RAO-BLACKWELLIZED RESAMPLING PARTICLE FILTER FOR REAL-TIME PLAYER TRACKING IN SPORTS

Nicolai v. Hoyningen-Huene and Michael Beetz
Intelligent Autonomous Systems Group, Technische Universität München, Boltzmannstr. 3, D-85748 Garching, Germany

Keywords: Multi-target tracking, Particle filter, Rao-Blackwellization, Kalman filter, Resampling.

Abstract: Tracking multiple targets with similiar appearance is a common task in computer vision applications, especially in sports games. We propose a Rao-Blackwellized Resampling Particle Filter (RBRPF) as an implementable real-time continuation of a state-of-the-art multi-target tracking method. Target configurations are tracked by sampling associations and solving single-target tracking problems by Kalman filters. As an advantage of the new method the independence assumption between data associations is relaxed to increase the robustness in the sports domain. Smart resampling and memoization is introduced to equip the tracking method with real-time capabilities in the first place. The probabilistic framework allows for consideration of appearance models and the fusion of different sensors. We demonstrate its applicability to real world applications by tracking soccer players captured by multiple cameras through occlusions in real-time.

1 INTRODUCTION

Tracking multiple targets is needed in a lot of computer vision applications like surveillance or sports analysis. The sports domain provides a challenging testbed for concurrent tracking of multiple targets with similar appearance through frequent occlusions measured from different views. In team sports the complex coordination of movements of different players are crucial to the success of a squad. For automated analysis thereof the correct association of players to movements is equally important as the recognition of the movement itself.

To achieve an automatic extraction of athlete positions during sports games from video streams, beside camera estimation and player segmentation, a robust and fast multi-target tracking method is needed. In sports games the number of players is usually known and constant. In contrast the number of observations for each player obtained by measurements from sensors or segmentation for videos varies; it ranges from zero in case of occlusion and oversight to several measurements in case of hallucination and inaccuracy of the player extraction. Players usually differ by appearance from the field to help viewers and referees to follow the game easily, so the association of players of one team with their individual name is the bigger problem.

In this paper we propose a Rao-Blackwellized Resampling particle filter (RBRPF) for real-time tracking of multiple targets. Particles are represented as configurations of all players to result in tracking a mixture of Gaussians, where the multi-modality is caused by possible mix-ups of associations and the Gaussian refers to the uncertainty of dynamics. Sampling of new target configurations is reduced to sampling associations and Rao-Blackwellized by using the Kalman filter. Taking advantage of the fact that the number of probable associations for given player positions and measurements are usually low, the particle filter focuses on the most likely associations and can avoid unnecessary computations by smart resampling and memoization. The Bayesian framework allows the integration of kinematic and appearance models to determine the most probable player locations through occlusions.

Our contributions are the enhancements of a state-of-the-art theoretical multi target tracking method towards an implementable real-time algorithm that performs well in the demanding sports domain. We relax the independence assumption of single measurement associations to suit the original method to the application domain and achieve more robustness. Further we invent a smart resampling procedure that allows real-time in the first place and adapts to the complexity of the tracking problem. The proposed memoization of

v. Hoyningen-Huene N. and Beetz M. (2009).
RAO-BLACKWELLIZED RESAMPLING PARTICLE FILTER FOR REAL-TIME PLAYER TRACKING IN SPORTS.
In *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pages 465-472
DOI: 10.5220/0001788804650472

repeatedly computed results additionally improves efficiency.

The method is developed as part of the ASPOGAMO system (Beetz et al., 2006; Beetz et al., 2007), that aims to extract knowledge from broadcasted soccer games, and is evaluated by applying it to real soccer games, showing robust real-time performance over challenging sequences.

The remainder of this paper is organized as follows. We briefly talk about related work in the next section. In section 3 we derive the Rao-Blackwellized Resampling particle filter. Section 4 describes the experiments we conducted. We finish in section 5 with our conclusions.

2 RELATED WORK

Multiple-target tracking algorithms can be differentiated by their data association methods. Multiple hypothesis tracking (MHT) (Bar-Shalom et al., 2001) builds a pruned tree of all possible association sequences of each measurement with close targets by the Hungarian algorithm. The assumption of single associations and the use of Kalman filtering allow computation in polynomial time, but inhibit to handle multiple or merged associations. Khan et al., 2006 (Khan et al., 2006) propose a real-time Rao-Blackwellized MCMC-based particle filter where associations are sampled by a Markov chain. The Markov chain allows also sampling of merged measurement assignments but demands computation time that reduces the number of particles. In their experiments real-time could only be provided for a small number of particles (less than 6) i.e. the tracker can cope with three parallel mix-ups of targets max. Interaction of targets are modeled as correlations between target positions which does not hold for many applications.

The Rao-Blackwellized particle filter approach by Särkkä et al., 2004 (Särkkä et al., 2004; Särkkä et al., 2007) samples the associations directly and handles dependencies between them by data associations priors. The performance of the method was demonstrated only on synthetic simulations without statements about computation time. Our approach is an extension of this method to real world applications introducing smart resampling and memoization that leads to real-time tracking in the first place and relaxation of the association independence assumption.

Tracking of soccer players is classified by Li et al., 2005 (Li et al., 2005) in category and identity tracking. Category tracking extracts trajectories with team affiliation where in the other case each single player is traced with its identity. Barceló et al., 2005 (Barceló et al., 2005) and Figueroa et al., 2006 (Figueroa et al., 2006) label the measurements by nearest neighbor assignment. In Gedikli et al., 2007 (Gedikli et al., 2007) MHT was applied, but Particle filters constitute the mostly used method in the literature for category tracking (i.e. (Yang et al., 2005; A.Dearden and Grau, 2006)). Du et al., 2006 (Du et al., 2006; Du and Piater, 2007) aim on combining local particle filters to fuse measurements captured from different views. A MCMC method for team labelling is proposed by Liu et al., 2007 (Liu et al., 2007) to link observations of soccer players over time.

Identity tracking is often performed in a second stage by consistent labelling of the trajectory graph generated by category tracking. Huang and Hilton, 2006 (Huang and Hilton, 2006) propose an assignment in batch mode by shortest path algorithm, Nillius et al., 2006 (Nillius et al., 2006) solve the association of the trajectory graph by Bayesian network inference, and Sullivan and Carlsson, 2006 (Sullivan and Carlsson, 2006) combine trajectories of unoccluded players in a graph structure by clustering. Barceló et al., 2005 (Barceló et al., 2005) resolve collisions of nearest neighbor Kalman tracking by constraints in the trajectory graph. To the best of our knowledge no real-time identity tracking method for soccer player that allows multiple measurements and fuses different camera views was proposed in the literature yet.

3 RAO-BLACKWELLIZED RESAMPLING PARTICLE FILTER

A particle filter for complete player configurations constitutes the base of our algorithm. New particles are predicted by sampling associations of players with current measurements considering dependencies between them. Computation time is spend mostly on the highly probable configurations and on ambiguous associations by memoization of precomputed samples and probabilities. Sampling and weighting is done by using the Kalman filter for Rao-Blackwellization of the particle filter.

3.1 Bayesian View of Tracking

The problem of tracking is to recursively estimate a state x_k knowing the evolution of the state sequence

$$x_k = f_k\left(x_{k-1}, v_{k-1}\right) \tag{1}$$

from measurements

$$z_k = h_k(x_k, n_k) \tag{2}$$

where f_k is called system or motion model and h_k is called measurement model, v_{k-1} and n_k denote the process and measurement noise, respectively. The tracked state x_k is represented as the configuration of all player states

$$x_k = \left\{ x_{j,k} = \mathcal{N}\left(\begin{pmatrix} x \\ y \\ \dot{x} \\ \dot{y} \end{pmatrix}; m^i_{j,k}, V_{j,k} \right) \right\} j = 1, \ldots, T \tag{3}$$

where $x_{j,k}$ contains the position and velocity of player j at time k. An individual target sample $x^i_{j,k}$ is assumed to be Gaussian with mean $m^i_{j,k}$ and corresponding covariance matrix $V^i_{j,k}$.

In a Bayesian framework, the problem of tracking can be formulated as one of estimating the *posterior* probability density function $p(x_k|z_{1:k})$ for the state x_k given a sequence of measurements $z_{1:k}$ up to time k.

3.2 Particle Filtering

In *Sampling Importance Sampling* (SIS) particle filtering, the posterior probability density function is approximated by a weighted sum of random samples x^i_k also called particles (Arulampalam et al., 2002). The weights are normalized such that $\sum_i w^i_k = 1$:

$$p(x_k|z_{1:k}) \approx \sum_i w^i_k \delta\left(x_k - x^i_k\right). \tag{4}$$

We draw samples x^i_k by importance sampling from a proposal $q(.)$ called an importance density. Doucet, 1998 (Doucet, 1998) showed that the optimal importance density function that minimizes the variance of the true weights conditioned on x^i_{k-1} and z_k is

$$q_{opt}\left(x_k|x^i_{k-1}, z_k\right) = \frac{p\left(z_k|x_k, x^i_{k-1}\right) p\left(x_k|x^i_{k-1}\right)}{p\left(z_k|x^i_{k-1}\right)}. \tag{5}$$

In our case the importance density is the probability distribution of data associations, while the actual sample is deduced from an association by the use of Kalman filtering.

3.3 Sampling New Configurations

For known associations between measurements z_k and players of the sample x^i_{k-1} the new sampled configuration is Gaussian and can be evaluated analytically as an optimal fusion between measurements and predicted player positions. The Kalman filter provides the method to find the Gaussian that equals both probabilities in the numerator of equation 5 and thus maximizes their product. The sampling problem reduces therefor to sample associations between measurements and the predicted player configuration and solving multiple single target tracking problems by Kalman filtering. The analytical sampling part forms the Rao-Blackwellization of the particle filter. To supply an optimal solution the Kalman filter assumes state and measurement noise to be zero-mean, white Gaussian and the measurement as well as the motion model to be linear. If the last assumption does not hold an extended or unscented Kalman filter could be applied for a suboptimal solution. Following this approach the posterior probability density function of configurations form a mixture of Gaussians, where the multi-modality originates from ambiguities in the associations.

3.3.1 Predicting by System Model

We can sample from $p\left(x'_k|x^i_{k-1}\right)$ analytically by the Kalman prediction step according to the system dynamics of eq. 1. Each player state is predicted independently using the discretized Wiener velocity model $A_{\Delta t}$ (Bar-Shalom et al., 2001) for time difference Δt between $k-1$ and k as a linear motion model:

$$m'_{j,k} = \begin{pmatrix} x'_{j,k} \\ y'_{j,k} \\ \dot{x}'_{j,k} \\ \dot{y}'_{j,k} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^i_{j,k-1} \\ y^i_{j,k-1} \\ \dot{x}^i_{j,k-1} \\ \dot{y}^i_{j,k-1} \end{pmatrix} \tag{6}$$

The covariance matrix evolves to

$$V'_k = A_{\Delta t} V_{k-1} A^T_{\Delta t} + \begin{pmatrix} \frac{\Delta t^3}{3} & 0 & \frac{\Delta t^2}{2} & 0 \\ 0 & \frac{\Delta t^3}{3} & 0 & \frac{\Delta t^2}{2} \\ \frac{\Delta t^2}{2} & 0 & \Delta t & 0 \\ 0 & \frac{\Delta t^2}{2} & 0 & \Delta t \end{pmatrix} \tilde{q} \tag{7}$$

with power spectral density $\tilde{q}$ as a constant factor.

3.3.2 Sampling Associations

We introduce associations

$$J_k : \{1, \ldots, |z_k|\} \to \wp(\{1, \ldots, T\}) \tag{8}$$

as mappings from all measurements at time k to a (possibly empty) subset of all targets. We denote $\hat{J}_k$ as the inverse mapping from targets to their assigned observations for convenience. The space of data associations equals the finite and discrete set of all possible associations of measurements to targets containing $2^{|z_k| \times T}$ elements. If we restrict the data associations J_k to assign a measurement to one target max, the number of possible associations reduce to $(T+1)^{|z_k|}$. We can further reduce this number to $\sum_{i=0}^{\min(T,|z_k|)} \binom{\min(T,|z_k|)}{i} \max(T,|z_k|)^{\min(T,|z_k|)-i}$ if we prohibit multiple measurements per target, also called exclusion principle (MacCormick and Blake, 1999). Enumerating this set and solving each single target tracking problem is still intractable even for a small number of targets and measurements. Fortunately only a few associations have high probability, but to sample them efficiently, we have to assume the associations for single measurements to be independently or the dependency between them to be determined fast.

Individual Independent Associations. If we look at sampling an individual association for measurement $z \in z_k$, we can enumerate all possible assignments easily as z can be clutter viz. a false alarm or assigned to one of the players. Thus the importance distribution $\pi(z)$ for an association of a specific measurement z can be evaluated by normalizing the probabilities $\hat{\pi}(z)$ for each possible association.

Clutter measurements are assumed to be independent from player positions and uniformly distributed in the measurement space with volume M

$$\hat{\pi}_{\varnothing}(z) = p\left(J_k(z) = \varnothing | z_k\right) \sim \frac{1}{M}. \tag{9}$$

The probability for a data association between target t and an observation z is up to a constant factor:

$$\begin{aligned}\hat{\pi}_t(z) &= p\left(t \in J_k(z) | z_k, x'_k\right) \\ &\sim p_{app}\left(t \in J_k(z)\right) N\left(z; H_z m'_{t,k}, H_z^T V'_{t,k} H_z + R_z\right)\end{aligned} \tag{10}$$

with measurement model $H_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ and R_z as measurement noise covariance. $p_{app}\left(t \in J_k(z)\right)$ denotes the propability of an association based on the appearance model only, which is independent from player and measurement positions. The Gaussian in the second part refers to the probability of the association by the kinematic model. We included the appearance model in difference to (Särkkä et al., 2007) to allow a realistic influence of additional information from segmentation beside spatial data only.

Importance Density. Utilizing the independence of single associations the importance density for a sampled state x_k^j can be computed as a product over probabilities of assignments for each single measurement that are given by the normalized importance distribution π of equations 9 and 10.

$$q\left(x_k | x_{k-1}^i, z_k\right) = \prod_{z_k} \pi \tag{11}$$

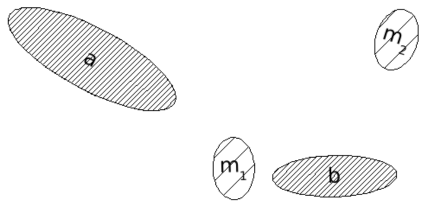

Figure 1: Association of a and m_1 increases the probability of b and m_2 being associated.

Relaxation of Independence. In the underlying method by Särkkä et al., 2004, the measurements are processed one at a time in sequential fashion based on the independence assumption of associations of individual measurements. This assumption does not always hold, the order of associations often do matter. This can be best exemplified by figure 1 assuming that measurements can be assigned to one target at max: If target a is assigned to measurement m_1, the probability of the association of m_2 and target b increases.

Särkkä et al., 2007 did not consider this problem at all but proposed the use of an data association prior. We follow this solution instead of establishing an additional Markov Chain as proposed by Khan et al., 2006 (Khan et al., 2006) in favor of computational efficiency but change the procedure slightly to improve robustness against violation of the mentioned assumption. To generate new particles x_k^j including the whole player configuration, we repeatedly sample an ordering on the measurements of one sweep uniformly at random, reducing the relevance of the ordering and the induced dependencies on the tracking result. With the randomly sampled ordering we draw an association for each measurement with the normalized importance distribution $\pi(z)$ one at a time. If a target was associated, it is excluded from further associations with the single detection probability p_{sd} and the importance distribution is renormalized. If the mentioned exclusion principle holds i.e. targets can be assigned to one measurement at max, p_{sd} should be set to one.

Determination of State from Associations. For sampled associations J_k the predicted player positions x'_k can be updated individually by Kalman update with their observations

$$x_{j,k}^i = x'_{j,k} + V'_{j,k} H^T \left(H V'_k H^T + R\right)^{-1} \left(\hat{J}_k(j) - H x'_{j,k}\right), \tag{12}$$

with H denoting the linear measurement model (2) as H_z stacked $|\hat{J}_k(j)|$ times and R as diagonal matrix of measurement covariances of observations $\hat{J}_k(j)$.

3.4 Weighting

For a good performance of the particle filter the computation of the weights of each sampled state is crucial. To approximate $p(x_k|z_{1:k})$ correctly the weights w_k^i have to be defined recursively as

$$w_k^i \propto w_{k-1}^i \frac{p\left(z_k|x_k^i\right) p\left(x_k^i|x_{k-1}^i\right)}{q\left(x_k^i|x_{k-1}^i, z_k\right)}. \tag{13}$$

The denominator was already computed in the sampling phase and was depicted in equation 11. The likelihood of the measurements given the sampled state x_k^i with known associations and the likelihood of x_k^i given the former state x_{k-1}^i and the dynamics can be computed for each player and measurement separately. The measurement likelihood can be computed analogously to eq. 10 but substituting x_k' by x_k^i and V_k' by V_k^i, respectively:

$$p\left(z_k|x_k^i\right) = \prod_{z \notin \hat{J}_k^i} \hat{\pi}_\varnothing(z) \prod_j p\left(\hat{J}_k^i(x_{j,k}^i)|x_{j,k}^i\right). \tag{14}$$

The likelihood of the new sample according to the motion model can be computed by reusing the already predicted state x_k' of eq. 6

$$p\left(x_k^i|x_{k-1}^i\right) = \prod_j N\left(m_{j,k}^i; m_{j,k}', V_{j,k}^i + V_{j,k}'\right). \tag{15}$$

3.5 Resampling

SIS particle filters suffer from the so called degeneracy phenomenon, where only a small amount of all particles have not negligible weights. This implies that most of the computation time will be spent on particles that contribute only marginally to the approximation of the posterior probability density function of equation 4. To reduce the degeneracy problem resampling has been proposed to eliminate particles with small weights and clone the others according to their weights. We include the resampling step by sampling $w_{k-1}^i \times N_{max}$ associations for particle x_{k-1}^i. Particles with larger weights will therefor allocate more particles in the next time step, while particles with small weights are dropped.

Sampling several times from the same particle the number of distinct sampled particles will approach the number of ambiguities in the associations because a specific assignment leads to the same sampled configuration. Due to their discreteness there are usually only a small number of distinct probable associations. This allows a chance for noticeable improvement in computation time by smart memoization. Caching and testing sampled associations for equality can save computation time considering not only the update to generate a new state of equation 12 but also the prediction in the next particle filtering step in equation 6.

After resampling the weights are usually reset to $w_k = 1/N_{max}$ to reflect the equal probability of all particles. In our case we count the times n_{J_k} the same association J_k was sampled for a specific particle and provide only one single particle for the next filtering step having the weight set to $w_k = n_{J_k}/N_{max}$. Then the weights are recursively updated as in equation 13 and normalized at the end of the filtering step. The actual number of particles can therefor vary between 1 and N_{max} using more particles in situations with high association ambiguities. This smart resampling reduces the computation time and allows real time in the first place.

3.6 Estimate of the State

An estimate of the player positions at time k i.e. of the state x_k can be found by either selecting the particle with maximum weight or by clustering the particles and taking the weighted mean of the most probable cluster. Calculating the weighted mean of all particles should not be considered here because it can lead to the so called ghost phenomenon for multi-modal distributions i.e. it leads to a state estimated as the mean of two modes that is known to be wrong.

3.7 Implementation

The complete algorithm is depicted in figure 2 following the derivation of the former section. The individual importance distributions π as well as $\hat{\pi}$ and the Kalman prediction and updates are cached for reuse in the next sampling iteration to improve efficiency. The importance distribution, all probabilities and weights are calculated in log-space to avoid numerical problems.

4 EXPERIMENTAL RESULTS

The proposed tracking method is evaluated as part of the ASPOGAMO system (Beetz et al., 2006; Beetz et al., 2007), that aims to extract knowledge from broadcasted soccer games. ASPOGAMO is able to track multiple dynamic pan-tilt-zoom cameras and segment the soccer players and referee by a combination of variance filter and color templates. Segmentation influences the tracking process as the Kalman filters smooth assigned measurements, quality evaluation of the used method can be found in (Beetz et al., 2007). However segmentation by background subtraction for

$\left[\{x_k^i, w_k^i\}_{i=1}^{N_k}\right] = RBRPF\left[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_{k-1}}, z_k\right]$
$N_k = 0$
FOR $i = 1 : N_{k-1}$
  Predict x_k' as in 6
  $C = \varnothing$
  FOR $j = 1 : (w_{k-1}^i \times N_{max})$
    Sample an association J_k:
      $\tau = \{1, \ldots, T\}$
      Init $J_k : \forall p \in \tau . \hat{J}_k(p) = \varnothing$
      Reorder measurements z_k randomly
      FOR $l = 1 : |z_k|$
        Compute $\hat{\pi}()$ as in 9 and 10
        $\pi =$ normalized $\hat{\pi}$
        Draw association for lth measurement with player $p \in \tau$ by π
        $J_k(p) = J_k(p) \cup \{l\}$
        IF random$(0,1) < p_{sd}$: $\tau = \tau \setminus p$ and renormalize π
      END FOR
    IF J_k not in C
      $N_k = N_k + 1$
      $n_{J_k} = 1$
      Compute $x_k^{N_k}$ by Kalman update if not done previously as in 12
      $\hat{w}_k^{N_k} = \frac{1}{N_{max}}$
      Update $\hat{w}_k^{N_k}$ as in 13
      $C = C \cup \{J_k\}$
    ELSE
      $n_{J_k} = n_{J_k} + 1$
      $\hat{w}_k = \hat{w}_k \frac{n_{J_k}}{n_{J_k} - 1}$
    END IF
  END FOR
END FOR
Calculate total weight: $t = \sum_{j=1}^{N_k} \hat{w}_k^j$
FOR $j = 1 : N_k$
  Normalize: $w_k^j = t^{-1} \hat{w}_k^j$
END FOR

Figure 2: Algorithm for one iteration of the proposed Rao-Blackwellized Resampling particle filter.

static cameras is usually of higher quality. Digital videos captured by two dynamic cameras with a frame rate of 25Hz provide the basic raw material. Tracking results in both camera perspectives are depicted in figures 3 and 4 and are presented quantitatively in table 2. The extracted players spatial measurements of each camera are fused by the proposed tracking algorithm as different measurement sweeps with same time stamps.

Player positions have been measured in meters and were initialized manually in the image with covariance $V_0 = 2I_4$, initial velocity was set to zero. The factor for the kinematic process noise $\tilde{q} = 0.0008$ is derived from maximal human speed. The probability of multiple observations for the same target was obtained experimentally to $p_{sd} = 0.92$. A confusion matrix between different categories was used as a simple appearance model p_{app} and is depicted in table 1. The measurement space is determined by the number of pixels in each camera frame and evaluates to $M = 720 \times 576$. We used $N_{max} = 50$ particles to track all of the 22 players and the referee.

There is no ground truth for broadcasted soccer games because players can be tracked only visually and camera parameters are unknown. We abandon to present a spatial error as this is influenced mainly by camera estimation and segmentation. Instead we tried to find a error measure that is related with the number of false associations. A failure was counted when the projected player position differed from the real player in the image by more than 10 pixels for longer than 3 frames. In this case the tracker was reset in the failed player positions and run again on the rest of the sequence. We tracked both camera views separately and also ran the same sequence fusing the measurements of both perspectives. Because the broadcasted high-angle camera shows only a part of the field and is panning and zooming fast, in average only 9.9 players are visible (with standard deviation of 3.2). We splitted the number of failures into association errors and assigning emerging players (second number) to be comparable with the other results. The second row of table 2 shows the number of frames that were tracked in the according experiment. The computation time was taken for one update step, where all experiments have been conducted on a 2.2 GHz Dual-core PC. We think the actual needed time is more significant than the theoretical complexity of the algorithm since the input data do not scale but stay in fixed boundaries (number of players is 22, number of measurements usually lower than 200). The last row depicts the average number of particles and the corresponding standard deviation. Table 2 clearly evidences the real-time tracking ability of the proposed method with low failure rate for single cameras. Fusion of different cameras reduces the occurrence of occlusions and therewith failure rate and number of particles even further.

The fourth experiment states a challenging sequence including several fouls and header duels where kinematic and appearance model have often been to weak to differentiate between players causing a higher number of failures. The amount of measurements (lower for the highangle view) correlates obviously with the number of particles and the computation time demonstrating the adaptiveness of the proposed method to the complexity of the tracking problem. Also we observed assignment errors if segmen-

Table 1: Confusion matrix between different categories.

	Italy	France	Referee
Italy	0.6	0.1	0.3
France	0.1	0.8	0.1
Referee	0.3	0.1	0.6

Table 2: Tracking performance on the final of the world cup 2006.

Game	Frames	Fail	Time(ms)	Particles
Tactical	1262	13	23.3± 4	43.5±10
Highangle	1262	7+54	8.5± 5	12.1±12
Fused	1262	11	30.2±20	33.3±16
Fused II	3202	98	33.4±18	34.1±14

tation could not extract a specific player for longer than 20 frames (e.g. fouled player on the ground).

We also implemented the method as proposed by Khan et al., 2006 (Khan et al., 2006) and tested it on the World Cup final. We encountered problems of two kind: low variance in sparse particles and misleading interaction handling. The real-time requirement allowed only a small number of particles (6 in our case) which had a low variance because the Markov chain converged to very similar associations. This misled the tracker to remember the most probable configuration only which often did not equal the true positions. Interactions are handled by dependencies in the positions via symmetric entries in the configuration covariance matrix. This modeling is inappropriate for interacting soccer players, where e.g. the player on the ball shows contrary motion to his competitor. Both drawbacks resulted in poor tracking performance for the inspected soccer game sequences.

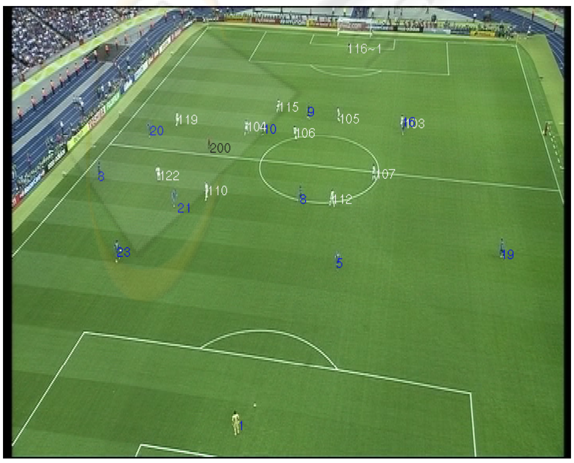

Figure 3: Tactical camera view of the World Cup final 2006.

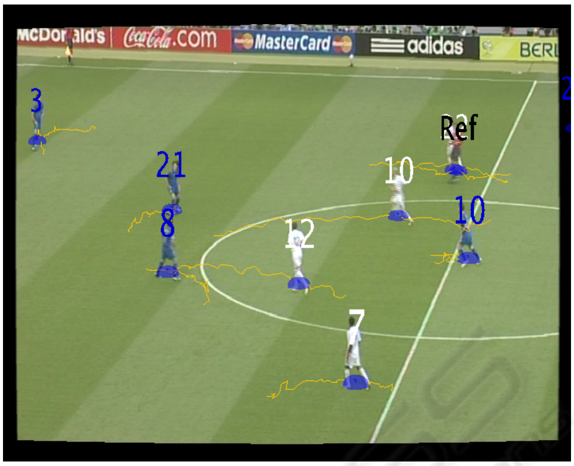

Figure 4: Identity tracking of soccer players in the broadcasted highangle camera view of the World Cup final 2006.

5 CONCLUSIONS

In this article we have proposed a real-time multiple target tracking method based on Rao-Blackwellized Resampling particle filtering for tracking soccer player identities. We presented the necessary extensions of an so far only theoretically evaluated state-of-the-art multi-target tracking method to handle real tracking problems being as challenging as in the sports domain. The first extension comprises the processing of measurements of one sweep instead of one at a time to relax the independence assumption of associations. Secondly, smart resampling and memoization was introduced to equip the tracking method with real-time capabilities. Experimental results demonstrate robustness and real-time performance of the developed method in challenging soccer game sequences including increased achievements by fusion of measurements from different cameras. A comparison with another recent multi-target tracking method explains the supremacy of our approach for the soccer domain. For future research we plan to examine more complex appearance models for automatic reinitialization of the identities especially regarding broadcasted single view sports videos.

ACKNOWLEDGEMENTS

This work was partially funded by the German Research Foundation DFG.

REFERENCES

A.Dearden, Y. and Grau, O. (2006). Tracking football player movment from a single moving camera using particle filters. In *European Conf. on Visual Media Production (CVMP 2006)*.

Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for on-line nonlinear/non-gaussian bayesian tracking. *IEEE Trans. on Signal Processing*, 50(2).

Bar-Shalom, Y., Li, X.-R., and Kirubarajan, T. (2001). *Estimation with Applications to Tracking and Navigation*. Wiley Interscience.

Barceló, L., Binefa, X., and Kender, J. R. (2005). Robust methods and representations for soccer player tracking and collision resolution. In *Proc. of Intl. Conf. on Image and Video Retrieval (CIVR 2005)*, pages 237–246.

Beetz, M., Bandouch, J., Gedikli, S., von Hoyningen-Huene, N., Kirchlechner, B., and Maldonado, A. (2006). Camera-based observation of football games for analyzing multi-agent activities. In *Proc. of Intl. Joint Conf. on Autonomous Agents and Multiagent Systems (AAMAS)*.

Beetz, M., Gedikli, S., Bandouch, J., Kirchlechner, B., von Hoyningen-Huene, N., and Perzylo, A. (2007). Visually tracking football games based on tv broadcasts. In *Proc. of Intl. Joint Conf. on Artificial Intelligence (IJCAI)*.

Doucet, A. (1998). On sequential Monte Carlo methods for Bayesian filtering. Technical report, Dept. End., Univ. Cambridge, UK.

Du, W., Hayet, J.-B., Piater, J., and Verly, J. (2006). Collaborative multi-camera tracking of athletes in team sports. In *Workshop on Computer Vision Based Analysis in Sport Environments (CVBASE)*, pages 2–13.

Du, W. and Piater, J. H. (2007). Multi-camera People Tracking by Collaborative Particle Filters and Principal Axis-Based Integration. In *Asian Conference on Computer Vision*, number 4843 in LNCS, pages 365–374. Springer.

Figueroa, P. J., Leite, N. J., and Barros, R. M. L. (2006). Tracking soccer players aiming their kinematical motion analysis. *Computer Vision and Image Understanding*, 101(2):122–135.

Gedikli, S., Bandouch, J., von Hoyningen-Huene, N., Kirchlechner, B., and Beetz, M. (2007). An Adaptive Vision System for Tracking Soccer Players from Variable Camera Settings. In *Proc. of Intl. Conf. on Computer Vision Systems (ICVS)*.

Huang, P. and Hilton, A. (2006). Football player tracking for video annotation. In *European Conf. on Visual Media Production*.

Khan, Z., Balch, T., and Dellaert, F. (2006). Mcmc data association and sparse factorization updating for real time multitarget tracking with merged and multiple measurements. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(12):1960–1972.

Li, Y., Dore, A., and Orwell, J. (2005). Evaluating the performance of systems for tracking football players and ball. In *IEEE Intl. Conf. on Advanced Video and Signal Based Surveillance*.

Liu, J., Tong, X., Li, W., Wang, T., Zhang, Y., Wang, H., Yang, B., Sun, L., and Yang, S. (2007). Automatic player detection, labeling and tracking in broadcast soccer video. In *British Machine Vision Conference*.

MacCormick, J. and Blake, A. (1999). A probabilistic exclusion principle for tracking multiple objects. In *Proc. of Intl. Conf. on Computer Vision (ICCV)*, pages 572–578.

Nillius, P., Sullivan, J., and Carlsson, S. (2006). Multi-target tracking - linking identities using bayesian network inference. In *Proc. of Computer Vision and Pattern Recognition*, pages 2187–2194.

Särkkä, S., Vehtari, A., and Lampinen, J. (2004). Rao-blackwellized monte carlo data association for multiple target tracking. In *Proc. of Intl Conf. on Information Fusion*, volume 7, Stockholm.

Särkkä, S., Vehtari, A., and Lampinen, J. (2007). Rao-blackwellized particle filter for multiple target tracking. *Information Fusion Journal*, 8(1):2–15.

Sullivan, J. and Carlsson, S. (2006). Tracking and labelling of interacting multiple targets. In *Proc. of European Conf. on Computer Vision*, pages 619–632.

Yang, C., Duraiswami, R., and Davis, L. (2005). Fast multiple object tracking via a hierarchical particle filter. In *Proc. of Intl. Conf. on Computer Vision*, volume 1, pages 212–219.