Effective Team-Driven Multi-Model Motion Tracking *

Yang Gu
Computer Science Department
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA, USA
guyang@cs.cmu.edu

Manuela Veloso
Computer Science Department
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA, USA
mmv@cs.cmu.edu

ABSTRACT

Autonomous robots use sensors to perceive and track objects in the world. Tracking algorithms use object motion models to estimate the position of a moving object. Tracking efficiency completely depends on the accuracy of the motion model and of the sensory information. Interestingly, when the robots can actuate the object being tracked, the motion can become highly discontinuous and nonlinear. We have previously developed a successful tracking approach that effectively switches among object motion models as a function of the robot's actions. If the object to be tracked is actuated by a team, the set of motion models is quite more complex. In this paper, we report on a tracking approach that can use a dynamic multiple motion model based on a team coordination plan. We present the multi-model probabilistic tracking algorithms in detail and present empirical results both in simulation and real robot test. Our physical team is composed of a robot and a human in a real Segway soccer game scenario. We show how the coordinated plan allows the robot to better track a mobile object through the effective interaction with its human teammate.

Keywords

Team-Driven, Multi-Model, Motion Modelling, Tracking

1. INTRODUCTION

There have been considerable investigations into the problem of tracking moving targets e.g. [6]. Within the robotics community, there has been a similar interest in tracking targets from robot platforms e.g. [9]. When tracking is performed by a robot executing specific tasks acting over the target being tracked, such as a Segway RMP soccer robot grabbing and kicking a ball, the motion model of the target becomes dependent on the robot's actions [8]. The robot's tactic provides valuable information in terms of the target behavior. We have introduced the tactic-based motion modelling and tracking in such scenarios [7].

However, for the environments in which the Segway RMP soccer robot operates in, there are multiple targets, besides the ball, e.g. the teammate and the opponents, which need to be tracked properly. All the players on the field can also actuate over the ball, namely grab and kick the ball according to the rules which makes the motion model of the ball even more complex.

When the robot is playing a game as a member of a human-robot team, the team coordination knowledge provides further information that can be incorporated into the motion modelling and tracking process. In this paper, we present an extension to the tactic-based tracking scheme introduced in [7] to solve a plan-dependent multi-target tracking problem.

The paper is organized as follows. We first give a brief description of the Segway RMP soccer robot. Next we show the team-driven play-based motion modelling for multiple targets and we incorporate the team coordination knowledge into the motion modelling. We then describe the multi-sensor multi-model tracking algorithm for multiple targets, leading to our experimental results, related work, conclusions and future work.

2. SEGWAY RMP SOCCER ROBOT

The Segway platform is unique due to its combination of wheel actuators and dynamic balancing. Segway RMP, or Robot Mobility Platform, provides an extensible control platform for robotics research [10]. It imbues the robot with the novel characteristics of a fast platform and travel long ranges, able to carry significant payloads, able to navigate in relatively tight spaces for its size, and provides the opportunity to mount sensors at a height comparable to human eye level.

In our previous work, we have developed a Segway RMP robot base capable of playing Segway soccer (Figure 1). We briefly describe the two major components of the control architecture, the sensor and the robot cognition, which are highly related to our motion modelling for efficient multi-target tracking.

2.1 Vision Sensor and Infrared Sensor

*(Produces the permission block, copyright information and page numbering). For use with ACM_PROC_ARTICLE-SP.CLS V2.6SP. Supported by ACM.

Human Robot Interaction '06 Salt Lake City, UTah USA

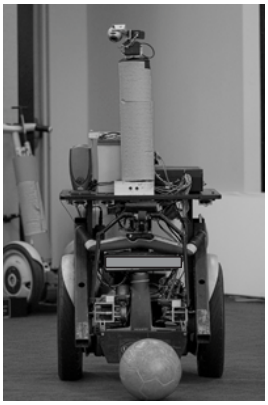

Figure 1: The Segway RMP soccer robot equipped with a kicker, a catcher, infrared sensors, and a camera mounted on a custom pan-tilt unit.

Over the years, a lot of different sensors such as vision sensors, infrared and ultrasound sensors have been used in the robotics community. For environments the Segway RMP operates in, there are few sensors that can compete with color vision for low cost, compact size, high information volume and throughput, relatively low latency, and promising usage for object recognition [5]. Thus, we choose vision as the primary sensor. The goal of vision is to provide as many valid estimates of targets as possible. Tracking then fuses this information to track the most interesting targets (the ball and the teammate, in this paper) of relevance to the robot. We do not discuss the localization of the robot in the sense that a lot of soccer tasks (known as tactics and plays in later sections) can be done by the Segway RMP robot independently of knowing where it is in the world. Also we use global reference in this paper (global position and velocity) which means it is relative to the reference point where the robot starts to do dead reckoning.

Recently, we have equipped each robot with infrared sensors to reliably detect the object which is in the catchable area of the robot. Its measurement is a binary value indicating whether or not an object is in that area. In most cases, this is the blind area of the vision sensor. Therefore, the infrared sensor is particularly useful when the robot is grabbing the ball.

2.2 Robot Cognition

A control architecture, called Skills-Tactics-Plays, was proposed in [3] to achieve the goals of responsive, adversarial team control. The key component of STP is the division between single robot behavior and team behavior.

A play, P, is a fixed team plan which consists of a set of applicability conditions, termination conditions, and N roles, one for each team member. Each role defines a sequence of tactics $T_1, T_2, \cdots$ and associated parameters to be performed by that role in the ordered sequence. Assignment of roles to team members is performed dynamically at run time. Upon role assignment, each robot i is assigned its tactic T_i to execute from the current step of the sequence for that role.

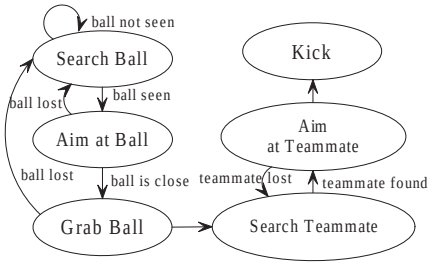

Figure 2: Skill state machines (SSMs) for an example tactic: *CatchKickToTeammate*. Each node is a skill and the edges show the transition between skills.

A tactic, T, encapsulates a single robot behavior. Each robot i executes its own tactic as created by the current play P. A tactic T_i determines the skill state machine SSM_i to be executed by the robot i.

A skill, S, is a focused control policy for performing some complex action. Each skill is a member of one, or more, skill state machines $SSM_1, SSM_2, \cdots$. Each skill S determines what skill it transitions to S' based upon the world state, the time skill S has been executing for, and the executing tactic for that robot.

We construct the robot cognition using a similar architecture. Plays, tactics, and skills, form a hierarchy for team control. Plays control the team behavior through tactics, while tactics encapsulate individual robot behavior and instantiate actions through sequences of skills. Skills implement the focused control policy for actually generating useful actions.

Figure 2 shows the $SSMs$ and transitions for an example tactic: *CatchKickToTeammate*, which contains six skills. The tactic starts from *SearchBall*, and when the ball is visible then transits to the skill *AimAtBall*. If the ball is lost, the state machine transits back to *SearchBall*. Else if the skill *GrabBall* is successfully executed, the state trasits to *SearchTeammate*, *AimAtTeammate* and the final *Kick* skill.

Segway soccer is a team sport, and therefore the building of our game strategy required not only execution of single robot behavior , but also coordination with the teammate, the human player. The current coordination is simple and basically based upon two fixed plays for offensive and defensive situation respectively. Our offensive play is shown as follows, in which the termination condition is either play aborted or the situation changed (a turn-over of ball possession announced by the coach). There are two roles in this play, one passes the ball to the other who positions down field and wait to receive a pass.

```
PLAY Naive Offense
APPLICABLE offense
DONE aborted !offense
ROLE 1
```

```
  pass 2
  none
ROLE 2
  position_down_field
  receive_pass
  none
```

Our current coordination is purely observation based. Each player assigns role from his own eyeshot without communication. For example, should the robot think the teammate is closer to the ball, the robot would choose to position and receive the ball (ROLE 2) from its teammate (ROLE 1). Furthermore, the robot knows which side gains possession of the ball from the coach announcement (whistle), therefore it tells offensive from defensive situation clearly and thus it has deterministic idea of which play the team is using. The robot makes an assumption that its teammate is performing the same game play as itself. The robot can infer what tactic the teammate is executing from the team play. For instance, after receiving the ball from the teammate, as a passer, the robot would assume the teammate go forward to a tactically advantageous position to receive a pass. The predefined play for team coordination provides useful information for motion modelling, which will be further discussed in section 3.

3. PLAY-BASED MOTION MODELLING

In this section, we take a multi-target tracking problem as a detailed example to show the extension of the tactic-based motion modelling method in general when the team coordination knowledge (play) is incorporated. First we give an introduction of the environment and targets under the Segway soccer setup. Second, we describe detailed motion models for both the teammate and the ball. Third, we extend the tactic-based motion modelling to the play level when both the ball and the teammate are included into the tracking. We show how we model the play-dependent interactions between the teammate, the robot and the ball and set up a base for giving the team-driven multi-model tracking algorithm in the next section.

3.1 Tracking Scenario

In a Segway soccer game, there are multiple moving targets on the field. e.g, the ball, the human teammate and the two opponents. Each team is identified by their distinct color. The ball is in orange [4]. We construct two single target trackers in the system, for the ball and the teammate respectively. We use two separate trackers instead of one multi-target tracker for both of them because we can differentiate the ball with the teammate thanks to the color-based vision system.

The general parameterized state-space system for the kth target $\mathbf{x}_{k,t}$ at time t is given by:

$$\mathbf{x}_{k,t} = f_k^m(\mathbf{x}_{k,t-1}, \mathbf{u}_{k,t-1}^m, \mathbf{v}_{k,t-1}^m) \quad (1)$$

$$\mathbf{z}_{k,t} = h_k^m(\mathbf{x}_{k,t}, \mathbf{n}_{k,t}^m) \quad (2)$$

where f_k^m and h_k^m are the parameterized state transition and measurement functions for the mth model of the kth target; $\mathbf{x}, \mathbf{u}, \mathbf{z}$ are the state, input and measurement vectors; $\mathbf{v}, \mathbf{n}$ are the process and measurement noise vectors of known statistics; m is the model index that can take any one of N_k values, where N_k is the number of models of the kth target being tracked (ball/teammate);

3.2 Ball Motion Modelling

In our Segway RMP soccer robot environment, we define five models to model the ball motion (in the rest of this paper, for simplicity, we use $\mathbf{x}_t$ to represent the ball state, and use $\mathbf{x}'_t$ to represent the teammate state).

- *Free-Ball.* The ball is not moving at all or moving straight with a constant speed decay d which depends on the environment surface.

$$\mathbf{x}_t = \mathbf{F}_t\mathbf{x}_{t-1} + \mathbf{v}_{t-1}^1 \quad (3)$$

$$\mathbf{z}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{n}_t^1 \quad (4)$$

  where $\mathbf{x}_t = (x_t, y_t, \dot{x}_t, \dot{y}_t)^T$, $\mathbf{z}_t = (x_t, y_t)^T$; x_t, y_t are the ball's x, y position in the global coordinate at time t; and $\dot{x}_t, \dot{y}_t$ are the ball's velocity in x and y direction in the global coordinate. The superscript "1" indicates the model index. $\mathbf{F}_t$ and $\mathbf{H}_t$ are known matrices as follows:

$$\mathbf{F}_k = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & d & 0 \\ 0 & 0 & 0 & d \end{bmatrix}, \mathbf{H}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

  where Δt is the time interval between vision frames.

- *Robot-Grab-Ball.* The ball is grabbed by the robot's catcher. In the case of robot grabbing ball, no vision is needed to track the ball, because we assume the ball moves with the robot. Therefore the ball has the same velocity as the robot (but plus the noise) and its global position at time t is just the robot's global position plus their relative position, which is assumed to be a constant, plus the noise.

- *Human-Grab-Ball.* The ball is held by the teammate. we can infer the ball position similarly if we know the teammate position well.

- *Robot-Kick-Ball.* The ball is kicked by the robot therefore its velocity is equal to a predefined initial speed plus the noise. The ball is supposed to move toward either the human teammate or the goal.

- *Human-Kick-Ball.* The ball is kicked by the teammate and it is supposed to be either a pass to the robot or a shoot at the goal .

3.3 Teammate Motion Modelling

We define four models to model the human teammate's motion.

- *Random Walk.* The teammate is wondering in the field. So the state at the new time is the state at the current time with some additive zero-mean (assumed Gaussian) noise.

- *Holding Ball.* The teammate is holding the ball without moving and waiting for the robot to receive the ball. Should the robot know the ball position well, it

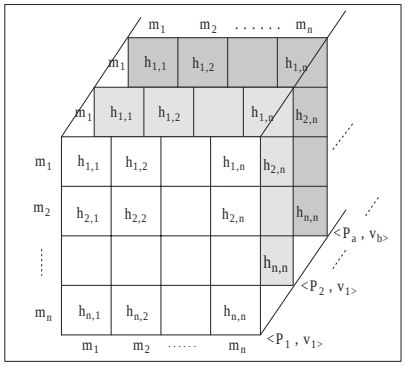

Figure 3: Play-Based motion modelling, where $m_1, m_2, \cdots, m_n$ are n models, P_a is the team play, v_b is the additional information. $h_{i,j}$ is the transition probability from model m_i to model m_j given m_i, and $\langle P_a, v_b \rangle$. Each layer in the graph is conditioned on a particular combination of the play executed and the additional information obtained.

can infer the teammate position by the ball position in a similar way as *Robot-Grab-Ball* for ball motion modelling.

- *Accelerating.* The teammate dashes and obtains a velocity in a short time.
- *Positioning.* The teammate is going to a predefined tactical position with a constant speed. This case happens mostly after the teammate passing the ball to the robot and moving down the field toward opponent's goal.

3.4 Play Based Model Transitions

Given the knowledge of the team coordination plan (the play P_{t-1} at time $t-1$), the robot can infer what tactic the teammate is executing (T'_{t-1}), which provides valuable information about the motion model of the teammate (m'_t). Both the robot and the teammate act over the ball in a Segway soccer game. The motion model of the ball (m_t) is therefore affected by what tactic the robot (T_{t-1}) and the teammate (T'_{t-1}) are executing.

From the previous subsection, we know that the model index m determines the present model being used. For our teammate tracking example, $m'_t = i, i = 1, \cdots, 4$. In our approach, it is assumed that the teammate motion model index, m'_t, conditioned on the current tactic executed T'_t by the teammate, and other useful information v'_t (such as ball state), is governed by an underlying Markov process, such that, the conditioning parameter can branch at the next time-step with probability.

$$p(m'_t = i | m'_{t-1} = j, T'_{t-1}, v'_t) = h'_{i,j} \tag{5}$$

where $i, j = 1, \cdots, N_{m'}$. Since T'_{t-1} can be determined by P_{t-1}, we get

$$h'_{i,j} = p(m'_t = i | m'_{t-1} = j, P_{t-1}, v'_t) \tag{6}$$

Since we can draw $p(m'_t = i | m'_{t-1} = j)$ in an $N_{m'} \times N_{m'}$ table, we can create a table for Equation 6 with a third axis

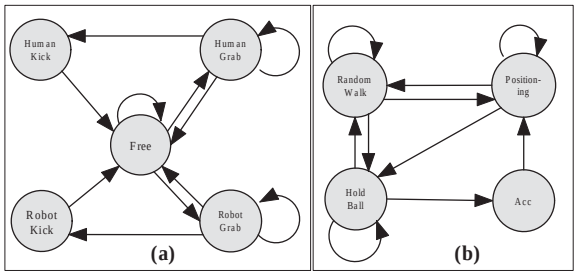

Figure 4: Object motion modelling based on the play: *Naive Offense*. Each node is a model. Models transit to one another according to the predefined probabilities (not shown in the figure). (a) Ball motion model. (b) Human teammate motion model.

which is defined by the tuple $\langle P_a, v_b \rangle$ as shown in Figure 3. Here the play P_a, is the primary factor that determines whether m_i transits to m_j and what the transition probability is, while the information v_b determines the prior condition of the transition. Each layer in the graph is conditioned on a particular combination of the tactic executed and the additional information obtained.

For our ball tracking example, $m_t = i, i = 1, \cdots, 5$. Similarly,

$$h_{i,j} = p(m_t = i | m_{t-1} = j, T_{t-1}, T'_{t-1}, v_t) \tag{7}$$

where $i, j = 1, \cdots, N_m$. Since T_{t-1}, T'_{t-1} can be determined by P_{t-1}, we get

$$h_{i,j} = p(m_t = i | m_{t-1} = j, P_{t-1}, v_t) \tag{8}$$

Suppose the current team play is the *Naive Offense* in Section 2.2, we can obtain the corresponding motion model transitions for both the ball and the teammate using the play-based method (Figure 4).

4. MULTI-SENSOR MULTI-MODEL TRACKING

Following the play-based motion model given in the previous section, we can use a dynamic Bayesian network (DBN) to represent the whole system for teammate and ball tracking in a natural and compact way as shown in Figure 5 and Figure 6 respectively. In this graph, the system state is represented by variables (play P, tactic T, infrared sensor measurement s, ball state $\mathbf{x}$, ball motion model index m, vision sensor measurement of ball $\mathbf{z}$, teammate state $\mathbf{x}'$, teammate motion model index m', vision sensor measurement of teammate $\mathbf{z}'$), where each variable takes on values in some space. The variables change over time in discrete intervals, so that m_t is the object state at time t.

Furthermore, the edges indicate dependencies between the variables. For instance, in Figure 6 the ball motion model index m_t depends on $m_{t-1}, T_{t-1}, T'_{t-1}, s_t$ and $\mathbf{x}_{t-1}$, hence there are edges coming from the latter five variables to m_t. Note that we use an approximation here. We assume the measurement of the infrared sensor is always the true value, so it does not depend on the ball state. Under this assumption, there is no edge from $\mathbf{x}_{t-1}$ to s_t, which greatly simplifies the ball-tracking DBN and the sampling algorithm as well.

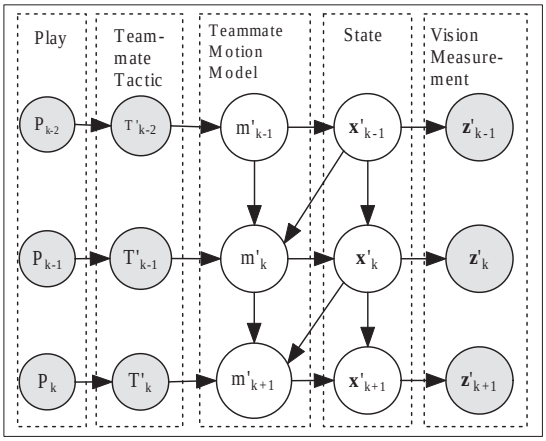

Figure 5: A dynamic Bayesian network for teammate tracking with a Segway RMP robot. Filled circles represent deterministic variables which are observable or are known as the tactic or the play that the robot is executing.

For the rest of this section, we give the ball-tracking algorithm following Figure 6. The teammate-tracking algorithm can be obtained similarly following Figure 5.

We use the sequential Monte Carlo method to track the motion model m and the object state $\mathbf{x}$. Particle filtering is a general purpose Monte Carlo scheme for tracking in a dynamic system. It maintains the belief state at time t as a set of particles $p_t^{(1)}, p_t^{(2)}, \cdots, p_t^{(N_s)}$, where each $p_t^{(i)}$ is a full instantiation of the tracked variables $\{p_t^{(i)}, w_t^{(i)}\}$, $w_t^{(i)}$ is the weight of particle $p_t^{(i)}$ and N_s is the number of particles. In our case, $p_t^{(i)} = \langle \mathbf{x}_t^{(i)}, m_t^{(i)} \rangle$.

The equations below follow from the ball-tracking DBN.

$$m_t^{(i)} \sim p(m_t | m_{t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, s_t, T_{t-1}, T'_{t-1}) \quad (9)$$

$$\mathbf{x}_t^{(i)} \sim p(\mathbf{x}_t | m_t^{(i)}, \mathbf{x}_{t-1}^{(i)}) \quad (10)$$

Note that T_{t-1} and T'_{t-1} are inferred deterministically from P_{t-1} instead of sampling. Also note that in Equation 10, the ball state is conditioned on the ball motion model $m_t^{(i)}$ sampled from Equation 9.

Then we use the Sample Importance Resampling (SIR) algorithm to update the state estimates. The sampling algorithm is as follows:

$[\{\mathbf{x}_t^{(i)}, m_t^{(i)}, w_t^{(i)}\}_{i=1}^{N_s}] = \textbf{\textit{SIR}}[\{\mathbf{x}_{t-1}^{(i)}, m_{t-1}^{(i)}, w_{t-1}^{(i)}\}_{i=1}^{N_s}, \mathbf{z}_t, s_t, T_{t-1}, T'_{t-1}]$

01 **for** $i = 1 : N_s$
02 draw $m_t^{(i)} \sim p(m_t | m_{t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, s_t, T_{t-1}, T'_{t-1})$.
03 draw $\mathbf{x}_t^{(i)} \sim p(\mathbf{x}_t | m_t^{(i)}, \mathbf{x}_{t-1}^{(i)})$.
04 set $w_t^{(i)} = p(\mathbf{z}_t | \mathbf{x}_t^{(i)})$
05 **end for**
06 Calculate total weight: $w = \sum[\{w_t^i\}_{i=1}^{N_s}]$
07 **for** $i = 1 : N_s$
08 Normalize: $w_t^i = w_t^i / w$
09 **end for**
10 ***Resample***.

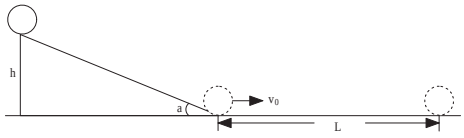

Figure 7: Test setup for estimating the ball speed decay d. The ball rolls off the ramp (with height h) with speed v_0 and it stops after it travels a distance of L.

The inputs of the algorithm are samples drawn from the previous posterior $\langle \mathbf{x}_{t-1}^{(i)}, m_{t-1}^{(i)}, w_{t-1}^{(i)} \rangle$, the present vision and infrared sensory measurement z_t, s_t, and the tactic t_{t-1}. The outputs are the updated weighted samples $\langle \mathbf{x}_t^{(i)}, m_t^{(i)}, w_t^{(i)} \rangle$. In the sampling algorithm, first, a new ball motion model index, $m_t^{(i)}$, is sampled according to Equation 9 at line 02. Then given the model index, and previous ball state, a new ball state is sampled according to Equation 10 at line 03. The importance weight of each sample is given by the likelihood of the vision measurement given the predicted new ball state at line 04. Finally, each weight is normalized and the samples are resampled. Then we can estimate the ball state based on the mean of all the $\mathbf{x}_t^{(i)}$. Similarly the state of the teammate $\mathbf{x}'_t$ can be obtained from the teammate tracker.

5. EXPERIMENT

In this section, we design experiments to estimate the ball speed decay in Δt (time interval between vision frames) on different surfaces. We profile the system and measurement noise. Finally we evaluate the effectiveness of our tracking system in both simulated and real-world tests.

5.1 Ball Motion Profiling

From previous work we know the initial speed and accuracy of the ball velocity after a kick motion.

And we use the setup shown in Figure 7 to estimate the ball speed decay d. In detail, we put the ball on the top of a ramp and let it roll off the ramp with initial speed

$$v_0 = \sqrt{2gh}$$

without taking the friction on the surface of the ramp into account, where g is the gravity and h is the height of the ramp. We record the distance the ball travelled (L) from the position the ball rolls off the ramp to the position it stops. Obviously, the ball speed decay can be approximated as

$$d = 1 - \frac{v_0 \Delta t}{L}$$

where $\Delta t \approx 0.033$ sec. Following the test result, we use $d = 0.99$ for the cement surface. From the test, we note that the faster the ball's speed, the smaller the system noise, hence the more the ball's trajectory forms a straight line. Based on the data we collected from experiments, we therefore model the system noise when the motion model is *Free-Ball* to be inverse proportional to the ball speed.

5.2 Measurement Noise Profiling

In order to profile the measurement noise, we put the ball on a series of known positions, read the measurement from

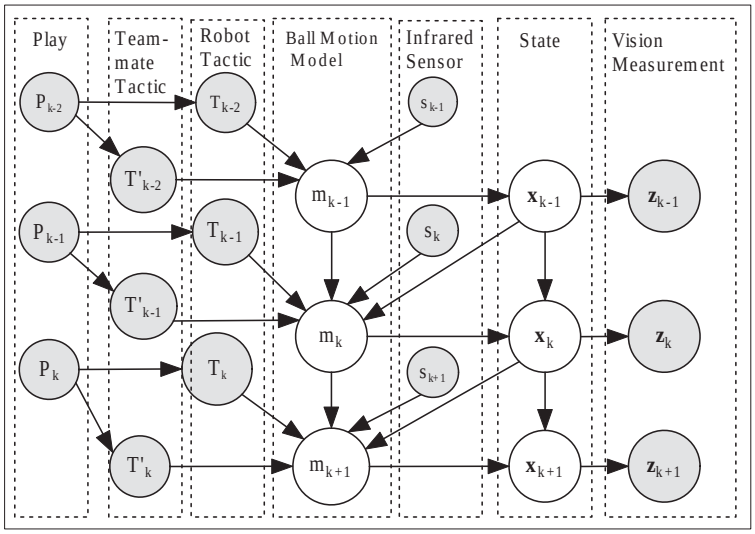

Figure 6: A dynamic Bayesian network for ball tracking with a Segway RMP robot.

Table 1: The average RMS error of position estimation and velocity estimation from human-trackers.

Motion Model	Single Model	Multi-Model
Position Est RMS (m)	0.0030	0.0014
Velocity Est RMS (m/s)	0.42	0.025

Table 2: The average RMS error of position estimation and velocity estimation from ball-trackers.

Motion Model	Single Model	Multi-Model
Position Est RMS (m)	0.0028	0.0017
Velocity Est RMS (m/s)	0.4218	0.0597

vision sensor, and then determine the error in that measurement. From the results, we know that the nearer the ball, the smaller the observation noise. Therefore we choose to approximate the error distribution as different Gaussians based on the distance from the robot to the ball.

5.3 Simulation Experiments

Because it is difficult to know the ground truth of the object's position and velocity in the real robot test, we do the simulation experiments to evaluate the precision of tracking.

Experiments are done following the *Naive Offense* play, in which the robot acts as the receiver and the human teammate acts as the passer. Noises are simulated according to the model profiled in the previous section. In the beginning, the teammate holds the ball. After a fixed amount of time, the ball is kicked towards the robot, and the teammate moves forward to a predefined location.

We implement both a single model tracker and a play-based multi-model tracker for the ball and the teammate. We simulate the experiment for 50 runs, and then compare the performance of the two trackers with different implementations. The average RMS error of position estimation and velocity estimation are shown in Table 1 and 2 respectively. The results show that the play-based multi-model scheme performs much better than the single model especially in terms of velocity estimation. Because with the play-based motion model, when the ball is being kicked, most particles evolving using the transition model determined by the play will change its motion model $m_t^{(i)}$ from *Free-Ball* to *Human-Kick-Ball*, and a velocity will be added to the ball accordingly.

Figure 8 and Figure 9 show the ball velocity estimation and the teammate velocity estimation during a short term for a given simulation test. In both figures, The left graph shows the x-component of the velocity (v_x) estimation through single model tracking and play-based multi-model tracking. The right graph shows the y-component of the velocity (v_y) estimation. The dotted line with x-mark represents the true value, the solid line with circle represents the the velocity estimation through play-based multi-model tracking, the solid line with cross represents the the velocity estimation through single model tracking. We note that the velocity estimation with multi-model trackers the true velocity in terms v_x and v_y much more consistent than with single model trackers.

5.4 Team Cooperation Test

In the real-world test, we do experiments on the Segway RMP soccer robot executing the offensive play and coordinating with the human teammate. The test setup is demonstrated in Figure 10, in which the digits along the lines show the sequence of the whole strategy, the filled circle at position B represents the robot, the unfilled circle at position E represent an opponent player, and the shaded circle represent the human teammate.

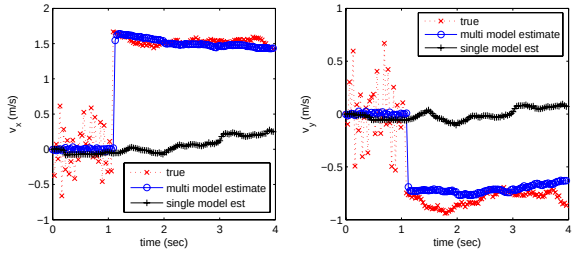

Figure 8: Ball velocity estimation. The left figure shows the x-component of the velocity (v_x) estimation through single model tracking and play-based multi-model tracking. The right figure shows the y-component of the velocity (v_y) estimation. The dotted line with x-mark represents the true value, the solid line with circle represents the the velocity estimation through play-based multi-model tracking, the solid line with cross represents the the velocity estimation through single model tracking.

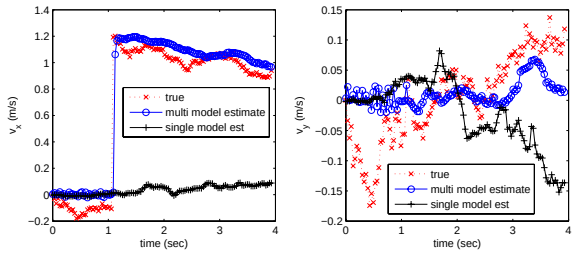

Figure 9: Human teammate velocity estimation. The left figure shows the x-component of the velocity (v_x) estimation through single model tracking and play-based multi-model tracking. The right figure shows the y-component of the velocity (v_y) estimation. The dotted line with x-mark represents the true value, the solid line with circle represents the the velocity estimation through play-based multi-model tracking, the solid line with cross represents the the velocity estimation through single model tracking.

Table 3: The average time taken over all the successful runs.

Motion Model	Single Model	Multi-Model
Mean Time (sec)	33.4	22.6

When each run begins, the human teammate is at position A . With this team cooperation plan (play), the robot chooses the tactic *CatchKickToTeammate* to execute, in which the robot starts with the skill *Search-Ball*. When the robot finds the ball, the teammate passes the ball directly to the robot and chooses a positioning point to go to either at C or D. The robot grabs the ball after the ball is in the catchable area and is detected by the infrared sensor (skill *Grab-Ball*). Next the robot searches for the teammate holding the ball with its catcher (skill *Search-Teammate*). After the robot finds the teammate, the robot kicks the ball to its teammate and the teammate shoots at the goal(skill *KickToTeammate*, completing the whole offensive play. Each run ends in one of the following conditions.

- succeed if the human receives the ball from the robot or the human does not receiver the ball but the pass can be considered as a "good" one.
- fail if the robot is in searching for the ball or the teammate for more than 30 seconds.
- fail if the ball is out of the field before the robot catches it.

In the experiment over 15 runs, the robot with single model trackers fails 5 of the total. While the robot with play-based multi-model trackers fails 2 of the total. We also keep track of the mean time taken in each successful run listed in Table 3. Using play-based multi-model tracking saves 32.3% time in terms of completing the whole play over single model tracking. During the experiment, we note that when using the single model tracking, most time were spent on searching the teammate. Incorporating the team cooperation knowledge known as play into the teammate motion modelling greatly improves the accuracy of the teammate motion model and therefore avoids taking time in searching a lost target from scratch.

6. RELATED WORK

Tracking moving objects using a Kalman filter is the optional solution if the system follows a single model, f and h in Equation 1 and 2 are known linear functions and the noise $\mathbf{v}$ and $\mathbf{n}$ are Gaussians [1]. Multiple model Kalman filters such as Interacting Multiple Model (IMM) are known to be superior to the single Kalman filter when the tracked object is maneuvering [2]. For nonlinear systems or systems with non-Gaussian noises, a further approximation is introduced, but the posterior densities are therefore only locally accurate and do not reflect the actual system densities.

Since the particle filter is not restricted to Gaussian densities, a multi-model particle filter is introduced. However,

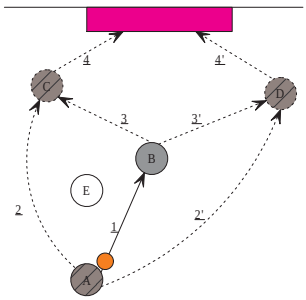

Figure 10: A demonstration of a naive team cooperation plan in offensive scenario. The digits along the lines show the sequence of the whole plan. The filled circle at position B represents the robot. The unfilled circle at position E represent an opponent player. The shaded circle represent the human teammate.

this approach assumes that the model index, m, is governed by a Markov process such that the conditioning parameter can branch at the next time-step with probability $p(m_t = i|m_{t-1} = j) = h_{i,j}$ where $i, j = 1, \cdots, N_m$. But the uncertainties in our object tracking problem do not have such a property due to the interactions between the robot and the tracked object. In this motivation, a tactic-based motion modelling method is proposed in [7]. Based on that approach, we further introduce the play-based motion modelling method when team coordination knowledge is available.

In [8], an approach were proposed for tracking a moving target using Rao-Blackwellised particle filter. In their experiments, the discrete states are the non-linear motion of the observing platform and the different motion models for the target. But they use a fixed transition table between different models. Our transition model is dependent on the play that the robot is executing and the additional information that matters. This play-based motion modelling can be flexibly integrated into our existing skills-tactics-plays architecture.

7. CONCLUSIONS AND FUTURE WORK

Motivated by the interactions between a team and the tracked object, we contribute a method to achieve efficient tracking through using a play-based motion model and combined vision and infrared sensory information. The team-driven motion modelling method gives the robot a more exact task-specific motion model when executing different tactics over the tracked object (e.g. the ball) or collaborating with the tracked object (e.g. the teammate). Then we represent the system in a compact dynamic Bayesian network and use particle filter to keep track of the motion model and object state through sampling. The empirical results from the simulated and the real experiments show the efficiency of the team-driven multi-model tracking over single model tracking.

Future work will include modelling the multi-target motion when each target has multiple hypothesis, which is caused by incorrect measurements originating from the clutter. We would like to see how the information from the tactic and the play can help to eliminate false alarms and achieve efficient resampling under the framework of the particle filter.

8. ACKNOWLEDGMENTS

We would like to thank Brett Browning for developing the whole architecture and vision system and Brenna Argall for developing the skill and game play for the Segway robots.

This work was supported by United States Department of the Interior under Grant No. NBCH-1040007. The content of the information in this publication does not necessarily reflect the position or policy of the Defense Advanced Research Projects Agency (DARPA), US Department of Interior, US Government, and no official endorsement should be inferred.

9. REFERENCES

[1] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, Feb. 2002.

[2] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc, 2001.

[3] B. Browning, J. Bruce, M. Bowling, and M. Veloso. STP: Skills, tactics and plays for multi-robot control in adversarial environments. *IEEE Journal of Control and Systems Engineering*, 219:33–52, 2005.

[4] B. Browning, P. Rybski, Y. Gu, and B. Argall. Segwayrmp robot football league rules. In *Coral Publications*, 2005.

[5] B. Browning, L. Xu, and M. Veloso. Skill acquisition and use for a dynamically-balancing soccer robot. *Proceedings of Nineteenth National Conference on Artificial Intelligence*, 2004.

[6] A. Doucet, N. D. Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.

[7] Y. Gu. Tactic-based motion modelling and multi-sensor tracking. *Proceedings of Twentieth National Conference on Artificial Intelligence*, 2005.

[8] C. Kwok and D. Fox. Map-based multiple model tracking of a moving object. *Proceedings of eight RoboCup International Symposium*, July 2004.

[9] D. Schulz, W. Burgrad, and D. Fox. People tracking with mobile robots using sample-based joint probabilistic data association filters. *International Journal of Robotics Research*, 22(2), 2003.

[10] J. Searock, B. Browning, and M. Veloso. Turning segways into robust human-scale dynamically balanced soccer robots. *Proceedings of eight RoboCup International Symposium*, July 2004.