2019-01603 - PhD Position F/M Physical complex Interactions and Multi-person Pose Estimation

Type de contrat : CDD de la fonction publique
Niveau de diplôme exigé : Bac + 5 ou équivalent
Fonction : Doctorant

Contexte et atouts du poste

This PhD thesis is proposed in the context of the PIMPE project: Physical complex Interactions and Multi-person Pose Estimation. The project is funded by the IDEX of the Communauté Université Grenoble Alpes, under the International Strategic Partnerships program. Therefore, the PhD funding is granted and the student will spend half of the PhD in Grenoble and half of the PhD in Barcelona. In Grenoble, the PhD student will be hosted at the Perception Team of Inria/LJK (https://team.inria.fr/perception/). In Barcelona, the PhD student will be hosted at IRI-UPC (https://www.iri.upc.edu/).

Mission confiée

The estimation of the full human body pose is a paramount computer vision low-level task, potentially applicable to a wide variety of fields including the entertainment industry, sports technology, physical therapy and medical diagnosis. Seminal works grounded the estimation of the pose of a single person [1,2], which is still a very active line of research [3-8]. More recent works on human pose estimation can be roughly split into two categories. On the one side, methods targeting multi-person pose estimation in non-interactive scenarios [9-11], meaning people acting almost independently of each other. On the other side, there are a few works devoted to extreme/uncommon body poses [12], which could have significant impact in detecting abnormal situations (e.g. helping to diagnose patients with partial motion impairments or detecting fights in surveillance cameras) or on understanding team activities (e.g. in sport and art performances). Full body pose estimation, is often cast into a multivariate regression problem [13], for which robust (deep) regression techniques are the state-of-the-art [14,15].

Nevertheless, we identify a lack of methodological approaches being able to capture the complexity of visual human pose in multi-person interactive scenarios. In other words, current methods do not provide any means for the analysis or synthesis of images describing multi-person complex interactions that are faithful and realistic in terms of human body pose. We hypothesize that this is in part due to the lack of large training sets (manually) annotated for such scenarios, and therefore to the impossibility of training learning models in general and deep neural architectures in particular.

In this project we aim to develop machine learning techniques able to jointly estimate the full body pose of several persons involved in the same physical interaction. We are interested in a plethora of applicative scenarios, from industrial worker cooperation, to sports and arts as well as anomaly detection. To overcome the data availability issue, we propose to design learning strategies for the generation of realistic and controllable multi-person full-body extreme pose datasets: realistic so as to minimize the data distribution gap between training and real scenarios, controllable so as to generate a wide variability of poses, thus allowing for generalization to complex body motions and inter-body physical interactions.

In order to tackle the main challenges of the project, we will consider two important methodological tools. Firstly, the use of deep neural architectures, that demonstrated great potential on visual data generation [17,18,19], and whose evaluation for regression tasks is still an active area of research [20]. Second the use of resources available from the two partners, namely the Kinovis room at Inria, and a Xsens suit at UPC.

References


**Principales activités**

In the first months the student will have three main tasks: state-of-the-art on physical interaction modeling (directly related to CH1), software exploration and dataset acquisition, both useful for the entire PhD. First, the student will need to parse the literature to understand what is the state-of-the-art in modeling humans and their physical interactions, in natural scenarios related to the applications mentioned in the project’s proposal. This will provide the student a good understanding of the methodological choices present in the literature. Secondly, the student will need to explore the existing software, both well established libraries as well as related software projects publicly available on the cloud. This will provide the student the necessary tools and practical knowledge for the rest of the thesis. Third, the student will need to devote time to data acquisition, both at Inria/LJK exploiting the Kinovis platform, and at UPC-IRI using the Xsens suit.

**Compétences**

Research Master's degree, or equivalent, in a discipline connected to computer vision and machine learning. A particular interest/experience in computer vision tasks and machine learning problems, specifically with deep learning is a must. Strong motivation for the research work. Would both...
independently and to collaborate within a small team. Computer skills: MATLAB, Python, Deep Learning Toolkits (e.g. Keras, Pytorch). Oral and written English communication skills are mandatory.

Avantages
- Partial reimbursement of public transport costs
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Social security coverage

Rémunération
1768,55€/month before taxes