problems of human cognition by care-robots in the in-house situation. Large variations in body shapes, motions, clothes, and frequent interactions with objects also contribute to the difficulty. Finally, viewing dynamics by moving robots and humans is another source of the problem. Learning-based methods relying on dataset inevitably show limited performance, as it is almost impossible to collect a large, annotated dataset that spans over all possible configurations of the real-world scene.

Model-based learning approaches are good alternatives to this problem, as have been confirmed by a considerable amount of previous works on human face and body recognition from images based on pre-constructed 3D models. However, many existing models consider only the geometry of human face or body in isolation, making it difficult to real-world situation involving obstacles or environmental objects. In this project[2], we will also adopt a model-based approach but with enhanced realism, additional dimension (i.e. time), and beyond. Our aim is to push the current limits of robot vision in human cognition by care-robots in the in-house situation. The specific goal is to make the performance of the vision-intelligence robust to large variations (in body shapes, motions,..) to occlusion (cloth, furniture, wall,..), and capable of understanding the interaction by developing a photo-realistic, physics-aware 4D human model.

Objectives

The proposed PhD study is articulated with three following tasks:

1. A photo-realistic 4D human model will be developed, by coupling a geometric human model and a color- and illumination components in a way that any desired shape, pose, and attribute can be generated. We will build our human model by grounding on our previous work as well as other standard geometric human models learned from multiple images of real people. Additional efforts will be made to consolidate the model by adding cloth and garment draping module, again based on our previous work.

2. The evolution of the human model over time due to motion will be learned by a deep neural network (DNN), over multiples of annotated motion datasets. We will focus on the combinations of RNNs and variational autoencoders, allowing a stochastic prediction of shape- and pose-sequences in a latent space. Different data representations and network hyperparameters will be experimented, to obtain best results.

3. An inverse problem will be formulated to reconstruct a 4D human model given an image, video, or a simple action command. All model components will be developed in a differentiable manner, so that an inverse problem can be formulated: Given an observation data, we will be able to build a model instance by an optimal model-fitting and even to predict its future state.

Candidate profile

− Master student in Computer Science, Electronic & Electrical Engineering, or in Applied Mathematics

− Solid programming skills: Python/C++

− Background in Geometric Modeling and Statistics

− Experience in Deep Learning

− Good communication skills

Application

Submit your application at "Candidate area" from https://emploi.cnrs.fr/.

---

[2] Real-World Human Cognition by Care Robots (Nov. 2021 − Nov. 2024), Bi-national project (with South Korea) with three institutional partners: CNRS, INRIA and ETRI (Electronics and Telecommunications Research Institute).