# Collaborative Multi-Camera Tracking of Athletes in Team Sports


Wei Du, Jean-Bernard Hayet, Justus Piater, and Jacques Verly

University of Liège, Department of Electrical Engineering and Computer Science,
Institut Montefiore, Sart Tilman Campus, Bldg. B28, B-4000, Liège, Belgium
weidu@montefiore.ulg.ac.be, {jean-bernard.hayet, justus.piater, jacques.verly}@ulg.ac.be



**Abstract.** A novel approach to tracking athletes in team sports using multiple cameras is proposed that addresses several issues including occlusions and propagation of wrong information. The strength of this approach lies in the use of belief propagation which enables good observations in some views to compensate for poor observations in other views due e.g. to occlusions. Each target is tracked in each view by a dedicated particle-filter-based local tracker. The trackers in different views interact with each other via belief propagation so that a local tracker operating in one view is able to take advantage of additional information from other views. By combining particle filters and belief propagation in a unified framework, we develop a sequential belief propagation algorithm to perform inference of multi-view target states collaboratively. We demonstrate the effectiveness of our method on sequences of soccer games.


## 1 Introduction

Athlete tracking is a basic task in sports video analysis that provides quantitative data for high-level processing such as tactics analysis and highlight extraction. Applications range from golf [1] and tennis [2] to team sports such as soccer [3, 4], american football [5], and hockey [6], many of which involve tracking athletes with multiple overlapping cameras separated by wide baselines. In this paper, we focus on team-sports scenarios where athletes move on a ground plane and pan-tilt-zoom (PTZ) cameras are used.

Our objective is to estimate the trajectories of athletes—in particular, players in soccer games—using observations from multiple cameras. Several issues need to be addressed:

- Occlusion: Occlusions often cause incomplete or missing image observations. In sports scenarios, due to spatial proximity of similar objects, the even more difficult situation of spurious observations arises. For instance, a soccer player is often occluded by teammates, and these players are in general not easily distinguishable from each other.


This work was sponsored by the Région Wallonne under DGTRE/WIST contract 031/5439 (TRICTRAC Project).

– Wrong information propagation: One important issue in multi-camera tracking is the collaboration among multiple views. With a proper collaboration scheme, good observations in some views can compensate for poor observations in other views due e.g. to occlusions. A less sophisticated collaboration scheme may cause wrong information to be propagated across views so that the failure of one tracker breaks the whole system.

A new stochastic approach is proposed to solve these problems. In the multi-camera tracking context, target states in different views and in 3D are represented by different but dependent random variables. The conditional dependence between them is expressed by graphical models [7]. To infer the multi-view target states based on the multi-view observations, Belief Propagation (BP) [8] is employed to solve the inference problem. Intuitively, a target is tracked in a view by a dedicated local tracker. The trackers in different views interact with each other via a message passing process, BP, so that each local tracker is able to take advantage of additional information from other views. By combining particle filters and BP in a unified framework, Sequential Belief Propagation (SBP) [9] is adopted to have a set of particle-filter-based local trackers collaborate and to perform the inference of the multi-view target states. In doing so, we largely overcome the occlusion problem as BP enables information to be exchanged across views, while the asymmetric property of the message passing in BP [10] guarantees that information propagation is mainly from high-confidence views to low-confidence views, so that the propagation of wrong information is avoided.

It is well known that single-camera tracking faces the difficulties of tracking 3D targets using only 2D information, and is particularly challenged by occlusions. Many algorithms address this problem by tracking multiple targets simultaneously [6, 11, 12]. However, as multiple cameras provide several highly disparate views, which is desirable for occlusion handling, multi-camera tracking has received growing attention in the field of visual tracking [3, 4, 13–18]. Most related work on multi-camera tracking uses a centralized fusion framework, e.g. for soccer player tracking [3, 4] and for video surveillance [13–16], where targets are tracked independently in different views and a fusion module integrates the tracks obtained from each view [19]. However, with no interaction between individual trackers, these methods suffer from the problem that the fusion results are biased by the tracking error in one view. To overcome this problem, many approaches perform occlusion reasoning and tracking performance assessment to obtain a confidence weight for each view [14–16]. Still, since the weight of a view is computed according to the quality of the tracking results, the risk of wrong target association is high if the targets being tracked are approached by other, similar objects, which is usual in sports scenarios.

Due to their tremendous success in visual tracking, particle filters were introduced into multi-camera tracking by Nummiaro et al. [17] and Wang et al. [18], among others. Both contributions are based on the best-view selection strategy: The real target states are estimated using the view that contains the most likely observations. In a sense, these methods exploit multi-view information by automatically switching observation models from one view to another. However, a problem is that the target of interest may not be very distinct from clutter in the chosen view. As a result, poor selection of the best view may cause the complete loss of tracks of targets.

These methods are thus not sufficient to solve the occlusion problem in that they do not properly model the correlation between the states of a target in different views. In contrast to previous work, we build a two-level, tree-structured graphical model to describe the dependence between the target states in different views and in 3D. The model contains a set of leaf nodes and a central node. Bidirectional belief propagation between them enables the exchange of information across views. An efficient sequential belief propagation algorithm is adopted for the collaboration of a set of particle-filter-based local trackers. SBP is a sequential version of the non-parametric belief propagation algorithm [20, 21], which was first introduced by Hua et al. [9] to multi-scale visual tracking. We borrow the idea and adapt it to our multi-camera tracking task. The proposed algorithm proves robust, particularly to occlusions introduced either by clutter in the background or by similar objects, without any explicit occlusion reasoning or tracking performance assessment.

The rest of the paper is organized as follows. Section 2 describes the multi-view target representation and the graphical models designed for the multi-camera tracking problem. The theory of SBP and the details of the tracking algorithm are introduced in Section 3. Results on soccer game sequences are illustrated in Section 4.

## 2 Model Description

The target state in view $j$ is denoted by $x_{t,j}$, $j = 1, \dots, L$, and the target state in 3D is denoted by $x_{t,0}$. Putting all states together results in a multi-view target state, denoted by $X_t = \{x_{t,0}, \dots, x_{t,L}\}$. A benefit of this representation is that it facilitates the integration of multi-view image observations, which helps overcome the occlusion problem. The image observation associated with $x_{t,j}$ is denoted by $z_{t,j}$ and $Z_t = \{z_{t,1}, \dots, z_{t,L}\}$ is the multi-view image observation. Note that there is no image observation associated with $x_{t,0}$.

Given the above definitions, a graphical model (Fig. 1) is built to model the dependence between different states. The model consists of a central node associated with $x_{t,0}$, a set of leaf nodes associated with $x_{t,j}$, and observation nodes associated with $z_{t,j}$, $j = 1, \dots, L$. We assume that target states in different views are independent given $x_{t,0}$ so that a tree-structured model is formed. One advantage of this model is that it is acyclic so that BP can perform exact inference [8]. Connecting the graphical models at different times results in a dynamic Markov model shown in Fig. 2, which describes the evolving process.

In both models, the undirected link between $x_{t,j}$, $j = 1, \dots, L$, and $x_{t,0}$ describes the mutual influence between each leaf node and the central node, and is associated with a potential function $\psi^t_{0,j}(x_{t,0}, x_{t,j})$. The directed link from $x_{t,j}$ to $z_{t,j}$, $j = 1, \dots, L$, represents the image observation processes and is associated with an image likelihood function $p_j(z_{t,j}|x_{t,j})$. In Fig. 2, the directed link from

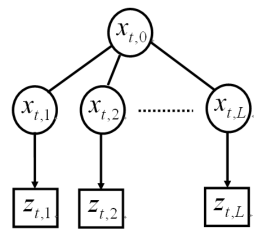



**Fig. 1.** Tree-structured graphical model.

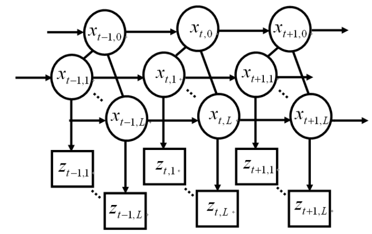



**Fig. 2.** Dynamic Markov model.

$x_{t-1,j}$ to $x_{t,j}$, $j = 0, \ldots, L$, represents the prediction process $p(x_{t,j}|x_{t-1,j})$ and is associated with a motion model.

According to Bayes' rule and the Markov assumption, the recursive inference of the marginal posterior $p(X_t|Z^t)$ is formulated as

$$p(X_t|Z^t) \propto p(Z_t|X_t) \int_{X_{t-1}} p(X_t|X_{t-1}) p(X_{t-1}|Z^{t-1}),$$

where $X_t$ is the multi-view state at time $t$, and $Z^t = \{Z_1, \ldots, Z_t\}$ denotes the multi-view observation up to time $t$.

Direct inference of $p(X_t|Z^t)$ is intractable due to the lack of a closed-form solution and the high dimensionality of the joint state space. In practice, we infer $p(x_{t,j}|Z^t), j = 0, \ldots, L$, collaboratively, as shown in the following sections.

## 3 SBP-based Multi-Camera Tracking

### 3.1 Sequential Belief Propagation

The basic idea is to calculate the inference of the multi-view target states through a message-passing process. The local message passed from view $j$ to the central node in the graphical model in Fig. 1 is

$$m_{0j}(x_{t,0}) \leftarrow \int_{x_{t,j}} p_j(z_{t,j}|x_{t,j}) \psi^t_{0,j}(x_{t,0}, x_{t,j}). \tag{1}$$

Here, $p_j(z_{t,j}|x_{t,j})$ is the image likelihood function of view $j$, and $\psi^t_{0,j}(x_{t,0}, x_{t,j})$ is the potential function mapping local information in view $j$ to the central node. It is symmetric, and its definition will be given later. Likewise, the local message passed from the central node to view $j$ is

$$m_{j0}(x_{t,j}) \leftarrow \int_{x_{t,0}} \prod_{l \neq j} m_{0l}(x_{t,0}) \psi^t_{0,j}(x_{t,0}, x_{t,j}). \tag{2}$$

As the central node is not associated with any image observation, it simply passes on the messages received from all the views.

Our goal is to infer the marginal posteriors $p(x_{t,j}|Z^t)$, $j = 0, \ldots, L$, based on the dynamic graphical model in Fig. 2. We assume independent motion models in each view, i.e.

$$p(X_t|X_{t-1}) = \prod_{j=0}^{L} p(x_{t,j}|x_{t-1,j}). \tag{3}$$

Given the marginal posteriors at the previous time instant $p(x_{t-1,j}|Z^{t-1})$, $j = 0, \ldots, L$, the above message-passing equations are updated as

$$m_{0j}(x_{t,0}) \leftarrow \int_{x_{t,j}} p_j(z_{t,j}|x_{t,j})\psi^t_{0,j}(x_{t,0}, x_{t,j}) \int_{x_{t-1,j}} p(x_{t,j}|x_{t-1,j})p(x_{t-1,j}|Z^{t-1}), \tag{4}$$

$$m_{j0}(x_{t,j}) \leftarrow \int_{x_{t,0}} \prod_{l \neq j} m_{0l}(x_{t,0})\psi^t_{0,j}(x_{t,0}, x_{t,j}) \int_{x_{t-1,0}} p(x_{t,0}|x_{t-1,0})p(x_{t-1,0}|Z^{t-1}). \tag{5}$$

Thus, the marginal posteriors of the state in each view $x_{t,j}$, $j = 1, \ldots, L$, and the state in 3D $x_{t,0}$ are given respectively by

$$p(x_{t,j}|Z^t) \propto p_j(z_{t,j}|x_{t,j})m_{j0}(x_{t,j}) \int_{x_{t-1,j}} p(x_{t,j}|x_{t-1,j})p(x_{t-1,j}|Z^{t-1}), \tag{6}$$

$$p(x_{t,0}|Z^t) \propto \prod_{l=1,\ldots,L} m_{0l}(x_{t,0}) \int_{x_{t-1,0}} p(x_{t,0}|x_{t-1,0})p(x_{t-1,0}|Z^{t-1}). \tag{7}$$

The above formulation shows that SBP involves both a particle filtering process that propagates the marginal posteriors over time, and a BP process that updates and passes messages. For the specific model in this paper, we update messages and marginal posteriors iteratively using Eqs. 4, 5, 6, 7, and $p(x_{t,0}|Z^t)$ contains the fusion of multi-view information. As the graphical model at each time instant is a two-level tree, theoretically it converges after two iterations [8]. However, due to the Monte Carlo simulation introduced below, a small number of additional iterations are required to produce robust results.

### 3.2 Monte-Carlo Implementation

In the Monte-Carlo implementation of the SBP-based multi-camera tracking algorithm, both the posteriors of the target states and the messages passed between leaf nodes and the central node are represented by weighted particles,

$$\begin{aligned} p(x_{t,j}|Z^t) &\sim \{s^{(n)}_{t,j}, \pi^{(n)}_{t,j}\}^N_{n=1}, \quad j = 0, \ldots, L, \\ m_{0j}(x_{t,0}) &\sim \{s^{(n)}_{t,0}, \omega^{(j,n)}_{t,0}\}^N_{n=1}, \; j = 1, \ldots, L, \\ m_{j0}(x_{t,j}) &\sim \{s^{(n)}_{t,j}, \omega^{(0,n)}_{t,j}\}^N_{n=1}, \; j = 1, \ldots, L, \end{aligned}$$

where $s^{(n)}_{t,j}$ denotes the sampled particles, $\omega^{(j,n)}_{t,0}$ and $\omega^{(0,n)}_{t,j}$ are the weights of the messages passed between leaf nodes and the central node, and $\pi^{(n)}_{t,j}$ is the

**Algorithm 1** SBP-based Multi-Camera Tracking

1. INITIALIZATION: $k \longleftarrow 1$

1.1 *Re-sampling*: re-sample $\{s_{t-1,j}^{(n)}, \pi_{t-1,j}^{(n)}\}_{n=1}^{N}$ to get $\{s_{t-1,j}^{(n)}, \frac{1}{N}\}_{n=1}^{N}$, $j = 0, \ldots, L$;

1.2 *Prediction*: generate $\{s_{t,j,k}^{(n)}\}_{n=1}^{N}$ from $p(x_{t,j}|x_{t-1,j})$, $j = 0, \ldots, L$;

1.3 *Message Initialization*: for $n = 1, \ldots, N$, $j = 1, \ldots, L$,

$$\omega_{t,0,k}^{(j,n)} = \frac{1}{N},\ \omega_{t,j,k}^{(0,n)} = \frac{1}{N}.$$

2. ITERATION: SBP

2.1 *Importance Sampling*: Sample $\{s_{t,j,k+1}^{(n)}\}_{n=1}^{N}$ from $p(x_{t,j}|x_{t-1,j})$, $j = 0, \ldots, L$;

2.2 *Message Re-weight*: for $n = 1, \ldots, N$, $j = 1, \ldots, L$,

$$\omega_{t,0,k+1}^{(j,n)} = \frac{\sum_{m=1}^{N}\{p_j(z_{t,j,k}^{(m)}|s_{t,j,k}^{(m)})[\frac{1}{N}\sum_{r=1}^{N} p(s_{t,j,k}^{(m)}|s_{t-1,j}^{(r)})]\psi_{0,j}(s_{t,0,k+1}^{(n)}, s_{t,j,k}^{(m)})\}}{\frac{1}{N}\sum_{r=1}^{N} p(s_{t,0,k+1}^{(n)}|s_{t-1,0}^{(r)})},$$

$$\omega_{t,j,k+1}^{(0,n)} = \frac{\sum_{m=1}^{N}\{[\prod_{l\neq j}\omega_{t,0,k}^{(l,m)}][\frac{1}{N}\sum_{r=1}^{N} p(s_{t,0,k}^{(m)}|s_{t-1,0}^{(r)})]\psi_{0,j}(s_{t,0,k}^{(m)}, s_{t,j,k+1}^{(n)})\}}{\frac{1}{N}\sum_{r=1}^{N} p(s_{t,j,k+1}^{(n)}|s_{t-1,j}^{(r)})}.$$

Normalize so that $\sum_n \omega_{t,j,k+1}^{(0,n)} = 1$ and $\sum_n \omega_{t,0,k+1}^{(j,n)} = 1$;

2.3 *Belief Re-weight*: for $n = 1, \ldots, N$,

$$\pi_{t,0,k+1}^{(n)} = [\prod_{l=1,\ldots,L}\omega_{t,0,k+1}^{(l,n)}][\frac{1}{N}\sum_{r=1}^{N} p(s_{t,0,k+1}^{(n)}|s_{t-1,0}^{(r)})],$$

$$\pi_{t,j,k+1}^{(n)} = p_j(z_{t,j,k+1}^{(n)}|s_{t,j,k+1}^{(n)})\omega_{t,j,k+1}^{(0,n)}[\frac{1}{N}\sum_{r=1}^{N} p(s_{t,j,k+1}^{(n)}|s_{t-1,j}^{(r)})],\ j = 1, \ldots, L.$$

Normalize so that $\sum_n \pi_{t,j,k+1}^{(n)} = 1$, $j = 0, \ldots, L$;

2.4 *Iteration*: $k \longleftarrow k + 1$, iterate several times.

belief of a particle. Note that the same particle sets are used to represent the messages and the marginal posteriors. The Monte-Carlo implementation of the SBP-based multi-camera tracking algorithm is given in Algorithm 1.

The occlusion problem is effectively solved by Algorithm 1 unless the targets are persistently occluded in all views. Our approach is superior to the centralized fusion strategy [3, 4, 13–16] and the best-view selection strategy [17, 18] proposed previously in that the full information in all views is taken into consideration during tracking. Even a view in which the target is completely occluded "contributes" to the tracking results by propagating uniformly-distributed belief to other views. Since this view is not informative, it will not affect the inference of the target states in other views. As pointed out by Sun et al. [10], the message passing in BP is asymmetric: the entropy of the messages from high-confidence nodes to low-confidence nodes is smaller than the entropy of the messages from low-confidence nodes to high-confidence nodes. Consequently, the propagation of incorrect information is avoided.


**Fig. 3.** A mosaic illustrating the incremental homography update.

### 3.3 Potential Function and Geometric Camera Calibration

One issue in Algorithm 1 is the proper definition of the potential function that describes the relationships between different target states. We model the target in a view as a rectangle, $x_{t,j} = (u_{t,j}, v_{t,j}, h_{t,j}, w_{t,j})$, where $(u_{t,j}, v_{t,j})$ is the image coordinate of the feet of the athlete and $(h_{t,j}, w_{t,j})$ is the 2D size, $j = 1, \ldots, L$, and the corresponding target in 3D as a cylinder, $x_{t,0} = (u_{t,0}, v_{t,0}, h_{t,0}, w_{t,0})$, where $(u_{t,0}, v_{t,0})$ is the position on the ground plane and $(h_{t,0}, w_{t,0})$ is its 3D size.

Under the assumption that the athletes move on a ground plane, full camera calibration is not necessary as $(u_{t,0}, v_{t,0})$ can be related to $(u_{t,j}, v_{t,j})$ through an image-to-ground homography $H_{t,j}$. In soccer scenarios where PTZ cameras are used, we update $H_{t,j}$ either by using known features such as border lines on the court where enough of them are visible, e.g. near the penalty area, or by cumulating small estimates of motion between consecutive frames where no or few features are visible [3]. Figure 3 illustrates the result of homography updates. The 3D size $(h_{t,0}, w_{t,0})$ is related to the 2D size $(h_{t,j}, w_{t,j})$ through some reference object in the sequence that has a known height, e.g. the goalpost.

The potential function $\psi^t_{0,j}$ is defined as

$$\psi^t_{0,j}(x_{t,0}, x_{t,j}) \;\propto\; \lambda N(x_{t,0}; \mu_{x_{t,0}}, \Lambda_0) + (1-\lambda) N(x_{t,0}; \Pi^t_j(x_{t,j}), \Gamma^t_j(x_{t,j})), \quad (8)$$

where the first term is a standard Gaussian outlier process, the second term models the spatial correlation between $x_{t,0}$ and $x_{t,j}$ and can intuitively be thought of as the distance between $x_{t,0}$ and $x_{t,j}$ after mapping $x_{t,j}$ from the image plane to the ground-plane reference frame using $\Pi^t_j$. $\Gamma^t_j$ propagates the uncertainties of $x_{t,j}$ from view $j$ to the ground plane using perturbation theory [22]. Figure 4 shows the results of the uncertainty propagation.

We find that the potential function is critical to the success of the multi-camera tracking algorithm. An important issue to stress is that the mapping of the target positions from the image plane to the ground plane has large uncertainty if the camera view direction is highly oblique. In this case, a relatively large number of particles is needed to model the target distribution on the ground plane. This motivates the use of more views to reduce the uncertainty.

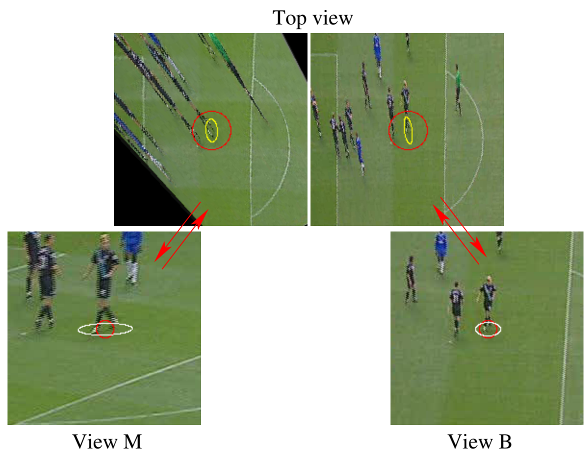



**Fig. 4.** Uncertainty propagation between the positions in each view and on the ground plane. Yellow ellipses depict the estimated covariance for a ground position given a target in View M or B (red). White ellipses give the result of the reverse propagation.

## 4 Results

The algorithm is tested on sequences of a soccer game taken from two uncalibrated cameras. Here, we assume the standard constant-velocity motion model for all states. Following Pérez et al. [23], a classical observation model based on HSV color histograms is adopted due to the advantage of its insensitivity to illumination effects. Thus, the observation process amounts to matching the color histograms in a set of sampled regions with a previously-learned reference model, where the Bhattacharyya coefficient is computed to measure the similarity. In all experiments, we manually initialize the regions of the athletes in the first frame of each camera and learn the reference color models. Some results are shown in Figs. 5 and 7.

To compare with classical work in this field, we implemented the Condensation algorithm [24]. As expected, the result of Condensation on the same sequence of Camera M in Fig. 5, shown in Fig. 6, are hampered by occlusion and nearby clutter of similar appearance, and, in the end, Condensation loses track and follows a wrong target. In contrast, our algorithm integrates information from other cameras to compensate for poor observations, and is thus able to keep track of the correct target.

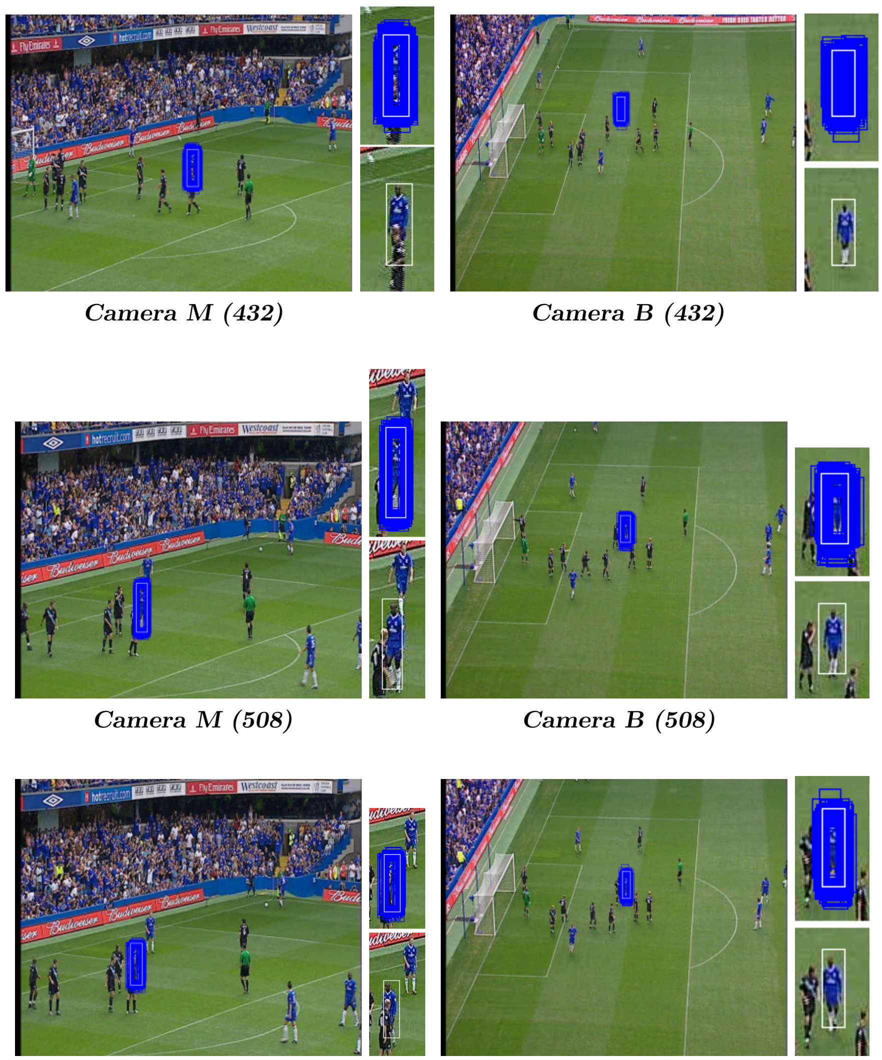
This comparison illustrates the strength of our algorithm. The proximity of similar objects poses the fundamental problem of selecting the correct mode that represents the true target from a multi-modal distribution produced jointly by multiple, similar objects. Neither the centralized fusion nor the best-view selection strategy are able to deal with this problem: the former will integrate the wrong results from camera M, and the latter method may select camera M as the best view, in which the wrong estimate happens to have a high match score. In contrast, our SBP-based multi-camera tracking algorithm enables the cameras to "talk" to each other so that the correct mode in a multi-modal distribution can be enhanced by receiving messages from other views.

## 5 Conclusion and Future Work

We presented a novel multi-camera tracking algorithm for tracking athletes in team sports. The strength of the method lies in the fact that each local tracker integrates information from multiple views via a message-passing process, BP, resulting in collaborative inference of multi-view target states. Technically, the method is insensitive to occlusions present in only some of the views. However, due to the use of SBP, the method itself has some capacity of dealing with total occlusions in all views as the SBP-based tracking algorithm generates target hypotheses according to a motion model. In practice, chances are that these hypotheses are not far from the truth if the duration of the occlusion is not too long. We are currently developing algorithms for multi-target, multi-camera tracking, which involves the association of the target observations both in time and across views. Another interesting extension is to adapt the algorithm to non-overlapping cameras.

## References

1. Urtasun, R., Fleet, D., Fua, P.: Monocular 3–D tracking of the golf swing. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 1., San Diego, CA (2005) 932–939
2. Pingali, G.S., Jean, Y., Carlbom, I.: Real-time tracking for enhanced tennis broadcasts. In: IEEE Conference on Computer Vision and Pattern Recognition. (1998) 260–265
3. Hayet, J.B., Mathes, T., Czyk, J., Piater, J., Verly, J., Macq, B.: A modular multi-camera framework for team sports tracking. In: International Conference on Advanced Video and Signal based Surveillance, Como, Italy (2005) 493–498
4. Xu, M., Lowey, L., Orwell, J.: Architecture and algorithms for tracking football players with multiple cameras. IEE Proceedings - Vision, Image, and Signal Processing **152** (April, 2005) 232–241
5. Intille, S., Davis, J., Bobick, A.: Real-time closed-world tracking. In: IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico (1997) 697–703
6. Okuma, K., Taleghani, A., de Freitas, N., Little, J., Lowe, D.: A boosted particle filter: multitarget detection and tracking. In: European Conference on Computer Vision, Prague, Czech Republic (2004) 28–39
7. Jordan, M.: Learning in Graphical Models. MIT Press, Cambridge, MA (1999)
8. Yedidia, J., Freeman, W., Weiss, Y.: Generalized belief propagation. Advances in Neural Information Processing Systems (NIPS) **13** (2000) 689–695

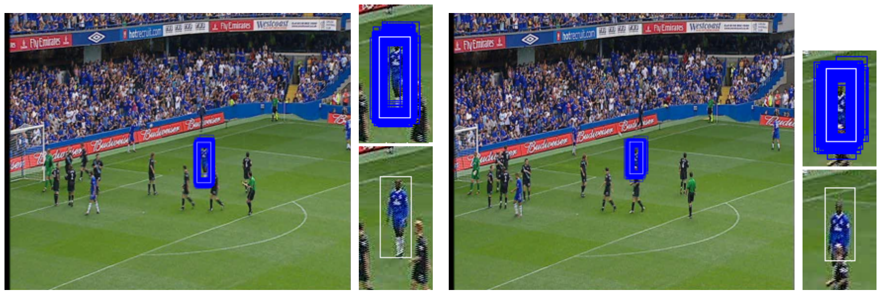


*Camera M (432)* *Camera B (432)*

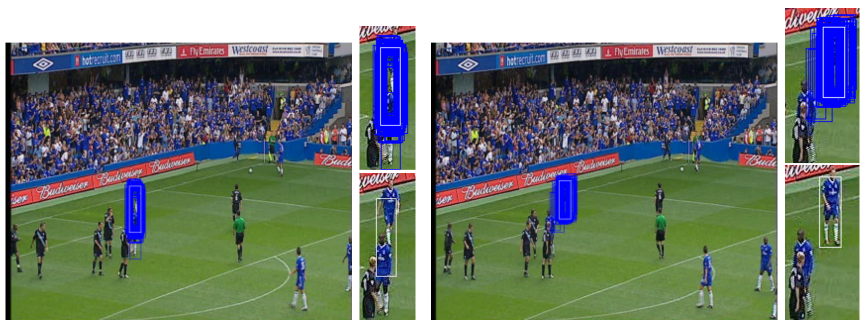
*Camera M (508)* *Camera B (508)*

*Camera M (519)* *Camera B (519)*

**Fig. 5.** Results of our SBP based multi-camera tracker. Although in Camera M the tracked player is occluded by his opponent and is approached by his teammate, he is visible mostly in Camera B so that a strong belief propagated from Camera B attracts the target to the true observations at Camera M. Numbers in parentheses are frame numbers.

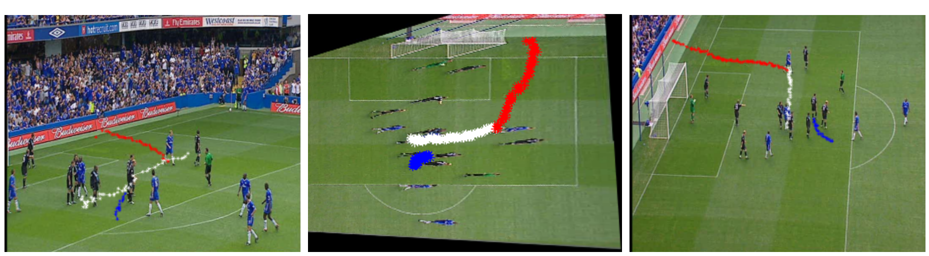


*Camera M (419)* *Camera M (432)*



*Camera M (508)* *Camera M (519)*

**Fig. 6.** Results of Condensation. Due to the occlusion and the teammate approaching the target of interest, Condensation loses the track at Frame 519.



*Camera M (400-600)* *Ground (400-600)* *Camera B (400-600)*

**Fig. 7.** Results of tracking several players simultaneously. Trajectories of players in different views and on a ground plane are shown in different colors.

9. Hua, G., Wu, Y.: Multi-scale visual tracking by sequential belief propagation. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 1., Washington, DC (2004) 826–833
10. Sun, J., Zheng, N., Harry, S.: Stereo matching using belief propagation. IEEE Transactions on Pattern Analysis and Machine Intelligence **25** (2003) 787–800
11. Rasmussen, C., Hager, G.D.: Probabilistic data association methods for tracking complex visual objects. IEEE Transactions on Pattern Analysis and Machine Intelligence **23** (2001) 138–150
12. Khan, Z., Balch, T., Dellaert, F.: MCMC-based particle filtering for tracking a variable number of interacting targets. IEEE Transactions on Pattern Analysis and Machine Intelligence **27** (2005) 1805–1918
13. Black, J., Ellis, T.: Multi camera image tracking. In: the Second IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Kauai, HI (2001)
14. Dockstader, S., Tekalp, A.: Multiple camera fusion for multi-object tracking. In: IEEE Workshop on Multi-Object Tracking, Vancouver, Canada (2001) 95–102
15. Beugnon, C., Singh, T., Llinas, J., Saha, R.: Adaptive track fusion in a multisensor environment. In: the Third International Conference on Information Fusion. Volume 1., Paris, France (2000) 24–31
16. Snidaro, L., Foresti, G.L., Niu, F., Varshney, P.K.: Sensor fusion for video surveillance. In: the Seventh International Conference on Information Fusion, Stockholm, Sweden (2004) 739–746
17. Nummiaro, K., Koller-Meier, E., Svoboda, T., Roth, D., Van Gool, L.: Color-based object tracking in multi-camera environment. In: 25th Pattern Recognition Symposium, DAGM, Magdeburg, Germany (2003)
18. Wang, Y., Wu, J., Kassim, A.: Multiple cameras tracking using particle filtering. In: IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Breckenridge, CO (2005) 33–39
19. Bar-Shalom, Y., Li, X.R.: Multitarget-Multisensor Tracking: Principles and Techniques. YBS Publishing (1995)
20. Sudderth, E., Ihler, A., Freeman, W., Willsky, A.: Nonparametric belief propagation. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 2., Madison, WI (2003) 605–612
21. Isard, M.: Pampas: Real-valued graphical models for computer vision. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 1., Madison, WI (2003) 613–620
22. Criminisi, A., Reid, I., Zisserman, A.: A plane measuring device. Image Vision and Computing **17** (1999) 625–634
23. Pérez, P., Hue, C., Vermaak, J., Gangnet, M.: Color-based probabilistic tracking. In: European Conference on Computer Vision. Volume 1., Copenhagen, Denmark (2002) 661–675
24. Isard, M., Blake, A.: Condensation – conditional density propagation for visual tracking. International Journal of Computer Vision **29** (1998) 5–28