Intra-Logistics with Integrated Automatic Deployment:
Safe and Scalable Fleets in Shared Spaces

H2020-ICT-2016-2017
Grant agreement no: 732737

DELIVERABLE 3.3

Report on human-robot spatial interaction and mutual communication of navigation intent

Due date: month 40 (April 2020)
Deliverable type: R
Lead beneficiary: UoL

Dissemination Level: PUBLIC

Main author: Manuel Fernandez-Carmona (UoL),Laurence Roberts-Elliott (UoL), Marc Hanheide (UoL), Ravi Teja Chadalavada (ORU)

Contents

Acronyms

AGV Autonomous Ground Vehicle

AMR Autonomous Mobile Robot

AOI-Robot Area of Interest around Robot

CTA Continuous Trajectory Assessment

EDA Electrodermal Activity

FOV field of view

HAN Human-Aware Navigation

HMM Hidden Markov Model

HR Heart Rate

HRI Human Robot Interaction

HRSI Human Robot Spatial Interaction

KL Kullback-Leibler

ORU Örebro University

PBL Passing By on the Left

PBR Passing By on the Right

PCL Path-Crossing from the Left

PCR Path-Crossing from the Right

QTC Qualitative Trajectory Calculus

QTC_C Qualitative Trajectory Calculus Version C

RMSH Robot Meets Stationary Human

ROL Robot Overtakes Left

ROR Robot Overtakes Right

ROS Robot Operating System

SAR Spatial Augmented Reality

SEEV Salience Effort Expectancy Value

UoL University of Lincoln

vSMU Vehicle Safety Motion Unit

WP Work Package

1 Introduction

The results covered in this report contribute to ILIAD objectives O4 (human-aware AGV fleets in shared environments), by ensuring a correct human-robot spatial interaction, and O6 (human safety), by supporting communication of navigation intents, to create a shared environment that is also safe for humans.

These objectives are part of Work Package (WP) 3: Human-aware AGV Fleets. WP 3 covers methods, theory and experiments supporting robot awareness of humans and socially normative behaviour in shared spaces. Within WP 3, this deliverable focuses on tasks T3.4 and T3.5, with the following two goals.

- Learn human spatial behaviour models and develop qualitative representations to track, predict and reason about human-robot joint motion (T3.4) with the goal of legible and collision-free paths. Conduct user studies to evaluate perceived legibility by users.
- Investigate means to communicate intended robot motion and equip the vehicles with such means including indicators or laser projection devices for graphic communication on the floor and surrounding surfaces (T3.5).

At a closer level, task T3.4 is highly coupled with tasks T3.3, T4.2, T4.3 and T5.2. Reliable human tracking (provided by T3.3) and current robot motion planning (from T5.2) are main inputs to asses human-robot interaction. This interaction may lead to safety constraints additional to the ones described by tasks T4.2 and T4.3 and eventually feed back to the robot motion planning.

This document describes the implementation and validation by user studies of the proposed frameworks for spatial interaction (T3.4) and navigation intent communication (T3.5). First, Section 2 gives an overview of the safety framework employed in ILIAD, and where the work of T3.4 and T3.5 fits in. The qualitative human-robot spatial interaction from T3.4 is detailed in Section 3 and the work on mutual communication of navigation intent ftom T3.5 is in Section 4.

2 Safety and Human Robot Interaction in ILIAD

In industrial applications with environments shared between humans and robots, Autonomous Mobile 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 [7], thus perceived as unsafe. We will describe in this section the ILIAD approach to safety in Human Robot Spatial Interaction (HRSI).

2.1 Definition

To increase safety and readability, in ILIAD we propose a safety stack with five different layers, each one focused on a different interaction level and designed to minimise the activation of the lower ones (see Table 1). Higher layers try to prevent unnecesary human robot interactions and potentially harmful manoeuvres. Middle layers will give relevance to Human Robot Interaction (HRI) at local, shorter timeframes and finally the lowest layer will ensure basic safety.

Our top-most safety layer (layer 5) is **flow-aware motion planning** using *maps of dynamics* (from T5.2 and T2.1, to be covered in detail in D5.3). This layer creates (global)

robot paths that are consistent with human patterns, reducing the amount of potential interference. It is based upon statistical information about human behaviours, trying to adapt robot paths to general *expected* motion flows, learned from site-specific data; as opposed to reacting to currently observed people, which is handled by the lower layers.

Below this one, there are the three middle safety layers. **Navigation intent communication** is the fourth safety layer proposed in ILIAD, as it makes robot behaviour more legible to humans. It creates an information loop between human and robot, and estimates what is the current worker focus depending on his gaze. These features may prevent most of the actions of the other layers, by allowing humans to avoid conflicting robots and making them aware of their whereabouts. This layer is the focus of T3.5 and is described in more detail in Section 4.

Next, safety layer 3 provides awareness to the presence of humans in the robot's path during path execution, which wasn't covered by the flow-aware motion planning and navigation intent communication layers. **Situation-aware planning** (see Section 3) allows to include human interactions into the path planning pipeline, aiming to increase perceived safety by humans while preventing any safety risk along the robot's trajectory. This layer will provide dynamic costmaps based on human interaction to the navigation planning, and will allow safe and legible spatial interactions.

In order to allow humans to carry on safely and undisturbed, a speed modulation of the robot's planned trajectory may be sufficient. Initially, the Vehicle Safety Motion Unit (vSMU) (to be reported in D4.3) will filter out velocities that could harm humans moving around the Autonomous Ground Vehicle (AGV). It creates speed constraints, derived from an extensive injury safety database (see D4.4), in order to produce speeds that are guaranteed not to cause injury. These speed constraints are then extended in the **Extended vSMU** safety layer to improve legibility and predictability of robot motion while still avoiding replanning as much as possible for efficiency reasons.

Finally, the lowest safety layer prevents the robot from getting close to any obstacle. **Safety stops** are the last resort. They halt the robot without taking into account any software condition but the readings of the safety laser sensor.

#	Layer	Role	WP
5	Flow-aware global motion planning	Create paths consistent with human behaviours	WP5
4	Navigation Intent Communication	Make robot motion legible by humans	WP3
3	Situation-aware planning	Include local interactions into planning	WP3
2	Extended vSMU: Speed Constraints	Adjust robot speed to human interactions	WP3, WP4
1	Safety stops	Avoid collisions	

Table 1: Safety layers in ILIAD. Highlighted ones are described in this deliverable.

This deliverable will focus on the safety features derived from local human interaction. That is, the three middle safety layers which deal with local information about human activity and use it to intelligently influence the behaviour of the robots and humans.

2.2 Implementation and Integration of the Safety Stack in ILIAD

Path planning modules centralize all the navigation information in ILIAD. These path planning modules take into account both global human patterns from flow-maps [3, 4, 5, 6] (layer 5) and local human information obtained from situational constraints (layer 3).

Layer 4 (see Section 4) uses navigation information from the robot to communicate navigational intent to humans, using a projector mounted on the robot. Additionally, we make use of eye-tracking devices for monitoring gaze and attention direction, which may

be used in this module to adjust the displayed information depending on the attention of surrounding humans.

Human tracking information (Task 3.3, see also D3.1 and D3.2) is the basic input to safety layer 3 in ILIAD. Several real-time human tracking systems can be used here, reducing the coupling between human tracking and safety in ILIAD: from the lightweight one described in [1] to the multimodal proposed by [2] and the further improvements developed in Task 3.3. These systems fuse data from the robot's on-board RGB-D camera and laser, and do not require any sensors to be installed in the environment.

We use tracking information and robot motion data to create *situational constraints*. These are generated from the the specific interaction described by human-robot relative trajectory using Qualitative Trajectory Calculus (QTC). The Qualitative Trajectory Calculus (QTC) module generates a discrete-space qualitative description of human-robot relative motion, a QTC state (see Section 3.3).

QTC state evolution is later used in the HRSI detection module. This module will estimate the most likely interaction from the QTC flow describing the relative trajectory, using the approach proposed in Section 3.5. Detected interaction will be one from the set of relevant ones previously defined in Section 3.4.

Depending on the detected spatial interaction, we define areas around the robot that should be avoided using cost map constraints. These maps have two uses. First, they shape valid paths generated by the path planning module, thus defining safety layer 3. But also, they are used by the Continuous Trajectory Assessment (CTA) module to continuously monitor the potential risk of remaining path. The CTA module guarantees that the proposed path is safe for humans by closely monitoring the robot's path execution in presence of humans. A safe path may turn unsafe if a human changes trajectory, a new actor appears in the scene or any other change occurs.

Depending on the risk level, the CTA module can trigger different actions. A lower risk level allows to stick to the original navigation plan, which is usually preferred, and relies on reducing robot speeds to allow humans to proceed undisturbed. This is addressed by the extended vSMU module, implementing safety layer 2. If the current path is unfeasible due to safety reasons, a recalculation of the current plan is requested. This will generate a completely new trajectory, taking into account current human interaction. This module will abort the current plan and request a new one. This is a last resort action, to be taken only if the situation at hand is considered harmful.

Finally, layer 1 has no software implementation, as it is a hardware component limiting the minimum distance detected from laser sensors that allows the robot to move.

3 Qualitative Human-Robot Spatial Interaction

This section 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) [8] 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.

3.1 Related Work

Common approaches to robot navigation consider humans the same as any other obstacle [9], resulting in movements that are inefficient and perceived as unsafe. Human-Aware Navigation (HAN) is required for legible paths which are more likely to be perceived as

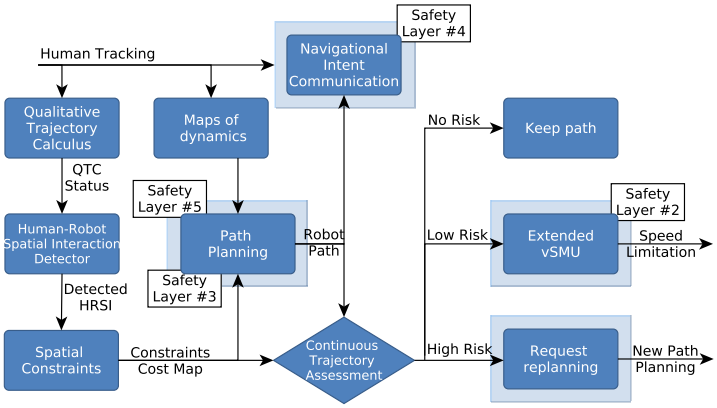

Figure 1: Interaction between ILIAD safety layers.

safe [10]. Approaches for HAN often consider Hall's proxemic zones [7], but neglect to consider the intentions of human's movement [9].

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 [11] and include desired features (that increase social normativity) [12] 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-like. 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 [8] in the same way that two humans walking on intersecting trajectories negotiate their movements without knowledge of their quantitative positions [13]. A set of per-situation Hidden Markov Models (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 [14]. Here we extend the HRSI situation classification of [13], modelling additional situations.

3.2 QTC Model

Our probabilistic model uses sequences of QTC_C states to develop a discrete transitions HMM for each of a set of HRSI situations, defined as classes in Section 3.3 and extending our previous work [1]. We encode pairs of human and robot trajectories using QTC version C (QTC_C) [8]. 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.

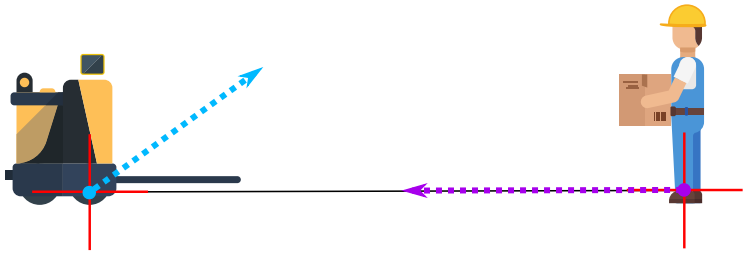

Figure 2: QTC_C state $(-,-,0,-)$: the human is moving directly towards the robot, while the robot is moving toward and on its left side.

The relations of the descriptor symbols are defined as follows, where t is the earlier of the two points in time, r is the robot's position, and h is the human's position:

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 $\overrightarrow{hr}$ at time t:

- $-$: h is moving to the left side of $\overrightarrow{hr}$
- 0 : h is moving along $\overrightarrow{hr}$ or not moving at all
- $+$: h is moving to the right side of $\overrightarrow{hr}$

r_2) movement of r w. r. t. the line $\overrightarrow{rh}$ at time t:

The same as h_2), but with h and r swapped

For example, Figure 2 shows an interaction: a human is moving towards the robot ($h_1 = -$) and the robot is approaching too ($r_1 = -$). The human is directly headed towards the robot ($h_2 = 0$) but robot is towards its left side ($r_2 = -$).

3.3 Classes of HRSI

In [15] 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 Figure 3) 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 from the Left (PCL): The robot has to slow or stop movement to allow the human to move across the robot's intended path from the robot's left side.
- Path-Crossing from the Right (PCR): The robot has to slow or stop movement to allow the human to move across the robot's intended path from the robot's right side.
- Rejection: Any situation in which a human is detected, but a qualitative constraint of the robot movement as described in Section 3.5.2 and 3.5.3 is not required. E .g., Robot Meets Stationary Human (RMSH): The robot moves toward the stationary human, and stops when close. We record trajectory pairs from examples of RMSH situations to test our multi-HMM classifier's ability to reject such situations. For these situations we apply a default constraint, i.e. Section 3.5.1, enforcing safe and comfortable distance between human and robot.

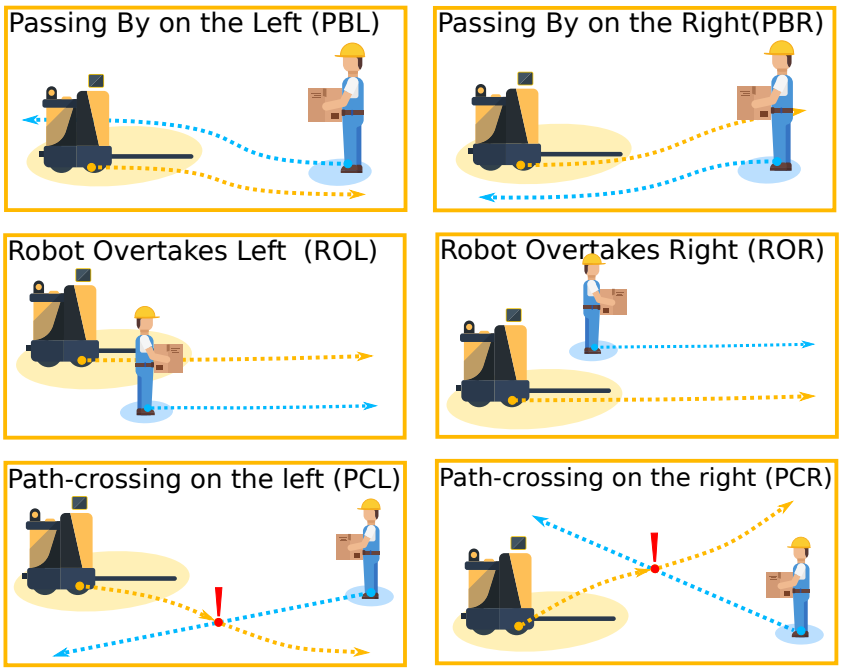

Figure 3: HRSI Classes

3.4 Creating HMM for our Multi-HMM Situation Classifier

We consider in this study 6 classes $C = (\text{'PBL'}, \text{'PBR'}, \text{'ROL'}, \text{'ROR'}, \text{'PCL'}, \text{'PCR'})$, exluding 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, a $|Q| \times |Q|$ 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 $Q = ((--\,--)\ldots(+\,+\,+\,+))$, as in [8] 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 initialise matrix B almost as an identity matrix with some noise, with diagonal $B[i,i] = t$ and the rest of elements $B[h,o] = \frac{1-t}{|Q|-1}\ |\ (b \neq o)$.

A list S of recorded QTC_C state sequences, 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

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

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 [16]. If the Kullback-Leibler (KL) divergence [17] of these log-likelihoods, normalised to sum to 1, from a uniform distribution of the same size is greater than a given threshold then the sequence is instead rejected.

3.5 Situational Constraints

The goal of our constraints is to ensure more socially legible trajectories and to reduce the number of times the robot stops due to proximity of humans to its safety laser(s). It is also relevant within the warehouse robotics context that robots adhere to their initial paths, for the sake of fleet coordination and legibility.

Situational as constraints described here are an extension of the proposal in [18]. Originally, constraints were defined by cost-maps in velocity space. We aim now for a more generic solution using metric cost-maps, allowing constraints to be applied to any motion planners which rely on hierarchical cost-maps, but still maintain speed modulation in place. Hierarchical cost-maps also allows to include other constraint sources such as local obstacles detected by sensors, multiple human constraints, etc.

Hence, constraints can work in two ways. Firstly they can impose speed restrictions on the robot proportional to the cost of movement. Secondly they can be used to update the most efficient path, triggering a request for replanning when the maximum or average cost of a path passes a set threshold.

3.5.1 Default Costmap Constraints

A default constraint costmap is defined in order to create a safety area around the closest human when none of the situations modelled by our multi-HMM situation classifier are detected. This default costmap is influenced by the classical proxemics approach. Cell costs in this default costmap are defined using the following two parameters:

- Distance to human (D): to penalise reductions in the relative distance between human and robot.

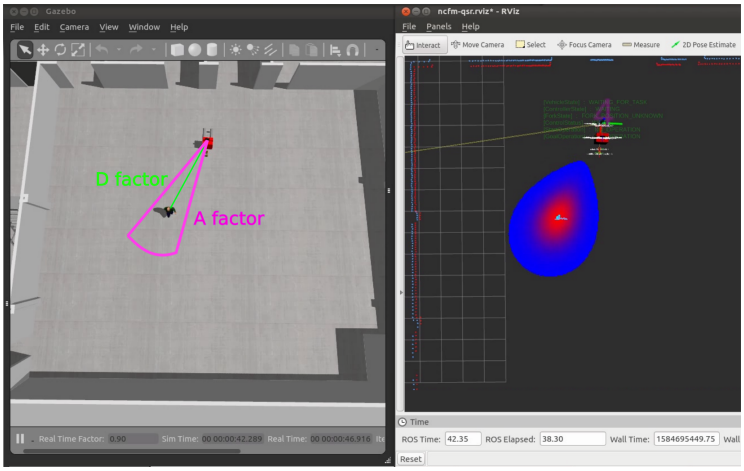

Figure 4: Default constraints, example in simulation.

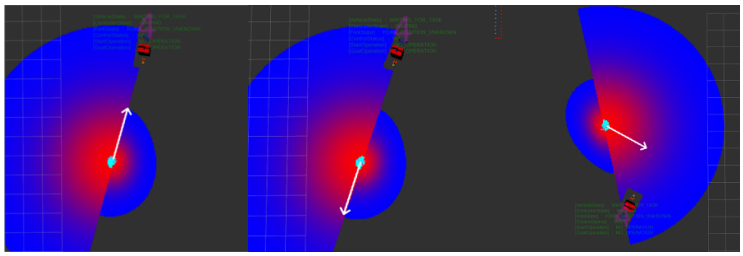

Figure 5: Block Left Constraint examples in simulation, for situations PBL, ROL, and PCL ordered from left to right. The white arrow indicates the human position and orientation.

- Arc around human-robot connecting line (A): to create a penalty arc around the human, this is, penalised spaces for the robot on the sides of the human.

Using these basic parameters, the cost map creates a basic exclusion area around humans that is always valid. Figure 4 illustrates an example of the default constraints implemented in the ILIAD safety stack.

3.5.2 Block Left

For the situations PBL, ROL, and PCL we assign higher costs to cells left of the line $\overrightarrow{hr}$, proportional to their distance to the human (D), but covering a much wider penalisation area than the default distance cost. We then apply the default constraint as described in Section 3.5.1 to cells which have not been assigned a cost, ensuring the robot maintains a comfortable distance from the human, even on the side which is not being 'blocked'. (See Figure 5.)

3.5.3 Block Right

For the situations PBR, ROR, and PCR, we generate the constraint cost-map using an identical approach to that described in Section 3.5.2, but instead assigning high costs to cells on the right side of the line $\overrightarrow{hr}$. (See Figure 6.)

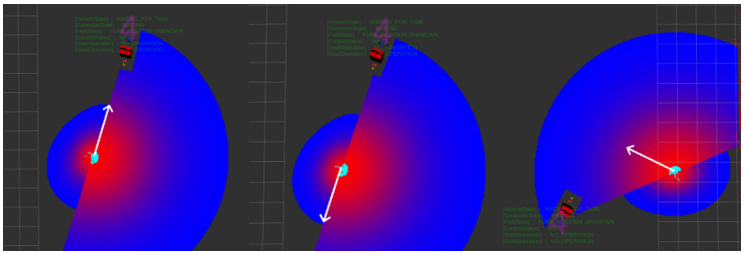

Figure 6: Block Right Constraint examples in simulation, for situations PBR, ROR, and PCR ordered from left to right. The white arrow indicates the human position and orientation.

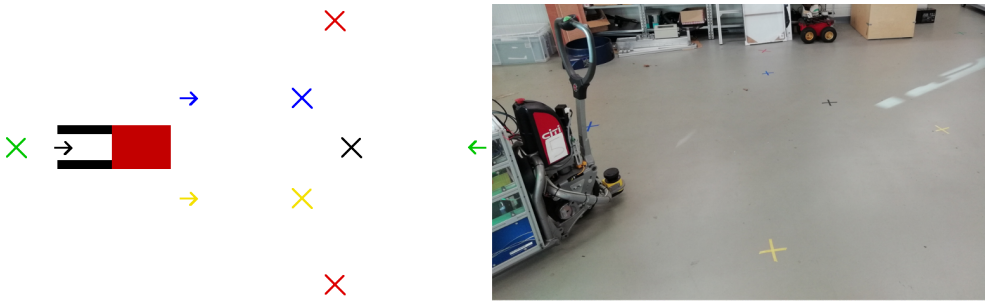

Figure 7: Illustration (left) and photograph (right) of laboratory setup for recording HRSI situations.

3.6 Validation: User Studies

3.6.1 Laboratory Setup for Recording HRSI Situations

The robot used for this validation is one of several automated pallet trucks belonging to the ILIAD Project. This one, 'robot5', is a modified Linde CitiTruck, equipped with a front-facing safety laser, a 3D lidar, a Kinect 2 RGBD camera, and a computer running Robot Operating System (ROS), interfacing with the sensors, and the pallet truck's motor controllers.

In the University of Lincoln (UoL) robotics lab we placed coloured tape on the floor, marking start and end positions for the human and the robot, as pictured in Figure 7. In the diagram on the left of Figure 7, 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, 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.

PCL – The human moves from the topmost red cross to the other red cross.

PCR – The human moves from the bottom red cross to the topmost red cross.

RMSH – The human stands stationary at the yellow cross, the robot moves from the black arrow to the black cross.

3.6.2 Training dataset

To train our multi-HMM classifier we recorded 35 interactions between a robotics expert and the robot for each of the 6 situation classes of Section 3.3, and 15 interactions for the rejection situation RMSH. A total of 225 HRSIs. We create the multi-HMM classifier's HMMs as described in Section 3.4, using QTC_C sequences generated from human and robot trajectories using *QSRlib* [19]. 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.

3.6.3 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 Section 3.6.1. These situations were enacted in a randomised order, with each participant also performing 1 of a set of 5 rejection situations, chosen at random. We realised in hindsight that only 1 of these situations, RMSH, would need to be rejected by our classifier as the others did not involve the robot moving. Thus with 11 participants, we recorded a total of 66 interactions, including 2 RMSH interactions. We created the multi-HMM classifier's **[!** ([**!**)HMM 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 50 ms to execute. Each classification took 60 ms to execute on average. The number of interactions recorded per class is detailed in the study's confusion matrix in Figure 8, which has its statistics explained in Section 3.6.4.

3.6.4 Results and discussion

Figure 8 is a confusion matrix containing metrics of the performance of the classifier at predicting the HRSI situation from QTC_C sequences, using training set sequences. 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.

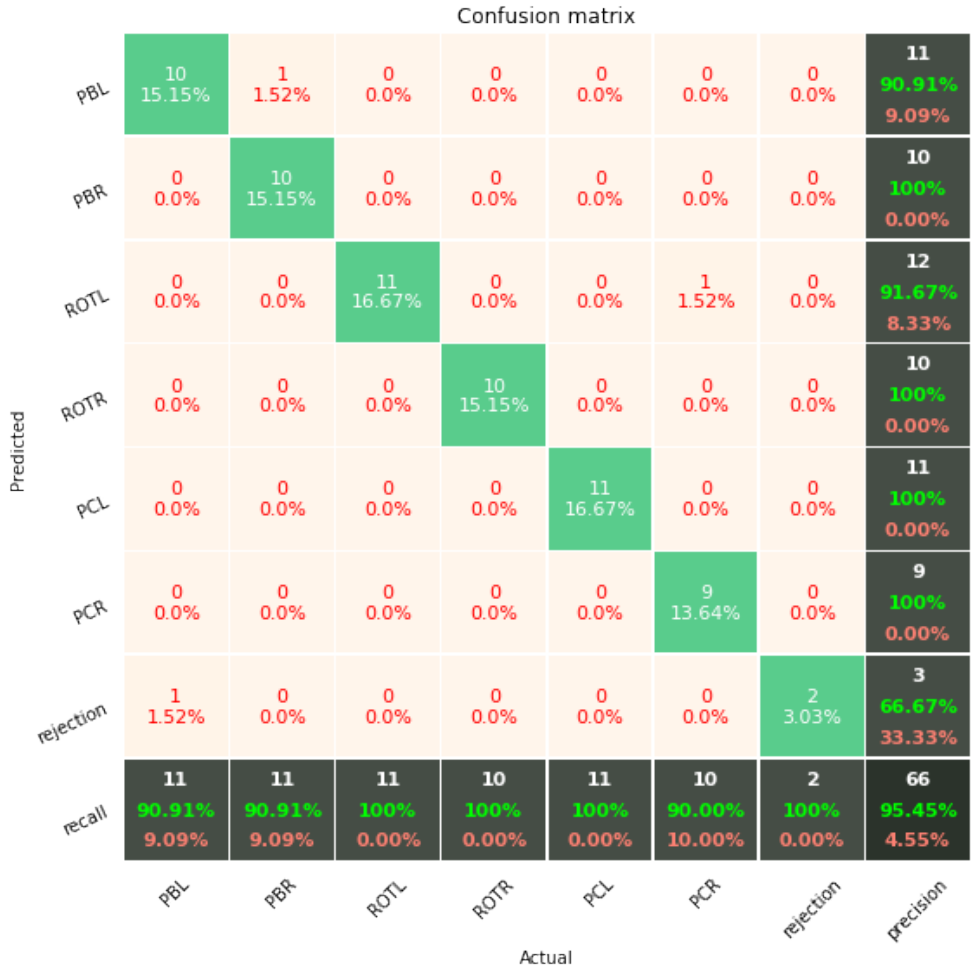

Figure 8: Confusion matrix for validation of our multi-HMM classifier using QTC_C sequences from study HRSIs as test data.

The classifier's overall accuracy is high at 95.45%, as it must be as a component of the safety focused HAN approach in ILIAD. The ability of our qualitative probabilistic model to accurately classify HRSIs with 11 non-experts, trained on HRSIs with 1 robotics expert, demonstrates the benefit of abstracting HRSIs to a qualitative description. It should be noted that while no rejection situations were misclassified, the precision of rejections is relatively low at 66.67%. This could be due to the small number of RMSH interactions recorded in our study, as explained in Section 3.6.3, as 3-fold Cross-Validation using the training dataset with its 15 RMSH interactions showed much higher rejection precision at 78.95%. The classifer is much less likely to confuse any of the other 6 classes, with precision and recall > 90% for all of these classes. We hope that we can improve the rejection precision and in turn the overall accuracy of our classifier by recording more examples of RMSH interactions and other rejection situations.

4 Mutual Communication of Navigational Intent

Navigation intent communication, the fourth safety layer in the ILIAD safety stack, ensures perceived safety in HRI. It comprises of bi-directional communication of navigation intent.

In the direction *robot-to-human*, we have developed a navigation intent communication system based on Spatial Augmented Reality (SAR) for AGVs. This enables the AGV to communicate its navigation intent using SAR such that humans can intuitively understand the robot's intention and feel safe in the vicinity of robots.

In the direction *human-to-robot*, eye gaze can convey information about intentions beyond what can be inferred from the trajectory and head pose of a person. Hence, we propose eye-tracking glasses as safety equipment in industrial environments shared by humans and robots. An implicit intention transference system was developed and implemented to recognize the human navigation intent and communicate it to the AGV. This allows pro-active safety approaches in HRI.

4.1 Implementation

The SAR based intention communication system was developed using a standard short-throw projector, Optoma X320UST, with 4000 ANSI lumens which was mounted pointing in the direction of the forks as shown in Figure 9. Ideally, the coverage of the projected floor space should enclose the area around the vehicle and be sufficiently large to allow displaying the intention of the vehicle over a time horizon of at least 3 s. It is, however, hard to realise full coverage of the whole 360 degrees around the robot. So, we selected the most important cone in the forward direction, which is sufficient as the robot drives only forward in our experiments and the projected area was always between the person and the robot during the encounters. The motivation behind the selection of the projector was to guarantee the visibility of projected intentions in brightly lit environments like warehouses and also to increase the field of view (FOV) so that the projected patterns are larger and clearly visible from a distance. The projector is connected to an on-board computer which renders images using an available pose estimate of the vehicle's location together with information regarding the current mission.

The rendering of images is done using the GLUT framework, using the reference frame provided by the navigation system. This allows us to render the image to be projected by updating the pose of a virtual camera (in the GLUT framework) using the localization estimate of the AGV along with extrinsic calibration parameters (i .e., the pose of the projector/virtual camera expressed in an AGV-fixed coordinate frame). The calibration approach for the projector display is described in [21].

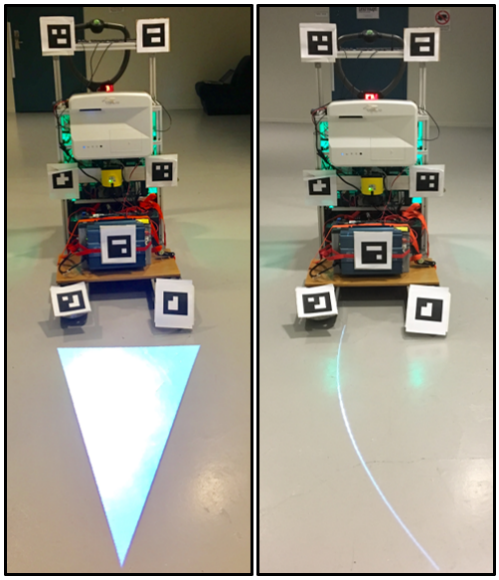

Figure 9: AGV communicating its future trajectory intentions. Four patterns were defined to be used in the experiments. 1: Line(Right), 2: Arrow(Left), 3: Blinking Arrow (Arrow that blinks at 1Hz) and 4: Nothing.

Eye gaze can convey information about intentions beyond what can be inferred from the trajectory and head pose of a person. We propose eye-tracking glasses as safety equipment in industrial environments shared by humans and robots. We have investigated the possibility of human-to-robot implicit intention transference solely from eye gaze data. Our results [20] have shown that, in the given scenario, a rather naive navigation intent predictor based on the simple rule, 'if people look more often to one side of the robot, they intend to go to that side,' would have predicted the correct navigation intention in 72.3 % of the encounters. Building upon this encouraging result that human navigation intention can be predicted from the eye gaze data, we developed an advanced implicit intention transference system, and implemented and tested it with workers at an industrial warehouse.

In the final system, the robot has access to human eye gaze data in real-time, and it responds in real-time to the received eye gaze data through SAR projections (see Figure 11). In order to achieve this functionality, we defined an Area of Interest around Robot (AOI-Robot) as shown in Figure 10 (label 5). This area of interest is used to decide when the person is looking at the robot, and includes an area that spans over the robot, projection and some area around the robot such that the robot would be in the field of view of the human. The eye gaze information is obtained through the eye-tracker worn by the human. We determine if the eye gaze is within the defined AOI-Robot or not using the Pupil Capture software, developed by the eye tracker manufacturer Pupil Labs. A network connection is established via ROS (see Figure 11) between the eye tracker and the robot and this is used to communicate the location of the eye gaze to the robot's SAR module and the robot responds to this information: if the eye gaze is within AOI-Robot, then the projected arrow remains static, and if the eye gaze is not within the AOI-Robot, then the projected arrow blinks to get the human's attention. (Video of the demonstration is available at https://youtu.be/lMEp6TcjDiw.)

This is intended as a proactive safety approach in HRI in industrial scenarios to ensure safety where the robot makes an attempt to get the human's attention when the human

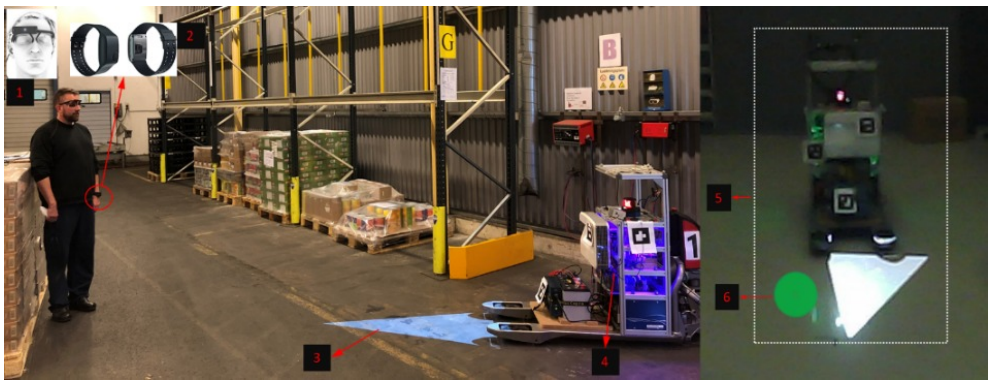

Figure 10: (1) Eye-tracker worn by a participant (2) Empatica E4 band used for measuring electrodermal activity. (3) SAR projection "Arrow" projected on the shared floor space to convey robot's navigation intent. (4) Robot (a retrofitted Automatic guided vehicle) with SAR intention communication system. (5) The defined area of interest AOI-Robot. (6) Eye gaze fixation.

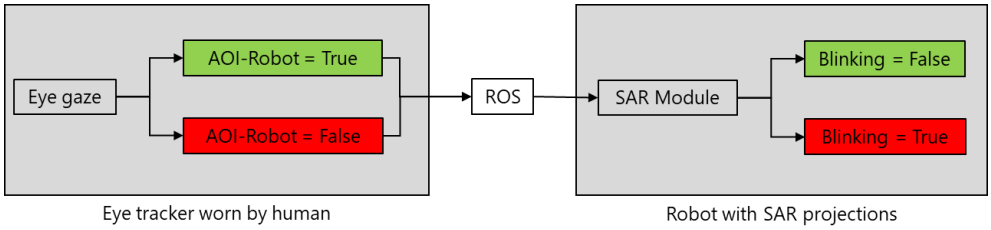

Figure 11: The methodology of implicit intention transference.

is in the vicinity of an AGV. The approach of blinking to get the attention is supported by our experiments that were recently published [22], which show that a blinking arrow immediately got the human's attention.

The developed system has won the first prize for software innovation at International Conference on Social Robotics 2019 held at Madrid, Spain.[1]

4.2 Validation: User studies

In order to evaluate the developed intention communication system, we conducted experiments where the humans encounter an AGV from different directions in different situattions. The robot used in these studies is the AGV 'robot1' located at Örebro University (ORU) and, as 'robot5' used in Section 3.6, is a modified Linde CitiTruck, but equipped with a projector as described above. The AGV projected various patterns on the shared floor space to convey its navigation intentions. We analysed trajectories, eye gaze patterns, Heart Rate (HR) and Electrodermal Activity (EDA) of humans while interacting with an autonomous forklift and carried stimulated recall interviews (SRI) and questionnaire based evaluations in order to identify desirable features for projection of robot intentions.

During these experiments, the robot followed pre-computed paths, not modified during the trials.

[1] Certificate available at https://orubox.box.com/s/fvvnj3o12od2y8ceyftt16j28h411067

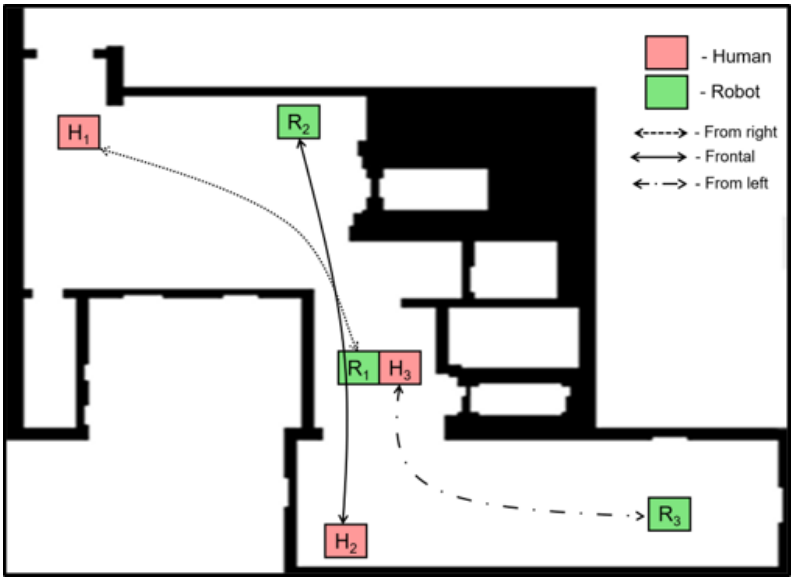

Figure 12: Experimental layout design: During the experiment, the humans (denoted by H_i) moved from H_i to R_i and the small forklift robot (denoted by R_i) moved from R_i to H_i.

4.2.1 Robot to Human Intention Communication study at ORU

Using the developed SAR based intention communication system, we studied how humans can intuitively understand the robot's intention and feel safe in the vicinity of robots. We conducted experiments using various patterns projected on the shared floor space to convey the robot's navigation intentions (for experimental layout see Figure 12). During the experiment every participant encountered the robot 4 times, each time with a different pattern being projected on the shared floor space. The four chosen patterns, shown in Figure 9, were:

A Line (indicating the path the robot intended to follow for the next 5 seconds)

B Arrow (indicating the current driving direction of the robot)

C Blinking Arrow (Arrow mentioned in (B) that blinks at 1Hz)

D Nothing (in order to observe the baseline behaviour when encountering robot)

Patterns A, B, C were supposed to convey the future trajectory of the mobile robot in an intuitive manner. They are not intended to replicate human-like intention communication, but rather exploring the possibilities that a robot has conveying its intentions to a human. Pattern A depicts the future trajectory over a time horizon of 5 seconds while patterns B and C communicate less detailed information using an arrow pointing along the instantaneous movement direction. The blinking arrow pattern C, blinks the B arrow pattern at a frequency of 1 Hz. The arrow was chosen for several reasons. Bertamini et al. [23] provide evidence that angles attract attention while Bar and Neta [24] suggest that the human brain can detect sharp features very fast as this can help to signal potential danger. Larson et al. [25] showed that a triangle with a downward-pointing vertex is recognised more rapidly than the identical shape with an upward-pointing vertex. Also the work in [26] used arrows to indicate the intention of their robot and concluded that their system is intelligible. Furthermore, people are used to arrows indicating directions

as in everyday life these are widely used. Accordingly, we believe that using an arrow to communicate the future path of a robot seems to be a good choice, as due to its angled v-shaped top it is supposed to attract the attention and might be detected faster than other symbols. Furthermore, it has already been used successfully and people already have a conception about the meaning of an arrow. This facilitates the understanding of this pattern and thus helps humans to understand the intention of the robot faster. In order to compare how a projection influences human behavior, the last condition D does not consist of any projection at all.

Using the eye gaze information, we were able to identify which pattern most strongly attracts attention. With the laser tracking data, we found that a mobile robot projecting its intentions on the shared floor space encourages humans to actively choose safer paths by not getting too close to the robot. From retrospective recall interviews, we found out that the 'ARROW' projection pattern was perceived being best suited for intention communication and a projected pattern on the floor made humans choose safer paths partially because it was perceived as a part of the robot. Our results also show that, in the given scenario, a navigation intent predictor based on the simple rule, 'If people look more often to one side of the robot they intend to go to that side', would have predicted the correct navigation intention in 72.3% of the encounters. Based on this result, we have further implemented a human-to-robot implicit intention transference system (as described in Section 4.1), and tested it with industrial workers at Orkla and at the ORU robot lab as described in the following.

4.2.2 Bi-directional Intention Communication Study at ORU Robot Lab

Experiments were conducted at the ORU robotlab where the participants interacted with the AGV when the implicit intention projections were enabled, and this was compared to the baseline condition of no projection and static projection. Data collected in the experiment[2] included the following:

1. Total time taken to finish the tasks.
2. Trajectory (recorded using a Qualisys motion capture system).
3. Eye tracking data using Pupil Pro glasses.
4. HR and EDA data from the Empatica E4, and also from the BioPac (recorded the data at a much higher frequency than the Empatica, the idea here is to compare the data from both the systems and recommend what system would be most suitable for evaluating such dynamic HRI scenarios).
5. Questionnaires.

This study was conducted in collaboration with the Psychology department at ORU, which includes a Master's thesis of two students from the psychology department titled 'Predicting attention allocation during human-robot interactions: An eye tracking study'. This study is a first attempt to apply the Salience Effort Expectancy Value (SEEV) attentional model to human-robot interactions in a simulated industry setting. The SEEV attentional model was originally developed to predict attention allocation in workspaces, taking into account the roles of Salience, Effort, Expectancy, and Value. Overt attention was measured via eye tracking in adult participants ($N = 18$) while they interacted with the AGV during a simple transportation task. The robot moved according to a pre-randomised pattern (turning left, turning right, or going straightforward) and projected an arrow cue

[2] Experiments were conducted in the end of November 2019 and analysis of this data is ongoing.

on the floor communicating its intent of movement (static projection, responsive projection, or no projection), resulting in nine conditions. Each participant underwent all nine conditions once. The model parameters were set for three predefined areas of interest (robot, robot periphery, projection) following previous applications of the SEEV model and taking into account the specificities of the current HRI. The results showed a poor model fit for the SEEV model when predicting the participants' eye gaze. These results suggest that the original SEEV-model for predicting attention allocation does not apply to the current HRI, and that a more adequate model needs to be developed. Processing of the trajectory, and psychophysiological data is ongoing in order to build spatio-temporal gaze patterns with the associated HR and EDA data for performing further analysis.

4.2.3 Implicit Intention Transference Study at Orkla

In order to validate the implicit intention transference system with the industrial workers, a study was conducted at Orkla Foods in Örebro, Sweden, where the industrial workers acted as the participants. During the experiment,participants were asked to observe the robot and its projections to identify the behavior of projections with respect to their eye gaze and were asked to say it out loud when they had guessed the behaviour (a link to the video from Orkla: https://youtu.be/ov8q_KXB2a4). The time that they were trying to guess the behaviour was timed using a stopwatch and during this time, an Empatica E4 band was used to measure the EDA of the participants. Seven participants have participated in this study. All the participants have understood the behaviour of the projections with respect to the eye gaze. The time they needed to understand the behaviour of projections with respect to eye gaze was 17±2.3 seconds ($N = 7$). We take this as an indication that the designed behavior of projections that respond to the eye gaze was intuitive and easy to comprehend. They were also verbally asked if a system like that would be useful when working with AGVs that are operating freely unlike the current AGVs which stick to a defined path. They have verbally opined that such a system could indeed be useful.

A significant increase in the EDA data was seen when the participants were trying to understand the behaviour of the projections. A rise in the EDA data is an indicator of an increase of cognitive load and stress [27] and considering the newly introduced intention communication system, the rise in the EDA is understandable. We suggest the usage of EDA as a potential training tool for industrial workers to measure their progress in training in a quantitative manner; i .e., by measuring how the stress levels and cognitive load vary during training. Apart from using the eye tracker as an intention communication tool, we have used it to record the pupil diameter during the experiment. We noticed that the pupils were dilated when the participants were doing the task despite being exposed to bright projections, which would normally result in a decrease of pupil diameter. This is another indicator of increased cognitive load [28]. However, further analysis needs to be conducted on the collected data to determine whether the increase in pupil dilation is due to increased cognitive load alone.

4.2.4 Focus Group Study at Orkla

A focus group study about modalities of communication for AGVs was conducted at Orkla foods on Nov 20, 2019. This focus group was conducted in collaboration with experts in human factors at Skövde University and RISE Viktoria in Göteborg. A total of 7 workers from Orkla foods have participated in this focus group. Two focus group sessions of 4 and 3 participants respectively were conducted. The participants were asked about their general challenges in working around the AGVs and investigated about what is working well and what could be improved, and then further discussed what the future holds in terms of communication modalities of AGVs. Audio recordings of approximately 4 hours

(in Swedish) were recorded during these sessions, which is now transcribed, translated to English and ready to be analysed. Another interesting aspect here is that the two groups had certain differences in terms of their adaptability towards technological developments – one of the groups was pro-technology and the other was not, which could be very interesting to investigate. The obtained insights from this focus group could be highly valuable for the future of intention communication abilities of AGVs, and the results from this study are expected to be submitted for publication in July 2020.

4.3 Ongoing and future work

4.3.1 Survey Article on Application of Eye Tracking in HRI

As a part of the literature study for the conducted work, a survey article on the topic 'Application of eye tracking in HRI' is in progress. The aim of the article is to give a literature overview, focusing on the evolution of interest in this area of research, types of hardware, applications, and analysis methods. The identified works are classified into following categories: (a) Using eye tracking to study gaze behavior, (b) Using eye tracking to predict human intent, and (c) Using eye tracking to control a robot. We expect that this work would be a useful contribution for the HRI research community. This work is expected to be submitted for a publication in May 2020.

4.3.2 Anthropomorphic Intention Communication

System development is ongoing for communicating intentions of AGV in an anthropomorphic manner by placing a humanoid robot on the AGV. A humanoid robot which is connected with the AGV would serve as a broadcaster of the AGV related intention communication to the human workers in its vicinity by using gestures, speech and head orientation. This is expected to be a natural way of intention communication for the workers cohabiting the environment. A Master student is working on this project full-time since March 2020 and experiments are planned for July-August 2020 in the robot lab at ORU.

5 Conclusions

This work has presented the proposed safety stack for ILIAD as a whole, describing the roles of each one and their implementations. Each layer in the stack is designed to address safety at a different scope, trying to reduce the activations of the layers below it. I .e., an efficient flow-aware mapping should minimize interactions and as a result reduce potential replannings or speed restrictions. Similarly, a reasonable speed profile should prevent emergency stops from the robot.

Some of the elements of this safety stack have already been introduced in other deliverables, leaving the focus of this work for the intermediate layers dealing with local Human Robot Spatial Interaction (HRSI) and navigational intent communication. Results in mutual communication of navigational intent show positive results attracting human's attention to the AGV in their field of view and an increase in the cognitive load. These facts point to a greater human awareness of robot intentions. Future works on the performed experiments will determine the level of acceptance and perceived safety obtained from this approach.

The Continuous Trajectory Assessment (CTA) module delivers both situation-aware planning and extended Vehicle Safety Motion Unit (vSMU) based on the outputs produced by the HRSI classifier. We have presented here the framework used in the ILIAD Project

to classify HRSI situations. If global planning is unable to create paths avoiding Human Robot Interaction (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. 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 Human-Aware Navigation (HAN) approach, with some room for improvement in the precision of rejection, which may be possible by taking the steps described in Section 3.6.4. Future work will evaluate the performance of this approach in the project's overall safety architecture, comparing the baseline system against the HAN system and measuring the the legibility of the chosen paths and the perceived safety of the participants.

References

[1] Christian Dondrup et al. Real-time multisensor people tracking for human-robot spatial interaction. In *Workshop on Machine Learning for Social Robotics at ICRA 2015*, Seattle, WA, May 2015. ICRA / IEEE.

[2] Timm Linder, Stefan Breuers, Bastian Leibe, and Kai O. Arras. On multi-modal people tracking from mobile platforms in very crowded and dynamic environments. In *2016 IEEE ICRA*, pages 5512–5519, Piscataway, New Jersey, USA, May 2016. IEEE.

[3] T. Vintr et al. Time-varying pedestrian flow models for service robots. In *2019 European Conference on Mobile Robots (ECMR)*, pages 1–7, September 2019.

[4] Luigi Palmieri, Tomasz Kucner, Martin Magnusson, Achim J. Lilienthal, and Kai O. Arras. Kinodynamic motion planning on Gaussian mixture fields. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 6176–6181, May 2017.

[5] Chittaranjan Srinivas Swaminathan, Tomasz Piotr Kucner, Martin Magnusson, Luigi Palmieri, and Achim Lilienthal. Down the cliff: Flow-aware trajectory planning under motion pattern uncertainty. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS)*, pages 7403–7409, 2018.

[6] Tomasz Piotr Kucner, Achim J. Lilienthal, Martin Magnusson, Luigi Palmieri, and Chittaranjan Srinivas Swaminathan. *Probabilistic Mapping of Spatial Motion Patterns for Mobile Robots*, volume 40 of *Cognitive Systems Monographs*. Springer International Publishing, 2020.

[7] Edward T. Hall et al. Proxemics. *Current Anthropology*, 9(2/3):83–108, 1968.

[8] Nico Van de Weghe et al. A Qualitative Trajectory Calculus and the Composition of Its Relations. In David Hutchison et al., editors, *GeoSpatial Semantics*, volume 3799, pages 60–76. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[9] Thibault Kruse et al. Human-aware Robot Navigation: A Survey. *Robotics and Autonomous Systems*, 61(12):1726–1743, December 2013.

[10] Manuel Fernandez Carmona, Tejas Parekh, and Marc Hanheide. Making the Case for Human-Aware Navigation in Warehouses. In Kaspar Althoefer et al., editors, *TAROS 2019*, LNCS, pages 449–453, Cham, 2019. Springer International Publishing.

[11] Christoforos I Mavrogiannis and Ross A Knepper. Multi-agent path topology in support of socially competent navigation planning. *The International Journal of Robotics Research*, 38(2-3):338–356, 2019.

[12] Henrik Kretzschmar, Markus Spies, Christoph Sprunk, and Wolfram Burgard. Socially compliant mobile robot navigation via inverse reinforcement learning. *The International Journal of Robotics Research*, 35(11):1289–1307, 2016.

[13] Marc Hanheide, Annika Peters, and Nicola Bellotto. Analysis of human-robot spatial behaviour applying a qualitative trajectory calculus. In *2012 IEEE RO-MAN*, pages 689–694, Paris, France, September 2012. IEEE.

[14] Christian Dondrup, Nicola Bellotto, and Marc Hanheide. A Probabilistic Model of Human-Robot Spatial Interaction Using a Qualitative Trajectory Calculus. In *AAAI Spring Symposium Series*, Palo Alto, California, USA, March 2014. AAAI.

[15] 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*, 4(1):63, 2015.

[16] Andrew M White. Sequence Classification - with emphasis on Hidden Markov Models and Sequence Kernels, September 2009.

[17] James M. Joyce. *Kullback-Leibler Divergence*, pages 720–722. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[18] Christian Dondrup and Marc Hanheide. Qualitative constraints for human-aware robot navigation using Velocity Costmaps. In *2016 IEEE RO-MAN*, pages 586–592, Piscataway, New Jersey, USA, August 2016. IEEE.

[19] Y. Gatsoulis et al. QSRlib: a software library for online acquisition of qualitative spatial relations from video. In *29th International Workshop on Qualitative Reasoning*, pages 36–41, California, USA, July 2016. IJCAI.

[20] R.T. Chadalavada, H. Andreasson, M. Schindler, R. Palm, and A.J. Lilienthal. Accessing your navigation plans! human-robot intention transfer using eye-tracking glasses. In *Advances in Transdisciplinary Engineering*, volume 8, 2018.

[21] R. T. Chadalavada, H. Andreasson, R. Krug, and A. J. Lilienthal. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In *2015 European Conference on Mobile Robots (ECMR)*, pages 1–6, 2015.

[22] R.T. Chadalavada, H. Andreasson, M. Schindler, R. Palm, and A.J. Lilienthal. 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.

[23] Marco Bertamini, Letizia Palumbo, Tamara Nicoleta Gheorghes, and Mai Galatsidas. Do observers like curvature or do they dislike angularity? *British Journal of Psychology*, 107(1):154–178, 2016.

[24] Moshe Bar and Maital Neta. Visual elements of subjective preference modulate amygdala activation. *Neuropsychologia*, 45(10):2191–2200, 2007.

[25] Christine L Larson, Joel Aronoff, Issidoros C Sarinopoulos, and David C Zhu. Recognizing threat: A simple geometric shape activates neural circuitry for threat detection. *Journal of cognitive neuroscience*, 21(8):1523–1535, 2009.

[26] Daisuke Matsui, Takashi Minato, Karl F MacDorman, and Hiroshi Ishiguro. Generating natural motion in an android by mapping human motion. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS)*, pages 3301–3308. IEEE, 2005.

[27] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Troster, and U. Ehlert. Discriminating Stress From Cognitive Load Using a Wearable EDA Device. *IEEE Transactions on Information Technology in Biomedicine*, 14(2):410–417, March 2010.

[28] Pauline van der Wel and Henk van Steenbergen. Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin and Review*, 25(6):2005–2015, December 2018.