Variational Bayesian PHD filter with Deep Learning Network Updating for Multiple Human Tracking

Pengming Feng[1], Wenwu Wang[2], Syed Mohsen Naqvi[1], Jonathon A. Chambers[1]
1. Communications, Sensors, Signal and Information Processing Group, Newcastle University, UK
Emails: {p.feng2, s.m.r.naqvi, jonathon.chambers}@newcastle.ac.uk
2. Center for Vision Speech and Signal Processing, University of Surrey, UK
Email: w.wang@surrey.ac.uk

***Abstract*—We propose a robust particle probability hypothesis density (PHD) filter where the variational Bayesian method is applied in joint recursive prediction of the state and the time varying measurement noise parameters. The proposed particle PHD filter is based on forming variational approximation to the joint distribution of states and noise parameters at each frame separately; the state is estimated with a particle PHD filter and the measurement noise variances used in the update step are estimated with a fixed point iteration approach. A deep belief network (DBN) is used in the update step to mitigate the effect of measurement noise on the calculation of particle weights in each frame. The deep learning network is trained based on both colour and oriented gradient histogram (HOG) features and then used to mitigate the measurement noise from the particle selection step, thereby improving the tracking performance. Simulation results using sequences from the CAVIAR dataset show the improvements of the proposed DBN aided variational Bayesian particle PHD filter over the traditional particle PHD filter.**

***Index Terms*—Multiple human tracking, PHD filter, deep learning, variational Bayesian**

I. Introduction

Unknown and varying number of targets cause the main problem in multiple target tracking (MTT); moreover the occlusion problem may occur which increases the challenge for reliable target tracking. A particular issue in MTT is that it is not always possible to associate measurements with particular targets and therefore false alarms and missed detections may be generated particularly in the presence of clutter, occlusion and noise [1], which can introduce detection in the number of targets and reduce the tracking accuracy in MTT.

Both the Kalman filter and particle filter have been widely used in tracking, however, in these approaches, the number of targets is assumed to be known and fixed. For a variable number of targets, the random finite set (RFS) [2] based probability hypothesis density (PHD) filter has been recently proposed for the MTT problem. The advantage of the PHD filter is that it can estimate both the number of targets and their locations, and thus avoids the need for data association techniques as part of the multiple target framework [3][4][5]. Moreover, it mitigates the computational complexity issue that often occurs in other multiple target tracking approaches such as the multiple hypothesis tracking (MHT) approach [5] by employing the intensity instead of the posterior distribution. However, the limitation of the PHD filter is that its performance can be easily affected by estimation errors caused by noise.

The particle PhD filter often assumes a priori knowledge of the measurement and dynamic model parameters, including the noise statistics. However, such knowledge is not always available in practical applications [6]. Variational Bayesian (VB) methods on the other hand have been used for a wide range of models to perform approximate posterior inference at low computational cost in comparison with the sample methods [7], thereby establishing an analytically tractable form for the joint posterior distribution of the state and measurement parameters. Moreover, in human tracking, the VB method can be used with a factor-free form of distribution of the state and measurement noise model.

In this paper, we therefore propose a novel robust PHD filter for multiple human tracking where the VB method is used to approximate the joint posterior distribution of the state and the noise variance with a factorized free form distribution. The variance of the measurement noise is also updated by this VB approach, thereby providing more stable updates for the weights in the particle PHD filter. Since in an enclosed environment, only limited human features can be extracted when using a single camera and accurate measurement of the humans can be difficult to obtain due to illumination and posture changes, we employ a deep belief network (DBN) to aid in calculating the weights for the particle based PHD filter, which utilizes the colour and oriented gradient histogram features. The DBN has the advantage that it is robust to background noise in the measurement due to the difference between human target and noises features. To evaluate the performance of our proposed robust PHD filter, we employ sequences from the CAVIAR dataset [8] which include appearance, occlusion, and disappearance of humans in the field of view of a camera. Next we introduce the proposed particle PHD filter.

II. Proposed particle PHD filter

A. Optimal Bayesian filtering

We choose the particle filter to form the fundamental framework of the PHD filter so that the weight for each particle can be calculated via a DBN to improve the accuracy in human tracking. We approximate the PHD filter with a set of weighted

978-1-4799-7444-3/15/$31.00 ©2015 IEEE

random samples using a sequential Monte Carlo method. The corresponding state and measurement model at time k can be written respectively as

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{w}_k \tag{1}$$

$$\mathbf{z}_k = \mathbf{H}\mathbf{z}_{k-1} + \mathbf{v}_k \tag{2}$$

where $\mathbf{x}_k$ denotes the state vector at time k including the 2-D position of the target, $\mathbf{z}_k$ is the measurement vector and $\mathbf{w}_k \sim N(\mathbf{0}, \mathbf{P}_k)$ is the Gaussian process noise, $\mathbf{v}_k \sim N(\mathbf{0}, \mathbf{R}_k)$ is the measurement noise with diagonal covariance matrix $\mathbf{R}_k$. $\mathbf{F}$ and $\mathbf{H}$ are respectively the state and measurement transition matrix.

We assume that the dynamic models of the states and the variance parameters within (1) and (2) are independent [9], and thus can be described as:

$$p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{x}_{k-1}, \mathbf{R}_{k-1}) = p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{R}_k|\mathbf{R}_{k-1}) \tag{3}$$

The goal of Bayesian optimal filtering of the above model is to compute the posterior distribution $p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_k)$, where $\mathbf{Z}_k = \{\mathbf{z}_k, k = 1, ..., N\}$. As described in [6], the filtering problem consists of the following steps:

1. Initialization: The recursion starts from the initial prior distribution $p(\mathbf{x}_0, \mathbf{R}_0)$.

2. Prediction: The predictive distribution of the state $\mathbf{x}_k$ and measurement noise covariance matrix $\mathbf{R}_k$ is given by the Chapman-Kolmogorov equation:

$$\begin{aligned} p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_{k-1}) = \int & p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{R}_k|\mathbf{R}_{k-1})\times \\ & p(\mathbf{x}_{k-1}, \mathbf{R}_{k-1}|\mathbf{Z}_{k-1})d\mathbf{x}_{k-1}d\mathbf{R}_{k-1} \end{aligned} \tag{4}$$

where the integral has the dimension equal to the sum of $dim(\mathbf{x}_k)$ and $dim(\mathbf{R}_k)$.

3. Update: Given the next measurement $\mathbf{z}_k$, the predictive distribution above is updated to a posterior distribution by Bayes' rule:

$$p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_k) \propto p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{R}_k)p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_{k-1}) \tag{5}$$

The integrations in the general solution are usually not analytically tractable [6]; in the following, the recursion steps are solved by using a variational Baysian approximation for the posterior update.

B. Variational approximation

As described in Section II-A, the goal of VB is to compute the posterior distribution $p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_k)$. Given that the inverse-Gamma distribution is the conjugate prior distribution for the variance of a Gaussian distribution [10], a product of inverse-Gamma distributions can be adopted to approximate the posterior distribution of $\mathbf{R}_k$; in this case, assuming the posterior distribution at time $k-1$ can be represented by

$$\begin{aligned} p(\mathbf{x}_{k-1}, \mathbf{R}_{k-1}|\mathbf{Z}_{k-1}) =& N(\mathbf{x}_{k-1}, \mu_{k-1}, \mathbf{P}_{k-1})\prod_{i=1}^{d}\times \\ & IG(\sigma^2_{k-1,i}|\alpha_{k-1,i}, \beta_{k-1,i}) \end{aligned} \tag{6}$$

where d is the dimension of the measurement noise vector and $N(\mathbf{x}_{k-1}, \mu_{k-1}, \mathbf{P}_{k-1})$ denotes a Gaussian probability density function for the random variable $\mathbf{x}_{k-1}$ with mean μ_{k-1} and covariance matrix $\mathbf{P}_{k-1}$, and $IG(\sigma^2_{k-1,i}|\alpha_{k-1,i}, \beta_{k-1,i})$ denotes the inverse-Gamma distribution, which has the degree of freedom parameter $\alpha_{k-1,i}$ and the scalar parameter $\beta_{k-1,i}$. With the assumption that the dynamic models of the state and measurement noise variance are independent, the joint predictive distribution remains a factored form [9] of a Gaussian distribution and an inverse Gamma distribution

$$\begin{aligned} & p_{k|k-1}(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_k) = p_{k|k-1}(\mathbf{x}_k|\mathbf{Z}_{k-1})p_{k|k-1}(\mathbf{R}_k|\mathbf{Z}_{k-1}) \\ & = N(\mathbf{x}_{k|k-1}, \mu_{k|k-1}, \mathbf{P}_{k|k-1})\prod_{i=1}^{d} IG(\sigma^2_{k|k-1,i}|\alpha_{k|k-1,i}, \beta_{k|k-1,i}) \end{aligned} \tag{7}$$

where in the joint posterior distribution, the state and measurement noise variance will be coupled with the likelihood function, which makes the exact posterior intractable.

The next step is to derive an analytical expression for the posterior distribution within the update equation; in order to make the computation tractable, an approximation to the posterior distribution is formed [7]. The standard VB approach is employed and a free form factored approximate distribution $p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_k)$ can be described as

$$p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_k) \approx Q_{\mathbf{x}}(\mathbf{x}_k)Q_{\mathbf{R}}(\mathbf{R}_k) \tag{8}$$

where $Q_{\mathbf{x}}(\mathbf{x}_k)$ and $Q_{\mathbf{R}}(\mathbf{R}_k)$ are respectively a Gaussian distribution and inverse Gamma distribution as follows:

$$Q_{\mathbf{x}}(\mathbf{x}_k) = N(\mathbf{x}_k, \mu_k, \mathbf{P}_k) \tag{9}$$

$$Q_{\mathbf{R}}(\mathbf{R}_k) = IG(\sigma^2_{k,i}|\alpha_{k,i}, \beta_{k,i}) \tag{10}$$

then the approximate posterior densities can be determined by minimizing the Kullback-Leibler(KL) divergence between the separable approximation and the true posterior density expressed as

$$\begin{aligned} & KL\{Q_{\mathbf{x}}(\mathbf{x}_k)Q_{\mathbf{R}}(\mathbf{R}_k)||p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_k)\} = \\ & \int Q_{\mathbf{x}}(\mathbf{x}_k)Q_{\mathbf{R}}(\mathbf{R}_k)\log\frac{Q_{\mathbf{x}}(\mathbf{x}_k)Q_{\mathbf{R}}(\mathbf{R}_k)}{p(\mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_k)}d\mathbf{x}_k d\mathbf{R}_k \end{aligned} \tag{11}$$

In order to minimize the KL-divergence, methods from calculus of variations [6] are employed. Using the alternating optimisation, the probability densities $Q_{\mathbf{x}}(\mathbf{x}_k)$ and $Q_{\mathbf{R}}(\mathbf{R}_k)$ are calculated in turn, while keeping the other fixed, yielding:

$$Q_{\mathbf{x}}(\mathbf{x}_k) \propto E\left\{\int \log p(\mathbf{z}_k, \mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_{1:k-1})Q_{\mathbf{R}}(\mathbf{R}_k)d\mathbf{R}_k\right\} \tag{12}$$

$$Q_{\mathbf{R}}(\mathbf{R}_k) \propto E\left\{\int \log p(\mathbf{z}_k, \mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_{1:k-1})Q_{\mathbf{x}}(\mathbf{x}_k)d\mathbf{x}_k\right\} \tag{13}$$

Since the two equations are coupled, they cannot be solved directly, however, computing the expectation of the first equation yields the following equation

$$\begin{aligned} & \int \log p(\mathbf{z}_k, \mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_{1:k-1})Q_{\mathbf{R}}(\mathbf{R}_k)d\mathbf{R}_k = \\ & -0.5(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k)^T\langle\mathbf{R}_k^{-1}\rangle_{\mathbf{R}}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k) \\ & -0.5(\mathbf{x}_k - \mathbf{F}_k\mathbf{x}_{k-1})^T(\mathbf{P}_k^{-1})(\mathbf{x}_k - \mathbf{F}_k\mathbf{x}_{k-1}) + C_1 \end{aligned} \tag{14}$$

where $\langle\cdot\rangle_{\mathbf{R}} = \int(\cdot)Q_{\mathbf{R}}(\mathbf{R}_k)d\mathbf{R}_k$ denotes the expected value with respect to the approximation distribution $Q_{\mathbf{R}}(\mathbf{R}_k)$ and C_1 denotes the terms independent of $\mathbf{x}_k$.
Similarly, the second expectation can be computed as follows

$$\int \log p(\mathbf{z}_k, \mathbf{x}_k, \mathbf{R}_k|\mathbf{Z}_{k-1})Q_{\mathbf{x}}(\mathbf{x}_k)d\mathbf{x}_k = -\sum_{i=1}^{d}(\frac{3}{2} + \alpha_{k,i})\ln(\sigma^2_{k,i}) - \sum_{i=1}^{d}\frac{\beta_{k,i}}{\sigma^2_{k,i}} - \frac{1}{2}\sum_{i=1}^{d}\frac{\langle(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k)^2_i\rangle_{\mathbf{x}_k}}{\sigma^2_{k,i}} + C_2 \quad (15)$$

where $\langle\cdot\rangle_{\mathbf{x}} = \int(\cdot)Q_{\mathbf{x}}(\mathbf{x}_k)d\mathbf{x}_k$, and the parameters $\mu_k, \mathbf{P}_k, \alpha_{k,i}$ and $\beta_{k,i}$ are the solutions to the following coupled set of equations:

$$\begin{aligned}
\mu_k &= \mu_{k-1} + \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k^T\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \hat{\mathbf{R}}_k)^{-1}(\mathbf{z}_k - \mathbf{H}_k\mu_{k|k-1}) \\
\mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k^T\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \hat{\mathbf{R}}_k)^{-1}\mathbf{H}_k^T\mathbf{P}_{k|k-1} \\
\alpha_{k,i} &= \hat{\alpha}_{k-1,i} + \frac{1}{2} \\
\beta_{k,i} &= \hat{\beta}_{k-1,i} + \frac{1}{2}[(\mathbf{z}_k - \mathbf{H}_k\hat{\mu}_k)^2_i + (\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^T)_i]
\end{aligned} \quad (16)$$

and the estimated covariance matrix $\hat{\mathbf{R}}_k$ is

$$\hat{\mathbf{R}}_k = diag\left\{\frac{\beta_{k,1}}{\alpha_{k,1}}, ..., \frac{\beta_{k,m}}{\alpha_{k,m}}\right\} \quad (17)$$

where ' $\hat{}$ ' denotes the estimate of the parameters. Following (16) and (17), the process of variational Bayes measurement parameter updating can be described as

1. Prediction: Compute the parameters of the predicted distribution as follows:

$$\begin{aligned}
\hat{\alpha}_{k-1,i} &= \rho_i\alpha_{k-1,i} \\
\hat{\beta}_{k-1,i} &= \rho_i\beta_{k-1,i} \\
\hat{\mu}_{k|k-1} &= \mathbf{F}_{k-1}\hat{\mu}_{k-1|k-1} \\
\mathbf{P}_{k|k-1} &= \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}
\end{aligned} \quad (18)$$

where $\rho_i \in (0, 1]$ is a scalar used for the extension of the noise fluctuations.

2. Update: In the update step, a fixed point iteration method is employed to achieve the best solution of the following equations. First set $\mu_k^0 = \hat{\mu}_{k|k-1}, \mathbf{P}_k^0 = \hat{\mathbf{P}}_{k|k-1}, \alpha_{k,i}^0 = \frac{1}{2} + \hat{\alpha}_{k-1,i}$ and $\beta_{k,i}^0 = \hat{\beta}_{k-1,i}$, for $i = 1, ..., d$. Then use the fixed-point iteration to achieve the solution of (16) and (17) for ℓ steps: then set $\beta_{k,i} = \beta_{k,i}^{\ell}, m_k = m_k^{\ell}$ and $P_k = P_k^{\ell}$, after obtaining the optimal solution for the equations, the parameters are updated within the measurement model of the particle filter which thereby helps calculate the weights of each particle. In the next section, the VB aided particle filter is combined with the DBN to update the weights of the particles.

C. Deep learning networks

Deep learning methods use a hierarchical learning method when training the classifier. As described in [11], hierarchical learning uses natural progression from low level to high level structure as seen in natural complexity, so it is easier to monitor what is being learnt and to guide the machine evolving to better subspaces. For example, when images are input into the system, the first layer of the system represents the 'edges' of the feature; the second layer represents the 'object parts' of the feature and the third layer represents the objects in the features [11]. As described in [12], wherein a restricted Boltzmann machine (RBM) is employed, DBNs are graphical models which learn to extract deep hierarchical representationS of the training data. DBNs provide the advantage of an intelligent method to use the limited information to train a classifier to calculate the weights of particles in human tracking, particularly with a single camera. The deep belief network we use is represented detailedly in [11], which is composed of one input layer, two RBMs based hidden layers and a one-class output layer. The process of training the RBM and DBN are described in [12] in detail. After training the DBNs following the process, the weight for each particle $\mathbf{x}_k^i$ can be calculated as

$$p(\mathbf{x}_k^i) = e^{(c\cdot\mathbf{1}\cdot\mathbf{W}^{(m)})} \quad (19)$$

where $\mathbf{1}$ is a $(1 \times j)$ all-one vector, c is a constant we set for calculating the weights for the particles and $\mathbf{W}^{(m)}$ is the weights from the last layer of DBN. In this way, the likelihood for each particle is obtained and these weights can then be taken as the input to the updating step of the PHD filter for MTT as discussed in the next section.

D. Particle PHD implementation

To formulate the PHD filter the RFS framework is employed [13]. We denote $\mathbf{D}_{k|k}(\mathbf{x})$ as the PHD filter at discrete time k associated with the multi-target posterior density $p_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k})$, where $\mathbf{X}_k = \{\mathbf{x}_k^m, m = 1, ..., M\}$ includes the 2D position of all the human targets, $\mathbf{x}_k^m$ denotes the state of the m^{th} target at time k, M is the number of targets and $\mathbf{Z}_{1:k}$ denotes the measurements up to time k. The PHD prediction step is defined as:

$$\mathbf{D}_{k|k-1}(\mathbf{x}_k^m) = \int \phi_{k|k-1}(\mathbf{x}_k^m, \xi)\mathbf{D}_{k-1|k-1}(\xi)d(\xi) + \Upsilon_k \quad (20)$$

where Υ_k is the intensity function of the new target birth RFS, $\phi_{k|k-1}(\mathbf{x}_k^m, \xi)$ is the analogue of the state transition probability in the single target case which is calculated from

$$\phi_{k|k-1}(\mathbf{x}_k^m, \xi) = e_{k|k-1}(\xi)f_{k|k-1}(\mathbf{x}_k^m|\xi) + \beta_{k|k-1}(\mathbf{x}_k^m|\xi) \quad (21)$$

in which $f_{k|k-1}$ is the multi-target transition density, $e_{k|k-1}(\xi)$ is the probability that the target still exists at time k and $\beta_{k|k-1}(\mathbf{x}_k^m|\xi)$ is the intensity of the RFS that a target is spawned from the state ξ. The PHD update step is defined as [14]:

$$\mathbf{D}_{k|k}(\mathbf{x}_k^m) = \left[p_M(\mathbf{x}_k^m) + \sum_{z\in\mathbf{Z}_k}\frac{\psi_{k,z}(\mathbf{x}_k^m)}{\kappa_k + \langle\psi_{k,z}, \mathbf{D}_{k|k-1}\rangle}\right]\mathbf{D}_{k|k-1}(\mathbf{x}_k^m) \quad (22)$$

where p_M is the missing detection probability, $\psi_{k,z}(\mathbf{x}_k^m) = (1 - p_M)g_k(\mathbf{z}|\mathbf{x}_k^m)$ is the single-target likelihood defining the probability that a measurement $\mathbf{z}$ is generated by a target with state $\mathbf{x}_k^m$, κ_k is the clutter intensity.

When the sequential Monte Carlo method is employed to approximate the PHD filter, two fundamental steps in the particle filter are sequential importance sampling and resampling. The basic principle of importance sampling is to represent a PDF $p(\mathbf{X}_k)$ by a set of random particles associated with the weights, where $\mathbf{X}_k^N = \{\mathbf{x}_k^i, i = 1, ..., N\}$, and N is the number of particles we employed in the particle filter. Given a set of particles [15]

$$\{w_{k-1}^i, \mathbf{x}_{k-1}^i\}_{i=1}^N \tag{23}$$

which are independently drawn from importance sampling density $q(\mathbf{X}_k)$ [15], the weight of each particle can be calculated as

$$w_k^i = p(\mathbf{x}_k^i)/q(\mathbf{x}_k^i) \tag{24}$$

thus $p(\mathbf{X}_k^N)$ can be approximated as

$$p(\mathbf{X}_k^N) \approx \sum_{i=1}^N w_k^i \delta(\mathbf{X}_k^N - \mathbf{x}_k^i) \tag{25}$$

where $\delta(\cdot)$ denotes the Dirac delta function.

Assuming the particles for the PHD filter are independently drawn from the PDF $p(\mathbf{X}_{k-1}^N|\mathbf{Z}_{1:k-1})$, the particles $\mathbf{x}_k^i, i = 1, ..., N$ are propagated and updated by the Gaussian distribution, which are approximately distributed as $p(\mathbf{X}_k^N|\mathbf{Z}_k)$ [16]. In this case, the proposed filter is an approximate implementation of the relationship between the prediction and updating step of the filter. The prediction and updating step can be described as follows.

1. Prediction: Draw particle $\mathbf{x}_{k-1}^i$ from $\mathbf{X}_{k-1}^N$ and feed it into the prediction step to obtain particles at time k. Thus the prediction model can be calculated as

$$p(\mathbf{X}_k^N|\mathbf{Z}_{k-1}) = \int p(\mathbf{X}_k^N|\mathbf{X}_{k-1}^N)p(\mathbf{X}_{k-1}^N|\mathbf{Z}_{k-1})d\mathbf{X}_{k-1}^N \tag{26}$$

2. Measurement update: Upon the receipt of the measurement $\mathbf{Z}_k$, the likelihood of each prior sample $\mathbf{x}_k^i, i = 1, ...N$, can be evaluated and drawn independently from importance sampling density $q(\mathbf{X}_k^N|\mathbf{Z}_k)$ [16]. The importance weight for each prior sample can be calculated as:

$$w_k^i = \frac{p(\mathbf{Z}_k|\mathbf{x}_k^i)p(\mathbf{X}_k^N|\mathbf{Z}_{k-1})}{q(\mathbf{x}_k^i|\mathbf{Z}_k)} \tag{27}$$

Equations (26) and (27) above form the basis of the proposed particle PHD filter. By introducing (26) into the PHD updating formula (22), we can obtain the particle PHD updating equation:

$$w_k^{(i)} = \left[p_M(x_k^{(i)}) + \sum_{z\in\mathbf{Z}_k} \frac{\psi_{k,z}(x)}{\kappa_k(z) + C_k(z)}\right] w_{k-1}^{(i)} \tag{28}$$

where

$$C_k(z) = \sum_{j=1}^{L_{k-1}+J_k} \psi_{k,z}(\widetilde{x}_k^{(i)})\widetilde{w}_{k-1}^{(i)} \tag{29}$$

The above work underpins the proposed particle PHD filter for human tracking; in the next subsection, human features can be extracted to train a deep belief network in the utilization of particle selection and the updating part of the particle PHD filter to mitigate the noise from the environment.

When the variational Bayes method is applied within the particle PHD filter, an iteration for updating the measurement noise should be added before the updating step for the PHD filter, in this case, the recursive particle PHD filter follows the following process

Algorithm 1 Processing of the variational Bayes particle PHD filter

At each time k, sample J_k particles for each new-born target from background subtraction, after setting the weight to the new-born target particles as $1/J_k$, we can obtain the particle set $\{\mathbf{x}_k, w_k\}_{i=1}^N$ where J_k is the new-born target number and N is the number of particles at time k.
For each particle $i = 1, ..., N$
Initialize the state $\mathbf{x}_k^i$ and hyper parameters: $\alpha_k^i, \beta_k^i, \mathbf{P}_k^i, \mathbf{R}_k^i$.
1 Predict the noise parameters using (18).
2 Sample the particles.
3 Update the measurement noise parameters using Variational Bayesian approach with fixed point iteration with (16) to (18) described in Section II-B.
4 Calculate the weights for each particle with DBN with (19) to aid the weights updating step.
5 Evaluate the weights using (22).
6 Resample the particles.
END

From the above process, we can obtain the Variational Bayes particle PHD filter; in the next section, simulation results and comparisons will be given.

III. Simulation Results

In order to evaluate the performance of the proposed robust particle PHD filter for multiple target tracking, we employed datasets from the EC Funded CAVIAR project [8], Video EnterExitCrossingPaths1cor and EnterExitCrossingPaths1front. There are four human targets appearing, occluding each other, and disappearing in the shopping mall environment. The number of particles is set to be 1000. When compared with the traditional PHD filter, our proposed PHD filter shows its improvement in reducing the mean of the error and standard deviation, where in scenario 1 which denotes the video EnterExitCrossingPaths1cor, the mean of the error is reduced from 13.45 to 11.89 with the standard deviation is reduced from 16.68 to 12.85; in scenario 2 which denotes the video EnterExitCrossingPaths1front, the mean of the error is reduced from 34.54 to 22.26 with the standard deviation is reduced from 19.87 to 11.85. To see this further the optimal subpattern assignment (OSPA) [17] is employed. The OSPA comparisons for both scenarios are shown in Figs. 1 and 2 respectively.

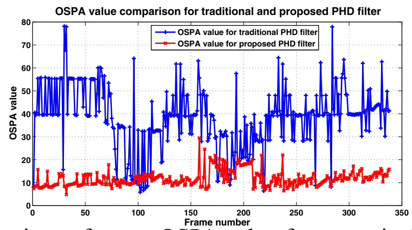

Fig. 1: Comparison of target OSPA value for scenario 1 between the proposed and the traditional PHD filter, where the blue line denotes the OSPA value for the traditional particle PHD filter with an average value of 38.25 and the red line denotes the OSPA value for the proposed particle PHD filter with an average value of 11.79.

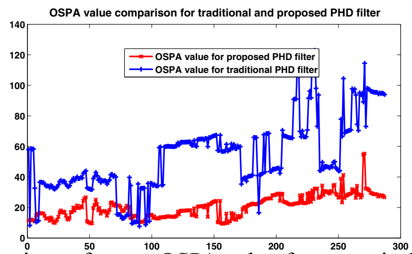

Fig. 2: Comparison of target OSPA value for scenario 2 between the proposed and the traditional PHD filter, where the blue line denotes the OSPA value for the traditional particle PHD filter with an average value of 53.55 and the red line denotes the OSPA value for the proposed particle PHD filter with an average value of 20.62.

To verify the performance improvement, a new performance metric for multiple target tracking based on optimal subpattern assignment, named OSPAMT [18] is employed, the OSPAMT includes both cardinality and localization errors versus time for the video 'EnterExitCrossingPaths1cor' are shown as From

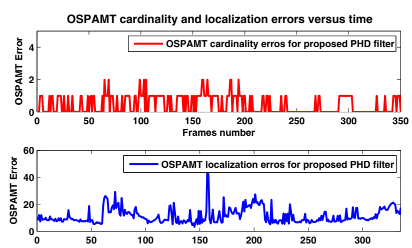

Fig. 3: OSPAMT cardinality and localization errors versus time for video 'EnterExitCrossingPaths1cor'.

the comparison of both OSPA and OSPAMT results, it can be observed that our proposed PHD filter performs much better than the baseline PHD filter since in both scenarios the OSPA value and OSPAMT are much reduced, which means false alarm and missing detection performance of our proposed PHD filter is likely to be much improved compared to the traditional PHD filter. The error for the PHD filter is mostly from the false alarm and miss detection caused by the noise from the background subtraction step, which can be mitigated by the DBN classifier in our proposed PHD filter. Both the mean of the error and the localization errors from OSPAMT are mitigated with the aid of our Variational Bayesian step for the update of the measurement parameters. This is because the parameters are adaptive to the posterior of the joint distribution of the measurements parameters and the state model. From the results, we can see that the proposed PHD filter improves the tracking accuracy over the traditional PHD filter.

IV. Conclusions and Future Work

In this paper, we have presented a new DBN aided variational Bayes particle PHD filter for human tracking, which is based on recursively forming separable approximations to the joint distribution of state and noise parameters by variational Bayesian methods and the noise mitigating step with the aid of DBN. The performance of the algorithm has been demonstrated by simulations. In our future work, more sequences will be used to evaluate the proposed PHD filter and an interaction model will be used to describe the state model of the targets.

V. Acknowledgment

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014307 and the MOD University Defence Research Collaboration in Signal Processing.

References

[1] D. E. Clark, "Multiple target tracking with the probability hypothesis density filter." PhD thesis, Heriot-Watt University, 2006.

[2] E. Maggio, M. Taj, and A. Cavallaro, "Efficient multi-target visual tracking using Random Finite Sets," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, pp. 1–12, 2008.

[3] D. E. Clark and J. Bell, "Data association for the PHD filter." *The Second International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, 2005, pp. 217–222.

[4] K. Panta, B.-N. Vo, and S. Singh, "Novel data association schemes for the probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 556–570, 2007.

[5] E. Maggio and A. Cavallaro, *Video Tracking*. John Wiley and Sons, Ltd, 2011.

[6] S. Sarkka and A. Nummenmaa, "Recurive noise adaptive Kalman filering by Variational Bayesian approximations," *IEEE Transactions on Automatic Control*, pp. 596–600, 2009.

[7] T. S. Jaakkola, "Tutorial on variational approximation methods," *In Advanced Mean Field Methods:Theory And Practice,* M. Oper and D. Saad, Eds. Cambidge, MA: MIT Press , pp. 129–159, 2001.

[8] R. Fisher. Caviar case scenarios. [Online]. Available: http://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1/

[9] G. Zhang, F. Lian, C. Han, and S. Han, "An improved PHD filter based on Variantional Bayesian method for multi-target tracking," *in Proc. IEEE 17th Conference on Information Fusion*, pp. 1–6, 2014.

[10] J. Christmas and R. Everson, "Robust autoregression: Student-t innovation using variational Bayes," *IEEE Transactions on Signal Processing*, vol. 59, pp. 48–57, 2011.

[11] P. Feng, M. Yu, , S. M. Naqvi, and J. Chambers, "Deep learning for posture analysis in fall detection," *in Proc. IEEE 19th of the International Conference on Digital Signal Processing*, pp. 1–6, 2014.

[12] G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, pp. 504 – 507, 2006.

[13] R. Mahler, "A theoretical foundation for the Stein-Winter Probability Hypothesis Density (PHD) multitarget tracking approach." DTIC Document, 2000.

[14] A. Doucet, B. Vo, C. Andrieu, and M. Davy, "Particle filtering for multi-target tracking and sensor management." *The International Conference on Informantion Fusion, Annaplis, MD*, 2002, pp. 474–481.

[15] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian bayesian tracking." *IEEE Transactions on Signal Processing*, vol. 50, pp. 2454–2467, 2002.

[16] S. Li, H. Wang, and T. Chai, "A t distribution based particle filter for target tracking," *Proceedings of the 2006 American Control Conference*, pp. 2191–2196, 2006.

[17] B. Ristic, B. N. Vo, and D. Clark, "Performance evaluation of multi-target tracking using the OSPA metric," *in Proc. IEEE 13th Conference on Information Fusion*, pp. 1–7, 2010.

[18] T. Vu and R. Evans, "A new performance metric for multiple target tracking based on optimal subpattern assignment," in *17th IEEE International Conference on Information Fusion (FUSION)*, 2014, pp. 1–8.