Benchmarking Human Motion Prediction Methods

Andrey Rudenko[*]
andrey.rudenko@oru.se

Tomasz Piotr Kucner
Tomasz.Kucner@oru.se

Chittaranjan Swaminathan
Chittaranjan.Swaminathan@oru.se

Ravi Chadalavada
raviteja.chadalavada@gmail.com

Kai O. Arras[†]
kaioliver.arras@de.bosch.com

Achim J. Lilienthal
achim.lilienthal@oru.se

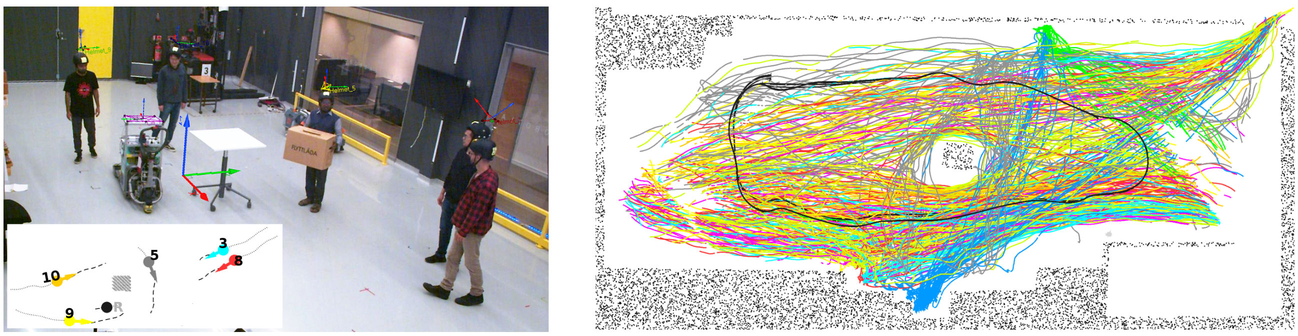

Figure 1: Data collection setup (left), visualisation of the subset of the collected trajectories (right).

ABSTRACT

In this extended abstract we present a novel dataset for benchmarking motion prediction algorithms. We describe our approach to data collection which generates diverse and accurate human motion in a controlled weakly-scripted setup. We also give insights for building a universal benchmark for motion prediction.

CCS CONCEPTS

• **Human-centered computing** → **Laboratory experiments**; *Interaction design process and methods*; • **Computer systems organization** → **Robotics**; • **General and reference** → *Evaluation.*

KEYWORDS

human motion prediction, benchmarking, datasets

ACM Reference Format:

Andrey Rudenko, Tomasz Piotr Kucner, Chittaranjan Swaminathan, Ravi Chadalavada, Kai O. Arras, and Achim J. Lilienthal. 2020. Benchmarking Human Motion Prediction Methods. In *Proceedings of HRI 2020 Workshop on Test Methods and Metrics for Effective HRI in Real World Human-Robot Teams (HRI'20).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

[*]A. Rudenko, T. Kucner, C. Swaminathan, R. Chadalavada and A. Lilienthal are with MRO Lab, Örebro University, Sweden

[†]A. Rudenko and K. Arras are with Bosch Corporate Research, Renningen, Germany

HRI'20, March 23, 2020, Cambridge, UK

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION: HUMAN MOTION IN HRI

Human motion plays a central role in human-robot interaction, especially for service robots, personal assistants and human-robot teams. It includes motion trajectories in navigation experiments [14], hand reaching motions or full body poses for collaborative robotics [7] action sequences [2], eye gaze directions [6], gestures, facial expressions etc. Robots operating in close proximity to humans benefit from processing such motion cues using a model of human motion, and reason about the future to improve safety and efficiency of collaborations and service activities. Data plays a key role in building such models, as it is used for hyperparameter estimation, motion patterns derivation or evaluation and comparison of methods. There are several aspects for a good training and benchmarking dataset: accuracy of the provided ground truth, diversity of recorded scenarios, and provided additional cues are among the important ones.

Following a rise of interest in human motion trajectories modeling and prediction [12], the insufficiency of existing datasets has triggered creation of new comprehensive benchmarks for outdoor and vehicle motion [3–5], driven by the progress in the automated driving domain. On the contrary, human motion recordings indoors and in pedestrian zones are still lagging behind. The drawbacks of commonly used datasets mainly come from two sources: (1) severe artifacts in ground truth estimation from noisy input (e.g. cameras or range scanners) and (2) recordings were collected in not challenging environments with homogeneous human motion. Both cases are illustrated in Fig. 2.

As an alternative to the traditional recordings of natural scenes, we propose a data collection procedure to generate a diverse and accurate human motion dataset in a controlled weakly scripted setup [11] surpassing the limitations of existing datasets.

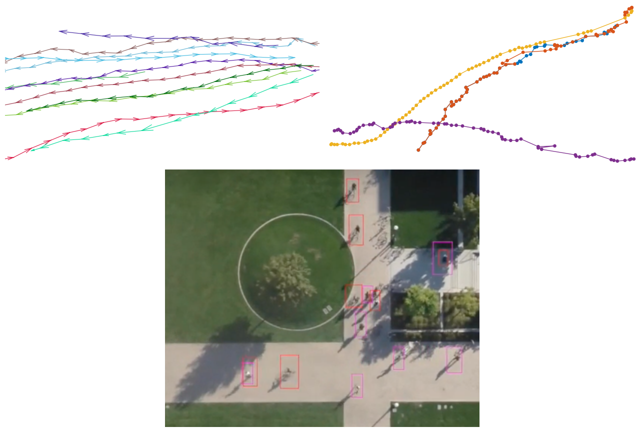

Figure 2: Prior art recordings of human motion trajectories: uniform motion in the ETH dataset [9] (top left), missing and incorrect detections in the Edinburgh dataset [8] (top right), rough position estimations with bounding boxes in the Standord Drone Dataset [10] (bottom). Our THÖR dataset addresses all of these issues, and offers various additional inputs (e.g. map of static obstacles, gaze directions).

2 THÖR: DIVERSE AND ACCURATE INDOOR MOTION TRAJECTORIES DATASET

The THÖR dataset ("Tracking Human motion in ÖRebro university") includes over 60 minutes of indoor human motion in a shared environment with a stationary and moving robot and static obstacles[1].

To tackle the issue of unreliable ground truth data, we have employed a high precision motion tracking setup Qualisys 7+ with 10 infrared cameras. The system tracks small reflective markers at 100 Hz with spatial discretization of 1mm. To reliably distinguish participants, each of them wore a bicycle helmet with the markers mounted in distinctive patterns. This system outputs 6D head position and orientation for each participant. Furthermore, we recorded eye gaze data with Tobii Pro Glasses for one participant.

To tackle the problem of not challenging enough setup, while retaining the controlled nature of the collected data, we have developed a framework for dynamical allocation of tasks for participants inspired by our experience in the ILIAD project, which investigates close human-robot collaboration in an industrial setup[2]. In contrast to the prior datasets containing monotonous observations of people following a very small number of motion patterns, the introduced framework allowed to develop a large number of intersecting maneuvers in both constrained areas and free space.

In an industrial setup, human behaviour is controlled through assigning each person goals, while leaving to the their own volition the method to achieve the said goals. In consequence, the employees generate most efficient paths according to their knowledge and criteria. In order to mimic these conditions we have generated a list of virtual tasks to be executed in different parts of the test environment. After completing each task, a new one is assigned to a participant. In consequence, the participants were forced to plan their own trajectories in a confined, dynamic environment.

Having instructed the participants to accomplish the tasks in dynamically allocated groups, we generated numerous interesting situations, such as accelerating to overtake another person; halting to let a large group pass or causing hindrance by walking towards each other in opposite directions. Here lies a specific benefit of our generative procedure for human motion, as in natural environments such interactions are comparatively rare, which contributes to unbalanced datasets and prevents the robot to correctly assess human interactions [13].

3 A MOTION PREDICTION BENCHMARK

THÖR is a first step of a larger effort for building a comprehensive motion prediction benchmark. The challenge lies in increasing scale and diversity, while maintaining a balance of interesting interactions. Firstly, a larger corpus of data is necessary for training modern deep learning methods [1]. Secondly, higher diversity in obstacle layout is required for generalizing human motion policies for avoiding static obstacles. Finally, varying robot behavior would allow researching human responses to robot's ego motion.

REFERENCES

[1] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese. 2016. Social LSTM: Human trajectory prediction in crowded spaces. In *Proc. of the IEEE Conf. on Comp. Vis. and Pat. Rec. (CVPR)*. 961–971.

[2] A. Bayoumi, P. Karkowski, and M. Bennewitz. 2017. Learning Foresighted People Following under Occlusions. (2017).

[3] Julian Bock, Robert Krajewski, Tobias Moers, Lennart Vater, Steffen Runde, and Lutz Eckstein. 2019. The inD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories at German Intersections. *arXiv preprint arXiv:1911.07602*.

[4] Markus Braun, Sebastian Krebs, Fabian B. Flohr, and Dariu M. Gavrila. 2019. EuroCity Persons: A Novel Benchmark for Person Detection in Traffic Scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019), 1–1. https://doi.org/10.1109/TPAMI.2019.2897684

[5] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. 2019. nuScenes: A multimodal dataset for autonomous driving. *arXiv preprint arXiv:1903.11027* (2019).

[6] Ravi Teja Chadalavada, Henrik Andreasson, Maike Schindler, Rainer Palm, and Achim J Lilienthal. 2020. Bi-directional navigation intent communication using spatial augmented reality and eye-tracking glasses for improved safety in human–robot interaction. *Robotics and Computer-Integrated Manufacturing* 61 (2020), 101830.

[7] Jim Mainprice, Rafi Hayne, and Dmitry Berenson. 2016. Goal set inverse optimal control and iterative replanning for predicting human reaching motions in shared workspaces. *IEEE Transactions on Robotics* 32, 4 (2016), 897–908.

[8] B. Majecka. 2009. Statistical models of pedestrian behaviour in the forum. *Master's thesis, School of Informatics, University of Edinburgh* (2009).

[9] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool. 2009. You'll never walk alone: Modeling social behavior for multi-target tracking. In *Proc. of the IEEE Int. Conf. on Computer Vision (ICCV)*. 261–268.

[10] A. Robicquet, A. Sadeghian, A. Alahi, and S. Savarese. 2016. Learning social etiquette: Human trajectory understanding in crowded scenes. In *Proc. of the Europ. Conf. on Comp. Vision (ECCV)*. Springer, 549–565.

[11] Andrey Rudenko, Tomasz P Kucner, Chittaranjan S Swaminathan, Ravi T Chadalavada, Kai O Arras, and Achim J Lilienthal. 2020. THÖR: Human-Robot Navigation Data Collection and Accurate Motion Trajectories Dataset. *IEEE Robotics and Automation Letters* 5, 2 (2020), 676–682.

[12] Andrey Rudenko, Luigi Palmieri, Michael Herman, Kris M Kitani, Dariu M Gavrila, and Kai O Arras. 2019. Human motion trajectory prediction: A survey. *arXiv preprint arXiv:1905.06113* (2019).

[13] Christoph Schöller, Vincent Aravantinos, Florian Lay, and Alois Knoll. 2019. The Simpler the Better: Constant Velocity for Pedestrian Motion Prediction. *arXiv:1903.07933* (2019).

[14] Dongfang Yang, Linhui Li, Keith Redmill, and Ümit Özgüner. 2019. Top-view trajectories: A pedestrian dataset of vehicle-crowd interaction from controlled experiments and crowded campus. In *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 899–904.

[1] Available at http://thor.oru.se
[2] https://iliad-project.eu