Robot's Workspace Enhancement with Dynamic Human Presence for Socially-Aware Navigation

Conference Paper · October 2017

DOI: 10.1007/978-3-319-68345-4_25

4 authors:

Ioannis Kostavelis
The Centre for Research and Technology, Hellas

Andreas Kargakos
The Centre for Research and Technology, Hellas

Dimitrios Giakoumis
The Centre for Research and Technology, Hellas

Dimitrios Tzovaras
The Centre for Research and Technology, Hellas

Robot's Workspace Enhancement with Dynamic Human Presence for Socially-aware Navigation

Ioannis Kostavelis, Andreas Kargakos,
Dimitrios Giakoumis, and Dimitrios Tzovaras

Centre for Research & Technology Hellas, Information Technologies Institute,
6th Km Charilaou-Thermi Road, 57001, Thermi-Thessaloniki, Greece
{gkostave,akargakos,dgiakoum,tzovaras}@iti.gr

Abstract. The incorporation of service robots in human populated environments gives rise to the adaptation of cruise strategies that allow robots to move in a natural, secure and ordinary manner among their cohabitants. Therefore, robots should firstly apprehend their space similarly with the people and, secondly, should adopt human motion anticipation strategies in their planning mechanism. The paper at hand introduces a closed-loop human oriented robot navigation strategy, where on-board a moving robot, multimodal human detection and tracking methods are deployed to predict human motion intention in the shared workspace. The human occupied space is probabilistically constrained following the proxemics theory. The impact of human presence in the commonly shared space is imprinted to the robot's navigation behaviour after undergoing a social filtering step based on the inferred walking pattern. The proposed method has been integrated with a robotic platform and extensively evaluated in terms of socially acceptable behaviour in real-life experiments exhibiting increased navigation capacity in human populated environments.

Keywords: robot navigation, leg and human skeleton tracker, social costmap, human motion intension prediction, robot path planning

1 Introduction

The interest of the film industry in making movies such as the *WALL-E* [2008], *Robot & Frank* [2012] and many others reveals the potential for broad acceptance, as well as comprehension of people for domestic and service robots in near future communities. An in-depth study of these movies discloses the human's expectations about the characteristics that the forthcoming artificial agents should retain so as to be broadly acceptable, while someone can infer that naturalness of robots' motion is the main attribute that renders their presence intimate to the humans. This challenge is broadly recognized by the respective scientific community which has already conducted a laborious research on robot motion planning [11]. Yet, the demand for mobile robots to operate in dynamic changing environments where humans are also involved, emerged the development of more sophisticated motion planning strategies where human factors should be considered [6], integrating contextual representations of the individual's presence [2].

Following this aspect, human-robot co-navigation strategies are foreseen when designing a service robot targeted to operate in human-centric environments [17]. Human aware navigation and planning include a significant variety of implementations and, therefore, we attempt to provide a categorization based on the methodologies utilized for the human space modelling and the navigation methods that consider human presence. The motivation behind this taxonomy it that these well known issues constitute the basic components of the proposed methodology.

Human space modelling: Proxemics theory early introduced by Edward T. Hall [5] comprised the cornerstone of human space apprehension by establishing the theory on how individuals' placement in space affects the quality of their interaction. This theory has been broadly accepted by the roboticists during last decades and embodied in various ways in the human space modelling mechanisms with robots. The work in [12] represents the social zones in terms of isocontours of an implicit function capable of describing complex social interaction. In more recent work, the authors in [15] marked the human space with a single Gaussian kernel parameterized with respect to human calculated velocity differentiating thus, in the modeled space the human presence from the obstacles; yet, this implementation was appropriate only for short term robot motion calculations. A relevant work which capitalizes on the proxemics theory for modelling the human personal space is the one described in [7], where used joint oriented Gaussian functions to model human presence targeting to be incorporated in global path planning level. Contrary to the aforementioned works, the human space modelling follows the notion of the proxemics theory, yet by considering a sequence of Gaussian kernels -instead of a single one- formulated along the estimated human paths, the amplitudes of which degrades considering the human's velocity allowing thus predictive long-term global path planning. Moreover, the proposed method can be extended to multiple human tracking and, hence, to their personal space modelling as well.

Path planning with human presence: Considering dynamic path planning, the work in [13] introduces a path planner that considers the presence of humans in terms of their vision field, their accessibility and their personal choices regarding the human-robot placement. However, this system considers solely static persons but the fast computation time of this module allows online path replanning. Dynamic human presence during path planning is considered in [14], where the location and movement of humans is modeled as potential fields and the most feasible trajectory is calculated using *Rapidly-exploring Random Tree* (RRT). The authors in [8] employ velocity obstacles to infer trajectories ample to avoid humans, while a probabilistic extension of RRT based on predictive Gaussian processes is employed in [4] and proved adequate for path planning in dynamic environments. Similar, yet more contemporary methods are also presented in [16] and [16]. Luber et al. [10] proposed a machine learning strategy that employs the measurements of walking people to solve an unsupervised learning problem. Specifically, the authors define the relative motion prototypes and cluster them hierarchically, by exploiting a distance function that relies on a modified *Dynamic Time Warping* (DTW) module. Afterwards, relative motion prototypes are used in a model selection to extract social context, which is further used for the formulation of the cost map. The authors in [17] introduced walking motion anticipatory features that when integrated with a learning scheme proved adequate navigation capacity to maneuver a robot around humans in dynamic environment. The paper at hand, Fig. 1 anticipates the walking motion of the human by inserting the concept of frequently visited areas in a well defined environment. The estimated human workspace is modeled as costmap ample to operate in real-time applications as it is integrated both in global and local planner level.

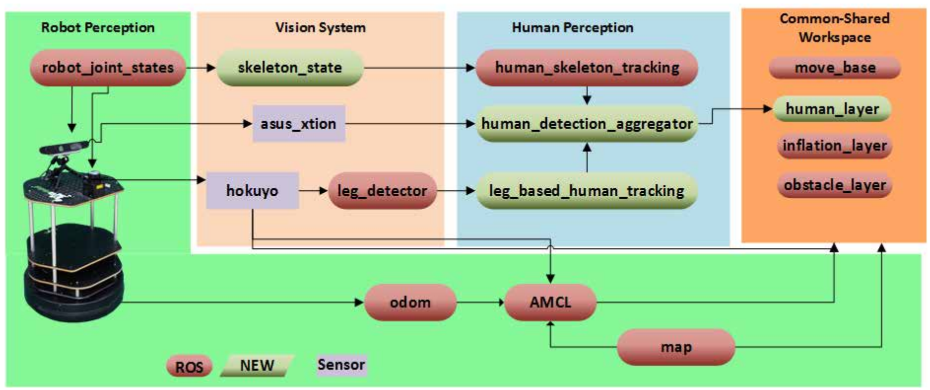


Fig. 1. Block diagram describing the software components involved in the proposed method

2 Human Presence Anticipation

2.1 Human Detection and Tracking

A basic prerequisite for human-aware autonomous robot navigation is that the robot should to detect the people present in its surroundings. For that reason two different human tracking algorithms have been adopted and fused in the present worked, based on LIDAR and RGB-D data obtaining this way constant human awareness during the robot operation, i.e. if the user is out of the field of view (FoV) of the RGB-D sensor it will be tracked with adequate confidence from the laser readings.

Nite skeleton tracker: Constant human detection with stationary robotic platform is addressed with the NiTE skeleton tracker. To keep track of the existing persons, which remain in the FoV of the robot's depth camera, the detected human IDs are shorted with respect to the returned confidence value. This information will be later used for the fusion with leg tracker. The advantage of this modality is that it performs reliable detection with increased confidence, but it cannot be used with moving robot.

Leg detector: Laser based human tracking is employed herein using the methodology introduced in [9], allowing human tracking while robot is moving. To shortly explain this method, the laser scans are clustered according to distance and a feature vector is extracted for each cluster using specific geometrical features. Next, a random forest classifier is trained, where negative examples are obtained by moving the robot in an environment devoid of people and positive examples are obtained by setting up the laser in an open area with significant people traffic. During inference procedure, all the detected pairs of legs that produce a high probability score are considered as potential humans. This method produces many false positive observations and, therefore, herein we extended this method so as to reduce false positives through a blacklisting procedure.
Fusion modality for human verification: NiTE skeleton tracker and leg based human detection are fused to constantly keep track of the human's location while retaining increased awareness also when the human is absent from the camera FoV. Specifically, a rule based fusion algorithm has been developed that takes in to account both observations. An algorithmic explanation is outlined in Alg. 1.

```
Data: LegTracks[ ], SkeletonTracks[ ], StaticMap, RobotTransformation
Result: FinalHumanPoses[ ]
if RobotIsNotMoving(RobotTranformation) then
    for i = 0 to LegTracks.size() do
        for j = 0 to SkeletonTracks.size() do
            if Dist(SkeletonTracks[j], LegTracks[i]) > DistThreshold or
             isCloseToObStacles(LegTracks[i] , StaticMap) then
                Blacklist ← LegTracks[i].ID ;
            else
                LegTracks[i].conf *= (1+ SkeletonTracks[j].conf);
            end
        end
    end
else
    FinalHumanPoses = getHumanPoses(LegTracks , Blacklist);
end
```

Algorithm 1: Nite human detection and leg detector fusion algorithm

2.2 Short-term walking paths prediction

In accordance to the theory of the human spatial experience [1], people tend to frequently revisit regions -indoors or outdoors- that stimulate them either with social experience or are related with specific interaction activities. Such regions can be spots with increased human traffic, e.g. doorways, or places with semantic meaning associated with the activities, e.g. a fridge in a kitchen. Based on this theory we identified such frequently visited human positions in a domestic environment and associated them with the target human locations. Then, we formulated the prediction about the human motion intention based on the rationale that when a person is moving s/he will probably approach one of the pre-defined frequently visited areas. By assuming this, the cases of random wandering are considered less probable since they constitute the less frequent walking patterns of people with normal behaviour [3]. Having said that, the next step comprises the modelling of the human motion intention on the robot's workspace. Firstly, we compute the pose deviation vector $\mathbf{D} = [D_i, D_{i+1}, D_N]$ as derived from the current human position $P_H = (x_h, y_h)$ and the N frequently visited positions $\mathbf{P} = [P_{i=1}, P_{i=2}, ..., P_{i=N}]$ in the environment. Note that $P_i = (x, y)$ where x, y comprise the explicit coordinates in the map, and $D_i = (t, r)$ where t, r corresponds to the computed Euclidean translational and rotational deviation respectively. To determine the most probable frequently visited position that the human will move towards to, we minimize the $argmin(\alpha \mathbf{d_r}_{i=1}^{N} + \beta \mathbf{d_l}_{i=1}^{N})$ criterion where α and β are regularization parameters that control situations where the modelled environment is congested, i.e with many furniture where the user has to follow curved paths to reach a standing position. Moreover, $\mathbf{d_r}$ and $\mathbf{d_l}$ are the vectors that retain the computed Euclidean distance and the rotational deviation of the humans's current position from the pre-defined frequently visited positions, respectively. The minimized values of the criterion are sorted from the most probable to the less probable one. Relying on the assumption that when a human walks towards a target location subconsciously selects the most shortest path, we adopted the D* Lite path planning algorithm in order to model the candidate human paths among the human's current position and the N frequently visited standing positions. The reason of the selection of the D*Lite is mainly due to the fact that this approach repeatedly determines shortest paths between the current position of the human and his/her goal position as the edge costs of a graph change while the human moves towards the goal position, allowing thus fast replanning ample to capture the unexpected changes in human's course. For each point that belongs to human path, an oriented Gaussian kernel is centered therein, the parameters σ_x and σ_y of which, model the personal space of the human in accordance with the proxemics theory; this results in a sequence of partial overlapping kernels. The amplitude A of the first Gaussian kernel in the sequence is reverse proportional to the values of the minimization criterion, indicating that the paths with less probability to be followed by the human have diminished weights. Additionally, the amplitude A of each kernel in the path decreases as the points in the path fend off the human current position and the degradation step is normalized to the total number of the calculated points in the path. This cycle is repeated anytime a human observation is received and due to the simplicity of the calculations and the fast execution rate of the D* algorithm, instantaneous estimations about the human motion are obtained anticipating thus the person's presence in the robot's workspace. To avoid unnecessary computational burden, in cases that human is standing still for a specific time i.e. 2secs, his/her presence is modelled with a sole Gaussian kernel.

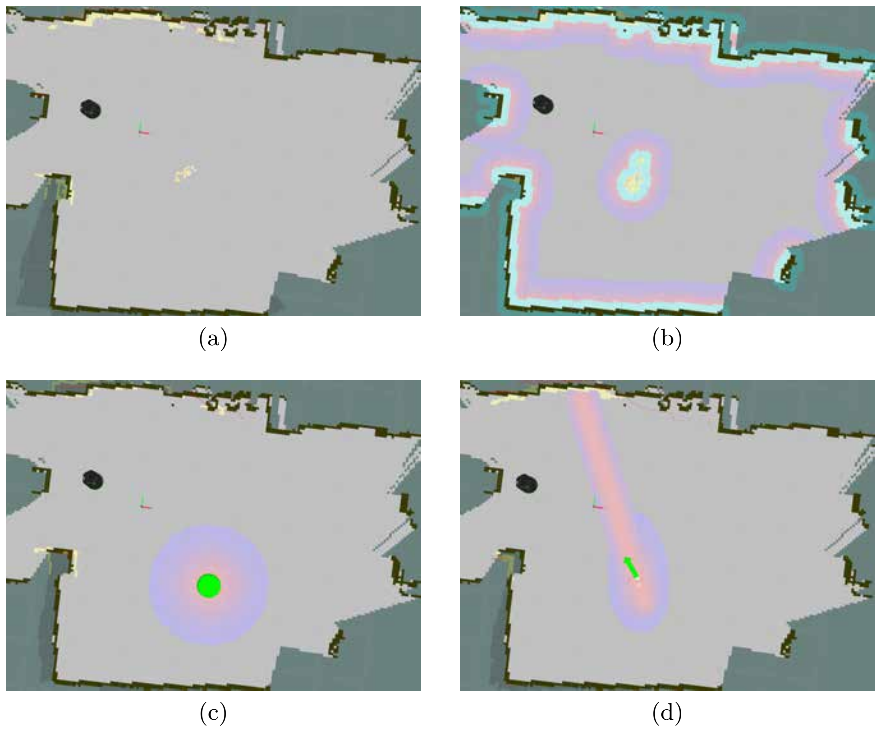


Fig. 2. a) Static map with obstacle layer where the human is detected in the middle of room and treated as obstacle, b) inflation layer, c) human layer with static human detected and d) human layer with moving human and predicted path

2.3 Human anticipation cost-map modelling

Following the aforemention methodology the anticipation of the human's presence within the robot's workspace, in terms of ROS infrastructure has been represented as separate costmap layer. Therefore, the existing costmap layers considered in our method are outlined as follows:

Static map layer: Represents the metric map, separating the obstacles from the free space and defines the width and the height of other layers.

Obstacle layer: Tracks the obstacles as observations obtained by the sensor data and marks them or clears the space by raytracing.

Inflation layer: Inflation is the process where the cost values decay while moving from obstacle cells to free cells. In this layer, the costmap is quantized with specific symbolic zones, i.e lethal (actual obstacle), inscribed (robot footprint in collision), freespace and unknown. All the costs are assigned depending on the distance from a lethal cell and decays to the free space cost.

Human layer: Depending on the observations of the human tracking module and the values of the above layers, extra costs are computed and formulated as final layer to the total costmap. The human layer retains the lower priority comparing to any other layer, since the robot firstly needs to avoid collisions, then to produce paths, and if possible to be also aware of the social cost of these paths. Initially, the human layer is inferred given the position, orientation and velocity of the human in space. If no human is detected, then no layer is generated. If the velocity of the detected human is low, then static human is assumed, and, thus static Gaussian kernel of costs is computed and superimposed to other costmap layers, centered in the detection coordinates. In situations where human velocity is above a threshold (i.e 2m/sec experimentaly defined herein compensating the glittering effect of laser measurements)then short-term walking path prediction is enabled as described in Sec. 2.

3 Methodology Evaluation

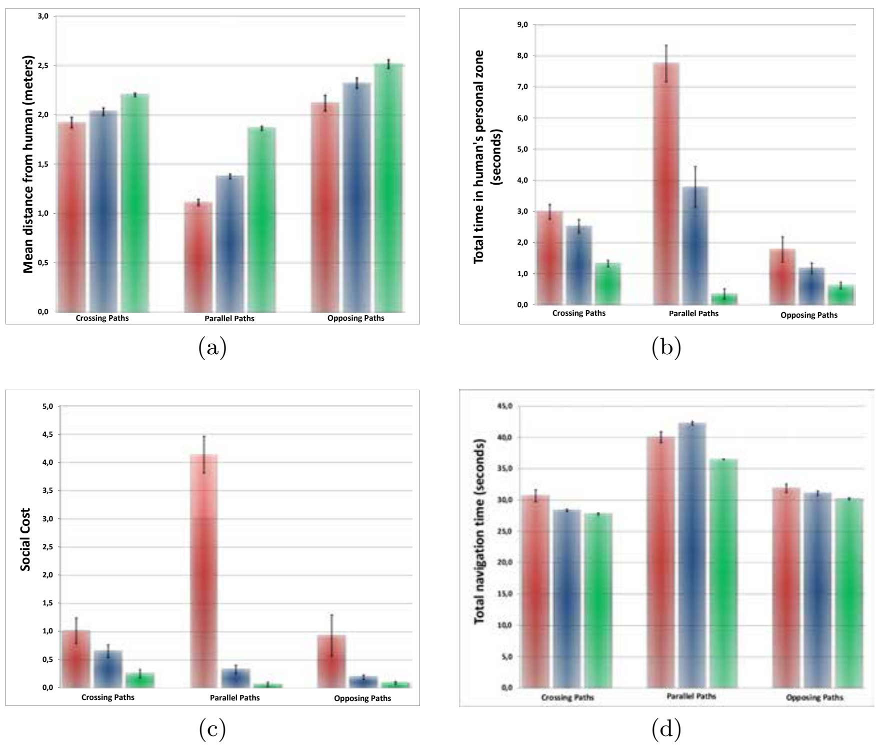
3.1 Implementation details

To facilitate autonomous robot navigation, specific components are required to be present, each of which is briefly discussed herein aiming to provide the means to the reader to reproduce the proposed method. For the metric mapping which is responsible to provide the geometrical representation of the robot's surroundings, a Gmapping implementation has been selected, since map construction slightly impacts on the method. To keep track of the robot's pose within the explored environment an AMCL localization method based on LIDAR scans has been adopted. The last required component is the path planning implementation which is responsible to navigate the robot from a current pose to a target location. The utilized global planning implementation is D^* Lite algorithm, while for the local planner the default Dynamic Window Approach suitable for short-wheel-axis differential platforms has been utilized. A diagram illustrating the connectivity among the aforementioned modules and the novel human presence anticipation one is exhibited in Fig. 1. The hardware setup on which the proposed method has been implemented and evaluated comprises a Turtlebot2 differential robotic platform equipped with an Asus Xtion PRO LIVE RGB-D sensor and a Hokuyo URG-04LX scanning laser rangefinder. The laser is placed on the top plate of the platform in a front facing orientation taking advantage of its 270ø FoV. The Asus Xtion is placed in the middle of the top plate of the platform yet, slightly lifted in the vertical axis to obtain better human view. The onboard computational unit is a notebook equipped with an Intel^R Core^{TM} i7-3632QM CPU operating at 2.20GHz with 8GB RAM.

3.2 Social Acceptable Behavior Assessment

For the evaluation of the robot's ability to retain socially acceptable behaviour, four different metrics already established in the community [15] have been utilized. However, these metrics have been extended to consider also moving human instead of a static one and the referenced time duration T corresponds to the time interval that the robot needs to navigate from its current pose to the target one:

- M_1: The mean distance D_{mean} among the moving human and robot which is computed in the the entire navigation process.
- M_2: The time spent in areas associated with the human personal zone
- M_3: The social cost which indirectly models the human discomfort by computing $m_3 = \sum_{t=1}^{T} Cs_t \dot{d}_t$, where Cs is the social cost of the cell where the robot's footprint is located and d_t is the duration that the robot operated in that cell. The sum of this product declares the overall time the robot spent within the human personal space during its navigation and indirectly models the human discomfort factor.
- M_4: Total navigation time needed for the robot to reach the desired goal.



(a) (b) (c) (d)

Fig. 3. a) M_1 mead distance from humans in meters, b) M_2 total time spent in human's personal zone, c) M_3 social cost and d) M_4 total navigation time

3.3 Experimental Results

The method evaluated in three different test cases namely, crossing, parallel and opposing direction movements among human and robot, while has also been compared with two other methods, no human aware navigation (ROS native implementation) and modeling of human presence with static Gaussian kernel model [15]. Each method underwent 30 repetitions for each scenario ,totaling to 270 experiments. The indoor area where the experiments have been conducted corresponds to $35m^2$ approximately, ensuring enough space for robot and human co-navigation. The area consists of three doors which have declared as the frequently visited regions of the space and were considered for the human motion intention prediction. The results of the experiments are illustrated in Fig. 3 for each scenario, indicating superior performance of the proposed method when compared to the rest two methods. In general, our dynamic approach in the M_1 mean human-robot distance achieved in total 2.21m, while static human model and pure navigation achieved 1.92m and 1.71m. This result is directly association with the fact that the time the robot spent in the humans personal zone, our dynamic model scored a mean of 0.7sec, while the static human model and the pure navigation approaches scored 2.49sec and 4.17sec respectively. In M_3 metric, social cost calculated from the previous measurements, the dynamic model scored 0.139 units, whereas static human model 0.393 and pure navigation achieved 2.029 units, respectively. Finally, the last metric M_4 does not measure the social behavior of the robot, however is a navigation performance indicator of the robot, where the dynamic human model achieved 31.4sec, the static human model and pure navigation scored 33.9sec and 34.2sec, respectively exhibiting thus, that the proposed method retains better social behavior and minimum robot travel time as well. This occurs due to the prediction of human motion and the earlier selection of the correct path that the robot needs to follow. The best results were achieved in the scenario where human and robot followed parallel with same direction paths, where both the competitive models had achieved high social scores. That was expected since there is a lot of overlap between human and robot path, however the proposed method successfully maintained better social behavior.

4 Conclusions and Discussion

In this work a human motion anticipation strategy has been introduced suitable for real-time robot navigation in human populated environments. The human presence in the robot's workspace is perceived by fusing multimodal perception modules, while the human walking intention is modeled based on the theory of the human frequently visited areas and the computation of probabilistic human-like paths. The predicted human presence in the robot's workspace is facilitated by adding a separate human-layer in the navigation framework which is parameterized so as to operate in real time scenarios where mutual or unilateral motion of human and robot is observed. The ability of the proposed method to preserve social acceptable behavior has been evaluated with a robotic platform in a real-life experiments where the designed methodology exhibited remarkable results. In our future work, we plan to extend the experimental procedure to consider multiple humans in the explored environment.

Acknowledgments

This work has been supported by the EU Horizon 2020 funded project namely: "Robotic Assistant for MCI Patients at home (RAMCIP)" under the grant agreement with no: 643433.

Bibliography

[1] Buttimer, A., Seamon, D.: The human experience of space and place. Routledge (2015)

[2] Charalampous, K., Kostavelis, I., Gasteratos, A.: Robot navigation in large-scale social maps: An action recognition approach. Expert Systems with Applications 66, 261–273 (2016)

[3] Flaherty, G.: Dementia and wandering behavior: Concern for the lost elder. Springer Publishing Company (2006)

[4] Fulgenzi, C., Tay, C., Spalanzani, A., Laugier, C.: Probabilistic navigation in dynamic environment using rapidly-exploring random trees and gaussian processes. In: IROS. pp. 1056–1062. IEEE (2008)

[5] Hall, E.T.: The hidden dimension, vol. 1990. Anchor Books New York (1969)

[6] Kim, B., Pineau, J.: Socially adaptive path planning in human environments using inverse reinforcement learning. Int. J. of Social Robotics 8(1), 51–66 (2016)

[7] Kirby, R., Simmons, R., Forlizzi, J.: Companion: A constraint-optimizing method for person-acceptable navigation. In: RO-MAN 2009. pp. 607–612. IEEE (2009)

[8] Large, F., Vasquez, D., Fraichard, T., Laugier, C.: Avoiding cars and pedestrians using velocity obstacles and motion prediction. In: Int. Vehicles Symposium. pp. 375–379. IEEE (2004)

[9] Leigh, A., Pineau, J.: Laser-based person tracking for clinical locomotion analysis. In: IROS-Rehabilitation and Assistive Robotics (2014)

[10] Luber, M., Spinello, L., Silva, J., Arras, K.O.: Socially-aware robot navigation: A learning approach. In: IROS. pp. 902–907. IEEE (2012)

[11] Mac, T.T., Copot, C., Tran, D.T., De Keyser, R.: Heuristic approaches in robot path planning: A survey. RAS 86, 13–28 (2016)

[12] Papadakis, P., Spalanzani, A., Laugier, C.: Social mapping of human-populated environments by implicit function learning. In: IROS. pp. 1701–1706. IEEE (2013)

[13] Sisbot, E.A., Marin-Urias, L.F., Alami, R., Simeon, T.: A human aware mobile robot motion planner. IEEE TRO 23(5), 874–883 (2007)

[14] Svenstrup, M., Bak, T., Andersen, H.J.: Trajectory planning for robots in dynamic human environments. In: IROS. pp. 4293–4298. IEEE (2010)

[15] Talebpour, Z., Navarro, I., Martinoli, A.: On-board human-aware navigation for indoor resource-constrained robots: A case-study with the ranger. In: Symposium System Integration. pp. 63–68. IEEE (2015)

[16] Trautman, P., Ma, J., Murray, R.M., Krause, A.: Robot navigation in dense human crowds: Statistical models and experimental studies of human–robot cooperation. IJRR 34(3), 335–356 (2015)

[17] Unhelkar, V.V., Pérez-D'Arpino, C., Stirling, L., Shah, J.A.: Human-robot co-navigation using anticipatory indicators of human walking motion. In: ICRA. pp. 6183–6190. IEEE (2015)