Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2009, Article ID 284525, 17 pages
doi:10.1155/2009/284525

*Research Article*

# Prioritized Multihypothesis Tracking by a Robot with Limited Sensing

**Paul E. Rybski and Manuela M. Veloso**

*School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

Correspondence should be addressed to Paul E. Rybski, prybski@cs.cmu.edu

Received 1 August 2008; Accepted 1 December 2008

Recommended by Matthijs Spaan

To act intelligently in dynamic environments, mobile robots must estimate object positions using information obtained from a variety of sources. We formally describe the problem of estimating the state of objects where a robot can only task its sensors to view one object at a time. We contribute an object tracking method that generates and maintains multiple hypotheses consisting of probabilistic state estimates that are generated by the individual information sources. These different hypotheses can be generated by the robot's own prediction model and by communicating robot team members. The multiple hypotheses are often spatially disjoint and cannot simultaneously be verified by the robot's limited sensors. Instead, the robot must decide towards which hypothesis its sensors should be tasked by evaluating each hypothesis on its likelihood of containing the object. Our contributed algorithm prioritizes the different hypotheses, according to rankings set by the expected uncertainty in the object's motion model, as well as the uncertainties in the sources of information used to track their positions. We describe the algorithm in detail and show extensive empirical results in simulation as well as experiments on actual robots that demonstrate the effectiveness of our approach.



## 1. Introduction

Robot perception processing consists of a mapping from sensory data to an estimate of the state of the elements of the environment that are of relevance to the task under execution. For example, a robot traversing a maze needs to estimate the area and position of open space and walls from its sensory data. Similarly in a team of soccer robots, each robot has the potential to estimate the state of the environment based on its own sensing and on the information communicated by its teammates. The complexity of state estimation greatly increases with the task, the dynamics of the environment, and the sensing capabilities of the robots.

In our work, we consider that robots have limited sensing and operate in complex and dynamic environments executing tasks that rely on multiple elements. We investigate robot state estimation as a result of the integration of sensory information obtained from a variety of sources, namely, the robot's own sensors and actions, models and communicated information from teammate robots' sensors and actuators, as well as models of the dynamics of the environment.

Concretely, we investigate the problem when robots have limited and narrow perceptual scope, such that they are only capable of observing a single object (or a reduced set of objects) at a time with their sensors. Thus, the relative size of the robot's sensor scope is small compared to the environment, and while the state of a single object is being updated by the sensors, the evolving state of all other nonsensed objects must be predicted from communicated information or from models learned from observations or provided a priori.

In addition to the complexity of the problem, not all sources of information about a single object can and should be handled equally, as in the traditional sense of weighting those estimates by their covariance. There are times when empirical evidence has proven that some modalities must be ignored as they are unreliable in certain circumstances.

Additionally, nondeterministic effects of actuators can create several distinctly different potential outcomes, each of which must be tracked and reasoned about separately.

To address this challenge, we define a method for reasoning over a disjoint hypothesis space whereby high-level domain knowledge is used to impose a strict ordering on estimates created by different sources of information. By segmenting the sources of information used to reason about the state of environmental quantities into different classes, each with different state dynamics and expected effect of robot actions, a prioritized hierarchy of state estimates can be inferred. Additionally, when tracking multiple objects simultaneously, the evolving states of those objects must be considered carefully when deciding where to task the robot's sensors.

We describe a hybrid state estimation algorithm that attempts to reduce the complexity of the generated probability density functions over a quantity of interest by factoring the problem into a series of small estimation problems that are tied to the different sources of model world information possessed by the robot. A high-level policy is used to determine where to task the robot's sensors to best track the objects in the environment. Such policies for creating hierarchies can be defined a priori, or they could potentially be learned from data. Using this policy, the decision process that governs each individual robot's actions can easily select the most informative state estimate to use as its input. The priorities are set by the expected uncertainty in the object's motion model as well as the uncertainties in the sources of information used to track their positions. Robot's actions directly affect its perception of the environment as well as the environment itself, and the best estimate is often one that will allow the robot to obtain more information about its surroundings to further clarify its estimate of quantities of interest. This, in turn, provides more information to the robot that further updates the ordered hierarchy of possible estimates.

This paper describes an active state estimation algorithm, as applied to a real-time adversarial multirobot domain, which combines action policies determined from high-level domain knowledge with multimodal probabilistic state estimators. In this work, we assume that each of the objects that are detected and tracked have unique sensor signatures whereby the additional complexity of the data association problem can be avoided so that we instead focus on analyzing the multiple hypothesis reasoning algorithms. Thus, we contribute an algorithm to address the problem of tracking a single object with multiple hypotheses. We have successfully applied this approach to the RoboCup Four-Legged league where a team of Sony AIBO robots autonomously play soccer against another team of AIBO robots, as shown in Figure 1.

Figure 1: Sony AIBO robots preparing to play robot soccer at a RoboCup competition.

## 2. Related Work

We discuss some related work along the three main aspects of our work: (i) probabilistic state estimation; (ii) object tracking; and (iii) reasoning about multiple hypotheses from multiple sensing sources.

Most probabilistic estimation techniques follow a Bayesian filtering approach [1] and have been successfully applied to robot state estimation (e.g., [2]). Object tracking using a Bayesian filter formalism relies on an a priori model of the object's motion that allows the algorithm to predict the object motion given noisy observations. One of the most widely used methods for state estimation is the Kalman filter [3], in which the system model is assumed linear and the noise is assumed Gaussian. When the linearity assumption becomes a limitation, the dimension of the state vector can be changed as the tracked object changes its perceived dynamics, such as with a variable state dimension (VSD) filter [4]. We also consider the object dynamics, but our approach changes the number of hypotheses, while the specific dimensions of those hypotheses' estimates do not change. Furthermore, we maintain multiple hypotheses independently as potential object locations.

An approach to reasoning about a complex motion model consists of maintaining multiple models. The interacting multiple model (IMM) filter [5] uses a weighted mixture of different process models. Our approach differs in that it maintains a disjoint set of hypotheses which are not merged or fused [6], but are prioritized and visited according to a specific policy. Similar approaches maintain separate estimations based on subjective sensing and other sources (e.g., sensing from robot teammates [7, 8]).

A more general approach is the Switching Kalman filter model [9], which represents multiple independent system state dynamics models and switches between them (or linearly combines them) to best fit the observed (or predicted) nonlinear dynamics of the system being modeled. Our approach creates multiple independent belief states (or hypotheses) rather than a single state with multiple potential models.

A multiple hypothesis tracking (MHT) [10, 11] approach uses multiple independent state estimators to estimate a multimodal probability density. This approach has been used successfully for challenging mobile robot localization problems [12], where nonparametric distributions are estimated through sampling techniques, such as the particle filtering [13]. The number of particles used can be dynamically adjusted as computational resources become available or are needed elsewhere [14]. Approaches that factor in a joint state estimation have been used successfully for tracking an object with a mobile robot [15], where the actions of the robot change the process characteristics of the tracked object. Our approach extends the MHT paradigm by reasoning about the different hypotheses as a function of the source of information that generated them.

Finally, object tracking is a complex problem addressed by different approaches that capture connected dynamics of the multiple objects. We address the problem of object tracking from a different perspective, namely, in terms similar to that of sensor planning [16]. Sensor (or actuator) planning generally requires that a policy be determined over a state space which dictates the appropriate action to take based on the state of the world and the robot. In [17], the reinforcement learning is used to find the a policy that avoids the problems of a state space explosion as well as the problems associated with missing sensor information. Our problem is defined over a continuous state space (e.g., the space of tracked object poses) whereby the effects of various actions are difficult to quantize into a state space on which a policy could be learned. In [18], a dynamic programming algorithm is proposed by which a static policy state over the entire field is determined, which dictates when the robot should stop its body and localize itself. In our model, the actions of the objects being tracked in the world are highly dynamic and are unlikely to be captured in a single policy over the entire space of poses. In [19, 20], the mechanisms for attention control are proposed that use expected information gain and cost to acquire the information as criteria to determine what the robot should do and when. A decision tree is learned to represent the policy of the robot. Our approach uses similar criteria to determine what the robot should do and when though in this work we do not discuss mechanisms for how the knowledge is obtained (e.g., learned offline or hand-coded) but rather focus on the utility of using such concepts and applying them to probabilistic state estimators that operate over a continuous state space.

We consider that the robot has a narrow sensor scope incapable of capturing more than one object at a time. Our algorithm includes a policy for directing the sensor machinery toward multiple objects. Furthermore, we consider different types of objects with different motion models which are used to update the confidence on the state estimation of each individual object.

## 3. Challenges of Dynamic World Modeling

We are interested in problems associated with having a robot autonomously build and maintain accurate world models in dynamic environments where the states of many objects must be estimated simultaneously. A robot will be able to make use of multiple sources of information that can describe the motion of objects in the environment. In any environment of reasonable complexity, a robot is incapable of viewing the entire environment at a single time with its sensors. In the extreme case, the robot can only track a single object at a time with its sensors. Figure 2 illustrates the general class of world modeling issues addressed in this work. We are primarily concerned with the issues involved with *object tracking* rather than issues involved with the complementary field of *map building* which is not part of our discussion.

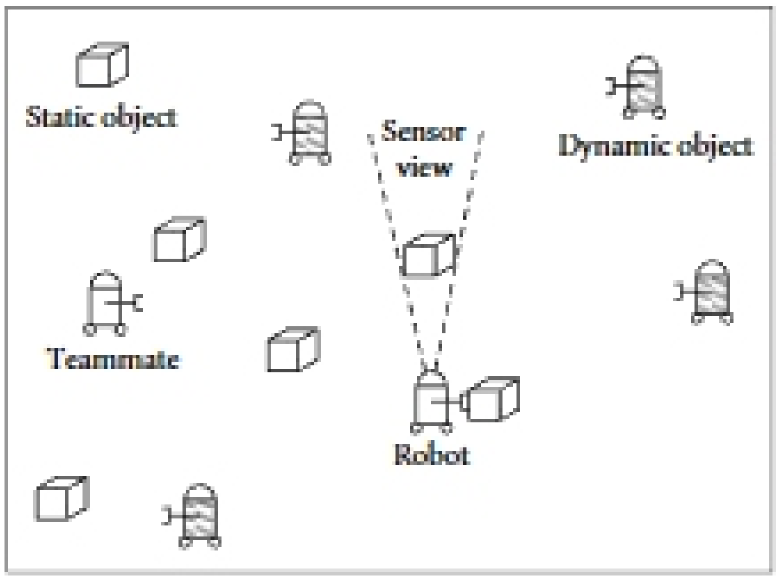


Figure 2: The general world modeling problem in a dynamic environment includes requiring a robot to use a narrow-scope sensor to track the positions of multiple (static and dynamic) objects in an environment. Determining when and how to use additional sources of information, such as from the effects of actuation, and teammate sensor information is a nontrivial task.

We consider the challenges of tracking multiple objects, where each object has multiple sources of sensor and model information that are available as a combined problem. In this work, we do not address the additional complexity of the data association problem where multiple objects have identical or ambiguous sensor signatures. In order to keep track of the positions of all objects in the environment, the robot must continually retask its sensors to refresh the models with more accurate position data. Deciding which object to track next is dependent on the expected uncertainty in the motion model for that object as well as the availability and quality of the different sources of information that can provide estimates for the expected position of the object.

To formally describe the problem, we define the following concepts:

$A$: the set of all actions, $a(t) \in A$, including the null action, that the robot is capable of performing at time $t$;

$O$: the set of all objects in the environment where $O_j$ is the $j$th object of which the robot must keep track, the set includes moving objects of which the robot must maintain an accurate estimate as well as stationary objects with which the robot must maintain periodic contact (such as landmarks for localization);

$X_{O_j}(t)$: the estimated position of object $O_j$ at time $t$;

$L_{O_j}(t)$: a sensor observation of object $O_j$ at time $t$ which can be null in the case that the robot does not perceive object $O_j$;