Proactive Motion Planning for Human-Robot Collaboration

Elena Basei, Edoardo Lamon, Matteo Saveriano, Daniele Fontanelli, Luigi Palopoli

***Abstract*—This abstract presents a real-time approach to human motion prediction and its integration into proactive dynamic human-aware motion planning for safe human-robot collaboration. We utilize a deep learning graph-based model to forecast future human movements. These predictions are then incorporated into a planning framework that employs a static roadmap and a time-variant A* algorithm to adapt the trajectory of a UR5e manipulator. This integration enhances human-robot interaction by combining accurate motion predictions with adaptive trajectory planning, enabling proactive collision avoidance.**

I. Introduction

Predicting human motion in real-time and proactively planning robot tasks is essential for safe human-robot collaboration. Traditional methods, such as Gaussian Mixture Models (GMMs) [1], function well in structured and repetitive scenarios but struggle with the unpredictability of human behavior. While conventional motion planning techniques prioritize safety, they often lack the adaptability needed for dynamic interactions. In contrast, deep learning (DL) methods for motion prediction, which utilize Recurrent Neural Networks (RNNs), Graph Convolutional Networks (GCNs), and Transformers, enhance the accuracy of predictions related to human tasks [2], [3]. However, their application in real-time settings is challenging due to high computational demands.

Traditional motion planning methods ensure safety but lack adaptability for dynamic interaction. Instead, proactive approaches, such as the Temporal Probabilistic Roadmap [4], improve collaboration by anticipating human actions. Our work seeks to incorporate these predictions into a human-aware planning framework to enhance safety and responsiveness. We employ a deep learning-based method to forecast human motion in real-time, and generate optimal and collision-free plans for the robot.

The contributions are the following:

1) a lightweight method for human motion prediction inference to be used in safety-critical contexts;
2) a real-time proactive planning method that utilizes predicted human motion in conjunction with a time-variant A* algorithm;
3) a parallelized version of the A* algorithm, optimized for real-time pathfinding in complex environments.

Interdepartmental Robotics Labs (IDRA), University of Trento, Trento, Italy {name.surname}@unitn.it

We acknowledge the support of the MUR PNRR project FAIR - Future AI Research (PE00000013), the project INVERSE (Grant Agreement n. 101136067),the project Magician (Grant Agreement n. 101120731) and the contribution of VRT foundation.

II. Methodology

A. Human Motion Prediction

To predict human motion, we utilize a graph-based model known as Graph-Mixer [2]. This model begins with an initial pose embedding module that converts input pose sequences into higher-dimensional features through adaptive spatial graph convolution. Following this, the spatial-temporal graph mixer combines both adaptive spatial and temporal graph convolutional networks to effectively capture human skeleton dynamics and the temporal relationships across different frames. Finally, the prediction head generates future motion predictions based on these spatial-temporal features, enhancing accuracy in long-term human motion forecasting. We trained this model for 20 epochs on the HA4M dataset [5]. After conducting a dedicated ablation study, we decided to use 12 frames of historical motion to predict 60 frames, equivalent to 2 seconds, of future motion.

B. Human-Aware Motion Planning

Another objective of this work is to incorporate real-time human predictions into the robot's planning process. In this way, we ensure that the robot can navigate without collisions while maintaining an acceptable level of performance. To address this challenge, we built our method upon the T-PRM framework proposed by Huppi et al. [4]. T-PRM was developed for planar holonomic robots with both static and moving obstacles. The roadmap graph was constructed using uniformly generated samples, which were deemed valid only if they did not collide with any static obstacles. Each edge in the graph was assigned a cost equal to its length, along with an estimated total time for traversal, calculated based on the assumption of a holonomic robot moving at a constant velocity. During the querying phase, the A* algorithm was utilized to identify the shortest paths in the graph. If any node of an edge was occupied by a dynamic obstacle during a specific time interval, that edge was assigned an infinite cost for that time interval. Our method expands upon T-PRM by incorporating human future motion as dynamic obstacles, enabling the generation of optimal collision-free trajectories for a manipulator. We create a uniformly spaced roadmap within the robot's Cartesian workspace, where each sample corresponds to a potential end-effector position. We then utilize the closed-form Inverse Kinematics (IK) of the manipulator to determine all possible joint configurations for each sample. These configurations are used to calculate travel times for each edge and to check for collisions with humans during specific time intervals. Furthermore, we introduce a human-robot distance condition in the collision-checking process to ensure that the robot generates trajectories that meet the safety requirements outlined in ISO/TS 15066:2016 for collaborative robots. To enable real-time computation of the final plan, we implemented multithreading to parallelize the execution of the A* algorithm. Once the optimal plan is identified, we continuously monitor the current and future positions of the human to dynamically adjust the robot's speed during its motion based on the distance between the robot and the human. If a potential collision is detected, we re-plan the robot's trajectory. If the predicted collision is expected to happen far in the future, the replanning can occur while the robot is still moving. However, if a collision is imminent, the robot will be stopped to prevent any accident as in ISO/TS 15066:2016. This replanning mechanism enables the robot to continuously adapt its motion according to predicted human movements, ensuring safe and smooth collaboration.

III. Results and Discussion

A. Human Motion Prediction Evaluation

We evaluated this method using the common Mean Per Joint Position Error (MPJPE), a standard metric in 3D human pose prediction. This metric calculates the average Euclidean distance between the predicted and actual joint positions in 3D space. The MPJPE is calculated by computing the error for each joint, averaging across all joints, and then averaging across all frames in a sequence. Formally:

$$MPJPE = \frac{1}{N}\sum_{i=1}^{N}\|\hat{J}_i - J_i\| \tag{1}$$

where N is the number of joints, $\hat{J}_i$ is the predicted 3D position of the i-th joint, J_i is the ground truth 3D position of the i-th joint, and $\|\cdot\|$ denotes the Euclidean distance between the predicted and ground truth joint positions. For the Graph-Mixer method, considering 2 s of prediction horizon, the MPJPE is 111.27 ± 101.89 mm, sufficient to understand the future occupancy area of the human operator (see Figure 1).

B. Human-Aware Motion Planning Evaluation

We conducted preliminary tests of our proposed framework using a Universal Robots UR5e robot to perform pick-and-place tasks while a human completed an assembly task (see Figure 2). A ZED2 RGB-D camera was employed to detect the current position of the human skeleton, and we utilized the Graph-Mixer network to predict the human's motion for the next two seconds at a frequency of 30 Hz. We compared our framework to a reactive planning approach that only considers the current position of the human in relation to the robot. Our results (see Table I) show that our proposed framework successfully completed nearly 30% more pick-and-place tasks than the reactive approach. Furthermore, the reactive method required the robot to stop its motion to avoid collisions nine times more often than our framework. When it came to re-planning while the robot was moving, our approach re-planned nearly five times more frequently than the reactive method.

The preliminary results suggest that incorporating human motion prediction into robot planning allows the manipulator to generate collision-free trajectories that are more resilient to human intrusion in the robot's workspace. These improved trajectories can enhance both safety and efficiency in collaborative industrial tasks.

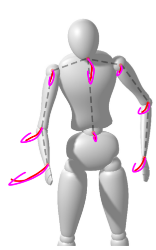

Fig. 1: Comparison of predictions (magenta lines) and ground truths (red lines) of the human's joint trajectories.

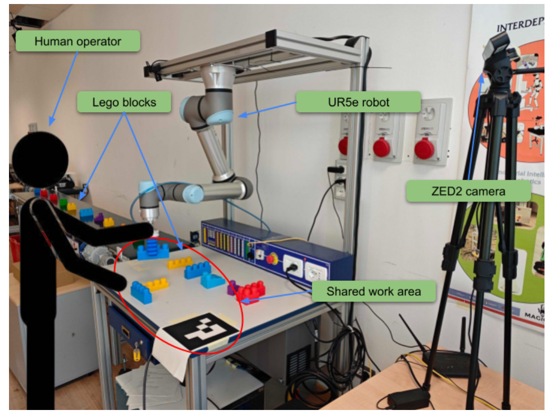

Fig. 2: Test environment with the Universal Robots UR5e robot and the ZED2 camera.

TABLE I: Comparison of proactive and reactive planning .

Method	Mean Planning Time (s)	Stop & Re-planning	Re-planning During Motion
Proactive	0.25	0.11	2.39
Reactive	0.19	0.93	0.5

References

[1] J. Mainprice and D. Berenson, "Human-robot collaborative manipulation planning using early prediction of human motion," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 299–306.

[2] S. Yang, H. Li, C.-M. Pun, C. Du, and H. Gao, "Adaptive spatial-temporal graph-mixer for human motion prediction," *IEEE Signal Processing Letters*, vol. 31, pp. 1244–1248, 2024.

[3] S. Tian, M. Zheng, and X. Liang, "Transfusion: A practical and effective transformer-based diffusion model for 3d human motion prediction," *IEEE Robotics and Automation Letters*, vol. 9, no. 7, pp. 6232–6239, 2024.

[4] M. Hüppi, L. Bartolomei, R. Mascaro, and M. Chli, "T-prm: Temporal probabilistic roadmap for path planning in dynamic environments," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 10 320–10 327.

[5] G. Cicirelli, R. Marani, L. Romeo, M. García-Domínguez, J. Heras, A. Perri, and T. D'Orazio, "The ha4m dataset: Multi-modal monitoring of an assembly task for human action recognition in manufacturing," *Scientific Data*, vol. 9, 12 2022.