Towards Benchmarking Human-Aware Social Robot Navigation: A New Perspective and Metrics

Phani-Teja Singamaneni, Anthony Favier, Rachid Alami

Towards Benchmarking Human-Aware Social Robot Navigation: A New Perspective and Metrics

Phani Teja Singamaneni[1], Anthony Favier[1,2] and Rachid Alami[1,2]

***Abstract*— Human-aware robot navigation planning enables robots to traverse human-occupied spaces socially. However, evaluating and benchmarking the 'human awareness' of such navigation schemes is challenging. With the growing necessity and research interest in the field, there is a need to define metrics to quantify and benchmark such qualities. In this regard, this paper proposes a set of metrics by looking at the problem from a new perspective. These proposals are made by inspecting the robot's navigation from the viewpoint of a human experiencing it and then defining proxies for the perceived human feelings. Analyses of some commonly occurring human-robot navigation scenarios using these metrics show their capability in benchmarking and differentiating human-aware robot navigation from standard robot navigation.**

I. Introduction

Benchmarking the performance of a **H**uman-**A**ware robot **N**avigation (HAN) (also called social robot navigation) planner is one of the open questions in the field. Although there are a number of metrics to evaluate the navigational performance of a robot [1], assessing the human-aware aspect remains challenging. The changing customs and societal rules could be partly attributed to the lack of standardisation. The existing approaches measuring the human's discomfort [1] are highly dependent upon Hall's Proxemics [2] theory, and it has been shown that the proxemic distances could change or be violated [3], [4] depending on place and context. In intricate scenarios like the robot crossing a narrow corridor or taking an elevator, the robot has to intrude into the personal space of humans. Given the circumstances, the human might not feel uncomfortable or even allow it to happen, but the proxemics might say otherwise. Therefore, it is necessary to study new ways to benchmark HAN and devise metrics that apply to a large set of human-robot navigation scenarios.

Navigation in the presence of humans requires an understanding of the dynamics of the situation at hand. As humans are social beings that perceive the environment and make choices, the robot's actions can affect these decisions and inflict different kinds of feelings. Hence, the robot should be more aware of how its actions can influence human partners and adjust its behaviour based on the reactions. For example, depending on how the velocity [5], [6] and the implicit gestures [7], [8], [9] are handled, the robot's behaviour could be more confusing or welcoming to the humans. Although such studies are used while modelling a HAN system, they are seldom used for evaluation. Taking these into consideration, we propose metrics that could act as proxies for different emotions experienced by humans when they interact with the robot in a navigational setting. The major contributions of this paper are twofold: 1) a new set of metrics are proposed for benchmarking HAN, and 2) a detailed analysis showing the application of these metrics to effectively benchmark the 'human awareness' of a system. This analysis is purely metric and is performed in simulation.

Before moving on to new metrics, we briefly present some of the commonly used ones in section II. Following this, we present the proposed metrics and their mathematical formulation in section III. In section IV, the ability of the proposed metrics for benchmarking HAN is put to test in different human-robot navigation scenarios using a human-aware planner and a standard planner. Finally, we present the discussion and conclusions in section V.

II. Existing Evaluation methods

Based on the classification proposed by Gao et al. [1], all the existing evaluation metrics can be broadly divided into four categories namely, navigation, naturalness, discomfort and sociability. The navigation metrics are used to test the robustness, stability and reliability of a planner. They include path length, path efficiency, relative throughput [10], time to reach the goal, rate of success, number of collisions, etc. A minimum expectation from a HAN planner is good performance even in the absence of humans and these metrics could be used to benchmark the performance. Many works in HAN include such evaluations [11], [12], [13] along with the other evaluation strategies.

Naturalness metrics are applied to measure the smoothness of the robot's trajectory and its similarity in comparison to a human trajectory in the same setting. The similarity in trajectory is quantified using metrics that measure the path deviation like Average Displacement Error (ADE) [14], Final Displacement Error (FDE) [15] and their variations. These are generally employed to assess the performance of human prediction algorithms that are used in HAN and seldom for measuring the robot's performance. However, the smoothness measures like cumulative heading changes [12] or the path irregularity [16], velocity and acceleration profiles, and the topological complexity [10], [17] are used to assess the robot's performance around humans. The smoothness metrics and the topological complexity index [18] are sometimes used to measure the legibility [19] of the robot's motion. Legibility comes with expressiveness, and it is necessary to define some metrics for measuring navigation intention expressiveness.

[1]LAAS-CNRS, Université de Toulouse, France.
[2]ANITI, Université de Toulouse, France
Email:{ptsingaman, anthony.favier, rachid.alami}@laas.fr

HAN is essentially a human-robot interaction in the context of robot navigation. Therefore, it is required to measure and quantify human-robot navigation interactions to benchmark the human awareness of the system. Discomfort metrics aim to do it by assessing such interaction and telling how well a system is performing. However, as mentioned previously, most of the existing metrics are distance based and largely rely on the proxemics theory [2]. The intrusions into personal and intimate spaces are usually taken as a measure of discomfort, and many works [20], [21] count the number of these intrusions to quantify the performance of a system. Some works define their own performance metrics [22], [23] based on proxemic rules.

Similar to individual humans, a group of humans also have various interactions zones like *o-space*, *p-space* and *r-space* [23], [24], and their shape depends on the *f-formation* [25] maintained by the members of the group. In such settings, the intrusion into *p-space* and *r-space* are considered as violations and HAN systems try to minimise these intrusions. Sometimes, the human-object interactions are also considered [11], [22] while measuring these intrusions. An alternative to the number of intrusions is to measure the time spent in the areas associated with the human's personal zone or the group's interaction zone [26]. The minimum distance from the human is another metric that is widely used for benchmarking HAN. One of the very useful metrics combining the velocity and distance is called *Time-to-Collision (TTC)* [27], and it represents how the robot's motion is relative to the human's motion. One of the metrics proposed in this paper is based on *TTC* which is detailed in the subsequent sections.

Human interactions and their psychological impressions and states cannot be quantified using just numbers. It requires some well-designed studies and questionnaires, and experimental evaluations on a real system or through videos. HAN does the same to evaluate the psychological safety (discomfort and stress) and the sociability of the system under study. The perceived psychological safety is commonly measured using questionnaires [28]. Some of the established questionnaires in social robotics like Godspeed [29] already include perceived safety and emotional states and the Robotic Social Attributes Scale (RoSAS) measures several psychological factors based on the Godspeed questionnaire. The social intelligence of a robot, sometimes called sociability, is not very easy to quantify. A robot's motion may be perfectly natural, satisfies all the comfort-based metrics and can still be perceived as not completely social. In several intricate scenarios, the robot might not reach the high-level expectations of a human and fail to convey its intention. Barchard et al. [30] proposed Perceived Social Intelligence (PSI) scales to evaluate around 20 aspects of the social intelligence of a robot. This was used in some of the recent works [31], [32] to evaluate the HAN systems. Some works [33] have employed custom questions apart from these scales to study sociability.

Unlike sociability, the human awareness of a robot could be quantified to some extent and discomfort measures try to measure it. The metrics proposed in this paper can be seen as discomfort metrics as they act as proxies for different emotional responses that are closely related to human comfort.

III. Proposed Discomfort Metrics

One of the recent studies by Joosse et al. [4] revealed that people are more lenient when a robot intrudes on their personal space than a human. Further, they showed that these intrusions could be mitigated by conveying the robot's intention to the humans. The velocity around the humans also plays a crucial role, as mentioned earlier [5], [6], [28]. These studies and observations lay the basis for our metrics based on velocity and distance, called the **Danger Metrics**. The second set of metrics is based on the human's visibility and recognition of the robot when it appears unexpectedly or when it is trying to approach a human for an interaction. These are called the **Surprise Metrics**. We introduce each of these metrics and provide their mathematical formulation in this section.

A. Danger Metrics

Depending on the velocities of humans and the robot, we define two metrics (or costs) that act as proxies for the emotions in danger. Fig. 1 shows two scenarios where the human and robot can approach each other, or the robot can pass by the human. In these situations, if the robot is moving head-on towards the human, there is a possibility of collision, and it also makes the human uncomfortable [28], [5]. To measure such cases, we define our first danger metric called $cost_{fear}$, which tries to measure the perceived feeling of fear of collision with the robot. It was also shown that humans prefer the robot to modulate its velocity [6] while crossing and prefer lower speeds [28], [5] in close vicinity. Therefore, if something or someone passes at a very high speed in close vicinity, it could cause distress and panic. The second danger metric, $cost_{panic}$, is defined to measure such perceived feelings of panic.

Mathematical Formulation: Suppose the circumscribed circle of the human has a radius of r_h and that of the robot has a radius of r_r. The human and the robot collide when the distance between their centres is less than or equal to the sum of these radii, $r_h + r_r = R$. If we assume the robot is a point and expand the radius of the human to R, the same conditions apply. This setting is shown in Fig. 2 along with the velocities of the human, $\overrightarrow{V_h}$ and the robot, $\overrightarrow{V_r}$. The relative velocity of the robot with respect to the human is given by $\overrightarrow{V_{rel}} = \overrightarrow{V_r} - \overrightarrow{V_h}$ and depending on where it falls

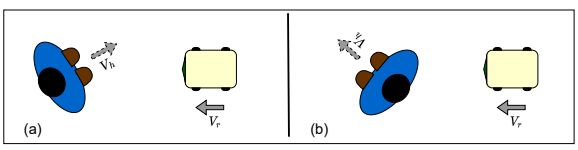


Fig. 1: Situations where Danger Metrics are important. (a) The human and the robot approach or cross each other. (b) The robot is behind the human and is about to overtake or pass by the human. In both situations, the human can be either static or moving, but the robot is always moving.

(see Fig. 2), we define the two metrics for danger. If it falls within the collision zone (cone formed by dotted lines), then there is a danger of collision, and we define the $cost_{fear}$ in this setting. If it falls outside the collision zone, the danger of collision no longer exists, but the robot may pass by the human, and so here, we define the $cost_{panic}$. Let the vector from the robot's position to the human's position be $\overrightarrow{P_{rh}}$ and θ be the angle between $\overrightarrow{V_{rel}}$ and $\overrightarrow{P_{rh}}$. While defining these costs, we use the effective distance between the human and the robot, $d_{rh_{eff}}$ and the perpendicular component of $\overrightarrow{P_{rh}}$ along $\overrightarrow{V_{rel}}$, $d_{\perp}$.

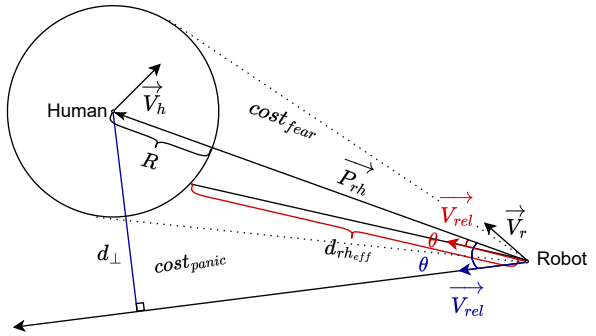


Fig. 2: Different vectors and the possible danger costs. If $\overrightarrow{V_{rel}}$ falls within the zone indicated by the dotted lines, there is $cost_{fear}$ and if it falls outside this zone, we have the $cost_{panic}$. Only one of the $\overrightarrow{V_{rel}}$ (blue or red) could exist at a time.

1) Cost of fear: The first metric for danger, $cost_{fear}$, is defined as the inverse of *TTC* and when $0 < TTC < \infty$, it is calculated as,

$$\boxed{cost_{fear} = \frac{1}{TTC} = \frac{\|\overrightarrow{V_{rel}}\|}{d_{rh_{eff}}}} \quad (1)$$
$$\ni \mathbb{P}.\mathbb{V} > 0 \text{ and } d_{rh_{eff}} > 0$$

where $\mathbb{P}.\mathbb{V} = \overrightarrow{P_{rh}} \cdot \overrightarrow{V_{rel}}$, $\cdot$ is the dot product and $\|\|$ is the magnitude of a vector. Note that $cost_{fear}$ is only defined when the $\overrightarrow{V_{rel}}$ falls within the collision zone. When the $\overrightarrow{V_{rel}}$ falls outside the collision zone, the $cost_{fear} = 0$ and as $d_{rh_{eff}}$ decreases or $\|\overrightarrow{V_{rel}}\|$ increases, $cost_{fear}$ increases. This cost should be maintained as low as possible to decrease the influence of the robot on humans. So, the HAN planners can be designed to have a certain threshold for this cost beyond which some mitigating actions are required.

2) Cost of panic: The second metric that we propose is valid when $\overrightarrow{V_{rel}}$ is outside the collision zone. In such a setting, the robot may cross or pass by the human, and we define the $cost_{panic}$ as,

$$\boxed{cost_{panic} = \frac{\|\overrightarrow{V_{rel}}\|}{d_{\perp} - R}|sin(\theta)|} \quad (2)$$
$$\ni \mathbb{P}.\mathbb{V} > 0 \text{ and } d_{\perp} > R$$

The $cost_{panic}$ increases as the robot approaches the human. Specifically, the cost increases as the perpendicular distance of the robot from the human decreases or the relative velocity increases. It also increases as θ increases, indicating that the robot is getting closer to the human. As this cost is trying to capture the panic caused when the robot passes by a human, a HAN system should try to minimise this cost whenever the robot is trying to overtake or cross a human.

B. Surprise Metrics

Human environments are highly dynamic, and humans rely on their senses to navigate and adapt to changing conditions. However, if there is not enough time and space to identify and react to such changes, it could be dangerous and might surprise or shock humans. Therefore, if something or someone appears suddenly and passes by closely, the human might not feel comfortable. Especially when a robot is entering the field of view (FoV) of the human from the back, it should use socially appropriate paths [34] and maintain a larger distance. To reduce the excess work and make humans more comfortable, the robot should enter the FoV at a lower angle and not be within the peripheral vision. Fig. 3 illustrates two scenarios where the robot can surprise the human. The depiction on the left shows a robot trying to enter the human's FoV, and the one on the right shows the sudden appearance of a robot in front of the human without having any knowledge about the whereabouts of the human.

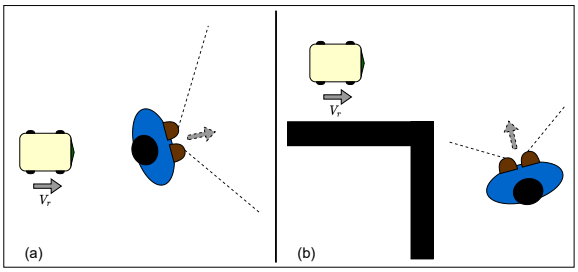


Fig. 3: Scenarios where a robot can surprise the human. (a) The robot needs to go from the back to the front to initiate the interaction or move forward. (b) The robot and human suddenly face each other without any prior knowledge about the other. The human can be static or dynamic in both these settings.

Encompassing all these, we define three metrics to quantify the surprise perceived by humans because of the actions of the robot. These metrics try to measure the feeling of surprise or shock occurring from different dimensions, unlike the danger metrics, which try to measure different emotions. The first cost, $cost_{visibility}$, measures the surprise based on the distance and angle, while the next two, $cost_{shock}$ and $cost_{react}$, measure it in terms of time. The idea behind $cost_{visibility}$ is to measure how well a HAN system adapts to the FoV of humans to reduce the surprise. The second set of metrics, $cost_{shock}$ and $cost_{react}$, measures the performance of HAN planners in handling the sudden emergence of occluded humans. Specifically, $cost_{shock}$ tries to measure the shock that can occur when something happens before it can be recognised, while $cost_{react}$ tries to check if there is enough time to take mitigating actions when surprised.

Mathematical Formulation: Considering a similar formulation as above, the robot is represented as a point, and the

circumscribed radius of the human is taken as $R = r_h + r_r$. The vector from the human's position to the robot's position is represented by $\overrightarrow{P_{rh}}$, and the unit vector in the direction of orientation of the human is given by $\hat{u}_h$. The angle between

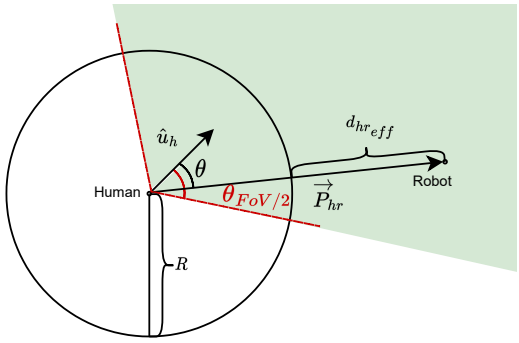


Fig. 4: Robot's appearance within the FoV of the robot. θ is the angle between the unit vector of human's direction, $\hat{u}_h$ and vector $\overrightarrow{P_{hr}}$. $\theta_{FoV/2}$ is the half angle of a human's field of view and $d_{hr_{eff}}$ is the effective distance between the human and the robot.

$\overrightarrow{P_{rh}}$ and $\hat{u}_h$ is represented by θ and the half angle of human's FoV by $\theta_{FoV/2}$. The effective distance between the human and the robot is given by $d_{hr_{eff}}$. When the robot enters the FoV of the human (green area in the Fig. 4), we estimate the surprise metrics using $d_{hr_{eff}}$, θ and some studies based on human perception.

1) Cost of visibility: As mentioned above, it is better to maintain a larger distance, $d_{hr_{eff}}$, when entering the FoV of a human and hence this cost should discourage close appearances. Further, it should increase as θ increases, indicating that more human effort is needed to look at the robot (head motion). Therefore, $cost_{visibility}$ is directly proportional to the angle and inversely proportional to the distance and can be calculated as below,

$$\boxed{cost_{visibility} = \alpha \left(\frac{\theta}{d_{hr_{eff}}} \right)} \tag{3}$$

where α is a constant which can be taken as $\alpha = \frac{d_{proxemics}}{\theta_{FoV/2}}$ and $d_{proxemics}$ is the defined proxemics-based distance that does not intrude the personal space ($> 0.45\,\mathrm{m}$). The lesser the cost of visibility, the better the behaviour of the robot. Thus, the cost should be maintained under a threshold by the HAN system while planning the robot's motion.

2) Costs of shock and react: When a human sees something, it takes a few milliseconds before it is registered by the brain. After the recognition, it takes more time to generate a response. The average recognition time for a human is around $t_{recognise} = 150$ milliseconds [35], [36] and the average time to react for visual stimuli is between $t_{react} = 400 - 600$ milliseconds [37], [38]. If something appears very close to a human before it can be recognised, it could result in a shock. Even after recognition, if there is not enough time to respond, it could be dangerous and causes distress apart from the surprise. Hence, it is necessary for a robot to be more aware of its surroundings to avoid the occurrence of dangerous situations.

The costs ($cost_{shock}$ and $cost_{react}$) are defined based on the recognition and reaction times and the robot's distance from a human when it enters his/her FoV. They follow similar formulations as before with some differences. We first define a linearly increasing function called the 'seen ratio' (SR) as given below,

$$SR = \begin{cases} \frac{t}{t_{react}}, & \text{if } 0 < t < t_{react} \\ 1, & \text{if } t \geq t_{react} \end{cases} \tag{4}$$

SR starts at zero the moment the robot enters the human's FoV, and as time, t, passes, this ratio slowly increases until it reaches one. When the value of SR is one, it means that the robot has been seen and identified by the human with enough time to react, if needed. Hence, we can say that the SR becomes one at around $t = t_{react} = 600$ milliseconds and continues to stay at one until the robot moves out of FoV of the human. The costs are now defined as follows,

$$\boxed{cost_{shock} = max\left(\frac{d_{proxemics}}{d_{hr_{eff}}} (1 - \gamma SR), 0 \right)} \tag{5}$$

$$\boxed{cost_{react} = \frac{d_{proxemics}}{d_{hr_{eff}}} (1 - SR)} \tag{6}$$

where $\gamma = \frac{t_{react}}{t_{recognise}}$. Both $cost_{shock}$ and $cost_{react}$ are high if the effective distance, $d_{hr_{eff}}$, is low when the robot enters the FoV of the human. $cost_{shock}$ lasts only for a short period, and if the robot comes very close to a human during this period, the human can be marked as surprised with a corresponding cost. As $cost_{react}$ measures how the robot's appearance affects the human, high cost means the HAN system has done a poor job in handling sudden appearances, and it needs to be improved.

C. Relevance of Metrics

The metrics can be calculated and used for evaluation in different ways depending on the scenario. One can study the evolution of the costs with time and see how the system responds to changing conditions. Another way is to study the maximum or minimum values over the course of a run or interaction. In this kind of evaluation, the relevant locations of occurrence of such costs give more information about the performance. For the costs proposed in this paper, a relevant location for the calculation of cost is required for it to be completely valid. The $cost_{fear}$ is valid as long as the human is in the FoV of the robot, but the $cost_{panic}$ is plausible only when the robot crosses the human or when it is at the shortest distance from the human during the run. For the surprise metrics, the location and the time when the robot first enters the FoV of the human are crucial. Hence, studying the first values of these costs could reveal more relevant information than their progression. Finally, for all the metrics (danger and surprise) proposed, it is ideal to set a threshold for acceptable behaviour, and these can be obtained based on real-world demonstrations and studies. However, even without such studies, we can perform a comparative analysis between

human-aware and non-human-aware planning systems using these costs. In the next section, we compare and analyse a standard and a human-aware navigation planner using the metrics calculated at relevant times and locations

IV. Experimental Analysis

Four different human-robot interaction scenarios in navigation are designed to test the ability of the proposed metrics in differentiating human-aware navigation from standard dynamic obstacle avoidance. We present a detailed analysis of each scenario to show how these metrics could be used to benchmark HAN. In all the experiments, $d_{proxemics} = 1.6\,\mathrm{m}$, $\theta_{FoV} = 120^\circ$ and $r_h = 0.3\,\mathrm{m}$. The interactions are simulated in the MORSE simulator, and the human agent is controlled using InHuS [39]. The robot is controlled using two different systems, CoHAN, a cooperative human-aware navigation planner [40], [41] and Simple Move Base (SMB), where humans are introduced as dynamic obstacles into the standard ROS navigation stack. It neither uses human motion prediction nor social norms while navigating. Hence, we expect our metrics to clearly differentiate SMB and CoHAN.

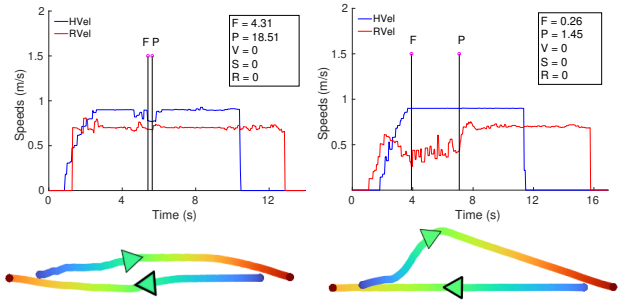
In all the figures used for the analysis, the multi-coloured paths show the progression of time from start (blue) to the end (red) of the run. The triangles with thin and thick borders are robot and human, respectively. The speed plots show the robot's velocity in red and that of the human in blue. These plots also show the relevant positions and values of the calculated metrics (F ($fear$), P ($panic$), V ($visibility$), S ($shock$) and R ($react$)) in one run of each case. Five such runs are performed for each scenario using SMB and CoHAN and the means of the costs are tabulated.

A. Scenario 1: Cross

In this setting, the robot and the human face each other at the start and cross paths as they move towards their goals at the other end. The averaged costs over 5 runs are shown in Table. I and Fig. 5 shows the paths and speed profiles for one of the runs. In this scenario, the robot is already in the FoV of the human as it starts moving. So, all the surprise metrics will be zero, and only the danger metrics are relevant here. Comparing the values of $cost_{fear}$ and $cost_{panic}$ from Table I and Fig. 5, the costs corresponding to CoHAN are significantly lower than costs for SMB. As CoHAN is a human-aware planner, it tries to provide more way for the human by moving away as shown in Fig. 5 (b) and slows down as it passes by the human. These behaviours result in low $cost_{fear}$ and $cost_{panic}$. On the other hand, SMB does not modulate its velocity much and takes only a small deviation to avoid the collision, as shown in Fig. 5 (a). The slight decrease in human speed and the path change can be seen in Fig. 5 (a). These behaviours may not be acceptable, and it is reflected by the high values in $cost_{fear}$ and $cost_{panic}$. Also, note that the times of the relevant costs are separated by large time in CoHAN and by a very small time in the case of SMB. This shows that CoHAN follows a well-planned trajectory while SMB does a last-minute collision avoidance.

Costs	*fear*	*panic*	*visibility*	*shock*	*react*
CoHAN	**0.33**	**1.38**	0.0	0.0	0.0
SMB	2.98	15.92	0.0	0.0	0.0

TABLE I: Averaged costs over 5 runs for cross scene

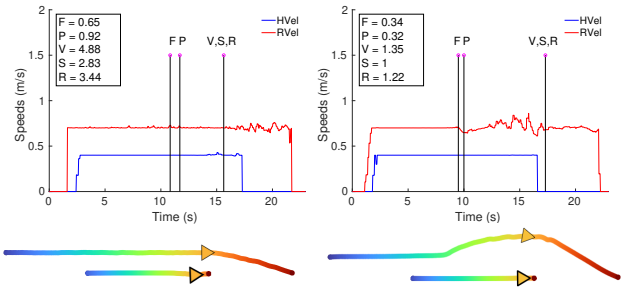


(a) SMB: Cross (b) CoHAN: Cross

Fig. 5: Cross scenario: (a) In the case of SMB, the robot moves close to the human, and the human path is slightly modified. As they cross each other, there is a slight decrease in the velocity of the human. (b) The robot running CoHAN moves away showing the intention and also keeping its distance from the human as it crosses. The human path is almost a straight line, and the velocity remains constant.

B. Scenario 2: Overtake

This scenario starts with the robot behind the human. During the run, the robot starts following a slow-moving human and finally overtakes the human to reach its goal as shown in Fig. 6. In this situation, a HAN system should have small values for all of the proposed metrics as the robot should pass by the human without colliding and not surprise the human as it overtakes him/her. Table II shows the mean values of the metrics calculated at relevant locations.



(a) SMB: Follow and Overtake (b) CoHAN: Follow and Overtake

Fig. 6: Overtake scenario: (a) The human and the robot move parallelly until the very end without much change in their velocity profiles. (b) The robot with CoHAN takes a larger deviation as it plans to enter the FoV of the human.

Costs	*fear*	*panic*	*visibility*	*shock*	*react*
CoHAN	**0.32**	**0.40**	**1.34**	**1.00**	**1.21**
SMB	0.47	0.94	4.72	2.75	3.34

TABLE II: Averaged costs over 5 runs for overtake scenario.

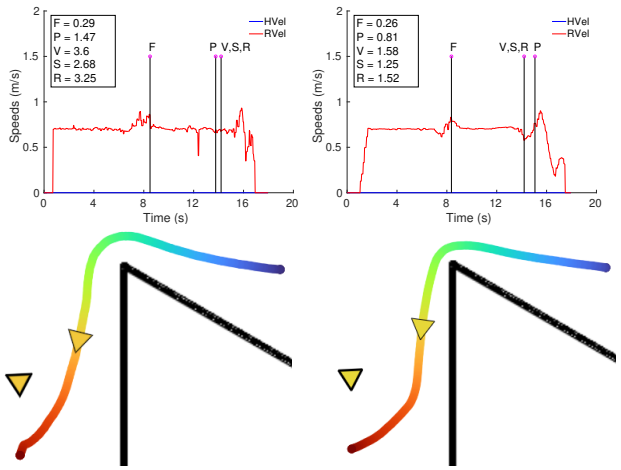
From the values in the table and the figure, it is evident that CoHAN performs significantly better than SMB. Specifically,

$cost_{shock}$, $cost_{react}$ and $cost_{visibility}$ are high for SMB, indicating that the robot might have entered the human's FoV suddenly and at a closer distance which is not ideal. It can be seen from Fig 6 (a) that the robot was very close as it overtook the human. The robot using CoHAN modifies its trajectory to accommodate the human and enters the human's FoV in a better manner, as seen in Fig. 6 (b) resulting in lower surprise metrics.

In this setting, the robot and the human have some perpendicular offset distance, unlike in the previous case, and they move almost parallelly for the most part (see Fig. 6). This is the reason for comparable values of $cost_{fear}$ and $cost_{panic}$ in SMB and in CoHAN. However, CoHAN maintains a larger distance as well as modulates the robot's velocity while passing by the human and this helps it to lower its costs of danger.

C. Scenario 3: Approach

This is a fairly simple setting with a static human. The robot has to approach and face the human for interaction. The robot starts somewhere behind the human and enters the FoV before aligning itself to face the human. The mean costs for the five runs are shown in Table III and as expected, a human-aware planner, CoHAN, has lower values for all the metrics. The value of $cost_{fear}$ is similar for CoHAN and



(a) SMB: Approach (b) CoHAN: Approach

Fig. 7: Approach scenario: (a) The robot moves straight towards the goal and slightly deviates to avoid a collision with the human. (b) The robot tries to move by the wall until it enters the FoV of the human. When the robot turns itself to face the human, CoHAN tries to keep the robot's velocity low.

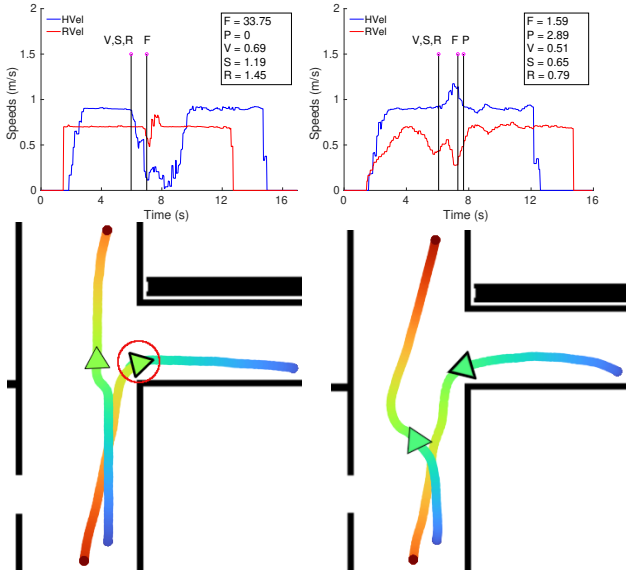
Costs	*fear*	*panic*	*visibility*	*shock*	*react*
CoHAN	**0.26**	**0.67**	**1.69**	**1.28**	**1.56**
SMB	0.28	1.51	3.75	2.71	3.29

TABLE III: Averaged costs over 5 runs for approach scenario.

SMB, but the $cost_{panic}$ for SMB is almost double both in Table III and Fig. 7. By observing the paths and the velocity profiles of the robot in Fig. 7, we can see that the robot using CoHAN moves closer to the wall and maintains its distance from the human until it enters the FoV. All the surprise metrics, therefore, remain low for the human-aware system than a standard navigation system. The speed profile in Fig. 7 (b) shows that the CoHAN moves the robot with a little lower velocity than SMB while aligning itself in front of the human.

D. Scenario 4: Appear

When a robot navigates in an environment, there could be several occluded regions from where a human might emerge. In this scenario, we put the planners to a test in the crossing scenario, where the human is initially occluded and becomes visible just before the cross. The human follows an L-shaped path as shown in Fig. 8. It is one of the most difficult cases to handle for a HAN system, as it has to prevent harm and shock to the human on top of avoiding the robot from freezing. As CoHAN already address such cases [41], it is expected to perform better in this case. From Table IV, we



(a) SMB: Appear (b) CoHAN: Appear

Fig. 8: Appear scenario: (a) The speed of the human drastically changes and the human even halts momentarily giving way for the robot (red circle on the path). The robot moves with almost a constant speed except for a little oscillation to avoid a collision. (b) The robot moves away from the corner and takes a deviated path in the anticipation of a human. The robot slows down as it sees a real human.

Costs	*fear*	*panic*	*visibility*	*shock*	*react*
CoHAN	**1.75**	3.16	**0.49**	**0.69**	**0.81**
SMB	30.85	**0.0**	0.73	1.21	1.48

TABLE IV: Averaged costs over 5 runs for appear scenario

can see that CoHAN performed better in terms of all the costs except $cost_{panic}$. The reason for this can be obtained by looking at Fig. 8. In the case of SMB, $cost_{panic}$ is not even calculated as the human lies in the collision zone all the time during the interaction. The very high value of $cost_{fear}$ clearly indicates that this system is performing really badly in the case of occluded humans. Further, the human has to slow down and wait for the robot to pass before continuing

the navigation as seen from the speed profile and the path of the human in Fig. 8 (a). By observing the region inside the red circle, one can see the discontinuity in the colour spectrum compared to the robot's path, indicating a halt (or oscillation) in human motion. CoHAN, however, handles this case by moving away from the corner anticipating a human and this results in a lesser possibility of collision and the feeling of fear ($cost_{fear}$). Although CoHAN performs better compared to SMB, the numerical values of both $cost_{panic}$ and $cost_{fear}$ are higher compared to all the previous cases. As the robot was modulating only its path to handle a sudden emergence, the velocity remains high, resulting in higher costs. This shows that CoHAN's performance needs to be improved in this context. Therefore, we can say that the proposed metrics can not only differentiate a HAN planner from a standard robot navigation planner but also can be used to benchmark HAN planners and help them improve their performance.

As the angle at which the robot enters the FoV of the human is low, both SMB and CoHAN have comparable values for $cost_{visibility}$. The other surprise metrics, $cost_{shock}$ and $cost_{react}$ clearly indicate that the human might be more surprised by the robot using SMB than the one with CoHAN. From the speed and path profiles in Fig. 8, we can also say that the human is disturbed less by CoHAN compared to SMB.

V. Discussion and Conclusions

With the growing research and scope for social robots in human environments, human-aware robot navigation is gaining more attention. Therefore, good metrics are required to evaluate HAN efficiently and benchmark the performance of the planners. Although, there are existing benchmarks like SocNavBench [27] and SEAN [42], the metrics used for the evaluation are largely related to navigation performance and proxemic violations. We believe that the metrics proposed in this paper from the perspective of a human experiencing the robot's navigation could improve the evaluation and benchmarking of HAN. It also provides us with a new direction from which new metrics can be designed. We think that the idea and the metrics are a good addition to the field.

Although the proposed metrics can benchmark HAN planners and differentiate them, they are not sufficient to completely assess a situation. During our analysis, we used the paths and the velocity profiles along with the metrics to provide a complete picture of what was happening. Hence, the proposed metrics should be used in conjunction with existing metrics to evaluate a situation better. One can also combine one or more of these costs to formulate a better metric. For example, the $cost_{shock}$ can be combined with $cost_{panic}$ or $cost_{fear}$ to check how admissible or undesirable the robot's behaviour is in a situation.

While designing the metrics, we have used the existing user studies on human-robot interaction and human perception. User studies play a great role in advancing the field of HAN and as a part of our immediate future work, we plan to evaluate the validity of these metrics as the proxies for emotions through a detailed user study. Further, we also plan to identify acceptable thresholds for the proposed metrics through this study to benchmark the navigation behaviours better.

Summarising, we have proposed some new metrics that measure the human-perceived value of danger and surprise as the robot navigates around them. We provided a detailed mathematical formulation of these metrics and talked about their relevance in different human-robot navigation scenarios. Finally, these metrics were used to benchmark the performance of a human-aware planner in comparison with a standard navigation planner that can avoid dynamic obstacles. The results demonstrated the capability of the proposed metrics in benchmarking a HAN system.

Acknowledgements

This work has been partially funded by the Agence Nationale de la Recherche through the ANITI ANR-19-PI3A-0004 grant and by the Horizon Europe Framework Programme through the euROBIN Grant 101070596.

References

[1] Y. Gao and C.-M. Huang, "Evaluation of socially-aware robot navigation," *Frontiers in Robotics and AI*, p. 420, 2021.

[2] E. T. Hall, *The Hidden Dimension: Man's Use of Space in Public and Private*. 1966.

[3] M. P. Joosse, R. W. Poppe, M. Lohse, and V. Evers, "Cultural differences in how an engagement-seeking robot should approach a group of people," in *Proceedings of the 5th ACM international conference on Collaboration across boundaries: culture, distance & technology*, pp. 121–130, 2014.

[4] M. Joosse, M. Lohse, N. V. Berkel, A. Sardar, and V. Evers, "Making appearances: How robots should approach people," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 1, pp. 1–24, 2021.

[5] D. Shi, E. G. Collins Jr, B. Goldiez, A. Donate, X. Liu, and D. Dunlap, "Human-aware robot motion planning with velocity constraints," in *2008 International Symposium on Collaborative Technologies and Systems*, pp. 490–497, IEEE, 2008.

[6] T. Kruse, A. Kirsch, H. Khambhaita, and R. Alami, "Evaluating directional cost models in navigation," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pp. 350–357, ACM, Mar. 2014.

[7] L. Takayama and C. Pantofaru, "Influences on proxemic behaviors in human-robot interaction," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5495–5502, IEEE, 2009.

[8] E. Pacchierotti, H. I. Christensen, and P. Jensfelt, "Human-robot embodied interaction in hallway settings: a pilot user study," in *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, pp. 164–171, IEEE, 2005.

[9] E. Senft, S. Satake, and T. Kanda, "Would you mind me if i pass by you?: Socially-appropriate behaviour for an omni-based social robot in narrow environment," in *15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 539–547, IEEE, 2020.

[10] C. Mavrogiannis, A. M. Hutchinson, J. Macdonald, P. Alves-Oliveira, and R. A. Knepper, "Effects of distinct robot navigation strategies on human behavior in a crowded environment," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 421–430, IEEE, 2019.

[11] X.-T. Truong, V. N. Yoong, and T.-D. Ngo, "Dynamic social zone for human safety in human-robot shared workspaces," in *2014 11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 391–396, IEEE, 2014.

[12] A. Vega-Magro, L. Manso, P. Bustos, P. Nunez, and D. G. Macharet, "Socially acceptable robot navigation over groups of people," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 1182–1187, IEEE, Aug. 2017.

[13] C. Chen, Y. Liu, S. Kreiss, and A. Alahi, "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 6015–6022, IEEE, 2019.

[14] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking," in *2009 IEEE 12th international conference on computer vision*, pp. 261–268, IEEE, 2009.

[15] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social lstm: Human trajectory prediction in crowded spaces," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 961–971, 2016.

[16] J. Guzzi, A. Giusti, L. M. Gambardella, G. Theraulaz, and G. A. Di Caro, "Human-friendly robot navigation in dynamic environments," in *2013 IEEE international conference on robotics and automation*, pp. 423–430, IEEE, 2013.

[17] C. I. Mavrogiannis, W. B. Thomason, and R. A. Knepper, "Social momentum: A framework for legible navigation in dynamic multi-agent environments," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 361–369, 2018.

[18] I. Dynnikov and B. Wiest, "On the complexity of braids," *Journal of the European Mathematical Society*, vol. 9, no. 4, pp. 801–840, 2007.

[19] A. D. Dragan, K. C. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 301–308, IEEE, 2013.

[20] B. Okal and K. O. Arras, "Learning socially normative robot navigation behaviors with bayesian inverse reinforcement learning," in *IEEE ICRA*, 2016.

[21] N. Pérez-Higueras, F. Caballero, *et al.*, "Teaching robot navigation behaviors to optimal rrt planners," *International Journal of Social Robotics*, 2018.

[22] X.-T. Truong and T. D. Ngo, "Toward socially aware robot navigation in dynamic and crowded environments: A proactive social motion model," *IEEE Transactions on Automation Science and Engineering*, no. 4, 2017.

[23] G. Ferrer, A. G. Zulueta, F. H. Cotarelo, and A. Sanfeliu, "Robot social-aware navigation framework to accompany people walking side-by-side," *Autonomous robots*, vol. 41, no. 4, pp. 775–793, 2017.

[24] J. Rios-Martinez, A. Spalanzani, and C. Laugier, "From proxemics theory to socially-aware navigation: A survey," *International Journal of Social Robotics*, vol. 7, no. 2, pp. 137–153, 2015.

[25] A. Kendon, "Spacing and orientation in co-present interaction," in *Development of multimodal interfaces: Active listening and synchrony*, pp. 1–15, Springer, 2010.

[26] I. Kostavelis, A. Kargakos, D. Giakoumis, and D. Tzovaras, "Robot's workspace enhancement with dynamic human presence for socially-aware navigation," in *international conference on computer vision systems*, pp. 279–288, Springer, 2017.

[27] A. Biswas, A. Wang, G. Silvera, A. Steinfeld, and H. Admoni, "Socnavbench: A grounded simulation testing framework for evaluating social navigation," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 11, no. 3, pp. 1–24, 2022.

[28] J. T. Butler and A. Agah, "Psychological effects of behavior patterns of a mobile personal robot," *Autonomous Robots*, vol. 10, no. 2, pp. 185–202, 2001.

[29] C. Bartneck, E. Croft, and D. Kulic, "Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots," 2008.

[30] K. A. Barchard, L. Lapping-Carr, R. S. Westfall, S. B. Banisetty, and D. Feil-Seifer, "Perceived social intelligence (psi) scales test manual," 2018.

[31] S. B. Banisetty and T. Williams, "Implicit communication through social distancing: Can social navigation communicate social norms?," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 499–504, 2021.

[32] K. A. Barchard, L. Lapping-Carr, R. S. Westfall, A. Fink-Armold, S. B. Banisetty, and D. Feil-Seifer, "Measuring the perceived social intelligence of robots," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 9, no. 4, pp. 1–29, 2020.

[33] A. Vega, L. J. Manso, R. Cintas, and P. Núñez, "Planning human-robot interaction for social navigation in crowded environments," in *Workshop of Physical Agents*, pp. 195–208, Springer, 2018.

[34] E. Avrunin and R. Simmons, "Socially-appropriate approach paths using human data," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pp. 1037–1042, IEEE, 2014.

[35] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *nature*, vol. 381, no. 6582, pp. 520–522, 1996.

[36] K. Rayner, T. J. Smith, G. L. Malcolm, and J. M. Henderson, "Eye movements and visual encoding during scene perception," *Psychological science*, vol. 20, no. 1, pp. 6–10, 2009.

[37] J. T. Eckner, J. K. Richardson, H. Kim, D. B. Lipps, and J. A. Ashton-Miller, "A novel clinical test of recognition reaction time in healthy adults.," *Psychological assessment*, vol. 24, no. 1, p. 249, 2012.

[38] B. Wolfe, B. Seppelt, B. Mehler, B. Reimer, and R. Rosenholtz, "Rapid holistic perception and evasion of road hazards.," *Journal of experimental psychology: general*, vol. 149, no. 3, p. 490, 2020.

[39] A. Favier, P. Singamaneni, and R. Alami, "An intelligent human simulation (inhus) for developing and experimenting human-aware and interactive robot abilities," 2021.

[40] P. T. Singamaneni, A. Favier, and R. Alami, "Human-aware navigation planner for diverse human-robot interaction contexts," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5817–5824, IEEE, 2021.

[41] P. T. Singamaneni, A. Favier, and R. Alami, "Watch out! there may be a human. addressing invisible humans in social navigation," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 11344–11351, IEEE, 2022.

[42] N. Tsoi, A. Xiang, P. Yu, S. S. Sohn, G. Schwartz, S. Ramesh, M. Hussein, A. W. Gupta, M. Kapadia, and M. Vázquez, "Sean 2.0: Formalizing and generating social situations for robot navigation," *IEEE Robotics and Automation Letters*, 2022.

APPENDIX

The extended mathematical formulation of TTC, $cost_{panic}$ and $cost_{visibility}$ in terms of the variables shown in Fig. 2 and Fig. 4 are given below.

$$
\begin{aligned}
TTC &= \frac{d_{rh_{eff}}}{\|\overrightarrow{V_{rel}}\|} \\
&= \frac{\mathbb{P}.\mathbb{V} - \sqrt{(\mathbb{P}.\mathbb{V})^2 - \left(\left\|\overrightarrow{V_{rel}}\right\|^2 \left(\left\|\overrightarrow{P_{rh}}\right\|^2 - R^2\right)\right)}}{\left\|\overrightarrow{V_{rel}}\right\|^2} \\
&\ni \mathbb{P}.\mathbb{V} > 0 \text{ and } (\mathbb{P}.\mathbb{V})^2 - \left(\left\|\overrightarrow{V_{rel}}\right\|^2 \left(\left\|\overrightarrow{P_{rh}}\right\|^2 - R^2\right)\right) > 0
\end{aligned}
$$

$$
cost_{panic} = \frac{\left\|\overrightarrow{V_{rel}}\right\| \left| \sqrt{\left\|\overrightarrow{V_{rel}}\right\|^2 \left\|\overrightarrow{P_{rh}}\right\|^2 - (\mathbb{P}.\mathbb{V})^2} \right|}{\left\|\overrightarrow{P_{rh}}\right\| \left(\sqrt{\left\|\overrightarrow{V_{rel}}\right\|^2 \left\|\overrightarrow{P_{rh}}\right\|^2 - (\mathbb{P}.\mathbb{V})^2} - \|\overrightarrow{V_{rel}}\| R \right)}
$$

$$
cost_{visibility} = \alpha \left(\frac{cos^{-1}\left(\hat{u}_h \cdot \overrightarrow{P_{hr}}\right)}{\left\|\overrightarrow{P_{hr}}\right\| \left(\left\|\overrightarrow{P_{hr}}\right\| - R\right)} \right)
$$

$$TTC = \frac{\overrightarrow{P_{rh}} \cdot \overrightarrow{V_{rel}} - \sqrt{\left(\overrightarrow{P_{rh}} \cdot \overrightarrow{V_{rel}}\right)^2 - \left(\left\|\overrightarrow{V_{rel}}\right\|^2 \left(\left\|\overrightarrow{P_{rh}}\right\|^2 - R^2\right)\right)}}{\left\|\overrightarrow{V_{rel}}\right\|^2} \tag{7}$$

$$cost_{passby} = \frac{\left\|\overrightarrow{V_{rel}}\right\| \left|\sqrt{\left\|\overrightarrow{V_{rel}}\right\|^2 \left\|\overrightarrow{P_{rh}}\right\|^2 - \left(\overrightarrow{P_{rh}} \cdot \overrightarrow{V_{rel}}\right)^2}\right|}{\left\|\overrightarrow{P_{rh}}\right\| \left(\sqrt{\left\|\overrightarrow{V_{rel}}\right\|^2 \left\|\overrightarrow{P_{rh}}\right\|^2 - \left(\overrightarrow{P_{rh}} \cdot \overrightarrow{V_{rel}}\right)^2} - \|\overrightarrow{V_{rel}}\| R\right)} \tag{8}$$