Incorporating Perception Uncertainty in Human-Aware Navigation: A Comparative Study

Zeynab Talebpour[1]   Deepak Viswanathan[2]   Rodrigo Ventura[3]   Alcherio Martinoli[4]

Abstract— In this work, we present a novel approach to human-aware navigation by probabilistically modelling the uncertainty of perception for a social robotic system and investigating its effect on the overall social navigation performance. The model of the social costmap around a person, has been extended to consider this new uncertainty factor, which has been widely neglected despite playing an important role in situations with noisy perception. A path planner based on the fast marching method has been used to demonstrate this idea and the effectiveness of the proposed approach has been tested in extensive experiments carried out with real robots and in simulation. Real experiments have been conducted, given noisy perception, in the presence of single/multiple, static/dynamic humans. Results show how this approach has been able to achieve trajectories that are able to keep a social distance to the people, compared to those of the basic navigation approach, and the human-aware navigation approach which relies solely on perfect perception, when the complexity of the environment increases. Accounting for uncertainty of perception is shown to result in smoother trajectories with lower jerk that are more natural from the point of view of humans.

I. INTRODUCTION

Navigation is one of the main required functionalities for enabling robots to be actively used in real social environments. Robots have to navigate in environments shared with humans and the quality of their movement strongly influences how their intelligence is perceived [1]. It is necessary for such robots to be aware of the people and their social interactions in a given environment. Conventionally, comfort, naturalness and scalability, have been the main focus of human-aware navigation techniques [2]. However, when the robots are deployed in real complex environments, a key assumption of many of these methods, which is having perfect information about the position of the people, is too simplistic and not robust to non-negligible perception uncertainty.

One important concept used in numerous studies [3]–[6] in this area, is the virtual space around a person that is mutually respected by other humans, called *proxemics* [7]. Based on this concept, depending on the relationship and the interaction that exists between humans, people choose different social distances relating to intimate, personal, social or public contexts. Social costmaps are a common way to

model this principle that and been used in various studies in the field. Many factors can be considered for shaping a social costmap, such as age and gender [8], velocity of the motion [9], etc. However, the proxemics distance has been the main factor when accounting for comfort in the literature.

Navigation exhibits human-awareness by modifying the plans given to the robot as the result of adopting socially-aware path planners. Such a planner should take into consideration individual people and possible social interactions taking place between them, when computing the optimal path. This information is provided through perceptual data and is translated to social costs using models such as social costmaps. However, state-of-the-art perception methods are not perfect and are affected by various elements such as the robot movement, people's movement, complexity of the environment in terms of occlusions, etc.

Due to the approximate nature of the models and the less than perfect human detectors available, we often can only provide estimates of the locations of the people with an associated uncertainty. Thus, any planning algorithm relying on real perception data, must be able to cope with this inherent uncertainty. Despite this fact, the assumption of having perfect information about the position of people at all times is common in the state-of-the-art research in this area and the main focus is on the planning itself. However, moving to real applications, poses serious challenges in terms of noisy perception information and high uncertainties, that need to be addressed and modelled in a human-aware approach.

In this work we propose a model that explicitly accounts for perception uncertainty by incorporating the uncertainty of position estimates of people into the social costmap. We aim to study, the effect of this factor on social costs and how taking this uncertainty into account in the extended model, can result in trajectories that are improved in terms of social acceptability. The proposed human-aware navigation system, uses uncertainty-based social costs along with a Fast Marching Method (FMM)-based [10] path planner for achieving socially-aware plans. It should be emphasized that our approach is planner-independent as long as the planner takes costmaps as input, *e.g.*, it could be used in combination with ROS's navigation stack [11].

The proposed approach is implemented in reality and tested in a stochastic environment with varying perception uncertainty, in an extensive set of experiments. Six quantitative evaluation metrics have been introduced and the results of applying them to the proposed methods have been reported. Real robot experiments show how accounting for uncertainty of perception improved the robot trajectories in terms of

[1,4]Distributed Intelligent Systems and Algorithms Laboratory, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne {zeynab.talebpour, alcherio.martinoli}@epfl.ch

[2]Deepak Viswanathan, Department of Informatics, University of Amsterdam D.geethaviswanathan@uva.nl

[3]Rodrigo Ventura, Institute for Systems and Robotics, Instituto Superior Técnico, Universidade de Lisboa, Portugal rodrigo.ventura@isr.ist.utl.pt

social distance and naturalness, as the complexity of the environment grows.

The remainder of this article is organized as follows: in section II we briefly introduce the related works in this area. Section III explains the probabilistic perception model and Section IV explains the model of social navigation. The robotic platform and the experimental setup will be explained in detail in section V. In section VI, we show the results of the real robot experiments and finally, section VII concludes this work.

II. BACKGROUND

Human-aware navigation focuses on the interaction dynamics between humans and robots that occur as a result of navigation [2]. In the literature, we can find several strategies for comfort ranging from appropriate approaching strategy [12], maintaining appropriate distance [4], control strategies to avoid noisiness [13] and use of planning for avoiding interference [14]. In this work we focus on the principle of proxemics which is the most common in the literature of human-aware navigation, with social costs encoded as costmaps similar to [15] along with an FMM-based path planner.

FMM has been proven to be successful in real domestic spaces with high complexity [16]. There have previously been a number of research papers addressing social path planning [15], [17] using FMM. In [17] a theoretical framework for a number of introduced sub-problems of social path planning is presented and an extended model for engaging groups of people is proposed by using a special version of fast marching square planning method [18]. Nonetheless, the information about humans are considered to be given and noiseless, while only simulations have been used to show the effectiveness of the method for static people. The same problem needs to be investigated in real-world scenarios with the challenges that exist therein. Particularly, in the case of moving people, while the perception is subject to uncertainty.

Uncertainty of perception in social robotics is an important topic which to our knowledge, has not been the subject of many notable studies. There is a dedicated chapter in [19] on local planning with uncertainty, however, this is not considered in a social context. The sources of uncertainty in [19] are the position of the robot and the obstacles and the partially known *motion* of moving obstacles that are considered to be people. However, perception of people and the uncertainty in person and group detection and tracking has not been investigated.

III. PERCEPTION MODEL

Presence of humans in an environment should be properly perceived by a robot. This is a must for any socially-aware path planner that needs to take into consideration, individual humans and possible social interactions taking place. This information can be obtained by an external source such as an overhead camera or can be attained using on-board sensors of the robot. Different perception sources for person detection and tracking, have different levels of uncertainty and accuracy in their detections, and are affected by various elements such as the movement of the robot, movement of the person, complexity of the environment in terms of occlusions, etc. While, there exist special-purpose expensive trackers able to perform this task with high level accuracy, others have a much larger uncertainty associated to them.

By taking the uncertainty of perception into account in a human-aware path planner, the same planning method could easily be reused even when the source of perception changes. As an example, while overhead cameras can provide position information of people with good accuracy, when moving to on-board perception for a mobile robot, this accuracy and the associated uncertainty will largely change. This is even more significant, if tracking is done with low-precision detectors, such as RFID and UWB. So, for more robustness and effectiveness, a human-aware navigation method, should be able to handle all these situations without undergoing major changes.

In the proposed framework, a probabilistic approach has been chosen to account for uncertainty in the tracking rather than deterministically reporting positions and dealing with the perception uncertainty inside the tracker module only. We are interested in the underlying *state* of the environment which is the position of the people. The *detectors* used to estimate these state variables have associated noise due to various factors such as occlusion, lighting conditions, different postures of people, motion of the robot and the people, etc. Coupled with this, there is also stochasticity in the state transitions, which makes it hard to compute an exact estimate of the location of the people. A principled approach to solve this problem, is to compute a *belief* (posterior distribution) over the states using recursive Bayesian estimation. We first describe the state representation of the system and then explain the tracking model formally.

A. Detector

The background-based detector proposed in [ref] is a very effective probabilistic method, which allows the automatic evaluation of the number of people in the scene and detection of their locations. This method has the following advantages: (1) It can incorporate prior knowledge, including which areas in the scene can contain people and how probable it is for people to be in those locations; a probability distribution over the number of people in the scene; a probabilistic model of how close together people tend to walk; etc. (2) The complexity of the algorithm depends linearly on the number of people in the scene. When many people are present in the frame, detecting all of them requires more than 40 *ms* with our current implementation of the algorithm. Although it still requires far less than a second, further optimizations could easily improve this performance. (3) The method is robust to changes in illumination, shadows and occlusions. It is adapted it to adjust to a non-static background automatically.

B. State representation

An occupancy grid-based approach is used for tracking the people in the environment. The environment is discretized

into *G* cells. Each cell is of size 25 *cm* by 25 *cm*. The size of cells have been chosen in such a way that each cell can be occupied at most with one person at any time. The occupancy of each cell is denoted by X_i where $i \in G$. The occupancy of all the cells at time t is the state of the world $\mathbf{X}_t$. At every time instance t, the observations from the detector for each cell i is given by O_i. The set of observations for the whole state is denoted by $\mathbf{O}_t$.

C. Tracking model

Let $\mathbf{X}_t$ be the state of the environment at time t.

$$P(\mathbf{X}_t|\mathbf{O}_t) = \frac{P(\mathbf{O}_t|\mathbf{X}_t)P(\mathbf{X}_t|\mathbf{X}_{t-1})}{P(\mathbf{X}_{t-1})} \qquad (1)$$

Where, $P(\mathbf{O}_t|\mathbf{X}_t)$ is the likelihood of the state given all our observations (detector outputs). Computation of this likelihood is best performed by using a learned model of how the detectors perform in different states.

$P(\mathbf{X}_t|\mathbf{X}_{t-1})$ is the *transition model* which models the evolution of the state variables. For a real multi person environment, an exact analytical model is intractable. So we choose to ignore this aspect of the environment and attribute it to the uncertainty we have in the state estimation. We assume that people move randomly and that there is an equal probability of motion in any direction.

In a multi person environment, the state space is extremely complex so as to compute exactly this probability distribution over the states. We use an MCMC-based sampling algorithm to approximately compute the belief. In the next section the implementation details of our probabilistic model for person detection and tracking are explained.

Although, the detector is modelled probabilistically, it is still needed to learn the distribution $P(\mathbf{O}|\mathbf{X})$ for the detector from data. Given the labelled location of the people in a data set, the uncertainty in the observations is learned for all the locations and configurations of the state space.

D. MCMC sampling

Markov Chain Monte Carlo is a widely used sampling algorithm for estimation of complex posterior distribution. It has been gaining popularity in multi-target tracking applications [20]. Compared to traditional particle filters, MCMC-based sampling leads to far less sample impoverishment and thus a much better estimate of the state over time.

The core idea of MCMC is to generate samples from a Markov chain. The samples are then evaluated using a *proposal distribution* and accepted or rejected based on an *acceptance probability*. The MCMC sampler creates hypothesis of the location of the people in the grids. Each sample is an estimate of the occupancy of all the cells taken collectively.

In this work, the occupancy of the grids are used as hypothesis. Each cell can either be occupied or not, with initially starting from a random distribution of occupancy and then generating samples using the following moves:

1. *Birth-Death proposals:* A cell is randomly selected, and the sample state of the cell is flipped. If the cell was occupied,

a proposal which makes the cell unoccupied is generated and vice-versa.

2. *Move proposals:* In this case, an occupied cell is selected and the occupancy is randomly moved to one of the 8 connecting neighbors.

Once the proposal sample is generated, we evaluate the original sample and the proposed sample with reference to a proposal distribution. In our case, we use a learned observation model of the detector output as the proposal distribution. We fold in the detector output $\mathbf{O}_t$ while evaluating the proposals using the proposal distribution. Every proposal is a hypothesis of the state $\mathbf{X}_t$. Evaluating the proposals will give us an acceptance probability. If the acceptance probability is greater than 1 the sample is accepted unconditionally. Otherwise, we randomly sample from a uniform distribution and then accept or reject the sample if the acceptance probability is greater than the sampled value of uniform distribution. Formally, The acceptance probability is computed as:

$$Acc(x|x^{'}) = \min\left\{ \left(\frac{P(\mathbf{O}|x)}{P(\mathbf{O}|x^{'})}\right), 1 \right\} \qquad (2)$$

Where x is the proposed sample of state $\mathbf{X}$, and $x^{'}$ is the initial sample.

If the sample is accepted, the currently proposed sample will be used as the initial sample for the next step of the MCMC sampling. If rejected, the sample is still added to the set of hypothesis, but we start sampling again from the old sample.

Sampling is successively repeated until N_s samples are accepted. The threshold for the number of accepted samples are set to be 100 in our experiments. Once N_s samples are accepted, they are collectively represented as the approximate representation of the multi modal posterior distribution.

For the purpose of social navigation, these samples are converted to a set of particles. Even though the set of samples are fixed, the set of particles can vary, since each sample is a joint sample which represents the whole occupancy grid. When there are multiple people in the environment, each one of the samples can be decomposed into a set of particles, proportional to the number of people in the environment.

IV. HUMAN AWARE NAVIGATION MODEL

In this section we will explain the navigation component of the system and describe the social cost computations methods that provide the information required for our socially-aware FMM-based path planner.

A. Navigation Framework

The robot navigation is based on the navigation system used in the MOnarCH project [21], detailed in [16]. As input, it uses the pose estimates provided by a standard AMCL self-localization system [22], given odometry, laser range finder readings, and a static map. The navigation system is based on FMM for motion planning, together with a Dynamic Window Approach (DWA) algorithm for guidance and obstacle avoidance [23]. DWA is essentially a maximization (over a discrete set of feasible velocity candidate commands)

of an evaluation function translating three guidance goals: (1) progress towards the goal, (2) clearance from obstacles, and (3) absolute speed.

The potential field output by FMM is minima free and yields an optimal path from a given initial to a final goal point. It is optimal in the sense that the integral of a costmap over the path is minimal, given the initial and final points as boundary conditions. However, we do not explicitly compute a path. Instead, we compute the progress towards the goal directly from the gradient of the potential field. FMM has been used in MOnarCH before, considering solely an increased cost near static map obstacles in its costmap. This keeps the resulting paths away from the obstacles. In this paper a *social* component is added to this costmap as well.

FMM and DWA run asynchronously. FMM is activated when either a new goal position is given, or when the costmap changes, and DWA is running in a closed loop with a fixed rate, using the last updated potential field from the FMM.

B. Social Costs

The personal space around a human can be defined as the mixture of two pseudo-Gaussian functions, one for the front and another one for the rear part of the area surrounding the person. The orientation and heading of the person will cause a corresponding rotation in these functions in such a way that the person is always in the center and the absolute orientation of the person matches that of the Gaussian functions. A Gaussian function ϕ, centered on p with covariance matrix Σ, is defined as follows:

$$\phi(q) = exp\left\{-\frac{1}{2}(q-p)\Sigma^{-1}(q-p)\right\} \qquad (3)$$

q indicates the position of a point and Σ is:

$$\Sigma = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \qquad (4)$$

σ_x and σ_y are used to modulate the shape of the Gaussian and are traditionally chosen in a way to respect the personal space of a person as indicated by the proxemics principle. Various factors can influence the size of this area, however σ_x and σ_y are commonly considered to be constant. Getting closer to a person, will cause an increase in the value of this function, and hence the social cost associated to that position will increase.

If the center of the costmap, which indicates the position of the person is not deterministically known, the costmap can not correctly model the social costs and hence the social path planning could fail in finding an appropriate socially compliant path. This problem becomes much more critical in real applications where robustness is vital for succeeding under different conditions. We believe probabilistic social costs can be a remedy to this problem.

In the following, we will go into the details of computing social costs with the particles obtained from the MCMC tracker. Using the concept of layered costmaps similar to [11], an FMM planner uses this information for performing the replanning.

1) Uncertainty-based Social Costs: The MCMC-based tracker provides samples of the *union* of the probability distributions for having a person in a given environment. More specifically, a predefined number of joint samples are reported at each time step, containing a set of particles for each person that the tracker detects. In the following, we propose two types of methods for using the particles given by these samples, to create social costmaps.

Convolution: The core idea we propose for incorporating uncertainty in the costmap is to compute an *expectation* based costmap. In other words, we treat the probability map of the presence of people in the environment as a two dimensional signal and convolve it with the conventional Gaussian costmap, to obtain the final probabilistic social costs. By convolving all the samples from the MCMC tracker with the social costmap, we compute an *expected costmap* incorporating all the uncertainty in the environment. This is a principled approach to solve the problem since we are not abstracting away any information provided by the perception system, and hence, in theory this approach should provide us with a costmap model that would be most informative for uncertainty-based human-aware navigation.

By taking this approach, the conventional 2D Gaussian shape of the social costmap is no longer mandatory, thus this costmap model is more flexible. Additionally, there is no need to know the number of people ahead of time as the samples given by tracker are reflecting this (refer to III-D).

Clustering: Another approach we propose is to abstract the uncertainty information in the samples by clustering the particles and then computing the uncertainty, based on the features of the cluster. Upon receiving the position samples from the tracker we compute the center of the social costmaps, and the σ values for all the people present in the environment. By clustering the particles and finding the centroids of the clusters, along with an uncertainty measure based on the spatial cluster scatter, we can adaptively compute the social costmaps at each time step. We will briefly describe the two clustering methods selected for our work in the following.

a) K-means clustering: K-means clustering [24] method can be adopted for computing the costmap centers and Σ. However, one requirement for using this method is knowing K which is the number of clusters and in our case the number of people, ahead of time. This is not a realistic assumption for dynamic environments with multiple people. Nonetheless, we will test this method as a baseline for comparing the performances of other human-aware navigation methods both in simulation and reality.

b) Mean shift clustering: To overcome the limitation of K-means clustering method for needing to know the number of clusters ahead of time, we chose another clustering method, mean shift clustering [25], that is able to determine the number of clusters given the sample set automatically. We used the median of all pairwise distances to estimate the bandwidth of the mean shift method.

2) Deterministic Social Costs: For the purpose of creating a baseline to compare our proposed model, we have also designed a deterministic model for the social costmap. Here,
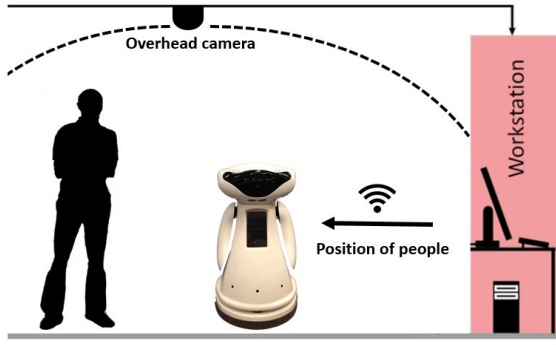
Fig. 1.   Experimental setup schematics and the MBot robotic platform.

we use a deterministic output from the tracker, without considering the particles, but just using the mode of the distribution as deterministic location of the people, along with the conventional costmap model.

This is the standard approach that is being used for social path planning in human-aware navigation, where it is assumed that the position and orientation of the person is deterministically known at every time step with no model of uncertainty. Once this information is known, the conventional Gaussian shaped costmap can be used to obtain the social costs. This costmap can be tuned according to the desired parameters. In this paper, we have taken an approach similar to [15] with $\sigma_x = 0.255$ *m*, for deterministic costmaps.

V. EXPERIMENTS

In this section we will briefly explain our robotic platform, experimental setup and the set of experiments along with the metrics which have been conducted to evaluate the performance of our system.

A. Robotic Platform

The robotic platform used in this work is shown in Fig. 1. This robot is called MBot [26] and is developed within the FP7 European project MOnarCH. It is an omni-directional drive robot with an approximately round footprint of 0.65 *m* in diameter and a height of 0.98 *m*. It is provided with two laser range finders, on both the front and the back for providing 360° coverage.

Two batteries give it an autonomy of approximately five hours, depending on the usage. The robot has two PCs inside its shell: one manages the sensors, navigation and actuators, while the second one is for other functions such as human-robot interaction functionalities, which are outside our interests for this work. The two on-board PCs, run Ubuntu desktop 12.04 and ROS Hydro.

B. Experimental Setup

We have used an extensive suite of experiments both in simulation and reality, for performance evaluation. A high fidelity robotic simulator Webots [27], with realistic models of the MBot and the testing environment has been used for evaluations in the initial step. The trackers have been emulated initially, but real recorded data bags of perception were used

in the next steps for more close to reality simulations. This was done to simplify the debugging process and speeding up the robotic experiments. Finally, we tested our methods in reality in three different scenarios, in a laboratory environment of 5 *m* × 7 *m* where the tracker was operational.

We used a networked omnidirectional overhead camera with a field of view of 180°, to track the positions of the people in the environment. This type of camera was chosen because: (1) it is less obtrusive, and can be left in the environment with less risk of making people feel uncomfortable about being watched; (2) it provides a global view of the area, with lower risk of occlusion than elevated side-view cameras and with more flexibility as to its positioning; (3) the number of cameras needed in the environment can be reduced, which has benefits in terms of equipment cost, installation cost and computational load of the perception algorithms.

The tracker outputs results at the rate of 3 *Hz*. The ground truth position of the robot is given by AMCL with 5-10 *cm* accuracy, and the person stands and walks on physically marked tracks to get the exact precise ground truth for the purpose of performance evaluation. The control rate of navigation is 20 *Hz* while the social costmap generation has a rate of approximately 3 *Hz*. This is to account for the low output frequency of the tracker.

C. Scenarios

We have studied three different scenarios, each having been tested five times. We started with 1) a single static person and incrementally increased the complexity to 2) one moving person and finally, 3) two static people in the arena. It should be emphasized that perception uncertainty is affecting the tracking performance and is not evident or quantifiable from just looking at the environment. This means, the person is not aware of what is happening on the tracker side, however, the information given by the tracker greatly affects the behavior of the robot, and therefore the social acceptability.

Since we aim to study the effect of perception uncertainty in human-aware navigation, we chose a task of point to point navigation for the robot in the vicinity of humans, which is the most general and basic navigation task.

In each experiment, the robot starts from a predefined starting point and is sent to one predefined goal. The robot then has to behave appropriately when it encounters people in the arena. For the static case, there is always a person standing between the robot and the straight line to the goal, and for the dynamic case the person moves along this line in the opposite direction, as the robot starts navigating towards the goal, causing a direct encounter with the robot. The following section will explain the metrics used for performance assessment.

D. Metrics

Six different metrics have been defined for performance evaluation. A subset of these metrics is chosen for each experiment based on the scenario of interest.

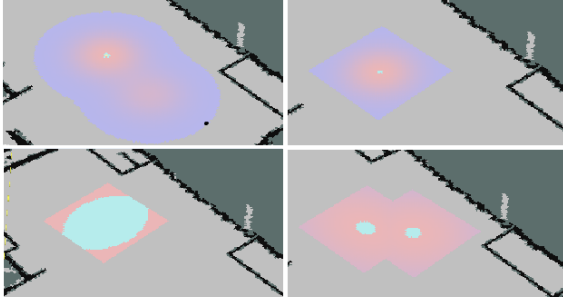- m_1: Measures the minimum distance that the robot has kept during the experiment to a human.

Fig. 2. Sample costmap shapes. Top right: standard 2D Gaussian, top left: Convolution method, bottom right: mean shift and bottom left: K-means.

- m_2: Evaluates how long the robot has been moving in areas associated with social costs, *i.e.*, a position with corresponding non-zero value in the costmap.
- m_3: Quantifies the accumulated social cost, this is to differentiate between being in different positions of the social costmap, which is not reflected in m_2. So if the robot is closer to a person, the social cost will be higher and this metric will increase. For more information on $m_1 - m_3$ refer to [28].
- m_4: Evaluates the smoothness of the robot trajectory. This is important from the human observer's point of view when perceiving the robot motion. Humans are known to prefer motion with minimum jerk [29], therefore we took the root mean squared error (RMSE) of the trajectory jerk in m_4:

$$r_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}, \ m_4 = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left|\frac{\mathrm{d}^3 r_t}{\mathrm{d}^3 t}\right|^2} \qquad (5)$$

r_t indicates the position of the robot at time t, and N indicates the total time of the experiment. It should be emphasized that we did not actively try to modify the robot control to get smoother trajectories, we are just interested to see which method results in a more natural and smooth path.

- m_5: Is the total time steps required to finish the navigation task.
- m_6: Measures the minimum of the average distance of the robot to all the people present in the environment. This is interesting since it shows how the robot is choosing to use the space. Does it choose to go through the area between the people or does it prefer to take a tour and increase this minimum average distance?

VI. RESULTS AND DISCUSSION

For each of the scenarios described in section V-C, we have compared the results obtained from the basic navigation (BN), deterministic (DHA), K-means clustering (KHA), mean shift clustering (SHA), and social cost convolution (CHA) human-aware (HA) navigation. We will only report the trajectories and metrics obtained from the results of our real experiments for the sake of conciseness. Larger values for m_1 and m_6, and smaller values for m_2-m_5 are preferable.
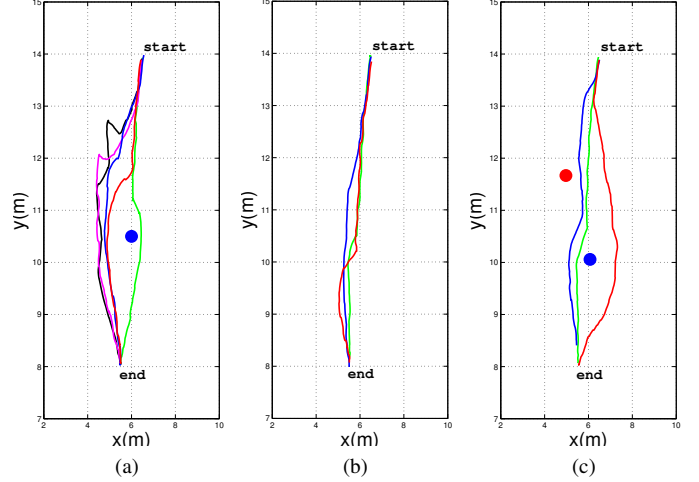

Fig. 3. Sample robot trajectories for different methods. Color coding: green for BN, red for DHA, black for KHA, pink for SHA and blue for CHA. Scenarios: (a) Single static person. (b) Single dynamic person. The person starts moving from the end point to the starting point of the robot's trajectory indicated by labels on the plots. (c) Two static people.

A. Results

Figure 2 shows sample costmaps of the different methods mentioned earlier. It can be seen that the clustering methods can end up with saturated costs when the uncertainty is high due to large σ values. The convolution costmap has a much more flexible shape and is not limited whereas other costmaps all have a cut off distance.

Sample trajectories of the robot are depicted in Fig. 3. It is clear from the plots that HA methods result in trajectories that preserve larger distances to the people. Additionally, they are smoother and therefore more natural from the point of view of a person, this is supported by Fig. 4d, 5a, and 6d. However, this may not be evident from the trajectory plots. This is due to the abrupt movement of non-HA navigation upon encountering a person which considerably affects the smoothness.

Clustering methods can cause the robot to modify its plan largely by enforcing a certain cluster shape upon finding cluster centroids: if the new probabilistic data leads to a new centroid that is not very close to the previous one, the costmap could change significantly and therefore the planned path. This is more severe for mean shift clustering due to adaptively modifying the number of clusters as well. This is to be expected given the probabilistic nature of the perception data, however the plan can be modified more smoothly using the convolution method. This method which outperforms all other methods in terms of smoothness in all our experiments, is shown to be a remedy to this problem based on our experimental results. Hence, for the second and third scenario, we only compare the results of 1) BN, 2) DHA, and 3) CHA methods.

When comparing the results of DHA with CHA we can see that the former is a more conservative method in terms of keeping distance to the people when receiving accurate data.
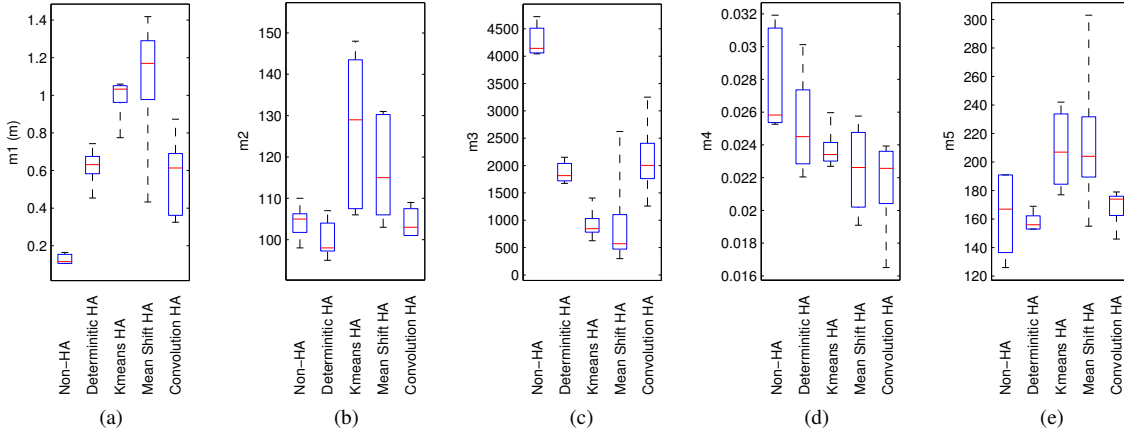
Fig. 4. Performance metrics obtained in the single static person scenario. HA stands for Human-Aware in the plot labels.
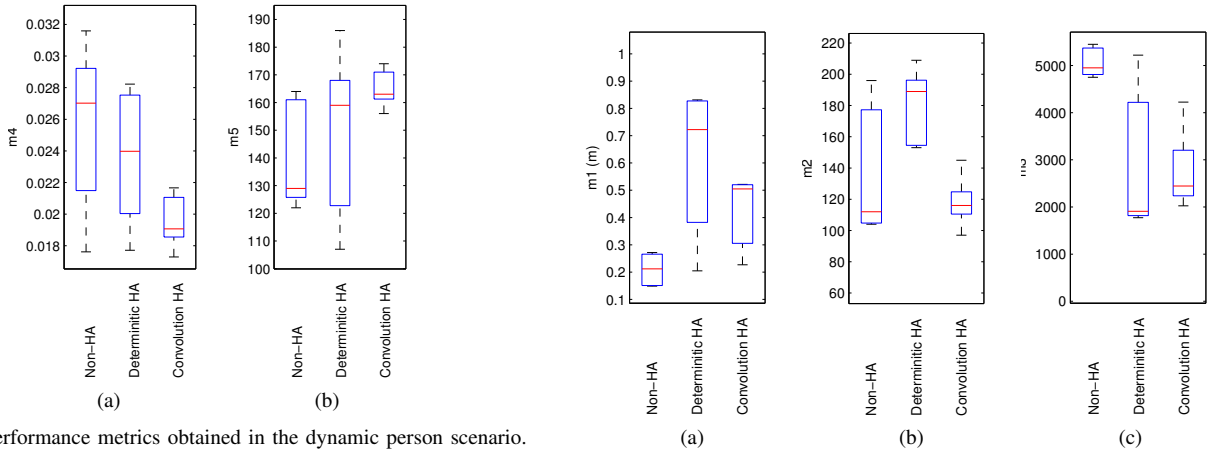

Fig. 5. Performance metrics obtained in the dynamic person scenario.

Fig. 6. Performance metrics obtained in the two static people scenario.

If DHA receives a perfect estimate of the person's position it can lead to the desired path, however this is seldom the case. Particularly, in the case of a moving person, the detector could not always keep up with the speed of the person, *i.e.*, the position estimates were reported with delay or the person was lost in some cases, and the robot was faced with the human while considering him an obstacle. This led to abrupt changes and getting too close to the person, see Fig. 3b. However, CHA by associating larger uncertainty to the estimates in this case, could lead to better plans in terms of proximity and smoothness. Unfortunately, due to our inaccurate ground truth of the moving person, we only rely on m_4 and m_5 for scenario 2, but we observed the delayed perception and lost person problem during our experiments.

Figures 4, 5 and 6 show the performance metrics for the three scenarios. We will discuss the results of each metric in the following. m_1 has increased for HA methods which shows the effectiveness of our FMM planner in social path planning. However, DHA is more conservative in this regard in the presence of good perception data. m_2 has increased for uncertainty-based methods in the simple scenario as the deterministic tracker is giving already good position estimates, but this is no longer the case when the complexity is increased as seen in Fig. 6b. m_3 is also reduced for HA methods and
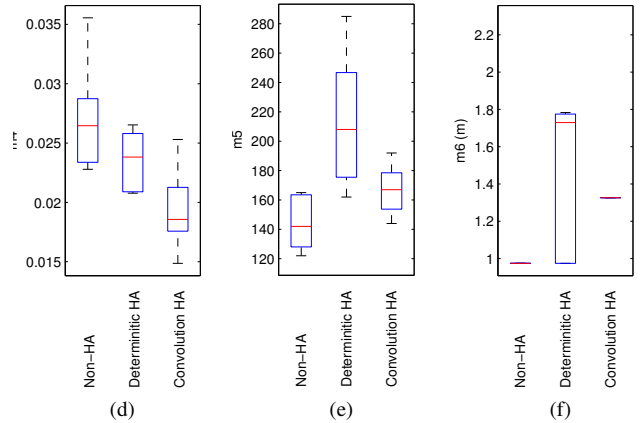
more so for uncertainty-based HA methods as the complexity of the problem increases, see Fig.6c.

m_4 is showing a very interesting result, we can see how uncertainty-based HA methods have managed to introduce smoothness into the trajectories without deliberately accounting for it, and CHA is a clear winner in this case. m_5 which shows the total navigation time is larger for KHA and SHA due to constant modifications of the path and taking longer

routes. For DHA this metric is comparable to CHA but it increases as the complexity of the environment grows. Lastly, m_6 which is only used in the third scenario, indicates how DHA has varying behavior (sometimes takes a detour) based on the tracker estimates whereas CHA has more or less the same behavior in our runs due to its probabilistic nature when computing the social costmap.

VII. Conclusion

In this work we have proposed a novel approach for extending the model of social costmaps to include uncertainty of perception. Experiments show how this extended model can lead to more natural robot trajectories that preserve a social distance from people. By combining the output of a probabilistic MCMC-based tracker with an expectation costmap computation method based on convolution, we introduce a principled approach to solve the social path planning problem in real environments with multiple people.

The idea presented in this paper can be extended to other types of perception sources. By providing probabilistic perception outputs, the proposed model can fulfill this task without loss of performance. As future steps for this research, we plan to test other types of perception, particularly on-board perception which allows for much more flexibility in terms of applications and fewer physical limitations. More tests with more accurate ground truth data can help to understand the problem and its challenges better. Moreover, quantifying the uncertainty of perception can be very useful in analyzing the behavior of expectation-based social costmap computation methods for further in-depth studies.

Acknowledgements

References

[1] P. Althaus, H. Ishiguro, T. Kanda, T. Miyashita, and H. Christensen, "Navigation for human-robot interaction tasks," in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 2, 2004, pp. 1894–1900.

[2] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1726 – 1743, 2013.

[3] J. Mumm and B. Mutlu, "Human-robot proxemics: Physical and psychological distancing in human-robot interaction," in *6th ACM/IEEE International Conference on Human-Robot Interaction*, 2011, pp. 331–338.

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

[5] M. Walters, M. Oskoei, D. Syrdal, and K. Dautenhahn, "A long-term human-robot proxemic study," in *IEEE RO-MAN*, 2011, pp. 137–142.

[6] G. Ferrer, A. Garrell, and A. Sanfeliu, "Robot companion: A social-force based approach with human awareness-navigation in crowded environments," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 1688–1694.

[7] E. T. Hall, "The hidden dimension," 1966.

[8] K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, A. Sisbot, R. Alami, and T. Siméon, "How may i serve you?: a robot companion approaching a seated person in a helping context," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, 2006, pp. 172–179.

[9] R. Kirby, R. Simmons, and J. Forlizzi, "Companion: A constraint-optimizing method for person-acceptable navigation," in *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 2009, pp. 607–612.

[10] J. A. Sethian, "Fast marching methods," *SIAM review*, vol. 41, no. 2, pp. 199–235, 1999.

[11] D. V. Lu, D. Hershberger, and W. D. Smart, "Layered Costmaps for Context-Sensitive Navigation," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014.

[12] K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, and C. L. Nehaniv, "How may i serve you?: a robot companion approaching a seated person in a helping context," in *1st ACM SIGCHI/SIGART conference on Humanrobot interaction*. ACM Press, 2006, pp. 172–179.

[13] E. Martinson, "Hiding the acoustic signature of a mobile robot," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 985–990.

[14] D. Vasquez, P. Stein, J. Rios-Martinez, A. Escobedo, A. Spalanzani, and C. Laugier, "Human Aware Navigation for Assistive Robotics," in *13th International Symposium on Experimental Robotics*, 2012.

[15] J. V. Gomez, N. Mavridis, and S. Garrido, "Social path planning: Generic human-robot interaction framework for robotic navigation tasks," in *2nd Intl. Workshop on Cognitive Robotics Systems: Replicating Human Actions and Activities*, 2013.

[16] R. Ventura and A. Ahmad, "Towards optimal robot navigation in domestic spaces," in *RoboCup 2014: Robot World Cup XVIII*. Springer, 2015, pp. 318–331.

[17] J. V. Gómez, N. Mavridis, and S. Garrido, "Fast marching solution for the social path planning problem," in *IEEE International Conference on Robotics and Automation*, 2014, pp. 2243–2248.

[18] A. Valero-Gomez, J. V. Gomez, S. Garrido, and L. Moreno, "Fast marching methods in path planning," *IEEE Robotics and Automation Magazine, no*, 2013.

[19] J. Correa Villanueva, "Uncertainty and social considerations for mobile assistive robot navigation," Ph.D. dissertation, Imperial College London, 2014.

[20] Z. Khan, T. Balch, and F. Dellaert, "An mcmc-based particle filter for tracking multiple interacting targets," in *Computer Vision-ECCV*. Springer, 2004, pp. 279–290.

[21] J. Sequeira, P. Lima, A. Saffiotti, V. Gonzalez-Pacheco, and M. Salichs, "Monarch: Multi-robot cognitive systems operating in hospitals," in *ICRA workshop on many robot systems*, 2013.

[22] D. Fox, "Kld-sampling: Adaptive particle filters," in *Advances in neural information processing systems*, 2001, pp. 713–720.

[23] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," *IEEE Robotics & Automation Magazine*, vol. 4, no. 1, pp. 23–33, 1997.

[24] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no. 1, pp. 100–108, 1979.

[25] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[26] J. Messias, R. Ventura, P. Lima, J. Sequeira, P. Alvito, C. Marques, and P. Carriço, "A robotic platform for edutainment activities in a pediatric hospital," in *IEEE International Conference on Autonomous Robot Systems and Competitions*, 2014, pp. 193–198.

[27] O. Michel, "Webots: Symbiosis between virtual and real mobile robots," in *Virtual Worlds*. Springer, 1998, pp. 254–263.

[28] Z. Talebpour, I. Navarro Oiza, and A. Martinoli, "On-board human-aware navigation for indoor resource-constrained robots: A case-study with the ranger," in *IEEE/SICE International Symposium on System Integration*, 2015.

[29] E. A. Sisbot, L. F. Marin-Urias, X. Broquere, D. Sidobre, and R. Alami, "Synthesizing robot motions adapted to human presence," *International Journal of Social Robotics*, vol. 2, no. 3, pp. 329–343, 2010.