

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

Informations générales

- **Thème/Domaine :** Vision, perception et interprétation multimedia Systèmes d'information (BAP E)
- **Ville :** Montbonnot
- **Centre Inria :** CRI Grenoble - Rhône-Alpes
- **Date de prise de fonction souhaitée :** 2019-09-02
- **Durée de contrat :** 3 ans
- **Date limite pour postuler :** 2019-06-30

Contacts

- **Equipe Inria :** PERCEPTION
- **Directeur de thèse :**
Alameda-pineda Xavier /
xavier.alameda-pineda@inria.fr

A propos d'Inria

Inria, l'institut national de recherche dédié aux sciences du numérique, promeut l'excellence scientifique et le transfert pour avoir le plus grand impact. Il emploie 2400 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3000 scientifiques pour relever les défis des sciences informatiques et mathématiques, souvent à l'interface d'autres disciplines. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 160 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.

L'essentiel pour réussir

The PhD student will start on September 2019, for a period of three years (until August 2022). The PhD will be split in two halves, first at the Perception Team of Inria and LJK (until April 2021) and then at IRI-UPC. The PhD student will be co-supervised by Dr. Moreno-Noguer

(<https://www.iri.upc.edu/people/fmoreno/index.htm>)

and Dr. Xavier Alameda-Pineda (<http://xavirema.eu>)

with the support of Dr. Antonio Agudo

(<http://www.iri.upc.edu/people/aagudo/>)

and Dr. Radu Horaud

(<https://team.inria.fr/perception/team-members/radu-patrice-horaud/>).

Consignes pour postuler

Sécurité défense :

- [1] Yang, Y., Ramanan, D. Articulated pose estimation with flexible mixtures-of-parts. In IEEE CVPR, 2011.
- [2] Ferrari, V., Marin-Jimenez, M., Zisserman, A. Progressive search space reduction for human pose estimation. In IEEE CVPR, 2008.
- [3] Bulat, A., Tzimiropoulos, G. Human pose estimation via convolutional part heatmap regression. In ECCV, 2016.
- [4] Toshev, A., Szegedy, C. Deeppose: Human pose estimation via deep neural networks. In IEEE CVPR, 2014.
- [5] Bogo, F., Kanazawa, A., Lassner, C., Gehler, P., Romero, J., Black, M.J.: Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image. In ECCV 2016
- [6] Martinez, J., Hossain, R., Romero, J., Little, J.: A simple yet effective baseline for 3D human pose estimation. In ICCV, 2017.
- [7] **Moreno-Noguer, F.**: 3D human pose estimation from a single image via distance matrix regression. In IEEE CVPR, 2017.
- [8] Pavlakos, G., and K. G. Derpanis, X.Z., Daniilidis, K.: Coarse-to-fine volumetric prediction for single-image 3D human pose. In IEEE CVPR 2017.
- [9] Cao, Z., Simon, T., Wei, S. E., Sheikh, Y. Realtime multi-person 2D pose estimation using part affinity fields. In IEEE CVPR, 2017.
- [10] Insafutdinov, E., Pishchulin, L., Andres, B., Andriluka, M., & Schiele, B. Deepcrut: A deeper, stronger, and faster multi-person pose estimation model. In ECCV, 2016.
- [11] Papandreou, G., Zhu, T., Kanazawa, N., Toshev, A., Tompson, J., Bregler, C., & Murphy, K. Towards accurate multi-person pose estimation in the wild. In IEEE CVPR, 2017.
- [12] Akhter, I., Black, M.: Pose-conditioned Joint Angle Limits for 3D Human Pose Reconstruction. In IEEE CVPR, 2015.
- [13] Belagiannis, V., Zisserman, A. Recurrent human pose estimation. In IEEE FG, 2017.
- [14] Belagiannis, V., Rupperecht, C., Carneiro, G., Navab, N. Robust optimization for deep regression. In IEEE ICCV, 2015.
- [15] Lathuilière, S., Mesejo, P., **Alameda-Pineda, X., Horaud, R.** DeepGUM: Learning Deep Robust Regression with a Gaussian-Uniform Mixture Model. In ECCV, 2018.
- [16] Sadeghian, A., Alahi, A., & Savarese, S. Tracking the untrackable: Learning to track multiple cues with long-term dependencies. In ICCV, 2017.
- [17] Wang, W., **Alameda-Pineda, X.**, Xu, D., Fua, P., Ricci, E., and Sebe, N., Every smile is unique: Landmark-guided diverse smile generation, in IEEE CVPR, 2018.
- [18] Pumarola, A., **Agudo, A.**, Sanfeliu, A., and **Moreno-Noguer, F.**, Unsupervised Person Image Synthesis in Arbitrary Poses, in IEEE CVPR, 2018.
- [19] Pumarola, A., **Agudo, A.**, Martinez, A. M. , Sanfeliu, A., and **Moreno-Noguer, F.** , Ganimation: Anatomically-aware facial animation from a single image, in ECCV, 2018.
- [20] Lathuilière, S., Mesejo, P., **Alameda-Pineda, X., Horaud, R.** A Comprehensive Analysis of Deep Regression. IEEE TPAMI, 2019.

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

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.

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.

learning problems, specifically with deep learning is a must. Strong motivation for the research work. Ability to work 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