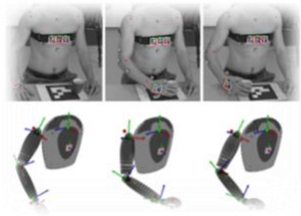
CALL FOR APPLICATION FOR A MASTER INTERNSHIP

REAL-TIME HUMAN JOINTS KINEMATICS ESTIMATION IN SPORTS APPLICATIONS USING A STEREO CAMERA AND DEEP LEARNING

ABSTRACT

The general objective of this internship is to develop and validate a new framework for the real-time determination of human pose performing sport, rehabilitation and industrial tasks. The proposed method will be based on several deep learning trained models, a constrained biomechanical model of the human, and a(several) low-cost RGB-camera(s).

The internship will open opportunities to pursue a PhD in robotics/biomechanics within the GEPETTO team at LAAS-CNRS.



PROJECT

This internship is a part of the SAVE-AI interdisciplinary project (https://miti.cnrs.fr/prime/save-ai/) that aims to analyze and enhance the performance of sports or industrial movements using a new video analysis framework. This system classifies the movement, estimates its kinematics and dynamics, and then suggests optimal postures to reduce injury risks. Applications include complex, dynamic, or industrial tasks, with an approach combining optimal control and deep learning.

Accurate kinematics estimation of human motion is the basis for motion dynamics estimation and movement classification, which can be useful in countless applications. In manufacturing industries it is crucial for the better understanding and the analysis of work related musculoskeletal disorders. These quantities should be known at the joint level, i.e. joint angles with the best accuracy to give direction to ergonomists and managing staff. These quantities cannot be directly measured but only estimated. They are usually obtained with a stereophotogrammetric system and one or two force plates, which provide accurate results; however, these equipment require considerable financial investment and a complex experimental protocol. The challenge for industrial applications to find easy-to-use and low-cost instruments and a method for straightforward interpretation of the numerical outcome, is a goal that has yet to be achieved [1]. For applications on the factory floor, it is desirable to avoid the use of embedded sensors so as to not disturb the workers movements.

Markerless approaches solely relying on the use of vision (RGB, RGB-D cameras) appear as a practical solution. The most popular and basic approach is the human pose estimation (HPE), which consists in estimating the 2D/3D position body keypoints from an image, generally using deep neural networks [2]. HPE can be performed in real-time and provides impressive results in terms of 2D keypoints identification [2]. However, measurements obtained from HPE (e.g. depth of the keypoints) are not reliable enough to accurately reconstruct the 3D joints kinematics [3]. One solution to this issue consists in using several calibrated cameras to perform a triangulation and obtain more reliable 3D keypoints coordinates. Still, even when using more than 6 cameras and a very controlled setup, Lahkar et al. obtained errors as high as 5deg for full body joints kinematics estimation [3]. Furthermore, HPE solely allows the identification of a sparse set of keypoints which cannot be used to fully reconstruct the complexity of 3D joints kinematics. To cope with these issues, recent approaches proposed to perform data augmentation through neural networks that are trained to estimate the position of more than 40 body markers, from the sparse and noisy set of triangulated landmarks [4]. This approaches claim to obtain an accuracy lower than 5deg when compared to a reference stereophotogrammetric system [4].

However, in its current state, it requires the use of two separated proprietary cameras, which can be cumbersome and requires a calibration step at each use. Moreover, the method does not implement a real-time pipeline for 3D joints kinematics estimation and is a black box for developers. Thus, the goal of this project is to develop a Python library (COSMIK, Open Source Markerless Constrained

CONTEXT

This 6 months internship will take place at LAAS-CNRS in the Gepetto team.

REQUIREMENTS

-MsC in Robotics/Machine Learning or related field

-Strong programming skills (Python)

- Pytorch/GitHub/ Caffe/ Onnx/TensorRT are a plus

-Ability to work in English

-Taste for experiments

LAAS CNRS

Inverse Kinematics) that relies on state-of-the-art HPE, data augmentation, and inverse kinematics algorithms for real-time estimation of 3D joints kinematics from a single stereo-camera. The intern will be assigned with the following tasks:

- Development of a real-time pipeline for 3D joints kinematics estimation from camera images. The pipeline will rely on the MMPose library which is a HPE algorithms aggregator. MMPose will then be integrated as a dependency to COSMIK. This will allow the continuous integration of most recent HPE algorithms in COSMIK. The student will integrate state-of-the-art triangulation and data augmentation methods in the pipeline. Finally, the intern will integrate real-time inverse kinematics algorithms using quadratic programming,
- Data acquisition during experiments with the whole pipeline, several low-cost stereo cameras and a gold standard motion capture system. The student will proceed to experiments with different stereo-camera models and will do a benchmark of the estimated joint kinematics from different camera models and the COSMIK pipeline, w.r.t. to reference joint angles obtained with the gold standard motion capture system.
- Data acquisition "on the pitch" using the camera chosen through the benchmark, with high-level athletes possibly from the Toulouse Football Club. This data will be processed through the COSMIK pipeline to feed a dataset with estimated joints kinematics, and labeled motions which will further be used for movement classification.

The first task was partially implemented by our team. Thus, the student will have to get familiar with what has already been developed and to integrate it in a proper way in COSMIK to allow further developments of the library. To validate the approach, it will be required to perform human experiments and create a new dataset composed of reference motion capture data, videos, HPE measurements, data augmentation measurements, and joint angles, both in controlled environment and outdoor "on the pitch" scenarios. An RGB video will be provided as input and processed by a Convolutional Neural Network (CNN), which will generate two distinct outputs through two separate branches. The network will need to capture more refined features and predict confidence maps for the location of various body parts, such as the right eye, left eye, right elbow, and others. This dataset could further be used to improve the approach, and also to develop movement classification algorithms. **The internship will open opportunities to pursue a PhD.**

REFERENCES

[1] Mohamed Adjel, Toward the development of a sparse, multi-modal and affordable motion analysis system. Applications to clinical motor tests. 2024

[2] Jiang et al., RTMPose: Real-Time Multi-Person Pose Estimation based on MMPose. 2023

[3] Lahkar et al., Accuracy of a markerless motion capture system in estimating upper extremity kinematics during boxing. 2022

[4] Uhlrich et al., OpenCap: Human movement dynamics from smartphone videos. 2023

CONTACT

Supervisor(s): Vincent Bonnet / Emmanuelle Claeys/ Kahina Chalabi / Mohamed Adjel

Email: vincent.bonnet@laas.fr / Emmanuelle.Claeys@irit.fr/ kchalabi@laas.fr/ madjel@laas.fr