VARIATIONAL BAYESIAN AND BELIEF PROPAGATION BASED DATA ASSOCIATION FOR MULTI-TARGET TRACKING

Ata-ur-Rehman[⋆], *Syed Mohsen Naqvi*[⋆], *Lyudmila Mihaylova*[†], *Jonathon A. Chambers*[⋆]

[⋆]School of Electronic, Electrical and Systems Engineering Loughborough University UK.
{a.ur-rehman, s.m.r.naqvi, j.a.chambers}@lboro.ac.uk
[†]School of Computing and Communications, Lancaster University UK.
mila.mihaylova@lancaster.ac.uk

ABSTRACT

A novel two stage data association technique for multi-target tracking is proposed which assigns multiple measurements to a target to mitigate information loss. At the first stage a variational Bayesian (VB) clustering technique is used which groups the measurements automatically into a determined number of clusters. In the second stage a belief propagation (BP) based cluster to target association method is proposed to assign multiple clusters to a target. This is achieved by exploiting the inter-cluster dependency information. The proposed technique is suitable to accommodate non-rigid targets such as humans. Both location and features of clusters are used to re-identify the targets when they emerge from occlusions. The proposed technique is compared with state of the art method due to Laet et al. and evaluations are presented on a real data set.

Index Terms**—** variational Bayesian methods, clustering, data association, belief propagation, multi-target tracking

1. INTRODUCTION

Multi-target tracking is a challenging problem which has occupied various researchers [1, 2, 3, 4, 5, 6]. It more generally has many applications, such as surveillance, intelligent transportation, submarine tracking, animal tracking for behavioral analysis, human computer interfacing, and etc. [1, 7, 8, 9]. A successful multi-target tracking system requires a reliable identification of targets, which can be achieved by adopting an appropriate data association technique. Most of the existing data association methods assume that the targets generate one measurement at a time [1, 10]. This is a strong assumption because in most tracking applications [5, 11, 12, 13], targets generate multiple measurements e.g. in the case of video tracking there are multiple pixels originating from one target. To deal with such problems, existing techniques extract features from the group of measurements [9, 14, 15]. These techniques can result in loss of information and degrade the tracking results. A clustering and data association based approach has been recently proposed in [5] to assign multiple measurements to one target. In this approach the measurements originating from the targets are first grouped into clusters by using the variational Bayesian (VB) [16] approach and then clusters are assigned to the different targets by using a joint probabilistic data association filter (JPDAF) [17]. This is a shape based approach in which clustering is performed on the basis of the structure of the targets. This approach is only appropriate for rigid shapes where targets do not change their shapes. Another drawback of this technique is its assumption that one target generates only one cluster at a time. This is again a strong assumption and results in data association failure when targets generate multiple clusters. Another limitation of this work is that it does not use different features of clusters and hence when targets emerge from occlusions, the tracker sometimes fails to re-identify the targets.

The main contribution of this work is that a more robust data association technique is proposed to overcome the limitations of the existing state of the art technique in [5]. The proposed technique assigns multiple measurements to a single target in a two stage process. In the first stage, measurements originating from all the targets are grouped by using the VB clustering technique and then at the second stage these clusters are assigned to targets by using a belief propagation (BP) [18] method. The advantage of using VB is that it automatically determines the number of clusters which can fit the measurements. This is very helpful in multi-target tracking: a case where the number of targets is unknown and remains changing. By using BP, the proposed algorithm describes a solution to assign multiple clusters to a target. Furthermore, the proposed algorithm does not completely rely on shape information and hence it provides a more suitable data association method for non-rigid targets such as humans.

The algorithm exploits the association between the clusters which helps to assign multiple related clusters to one target. It uses both location and features of clusters to perform the data association and thereby achieves more reliable data association results. Unary potential in BP is defined on the basis of the location information of clusters while pairwise potentials are defined on the basis of cluster features. This helps to correctly re-identify the targets emerging from occlusions and hence overcomes the tracking failures. Furthermore, a technique is also proposed to prioritise these clusters. Clusters which are more reliable or in other words which are more certain about their identity are given priority as compared to the clusters which are ambiguous. This helps to simplify the cluster labeling problem.

The remainder of the paper is organized as follows: Section 2 describes the VB clustering technique to group the measurements; cluster to target association by using BP is explained in Section 3; experimental validation is presented in Section 4; finally, conclusions and relation to the prior work are discussed in Section 5.

2. VARIATIONAL BAYESIAN CLUSTERING

At the first stage of the data association process, measurements are grouped into clusters. This clustering process needs to be done in a way that each cluster contains measurements originating from one

Thanks to EPSRC UK (Grant number EP/H049665/1) and UET Lahore, Pakistan for supporting the research.

target. If we assume M number of measurements originating from N targets, then the clustering process groups them into K clusters. Each cluster is represented by its center $\boldsymbol{\mu}_q$ where $q = 1, \cdots, K$ represents the cluster index. We assume that $\mathbf{y}_k^j$ represents the j^{th} measurement at discrete time step k. The goal of the clustering is to find the center points of clusters and measurement to cluster association for every single measurement. A binary association indicator $b_k^{j,q} \in \{0, 1\}$ is used to represent the association of the j^{th} measurement to the q^{th} cluster. If for example the measurement $\mathbf{y}_k^j$ is assigned to the q^{th} cluster then $b_k^{j,q} = 1$ and $b_k^{j,l} = 0$ for $l \neq q$. The vector $\mathbf{b}_k^j = [b_k^{j,1}, b_k^{j,2}, \ldots, b_k^{j,q}, \ldots, b_k^{j,K}]$ indicates, at discrete time k, to which cluster $\mathbf{y}_k^j$ corresponds, and the matrix $\mathbf{B}_k = [\mathbf{b}_k^1, \ldots, \mathbf{b}_k^j, \ldots, \mathbf{b}_k^M]$ represents all the association indicators at time step k.

Clustering can be viewed as fitting mixtures of Gaussians to the measurements [16]

$$p(\mathbf{y}_k^j) = \sum_{q=1}^{K} C_k^{j,q} N(\mathbf{y}_k^j | \boldsymbol{\mu}_q, \boldsymbol{\Sigma}_q) \tag{1}$$

where each Gaussian density $N(\mathbf{y}_k^j|\boldsymbol{\mu}_q, \boldsymbol{\Sigma}_q)$ is a component of the mixture with mean $\boldsymbol{\mu}_q$ and covariance $\boldsymbol{\Sigma}_q$. The mixing coefficient $C_k^{j,q}$ represents a prior probability of picking the q^{th} component of the mixture, which can be represented as

$$C_k^{j,q} = p(b_k^{j,q} = 1) \tag{2}$$

In our particular problem of clustering, the main objective is to eveluate the probability distribution of the unknown parameters $\mathbf{B}_k$, $\mathbf{C}_k$, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ given the observed measurements $\mathbf{Y}_k$. VB clustering is one of the techniques which approximates the probability distribution over such unknown parameters. The joint distribution of all the known and unknown parameters can be decomposed as [16]

$$\begin{aligned} &p(\mathbf{B}_k, \mathbf{C}_k, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{Y_k}) = \\ &p(\mathbf{Y_k}|\mathbf{B}_k, \mathbf{C}_k, \boldsymbol{\mu}, \boldsymbol{\Sigma}) p(\mathbf{B}_k|\mathbf{C}_k) p(\mathbf{C}_k) p(\boldsymbol{\mu}|\boldsymbol{\Sigma}) p(\boldsymbol{\Sigma}). \end{aligned} \tag{3}$$

and VB is a two step process to evaluate the optimum forms of approximations to these distributions, defined as $q^*(\mathbf{B}_k)$, $q^*(\mathbf{C}_k)$, $q^*(\boldsymbol{\mu})$ and $q^*(\boldsymbol{\Sigma})$ over the unknown parameters as explained in [16].

3. DATA ASSOCIATION USING BELIEF PROPAGATION

The main contribution of the paper is a new BP based approach to associate multiple measurements (clusters) to each target. This new approach helps to overcome the limitations of JPDAF based data association used in [5]. In this section, the standard BP is first explained. We then describe the construction of the BP graph in the context of the cluster to target association problem, which includes the method of defining the unary and pairwise potentials and priority scheduling of nodes.

3.1. Standard Belief propagation

A pairwise Markov random field (MRF) [16] provides an appropriate theoretical graphical model for the cluster to target association problem. An MRF is an undirected graph $G = (V, E)$, where V represents nodes of the graph and E represents the undirected edges between them. N_v represents the neighboring nodes of node $v \in V$. Each node v has a hidden node (label) $h_v \in H$ attached to it and the purpose of BP is to assign labels to the nodes to maximize the joint distribution over the hidden nodes (labels)

$$p(H) = \prod_{v \in V} \phi_v(h_v) \prod_{u \in N_v} \psi_{uv}(h_u, h_v) \tag{4}$$

where $\phi_v(h_v)$ is the unary potential which represents the prior probability of node v having the label h_v. The pairwise potential $\psi_{uv}(h_u, h_v)$ represents the probability that node v takes label h_v and its neighboring node u takes the label h_u. The pairwise potential helps to model the joint behavior of neighboring nodes.

In BP we calculate beliefs at every node of the graph. Belief $f_v(h_v)$ at node v represents the posterior probability of node v having the label h_v. The belief vector $\mathbf{f}_v = [f_v(h_1), \cdots, f_v(h_v), \cdots, f_v(h_N)]$ represents the beliefs at node v for taking labels $h_1 \cdots h_N$. Belief at node v is the product of unary potential at that node and all the messages coming from the neighboring nodes

$$f_v(h_v) = \kappa \phi_v(h_v) \prod_{u \in N_v} m_{uv}(h_v) \tag{5}$$

where κ is the normalizing factor and $m_{uv}(h_v)$ is the message to node v coming from its neighbor u. This message represents how likely the node u thinks that the node v will take the label h_v. The messages are evaluated by using the message update rule

$$m_{uv}(h_v) \propto \sum_{h_u \in H} \psi_{uv}(h_u, h_v) \phi_u(h_u) \prod_{r \in N_u \backslash v} m_{ru}(h_u) \tag{6}$$

where

$$g_u(h_u) = \prod_{r \in N_u \backslash v} m_{ru}(h_u) \tag{7}$$

is know as a pre-message at node u which it calculates by taking messages from the neighboring nodes except v. The max product form of the message passing equation can be achieved by replacing the summation term by the max term. The next section describes how the BP algorithm is constructed to use it in the cluster to target association problem.

3.2. Belief propagation for cluster to target association

The VB clustering described in Section 2 gives K clusters for N targets. Now the data association problem simplifies to assigning labels (identities) to clusters. Each cluster can originate from one target but BP based data association helps to overcome the limitation of [5] and allows us to associate multiple clusters to a single target.

This cluster to target association problem is represented as a graph $G = (V, E)$. Nodes V represent clusters, and edges E represent inter cluster dependency. The edge between two nodes shows that the identity of one node gives some knowledge about the identity of the other node e.g. if the features of two clusters are very similar it means that there is higher probability that both the clusters originate from the same target and it is highly probable that both will take the same label.

Unary and pairwise potentials of the graph are designed such that we can exploit both features and location information of clusters. The unary potential $\phi_v(h_v)$ is formulated by using the distance of the measurements in a cluster from the estimated location of the target. It is defined such that the probability of the node v having identity h_v decreases with the increase in the distance between cluster v and the estimated location of target h_v and vice versa. Locations of targets at each state can be estimated with the help of Kalman or a particle filters [7]. Because the nature of the problem is non-linear and non-Gaussian, we have used a particle filter to estimate the locations of targets. Focus of the paper is on the data association part of tracking therefore particle filtering and other tracking related details are not provided, see [19] for more details. The particle filter based unary potential term is designed as

$$\begin{aligned}\phi_v(h_v) &= p(\mathbf{Y}_v|\mathbf{x}(h_v)) \\ &= \prod_{i\in\mathbf{Y}_v} \frac{1}{N_s}\sum_{s=1}^{N_s} p(\mathbf{y}_v^i|\mathbf{x}^s(h_v))\end{aligned} \tag{8}$$

where $\mathbf{Y}_v$ represents the measurements in cluster v, $\mathbf{x}(h_v)$ represents the estimated location of target h_v and N_s is the number of samples in a particle filter. An advantage of using this unary potential is twofold: one is that we are utilizing the position information of the cluster and second we do not need any training data.

Next we define the pairwise potential term $\psi_{uv}(h_u, h_v)$. In the proposed work this term is defined such that if two clusters are similar then they should have the same identity. We have exploited the features of clusters to evaluate similarity between them. The pairwise potential term is defined as

$$\psi_{uv}(h_u, h_v) = \begin{cases} exp(-d_{uv}/\vartheta) & if \quad h_v = h_u \\ 1 - exp(-d_{uv}/\vartheta) & \text{otherwise} \end{cases} \tag{9}$$

where ϑ is a constant and d_{uv} is the distance between nodes u and v. One possible choice of calculating this distance is to calculate the Bhattacharyya distance [20] between the features. These unary and pairwise potentials are used in equations (5) and (6) to calculate beliefs at each node. After L iterations label h_v is assigned to node v if it produces the maximum belief at that node i.e $f_v(h_v) > f_v(h_p)$ where

$$f_v(h_p) = \max_{h_p \in H\setminus h_v} f_v(h_p) \tag{10}$$

The standard BP algorithm randomly selects the nodes to which to send messages. In the cluster to target association problem, all nodes do not have the same level of initial belief about their identities e.g. clusters which are at a higher distance from the other nodes and have more discriminating features are less ambiguous as compared to others. In the proposed algorithm we priorities nodes which are less ambiguous and these are allowed to send their messages first. A similar approach is used in [21] to assign labels to tracklets.

We have used the entropy of the belief vector to define the ambiguity of the node, as used in [21]

$$S(v) = -\sum_{h_v\in H} f_v(h_v) log(f_v(h_v)). \tag{11}$$

Higher entropy of node v shows a higher ambiguity of the node, and hence results in a lower priority and vice versa. To construct a belief at node v, the neighboring node u sends a message to node v by gathering messages from all its neighbors except v. However the messages received by node u from the more ambiguous nodes do not provide useful information to solve the labeling problem. Hence in the proposed work the message $m_{uv}(h_v)$ is constructed at node u by considering only those messages which are less ambiguous than the node u. This priority scheduling avoids loops in the graphs and provides a tree type structure in which nodes are arranged according to their priority. The complete data association algorithm is summarized in Algorithm 1.

Algorithm 1 VB and BP based data association algorithm

Input: M measurements
N_s particles for every target, particles are the estimated locations of targets.

$\mathbf{B}_k$, K $\leftarrow$ Labels $\mathbf{B}_k$ of M measurements from clustering process results in K clusters.
$\phi_v(h_v)$ $\forall v \in V$, $\forall h_v \in H \leftarrow$ Evaluate unary potentials by using equation (8)
Initialize beliefs $f_v(h_v) \leftarrow \phi_v(h_v)$ $\forall v \in V$, $\forall h_v \in H$
for $l = 1 : L$ **do**
  $\mathbf{S} \leftarrow$ Evaluate entropy of all the nodes by using equation (11).
  $E \leftarrow$ Prioritise the nodes according to the entropy
  for $w = 1 : V - 1$ **do**
    $u \leftarrow$ Pick the most unambiguous node in E
    $I \leftarrow$ Pick all the less ambiguous neighbors of u
    $g_u \leftarrow$ Evaluate the pre-message at u by using less ambiguous nodes I in equation (7)
    $J \leftarrow$ Pick all the more ambiguous neighbors of u
    for $v \in J$ **do**
      $m_{uv}(h_v)$ $\forall h_v \in H \leftarrow$ Evaluate message from node u to v by using equation (6)
      $f_v(h_v)$ $\forall h_v \in H \leftarrow$ Evaluate belief at node v
    end for
    $E \leftarrow E\setminus u$ Resize E by eliminating the node u.
  end for
end for
Assign labels to all the nodes according to rule explained in Section 3.2

4. EXPERIMENTAL RESULTS

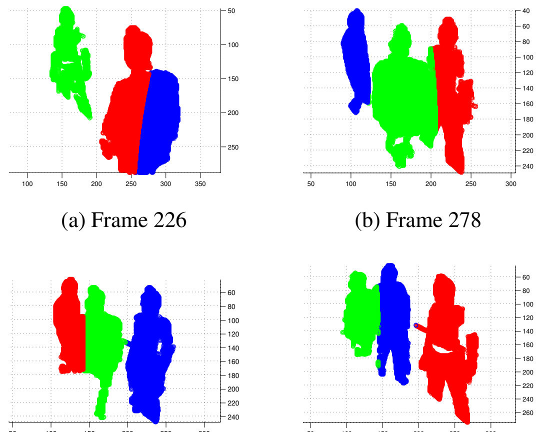
The algorithm is evaluated by tracking people in the video recording (seq45-3p-1111_cam3_divx_audio.avi) taken from the AV16.3 corpus [22] available at http://glat.info/ma/av16.3/. The test sequences are recorded at a resolution of 288 $\times$ 360 at 25 frames/sec showing up to three people moving in a room environment in a closed arena. All the parameters have been chosen empirically to yield best results. To investigate the performance of the proposed algorithm, results are compared with [5].

4.1. Clustering Results

Clustering is based on the VB technique, advantage of using this technique is that we do not have to define the number of clusters K. Clustering results for a few of the video frames are shown in Fig.1. Blue, red and green colours represent first, second and third cluster respectively. It is clear from the clustering results that because of the non rigid nature of targets they change their shapes, and hence the number of clusters per target are not always the same. It can be seen in the frame 226 of Fig.1 that one of the targets is producing two clusters. The clustering results fulfil the condition that even during the close interactions, one cluster originates from a single target. At this level of the algorithm, identities of the clusters are not defined.

4.2. Clusters to Targets Association Results

Clusters to target association is the most crucial part of the algorithm. In [5] the JPDAF is used to assign clusters to the targets. To compare, we have implemented the JPDAF technique by considering all possible hypothesis. Table 1 shows the ground truth of the cluster to target association. The data association results with JPDAF are



(a) Frame 226 (b) Frame 278

(c) Frame 288 (d) Frame 313

Fig. 1. Clustering results: First, second and third clusters are represented by blue, red and green colours respectively. (a) Frame 226 shows that one of the targets produces two clusters. (b),(c) and (d) show that each cluster is generated from a single target even during close interactions.

Table 1. Cluster to target association ground truth

Frame No.	Target No.	Cluster Colour	Cluster No.
226	1	green	3
	2	blue and red	1 and 2
278	1	red	2
	2	green	3
	3	blue	1
288	1	blue	1
	2	green	3
	3	red	2
313	1	red	2
	2	green	3
	3	blue	1

shown in Table 2. In JPDAF we do not know how many clusters are originating from a target, hence we assign a cluster with the highest probability to the respective target. For frame 226 only cluster 1 is assigned to target 2 and the second cluster is ignored. Moreover, data association failures can also be seen in frame 278 where target 1 is assigned cluster 3 and target 3 is assigned cluster 2. The similar failures can be seen in association results for frames 288 and 313. This is because the targets are coming out of occlusions and JPDAF does not use the feature information hence fails to correctly re-identify the targets. These failures are corrected with the help of the proposed BP approach.

Data association results of the proposed BP based cluster to target association technique are shown in Table 3. We used the simple colour histograms feature to calculate the pairwise potential terms $\psi_{uv}(h_u, h_v)$. Bhattacharyya distance between the colour histograms of clusters is used to calculate the distance between them. It can be seen for frame 226 multiple clusters are successfully assigned to target 2. By exploiting the feature information in pairwise potentials, we achieved a successful re-identification of targets when they emerge from occlusions. This can be seen in frame 278 where target 1 is correctly assigned cluster 2 and target 3 is assigned cluster 1. The similar corrections can be seen in association results for frames 288 and 313. This confirms the value of applying BP in data association.

Table 2. Cluster to target association with JPDAF. Clusters with the highest probability are assigned to the respective targets

Frame No.	Target No.	Probability of cluster 1	Probability of cluster 2	Probability of cluster 3
226	1	0.080	0.103	0.816
	2	0.496	0.360	0.144
278	1	0.152	0.307	0.541
	2	0.212	0.357	0.431
	3	0.196	0.412	0.392
288	1	0.136	0.332	0.532
	2	0.306	0.191	0.501
	3	0.384	0.195	0.421
313	1	0.422	0.087	0.491
	2	0.435	0.141	0.424
	3	0.494	0.125	0.381

Table 3. Cluster to target association results with BP. Clusters with labels 1 are assigned to the respective targets

Frame No.	Target No.	Label of cluster 1	Label of cluster 2	Label of cluster 3
226	1	0	0	1
	2	1	1	0
278	1	0	1	0
	2	0	0	1
	3	1	0	0
288	1	1	0	0
	2	0	0	1
	3	0	1	0
313	1	0	1	0
	2	0	0	1
	3	1	0	0

5. RELATION TO PRIOR WORK AND CONCLUSIONS

A novel data association technique was proposed to assign multiple measurements to a target to mitigate the information loss in multi-target tracking. The work has been built on a recently proposed state of the art technique [5] in which data association was performed with the help of a clustering and a JPDAF based approach. In our proposed technique VB based clustering was used which grouped the measurements automatically into determined number of clusters. The JPDAF used in [5] for cluster to target association was replaced with a new BP based approach, which helped to assign multiple clusters to a target of non-rigid shapes. Unary and pairwise potentials in BP were designed to exploit both locations and features of clusters to re-identify the targets when they come out of occlusions. Results showed that the proposed technique successfully assigned multiple clusters to a target and re-identifications after occlusions was also successfully achieved.

6. REFERENCES

[1] Z. Khan, T. Balch, and F. Dellaert, "MCMC-Based particle filtering for tracking a variable number of interacting targets," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, pp. 1805–1819, 2005.

[2] M. Montemerlo, W. Whittaker, and S. Thrun, "Particle filters for simultaneous mobile robot localization and people-tracking," *IEEE Int. Conf. Robotics and Automation*, pp. 695–701, 2002.

[3] M. Lindstrom and J.O. Eklundh, "Detecting and tracking moving objects from a mobile platform using a laser range scanner," *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pp. 1364–1369, 2001.

[4] X. Xu and B. Li, "Head tracking using particle filter with intensity gradient and color histogram," *IEEE Int. Conf. Multimedia and Expo*, pp. 888–891, 2005.

[5] T. D. Laet, H. Bruyninckx, and J. D. Schutter, "Shape-based online multitarget tracking and detection for targets causing multiple measurements: Variational Bayesian clustering and lossless data association," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, pp. 2477–2491, 2011.

[6] D. Gatica-Perez, G. Lathoud, J. Odobez, and I. McCowan, "Audiovisual probabilistic tracking of multiple speakers in meetings," *IEEE Trans. Audio, Speech and Language Processing*, vol. 15, no. 2, pp. 601–616, 2007.

[7] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filter for Tracking Applications*, Boston—London: Artech House Publishers, 2004.

[8] S. M. Naqvi, M. Yu, and J. A. Chambers, "A multimodal approach to blind source separation of moving sources," *IEEE J. Selected Topics in Signal Processing*, vol. 4, no. 5, pp. 895–910, 2010.

[9] Ata ur Rehman, S. M. Naqvi, R. Phan, W. Wang, and J. A. Chambers, "MCMC-PF based multiple head tracking in a room environment," *British Machine Vision Workshop*, 2012.

[10] J. Vermaak, S. J. Godsill, and P. Perez, "Monte Carlo filtering for multi-target tracking and data association," *IEEE Trans. Aerospace and Electronic systems*, pp. 309–332, 2005.

[11] S. M. Naqvi, W. Wang, M. S. Khan, M. Barnard, and J. A. Chambers, "Multimodal (audio-visual) source separation exploiting multi-speaker tracking, robust beamforming, and time-frequency masking," *IET Signal Processing*, vol. 6, pp. 466–477, 2012.

[12] H. K. Maganti, D. Gatica-Perez, and I. McCowan, "Speech enhancement and recognition in meetings with an audio-visual sensor array," *IEEE Trans. on Audio, Speech and Language Processing*, vol. 15, pp. 2257–2269, 2007.

[13] Y. Liang, S. M. Naqvi, and J. A. Chambers, "Audio-video based fast fixed-point IVA for multisource separation in a room environment," *EURASIP Journal on Advances in Signal Processing*, (Accepted for publication).

[14] P. Brasnett, L. Mihaylova, D. Bull, and N. Canagarajah, "Sequential Monte Carlo tracking by fusing multiple cues in video sequences," *Image and Vision Computing*, vol. 25, no. 8, pp. 1217–1227, 2007.

[15] A. M. Elgammal and L. S. Davis, "Probabilistic framework for segmenting people under occlusion," *IEEE Int. Conf. Computer Vision*, vol. 2, pp. 145–152, 2001.

[16] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[17] I. J. Cox, "A review of statistical data association techniques for motion correspondence," *Int. J. Computer Vision*, vol. 10, pp. 53–66, 1993.

[18] P. Felzenszwalb and D. Huttenlocher, "Efficient belief propagation for early vision," *Int. J. Computer Vision*, pp. 70:41–54, 2006.

[19] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[20] A. Bhattacharayya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bulletin of Calcutta Mathematical Society*, vol. 35, pp. 99–110, 1943.

[21] A. Kumar and C. D. Vleeschouwer, "Prioritizing the propagation of identity beliefs for multi-object tracking," *British machine vision conference*, 2012.

[22] G. Lathoud, J. M. Odobez, and D. Gatica-Perez, "AV16.3: an audio-visual corpus for speaker localization and tracking," *Proc. the MLMI'04 Workshop*, 2004.