Decision-Theoretical Navigation of Service Robots Using POMDPs with Human-Robot Co-Occurrence Prediction

Regular Paper

Kun Qian1,2,*, Xudong Ma1,2, Xianzhong Dai1,2, Fang Fang1,2 and Bo Zhou1,2

1 Key Lab of Measurement and Control of Complex Systems of Engineering, Ministry of Education, China
2 School of Automation, Southeast University, Nanjing, China
* Corresponding author E-mail: kqian@seu.edu.cn

Received 28 Apr 2012; Accepted 22 Jan 2013
DOI: 10.5772/55926
© 2013 Qian et al.; licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract To improve the natural human-avoidance skills of service robots, a human motion predictive navigation method is proposed, namely PN-POMDP. A human-robot motion co-occurrence estimation algorithm is proposed which incorporates long-term and short-term human motion prediction. To improve the reliability of probabilistic and predictive navigation, the POMDP model is utilized to generate navigation control policies through theoretically optimal decisions. A layered motion control structure is proposed that combines global path planning and reactive avoidance. Multiple comity policies are integrated with a decision-making module that generates efficient and human-compliant navigational behaviours for robots. Experimental results illustrate the effectiveness and reliability of the predictive navigation method.

Keywords Predictive Navigation, Motion Estimation, Uncertainty, POMDPs

1. Introduction

As service robots have been designed to provide interactive tasks in domestic and office environments, they must reliably navigate around a populated room. When robots and people encounter each other, human-aware motion planners[1] help robots treat people as social entities and aim to endow robots with safe and human-friendly navigational behaviours[2][3].

Predicting the motion of moving people is an effective way for compliant robot navigation in dynamic environments[4]. Many researchers[5][6][7][8] have developed efficient replanning algorithms to cope with the environmental dynamic and satisfy the real-time requirement by feeding the updated information to a grid map and optimizing the robot’s path to minimize the expected time to its destination. Although reactive motion planners[9][10][11][12] are able to rapidly query the next appropriate action, they are prone to getting robots...
blocked in complex environments because of their greedy property and the uncertainty of human motion. According to research on human indoor motion modelling and understanding, a human’s daily motions in a specific room environment presents certain long-term patterns. Nevertheless, uncertainties are pervasive in the velocity and heading direction of people’s movement. Several studies have exploited the spatial-temporal nature of human motion using a chain of Gaussian distributions\cite{13}, clustering the trajectories with K-means\cite{14}, and learning human motion patterns from tracking data using an EM algorithm\cite{16}. However, most past research ignores the combination of human motion uncertainty prediction with motion pattern prediction.

Another key factor of predictive navigation is inherent uncertainties\cite{15,16}, which typically necessitate intelligently reacting to unknown moving people. Due to the non-linear nature of human and robot motion, as well as sensor noise, the popular robot localization and people-tracking algorithm based on Bayes filtering can only estimate position distributions. The probabilistic human motion prediction algorithm is also likely to produce larger errors when making motion prediction.

Many researchers have already pointed out that probabilistic representation and reasoning is appropriate and very effective for navigating in noisy environments in the real world. In situations of probabilistic decision-making, Partially Observable Markov Decision Processes (POMDPs)\cite{17,18} have already been widely used in robot navigation and interacting with people. Robots such as Flo\cite{19} or Pearl\cite{20} use POMDPs at all levels of decision-making, and not only in low-level navigation routines. But since finding optimal control strategies in POMDP cases is computationally intractable due to the continuous and high-dimensional beliefs space, POMDPs have usually been applied to topological navigation. For example, Foka\cite{22} proposes a method for combining the prediction of the destination of a moving obstacle and its one-step-ahead positions. However, the proposed method relies on complex hierarchical decomposition of the environment and has only been implemented and successfully tested in a simulation application.

In this paper, a novel approach called POMDPs for Predictive Navigation (PN-POMDP) is proposed. The idea of predictive navigation is largely inspired by Sisbot\cite{5} human-aware planner, which focuses on providing human-friendly robot behaviours that imitate human motion habits. However, the human-aware planner does not take into account uncertainties when robots work in unstructured environments. This paper deals with the robustness and reliability requirements of a navigation system using the probabilistic reasoning (POMDP) method. We compute the uncertainties of human motion in two parts: the ambiguity in path selection and the motion uncertainty along each path. Then, the human-robot co-occurrence probability is estimated by analysing two situations: conflict and obstruction. Our major contribution is the PN-POMDP framework that coordinates the global path planner, the motion reactor and the speed controller in the context of probabilistic decision-making under conditions of multiple uncertainties. The control framework combines the objectives of goal-guiding, reducing the probability of human-robot conflict. More importantly, by considering high perceptual aliasing and other uncertainty factors, it combines the probabilistic robot localization, people-tracking and human motion prediction in a natural probabilistic decision-making framework to generate robot control policies resulting in efficient and polite navigation behaviours.

This paper is organized as follows. After an overview of the navigation system framework in Section 2, Section 3 describes the human motion prediction method and Section 4 introduces the human-robot co-occurrence estimation in the spatial-temporal aspect. Section 5 describes the decision-making mechanism of predictive navigation and the POMDP. Finally, some experimental results are reported in Section 6, followed by a final discussion and conclusion that summarizes the paper.

2. Navigation System Architecture

2.1 Uncertainties in predictive navigation

Firstly, the trajectories of human motion are uncertain; velocity and direction usually vary within a range when they are engaged in specific motion patterns. Secondly, the localization errors are pervasive. A Simultaneous robot Localization And People-tracking (SLAP) system using global cameras and an onboard laser range finder has been developed in our previous work\cite{23}. This jointly estimates a robot’s pose $r_i$ and people’s ground-plane position $h_i = (x_i, y_i, r_i)$ in the global coordinate frame using two sets of particles, as shown in Figure 1. But in clustering environments with table and chair legs, the localization errors are prone to deteriorate the human motion prediction. Thirdly, the control uncertainty\cite{14} caused by wheel slip, time-delay and other unexpected factors is commonly reported in the robot navigation domain.

![Figure 1. Uncertainty of robot localization and people-tracking](image-url)
2.2 PN-POMDP control framework

The navigation system uses the robot’s abstracted positional relation with the human as the states of the decision-making subsystem. The PN-POMDP system is modelled by a sextuple, \( <S, A, T, R, O, B> \). More specifically, \( S = \{s_1, \ldots, s_n\} \) is the agent state set and \( A = \{a_1, \ldots, a_l\} \) is the agent action set. \( T : S \times A \times S \rightarrow S \) is a state transition matrix, in which \( T(s', a, s) \) specifies the conditional probability distribution of transition from state \( s \) to \( s' \) by executing action \( a \). \( R(s, a) \) is the reward function that determines the immediate utility of executing action \( a \) at state \( s \). \( O(s', a, o) \) and \( O : S \times A \times O \) form the observation model that computes the probability of obtaining observation \( o \) in state \( s' \) when executing action \( a \). \( b(s) \) is the probability that agent is in state \( s \) with \( s_i \) as the initial state. The collection of all states formulates an agent’s belief space \( B \) of \(|S| \) dimension. The PN-POMDP solution is to seek an optimal navigational policy \( \pi \) to specify the corresponding action of the robot for each possible state, i.e., to find \( \pi^{\text{POMDP}} : b \rightarrow a \).

Finding optimal policies in the POMDP case is computationally intractable because the beliefs space is continuous and high-dimensional. The solution adopted in this work is a hybrid control structure that combines the reactive motion control and the probabilistic strategy selection for generating optimal navigational behaviours, as shown in Figure 2. In the system, sensory data obtained from laser, global cameras and other sensors are processed by the SLAP module and fed to the Perception module. Human motion patterns learned by the Modelling module are also inputted to the Perception module, in which the future motion tendencies of people are predicted in both their long-term and short-term aspects. Then, the human and robot motion states are abstracted and three types of abstracted observations are formatted, namely People’s Action Observation (PAO), People-Robot Relation Observation (PRO) and Robot State Observation (RSO). These abstracted observations are inputted to the Control module.

![Figure 2. PN-POMDP system framework](image)

In the Control module, the POMDP-based decision-making sub-module generates a suitable predictive navigational (PN) policy that minimizes the risk of conflict with humans. We have designed four types of action: detour, slow-down, speed-up, and halt, which will be explained in Section 5. In order to ensure goal-directed and predictive navigation performance, the motion controller is constructed with a two-layered architecture augmented by policies generated from the POMDP controller. The wavefront-based global path planner calculates the optimal path and the reference points along the path based on mapped obstacles. The Nearness Diagram[12] based local reactive obstacle avoidance controller computes actual translational velocity \( v \) and rotational velocity \( \omega \) based on the reference points and real-time sensory data.

A more detailed illustration of the POMDP controller structure is depicted in the right part of Figure 2. The POMDP controller contains a state estimator (SE) and a policy generator. The state estimator computes probabilistic distributions upon the belief \( b_1 \) according to \( o, a \) and \( b_{t-1} \). Meanwhile the policy generator maps the belief onto an optimal behaviour of the robot, i.e., \( a = \pi(\text{bel}) \).

3. Human motion prediction

3.1 Long-term modelling of motion pattern

Based on the collection of the tracking trajectories, a set of motion patterns of people are clustered hierarchically using a fuzzy K-means algorithm based on the spatial and temporal information. This algorithm results in a set
motion pattern \( \Psi_m \) is approximated by a mixture of \( K \) Gaussian distributions \( \Psi_m = \Psi_{m1}, ..., \Psi_m^K \). Spatial probability of the person located at \( h_t \), given step \( k \) of the motion pattern \( \Psi_m \) is computed according to the Gaussian distribution and denoted as \( p(h_t | \Psi_m^k) \):

\[
p(h_t | \Psi_m^k) = \frac{1}{\sqrt{2\pi \Sigma_m^k}} \exp \left( -\frac{\|h_t - \mu_m^k\|^2}{2\Sigma_m^k} \right),
\]

(1)

Examples of learned human motion patterns are shown in Figure 3, which indicates that people typically move between places with important objects to manipulate: fridge, a printer, a washing machine, etc.

3.2 Short-term motion prediction

To account for the short-term uncertainty of the movement along the path of \( \Psi_m \), the variation in velocity and heading orientation of the person are modelled.

We assume the following on the movement of a person \( \Psi_m \):

- \( \Delta T \) is the time step in which a person keeps to a certain velocity and heading direction; the possible ranges of his/her motion velocity and orientation are represented as \([v_{\min}, v_{\max}]\) and \([\theta_{\min}, \theta_{\max}]\), respectively; he/she changes the velocity and orientation only at every time step \( \Delta T \). In this sense, the velocity (orientation) of a time step is constant, and is randomly and independently selected within the above range. According to these assumptions, the sequence of velocity (orientation) along time contains a list of velocities (orientations) that are independently distributed random variables. This assumption describes a common indoor motion style where people move smoothly between two places.

The probability \( p(h_t | \Psi_m^k, \theta_{\min}, \theta_{\max}, \Delta T) \) can be decomposed according to the Bayes rule:

\[
p(h_t | \Psi_m^k, \theta_{\min}, \theta_{\max}, \Delta T) = \int p(h_t | \Psi_m^k, \theta_{\min}, \theta_{\max}, \Delta T) \, d\theta_{\min} \, d\theta_{\max}
\]

(2)

The probability \( p(h_t | \Psi_m^k, \theta_{\min}, \theta_{\max}, \Delta T) \) evaluates the probability of the person that covers the point \( h_t \) at time \( t \), given a sequence of observations \( \theta_{\min} \) and given that \( \theta_{\min} \) starts at \( \Psi_m^k \) and ends at \( \Psi_m^k \). The probability \( p(\Psi_m, k | \zeta_{t+1}) \) can be approximated by a Gaussian distribution of the form:

\[
p(\Psi_m, k | \zeta_{t+1}) = \eta p(\zeta_{t+1} | \Psi_m^k, k, \zeta_{t+1}) p(\Psi_m^k) p(k | \Psi_m^k)
\]

(3)

where \( \eta \) is a normalizer, \( p(\zeta_{t+1} | \Psi_m^k, k, \zeta_{t+1}) \) is the observation likelihood of \( \zeta_{t+1} \) and \( p(\Psi_m^k) \) and \( p(k | \Psi_m^k) \) are two prior probability distributions.
at the current position of the person, \( h_p = (x_p, y_p, \theta_p) \), and takes the Person’s Instantaneous Orientation (PIO) as the symmetry axis. The maximum angular distance from the PIO is \( \Lambda = \theta_{\text{max}} \) and is defined by the size of the field of view. Within the field of view area, a point \( h = (r, \alpha) \) has a probability of

\[
P_{\text{Ori}}(h | h_p) = \exp(-\alpha^2),
\]

to be the goal of movement in the next step. Eq. 4 indicates that the higher the value of \( \alpha \), the less likely it is that the person will head in that direction.

Secondly, the velocity variance is modelled by a distribution \( p_{\text{vel}}(h_i; t) \) that calculates the probability of reaching a point \( h_i \) along a straight line path, as demonstrated in Figure 5. Let \( \theta_{0i} \) and \( \sigma^2_{\text{vel}} \) be the current heading direction and the positional variance of the person, respectively. Then the position \( h_i = (x_i, y_i) \) after \( t \) time steps is given by

\[
x_i = x_0 + \sum_{i=1}^{t} v_i \Delta T \cos \theta_0, \quad (5)
\]
\[
y_i = y_0 + \sum_{i=1}^{t} v_i \Delta T \sin \theta_0. \quad (6)
\]

Since each \( v_i \) (\( i = 0, \ldots, t \)) follows the same but independent uniform distribution of \( v_i \sim \mathcal{U}(v_{\text{min}}, v_{\text{max}}) \), the variance \( \sigma^2_{\text{vel}} \) of the movement (i.e., the velocity variance) added by one time step is calculated as:

\[
\sigma^2_{\text{vel}} = \int_{-\infty}^{\infty} (v - \mu)^2 f(v) dv = \frac{1}{v_{\text{max}} - v_{\text{min}}} \int_{v_{\text{min}}}^{v_{\text{max}}} (v - \frac{v_{\text{max}} + v_{\text{min}}}{2})^2 dv.
\]

Let \( \mu = \mu_i \) and \( \sigma^2 = \sigma^2_{\text{vel}} \) be the mathemathical expectation and variance of \( v_i \), respectively. According to the central limit theorem, \( \sum_{i=1}^{t} v_i \) approximately follows \( N(\sum_{i=1}^{t} \mu_i, \sum_{i=1}^{t} \sigma_i^2) \), and so does \( x_i \). Thus the probability \( p_{\text{vel}}(h_i; t) \) is computed as:

\[
p_{\text{vel}}(h_i; t) = \frac{1}{2\pi \sigma_{x,i}\sigma_{y,i}} \exp \left\{ -\frac{1}{2} \left( \frac{(x_i - \bar{x}_i)^2}{\sigma_{x,i}^2} + \frac{(y_i - \bar{y}_i)^2}{\sigma_{y,i}^2} \right) \right\},
\]

where

\[
\bar{x}_i = x_0 + \sum_{i=1}^{t} \mu_i \Delta T \cos \theta_0, \quad (9)
\]
\[
\bar{y}_i = y_0 + \sum_{i=1}^{t} \mu_i \Delta T \sin \theta_0. \quad (10)
\]

Moreover, according to the assumption that \( v_i \) (\( i = 0, \ldots, t \)) follows the same but independent uniform distribution, the sequence of variables \( v_{i,1}, \ldots, v_{i,t} \) have the same mean and variance. This indicates that

\[
\mu_i = \bar{v} = \frac{(v_{\text{max}} + v_{\text{min}})}{2},
\]
\[
\sigma_{i}^2 = \sigma^2 = \frac{\sigma^2_{\text{vel}}}{t} + t \sigma^2_{\text{vel}}.
\]

To combine the long-term and short-term prediction, the heading orientation probability \( p_{\text{Ori}}(h | h_i) \) is used as the exponent discount factor to the velocity probability, and the probability of the motion pattern that the person is involved in at current position \( h_i \) is also normalized by a normalization factor \( \eta \) for all M motion patterns. Finally, the probability of reaching \( h_i \) at time \( t \) is computed as:

\[
p_{\text{predict}}(h_i; t) = \eta p_{\text{Ori}}(h_i | h) p_{\text{vel}}(h_i; t)^\gamma,
\]
\[
\gamma = p_{\text{Ori}}(h_i | h_i).
\]

4. Human-robot Co-occurrence Estimation

The probability of human and robot co-occurrence is estimated in the spatial-temporal aspect according to the robot’s travelling route obtained from the global path planner and the human motion prediction. In the POMDP system, situations of human-robot co-occurrence are classified into two types.

The first type is human-robot motion conflict, which is denoted as \( \theta = |\theta_0 - \theta_r| \), where \( \theta_0 \) and \( \theta_r \) are the orientation of robot and human, respectively. In wide open indoor areas, a situation that satisfies \( \theta_0 < \theta < \pi \) indicates that the robot’s movement along its planned path and the person’s future movement are likely to encounter both spatial and temporal factors. Let \( L_{\text{safe}} \) be a safe distance between person and robot, and \( P_s \) denote the intersection point of the robot’s path and the person’s predicted motion trajectory. At current time \( t_s \), the predicted human-robot motion conflict at a future time has to satisfy the condition that the person is now at a position within \( [P_c - D_{\text{safe}}^0, P_c + D_{\text{safe}}^0] \) (as shown in Figure 6), where

\[
D_{\text{safe}}^0 = \frac{L_{\text{safe}}}{\sin \theta} \sqrt{v_r^2 + v_h^2 - v_r v_h \cos \theta + \frac{v_h^2}{v_r}},
\]
This is because if the person and the robot are moving along their respective paths at constant speeds of $v_h$ and $v_r$, the robot will arrive at the place $P_c$ at time $t = t_0 + D_{\text{safe}}/v_h$, and then its distance to the person will be less than $L_{\text{safe}}$. If the motion uncertainty of the person is taken into account, the probability that he/she will arrive at place $P_c$ at a future time $t$ is computed according to Eq. 18:

$$P_{\text{conflict}} = P_{\text{predict}}(P_c, t),$$

where $t = (L_{\text{safe}}/\sin \theta)/v_r$. Moreover, in a special situation where $\theta = \pi$, which indicates that the robot is moving towards the person, it will be impossible to calculate variables $P_c$, $D_{\text{in}}$, and $D_{\text{out}}$. In this case, a face-to-face human-robot motion conflict is accounted.

The second situation is human-robot motion obstruction (as shown in Figure 7). This represents a situation where the robot’s path will block a human’s intentional trajectory, which happens to traverse a narrow passage. In this type of situation, a detected $\theta < \theta_{\text{th2}}$ event indicates that the paths of robot and human are generally parallel in opposite directions. To calculate possible spatial-temporal obstruction, the first and the last point along the robot’s planned path that satisfies the condition of Eq. 19 and Eq. 20 are computed:

$$P_{\text{pattern}}(s | \Psi_m) \leq x^k_m$$

$$k = \text{arg min}_{k=1,...,K} \{ P_{\text{pattern}}(s | \theta_{m_k}, \pi_{m_k}) \}$$

These two points are denoted as the InPoint $s_{\text{in}}$ and the OutPoint $s_{\text{out}}$ of their path intersection, respectively. Suppose that human and robot are moving along their corresponding path at constant speeds of $v_h$ and $v_r$, respectively. The person will then arrive at the points $s_{\text{in}}$ and $s_{\text{out}}$ at future moments $t_{\text{in}}^k$, $t_{\text{out}}^k$, respectively. Similarly, the robot will arrive at the points $s_{\text{in}}$ and $s_{\text{out}}$ at future moments $t_s^k$ and $t_r^k$, respectively.

If $|t_{\text{in}}^k - t_{\text{out}}^k| < \theta_{\text{th3}}$ and the motion uncertainty of the person is taken into account, the probability of human-robot motion conflict obstruction at point $s_{\text{in}}$ is computed as:

$$P_{\text{obstruct}} = \eta P_{\text{pattern}}(s_{\text{in}} | \Psi_m) P_{\text{vel}}(s_{\text{in}}; t_{\text{in}}^k, \gamma_{\text{vel}}) P_{\text{vel}}(s_{\text{out}}; t_{\text{out}}^k, \gamma_{\text{vel}})$$

$$\gamma_1 = P_{\text{orien}}(s_{\text{in}} | r_{l-1}), \quad \gamma_2 = P_{\text{orien}}(s_{\text{in}} | h_{l-1})$$

Similarly, if $|t_{\text{in}}^k - t_{\text{out}}^k| < \theta_{\text{th4}}$ and the motion uncertainty of the person is taken into account, the probability of human-robot motion conflict obstruction at point $s_{\text{out}}$ is computed as:

$$P_{\text{obstruct}} = \eta P_{\text{pattern}}(s_{\text{out}} | \Psi_m) P_{\text{vel}}(s_{\text{out}}; t_{\text{in}}^k, \gamma_{\text{vel}}) P_{\text{vel}}(s_{\text{out}}; t_{\text{out}}^k, \gamma_{\text{vel}})$$

$$\gamma_1 = P_{\text{orien}}(s_{\text{out}} | r_{l-1}), \gamma_2 = P_{\text{orien}}(s_{\text{out}} | r_{l-1})$$

Figure 6. Human-robot motion conflict

Figure 7. Human-robot motion obstruction

5. POMDP-based decision-making for polite avoidance

5.1 The Elements: States, Actions and Observations

To automatically compile the POMDP model $< S, A, T, R, O, B >$, it is necessary to define some action and observation uncertainties.
In this case, the states of the POMDP model are defined based on the abstraction of a set of variables.

- Human motion state, which takes values of move(PM) and stay(PS);
- Human motion tendency, which takes values of move(GM) and stay(GS);
- Human-robot distance, which is obtained by discretization of metric distance between human and robot, taking values of safe(DS), minor danger(DB), major danger(DJ) and high danger(DE);
- Robot motion state, which takes values of normal path following(RN), accelerating along path(RF), decelerating along path(RS), and dynamic replanning(RR).

A number of 64 states $s_1, \ldots, s_{64}$ can be composed; in analysis, 32 of these seldom appear, and the remaining 32 situations are selected as system states for the POMDP model.

Observations ($O$) in our model are the abstractions of the movements and positional relation between human and robot. In each state, the robot makes three kinds of observation:

- People's Action Observation (PAO), which can be abstracted as \{Move,Stay\}.
- People-robot Relation Observation (PRO), which can be classified as eight possible values \{PRO1,\ldots, PRO8\} according to the abstraction of relative position between human and robot.
- Robot State Observation (RSO), which can be abstracted as \{Path-following, Accelerating, Decelerating, Replanning\}.

Actions ($A$) are human-compliant avoidance behaviours that the robot can execute to give way to human in a polite manner:

- Normal path following $a_\text{nh}$;
- Accelerating along path $a_\text{uh}$;
- Decelerating along path $a_\text{sh}$;
- Dynamic replanning for detour $a_\text{rh}$.

The first three actions indicate that the robot follows the planned path with only velocity changes. These actions are usually more efficient for avoiding conflict with or obstruction of humans. The fourth action indicates that the robot replans a new path according to the updated environmental map by incorporating the probability distribution function (PDF) of $p(k; t)$ with the occupancy grid map.

The reward ($R$) defines the reward function that determines the immediate utility of executing action $a$ at state $s$. In our system, the reward matrix is manually specified based on a criterion that behaviours ensuring more safety and politeness receive higher rewards. Nevertheless, the optimum choice of settable parameters can be adjusted through a user-supervised learning system when a robot is installed in a new environment and performs a daily room exploration task, as suggested by Lopez [22].

5.2 POMDP Compilation

The transition model $T(s',a,s)$ specifies the conditional probability distribution of transitioning from state $s$ to $s'$ by executing action $a$. $O(s',a,o)$ is the observation model that computes the probability of obtaining observation $o$ in state $s'$ when executing action $a$.

Usually, the transition model and observation model can be rewritten as

$$T(s',a,s) = p(s'|s,a), \quad O(s',a,o) = P(o|s',a)$$

where $o \in O, s \in S, a \in A$.

![Figure 8](https://www.intechopen.com) Rules for constructing the POMDP model from empirical uncertainties

To automatically learn the observation model and the observation model, initial values are given (examples are shown in Figure 8) and the EM algorithm is employed for learning from collected data until the output parameter converges. The Randomized Point-based Value Iteration algorithm [18] is utilized to solve the above-defined POMDP model. In our system, 32 iterations and 63.578 seconds are required for offline model compiling. Figure 9 shows the errors during the iteration. The error between two successive iterations is plotted in the $y$-axis, with higher value indicating faster convergence rate.
The proposed approach was validated in a real office environment of size 12 m × 7 m. An ActiMedia Peoplebot was used in the experiments. We assumed that participants in the experiment walked at a smooth speed and intended to follow certain motion patterns. The sensory system for robot localization and people-tracking consists of five stationary CCD cameras mounted on each side of the room above head level and the robot’s onboard laser range finder. The environmental grid map was previously built by a SLAM algorithm with a resolution of 0.1 m. Based on the collection of tracking trajectories, typical indoor motion patterns of humans are learned, as presented in our previous work[23][24].

6.1 Predictive Navigation

During online predictive navigation, the robot collected laser scan data with a time period of 200 ms, and updated the local map gridlocal with a time period of 80 ms, according to the positions of detected human legs. In the PN-POMDP algorithm, $\theta_{th1}$, $\theta_{th2}$, $\theta_{th3}$ and $\theta_{th4}$ are the threshold values that can be set and adjusted in the experiment.

In the first three testing scenarios, the robot and the human were initially located in the same room area within a short distance of each other (less than 5 m). In the first case, where the robot was moving towards the human, the system predicted human motion 5 seconds ahead of time. Figure 10(a) shows that the robot began to avoid possible human-robot conflict using the detour policy when it was still about 3 m away from the human. In the second scenario (Figure 10(b)), when the robot had predicted that its path would intersect with the predicted human trajectory from one side, it selected the slow down action $a_s$. This policy was efficient because frequent replanning was avoided and the robot would continue to move along the path with normal speed when the human had passed the predicted intersection point. To test the reliability of the algorithm, in the third scenario (Figure 10(c)) the robot was following a person through a narrow corridor. In this case, the robot made the prediction that it would not interfere with the human’s motion if it did not overtake him/her. Thus the robot followed its planned path with regular speed.

The fourth testing scenario involved predictive navigation. The robot and the human were initially positioned in two different rooms and global cameras were utilized for people-tracking. In the experiment, the robot was initially positioned at place A as in Figure 11(a), and it planned to navigate through a narrow passage (passage II) to place B. In the meantime, a person intended to walk through the same passage in the opposite direction.

Before the robot approached the passage entrance, the probability of human-robot motion conflict at the entrance of the passage (place C in Figure 11 (b)) was estimated. More specifically, the tendency likelihood of the person’s temporary motion to be engaged in the motion pattern ending at place C was estimated at as high as 97.5%. However, since the predicted occupancy probability of the grid cells within the passage was
volatile, the traditional replanning method caused the robot to switch paths between two candidate routes (plan1, detour via passage I and plan2, continue via passage II). This method guided the robot to unnecessarily move back and forth at the passage entrance and finally reach the goal, requiring as long as 633 time periods. In comparison, the PN-POMDP method supporting multiple comity policies generated a highly efficient and human-friendly behaviour. When the human-robot conflict probability within the passage II was predicted, the robot consequently drove to a free space outside the entrance of the passage and waited. After the person had passed through the passage, the robot proceeded to cross the doorway and continued on its route. The resulting behaviour of the robot improved the navigation efficiency (only 342 time periods for reaching the goal) by avoiding unnecessary repeated zigzagging and wandering before entering the passage. The pose and translational velocity of the robot during the navigation test is shown in Figure 12. As shown in Figure 12(a), the replanning method caused the robot to switch between the two candidate paths during the interval (2) to (5). In contrast, Figure 12(b) shows that the PN-POMDP method ensures smooth and efficient robot navigation. More importantly, the polite navigation behaviour is comprehensible to humans and shows full respect to the human.

![Image](image1.png)

Figure 11. Comparison in predictive navigation scenario

(a) Replanning method result

(b) PN-POMDP method result
Figure 12. Robot pose and velocity change during the navigation scenario

6.2 Predictive navigation in highly populated areas

Figure 13 shows the experimental result in crowded environments with three participants walking around the robot. Since the PN-POMDP method supports multiple policies of predictive navigation, the robot frequently adjusted the policy according to the predicted human-robot co-occurrence situations. In fact, after raising the reward function of the deceleration action $a_s$, the robot tended to slow down for the human to pass first. This indicates that the PN-POMDP method is feasible to be applied to service robots that work in crowded environments such as exhibition halls and museums.

6.3 Trial study

A statistical trial study was also conducted to verify the success rate of the PN-POMDP method. We invited 12 participants (eight male and four female), ranging in age between 21 and 34. 33% of them were from non-technological fields, while 67% worked in technology-related areas. The trial tests involved different types of situation as described above. In the trial tests, the following situations were treated as “failure”: (i) The robot blocked the human’s intended route of movement (subjective scoring); (ii) The robot failed to reach the goal because of getting trapped or localization failure; (iii) The robot reached the goal with time consumption as high as four times that needed in situations without humans moving around. Figure 14 shows that the PN-POMDP method achieved a higher success rate than the traditional real-time replanning method.
7. Conclusion

In this paper, we have presented a predictive navigation method for service robots in the POMDP framework. By learning human motion patterns and combining long-term and short-term human motion prediction, space-time estimation of human-robot co-occurrence is achieved. In order to execute tasks in typical partially observable environments, POMDP-based probabilistic decision-making is incorporated to generate a theoretically optimal policy that allows the robot to behave in an efficient and polite manner. Thus the risk of conflict with human motion is minimized. The feasibility of the proposed methodology is validated by navigation experiments as well as user trials, in which the robot's navigational behaviour is interpreted by humans as safe, comprehensible and polite.

Although the system makes use of external cameras for human tracking, the proposed methodology framework does not rely on specific means for the acquisition of human motion. In situations where robots are not close to people, we suggest the utilization of global cameras to ensure seamless and reliable human-tracking, which improves the performance of predictive navigation.

8. Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant no. 61105094, No. 61075090, No.61005092 and No. 60805032) and the open fund of Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education (no. MCCSE2012B02).

9. References

[1] E.A. Siebot, L.F. Marin-Urias, R. Alami, T. Simeon, “A Human Aware Mobile Robot Motion Planner,” IEEE Transactions on Robotics, Vol. 23, pp. 874-883, 2007.
[2] P. Althaus, H. Ishiguro, K. Kanda, T. Miyashita and H.I. Christensen, “Navigation for human-robot interaction tasks,” Proceedings of IEEE International Conference on Robotics and Automation, New Orleans, LA, USA, pp. 1894-1900, 2004.
[3] R. Gockley, J. Forlizzi, R. Simmons, “Natural person-following behavior for social robots,” Proceedings of Human-Robot Interaction, pp.17-24, 2007.
[4] S. Y. Chung and H. P. Huang, “Predictive Navigation by Understanding Human Motion Patterns”, International Journal of Advanced Robotic Systems, Vol. 8, pp. 52-64, 2011.
[5] M. Bennewitz, W. Burgard, G. Cielniak and S. Thrun, “Learning motion patterns of people for compliant robot motion,” The International Journal of Robotics Research, Vol. 24, pp. 31-48, 2005.
[6] S. Osentoski, V. Manfredi, S. Mahadevan, “Learning Hierarchical Models of Activity,” Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan, pp. 891-896, 2004.
[7] J. Miura, Y. Shirai, “Modeling motion uncertainty of moving obstacles of robot motion planning,” Proceedings of the 2000 IEEE International Conference on Robotics and Automation, San Francisco, CA, USA, pp. 2258-2263, 2000.
[8] F. Hoeller, D. Schulz, M. Moors, F. E. Schneider, “Accompanying Persons with a Mobile Robot using Motion Prediction and Probabilistic Roadmaps,” Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, California, USA, pp.1260-1265, 2007.
[9] A. O. Djekoune, K. Achour and R. Toun, “A sensor based navigation algorithm for a mobile robot using the DVFF approach”, Int. J. of Advanced Robotic Systems, Vol. 6, pp.97-108, 2009.
[10] S. Koenig and M. Likhachev, “Fast replanning for navigation in unknown terrain,” IEEE Trans. on Robotics, Vol. 21, pp. 354-363, 2005.
[11] A. Poncea, C. Urdiales, E. J. Perez and F. Sandoval, “A new efficiency-weighted strategy for continuous human/robot cooperation in navigation,” IEEE Trans. on Systems, Man and Cybernetics Part A: Systems and Humans, Vol. 39, pp. 486-500, 2009.
[12] J. Minguez, L. Montano, “Sensor-based robot motion generation in unknown, dynamic and troublesome scenarios,” Robotics and Autonomous Systems, Vol. 52, pp. 290-311, 2005.
[13] W. Hu, M. Hu, X. Zhou, T. Tan, et al., “Principal Axis-Based Correspondence between Multiple Cameras for People Tracking,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, pp. 663-671, 2006.
[14] H. C. Yen, H. P. Huang and S. Y. Chung, “Goal-Directed Pedestrian Model for Long-Term Motion Prediction with Application to Robot Motion Planning,” Proc. IEEE Int. Conf. on Advanced Robotics and its Social Impacts, Taipei, Taiwan, pp. 1-6, 2008.
[15] M. E. Lopez, L. M. Bergasa, R. Barea, M. S. Escudero, “A Navigation System for Assistant Robots Using Visually Augmented POMDPs,” Autonomous Robots, Vol. 19, pp. 67-87, 2005.
[16] Foka A, Trahansis P, “Real-Time Hierarchical POMDPs for Robot Navigation,” Robotics and Autonomous Systems, Vol. 55, pp. 561-571, 2007.
[17] J. Pineau, G. Gordon, S. Thrun, “Anytime point-based approximations for large POMDPs,” Journal of Artificial Intelligence Research, Vol. 27, pp.335-380, 2006.
[18] M. Spaan, N. Vlassis, “Perseus: Randomized Point-based Value Iteration for POMDPs,” Journal of Artificial Intelligence Research, Vol. 24, pp. 195-220, 2005.
[19] N. Roy, G. Baltus, D. Fox, F. Gemperle, J. Goetz, T. Hirsch, D. Magaritis, M. Montemerlo, J. Pineau, J.
Schulte and S. Thrun, “Towards personal service robots for the elderly,” Proc. Of the Workshop on Interactive Robotics and Entertainment (WIRE), Pittsburgh, PA, 2000.

[20] M. Montemerlo, J. Pineau, N. Roy, S. Thrun and V. Verma, “Experiences with a mobile robotic guide for the elderly,” Proc. of the AAAI National Conference on Artificial Intelligence. Edmonton, Canada, pp. 587-592, 2002.

[21] K. Qian, X. D. Ma and X. Z. Dai, “Simultaneous Robot Localization and Person Tracking Using Rao-Blackwellised Particle Filters With Multi-modal Sensors,” Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France, pp. 3452-3457, 2008.

[22] M. E. Lopez, R. Barea, L. M. Bergasa, M. S. Escudero, “A Human–Robot Cooperative Learning System for Easy Installation of Assistant Robots in New Working Environments,” Journal of Intelligent and Robotic Systems, Vol. 40, pp. 233-265, 2004.

[23] K. Qian, X. D. Ma and X. Z. Dai and F. Fang, “Socially Acceptable Pre-collision Safety Strategies for Human-Compliant Navigation of Service Robots,” Advanced Robotics, Vol. 24, pp. 1813-1840, 2010.

[24] K. Qian, X. D. Ma and X. Z. Dai and F. Fang, “Robotic Etiquette: Socially Acceptable Navigation of Service Robots with Human Motion Pattern Learning and Prediction,” Journal of Bionic Engineering, Vol. 7, pp. 150-160, 2010.