Interacting multiple model-based human motion prediction for motion planning of companion robots*

Donghan Lee[1], Chang Liu[2] and J. Karl Hedrick[3]

Abstract— **Motion planning of human-companion robots is a challenging problem and its solution has numerous applications. This paper proposes a motion planning method for human-companion robots to accompany humans in a socially desirable manner, which takes into account the safety, comfortableness and naturalness requirements. A unified Interacting Multiple Model (IMM) framework is developed to estimate human motion states from noisy sensor data and predict human positions in a finite horizon. The robot motion planning is formulated as a model predictive control problem to generate socially desirable motion behavior based on the predicted human positions. The effectiveness of the proposed motion planning method in facilitating the socially desirable companion behavior is evaluated through simulations and the advantage of IMM framework for human motion estimation and prediction compared to traditional single-model approaches has been demonstrated.**

[Note:] add more plots for explanation; add comparison of single-model, IMM and no-prediction (maybe) results

[Note:] dh:in above note, does no-prediction means reactive method?

I. INTRODUCTION

The application of autonomous robots for search and rescue missions have received considerable attentions in last decades. One particular scenario that is interesting to us is allowing autonomous robots to accompany humans in the search and rescue mission. Robot companions can follow and assist humans by carrying heavy apparatus, exploring dangerous areas or detecting signals of survivors.

To enable socially desirable companion behavior, certain requirements need to be satisfied, including safety, comfort, naturalness and sociability [1]. Among these requirements, safety serves as the fundamental guideline, requiring that robots avoid collision with accompanied humans under all circumstances. Safety is usually enforced by preventing robots from entering a 'safe region' around the target humans [**?**], [**?**], [**?**], [**?**]. (**TODO:** use lit from the survey paper. explain many aspects of social navigation. However, in problem formulation, just list what is done in this paper: focus on comfortableness (distance using proxemics and velocity similarity)) Comfort composes a social requirement on the robot behavior, which requires robots to pose little annoyance and stress for the accompanied humans [1]. It is mostly formulated regarding the distance that a robot needs to keep from persons. For example, [**?**], [**?**] proposed the concept of "Proxemics" to denote the virtual zone around a person that a robot should avoiding entering in order to prevent the discomfort the person may feel. (**TODO:** rephrase the following sentence:) In [**?**] extract robot design principles for approaching distance and gaze depending on social context.

(**TODO:** move the following sentence to the method review section) By incorporating these requirements into account, researchers have developed several human-companion robots. In [**?**], a generalized framework for representing social conventions as components of a constrained optimization problem was presented and it was used for path planning and navigation.

Cosgun et al. [8] develops a robot for telepresence usage, which uses the laser scanner to track human movement and predicts human path by extrapolating based on the estimated human position and speed, assuming a constant velocity human motion model. The robot utilizes the depth-limited breadth-first search approach to find the waypoints around the predicted human positions so that the robot can always face the accompanied human with a similar velocity as the human.

[Note:] dh: do we need to explain more detail about the depth-limited breadth-first search?

[Note:] ch: I think it's fine to just mention the name of depth-limited breadth-first search since it's a basic technique for people in robotics community.

Hoeller et al. [9] presents a local path planning approach for the companion robot to stay in proximity to the target person and simultaneously prevent colliding with any passersby. The probabilistic roadmaps is used to plan collision-free paths to a given target location, chosen based on the predicted adjacent humans' positions. A laser-based people tracking technique is used to estimate the motions of humans and a potential field method is applied for predicting the humans future trajectories.

An effective motion planner that generates socially desirable motion behavior necessitates accurate human motion estimation and prediction, which is then utilized for robot motion planning. To be specific, a robot needs to estimate motion states of the accompanied human in real time based on measurements from equipments such as GPS or cameras and then predict human future trajectory. There exist several

*This work was not supported by any organization

[1]Donghan Lee with the Vehicle Dynamics & Control Lab, Department of Mechanical Engineering, University of California at Berkeley, California 94720, USA donghan.lee@berkeley.edu

[2]Chang Liu with the Vehicle Dynamics & Control Lab, Department of Mechanical Engineering, University of California at Berkeley, California 94720, USA changliu@berkeley.edu

[3]J. Karl Hedrick is with Faculty of Mechanical Engineering, University of California at Berkeley, California 94720, USA khedrick@berkeley.edu

widely-adopted estimation methods, such as the Kalman filter [2], which provides the minimum mean square error estimation of states for linear systems with Gaussian noise. (**TODO:** find references for its application to motion estimation and prediction) The Extended Kalman filter (EKF) and Unscented Kalman filter (UKF) [12], which are extensions of the KF for nonlinear systems, have also been applied for motion estimation and prediction [**?**]. EKF can be obtained by applying the KF to the first-order Taylor expansion of nonlinear systems while UKF uses a deterministic sampling technique, called the Unscented Transform, to perform a stochastic linearization through the use of a weighted statistical linear regression process [3]. (**TODO:** if the references for motion estimation and prediction using xKF methods can be found, then remove the following part.) Based on the estimated states, extrapolation assuming known human motion models, such as the constant velocity model (**TODO:** lit review)

[Note:] dh: constant velocity and constant turn model

, can be applied for predicting human positions in a finite horizon.

[Note:] try to merge the following sentence with the previous one.

These methods usually assumes a single motion model for estimation and prediction and have been widely applied in motion planning and localization in the field of robotics [3].

In recent years, machine learning techniques have been utilized for human motion prediction and estimation. In [6], Wu et al. specifically focuses on the human motion estimation and prediction for robot motion planning. It develops an effective (**TODO:** describe what the filter actually is) filter for human motion estimation that can deal with abrupt human direction change when he/she makes turns. (**TODO:** more references on human motion estimation and prediction.)

These approaches have achieved success in human motion prediction. However, they usually involves the collection of large training data (human trajectory) and are scenario specific. In fact, human motion are well predicted in the scenario in which much human trajectories are used for training. This renders such methods less applicable for search and rescue missions since disaster sites are rarely similar to each other and therefore robots need to work in unfamiliar environments. In fact, human motion usually involves different identified motion models, such as straight-line movement, making turns and change of speed. Therefore, we use IMM approach, a model-based prediction approach that takes into several system dynamics into account and dynamically adjusts the mode probabilities, for human motion estimation and prediction.

To deal with the nonlinear dynamics of the human motion (such as making turns), Unscented Kalman filter (UKF) is applied to each model in the IMM framework, resulting in the so-called IMM-UKF approach. Such approach can achieve higher accuracy compared to traditional estimation and prediction methods that utilize a single motion model. Additionally, no training is needed and thus scenario-independent, which is advantageous over machine learning methods for search and rescue missions. Utilizing the predicted human trajectory, a model predictive control (MPC)-based motion planner is developed for the companion robot to autonomously accompany the target person in a socially desirable manner, which takes into consideration the safety, comfortableness and naturalness requirements. The proposed motion planning method differs from previous works in several aspects. First, different from the single-model based estimation and prediction approaches used in [**?**], [**?**], [8], the IMM-UKF method incorporates different human movement models into account, which results in higher estimation and prediction accuracy, especially when humans motion involves complex patterns. Second, the model-based IMM-UKF approach does not require training data and is scenario-independent, which is advantageous over the machine learning methods, such as the ANN method in [6], [**?**], [**?**]. (**TODO:** emphasize on the benefit of MPC) Additionally, MPC-based planner solves a nonlinear program for the optimal trajectory over finite horizon. Motion requirements such as the safety and comfort can be easily incorporated in the constraints of the nonlinear program.

[Note:] need to find more ref about motion planning methods for navigating in human environments

This is different from potential field based methods, such as in [9], which needs careful design of potential functions to enforce all these requirements.

The remainder of this paper is organized as follows: first, the problem of motion planning of a human-companion robot is formulated for a search and rescue mission; second, the IMM-UKF estimation and prediction method is described, followed by the MPC planner; next, simulation setup and results on evaluating the proposed approach are presented; lastly, concluding remarks and ideas of future work are presented.

II. PROBLEM FORMULATION

Consider

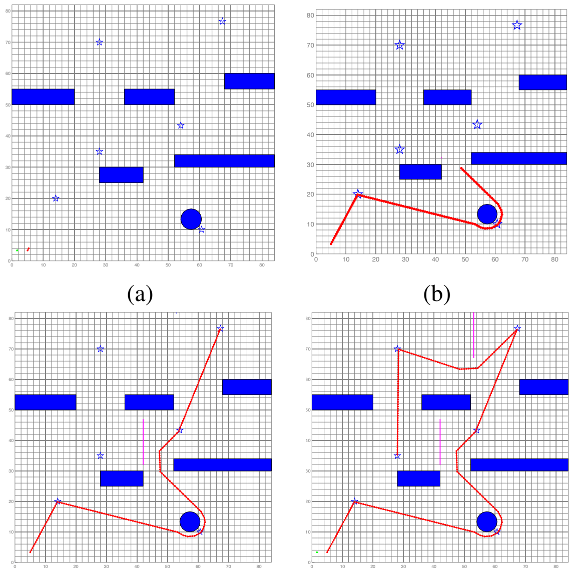
[Note:] dh: Considering?

a scenario (**??**) in which a companion robot, denoted as the red dot, needs to accompany a human, represented by the green triangle, who will move sequentially to several destinations, shown as blue hollow stars, in the field. The robot has no knowledge about the positions of destinations. However, it can measure human positions in real-time from a GPS sensor via cellphone communication.

[Note:] dh: sensors like GPS or Camera

There exist several obstacles in the field, the positions of which are known and are shown as the blue circle and semicircles.

[Note:] dh: I guess we need to explain more detailed about obstacles in the sceniario to show the effort that mimic the real environment. for example, we can say that there are two moving and six stationay obstacles in the scenario



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

Fig. 1: (a) A scenario for the companion robot to accompany a human. The red circle and the green triangle represent the human and the robot, respectively. Stars denote human destinations and the blue circle and semicircles stand for obstacles. (b) An obstacle starts moving. (c) Another obstacle starts moving. (d) The full human trajectory.

[Note:] dh: we need to show more clear plots, especially lines for the moving obstacles. And I guess it had better to use '1st','2nd' instead of 'an','another'

To make socially desirable behavior, the robot should satisfy the aforementioned safety, comfortableness and naturalness requirements(**TODO:** add formal definitions for these concepts). In this scenario, the safety concern requires the robot to avoid colliding with either the human or any obstacles. Comfortableness requires that the robot maintain a proper distance from the human and naturalness necessitates that the robot keep the velocity similar to that of the human. (**TODO:** make it clear that how these requirements are instantiated in this paper.)

III. METHODS

A. Human Motion Estimation

The Interacting Multiple Model (IMM) approach are applied for estimating the human motion from the noisy sensor data. The IMM approach has been generally considered as the mainstream method for maneuvering target estimation. It utilizes a bank of r number of filters, each designed to model a different dynamics. In the IMM estimator, state estimate at time k is computed with mode probability and state estimates from each possible current model associated with one of the r filters, using the following formula:

$$\hat{x}(k|k) = \sum_{j=1}^{r} \mu_j(k)\hat{x}^j(k|k)$$

where r denotes the number of models; $\hat{x}^j(k|k)$ represents the state estimate from the j^{th} filter; $\mu_j(k)$ stands for the mode probability and is computed as follows:

$$\mu_j(k) = \frac{1}{c}\sum_{i=1}^{r} L_{ij}(k)p_{ij}\mu_j(k-1)$$

where c denotes the normalizing factor; $L_{ij}(k)$ stands for the likelihood function and p_{ij} represents the mode transition probability from the i^{th} to the j^{th} model. Each filter uses the mixed initial state estimate and covariance from different combination of the previous model. Readers interested in the details of the IMM approach can refer to [10].

Two different dynamic models are used in the IMM framework. One is the coordinated turn motion model and the other is the uniform motion model. The equation of motion for the both the coordinated turn motion and uniform motion model becomes the following

$$x^h(k+1) = f(x^h(k)) + Gw(k) \tag{1a}$$

$$f(x^h(k)) = \begin{bmatrix} p_1^h + \frac{\sin(\omega^h T)}{\omega^h}v_1^h - \frac{1-\cos(\omega^h T)}{\omega^h}v_2^h \\ \cos(\omega^h T)v_1^h - \sin(\omega^h T)v_2^h \\ p_2^h + \frac{1-\cos(\omega^h T)}{\omega^h}v_1^h + \frac{\sin(\omega^h T)}{\omega^h}v_2^h \\ \sin(\omega^h T)v_1^h + \cos(\omega^h T)v_2^h \\ \omega^h \end{bmatrix} \tag{1b}$$

$$G = \begin{bmatrix} \frac{T^2}{2} & 0 & 0 \\ T & 0 & 0 \\ 0 & \frac{T^2}{2} & 0 \\ 0 & T & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1c}$$

$$w \sim \mathcal{N}(0, Q) \tag{1d}$$

where $x^h(k)$ represents the human motion state including five elements : $p_1^h, v_1^h, p_2^h, v_2^h, \omega^h$, where p_1^h, p_2^h denote the longitudinal and lateral position of the human, v_1^h, v_2^h the corresponding velocity and ω^h the turn rate of the human; $w(k)$ represents process noise; T represents the sampling time; Q is the covariance matrix of the process noise.
The uniform motion model is nothing more than the limiting case of the coordinated turn motion model as turn rate, ω, gose to zero while the coordinated turn motion model consider that turn rate is modeled as time-varying. If all motions are covered by a single model, the need for the IMM is eliminated. Although computations are reduced by using a single model, it happens to lose the ability to quickly detect the change of motions. This trade-off between computation and motion-change detection is main argument when choosing a single or a multiple model approach. The decision is a bit clearer in linear case since the turn rate is fixed in linear dynamic models. With a single nonlinear model, it is possible to provide accurate state estimates. However, it is common to include one uniform motion model and one coordinated turn motion model for quick motion-change detection. Since human motion usually involves different motions in the real world, the ability to quickly dectect the change of the motion is one of the important properties for the estimation and prediction.

The equation of the uniform motion model is shown as follows:

$$\dot{x}^h(k+1) = Ax^h(k) + Bw(k) \quad (2a)$$

$$A = \begin{bmatrix} 1 & T & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & T & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \; B = \begin{bmatrix} \frac{T^2}{2} & 0 & 0 \\ T & 0 & 0 \\ 0 & \frac{T^2}{2} & 0 \\ 0 & T & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2b)$$

The observation equation is represented as :

$$y^h(k) = Cx^h(k) + v(k) \quad (3a)$$

where $y^h(k)$ denotes the human motion observation at the time step k; $v(k)$ stands for measurement noise

We assume that only the human position can be measured. Therefore, the parameters in observation model Eq. (3a) can be defined as:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad (4a)$$

$$v \sim \mathcal{N}(0, V) \quad (4b)$$

where V is the convariance matrix of the measurement noise. The above two models are used for human motion state estimation using the Unscented Kalman Filter.

B. Unscented Kalman Filter

The Unscented Kalman Filter (UKF) is an effective state estimation technique for nonlinear systems by implementing the Unscented Transformation (UT) to calculate the statistics of a random vector that undergoes a nonlinear transformation [12] Given an L-dimensional Gaussian Random Vector (GRV) x with mean $\hat{x}$ and covariance P_x, the statistics of $z = f(x)$ are approximated by the selection of $2L+1$ discrete sample points $\left\{\chi^{(i)}\right\}_{i=0}^{2L} = \{\hat{x} \text{ and } \hat{x} \pm \sigma_j, j = 1, ..., L\}$ where σ_i is the i^{th} column of the matrix $\sqrt{(L+\lambda)P_x}$. λ is a scaling parameter, defined below.

$$\lambda = \alpha^2(L+\kappa) - L \quad (5a)$$

$$W_0^{(m)} = \frac{\lambda}{L+\lambda} \quad (5b)$$

$$W_0^{(c)} = \frac{\lambda}{L+\lambda} + 1 - \alpha^2 + \beta \quad (5c)$$

$$W_i^{(m)} = W_i^{(c)} = \frac{1}{L+\lambda}, \quad i = 1, ..., 2L \quad (5d)$$

where α determines the spread of sigma points about the mean $\hat{x}$; κ is a secondary scaling parameter; β is used to incorporate prior knowledge of the distribution. (our simulation uses $L = 5, \alpha = 0.001, \kappa = 0, \beta = 2$)

Once the discrete sample points $\left\{\chi^{(i)}\right\}_{i=0}^{2L}$, called *sigma points*, have been generated, each point is passed through the nonlinear function $z = f(x)$, i.e. each column of the sigma points is propagated through the non-linearity, as in $\zeta = f(\chi), i = 0, ..., 2L$. The mean $\hat{z}$ and the covariance P_z are approximated as $\hat{z} \simeq \Sigma_{i=0}^{2L} W_i^{(m)} \zeta^{(i)}$ and $P_z \simeq \Sigma_{i=0}^{2L} W_i^{(c)} (\zeta^{(i)} - \hat{z})(\zeta^{(i)} - \hat{z})^{\mathsf{T}}$, are calculated as given in above equations of the weights and parameters [13]. Readers can refer to [12] for more details of UKF algorithm.

C. Human Motion Prediction

The estimated human motion states and the mode probabilities are utilized for predicting human future states. Let $\hat{x}^{h,j}(k|k)$ and $\tilde{x}^{h,j}(k+i|k)$ represent the estimated and predicted human states associated with the j^{th} model at time k and $k+i\,(i \geq 0)$, respectively, based on the observation up to time k. Using the uniform motion model and the turn motion model, human positions for each model can be extrapolated and combined with the mode probabilities. To be specific, the prediction procedure works as follows:

$$\tilde{x}^h(k+l+1|k) = \sum_{j=1}^{r} \mu_j \tilde{x}^{h,j}(k+l+1|k) \quad (6a)$$

$$l = 0, \ldots, N-1 \quad (6b)$$

$$\tilde{x}^{h,j}(k+l+1|k) = \hat{x}^{h,j}(k+l+1|k) \quad j = 1, \ldots, r \quad (6c)$$

$$= \Sigma_{i=0}^{2L} W_i^{(m)} \chi_{k+l+1|k}^{(i)} \quad (6d)$$

$$\chi_{k+l+1|k}^{(i)} = f(\chi_{k+l|k}^{(i)}) \quad i = 0, \ldots, 2L \quad (6e)$$

where N denotes the prediction horizon; r represents the number of models; L is the dimension of $x^{h,j}$. For the purpose of simplicity, we define $p^h(k) = \begin{bmatrix} p_1^h(k) \\ p_2^h(k) \end{bmatrix}$ and $v^h(k) = \left\| \begin{bmatrix} v_1^h(k) \\ v_2^h(k) \end{bmatrix} \right\|_2$ to represent the position vector and the speed of the human at time k, respectively. Notations for the estimated and predicted position vector and speed can be defined as $\hat{p}^h(k|k), \tilde{p}^h(k+i|k), \hat{v}^h(k|k), \tilde{v}^h(k+i|k)$, respectively.

D. Robot Path Planning

The model predictive control (MPC) method provides an effective framework for incorporating the safety, comfortableness and naturalness requirements into the robot motion planning. MPC iteratively solves a finite time constrained optimal control problem. After obtaining the optimal series of control inputs at current state, it implements the first input and then computes for a new series of control inputs, starting from the new state. Use $(p^r(k), v^r(k), \theta^r(k))$ to denote the robot state at time k, representing the position, speed and heading angle, respectively. The control input consists of the acceleration $a^r(k)$ and the angular velocity $\omega^r(k)$. The robot motion planning problem is formulated as an MPC problem that incorporates the kinematics of the robot and the aforementioned requirements:

$$\min_{\mathbf{A}_k,\mathbf{\Theta}_k} \quad \sum_{i=1}^{N} q_1||\bar{p}^r(k+i|k)-\tilde{p}^h(k+i|k)||_2^2 + q_2|\bar{v}^r(k+i|k)-\tilde{v}^h(k+i|k)|^2 \tag{7a}$$

subject to

$$\bar{p}^r(k+i+1|k) = \bar{p}^r(k+i|k) + \bar{v}^r(k+i|k)\begin{bmatrix}\cos\bar{\theta}(k+i|k)\\ \sin\bar{\theta}(k+i|k)\end{bmatrix}T \tag{7b}$$

$$\bar{v}^r(k+i+1|k) = \bar{v}^r(k+i|k) + \bar{a}(k+i|k)T \tag{7c}$$

$$\bar{\theta}^r(k+i+1|k) = \bar{\theta}^r(k+i|k) + \bar{w}(k+i|k)T \tag{7d}$$

$$a_{lb} \leq \bar{a^r}(k+i|k) \leq a_{ub} \tag{7e}$$

$$w_{lb} \leq \bar{\omega^r}(k+i|k) \leq w_{ub} \tag{7f}$$

$$||\bar{p}^r(k+i+1|k)-\tilde{p}^h(k+i+1|k)||_2 \geq d_s \tag{7g}$$

$$||\bar{p}^r(k+i+1|k)-p_l^{obs}||_2 \geq r_l^{obs} \tag{7h}$$

$$||\lambda\bar{p}^r(k+i|k)+(1-\lambda)\bar{p}^r(k+i+1|k)-p_l^{obs}||_2 \geq r_l^{obs},$$

$$\forall l=1,\ldots,m,\ 0\leq\lambda\leq 1 \tag{7i}$$

$$\bar{p}^r(k|k) = p^r(k) \tag{7j}$$

$$\bar{v}^r(k|k) = v^r(k) \tag{7k}$$

$$\bar{\theta}^r(k|k) = \theta^r(k) \tag{7l}$$

where $\bar{p}^r(k+i|k)$, $\bar{v}^r(k+i|k)$ and $\bar{\theta}(k+i|k)$, $0\leq i\leq N$ represent the planned positions, velocities and heading angles of the robot at time $k+i$, respectively; m is the number of obstacles and p_l^{obs} and r_l^{obs} denote the position and the radius of the l^{th} obstacle; $(\mathbf{A}_k, \mathbf{\Theta}_k)$ stand for the set of optimal acceleration and angular velocity in the prediction horizon $[k, k+N-1]$, obtained by solving the MPC problem at time k.

The objective function Eq. (7a) consists of two terms, standing for the square sum of position and velocity differences between planned robot states and predicted human states over the horizon N. This reflects the comfortableness and naturalness requirements that the robot stay close to the accompanied human and keep similar pace. q_1 and q_2 denote the weights for these two terms. Eqs. (7b) to (7f) represent the roboot's discrete-time kinematic model with limited acceleration and angular velocity, with a_{lb}, w_{lb} being the corresponding lower bounds and a_{ub}, w_{ub} the upper bounds. Eq. (7g) imposes the safety constraints that the robot should maintain a minimum distance d_s from the human in order to avoid collision. Eqs. (7h) and (7i) enforces the requirements on collision avoidance with obstacles. In particular, Eq. (7h) demands that each way point of the robot be kept outside of obstacles and Eq. (7i) requires that the trajectory connecting the adjacent way points not intersect with obstacles. Eqs. (7j) to (7l) shows the constraints on robot's initial state at time k.

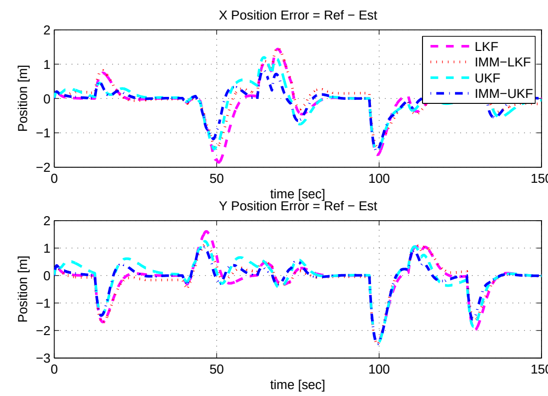


Fig. 2: Comparison of position estimation error with LKF, IMM-LKF, UKF and IMM-UKF

IV. SIMULATION RESULTS & DISCUSSION

A. Simulation setup

Simulations have been run to evaluate the proposed robot motion planning approach. There are one human and one robot with five targets and four obstacles in a $84m \times 82m$ field. The human speed in the simulation is set to be constant at $1.5m/s$. The human will follow the trajectory shown in Fig. 1d. However, this trajectory is unrevealed to the robot. The safety distance d_s is chosen as $2m$. The sampling rate of GPS sensor is $20Hz$ and the variance of sensor measurement noise is considered as $2m$. The robot's maximum acceleration and deceleration are set to be $1m/s^2$ and $-3m/s^2$ respectively and the angular velocity range is chosen to be $[-30°/s, 30°/s]$. In the IMM estimator, the process noise and the measurement noise are set to be 1.5×10^{-2} and 1.5, respectively. The prediction horizon for the human motion is chosen as $500ms$ and the robot recomputes the MPC problem every $500ms$.

B. Simulation results

We evaluate the performance of human motion estimation and prediction and robot motion planning methods, respectively.

1) Human motion estimation: The error between the estimated and the actual human position and velocity at each time step are compared to evaluate the estimation accuracy. The position error vector can be formulated as:

$$\Delta_p^t(k) = p^h(k) - \hat{p}^h(k|k)$$

where $p^h(k)$ denotes the actual human position at time k.

Fig. 2 shows the position estimation error on longitudinal and lateral directions using four different estimators: Linear Kalman Filter (LKF), IMM-LKF, UKF and IMM-UKF. In the simulation, the same turn motion model in IMM-UKF is adopted for the system dynamics in UKF and the uniform motion model in IMM-LKF is also applied to the dynamic model in LKF. There are two observations in this figure. First, the responses of the nonlinear estimators such as UKF and IMM-UKF are faster than the linear

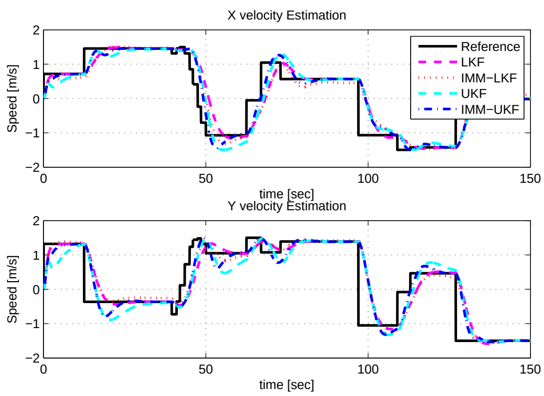


Fig. 3: Comparison of the estimated velocity using the IMM-based and the single-model approaches

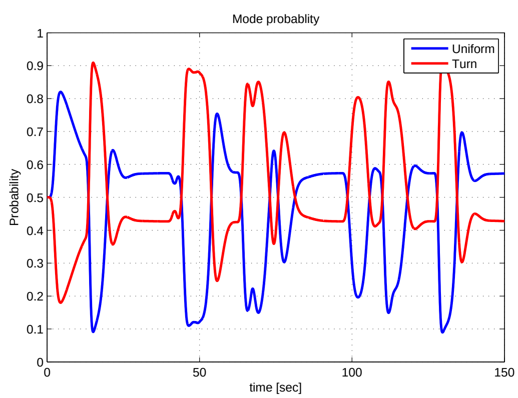


Fig. 4: Model probabilities of two models in the IMM-UKF estimator

estimators. Second, the IMM-based approaches show better performance in accuracy than the single-model approaches. Besided, IMM-UKF achieves the fastest response and best accuracy compared to other methods, especially when the human turns around the circular obstacle at time 50 and makes sharp turn after arriving at a destination at time 110. Fig. 3 compares the velocity estimation using four estimators. Overall, the nonlinear estimators (UKF and IMM-UKF) show faster response compared to the linear estimators (LKF and IMM-LKF), though they have overshoots due to the fast response. It is worth noting that the overshoots of IMM-UKF are smaller than UKF while keeping the fast response at time 53 and 118 when the velocity changes abruptly. This makes sense as the UKF-IMM estimator incorporates uniform motion model that can capture the sudden velocity changes. Fig. 4 shows the mode probabilities of the uniform motion model and turn motion model in IMM-UKF over time. When the human speed changes, the mode probability of the turn motion becomes higher than that of the uniform motion. These changes illustrates the reason that IMM-based estimators achieve more accurate and faster estimation than the single-model estimator at the sharp turn and circular

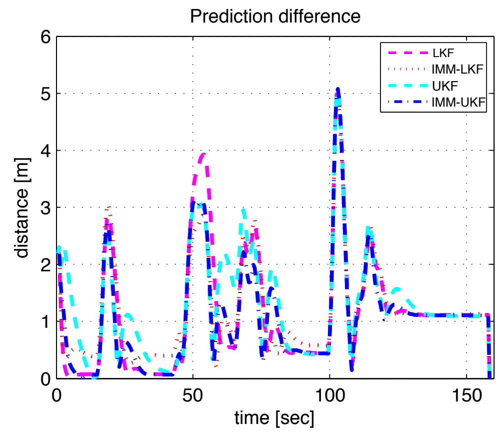


Fig. 5: Comparison of prediction error between the IMM-based and single-model approaches

turn, thus demonstrating the necessity of applying IMM-UKF estimator for human tracking.

2) Human motion prediction: To evaluate the IMM-based prediction approaches, the average prediction error over the prediction horizon is computed and compared with the single-model approaches. At time k, the prediction error is defined as:

$$\Delta_p(k) = \frac{1}{N}\sum_{i=1}^{N} ||\tilde{p}^h(k+i|k) - p^h(k+i)||_2 \tag{8}$$

Different from the IMM-based prediction approaches that extrapolate the human position by a weighted sum of the predicted positions from each model, the single-model methods only utilize the uniform motion model for prediction.

Fig. 5 shows the comparison of prediction error using two single-model approaches (LKF, UKF) and two IMM-based methods (IMM-LKF, IMM-UKF). It can be noticed that single-model approaches generate larger prediction error than IMM-based methods, especially when the human makes turns, such as at time 50. This makes sense as IMM-based method considers different dynamic models related to the human motion. Based on the simulation results, IMM-UKF outperforms the other three methods for prediction.

3) Robot motion planning: Fig. 6 shows a screenshot of the simulation that the robot accompanies the target person moving in the field. The performance of the robot motion planning is evaluated using the criterion of safety, comfortableness and naturalness. To be specific, we measure the distance and speed difference between the robot and the human at each time step. At time k, they can be defined as:

$$\Delta_d(k) = ||z^r(k) - z^h(k)||_2 - d_s \tag{9a}$$

$$\Delta_v(k) = |v^r(k) - v^h(k)| \tag{9b}$$

To evaluate the accompanying performance using single-model predictor and multiple-model predictor, the UKF and IMM-UKF are used in conjunction with MPC to generate two sets of simulation, the results of which are compared.

Fig. 7 compares the distances between the human and the robot using these two predictors. Notice that the distances

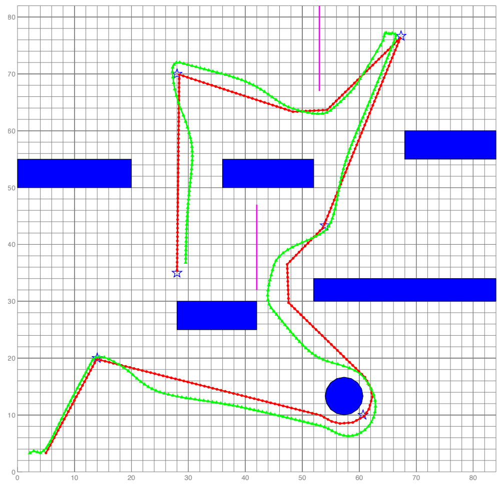


Fig. 6: A screenshot of the simulation. The red line represents the human trajectory and the green on shows the companion robot's trajectory. For most of the time, the robot follows the human from behind

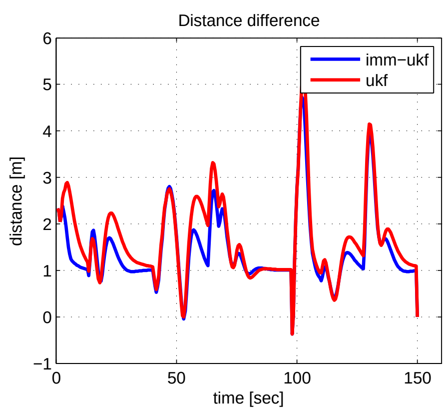


Fig. 7: Comparison of distance (substracted by the safety distance) between the human and the robot using the MPC and the reactive methods

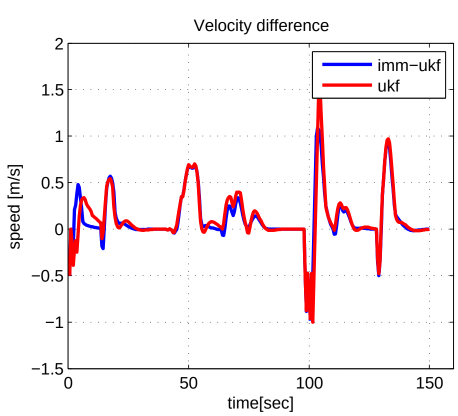


Fig. 8: Comparison of velocity difference between the human and the robot using the MPC and the reactive methods

in the plot has been substracted by the safety distance d_s. Therefore, the ideal distance is zero at each time step. We can notice that the IMM-UKF predictor generates smaller distance difference than the UKF predictor. **Table xxx shows the statistics of the performance. The IMM-UKF method achieves xxx% improvement on average distance, xxx% on maximum distance and also minimum distance. In addition, the robot using IMM-UKF as the predictor can achieve less unsafe distance than the UKF method. Such improvement is desirable for an accompanying robot as it keeps proper distance from the human and only moves within the unsafe distance for xxx long. Such unsafty is caused when the human makes sharp turns and the prediction deviates from the actual positions, which has been discussed in Sections IV-B.1 and IV-B.2. The UKF and IMM-UKF predictor result in similar velocity behavior, as shown in Fig. 8. The statistics also shows smaller velocity difference using the IMM-UKF predictor**.

The simulation results shows the superiority of using the MPC approach for motion planning.

V. CONCLUSION

We have developed a model predictve control (MPC)-based motion planning approach for human-companion robots to accompany a target person in a socially desirable manner, which considers the safety, comfortableness and naturalness. The IMM-UKF approach that incorporates the uniform motion model and the coordinated turn motion models is developed for human position estimation and prediction. Based on the predicted human positions, an MPC problem is formulated for robot motion planning. The estimation and prediction accuracy using IMM-UKF is compared with two single-model methods (LFK and UKF) and IMM-LKF. Comparison results show superior accuracy and response in estimation and prediction using IMM-UKF approach, especially when the human makes circular motion or sharp turns. The MPC motion planning approach is compared with a reactive method and shows that the MPC method achieves better performance in generating smaller human-robot distance while similar performance in the velocity difference.

In the future work, we plan to investigate other motion prediction methods, such as the autoregressivemoving-average method, to compare with IMM-UKF method. Besides, enabling the robot to learn human motion model in real time is an attractive topic and may provide more accurate human motion prediction and results in better human-companion behaviors.

VI. APPENDIX

Similar to IMM-UKF, IMM-LKF works with two dynamic models: one is the uniform motion model; the other is the coordinated turn motion model. If the turn rate is a known constant in the coordinated turn motion model in **??**, the human estimation procedure can be modeled with the discrete time linear state space system as follows:

$$x^h(k+1) = Ax^h(k) + B_w w(k) \tag{10a}$$

$$y^h(k) = Cx^h(k) + v(k) \tag{10b}$$

where $x^h(k)$ and $y^h(k)$ represent the human motion state and the observation, respectively, at the time step k; $w(k)$ and $v(k)$ represent process noise and measurement noise, respectively. $x^h(k)$ consists of four elements: $p_1^h, v_1^h, p_2^h, v_2^h$, where p_1^h, p_2^h denote the longitudinal and lateral position of the human and v_1^h, v_2^h the corresponding velocity. We use two Linear Kalman Filters in the IMM for human tracking, each corresponding to a different dynamics model: the uniform motion model and the turn motion model. Two models differ in the A matrix and w in Eq. (10a) while sharing the same B_w. In particular, we define the matrices as follows:

$$A_U = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{11a}$$

$$A_T = \begin{bmatrix} 1 & \frac{\sin(\omega T)}{\omega} & 0 & \frac{1-\cos(\omega T)}{\omega} \\ 0 & \cos(\omega T) & 0 & -\sin(\omega T) \\ 0 & \frac{1-\cos(\omega T)}{\omega} & 1 & \frac{\sin(\omega T)}{\omega} \\ 0 & \sin(\omega T) & 0 & \cos(\omega T) \end{bmatrix}, \tag{11b}$$

$$B_w = \begin{bmatrix} \frac{T^2}{2} & T & 0 & 0 \\ 0 & 0 & \frac{T^2}{2} & T \end{bmatrix}, \tag{11c}$$

$$w_U \sim \mathcal{N}(0, Q_U) \; w_T \sim \mathcal{N}(0, Q_T) \tag{11d}$$

where A_U and A_T stand for the A matrices of the uniform motion model and turn motion model, respectively; w_U and w_T denote the process noise of the uniform motion model and turn motion model, respectively; T represents the sampling time; ω represents the constant turn rate.

We assume that only the human position can be measured. Therefore, the parameters in observation model Eq. (10b) can be defined as:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{12a}$$

$$v \sim \mathcal{N}(0, V) \tag{12b}$$

Above linear state space models are used in LKF and IMM-LKF in this paper. Moreover, we set the turn rate ω to be 0.1 rad/s as a known constant, in the turn motion model in IMM-LKF.

References

[1] 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.

[2] G. Welch and G. Bishop, "An introduction to the kalman filter," 1995.

[3] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.

[4] P. Trautman, "Probabilistic tools for human-robot cooperation," in *Human Agent Robot Teamwork Workshop HRI*, 2012.

[5] D. Vasquez, B. Okal, and K. O. Arras, "Inverse reinforcement learning algorithms and features for robot navigation in crowds: An experimental comparison," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pp. 1341–1346, IEEE, 2014.

[6] Z. Wu, M. Hashimoto, B. Guo, and K. Takahashi, "A path prediction method for human-accompanying mobile robot based on neural network," in *Intelligent Science and Intelligent Data Engineering*, pp. 35–42, Springer, 2012.

[7] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *The international journal of robotics research*, vol. 5, no. 1, pp. 90–98, 1986.

[8] A. Cosgun, D. A. Florencio, and H. I. Christensen, "Autonomous person following for telepresence robots," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pp. 4335–4342, IEEE, 2013.

[9] F. Hoeller, D. Schulz, M. Moors, and F. E. Schneider, "Accompanying persons with a mobile robot using motion prediction and probabilistic roadmaps," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pp. 1260–1265, IEEE, 2007.

[10] B.-S. Yaakov, K. Thiagalingam, and L. Rong, "Estimation with applications to tracking and navigation,", 2002.

[11] D.S.Caveney, "Multiple Model Techniques in Automotive Estimation and Control", Thesis (Ph.D.)–University of California, Berkeley, pp. 42–50, 2004

[12] S. Haykin, "Kalman Filtering and Neural Networks". John WileySons, Apr. 2004.

[13] S.Hong, T.Smith, F.Borrelli and J.K.Hedrick, "Vehicle Inertial Parameter Identification using Extended and Unscented Kalman Filters" in*Intelligent Transportation Systems, 2013. ITSC 2013*, pp. 1436–1441, IEEE, 2013.