Supplementary material: Prioritizing the Propagation of Identity Beliefs for Multi-object Tracking

Amit Kumar K.C.
http://www.uclouvain.be/amit.kc
Christophe De Vleeschouwer
http://www.uclouvain.be/christophe.devleeschouwer
ICTEAM Institute
Université catholique de Louvain
Louvain-la-neuve, Belgium

1 Performance metrics for all configurations of the graph

Table 1 presents the results of all three message passing techniques, described above. We can see that the yellow team is recognized better than the blue team. This might be accredited to asymmetry of the position of the cameras around the court.

Team	Metrics	No BP	Std. BP	Priority BP		
				Mutex	Temporal	Both
Blue	Accuracy(%)	48.74	55.54	70.11	77.45	86.37
	FP(%)	0.24	0.24	0.24	0.24	0.24
	WI(%)	0.71	1.04	1.82	2.26	2.42
	MS(%)	50.31	43.18	27.83	20.05	10.91
Yellow	Accuracy(%)	87.54	91.64	90.53	91.54	91.71
	FP(%)	0.03	0.07	0.07	0.07	0.07
	WI(%)	2.25	2.48	2.25	3.16	2.65
	MS(%)	10.18	5.81	7.15	5.23	5.58
Avg.	Accuracy(%)	68.14	73.59	80.32	84.49	**89.04**
	FP(%)	0.13	0.15	0.15	0.15	**0.15**
	WI%)	1.48	1.76	2.03	2.71	**2.54**
	MS(%)	30.25	24.50	17.50	12.65	**8.27**

Table 1: Performance metrics for all configurations.

Figure 1 depicts the performance metrics for the targets across time. Even though there are 21 targets in the ground truth, there are 2 referees and 4 spectators (which correspond to ID 1, 10, 12, 14, 15 and 16) for which there is no appearance features. Therefore, the identities for these targets have not been computed.

2 Performance with respect to the entropy threshold, τ_{TH}

This threshold is used in the definition of the pairwise potential to determine whether the current identity estimate is reliable or not. If the identities of two nodes are reliable, then the potential function is defined by the Bhattacharyya distance between their beliefs. Otherwise, the position feature is used for the definition.

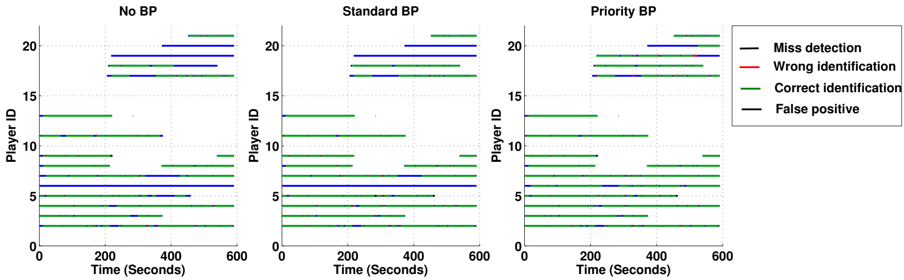


Figure 1: Performance metrics across time for each target. Each line corresponds to a single player track. **Left**: Baseline system where no belief propagation is performed (accuracy = 68.14%). **Middle**: nodes are scheduled in the order they are stored in memory and hence no priority (standard BP) (accuracy = 73.59%). **Right**: Nodes are prioritized according to their level of ambiguity (priority BP) (accuracy = 89.04%).

When this threshold is small, the potential function between the nodes is governed by the identity estimate only when the nodes have small entropy. That is, these nodes are already unambiguous, and the exchange of message between them does not bring extra information. This can be observed from the Table 2, in which the metrics do not significantly differ for $\tau_{\text{TH}} \leq 0.3$.

Conversely, when the threshold is high, many temporal edges will be driven by the Bhattacharyya distance, and even if the spatio-temporal information is discriminant, ignored and replaced by the ambiguous identity information. Consequently, the system will not be able to exploit the position features. More importantly, unreliable messages (and possibly wrong messages) might be exchanged between the nodes, which, in turn, might degrade the performance. This can be seen from the Table 2 in which the metrics get worse for $\tau_{\text{TH}} > 0.5$.

τ_{TH}	**0.05**	**0.1**	**0.3**	**0.5**	**0.6**	**0.8**
Accuracy(%)	87.79	87.79	87.79	89.04	87.08	84.07
FP(%)	0.15	0.15	0.15	0.15	0.15	0.27
WI(%)	2.54	2.54	2.54	2.54	2.54	2.57
MS(%)	9.52	9.52	9.52	8.27	10.23	13.09

Table 2: Performance with respect to the threshold on entropy τ_{TH}.

3 Impact of the node ambiguity level estimation method

The results presented so far are obtained by scheduling the nodes according to their entropies. Our objective is to prioritize the nodes that have a peaked distribution over the ones with flat distribution. Thus, we can define various metrics that estimate the measure of peakedness of the identity distribution. Apart from the entropy, we have tested two methods to estimate the level of ambiguity of the nodes. They are the cardinality of the confusion set and the kurtosis of the belief vector.

Confusion set

The notion of confusion set has been introduced in [9]. Given the belief vector $\mathbf{b}_v$, one way to count the confidence of this node is simply by counting the number of likely labels, e.g., those whose belief value exceed a certain threshold τ_b. More number of labels imply that the node identity is more ambiguous. labels with high probability exist for that node and hence it is more ambiguous. Of course, the relative beliefs $b_v^{(\mathrm{rel})}(l_v) = b_v(l_v) - b_v^{(\mathrm{min})}$ (where, $b_v^{(\mathrm{min})} = \min_{l_v} b_v(l_v)$) matter in this case. Thus, the set $\mathcal{CS}(v) = \{l_v | b_v^{(\mathrm{rel})}(l_v) \geq \tau_b\}$ corresponds to the confusion set of v, and its cardinality $|\mathcal{CS}(v)|$ is an estimate of the ambiguity of v.

The performance of the system with various values of the τ_b is presented in Table 3.

τ_b	0.01	0.03	0.1	0.3
Accuracy(%)	90.16	89.05	86.23	69.84
FP(%)	0.15	0.15	0.15	0.13
WI(%)	2.52	2.45	2.24	1.85
MS(%)	7.17	8.35	11.38	28.18

Table 3: Performance metrics with respect to τ_b.

From the table, it can be deduced that the larger value of τ_b implies that the cardinality of the confusion set will be low and thus many nodes will have similar priorities, which, in turn, deteriorate the performance of the system.

Kurtosis of the belief vector

Another method to estimate the peakedness of the belief vector is to compute its kurtosis. The kurtosis is the fourth central moment, divided by fourth power of the standard deviation of a distribution. It is a descriptor of the shape of the distribution and higher kurtosis values correspond to peaky distributions. Thus, we use the kurtosis as the estimate of the level of ambiguity of the identity distribution and schedule the nodes in decreasing order of kurtosis. The results of kurtosis for assigning priority to nodes are as follows:

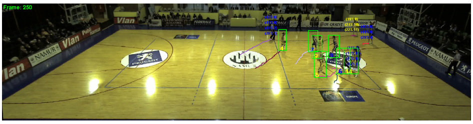
Accuracy=89.09% FP=0.15% WI=2.54% MS=8.22%.

4 Algorithm for message construction

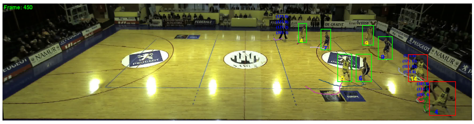
Algorithm 1 ComputeMessage

Input: souce node u, target node v, pre-message $\mathbf{h}_u$, node separation time $\tau_{\max}$
Output: message from u to v at time t: $\mathbf{m}_{u\to v}^{(t)}$
Procedure:
if Edge (uv) is mutex **then**
  $m_{u\to v}^{(t)}(l_v) \leftarrow$ compute according to Equations 3 and 5.
else
  if $\mathcal{E}(u) < \tau_{\mathrm{TH}}$ AND $\mathcal{E}(v) < \tau_{\mathrm{TH}}$ **then**
    $d_{uv} \leftarrow$ compute Bhattacharyya distance between $\mathbf{b}_u$ and $\mathbf{b}_v$
    $m_{u\to v}^{(t)}(l_v) \leftarrow$ compute according to Equations 3 and 6
  else
    $\Delta t_{uv} \leftarrow |t_v^{(s)} - t_u^{(e)}|$
    if $\Delta t_{uv} < \tau_{\max}$ **then**
      $d_{uv} \leftarrow$ compute spatial distance between $\mathbf{x}_u^{(e)}$ and $\mathbf{x}_v^{(s)}$
      $m_{u\to v}^{(t)}(l_v) \leftarrow$ compute according to Equations 3 and 6
    end if
  end if
end if
Normalize $\mathbf{m}_{u\to v}^{(t)}$

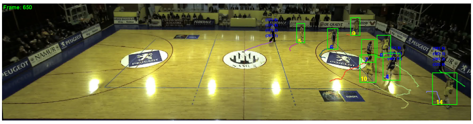
5 Sample frames



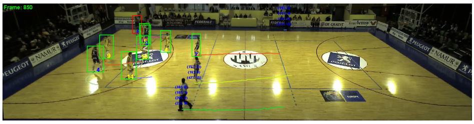
Frame 250



Frame 450

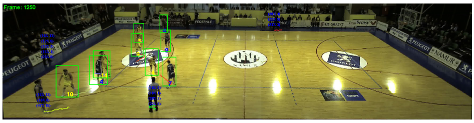


Frame 650

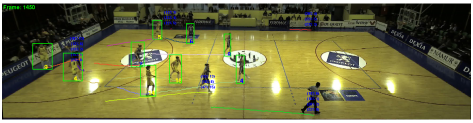


Frame 850

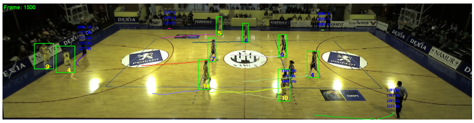
Frame 1050



Frame 1250



Frame 1450



Frame 1500

Table 4: Sample frames. At each frame, estimated IDs of the players are shown, along with a 50-frame long track. **Green**- correct recognitions; **Red**-wrong identifications; **Blue**-misses. As we can see, referees do not have distinct appearance features and thus their identities are almost uniformly distributed among the blue players.