A comparison of optical flow and skeleton based human activity recognition in omnidirectional images

Research Project or Master Thesis

June 23, 2022

Abstract

Figure 1: Example flow visualisation with superimposed keypoints from OmniFlow [1]

The professorship of Digital- and Circuit design (DST) is working in the field of Convolutional Neural Networks (CNNs) for various computer vision tasks. In this context the aim is to learn activities from a time series of motion vectors by temporal segment networks (TSNs) [2]. To detect 'activities' from motion vectors (optical flow) a evaluation is necessary. In this project, the optical flow based activities shall compared with the key-point related activities in omnidirectional images. For this, OmniFlow has to be used to compare both human activity recognition techniques.

The students work isn't limited to this work but should at least contain the following steps:

- Investigation in skeleton-based human activity recognition (HAR) approaches
- implementation of a method to automatically annotate skeleton points in MS COCO format (*.json) from existing armature in blender (bvh-files, e.g. from `https://sites.google.com/a/cgspeed.com/cgspeed/motion-capture/the-motionbuilder-friendly-bvh-conversion-release-of-cmus-motion-capture-database` in the OmniFlow dataset [1] OR
- using a annotation tool for the existing OmniFlow dataset for labeling skeleton points

- training of a 2D human pose estimation CNN (to be discussed in the project, framework: `https://github.com/open-mmlab/mmaction2`)
- evaluation of the results

Requirements

- basic understanding in computer vision
- at least a ‘good‘ result in Computer Vision I
- good programming skills in Python and Blenders Python-API bpy

Contact Information

If you’re interested in this topic please send a Mail to roman.seidel@etit.tu-chemnitz.de or +49 371 531 30527)

References

[1] Roman Seidel, André Apitzsch, and Gangolf Hirtz. Omniflow: Human omnidirectional optical flow. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3678–3681, 2021.

[2] Limin Wang, Yuanjun Xiong, Zhe Wang, Yu Qiao, Dahua Lin, Xiaoou Tang, and Luc Van Gool. Temporal segment networks: Towards good practices for deep action recognition. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision – ECCV 2016*, pages 20–36, Cham, 2016. Springer International Publishing.