Probabilistic Vision-Based Opponent Tracking in Robot Soccer

Thorsten Schmitt, Robert Hanek, Sebastian Buck, and Michael Beetz

TU München, Institut für Informatik, 80290 München, Germany
{schmittt,hanek,buck,beetzm}@in.tum.de
http://www9.in.tum.de/agilo

Abstract. Good soccer players must keep their eyes on their opponents in order to make the right plays and moves. The same holds for soccer robots, too. In this paper, we apply probabilistic multiple object tracking to the continual estimation of the positions of opponent players in autonomous robot soccer. We extend MHT [3], an existing tracking algorithm, to handle multiple mobile sensors with uncertain positions, discuss the specification of probabilistic models needed by the algorithm, and describe the required vision-interpretation algorithms. The tracking algorithm enables robots to estimate the positions and motions of fast moving robots both accurately and robustly. We have applied the multiple object tracking algorithm throughout the RoboCup 2001 world championship. Empirical results show the applicability of multiple hypotheses tracking to vision-based opponent tracking and demonstrates the advantages for crowded environments.

1 Introduction

Good soccer players must keep their eyes on their opponents in order to make the right plays and moves. The same holds for soccer robots, too [2]. Unfortunately, object tracking systems are difficult to realize. Observations of the robots are inaccurate and incomplete. Sometimes the sensors hallucinate objects. Often the robots cannot perceptually distinguish the individual objects in their environments. To reliably estimate the positions and motions of the objects despite these perturbations, researchers have proposed object tracking algorithms that are capable of tracking multiple objects. Tracking algorithms use motion models of the objects and sequences of observation to distinguish real object observations from clutter and can thereby keep track of object positions both more reliably and more accurately.

Multiple object tracking is particular difficult for autonomous robot soccer, where the state is to be estimated by multiple mobile sensors with uncertain positions, the soccer field is only partly visible for each sensor, occlusion of robots is a problem, the robots change their direction and speed very abruptly, and the models of the dynamic states of the robots of the other team are very crude and uncertain.

G.A. Kaminka, P.U. Lima, and R. Rojas (Eds.): RoboCup 2002, LNAI 2752, pp. 426–434, 2003.

Many robots employ probabilistic state estimation algorithms for keeping track of the moving objects in their environments [12], such as Multiple Hypothesis Tracking (MHT) [8,3] and Joint Probabilistic Data Association Filter (JPDAF) [1,11]. Using probabilistic motion and sensing models these algorithms maintain probabilistic estimates of the objects' positions and update these estimates with each new observation. Probabilistic tracking algorithms are attractive because they are concise, elegant, well understood, and remarkably robust.

In this paper we show how the MHT algorithm can be applied to opponent tracking in autonomous robot soccer. This application requires programmers to equip the robots with sophisticated mechanisms for observing the required information, and to provide probabilistic domain descriptions that the algorithm needs for successful operation. These probabilistic descriptions include motion models and sensing models, such as the probability of the robot detecting an object within sensor range. We show that such mechanisms enable the MHT to reliably and accurately estimate the positions of opponent robots using passive vision-based perception where the cameras have a very restricted field of view. In addition, we will show that the cooperation between robots provides the robots with a more complete estimate of the world state, a substantial speed up in the detection of motions, and more accurate position estimates.

In the remainder of the paper we proceed as follows. The next section describes the MHT algorithm. In the subsequent section we provide a detailed account of how to apply the MHT to autonomous robot soccer. We conclude with empirical results and a discussion of related work.

2 Multiple Hypothesis Tracking

Multiple hypothesis tracking considers the following state estimation problem. The world is populated with a set of stationary and moving objects. The number of objects may vary and they might be occluded and out of sensor range. Robots are equipped with sensing routines that are capable of detecting objects within sensor range, of estimating the positions of the detected objects, and of assessing the accuracy of their estimate.

The objective of the MHT algorithm is to keep a set of object hypotheses, each describing a unique real object and its position, to maintain the set of hypotheses over time, and to estimate the likelihood of the individual hypotheses.

The basic data structure used by the MHT algorithm is the object hypothesis. An object hypothesis consists of an estimated position, orientation, and velocity of an object, a measure of uncertainty associated with the estimation, and a second measure that represents the degree of belief that this hypothesis accurately reflects an existing object. Because the number of objects might vary new hypotheses might have to be added and old ones might have to be deleted.

Before we dive into the details of the MHT algorithm let us first get an intuition of how it works. The MHT algorithm maintains a forest of object hypotheses, that is a set of trees. The nodes in the forest are object hypotheses and represent the association of an observed object with an existing object hy-

```
algorithm MULTIPLEHYPOTHESISTRACKING()
1   let Ĥ^k = {ĥ^k_1, ..., ĥ^k_{m_k}}  % predicted hyps.
2       Z(k) = {z_1(k), ..., z_{n_k}(k)}  % observed features
3       H^k = {h^k_1, ..., h^k_{o_k}}  % new hyps.
4       X^{k-N}  % world state at time step k-N.
5   do for k ← 1 to ∞
6       do Z(k) ← INTERPRETSENSORDATA(GETSENSORDATA());
7          Ĥ^k ← APPLYMOTIONMODEL(H^{k-1}, M);
8          for i ← 1 to n_k
9          do for j ← 1 to m_k
10             do h^k_{ij} ← ASSOCIATE(ĥ^k_j, z_i(k));
11                COMPUTE(P(h^k_{ij}|Z(k)))
12         for j ← 1 to n_k
13         do H^k ← H^k ∪ {GENERATENEWHYP(z_j(k))};
14         PRUNEHYPOTHESES(H^k);
15         X^{k-N} ← {x^{k-N}_1, ..., x^{k-N}_{o_{k-N}}}
```

Fig. 1. The multiple hypothesis tracking algorithm.

pothesis. Each hypothesis has an association probability, which indicates the likelihood that observed object and object hypothesis refer to the same object. In order to determine this probability the motion model is applied to the object hypothesis of the previous iteration, in order to predict where the object will be now. Then the association probability is computed by weighing the distance between the predicted and the observed object position. Thus in every iteration of the algorithm each observation is associated with each existing object hypothesis.

Our MHT algorithm is an extension of Reid's algorithm [8]. It extends Reid's version in that it can handle multiple mobile sensors with uncertain positions. The computational structure of the algorithm is shown in Fig. 1. An iteration begins with the set of hypotheses of object states $H^k = \{h_1^k, \ldots, h_m^k\}$ from the previous iteration k. Each h_i^k is a random variable ranging over the state space of a single object and represents a different assignment of measurements to objects, which was performed in the past. The algorithm maintains a Kalman filter for each hypothesis.

With the arrival of new sensor data (6), $Z(k+1) = \{z_1(k+1), \ldots, z_{n_{k+1}}(k+1)\}$, the motion model (7) is applied to each hypothesis and intermediate hypotheses $\widehat{h}_i^{k+1}$ are predicted. Assignments of measurements to objects (10) are accomplished on the basis of a statistical distance measurement, such as the Mahalanobis distance. Each subsequent child hypothesis represents one possible interpretation of the set of observed objects and, together with its parent hypothesis, represents one possible interpretation of all past observations. With every iteration of the MHT probabilities (11) describing the validity of an hypothesis are calculated. Furthermore for every observed object an new hypothesis with associated probability is created (13). The equations used for the computation of these probabilities can be found in [9].

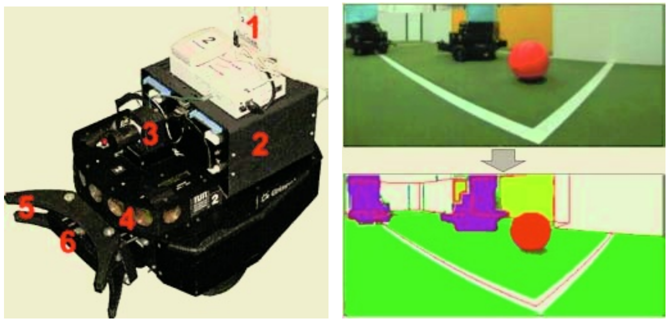
In order to constrain the growth of the hypothesis trees the algorithm prunes improbable branches (14). Pruning is based on a combination of ratio pruning, i.e. a simple lower limit on the ratio of the probabilities of the current and best hypotheses, and the N-scan-back algorithm [8]. This algorithm assumes that any ambiguity at time k is resolved by time $k + N$. Consequently if at time k hypothesis h_i^{k-1} has m children, the sum of the probabilities of the leaf notes of each branch is calculated. The branch with the greatest probability is retained and the others are discarded. After pruning the world state of X^{k-N} can be extracted (15). Please note that this world state is always N steps delayed behind the latest observations. However, this delay can be overcome by N observers performing observations in parallel.

3 Applying MHT to Autonomous Robot Soccer

Autonomous robot soccer confronts object tracking mechanisms with challenging research problems. The camera system with an opening angle of 90^o and pointed to the front gives an individual robot only a very restricted view of the game situation. Therefore, the robot needs to cooperate to get a more complete picture of the game situation. Vibrations of the camera, spot light effects, and poor lighting conditions cause substantial inaccuracies. Even small vibrations that cause jumps of only a few pixel lines cause deviations of more than half a meter in the depth estimation, if the objects are several meters away. The opponent robots change their speed and moving directions very quickly and therefore an iteration of the tracking algorithm has to be very fast such that the inaccuracies of the motion model does not have such a huge effect.

The information needed for object tracking is provided by the perception system and includes the following kinds of information: (1) partial state estimates broadcasted by other robots, (2) feature maps extracted from captured images, and (3) odometric information. The estimates broadcasted by the team mates comprise the respective robot's location and the locations of the opponents. From the captured camera images the feature detectors extract problem-specific feature maps that correspond to (1) static objects in the environment including the goal, the borders of the field, and the lines on the field, (2) a color blob corresponding to the ball, and (3) the visual features of the opponents.

The working horse of the perception component are a color classification and segmentation algorithm that is used to segment a captured image into colored regions and blobs (see Fig. 2b). The color segmented image is then processed by a feature extraction algorithm (see Fig. 3) that estimates the 2D positions and the covariances of the objects of interest. At present it is assumed that the objects are colored black and have approximately circular shape. Object detection is performed on the basis of blob analysis. The position of an object is estimated on the basis of a pinhole camera model. Due to rotations and radial distortions of the lenses this model is highly non-linear. The uncertainty estima-



(a) (b)

Fig. 2. An AGILO soccer robot (a) and an image captured by the robot and the feature map that is computed for self, ball, and opponent localization (b).

---

algorithm INTERPRETSENSORDATA($\widehat{\Phi}$, $\widehat{\mathbf{C}}_\phi$)

let $\widehat{\Phi}$ `% robot pose`
  I `% image data`
  $R = \{r_1, \ldots, r_{n_k}\}$ `% set of regions`
  ω `% augmented mean`
  $\mathbf{C}_\omega$ `% augmented covariance`
  $Z(k) = \{z_1(k), \ldots, z_{n_k}(k)\}$ `% observed feat.`
do $I \leftarrow$ GETSENSORDATA();
  $R \leftarrow$ EXTRACTBLACKREGIONS(I);
  $R \leftarrow$ CHECKCONSTRAINTS(R);
  $R \leftarrow$ EXTRACTCASCADEDROBOTS(R);
  for $i \leftarrow 1$ *to* $|R|$
  do $(row, col, width) \leftarrow$ EXTRACTFEATURES(r_i);
    $\omega \leftarrow \left[\widehat{\Phi}, row, col, width\right]^T$;
    $\mathbf{C}_\omega \leftarrow \begin{pmatrix} \widehat{\mathbf{C}}_\phi & 0 & 0 & 0 \\ 0 & \sigma_{row} & 0 & 0 \\ 0 & 0 & \sigma_{col} & 0 \\ 0 & 0 & 0 & \sigma_{width} \end{pmatrix}$;
    $z_i(k) \leftarrow$ UNSCENTEDTRANSFORM($\omega, \mathbf{C}_\omega, opp$);

Fig. 3. The Algorithm used for feature extraction and uncertainty estimation.

tion process is based on the unscented transformation [7]. This allows the use of non-linear measurement equations, the incorporation of parameters describing the measurement uncertainty of the sensor at hand as well as an efficient way of propagating the uncertainty of the observing robots pose. A detailed description of the feature extraction algorithm and uncertainty estimation process can be found in [10].

4 Empirical Investigation

The multiple object tracking algorithm described in this paper has been employed by our AGILO robot soccer team in the fifth robot soccer world championship in Seattle (2001). Our RoboCup team consists of four Pioneer I robots (see Fig. 2a). The robot is equipped with a single on board linux computer (2), a wireless Ethernet (1) for communication, and several sonar sensors (4) for collision avoidance. A color CCD camera with an opening angle of 90^o (3) is mounted fix on the robot. The robot also has a dribbling (5) and a kicking device (6) that enable the robot to dribble and shoot the ball. In Seattle the team has played six games for a total of about 120 minutes and advanced to the quarter finals.

Unfortunately, in midsize robot soccer there is no external sensing device which records a global view of the game and can be used as the ground truth for experiments. Thus for the experimental results in this section we can only use the subjective information of our robots and argue for the plausibility of their behavior and belief states. To do so, we have written log files and recorded the games using video cameras in order to evaluate our algorithm. The analysis of the log files from RoboCup 2001, revealed that an average MHT update takes between 6 to 7 msecs. This allows our implementation to process all observations of all robots (max. frame rate: 25Hz) in real time. The minimum and maximum iteration times were measured to be 1.1 msecs and 86 msecs respectively. On average the MHT tracked 3.2 opponents. This is a reasonable number since there are maximal 4 opponent players and players can be send off or have to be rebooted off field. In breaks of the games (when people get on to the field) or when there are crowds of robots the MHT successfully tracked up to 11 objects.

A typical result of the AGILO game state estimator is shown in Fig. 3. The upper picture shows the positions of the AGILO players of the own team, computed through vision-based self localization [10,5]. The middle picture shows the individual observations of the opponent robots. The opponent observations performed by the AGILO robots are indicated with circles, crosses, diamonds, and triangles. In the lower picture the tracks as they were resolved by the MHT are displayed. They are divided into subsections. The number of the robot that contributed the most observations to this part of the track is denoted next to the track.

Qualitatively, we can estimate the accuracy of the game state estimation by looking for the jumps in the tracked lines. The tracks of the opponents look very reasonable. They are less accurate and sometimes only partial. This is due to the high inaccuracy and incompleteness of the sensory data. However, it is observable that several tracks resulted from merging the observations of different robots. In addition, the merging of the different observations results in fewer hallucinated obstacles and therefore allows for more efficient navigation paths. Several wrong opponent observations made by the goal keeper (1) were correctly omitted by the MHT and not assigned to a track. We have cross checked the tracks computed by the algorithm using video sequences recorded during the matches. The tracks are qualitatively correct and seem to be accurate. A more thorough evaluation is only possibly based on the ground truth for the situations. We are currently implementing tracking software for a camera mounted above the field that allows us to compute the ground truth for the next RoboCup championship.

The cooperation of the different robots increases both, the completeness and the accuracy of state estimation. Accuracy can be substantially increased by fusing the observations of different robots because the depth estimate of positions are much more inaccurate than the lateral positions in the image. This can be accomplished through the Kalman filter's property to optimally fuse observations from different robots into global hypotheses with smaller covariances. The completeness of state estimation can be increased because all the robots can see only parts of the field and can be complemented with observations of the team mates. The other effect we observed was that cooperation allowed to maintain the identity of opponent players over an extended period of time, even though the field of view of the observing robots is limited. This point is well illustrated in Fig. 3. The three opponent field players were tracked successfully over a period of 30 seconds.

5 Related Work

Related work comprises work done on object tracking in the robot soccer domain and probabilistic and vision-based tracking of moving targets. To the best of our knowledge no probabilistic state estimation method has been proposed for tracking the opponent robots in robot soccer or similar application domains. Dietl et al. [4] estimate the positions of the opponents and store them in the team world model but they solve the correspondence problem on a rather coarse level. Probabilistic tracking of multiple moving objects has been proposed by Schulz et al. [11]. They apply sample-based JPDAF estimation to the tracking of moving people with a moving robot using laser range data. The required computational power for the particle filters is opposed by the heuristic based pruning strategies of the MHT algorithm. Hue et al.[6] are also tracking multiple objects with particle filters. In their work data association is performed on the basis of the Gibbs sampler. Our approach to multiple hypothesis tracking is most closely related to the one proposed by Cox and Miller [3]. We extend their work on multiple hypothesis tracking in that we apply the method to a much more challenging application domain where we have multiple moving observers with uncertain positions. In addition, we perform object tracking at an object rather than on a feature level.

6 Conclusions

In this paper, we have extended and analyzed a probabilistic object tracking algorithm for a team of vision-based autonomously moving robots. Our results suggest that purely image-based probabilistic estimation of complex game states is feasible in real time even in complex and fast changing environments. We have also seen that maintaining trees of possible tracks is particularly useful for estimating a global state based on multiple mobile sensors with position uncertainty.

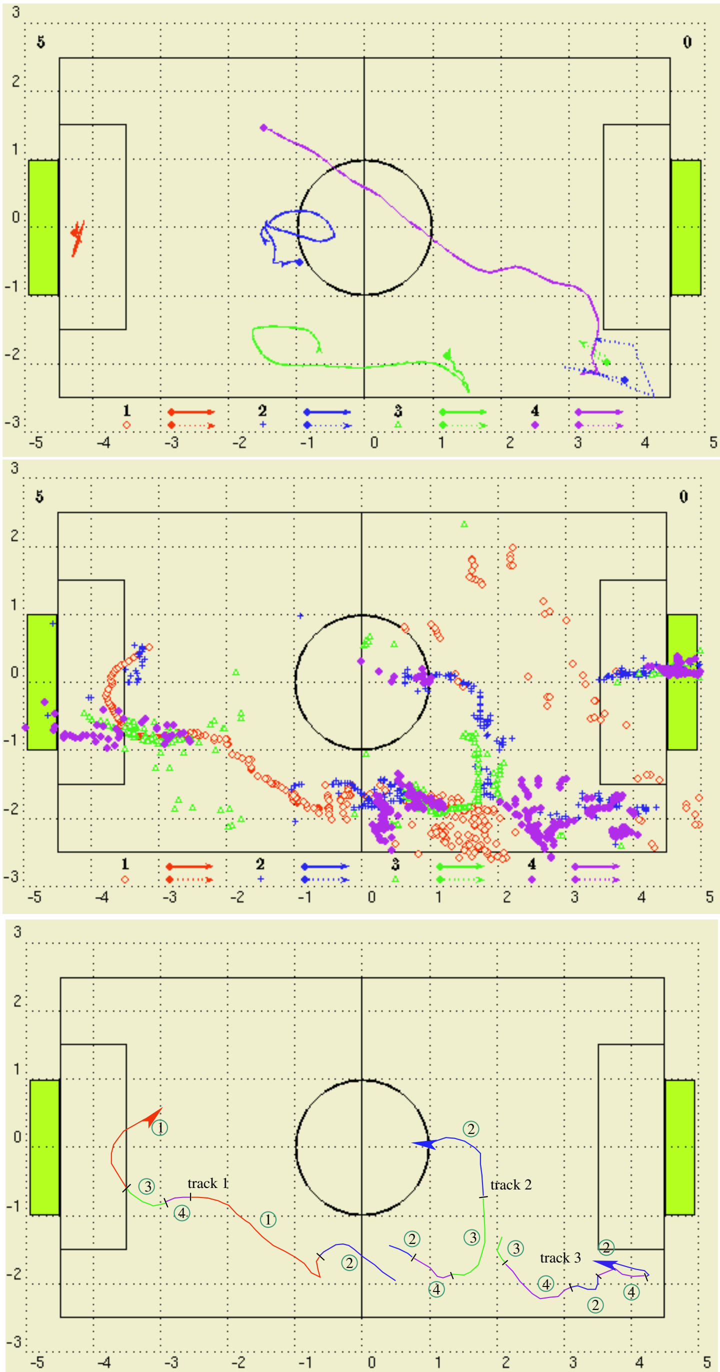


Fig. 4. Opponent observations and resolved tracks.

Finally, we have seen how the state estimation modules of individual robots can cooperate in order to produce more accurate and reliable state estimation. Besides an empirical analysis of the parameter settings and learning accurate sensing models, we intend to compare in future work the MHT algorithm with the JPDAF implementation of [11].

References

1. Y. Bar-Shalom and T. Fortmann. Tracking and data association. Academic Press., 1988.
2. M. Beetz, S. Buck, R. Hanek, T. Schmitt, and B. Radig. The AGILO autonomous robot soccer team: Computational principles, experiences, and perspectives. In *Procs. of the First International Conference on Autonomous Agents and Multi-agent Systems*, 2002. to appear.
3. I.J. Cox and S.L. Hingorani. An Efficient Implementation of Reid's Multiple Hypothesis Tracking Algorithm and Its Evaluation for the Purpose of Visual Tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(2):138–150, February 1996.
4. M. Dietl, J.-S. Gutmann, and B. Nebel. Cooperative sensing in dynamic environments. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 1706–1713, Maui, Hawaii, 2001.
5. R. Hanek and T. Schmitt. Vision-based localization and data fusion in a system of cooperating mobile robots. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 1199–1204, Takamatsu, Japan, 2000.
6. C. Hue, J.-P. Le Cadre, and P. Perez. Tracking multiple objects with particle filtering. Technical Report 1361, IRISA, 2000.
7. S. Julier and J. Uhlmann. A new extension of the kalman filter to nonlinear systems. The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls., 1997.
8. D. Reid. An algorithm for tracking multiple targets. IEEE Transactions on Automatic Control, 24(6):843–854, 1979.
9. T. Schmitt, M. Beetz, R. Hanek, and S. Buck. Watch their moves: Applying probabilistic multiple object tracking to autonomous robot soccer. In *AAAI National Conference on Artificial Intelligence*, page to appear, Edmonton, Canada, 2002.
10. T. Schmitt, R. Hanek, S. Buck, and M. Beetz. Cooperative probabilistic state estimation for vision-based autonomous mobile robots. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 1630–1638, Maui, Hawaii, 2001.
11. D. Schulz, W. Burgard, D. Fox, and A.B. Cremers. Multiple object tracking with a mobile robot. In *Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 371–377, Kauai, Hawaii, 2001.
12. S. Thrun. Probabilistic algorithms in robotics. *AI Magazine*, 21(4):93–109, 2000.