

## 2018-01170 - PhD Position F/M Learning-Based Human Character Animation Synthesis for Content Production

Contract type : Public service fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : PhD Position

### About the research centre or Inria department

Inria, the French national research institute for the digital sciences, promotes scientific excellence and technology transfer to maximise its impact.

It employs 2,400 people. Its 200 agile project teams, generally with academic partners, involve more than 3,000 scientists in meeting the challenges of computer science and mathematics, often at the interface of other disciplines.

Inria works with many companies and has assisted in the creation of over 160 startups.

It strives to meet the challenges of the digital transformation of science, society and the economy.

### Context

This PhD will be in the context of a CIFRE collaboration between Technicolor and the MimeTIC team (Inria Rennes). Technicolor is a leading company in the VFX world, combining their R&D expertise in Computer Vision and Computer Graphics with the artistic expertise from their studios, such as The Mill, Moving Picture Company, Mikros Image, etc. Inria is a French leading research centre in Computer Sciences, where research activities in MimeTIC focus on simulating virtual humans that behave in a natural manner and act with natural motions.

### Assignment

This PhD will be conducted in the context of a collaboration between Technicolor and Inria. The starting date of the PhD is flexible, and could be as soon as 1st of February 2019.

## Introduction and context

Content production for film and advertising increasingly relies on computer-generated imagery to lower costs and enhance creative possibilities. In particular, many of today's movies and advertisements feature synthetic human characters. The animation of the characters' bodies is driven by the dynamics of an underlying skeleton, built from the main joints of the human body. The skeleton is later fleshed into a 3D mesh by a process known as skinning, whereby the displacement of each vertex of the mesh is computed from the displacement of the neighbouring skeleton joints it is bound to. Accurately capturing the naturalness of human motion in the dynamics of the skeleton is key to the perceptual plausibility of the rendered animation.

Creating animations for photorealistic computer-generated movies is a highly demanding complex part of the film production workflow that requires an insane amount of manual work. Keyframing and motion capture are the two dominant techniques used in the industry today. Keyframing refers to a purely manual editing process wherein artists draw the skeletons at selected temporal frames ("keyframes"), and further define non-linear interpolation paths for joints locations in-between the keyframes. Motion capture is performed in a green room with specialized hardware, with marker-based setups that requires some involvement on the part of the actors, as well as manual post-processing to incorporate artistic edits into the animations. In both cases, the amount of human intervention and hence the production costs are very high. Thus, there is a strong business justification in the automation of the non-creative parts of the animation process.

Advances in machine learning and particularly deep learning in recent years have boosted the research effort towards obtaining skeletal animations from the analysis of videos. The idea is to learn a mapping between the image of a human character and the 2D or even 3D locations of the joints of the character body. However, due in part to the difficulty of the problem and in part to the lack of 3D annotated training data, the accuracy on joint location estimates is often poor, especially in the depth direction that is not observable in the image. Besides, the estimated skeletons consist of only a few joints and often fail to cover the hands and the feet.

The generation of animations from videos offers promising prospects for optimizing the animation workflow in the content production industry. Still, a lot of work is needed to improve the resolution and accuracy of the produced animations, and to adapt the technology to make it usable in an interactive way by animation artists. Advancing towards these goals is the main purpose of the proposed PhD.

### Main activities

## Existing techniques and limitations

The estimation of animation skeletons, a.k.a. human poses, in images and videos is an active research area, dominated by supervised machine learning approaches that leverage databases of images annotated with human joint locations. The initial target of 2D pose estimation [1] has now been extended to 3D, see for instance [2, 3]. Inferring the depth components of the skeleton joints turns out to be a challenging ill-posed problem. Even though various regularization strategies have been proposed, the estimated joint locations are still quite noisy, especially in the depth direction orthogonal to the plane of the observed image. This is also, to some extent, a consequence of the scarcity of 3D skeleton annotations, which are difficult to generate in "in-the-wild" environments [4]. A further issue with annotations, and as a result human pose estimates, is that they are limited to a

### General Information

- **Theme/Domain** : Interaction and visualization  
Instrumentation et expérimentation (BAP C)
- **Town/city** : Rennes
- **Inria Center** : CRI Rennes - Bretagne Atlantique
- **Starting date** : 2019-02-01
- **Duration of contract** : 3 years
- **Deadline to apply** : 2019-09-30

### Contacts

- **Inria Team** : MIMETIC
- **PhD Supervisor** :  
Hoyet Ludovic / ludovic.hoyet@inria.fr

### About Inria

Inria, the French national research institute for the digital sciences, promotes scientific excellence and technology transfer to maximise its impact. It employs 2,400 people. Its 200 agile project teams, generally with academic partners, involve more than 3,000 scientists in meeting the challenges of computer science and mathematics, often at the interface of other disciplines. Inria works with many companies and has assisted in the creation of over 160 startups. It strives to meet the challenges of the digital transformation of science, society and the economy.

### The keys to success

There you can provide a "broad outline" of the collaborator you are looking for what you consider to be necessary and sufficient, and which may combine :

- tastes and appetencies,
- area of excellence,
- personality or character traits,
- cross-disciplinary knowledge and expertise...

This section enables the more formal list of skills to be completed and 'lightened' (reduced) :

- "Essential qualities in order to fulfil this assignment are feeling at ease in an environment of scientific dynamics and wanting to learn and listen."
- " Passionate about innovation, with expertise in Ruby on Rails development and strong influencing skills. A thesis in the field of \*\*\*\* is a real asset."

### 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.

**Warning** : you must enter your e-mail address in order to save your application to Inria. Applications must

small number of body joints, excluding hands and feet. Overall, the accuracy and resolution of state-of-art "video to analysis" techniques is still unsuitable for animating even secondary characters in photorealistic films and movies.

In parallel to human pose estimation, some research effort has been devoted to the characterization of human motion kinematics using learning-based approaches. The seminal work of Holden [5] leverages an autoencoder framework to learn a "manifold" of human motion. It further proposes methods for editing animations in this manifold and mapping the editing controls to human-understandable high-level parameters. The learnt parameters of the encoder can be used to characterize the style of the motion and perform style transfer on animations. This technique could be extended to learn a specific motion model for a given character, perhaps based on initially produced animation sequences for this character, and further improve the generation of subsequent animations for this same character based on the learnt model.

## Directions for research

Directions of research are flexible within the proposed context, but will explore areas related to improving animation quality for production usages.

## References

- [1] 6. Newell, K. Yang and J. Deng, "Stacked Hourglass Networks for Human Pose Estimation," in *European Conference on Computer Vision*, 2016.
- [2] 7. Mehta, S. Sridhar, O. Sotnychenko, H. Rhodin, M. Shafiei, H. Seidel, W. Xu, D. Casas and C. Theobalt, "VNect: Real-Time 3D Human POse Estimation with a Single RGB Camera," *ACM Transactions on Computer Graphics*, vol. 36, no. 4, pp. 44:1 - 44:14, 2017.
- [3] 6. Tekin, A. Rozantsev, V. Lepetit and P. Fua, "Direct Prediction of 3D Body Poses from Motion Compensated Sequences," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [4] 7. Zhou, Q. Huang, X. Sun, X. Xue and Y. Wei, "Towards 3D Human Pose Estimation in the Wild: a Weakly-supervised Approach," in *IEEE International Conference on Computer Vision (ICCV)*, 2017.

### Skills

### Requirements for candidacy

- Strong programming skills (C/C++ recommended)
- Strong knowledge of machine learning
- Basic knowledge of computer animation and graphics

### Benefits package

- Subsidised catering service
- Partially-reimbursed public transport

### Remuneration

Monthly gross salary amounting to 1 982 euros for the first and second years and 2 085 euros for the third year

be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.