Institut für Photogrammetrie
und GeoInformation

Leibniz Universität Hannover

Augmentation for End-to-End deep learning-based multiple-view multi-object tracking

Proposal for a Master thesis topic (EN)

Nienburger Straße 1, 30167 Hannover
Fakultät für Bauingenieurwesen und Geodäsie

Institut für Photogrammetrie
und GeoInformation

Prof. Dr.-Ing. habil. Christian Heipke

Rasho Ali

Tel. +49 511 762 2488
Fax +49 511 762 2483
E-Mail:
ali@ipi.uni-hannover.de

3D multi-view multi-object tracking (MV-MOT) plays a crucial role in various real-world applications, including pedestrian safety, autonomous driving, and sports analysis. 3D MV-MOT involves using images from multiple cam-eras with known intrinsic and extrinsic parameters to detect and track pedestrians in a scene. This task is typically performed in 3D, where the positions of pedestrians feet are represented in the bird's eye view (BEV), as shown in Fig 1. Although there is extensive research in the field of single-view multi-object tracking, there is limited research on MV-MOT. Moreover, there has been little to no research specifically focused on augmentation methods for enhancing MV-MOT performance.

In the framework of this master's thesis, the student will investigate augmentation methods used for multi-view multi-object detection, such as 3D Random Occlusion (Qiu, Rui, et al., 2022) and the view-coherent data augmentation method proposed by Hou and Zheng (2021), for the purpose of the MV-MOT task. The student will explore the impact of different augmentation techniques on 3D MV-MOT, particularly in challenging scenarios involving occlusion and variations in viewpoint.

In addition to the conceptual elaboration, this master's thesis also includes a practical component, where the student will test the different augmentation methods on an existing MV-MOT model. The tests will be conducted using the WildTrack dataset (Chavdarova, Tatjana, et al., 2018), which will be made available to the student. The evaluation of the integrated augmentations will be carried out using standard evaluation metrics for MV-MOT. The evaluation should also include an analysis of failure cases of the methodology, along with initial ideas for potential adaptations to overcome these limitations.

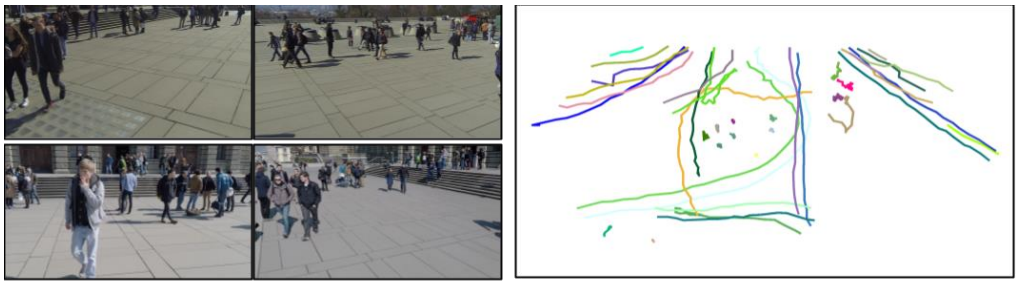

Figure 1. The input to the MV-MOT model and the expected output.

References

Qiu, Rui, et al. "3d random occlusion and multi-layer projection for deep multi-camera pedestrian localization." European Conference on Com-puter Vision. Cham: Springer Nature Switzerland, 2022. (pp. 695-710)

Hou, Yunzhong, and Liang Zheng. "Multiview detection with shadow transformer (and view-coherent data augmentation)." Proceedings of the 29th ACM International Conference on Multimedia. 2021. (pp.1673-1682)

Besucheradresse:
Nienburger Straße 1
30167 Hannover
www.ipi.uni-hannover.de

Chavdarova, Tatjana, et al. Wildtrack: A multi-camera hd dataset for dense unscripted pedestrian detection. In: Proceedings of the IEEE confer-ence on computer vision and pattern recognition. 2018. (pp. 5030-5039)