Understanding socially aware robot navigation

Kiran Jot Singh*, Divneet Singh Kapoor and Balwinder Singh Sohi

Embedded Systems & Robotics Research Group, Chandigarh University, Punjab, India
* Corresponding Author: kiranjot.693@gmail.com

ABSTRACT

Navigation is an essential skill for autonomous robots, and it becomes a cumbersome task in human populated environments. Robots need to perform the tasks without disturbing the humans around them and ensure comfort and safety of humans as well. It is further influenced by various factors like social norms, geometry of environment, and surrounding people. It is essential to comprehend three components, i.e., social conventions (SC), human motion (HM), and context aware mapping (CAM), for establishing effective socially aware robot navigation (SARN). This article discusses these different aspects of these components, which should be taken into consideration while designing an efficient and optimal SARN. Further, it reports recent experiments conducted by different institutes pertaining to these components.

Keywords: Socially aware robot navigation; Social conventions; Human motion; Context aware mapping.

INTRODUCTION

The applications of robotics are growing day by day. Robots are no more limited to industry environments. With advancements in technology, the robots are moving into social environments like homes and workplaces. The advent of robots in social structures will bring wide range of opportunities for social robotics. However, the success of these applications will be highly dependent on how robots adapt their behavior as per the social expectations. One of the crucial factors for how robot will adapt to its surroundings as per the social expectations is navigation. Human aware robot navigation can be considered as intersection between robot motion planning and human robot interaction (HRI). As humans and robots are going to share the same physical space, robots' navigation not only needs to be safe, but needs to be sociable as well.

In any case, if the robot does not perform navigation properly as per social navigation norms, it can cause an unfavorable social response (Hamilton, 2018; Mutlu and Forlizzi, 2008). Here, social navigation norms refer to robots behaving in a civilized way while navigating within humans. For example, altogether for a robot to explore inside a workplace setting, the robot must be completely customized, programmed, and aware as per the environmental factors. But the equation completely changes when humans become part of that environment setting. Without a careful comprehension of its condition or the capacity to recognize its condition, a robot can experience some significant obstructions by which individuals might get injured. That is why the navigation of any mobile robot in a social environment plays an extremely important role and needs to be understood completely while designing a socially aware robot navigation (SARN) framework for any environment.

Humans are often driven by some inner motivation towards a goal. Furthermore, predicting accurate human trajectory in real time is a complex task as humans follow a nonlinear pattern. Prediction is challenged by a number of factors like complexity of human behavior, surrounding people, social rules, and environment. Most of these factors are not directly visible and need to be inferred from context. As human environments are designed for human convenience, robots need to understand the context in the form of Context Aware Mapping (CAM) for effective SARN. Human Motion

(HM) has further many forms like walking through space or full body motion, facial expressions, gestures, etc. So, basically, human motion is influenced by both physical and social constraints in a particular environment (indoor/outdoor) and context. We need to cater to both while performing human motion prediction for SARN.

Here, we propose a framework that caters to reactive and deliberate components for effective SARN. In social environment setting, a robot should be reactive enough to deal with the environment constraints in which it is situated, and it should also be deliberate, so that it can plan the complete motion before execution. Reactive paradigm is more focused towards local planning (a little time ahead of real execution), which is taken care of by the social convention (SC) block that caters to the norms of the society, which provides the support to civil order and solution to the conflicting situations. Different situations like respecting and following the queue, moving left-/ right-hand as per country, asking for pass permission wherever required, etc. are the examples of the situations that are catered by social conventions, which are used to plan little time ahead of real execution. The deliberate paradigm plans the complete motion, which is taken care of by human motion (HM) block that caters to the velocity and geometry of the robot path, which is primarily focused on the energy optimization. As proxemics deals with distance, motion is more vital than distance, as it caters to the naturalness in robot motion, which comes via predicting human motion in advance, by adapting to velocity and directions of humans. Further, context aware mapping (CAM) represents the movement of humans and robots in the form of metric maps. This data (in the form of metric maps) helps calculate the cost function based upon the state of objects, humans, and current activity in the environment, which is used by both reactive and deliberate paradigms to improve the system tolerance for unseen environments. It is also to be noted that tracking humans and obstacles is a cumbersome job for a robot in motion. So, it is always beneficial to predict the human trajectories in order to plan the robot navigation in advance. Also, the robot needs to be equipped for the overcrowded environments where it has to navigate through the congested spaces. So, the integration of SC, HM prediction, and CAM can lead to social awareness (Chik et al., 2016), which in turn boosts the efficiency of SARN. Figure 1 illustrates the proposed framework that incorporates the aforementioned components for designing SARN strategies.
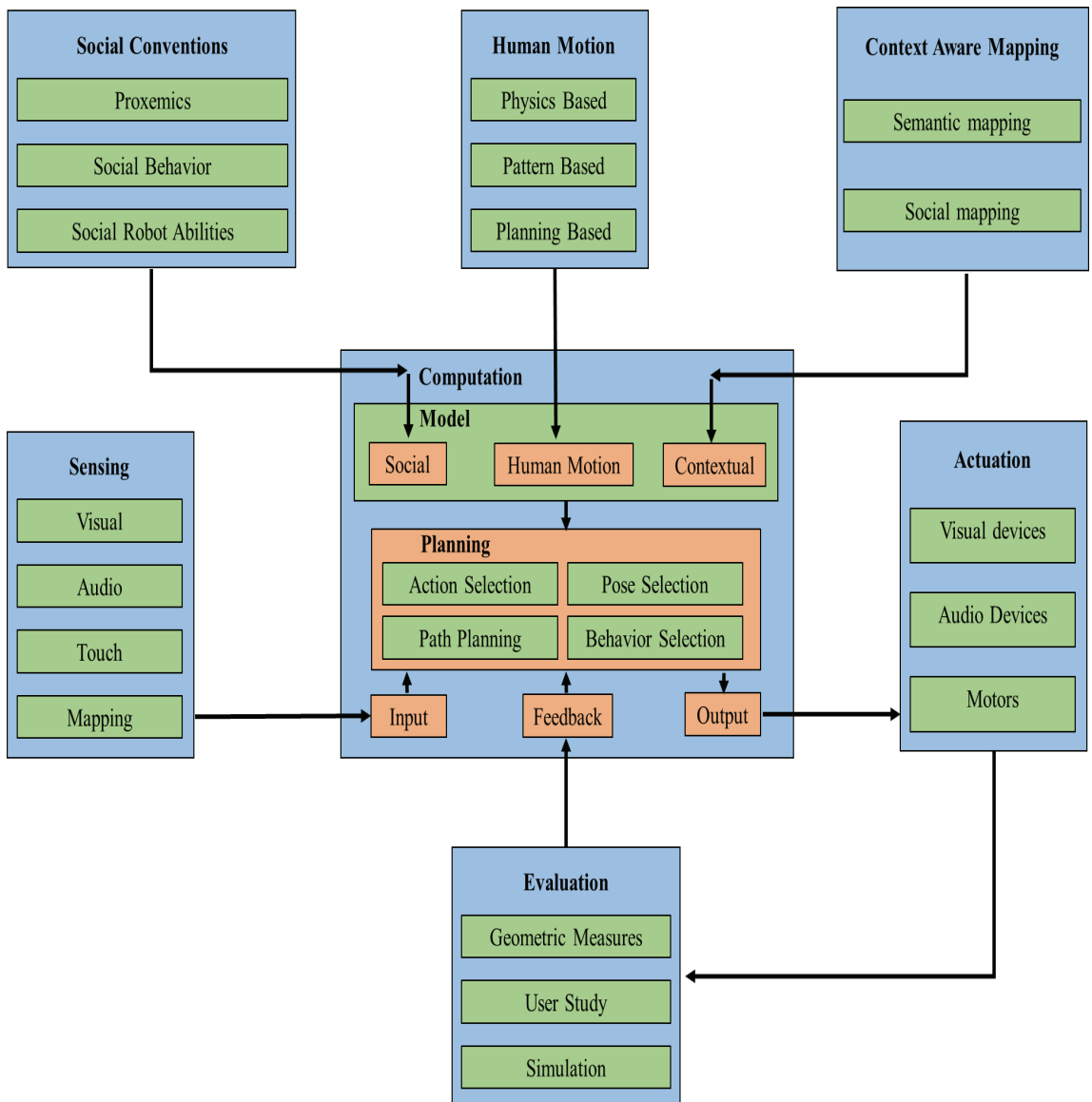
Figure 1. Components of socially aware robot navigation.

Every component is connected to computation block, which ultimately processes every piece of data to actuate output devices to perform SARN. Although every component has its own significance, in order to get acceptance in human society, robots need to respect the social conventions, understand human motion, and follow context aware mapping. So, this article discusses different aspects of SC, HM, and CAM, which should be taken into consideration while designing SARN, and reports different experiments conducted by different institutes pertaining to these components. Further, Table 1 concisely depicts different articles that have discussed various factors of SC, HM, and CAM.

Table 1. Concise literature review.

Articles	Social Convention			Huma Motion			Context aware mapping	
	Proxemics	Social behavior	Social robot abilities	Physics based	Pattern based	Planning based	Semantic mapping	Social mapping
Ballan et al., 2016					✓			
Bartoli et al., 2018					✓			✓
Bera et al., 2016				✓				
Charalampous et al., 2017							✓	✓
Che et al., 2020								✓
Chen et al., 2017	✓	✓			✓			
Hu et al., 2019			✓					
Jarvis, 2019	✓							
Jeong et al., 2018							✓	
Khelloufi et al., 2020							✓	
Kooij et al., 2019				✓			✓	
Macenski et al., 2020								✓
Mehta et al., 2016	✓		✓					
Narayanan et al., 2020	✓	✓						
Radwan et al., 2018					✓			✓
Taylor et al., 2020					✓			✓
Tolani et al., 2020							✓	
Truong and Ngo, 2017	✓	✓						
Vasishta et al., 2017						✓	✓	
Vasquez, 2016						✓		
Vega et al., 2019b	✓	✓						

Therefore, in this article, we focus on the following:
- A framework for designing SARN that incorporates the three main components, i.e., social convention, human motion prediction, and context aware mapping, the integration of which is vital for increasing efficiency of SARN.
- Discussion of recent experiments conducted by various institutes in the domain of social convention, human motion, and context aware mapping, which supported the applicability of the framework.

SOCIAL CONVENTIONS

Behaviors created and acknowledged by society that help humans to comprehend goals of others and encourage the correspondence are called social conventions (Parry and le Roux, 2020). The conventions required for robot navigations should possess safe and understandable behavior. We have further divided them into three areas (Rossi et al., 2020; Schreck et al., 2019; Vega et al., 2019a), i.e., social behavior, proxemics, and social robot abilities as discussed below.

Social Behavior

There are different perspectives from which human behavior can be studied, for example, anthropology, psychology, and sociology. It is very important to understand how surrounding space is managed by humans while navigating, in order to design a navigation strategy for robots in human populated environment.

When the personal space of an individual is invaded by another individual, they do not try to escape or attack other individuals but inform others via social cues like face expressions, gestures, etc. (Jarvis, 2019). So, the term social cue illustrates the verbal and nonverbal messages. These cues can be derived from proximity, body posture, facial expressions, etc. Current sensing technologies are already performing well in detecting social cues (Daily et al., 2017). The real challenge lies in connecting right social cue to right social signal. Furthermore, context plays a crucial role in establishing this connection. Many researchers have tried to connect social signals to social cues as shown in Table 2 (Singh et al., 2019; Vinciarelli et al., 2008).

Table 2. Linkage of social signals to social cues.

Social Signals	Social cues							
	Height	Body Shape	Hand Gestures	Walking	Facial Expressions	Gaze Behavior	Vocalization	Distance
Emotion			✓		✓	✓	✓	
Personality		✓		✓	✓	✓	✓	✓
Status	✓			✓	✓	✓		✓
Dominance	✓	✓		✓	✓	✓	✓	
Persuasion			✓		✓	✓	✓	✓
Regulation			✓		✓	✓	✓	

The navigation behavior of robot must lead to social comfort, where social comfort can be defined as the absence of stress and irritation while interacting with robots. As per the above discussion, social comfort is very subjective matter and cannot be measured directly. There are various studies that make use of distance, visual behavior, and more specifically proxemics to empower the SARN system to cater to social comfort (Jicol et al., 2019; Narayanan et al., 2020).

Proxemics

The study of maintaining spatial distances in various interpersonal and social spaces is called proxemics. A robot in socially aware navigation should be able to manage four kinds of spaces, i.e., space related to individuals, space related to groups, interaction space between humans and objects, and human robot interaction space (Lindner, 2015). The personal space related to individuals can be defined as the minimum space that individuals maintain in order to avoid the discomfort. A classification of space around an individual is shown in Figure 2.
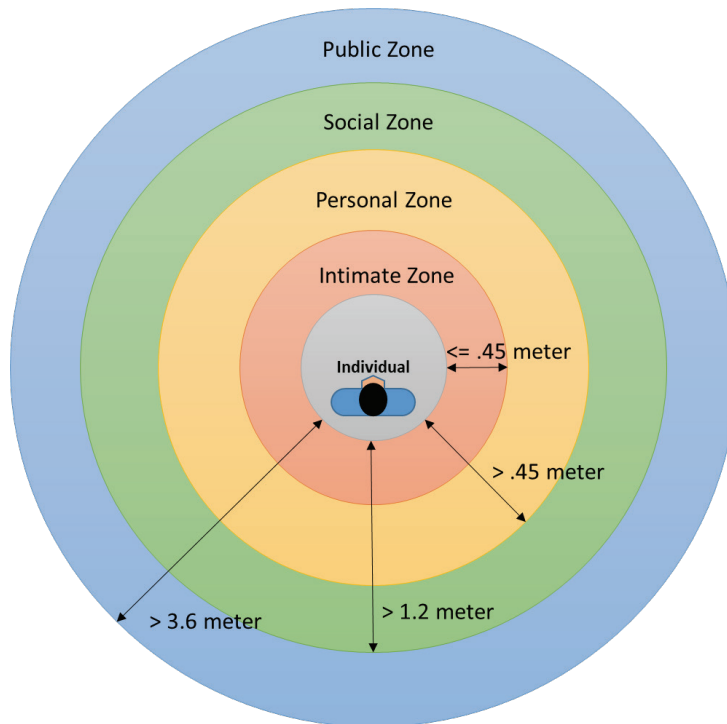

Figure 2. Classification of personal space (Walters et al., 2005).

As far as the space related to groups is concerned, it is the coverage of space when two or more individuals are engaged in conversation. The communication between the individuals can be verbal or nonverbal. Usually, group space configurations can be divided into seven types (Vega et al., 2019b) as shown in Figure 3. One also needs to consider that the space around an individual is situation-dependent and is dynamic in nature. Another aspect of space is information process space, which is nothing but the space that an individual maintains from the obstacles while planning its trajectory.
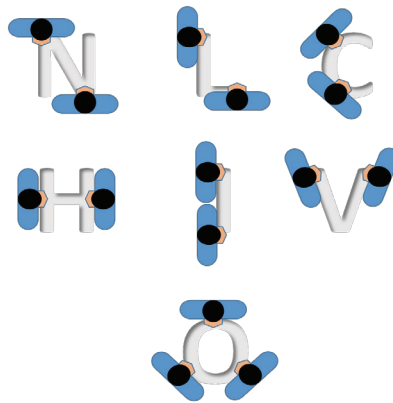
Figure 3. Different types of group space configurations.

It has a strong relation with eye contact, which is further affected by culture and gender as well. Proxemics goes hand in hand with other factors like robot appearance, speed, and approach direction. Many experiments (Narayanan et al., 2020; Rios-Martinez et al., 2015; Singh et al., 2019; Vega et al., 2019b) have been conducted for understanding human aware robot navigation. Although it depends upon situations, in many experiments, it has been demonstrated that slower speed robots are preferred. While closer distances are acceptable, the robot that looks like a humanoid and the robot rotating its head before initiating motion are considered more effective.

Social Robot Abilities

In order to effectively deploy SARN, robots must establish natural interaction similar to humans. Robots must consider the social requirement while navigating like human comfort, need, and preferences. Robots must be aware of the surrounding and treat humans differently from objects while navigating. Robots navigations should be predictable and easily understood by humans; then, it will only be considered as SARN in real terms. Social robot abilities can be further classified into four types, i.e., avoiding collision, passing humans, following humans, and moving along with humans (Hu et al., 2019) as shown in Figure 4.
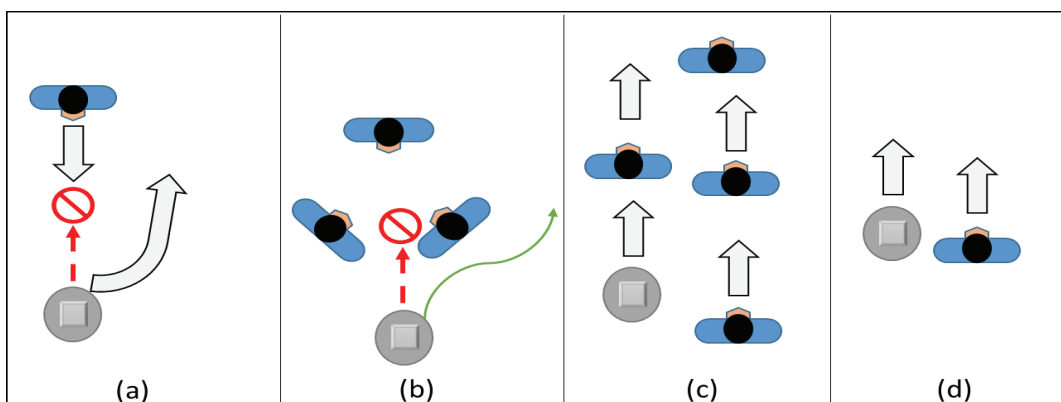

Figure 4. (a) Avoiding collision. (b) Passing humans. (c) Following humans. (d) Moving along with humans.

Collision avoidance is established by social force model, which constantly monitors human trajectory and personal space when humans approach the robot from the opposite direction. In passing humans, robots calculate the cost for every trajectory it can select by considering parameters like distance and choosing between left or right pass when humans are stationary.

Table 3. SARN experiments by different institutes.

Institutes	Study design	Experiment Type	Environment	Robot abilities
• Massachusetts Institute of Technology • IBM Research Center (Chen et al., 2017)	*Subjects* • Many pedestrians *Measure* • Time to goal • Minimum separation distance	• Simulation • Real World	Crowed pedestrian rich environment moving everywhere	• Avoiding collision • Passing humans • Following humans • Moving along with humans
• University of Brunei (Truong and Ngo, 2017)	*Subjects* • Three participants *Measure* • Distance • Human velocity • Robot Velocity	• Simulation • Real World	Laboratory room with stationary individuals in group space configurations	• Avoiding collision • Passing humans
• Universidad de Extremadura • Aston University (Vega et al., 2019b)	*Subjects* • Two participants *Measure* • Euclidean distance • Crowd density	• Real World	Laboratory room with stationary individuals in group space configurations	• Avoiding collision • Passing humans
• University of Michigan (Mehta et al., 2016)	*Subjects* • Nine participants *Measure* • Relative distance • Robot velocity • Time to goal	• Simulation • Real World	Corridor setting with moving individuals and stationary individuals in group space configurations	• Avoiding collision • Passing humans • Following humans

While following humans, navigation is carried out in crowded scenarios by robots considering a target location and selection of a human leader to be followed. Moving along with humans is carried out in environments like museums, airports, etc., while taking optimum robot speed and braking into consideration.

Overall, the detailed knowledge of human representation and space affordances can help plan an effective SARN. Few real world experiments (Chen et al., 2017; Mehta et al., 2016; Truong and Ngo, 2017; Vega et al., 2019b) were conducted by different institutes in the domain of SARN while taking social conventions into account, as illustrated in Table 3. They report the study design depicting the number of participants and quantities measured during the experiment. Further, they report the type of experiment whether it is simulation-based or a real-world demonstration, and environment depicting the type of surroundings and group space configurations as discussed in proxemics section. Lastly, they also report the abilities of the robot deployed in the respective experiments, as discussed in the preceding section. Truong et al. (2017) and Vega et al. (2019b) proposed socially aware robot navigation systems by taking proxemics and social behavior into consideration. They carried out their experiments in a laboratory setup, where both of the systems performed well on social acceptability but lacked energy optimization. Mehta et al. (2016) proposed an autonomous navigation model by taking social robot abilities like avoiding collision, passing humans, following humans, and moving along with humans into consideration while respecting personal space. Their model was based on multipolicy decision making, but no discussion was made on energy optimization. Chen et al. (2017) proposed an autonomous navigation model using deep reinforcement learning while taking care of social conventions in a time efficient manner. So, they combined social conventions and pattern based modeling, which in turn led to better socially aware motion planning.

MODELLING HUMAN MOTION

Modelling human motion is a complex procedure that aims towards generation of future frames based upon the observed human motion state. It is one of the crucial and vital parts of robot intelligence, which enables the robot to efficiently plan its path (Ferrer, 2020). Different human motion modelling techniques use various approaches to represent, learn, and solve the tasks. It can be further subdivided in three types, i.e., Physics, Pattern, and Planning based approaches (Ferrer, 2020) as discussed below.

Physics Based Approaches

Physics based approach makes use of Newton laws of motion to predict human trajectory. A variety of physics based approaches perform decently in predicting how humans move and interact within environments (Møgelmose et al., 2015). Physics based approaches are further divided into Single- and Multimodel approaches involving different modes of dynamics.
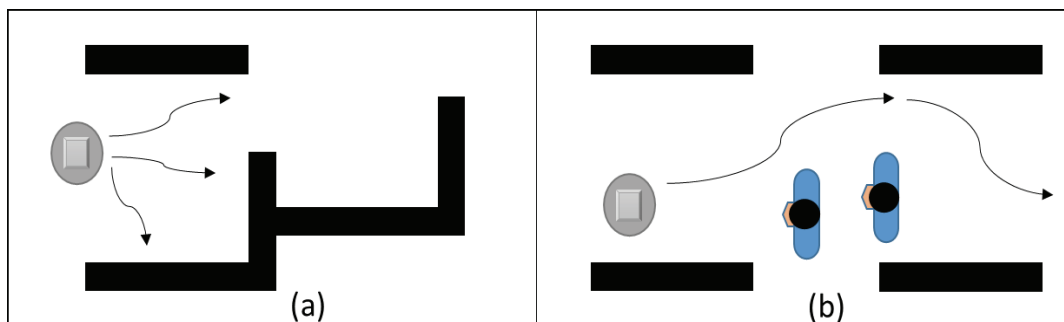

Figure 5. (a) Map based model. (b) Dynamic environment based model.

Single-model approach refers to representing human motion state in the form of position and velocity. It is accomplished by representing the motion state in the form of kinematic model excluding the forces that govern the

motion. Considering external forces into account and in order to cater to context of the environment, the model will be changed into a dynamic model, i.e., map based model or dynamic environment based model, as shown in Figure 5.

Map based model is an extension to single model by considering the map information into the system, where a path planning robot extracts the location of walls, turns, diversions, and other static environment cues from the map in order to improve its navigation (Batkovic et al., 2018). Furthermore, inclusion of dynamic agents, i.e., humans into the system along with static environment cues, leads to dynamic environment based model. In this model, the robot takes into consideration the attractive forces of the desired location where it has to reach repulsive forces of obstacles and dynamic agents on the way, in order to achieve effective SARN. Including different motion modes, along with map and dynamic environment based models, leads to multimodel approaches (Kooij et al., 2019). Motion mode may refer to turns, linear movement, maneuvers, human velocity, human head orientation, etc. in order to describe the complex motion behavior of dynamic agents.

Pattern Based Approaches

Pattern based approaches exploit the power of data by deploying Sense-Learn-Predict paradigm, which utilizes techniques like neural networks, Markov models, etc. Broadly, these approaches can be classified into two types, that is, sequential and nonsequential models.

Sequential models are built on the belief that the current position and velocity can be explained, based on the data from the previous-state's statistical observations. One of the methods to implement sequential model is local transition patterns, i.e., plotting motion pattern probabilities on a grid map (Vasishta et al., 2018). Furthermore, various other methods such as instantaneous topological map, Voronoi graphs, etc. are combined with local transition patterns to predict human motion (Ballan et al., 2016). Recent sequential models make time series prediction of human motion by utilizing neural networks. Another popular approach is long short term memory (LSTM) to predict human activity patterns via long-term observations by observing the motion patterns for several weeks and including contextual information as well (Sun et al., 2018). On the other hand, nonsequential approaches intend to learn human motion pattern by clustering the trajectories, observed over a long term (Bera et al., 2017). It is commonly realized by employing unsupervised clustering techniques in order to determine the future trajectories of humans in a social setting. In order to have a better understanding, a pictorial representation of sequential and nonsequential approaches is shown in Figure 6.
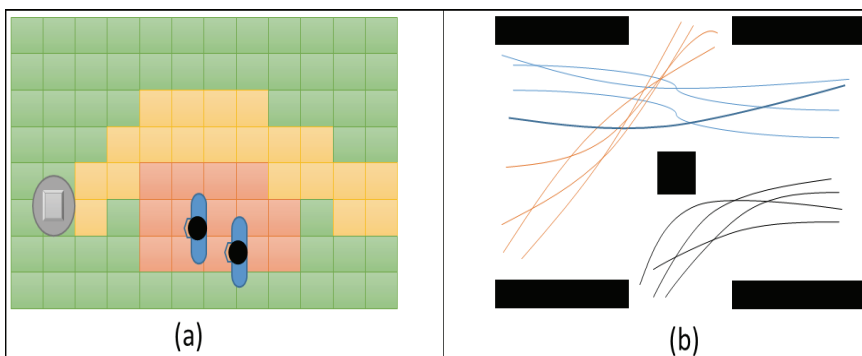

Figure 6. (a) Sequential model. (b) Nonsequential model.

Planning Based Approaches

Planning based approaches build on the paradigm of Sense-Reason-Act, which is reasoning the current state of human motion and acting accordingly, to choose the best possible navigation path for robot. These approaches can be classified into two types, that is, forward planning based and inverse planning based approaches as discussed below.
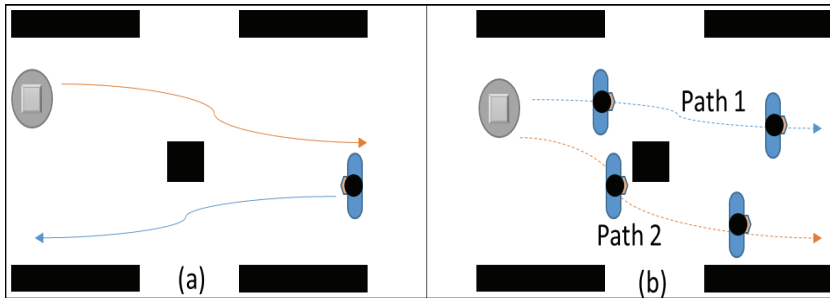

Figure 7. (a) Forward planning. (b) Inverse planning.

Forward planning based approaches use a predefined function for planning motion and path. A probabilistic dynamic model utilizes distance between robot and destination as a metric to predict the next state of humans (Vasishta et al., 2017). In this process, custom cost functions are developed for various environmental entities such as road and crosswalk, whereas, in inverse planning based approaches, the cost function is calculated by observing the humans navigating in the current setting by utilizing imitation techniques (Osa et al., 2018). These approaches select the robot trajectory by transforming it into optimization problem, while taking environment semantics, cost function, and multihuman interaction into consideration. A pictorial representation of forward and inverse planning based approaches is shown in Figure 7.

In the last decade, the field of human motion prediction has seen tremendous improvements in terms of algorithm performance and quantum of application scenarios catered. Different approaches as discussed above are fruitful in different scenarios. Physics based approached are very helpful in static environments or where modelling of human motion is possible via mathematical functions. Pattern based approaches work well for large environments, e.g., public places, where dynamics are unknown, and multiple pedestrians/humans are involved, whereas when robots and human goals can be explicitly defined in an environment, then planning based approaches are extremely suitable. Few real world experiments (Bartoli et al., 2018; Bera et al., 2016; Radwan et al., 2018; Vasquez, 2016) conducted by different institutes in the domain of human motion prediction are illustrated in Table 4.

Table 4. Human motion prediction experiments by different institutes.

Institutes	Study design	Environment	Approach used
• Centre Inria Grenoble, Europe (Vasquez, 2016)	*Subjects* • Many pedestrians *Measure* • Modified Hausdorff Distance	16 urban scenarios *Experiment Type* • Real World	Forward Planning (Planning based approach)

	Subjects • Many pedestrians *Measure* • Final Displacement Error	Traffic light scenarios *Experiment Type* Real World	LSTM – Sequential Model (Pattern based approach)
• University of Freiburg (Radwan et al., 2018)			
• University of North Carolina (Bera et al., 2016)	*Subjects* • Many pedestrians *Measure* • Prediction Accuracy	Different public places *Experiment Type* Simulation	Dynamic environment based model (Physics based approach)
• University of Florence • University of Padova (Bartoli et al., 2018)	*Subjects* • Visitors at museum *Measure* • Mean Euclidean Distance	Museum *Experiment Type* Real World	LSTM – Sequential Model (Pattern based approach)

It reports the study design depicting the number of participants and quantities measured during the experiment. Further, it reports the type of environment depicting the type of surroundings and experiment, whether it is simulation-based or a real-world demonstration. Lastly, it also reports the type of approach used for human motion prediction as discussed in the above sections. Vasquez (2016) proposed a novel planning based algorithm for the prediction of human motion, which was based on forward planning. This algorithm used fast marching method in parallel fashion and was able to reduce the computation time by a factor of 30 as compared to Markov decision process based models. Bera et al. (2016) proposed path prediction based on global and local movement patterns on different videos captured by overhead cameras, but their experiment lacked testing their model with the robots. Bartoli et al. (2018) and Radwan et al. (2018) models performed well by combining human motion approaches with context aware mapping by utilizing LSTM sequential model to achieve energy optimization as well as offered better system tolerance to new unseen scenarios.

CONTEXT AWARE MAPPING

Various service industries built around a human centered social setting like hotels, offices, hospitals, and even homes are going to utilize more and more robots for package delivery, elderly care, cleaning, etc. A big step in effective SARN is context aware mapping technology to bridge the gap between robot's geometrical measures interpretation and understanding of spatial relations in human environments. The context aware mapping is further divided into two parts, that is, semantic mapping and social mapping as described below.

Semantic Mapping

A robot with semantic mapping capability makes the use of high-level modalities such as optical character recognition (OCR), object recognition, and distance measurements in order to formulate its geometrical interpretation. It appends the features understandable by humans on top of metric maps in order to enhance its geometrical interpretation and further communicating the same in human compatible manner.

One of the methods to perform indoor single scene semantic mapping is to use a color camera integrated with depth data, e.g., red-green-blue-depth (RGB-D). Further, this data is transformed into a different color space like La*B* and made use of conditional random field (CRF) model and random forest model to perform segmentation of the scene (Harada et al., 2014). Another research performs context aware semantic mapping by utilizing metric coordinates to represent high-level feature, e.g., door signs, objects, etc. Also, a multimodel interaction was used to represent semantic landmarks to demonstrate context aware navigation by making use of annotated landmarks and human tracking (Cosgun and Christensen, 2018).

On the other side, outdoor scene semantic mapping makes use of stereo cameras for RGB and depth data to perform feature extraction and segmentation at multiple scales. Usually, multi model approach is used for the fusion of different field of views (FoVs) in conjunction with CRF to accurately represent the outdoor scene, which is also understandable by humans as well, as shown in Figure 8 (Jeong et al., 2018).

Figure 8. Semantic segmentation with class labels (Jeong et al., 2018).

Social Mapping

The most anticipated step after establishing semantic mapping is to take human factors like sociability, naturalness, safety, and comfort into consideration, which leads to the field of social mapping. For example, a robot is navigating in a human centered environment, and it chooses not to pass through a group of people chatting; then, it is following social mapping.

Safety and visibility are the two key criteria to form a social map for indoor environments. Safety caters to the distance between human and robot. Researchers usually make use of cost function approach in conjunction with theory of proxemics to evaluate the distance between human and robot and take desired action (Charalampous et al., 2017). The visibility helps form the social map, building on the criterion that robot must be in line of sight of human

in order to eliminate the element of surprise (Khelloufi et al., 2020; Macenski et al., 2020), as shown in Figure 9. Furthermore, one more approach is to develop a perceptual model that also considers cost functions for human gestures, speech levels, and relative poses along with the distance parameter (Che et al., 2020). This strategy reduced user effort and built the human trust for robot.

Figure 9. Social mapping experiment in an indoor environment (Macenski et al., 2020).

Many social mapping algorithms utilize social force model (SFM) in outdoor environments to predict the human motion. SFM builds on the concept that human trajectories can be explained by social field and forces. The individual trajectory of a human is seen in conjunction with neighboring obstacles acting as repulsive forces and nearby humans. So, based on the forces exerted by nearby objects and humans, social force analysis shapes social maps (Taylor et al., 2020).

Table 5. Context aware mapping experiments by different institutes.

Institutes	Study design	Environment	Approach used
• Samsung Research • Rey Juan Carlos University • Contextual Robotics Institute • Rey Juan Carlos University (Macenski et al., 2020)	*Subjects* • Many pedestrians *Measure* • Timestamp • Distance • Velocity • Recoveries executed	Human-filled environment in a University setting (Indoor setting) *Experiment Type* • Real World	Social mapping (Behavior tree)

• University of California • Google Brain Research (Tolani et al., 2020)	*Subjects* • Single human *Measure* • Time • Acceleration • Jerk	Human navigating in indoor setting *Experiment Type* • Real World • Simulation	Semantic mapping (Convolutional neural network based perception model)
• University of California (Taylor et al., 2020)	*Subjects* • Many pedestrians *Measure* • Accuracy • Precision • Recall • Depth	Human navigating in outdoor setting at different public places *Experiment Type* • Real World	Social mapping (Hierarchical tree)
• Center for Development of Advanced Technologies • University Of Science And Technology • Universite de Montpellier (Khelloufi et al., 2020)	*Subjects* • Single human *Measure* • Distance • Orientation • Velocity	Different indoor settings *Experiment Type* • Real World • Simulation	Semantic mapping (Tentacles-based approach)

Furthermore, SFM is also deployed in conjunction with multihypothesis approach to predict human trajectory (Senft et al., 2020). These hypotheses are the combinations revolving around human interactions, which further utilize joint probability functions to evaluate humans in the scene.

Overall, a lot of research is going on to bridge the gaps in human robot interaction by enhancing the capabilities of robots in order to make robots obey the social norms. The social mapping emphasizes on the aspect of human robot coexistence and robot navigation in human centered environments. Motion pattern of humans is considered in outdoor situations, whereas, in indoor setting, different parameters like positioning of humans and limited space for robot navigation come into play. Few recent real world/simulation experiments (Khelloufi et al., 2020; Macenski et al., 2020; Taylor et al., 2020; Tolani et al., 2020) conducted by different institutes in the domain of context aware mapping are shown in Table 5.

It reports the study design depicting the subjects and quantities measured during the experiment. Further, it reports the type of environment depicting the type of surroundings and experiment, whether it is simulation-based or a real-world demonstration. Lastly, it also reports the type of approach used for mapping as discussed in above sections. Tolani et al. (2020) and Khelloufi et al. (2020) used a semantic mapping based approach to detect the surrounding obstacles during navigation, but both models lacked adapting to social norms. Macendki et al. (2020) proposed a secure message passing framework specifically for safety critical applications. The model lacked dynamic obstacle tracing and planning, which will be taken as future work by the authors. Taylor et al. (2020) proposed a robot centric group estimation model by utilizing social mapping in conjunction with motion estimation. The proposed model performed better in terms of tracking of humans and energy optimization.

In a nutshell, we have seen that social convention block makes the robot follow social norms, which in turn elevates its acceptability among humans. Human motion prediction supports the robot in path planning in advance, which leads to time saving as well as energy optimization. Lastly, context aware mapping improves the accuracy and tolerance of the robot in unseen environments. Hence, integration of these three components will bring all the benefits together to enhance the acceptability of the robot for various robotics applications as well.

CONCLUSION

Scientific community is working really hard in robotics, and soon, robots will be the part of our everyday life. Making robots understand social conventions, human motion, and context aware mapping will play a key role to turn this into reality. This article discusses different aspects of social conventions, human motion, and context aware mapping that are essential to establish effective SARN. Furthermore, it gives a brief comparison of experiments, conducted by various institutes, in terms of the study design, environment, and approaches used. Artificial intelligence and machine learning algorithms are quite popular and used extensively for estimating human motion patterns, which are dependent on context. Understanding of context in respect to the semantics of environment for efficient trajectory prediction is still an unexplored area. Real human trajectories can be motivated by their personal goals, and cooperative behavior of humans is taken for granted, which is not always applicable. So, it is required to design the SARN models by taking intention or goal directed human behavior into consideration, in order to cater to long-term predictions. In such scenarios, game theory approach can be explored for efficient modelling of human behavior. Roboticists need to work along with people from other domains of design, artificial intelligence, and psychology in order to better define social conventions, human motion, and context aware mapping in order to realize effective SARN. Finally, these components have to be mapped in suitable software process to achieve the desired SARN. Future scope includes the incorporation of this framework on the robot MAI (Singh et al., 2018) and to further test it in the different social environment settings.

REFERENCES

Ballan, L., Castaldo, F., Alahi, A., Palmieri, F., Savarese, S., 2016. Knowledge transfer for scene-specific motion prediction. In: European Conference on Computer Vision. pp. 697–713.

Bartoli, F., Lisanti, G., Ballan, L., Del Bimbo, A., 2018. Context-aware trajectory prediction. In: 2018 24th International Conference on Pattern Recognition (ICPR). pp. 1941–1946.

Batkovic, I., Zanon, M., Lubbe, N., Falcone, P., 2018. A computationally efficient model for pedestrian motion prediction. In: 2018 European Control Conference (ECC). pp. 374–379.

Bera, A., Kim, S., Randhavane, T., Pratapa, S., Manocha, D., 2016. GLMP-realtime pedestrian path prediction using global and local movement patterns. In: 2016 IEEE International Conference on Robotics and Automation (ICRA). pp. 5528–5535.

Bera, A., Randhavane, T., Manocha, D., 2017. Aggressive, Tense or Shy? Identifying Personality Traits from Crowd Videos. In: IJCAI. pp. 112–118.

Charalampous, K., Kostavelis, I., Gasteratos, A., 2017. Recent trends in social aware robot navigation: A survey. Rob. Auton. Syst. 93, 85–104.

Che, Y., Okamura, A.M., Sadigh, D., 2020. Efficient and Trustworthy Social Navigation via Explicit and Implicit Robot--Human Communication. IEEE Trans. Robot. 36, 692–707.

Chen, Y.F., Everett, M., Liu, M., How, J.P., 2017. Socially aware motion planning with deep reinforcement learning. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 1343–1350.

Chik, S.F., Yeong, C.F., Su, E.L.M., Lim, T.Y., Subramaniam, Y., Chin, P.J.H., 2016. A review of social-aware navigation frameworks for service robot in dynamic human environments. J. Telecommun. Electron. Comput. Eng. 8, 41–50.

Cosgun, A., Christensen, H.I., 2018. Context-aware robot navigation using interactively built semantic maps. Paladyn, J. Behav. Robot. 9, 254–276.

Daily, S.B., James, M.T., Cherry, D., Porter III, J.J., Darnell, S.S., Isaac, J., Roy, T., 2017. Affective computing: historical foundations, current applications, and future trends. In: Emotions and Affect in Human Factors and Human-Computer Interaction. Elsevier, pp. 213–231.

Ferrer, 2020. Human Motion Prediction for Social Robot Navigation. In: IEEE International Conference on Robotics and Automation. IEEE.

Hamilton, I.A., 2018. People kicking these food delivery robots is an early insight into how cruel humans could be to robots.

Harada, K., Tsuji, T., Uto, S., Yamanobe, N., Nagata, K., Kitagaki, K., 2014. Stability of soft-finger grasp under gravity. In: 2014 IEEE International Conference on Robotics and Automation (ICRA). pp. 883–888.

Hu, Z., Pan, J., Fan, T., Yang, R., Manocha, D., 2019. Safe navigation with human instructions in complex scenes. IEEE Robot. Autom. Lett. 4, 753–760.

Jarvis, C.M., 2019. Navigating the public-place/private-market tension: toponmyic and material reading of Broncos Stadium. Sport Soc. 1–17.

Jeong, J., Yoon, T.S., Park, J.B., 2018. Towards a meaningful 3D map using a 3D lidar and a camera. Sensors 18, 2571.

Jicol, C., Taulo, G., Goldie, C., Esenkaya, T., Hynes, R., Paradise, C., Proulx, M., de Sousa, A.A., 2019. Exploring the Effects of Environmental Cues on Perceived Personal Space in the Virtual Workplace.

Khelloufi, A., Achour, N., Passama, R., Cherubini, A., 2020. Sensor-based Navigation of Omnidirectional Wheeled Robots Dealing with both Collisions and Occlusions. Robotica 38, 617–638.

Kooij, J.F.P., Flohr, F., Pool, E.A.I., Gavrila, D.M., 2019. Context-based path prediction for targets with switching dynamics. Int. J. Comput. Vis. 127, 239–262.

Lindner, F., 2015. A conceptual model of personal space for human-aware robot activity placement. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 5770–5775.

Macenski, S., Mart\'\in, F., White, R., Clavero, J.G., 2020. The Marathon 2: A Navigation System. arXiv Prepr. arXiv2003.00368.

Mehta, D., Ferrer, G., Olson, E., 2016. Autonomous navigation in dynamic social environments using multi-policy decision making. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 1190–1197.

Møgelmose, A., Trivedi, M.M., Moeslund, T.B., 2015. Trajectory analysis and prediction for improved pedestrian safety: Integrated framework and evaluations. In: 2015 IEEE Intelligent Vehicles Symposium (IV). pp. 330–335.

Mutlu, B., Forlizzi, J., 2008. Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction. In: 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 287–294.

Narayanan, V., Manoghar, B.M., Dorbala, V.S., Manocha, D., Bera, A., 2020. ProxEmo: Gait-based Emotion Learning and Multi-view Proxemic Fusion for Socially-Aware Robot Navigation. arXiv Prepr. arXiv2003.01062.

Osa, T., Pajarinen, J., Neumann, G., Bagnell, J.A., Abbeel, P., Peters, J., 2018. An algorithmic perspective on imitation learning. arXiv Prepr. arXiv1811.06711.

Parry, D.A., le Roux, D.B., 2020. Introducing the Media Use Behaviour Conceptual Framework. In: Conference on E-Business, e-Services and e-Society. pp. 175–186.

Radwan, N., Burgard, W., Valada, A., 2018. Multimodal interaction-aware motion prediction for autonomous street crossing. arXiv Prepr. arXiv1808.06887.

Rios-Martinez, J., Spalanzani, A., Laugier, C., 2015. From proxemics theory to socially-aware navigation: A survey. Int. J. Soc. Robot. 7, 137–153.

Rossi, S., Rossi, A., Dautenhahn, K., 2020. The Secret Life of Robots: Perspectives and Challenges for Robot's Behaviours During Non-interactive Tasks. Int. J. Soc. Robot. 1–14.

Schreck, J.L., Newton, O.B., Song, J., Fiore, S.M., 2019. Reading the Mind in Robots: How Theory of Mind Ability Alters Mental State Attributions During Human-Robot Interactions. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. pp. 1550–1554.

Senft, E., Satake, S., Kanda, T., 2020. Would You Mind Me if I Pass by You? Socially-Appropriate Behaviour for an Omni-based Social Robot in Narrow Environment. In: Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. pp. 539–547.

Singh, K.J., Kapoor, D.S., Sohi, B.S., 2018. The MAI: A Robot for/by Everyone. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. pp. 367–368.

Singh, P., Pisipati, N., Krishna, P.R., Prasad, M.V.N.K., 2019. Social Signal Processing for Evaluating Conversations Using Emotion Analysis and Sentiment Detection. In: 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP). pp. 1–5.

Sun, L., Yan, Z., Mellado, S.M., Hanheide, M., Duckett, T., 2018. 3DOF pedestrian trajectory prediction learned from long-term autonomous mobile robot deployment data. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). pp. 1–7.

Taylor, A., Chan, D.M., Riek, L.D., 2020. Robot-centric perception of human groups. ACM Trans. Human-Robot Interact. 9, 1–21.

Tolani, V., Bansal, S., Faust, A., Tomlin, C., 2020. Visual Navigation Among Humans with Optimal Control as a Supervisor. arXiv Prepr. arXiv2003.09354.

Truong, X.-T., Ngo, T.D., 2017. Toward socially aware robot navigation in dynamic and crowded environments: A proactive social motion model. IEEE Trans. Autom. Sci. Eng. 14, 1743–1760.

Vasishta, P., Vaufreydaz, D., Spalanzani, A., 2017. Natural vision based method for predicting pedestrian behaviour in urban environments. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). pp. 1–6.

Vasishta, P., Vaufreydaz, D., Spalanzani, A., 2018. Building prior knowledge: A Markov based pedestrian prediction model using urban environmental data. In: 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV). pp. 247–253.

Vasquez, D., 2016. Novel planning-based algorithms for human motion prediction. In: 2016 IEEE International Conference on Robotics and Automation (ICRA). pp. 3317–3322.

Vega, A., Cintas, R., Manso, L.J., Bustos, P., Núñez, P., 2019a. Socially-Accepted Path Planning for Robot Navigation Based on Social Interaction Spaces. In: Iberian Robotics Conference. pp. 644–655.

Vega, A., Manso, L.J., Macharet, D.G., Bustos, P., Núñez, P., 2019b. Socially aware robot navigation system in human-populated and interactive environments based on an adaptive spatial density function and space affordances. Pattern Recognit. Lett. 118, 72–84.

Vinciarelli, A., Pantic, M., Bourlard, H., Pentland, A., 2008. Social signal processing: state-of-the-art and future perspectives of an emerging domain. In: Proceedings of the 16th ACM International Conference on Multimedia. pp. 1061–1070.

Walters, M.L., Dautenhahn, K., Koay, K.L., Kaouri, C., Boekhorst, R. te, Nehaniv, C., Werry, I., Lee, D., 2005. Close encounters: Spatial distances between people and a robot of mechanistic appearance. In: 5th IEEE-RAS International Conference on Humanoid Robots, 2005. pp. 450–455.