

Towards Safer Robot Motion: Using a Qualitative Motion Model to Classify Human-Robot Spatial Interaction ^{*}

Laurence Roberts-Elliott¹[0000-0001-9877-0553], Manuel Fernandez-Carmona¹[0000-0002-0512-8594], and Marc Hanheide¹[0000-0001-7728-1849]

University of Lincoln, Lincoln LN67TS, UK
{laelliott, mfernandezcarmona, mhanheide}@lincoln.ac.uk

Abstract. For adoption of Autonomous Mobile Robots (AMR) across a breadth of industries, they must navigate around humans in a way which is safe and which humans perceive as safe, but without greatly compromising efficiency. This work aims to classify the Human-Robot Spatial Interaction (HRSI) situation of an interacting human and robot, to be applied in Human-Aware Navigation (HAN) to account for situational context. We develop qualitative probabilistic models of relative human and robot movements in various HRSI situations to classify situations, and explain our plan to develop per-situation probabilistic models of socially legible HRSI to predict human and robot movement. In future work we aim to use these predictions to generate qualitative constraints in the form of metric cost-maps for local robot motion planners, enforcing more efficient and socially legible trajectories which are both physically safe and perceived as safe.

Keywords: HRI · HRSI · spatial reasoning · Human-Aware Navigation · Hidden Markov Models · classification

1 Introduction

In industrial applications with environments shared between humans and robots, AMRs must move safely around humans and in a way which humans perceive to be safe. Physical human safety in robot navigation can be all but assured by simply stopping robot motion when anything is detected closer than a minimum safe distance to a robot’s safety laser(s). This is highly inefficient and usually conflicts with personal space [1], thus perceived as unsafe.

The work discussed in this paper contributes to the safety stack of the human collaborative warehouse robots of the EU ILIAD Project [2], aiming to increase perceived safety by humans while preserving a critically physically safe system. The safety stack is implemented at four levels. At the highest level, global planning, robots take human flows into account to generate paths that

^{*} Supported by the EU H2020 ILIAD Project (grant #732737).

don't interfere with human motion patterns [3]. The second level affects robot maximum speeds. If a robot detects a human in its vicinity, it will adapt its speed to the HRSI at hand. However, if the global path seems unfeasible given current HRSI, a new path will be triggered to account for the new interaction. Finally, a safety stop will be triggered if human gets too close to the robot's laser anyways.

This work, involved in the second level of the ILIAD safety stack, addresses HRSI awareness with per-situation probabilistic models of qualitative abstractions of human and robot movement. Our probabilistic models use Qualitative Trajectory Calculus version C (QTC_C) [4] to encode the movements of two positions in space from one point in time to a subsequent point in time. This lets us represent pairs of trajectories in HRSI as sequences of qualitative states, where each state describes the relations of the movements of both human and robot.

The main contribution of this work is the multi-HMM classifier of the HRSI situation of sequences of qualitative descriptions of human and robot movement, extending [5] with multi-HMM classification and modelling of additional situations. The classifier is very fast to train, and fast and accurate enough at classification to be used in real-time in a safety critical situational HAN approach. The use of QTC, described in detail in Sec. 3.2, abstracts pairs of trajectories, decreasing the impact of sensor and tracking error. We train and test our classifier using data recorded from HRSIs with an automated Linde CiTiTruck pallet truck, making it more suitable for industrial applications than approaches which model robot movement after human-human interaction.

We plan to use the classifier, along with per-situation HMMs of QTC_C state transitions from socially legible HRSI, to predict the socially legible next state for a HRSI in realtime, given the HRSI's situation and QTC_C sequence. In future work, we aim to constrain the robot's local planner to enforce the transition to this state.

The rest of the paper is structured as follows. We first present other works within the field of HRSI in Sec. 2. Later on, we describe in Sec. 3 the theoretical foundations of our system. Implementation and analysis of the system is made in Sec. 4. Finally we summarise the results and draft future work in Sec. 5.

2 Related Work

Common approaches to robot navigation consider humans the same as any other obstacle [6], resulting in movements that are inefficient and perceived as unsafe. HAN is required for legible paths which are more likely to be perceived as safe [7]. Approaches for HAN often consider Hall's proxemic zones [1], but neglect to consider the intentions of human's movement [6].

Predictive models address this limitation by identifying and forecasting human trajectories. They rely on studying human motion in social environments, so that most likely paths are known when a similar situation is

matched. These works propose the use of qualitative domains and symbols to reduce the complexity of the task at hand [8] and include desired features (that increase social normativity) [9] with good performance in crowded scenarios.

However, these approaches do not capture in their motion models the presence of a robot. There is the implicit assumption that humans will move as if the robot was just another pedestrian. This is not appropriate for navigation of heavy industrial robots, which cause humans to feel unsafe when the robot moves close to them. Also, using these models as path planners for the robot may not provide the safest route, but instead the most human alike. This is particularly relevant for the domain at hand: shared warehouse environments. Within industrial applications, robots need to ensure safety over any other requirement, so mimicking human trajectories (e.g. cutting through a crowd) may be discouraged.

QTC used in our HRSI model describes relative movements of both human and robot [4] in the same way that two humans walking on intersecting trajectories negotiate their movements without knowledge of their quantitative positions [10]. A set of per-situation HMMs of transitions between QTC states can be used to classify the HRSI situation from QTC sequences generated from pairs of human and robot trajectories [5]. Here we extend the HRSI situation classification of [10], modelling additional situations.

3 System definition

3.1 Human Position Tracking

Warehouses are a challenging scenario for people tracking (multiple occlusions, dim lights, people sitting, kneeling, etc.), which may be a limiting factor for qualitative HRSI analysis. Our probabilistic HAN model uses human trajectories obtained with the real-time first-person people tracking system described in [11]. This tracking fuses data from the robot’s on-board RGB-D camera and laser, which has a limited reach and requires extra computing power, but does not require any sensors to be installed in the warehouse.

3.2 Probabilistic QTC Model

Our probabilistic QTC model, uses sequences of QTC_C states to develop a Hidden Markov Model (HMM) for each of a set of HRSI situations, defined as classes in Sec. 3.3 and extending our previous work in [12]. We encode pairs of human and robot trajectories using QTC version C (QTC_C) [4]. In QTC_C , movements of two agents in space are represented by a 4-tuple of state descriptors (h_1, r_1, h_2, r_2) . Each descriptor expresses a qualitative spatial relation using a symbol $\in \{-, 0, +\}$. With this 4-tuple of descriptors comprised of 3 symbols, there a total of $3^4 = 81$ possible QTC_C states.

The relations of the descriptor symbols are defined as follows:

- h_1) movement of h w.r.t. r at time t :
- : h is moving towards r
 - 0 : h is neither moving towards nor away from r
 - + : h is moving away from r
- r_1) movement of r w.r.t h at time t :
- The same as h_1), but with h and r swapped
- h_2) movement of h w.r.t the line \vec{hr} at time t :
- : h is moving to the left side of \vec{hr}
 - 0 : h is moving along \vec{hr} or not moving at all
 - + : h is moving to the right side of \vec{hr}
- r_2) movement of r w.r.t the line \vec{rh} at time t :
- The same as h_2), but with h and r swapped

where t is the earlier of the two points in time, r is the robot's position, and h is the human's position.

For example, Fig. 1 shows an interaction: human is moving towards the robot ($h_1 = -$) and robot is approaching too ($r_1 = -$). Human is directly headed to the robot ($h_2 = 0$) but robot is to its left side ($r_2 = -$).

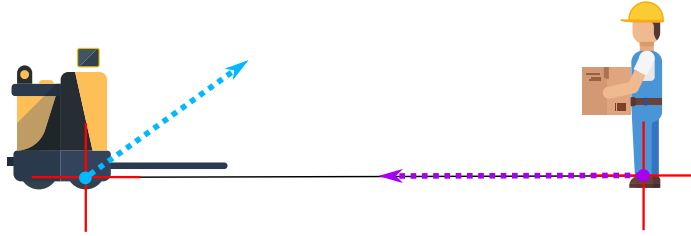


Fig. 1. QTC_C state $(-, -, 0, -)$: human is moving directly towards the robot, while the robot is moving toward and on its left side.

3.3 Classes of HRSI

In [12] there were two relevant classes in human robot spatial interaction. We extend this initial classification to account for the interactions in warehouse environments. Specifically we focus on interactions that require a robot to adjust its movements to accommodate the human's. In order to account for these, we model the following situations (see Fig. 2) with HMMs for our classifier:

- Passing By on the Left (PBL): Both actors pass each-other on the left side from their perspective, moving in opposite directions.
- Passing By on the Right (PBR): Both actors pass each-other on the right side from their perspective, moving in opposite directions.

- Robot Overtakes Left (ROL): The robot passes on the left of the human while both move in the same direction.
- Robot Overtakes Right (ROR): The robot passes on the right of the human while both move in the same direction.
- Path-crossing (PC): The robot has to slow or stop movement to allow the human to move across the robot's intended path.
- Rejection: A set of situations defined below which do not require the robot to alter its movement.

Rejection situations:

- Human and Robot are Stationary (HRS)
- Human Meets Stationary Robot (HMSR): The human moves toward the stationary robot, and stops when close.
- Robot Meets Stationary Human (RMSH): The robot moves toward the stationary human, and stops when close.
- Human Passing By Stationary Robot on the Left (HPBSRL): The same as PBL, except that the robot is stationary.
- Human Passing By Stationary Robot on the Right (HPBSRR): The same as PBR, except that the robot is stationary.

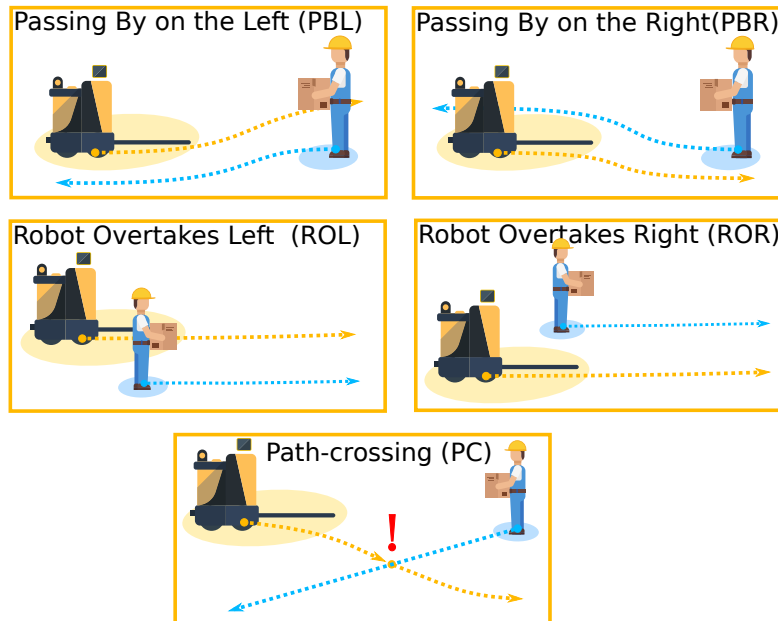


Fig. 2. HRSI Classes

3.4 Creating HMMs for our Multi-HMM Situation Classifier

We consider in this study 6 classes $C = (\text{'PBL'}, \text{'PBR'}, \text{'ROL'}, \text{'ROR'}, \text{'PC'}, \text{'Rejection'})$, thus we will model 6 different HMMs, where the observation corresponds to a new state. Each HMM is comprised by $|Q| \times |Q|$ transition matrices listed in A_i , an $|Q| \times |Q|$ observation matrix B , and the $1 \times |Q|$ initial state vectors listed in I_i . These account for all possible transitions in QTC_C states $Q = ((- - - -) \dots (+ + + +))$, as in [4] on each class.

Our system is composed by the collection of transition matrices $A = (A_1 \dots A_{|C|})$, and initial state vectors $I = (I_1 \dots I_{|C|})$, indexed by class number. Each element of the list of per-class HMMs $H = (H_1 \dots H_{|C|})$ is a tuple composed by (A_c, B, I_c) with c indexing the class's name in C , fully describes our system.

All of the classifier's HMMs share B as their observation matrix. We use B to account for the possibility of generating incorrect QTC_C states due to sensor and tracking error, assuming a probability $t = 0.95$ that the true (hidden) QTC_C state matches the emitted QTC_C state generated from tracked human and robot positions. So, we initialize matrix B almost as an identity matrix with some noise, with diagonal $B[i, i] = t$ and the rest of elements $B[h, o] = \frac{1-t}{|Q|-1}$ ($b \neq o$).

Recorded QTC_C state sequences list S generated from human-robot trajectory pairs is used to obtain A , and I . First, we map each QTC_C state in each sequence S_i to its index in Q . Each state sequence in S will have a class label assigned $l_i \in [1 \dots |C|]$, so that the list of labels will be $L = (L_1 \dots L_{|S|})$ and L_s is the class label for sequence S_s . Initially A , and I are assigned uniform probabilities. Then we use the recorded state sequence list S to model the probabilities:

$$I_{L_n}[S_n[1]] = I_{L_n}[S_n[1]] + 1 \text{ for } n = 1 \text{ to } |S|.$$

$$A_{L_n}[S_n[q], S_n[q+1]] = A_{L_n}[S_n[q], S_n[q+1]] + 1 \text{ for } n = 1 \text{ to } |S|, \text{ and } q = 1 \text{ to } |S_n| - 1.$$

Finally we normalise matrix B , the matrices of A , and the vectors of I , such that each row sums to 1. With these HMMs, we can classify a QTC_C sequence as the class of the HMM that estimates the highest log-likelihood of the given sequence being observed [13].

4 Experiments

4.1 Laboratory Setup for Recording HRSI Situations

Our robot is a modified Linde CitiTruck, equipped with front-facing laser, LiDAR, Kinect 2, and PC running ROS, interfacing with the sensors, and the pallet truck's motor controllers. One of several automated pallet trucks belonging to the ILIAD Project. In our robotics lab we placed coloured tape on the floor, marking start and end positions for the human and the robot, as pictured in Fig. 3. In the diagram on the left of Fig. 3, arrows indicate start positions, and crosses indicate end positions. The robot follows the path between the black positions. In HRSI situations where the robot is stationary,

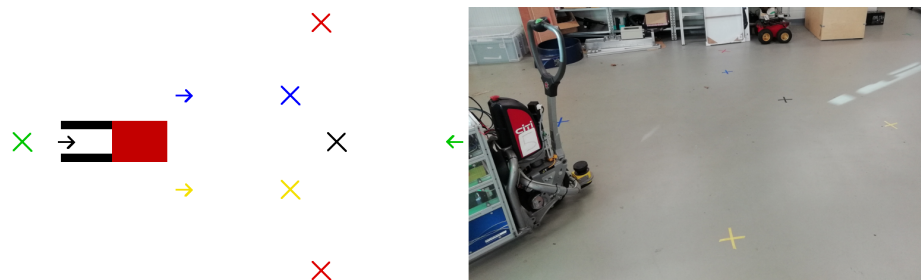


Fig. 3. Illustration (left) and photograph (right) of laboratory setup for recording HRSI situations.

the robot stays at the start position. When the robot begins moving it emits a click sound, which we use to signal the human to move. In conditions where the robot does not move, the experimenter speaks the signal 'go'. We record the robot and nearest human position on the robot's metric map, from when the robot begins moving, to when the human reaches their end position. Human positions are tracked within an 'active area' to reduce the risk of the experimenter being tracked instead of the interacting human.

The human moves as follows for the different situations:

PBL - The human moves from the green arrow to the green cross, moving to their left to pass the robot.

PBR - The same as PBL, but the human moves to their right to pass the robot.

ROL - The human moves from the yellow arrow to the yellow cross, moving as slowly as possible to allow the robot to overtake at a safe speed.

ROR - The same as ROL, but the human moves from the blue arrow to the blue cross.

PC - the human moves from the topmost red cross to the other red cross, we also record the human moving from the bottom red cross to the topmost red cross.

HRS - Human stands stationary at the yellow cross, with robot stationary at the black arrow.

HMSR - The same as HRS, but the human moves from the yellow cross to the yellow arrow.

RMSH - The same as HRS, but the robot moves from the black arrow to the black cross.

HPBSRL - The same as PBL, but the robot stays at the black arrow.

HPBSRR - The same as PBR, but the robot stays at the black arrow.

4.2 Training Dataset

To train our multi-HMM classifier we recorded 35 interactions between a robotics expert and the robot for each situation of the 5 classes of Sec. 3.3, and 15

interactions for each of the 5 rejection situations. A total of 250 HRSIs. We create the multi-HMM classifier’s HMMs as described in Sec. 3.4, using QTC_C sequences generated from human and robot trajectories using QSRlib [14]. We use QSRlib’s ‘collapse’ feature when generating all of our QTC_C sequences from HRSIs, to remove repeating states, reducing the variance between sequences from HRSIs of differing length. We used 3-fold Cross Validation for preliminary estimation of our classifier’s performance, to measure the likely impact of changes to our model to its ability to classify HRSI situations beyond those recorded in this training set.

4.3 Test Dataset

To test the ability of our multi-HMM classifier to classify the situation of spatial interactions between the robot and non-experts of varied age, gender and cultural background, we conducted a study. In this repeated measures study participants enacted HRSI situations using the methods described in Sec. 4.1. These situations were enacted in a randomised order, with each participant also only performing 1 of the 5 rejection situations, chosen at random. With 11 participants, we recorded a total of 75 interactions. We created the multi-HMM classifier’s HMMs using QTC_C sequences from the training set, and evaluated its performance in classifying the HRSI situation of QTC_C sequences from the study’s HRSIs. Training of the classifier took only 50ms to execute. Each classification took 60ms to execute on average. The number of interactions recorded per class is detailed in the study’s confusion matrix in Fig. 4, which has its statistics explained in Sec. 4.4.

4.4 Results and Discussion

Fig. 4 is a confusion matrix containing metrics of the performance of the classifier at predicting the HRSI situation from QTC_C sequences in the test set described in Sec. 4.3, trained on sequences from the training set described in Sec. 4.2. The cells of the confusion matrix with a green or pale red background contain the count of classifications for the predicted class and actual class given by the cell’s row and column respectively. Below each of these counts is the count as a percentage of the total number of classifications. The rightmost column of the matrix contains, in this order, the count of QTC_C sequences classified as the row’s class, and the precision and False Positive Rate of classifications as the row’s class. The bottom row of the matrix contains, in this order, the count of QTC_C sequences that are labelled with the column’s class, and the recall and False Negative Rate of classifications as the column’s class. The bottom-right cell contains, in this order, the total count of all classifications, the overall accuracy, and the overall misclassification rate.

The classifier’s overall accuracy is high at 92%, as it must be as a component of the safety focused HAN approach which we plan to develop, the details of which are elaborated on in Sec. 5. The ability of our qualitative probabilistic model to accurately classify HRSIs with 11 non-experts, trained

Confusion matrix

Predicted	PBL	8 10.67%	2 2.67%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 80.00%	20.00%					
	PBR	2 2.67%	9 12.00%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	11 81.82%	18.18%					
	ROL	1 1.33%	0 0.0%	11 14.67%	0 0.0%	1 1.33%	0 0.0%	13 84.62%	15.38%					
	ROR	0 0.0%	0 0.0%	0 0.0%	10 13.33%	0 0.0%	0 0.0%	10 100%	0.00%					
	PC	0 0.0%	0 0.0%	0 0.0%	0 0.0%	20 26.67%	0 0.0%	20 100%	0.00%					
	Rejections	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	11 14.67%	11 100%	0.00%					
	recall	11 72.73%	11 81.82%	11 100%	10 100%	21 95.24%	11 100%	75 92.00%	27.27%	18.18%	0.00%	0.00%	4.76%	0.00%
		PBL	PBR	ROL	ROR	PC	Rejections	precision						
		Actual												

Fig. 4. Confusion matrix for validation of our multi-HMM classifier using QTC_C sequences from study HRSIs as test data.

on HRSIs with 1 robotics expert, demonstrates the benefit of abstracting HRSIs to a qualitative description. It should be noted that none of the situations which require changes to robot motion were misclassified as rejections, though the classifier does confuse some of these situations with others. The 2 situations which are most often confused by our classifier are PBL and PBR. Some of the QTC_C sequences generated for these 2 situations are nearly identical between classes. In a small number of the trajectory pairs recorded for these situations, the human trajectory passes the wrong side of the robot for the interaction’s labelled situation. This could be the result of human error on the part of the experimenter in labelling the study, or could be due to error in the human tracking, which we’ve found to occasionally track inanimate objects or the experimenter instead of the interacting human. We planned to use the more robust tracker of [15], but it was too computationally expensive given our robot’s hardware. If we can optimise this tracker, or use another tracker more robust to False Positives than [11], and label interactions more vigilantly, we may see less confusion between PBL and PBR, resulting in higher overall accuracy.

5 Conclusion and Future Work

We have presented here the framework used in the ILIAD Project to classify HRSI situation. If global planning is unable to create paths avoiding HRI, the classified situation will define how the intermediate safety layers will react in ILIAD. An efficient and accurate prediction will have a direct impact in the number of safety stops triggered by the lowest safety layer. Future work will evaluate the performance of this approach in the project’s overall safety architecture.

The high accuracy of our multi-HMM HRSI situation classifier when tested on the HRSIs recorded in our experiment demonstrates its suitability for use in a situational HAN approach, with some room for improvement in recall and precision of PBL and PBR, which may be possible by taking the steps described in Sec. 4.4.

While this fully describes the applicability of our work in a qualitative HAN approach, the systems in this paper need to be applied to real-time robot navigation in the presence of humans. Then the real-time HAN system should be tested on an industrial AMR in its ability to improve perceived safety and social legibility of robot movement in HRSI, while minimising safety-laser stops. For this real-time system, we aim to predict the next QTC_C state in a given sequence using per-situation HMMs from QTC_C sequences generated from socially legible HRSIs, and provide qualitative constraints for robot motion planner cost-maps, in order to improve perceived safety with socially normative movement, extending [16].

We plan to develop this real-time system, recording QTC_C sequences from robot and human positions as they are detected by the robot’s human tracking. Once a human position is within a variable ‘interacting distance’ of the robot, the classifier defined in this paper will classify the HRSI situation from the QTC_C sequence generated from the human and robot positions, and predict the most likely next QTC_C state using a HMM constructed from QTC_C sequences from socially legible HRSIs in the classified situation. We continue to record the QTC_C sequence of the interacting human and robot, and, at a variable frequency, repeat the process of predicting the socially legible next state and constraining the cost-map accordingly while the human is within ‘interacting distance’. When multiple constraint cost-maps are to be applied simultaneously, the union of these is used to constrain the robot motion appropriately for simultaneous HRSIs with any number of humans.

For data to train the HMMs and validate this approach, we look to label and process trajectory pairs in larger datasets, e.g. the Thör human-robot trajectory dataset [17], then test in real-time on our robot in a real warehouse, interacting with human workers. We hope to use the qualitative constraint cost-maps to communicate the robot’s navigation intention, projecting on the floor in green on areas of high-cost, and red on areas of low-cost as conditioned symbols for ‘go’ and ‘don’t go’.

References

1. E. T. Hall *et al.*, “Proxemics,” *Current Anthropology*, vol. 9, no. 2/3, pp. 83–108, 1968.
2. ILIAD Project, “ILIAD Project – An EU-funded research project on Intra-Logistics with Integrated Automatic Deployment for safe and scalable fleets in shared spaces,” Feb. 2020.
3. T. Vintr *et al.*, “Time-varying pedestrian flow models for service robots,” in *2019 European Conference on Mobile Robots (ECMR)*, pp. 1–7, Sep. 2019.
4. N. Van de Weghe *et al.*, “A Qualitative Trajectory Calculus and the Composition of Its Relations,” in *GeoSpatial Semantics* (D. Hutchison *et al.*, eds.), vol. 3799, pp. 60–76, Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.
5. C. Dondrup, N. Bellotto, and M. Hanheide, “A Probabilistic Model of Human-Robot Spatial Interaction Using a Qualitative Trajectory Calculus,” in *AAAI Spring Symposium Series*, (Palo Alto, California, USA), AAAI, Mar. 2014.
6. T. Kruse *et al.*, “Human-aware Robot Navigation: A Survey,” *Robotics and Autonomous Systems*, vol. 61, pp. 1726–1743, Dec. 2013.
7. M. Fernandez Carmona, T. Parekh, and M. Hanheide, “Making the Case for Human-Aware Navigation in Warehouses,” in *TAROS 2019* (K. Althoefer *et al.*, eds.), LNCS, (Cham), pp. 449–453, Springer International Publishing, 2019.
8. C. I. Mavrogiannis and R. A. Knepper, “Multi-agent path topology in support of socially competent navigation planning,” *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 338–356, 2019.
9. H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard, “Socially compliant mobile robot navigation via inverse reinforcement learning,” *The International Journal of Robotics Research*, vol. 35, no. 11, pp. 1289–1307, 2016.
10. M. Hanheide, A. Peters, and N. Bellotto, “Analysis of human-robot spatial behaviour applying a qualitative trajectory calculus,” in *2012 IEEE RO-MAN*, (Paris, France), pp. 689–694, IEEE, Sept. 2012.
11. C. Dondrup *et al.*, “Real-time multisensor people tracking for human-robot spatial interaction,” (Seattle, WA), ICRA / IEEE, May 2015.
12. Christian Dondrup, Nicola Bellotto, Marc Hanheide, Kerstin Eder, and Ute Leonards, “A Computational Model of Human-Robot Spatial Interactions Based on a Qualitative Trajectory Calculus,” *Robotics*, vol. 4, no. 1, p. 63, 2015.
13. A. M. White, “Sequence Classification - with emphasis on Hidden Markov Models and Sequence Kernels,” Sept. 2009.
14. Y. Gatsoulis *et al.*, “QSRlib: a software library for online acquisition of qualitative spatial relations from video,” in *Qualitative Reasoning 29th International Workshop on Qualitative Reasoning*, (California, USA), pp. 36–41, IJCAI, July 2016.
15. T. Linder, S. Breuers, B. Leibe, and K. O. Arras, “On multi-modal people tracking from mobile platforms in very crowded and dynamic environments,” in *2016 IEEE ICRA*, (Piscataway, New Jersey, USA), pp. 5512–5519, IEEE, May 2016.
16. C. Dondrup and M. Hanheide, “Qualitative constraints for human-aware robot navigation using Velocity Costmaps,” in *2016 IEEE RO-MAN*, (Piscataway, New Jersey, USA), pp. 586–592, IEEE, Aug. 2016.
17. A. Rudenko *et al.*, “Thör: Human-robot indoor navigation experiment and accurate motion trajectories dataset,” *arXiv preprint arXiv:1909.04403*, 2019.