A ROBUST STUDENT'S-T DISTRIBUTION PHD FILTER WITH OCSVM UPDATING FOR MULTIPLE HUMAN TRACKING

Pengming Feng[1], *Miao Yu*[2], *Syed Mohsen Naqvi*[2], *Wenwu Wang*[1], *Jonathon A. Chambers*[2]

1. Center for Vision Speech and Signal Processing, Department of Electronic Engineering,
University of Surrey, UK
{p.feng, w.wang}@surrey.ac.uk
2. Advanced Signal Processing Group, Loughborough University, UK
{m.yu, s.m.r.naqvi, j.a.chambers}@lboro.ac.uk

ABSTRACT

We propose a novel robust probability hypothesis density (PHD) filter for multiple target tracking in an enclosed environment, where a one-class support vector machine (OCSVM) is used in the update step for combining different human features to mitigate the effect of measurement noise on the calculation of particle weights. A Student's-t distribution is employed to improve the robustness of the filters whose tail is heavier than the Gaussian distribution and thus has the potential to cover more widely-spread particles. The OCSVM is trained based on both colour and oriented gradient (HOG) histogram features and then used to mitigate the measurement noise from the particle selection step, thereby improve the tracking performance. To evaluate the proposed PHD filter, we employed two sequences from the CAVIAR dataset and used the optimal subpattern assignment (OSPA) method as an objective measure. The results show that the proposed robust PHD filter outperforms the traditional PHD filter.

***Index Terms*—** Multiple human tracking, PHD filter, Student's-t distribution, OCSVM

1. INTRODUCTION

In multiple target tracking (MTT), the number of targets is often unknown and varies with time. In addition, the occlusion problem may occur and this further increases the challenges 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 detection may be generated particularly in the presence of clutters, occlusion and noise [1].

Several methods could be used to address these challenges in MTT. Earlier methods include the Kalman filter and particle filter where 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 as often occurs in other multiple target tracking approaches such as the multiple hypothesis tracking (MHT) [5]. However, the limitation of the PHD filter is that its performance can be easily affected by estimation errors caused by the noise.

In this paper, we propose a novel robust PHD filter for multiple human tracking, where we employ a Student's-t distribution for the state and the observation models. This distribution has an advantage over the traditional Gaussian distribution in the sense that its tail is heavier and can potentially cover more widely-spread particles. In an enclosed environment, accurate measurement of the humans can be difficult to obtain due to illumination and posture changes. Therefore, we employ a one class support vector machine (OCSVM) to calculate the weights for the particle based PHD filter, which utilizes the colour and oriented gradient histogram features due to their accuracy in describing the human targets. The OCSVM is shown to be robust against 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 two sequences from the CAVIAR dataset which includes appearance, occlusion, disappearance of humans in the field of view of a camera. It is shown that our proposed PHD filter can obtain more accurate results and perform better when tracking a variable number of humans in an enclosed environment.

2. ROBUST STUDENT'S-T DISTRIBUTED PHD FILTER WITH OCSVM AIDED PARTICLE UPDATE

To formulate the PHD filter the RFS framework is employed [6]. We denote $\mathbf{D}_{k|k}(\mathbf{x})$ as the PHD 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

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

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:

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

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

$$\phi_{k|k-1}(\mathbf{x}_k^m, \xi) = e_{k|k-1}(\xi)f_{k|k-1}(\mathbf{x}_k^m|\xi) + \beta_{k|k-1}(\mathbf{x}_k^m|\xi) \tag{2}$$

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 [7]:

$$\mathbf{D}_{k|k}(\mathbf{x}_k^m) = \left[p_M(\mathbf{x}_k^m) + \sum_{z\in\mathbf{Z}_k} \frac{\psi_{k,z}(\mathbf{x}_k^m)}{\kappa_k + \langle\psi_{k,z}, \mathbf{D}_{k|k-1}\rangle}\right]\mathbf{D}_{k|k-1}(\mathbf{x}_k^m) \tag{3}$$

where p_M is the missing detection probability, $\psi_{k,z}(\mathbf{x}_k^m) = (1 - p_M)g_k(z|\mathbf{x}_k^m)$ is the single-target likelihood defining the probability that a measurement z is generated by a target with state $\mathbf{x}_k^m$, κ_k is the clutter intensity, and $\langle f, g\rangle = \int f(x)g(x)dx$ [5].

2.1. The Student's-t distributed Sequential Monte Carlo (SMC) method

There are numerical solutions [8] for the integrals in (1) and (3), one of which is obtained using a sequential Monte Carlo method that approximates the PHD with a set of weighted random samples, which is called the particle PHD filter and is the focus of this paper. We use this method because it performs well in the non-Gaussian noise and non-linear model framework. 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 $\mathbf{X}_k = \{\mathbf{x}_k^i, i = 0, ..., N\}$, N is the number of particles we employed in the particle filter. Given a set of particles [9]

$$\{w_{k-1}^i, \mathbf{x}_{k-1}^i\}_{i=1}^N \tag{4}$$

which are independently drawn from importance sampling density $q(\mathbf{X}_k)$ [9], the weight of each particle can be calculated as

$$w_k^i = p(\mathbf{x}_k^i)/q(\mathbf{x}_k^i) \tag{5}$$

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

$$p(\mathbf{X}_k) \approx \sum_{i=1}^N w_k^i \delta(\mathbf{X}_k - \mathbf{x}_k^i) \tag{6}$$

where $\delta(\cdot)$ denotes the Dirac delta measure. Since the efficiency of the particle filter is highly dependent on the choice of the importance sampling distribution, many attempts have been made to its construction. As described in [8], previous approaches to density estimation have mostly focused on Gaussian filters, but these are known to be sensitive to outliers. Hence a Student's-t distribution method is proposed here to improve the robustness of the PHD filter since its tails are heavier than a Gaussian distribution, which helps to potentially cover more widely-spread particles.

Assuming the particles for the PHD filter are independently drawn from the PDF $p(\mathbf{X}_{k-1}|\mathbf{Z}_{1:k-1})$, we employ the Student's-t distribution to propagate and update the particles $\mathbf{x}_k^i, i = 1, ..., N$, which are approximately distributed as $p(\mathbf{X}_k|\mathbf{Z}_k)$ [8]. 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 $\mathbf{X}_{k-1}$ and feed it into the prediction step to obtain particles at time k. If the distribution of $\mathbf{X}_{k-1}$ has a Student's-t distribution or stays close to it, then $p(\mathbf{X}_k|\mathbf{Z}_k)$ can be approximated as a Student's-t distribution. Thus the prediction model can be calculated by

$$p(\mathbf{X}_k|\mathbf{Z}_{k-1}) = \int p(\mathbf{X}_k|\mathbf{X}_{k-1})p(\mathbf{X}_{k-1}|\mathbf{Z}_{k-1})d\mathbf{X}_{k-1} \tag{7}$$

$$\approx t(\mathbf{X}_k|\mathbf{\Omega}_k) \tag{8}$$

where $t(\cdot)$ is the predicted PDF and $\mathbf{\Omega}_k$ is the importance density for the Student's-t distribution of the PHD filter.

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 $q(\mathbf{X}_k|\mathbf{Z}_{1:k})$ [8]. The importance weight for each prior sample can be calculated as:

$$w_k^i = \frac{p(\mathbf{Z}_k|\mathbf{x}_k^i)t(\mathbf{X}_k|\mathbf{\Omega}_k)}{q(\mathbf{x}_k^i|\mathbf{Z}_k)} \tag{9}$$

Equations (8) and (9) form the basis of the proposed robust Student's-t distribution particle PHD filter.

2.2. Human feature extraction

We detect the target in video using the background subtraction method [10]. The new-born target for the PHD filter can thereby be obtained more accurately than by selecting particles in the whole frame randomly. In this paper, we used a codebook method for background subtraction [11]. There is no parametric assumption on the codebook model and it has several advantages: it has the capability of coping with illumination changes, and the potential to capture structural background motion over a long period of time under limited memory. The detail of the algorithm is described in [11].

The resulting raw background subtraction results generally contain many noise artifacts, which include small 'salt and pepper' [11] and large noises caused by the problem of poor illumination and similar colour between the foreground and background information. This may be regarded as a new born target in the prediction step of the PHD filter and cause the occurrence of a false alarm. Since the noise can be distinguished from the human target by a classifier; in this paper, we used an OCSVM classifier with the aid of colour and oriented gradient histogram features of a human target to perform classification and assist the calculation of particle weights.

2.3. One class support vector machine

The OCSVM scheme is proposed in [12]. The basic idea is that given a data set drawn from an underlying probability distribution p, the OCSVM estimates a function f to describe its 'support region' (where a sample of p most likely comes from), where the corresponding values of the function f are larger than a particular threshold value.
To design the classifier, the following quadratic problem needs to be solved based on the dataset $\mathbf{S} = [\mathbf{s}_1, ..., \mathbf{s}_N]$ (assuming the dataset comes from a distribution p as mentioned previously) as:

$$\min_{\mathbf{w},\varsigma,\rho} \quad \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu N}\sum_{i=1}^{N}\varsigma_i - \rho$$
$$\text{subject to } (\mathbf{w}\cdot\Phi(\mathbf{x}_i)) \geq \rho - \varsigma_i, \quad \xi_i \geq 0 \qquad (10)$$

where $\nu \in (0, 1]$ and the nonzero slack variables $\varsigma = [\varsigma_1, ..., \varsigma_N]$ are introduced to allow for the possibility of outliers (the data points which are not drawn from the supporting region) and $\Phi(\cdot)$ is a kernel function which maps the original data into a higher dimensional space for better separation as in [21]. For a new test point $\mathbf{x}$, the decision function for estimating whether it comes from the distribution determined by $\mathbf{X}$ is:

$$f(x) = (\mathbf{w}\cdot\Phi(\mathbf{x})) - \rho \qquad (11)$$

In the application of multiple human tracking, the colour and oriented gradient features from multiple human regions are applied as the dataset for training the OCSVM classifier, which can be used to estimate the likelihood function value for each particle. For each particle i at time instance k, the corresponding feature value $\mathbf{x}_k^i$ is extracted and the corresponding likelihood function based on each measurement z, $\psi_{k,z}(\mathbf{x}_k^i)$ could be estimated as:

$$\psi_{k,z}(\mathbf{x}_k^i) = e^{(c\cdot f(\mathbf{x}_k^i))} \qquad (12)$$

where c is a constant we set for calculating the weights for the particles. 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.

2.4. The robust Student's-t distribution based PHD filter with OCSVM aided particle update

The construction of the proposed PHD filter can thus be built as shown by the flowchart in Fig. 1, where the main steps are described below in detail.

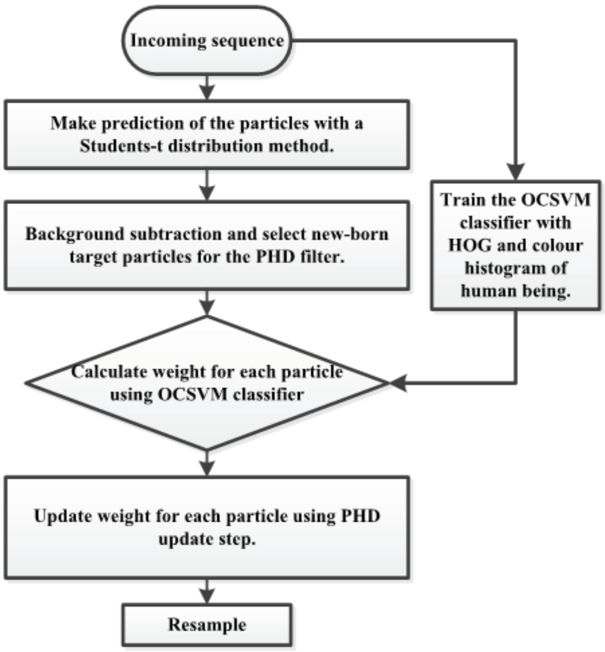

Fig. 1: The flowchart for the proposed human tracking method, which is separated by background subtraction, prediction, weight calculation, updating and resampling part.

1. Background subtraction. By employing the method described in a previous section, some background subtraction results are shown in Fig. 2.

Fig. 2: Background subtraction results for three frames in the sequence, where the green part in the first figure shows the occlusion of two human targets, which may cause miss detection; the red part in the second figure shows the appearance of another human target and the yellow part in the third figure shows the disappearance of a human target; as shown in the figure, there are noises in the background subtraction results, including the patches noises and salt and pepper noise, which may cause false alarm in the multiple tracking work.

From the results we found there are many noise patches if we only use background subtraction to provide the human

features, which may cause false alarms and miss detection. In order to avoid this, we use an OCSVM classifier to classify the noise and the foreground part of the human targets since the features extracted from them are different, and we can also calculate the weight for each particle at the same time.
2. OCSVM classifier. To train the OCSVM classifier, firstly, we extract the colour and oriented gradient histograms of 34 human being from the dataset, including different human targets, posture and colours of clothing to train the OCSVM classifier. After this, the OCSVM classifier can be used to obtain the weights for the predicted particles in the PHD filter.
3. PHD updating. After obtaining the particles and the weights of the surviving targets from Equations (8) and (9), the particles of the new born targets are drawn from foreground objects using background subtraction, whose weights are given as

$$\widetilde{w}_k^i = 1/J_k \tag{13}$$

where J_k is the number of new born targets at time k and '$\widetilde{\cdot}$' denotes the value from estimation. Once the new set of observations is available, we can substitute the approximation of $\mathbf{D}_{k|k-1}(\mathbf{x}_k^i)$ into (3) and the weights of each particles are updated as

$$\widetilde{w}_k^i = \left[p_M(\widetilde{\mathbf{x}}_k^i) + \sum_{\forall z \in \mathbf{Z}_k} \frac{\psi_{k,z}(\widetilde{\mathbf{x}}_k^i)}{\kappa_k(z) + C_k(z)} \right] \widetilde{w}_{k-1}^i \tag{14}$$

where

$$C_k(z) = \sum_{j=1}^{N+J_k} \psi_{k,z}(\widetilde{\mathbf{x}}_k^i) \widetilde{w}_{k-1}^i \tag{15}$$

At each iteration k, J_k new particles are added to the old N particles for the new born targets; to limit the growth of the number of particles, a resampling step is performed after the update step, the algorithm for the resampling step is described as Algorithm 1 below [10]:

Algorithm 1 Resampling step of the particle PHD filter

$\{\{\widetilde{w}_k^i, \widetilde{\mathbf{x}}_k^i\}_{i=1}^{N+J_k}\} \rightarrow \{w_k^i, \mathbf{x}_k^i\}_{i=1}^{N}$
Compute the target number at time k
$\hat{N}_k = \sum_{i=1}^{N+J_k} \widetilde{w}_k^i$
Initialize the cumulative probability $c_1 = 0$
$c_i = c_{i-1} + \frac{\widetilde{w}_k^{(i)}}{\hat{N}_k}, i = 2, ..., N + J_k$
Draw a starting point $\mu_1 \sim [0, N^{-1}]$
For $j = 1, ..., N$
$\mu_j = \mu_1 + N^{-1}(j-1)$
while $\mu_j > c_i$, $i = i + 1$. End while
$w_k^{(i)} = \widetilde{w}_k^{(i)}$
$\mathbf{x}_k^{(i)} = \widetilde{\mathbf{x}}_k^{(i)}$
End for
Rescale the weights by $\hat{N}_k$ to get $\{\mathbf{x}_k^{(i)}, \frac{\hat{N}_k}{N}\}$

3. SIMULATION EXPERIMENTS

In order to evaluate the performance of the proposed robust particle PHD filter for multiple target tracking, we employed the dataset of the EC Funded CAVIAR project [13], Video EnterExitCrossingPaths1cor and EnterExitCrossingPaths1front which have 383 frames. There are four human beings appearing, occluding each other, and disappearing in the shopping mall environment. We found empirically that the freedom rate for Student's-t distribution $v = 3$ is the appropriate value for the proposed PHD filter and the number of particles is set to be 1000. In order to evaluate the performance of our proposed PHD filter, we employ the optimal subpattern assignment (OSPA) [14] and the mean error (ME) for each target metric, the comparison result is shown as Fig. 3 and Table 1

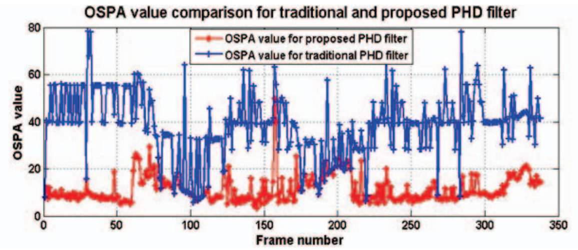

Fig. 3: Comparison of target number from the proposed method and the ground truth, where the blue line denotes the results from the proposed PHD filter and the red line denotes the target number from ground truth information.

Table 1: Comparisons of different tracking results

	Scenario 1		Scenario 2	
	PHD	t-SVM-PHD	PHD	t-SVM-PHD
OSPA(pixel)	38.25	11.45	34.54	22.26
ME (pixel)	11.27	4.44	19.87	11.85

where the PHD in the table denotes the human tracking work that only used background subtraction as the feature as proposed in [10] and the t-SVM-PHD denotes our proposed robust PHD filter. From the comparisons, it can be observed that our proposed PHD filter performs much better than the baseline PHD filter, where in both scenarios the OSPA value and the mean error are reduced. The error by the PHD filter is mostly from the false alarm and missed detection caused by the noise from the background subtraction step, which can be mitigated by the OCSVM classifier in our proposed PHD filter; so from the comparison, we can deduce that our proposed PHD filter is more robust and can improve the accuracy of multiple human tracking work.

4. CONCLUSION

We have presented a new robust particle PHD filter based on the Student's-t distribution and an OCSVM. In order to overcome the problem of false alarms and miss detection caused by the noise from background subtraction, we extracted colour and oriented gradient histogram features of human targets and used them with an OCSVM classifier to calculate the weights for particles in the prediction step of the PHD filter. From the results we found the proposed method achieved is more accurate than the baseline method in estimating the target number. In future work, to further improve the accuracy of the tracking result, we will employ a variational Bayesian step to predict the parameters in the Student's-t distribution [15]; and to make the classifier adaptable to the change of the environment and targets. An online OCSVM [16] can also be used to improve the robustness of the multiple human tracking system.

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

[7] A. Doucet, B.N. 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.

[8] 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.

[9] 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.

[10] Y. Wang, J. Wu, A. A. Kassim, and W. Huang, "Tracking a variable number of human groups in video using probability hypothesis density," *in Proc. IEEE 18th of the International Conference on Pattern Recognition.*, 2006.

[11] M. Yu, A. Rhuma, S. M. Naqvi, and J. Chambers, "Fall detection for the elderly in a smart room by using an enhanced one class support vector machine," *in Proc. IEEE 17th of the International Conference on Digital Signal Processing*, pp. 1–6, 2011.

[12] B. Scholkopf, J. Platt, J. Taylor, A. Smola, and R. Williamsom, "Estimating the support of a high dimensional distribution," *Neural Computation*, vol. 13, pp. 1443–1471, 2001.

[13] R. Fisher, "Caviar case scenarios [online]. Available:http://groups.inf.ed.ac.uk/vision/caviar/caviardata1/," .

[14] 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.

[15] D. Xu, C. Shen, and F. Shen, "A robust particle filtering algorithm with non-Gaussian measurement noise using Student's-t distribution," *IEEE Signal Processing Letters*, vol. 21, pp. 30–34, 2014.

[16] M. Yu, Y. Yu, A. Rhuma, S. M. Naqvi, L. Wang, and J. A. Chambers, "An online one class support vector machine-based peron-specific fall detection system for monitoring an elderly individual in a room environment," *IEEE Journal of Biomedical and Health Informatic*, vol. 17, pp. 1002–1014, 2013.