Simplified Multitarget Tracking Using The PHD Filter for Microscopic Video Data

Trevor M. Wood, *Student Member, IEEE*, Christian A. Yates, David A. Wilkinson, Gabriel Rosser

Abstract

The probability hypothesis density (PHD) filter from the theory of random finite sets is a well-known method for multitarget tracking. We present the Gaussian mixture (GM) and improved sequential Monte Carlo (SMC) implementations of the PHD filter for visual tracking. These implementations are shown to provide advantages over previous PHD filter implementations on visual data by removing complications such as clustering and data association and also having beneficial computational characteristics.

The GM-PHD filter is deployed on microscopic visual data to extract trajectories of free-swimming bacteria in order to analyse their motion. Using this method, a significantly larger number of tracks are obtained than was previously possible. This permits calculation of reliable distributions for parameters of bacterial motion. The PHD filter output was tested by checking agreement with a careful manual analysis. A comparison between the PHD filter and alternative tracking methods was carried out using simulated data, demonstrating superior performance by the PHD filter in a range of realistic scenarios.

Index Terms

Multitarget tracking, random finite sets, PHD filter, sequential Monte Carlo, bacterial motion

I. Introduction

Multitarget tracking is the problem of jointly determining the number of targets present and their locations in noisy data. In video data, multitarget tracking comprises two stages: i) detection of objects of interest within each frame, ii) maintaining continuous tracks for individual objects over a number of frames. Henceforth 'objects of interest' will be referred to as 'targets' in keeping with the standard terminology.

A number of problems may occur at the detection step, especially in noisy video data or situations where targets may be difficult to detect. Missed detections may occur if a target fails to generate a detection or is occluded by another target. False detections, or 'false alarms', may also occur and the true detections may be corrupted by

T.M.W.'s work is supported by an EPSRC CASE award with Thales Underwater Systems and Thales Aerospace. C.A.Y and G.R.'s work is supported by an EPSRC award via the Systems Biology and Life Sciences Interface Doctoral Training Centres. D.W's work is supported by BBSRC.

T.M.W. is with the Oxford Centre for Industrial Applied Mathematics, University of Oxford, OX1 3LB, U.K., e-mail: woodtm@maths.ox.ac.uk.; C.A.Y. and G.R. are with the Centre for Mathematical Biology, University of Oxford, OX1 3LB, U.K., email: {yatesc@maths.ox.ac.uk, rosser@maths.ox.ac.uk}; D.W. is with the Department of Biochemistry, University of Oxford, U.K., email: david.wilkinson@bioch.ox.ac.uk

measurement noise. In addition to these difficulties, the multitarget tracker must also handle the appearance and disappearance of new targets as well as uncertain target dynamics.

Finite Set Statistics (FISST), developed by Mahler [18], provides a theoretical framework for solution of the multitarget tracking problem. In this framework, the set of target states are treated as a random finite set; a set with a random number of elements which are themselves random. Within this framework, the multitarget Bayes filter has been derived [18]. This is the optimal Bayesian solution to the multitarget tracking problem but is computationally intractable. The PHD filter, also developed by Mahler [17], propagates only the first moment of the posterior from the multitarget Bayes filter, known as the probability hypothesis density (PHD). The PHD has an interpretation as a *density of expected targets* so that integration of the PHD over a region S gives the expected number of targets in S. The PHD can be integrated over the whole domain to give the total expected number of targets present, and individual peaks extracted to give likely target locations. The PHD filter gives a theoretically sound and computationally tractable solution to the multitarget tracking problem.

PHD filter methods for multitarget tracking in visual data have previously been presented by Maggio et al. [16] and Wang. et al [35], both of which show promising results for visual tracking on real data. However, in both [16] and [35], the implementation used is of the sequential Monte Carlo PHD (SMC-PHD) filter first presented in [32]. Such an implementation requires some additional heuristics for partitioning particles for state extraction, data association/track continuity and target initialisation which may make application difficult or impractical in general. This will be discussed further in Section II-B. As an alternative, we suggest using either the Gaussian-Mixture PHD (GM-PHD) filter [33] which removes the need for additional heuristics and is also able to significantly reduce computational cost. The GM-PHD filter is only valid in the case where dynamical and measurement models are both linear and Gaussian, or may be approximated as such, therefore SMC-PHD may still be required in some visual tracking scenarios. For these cases, the improved SMC-PHD filter originally presented in [25] and [36] may be used. In Section II the GM-PHD filter and improved SMC-PHD filters are presented to allow a simplified PHD filter implementation for visual data obviating the need for additional heuristics for particle partitioning, track continuity and target initialisation. The PHD filter implementations presented here might be a practical solution for tracking problems such as analysing video surveillance and infra-red visual data.

In Section III the simplified visual PHD filter is deployed on a difficult multitarget tracking problem relevant to current research in biochemistry [14] [15] [22]: tracking multiple bacteria in noisy video data taken through a microscope in order to gain information about bacterial swimming patterns. Use of this method permits bacterial paths to be extracted at a significantly faster rate than was previously possible. Tracker results are validated by checking agreement with a careful manual analysis and a comparison is performed with alternative methods for tracking bacteria using simulations. A summary is given in Section IV.

II. PHD FILTER

A. Background

Suppose that at time step k there are n_k targets present with states $\mathbf{x}_{k,1}, ..., \mathbf{x}_{k,n_k}$ and m_k measurements with states $\mathbf{z}_{k,1}, ..., \mathbf{z}_{k,m_k}$. The multitarget states and multitarget measurements are the random finite sets:

$$\mathbf{X}_k = \{\mathbf{x}_{k,1}, ..., \mathbf{x}_{k,n_k}\}, \tag{1}$$

$$\mathbf{Z}_k = \{\mathbf{z}_{k,1}, ..., \mathbf{z}_{k,m_k}\}, \tag{2}$$

respectively. $\mathbf{X}_k$ and $\mathbf{Z}_k$ are random finite sets because the number of members of each set is random and so are the set elements themselves.

At each time step, targets may appear/disappear and surviving targets may change state according to a target motion model. Measurements due to the targets (possibly corrupted by noise) are received. False measurements may also be received and detections of targets may be missed due to difficulty of detection or occlusions. Taking all of these effects into account, the aim is to calculate the multitarget posterior $f_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k})$ recursively for each time step.

Assuming that $f_{k-1|k-1}(\mathbf{X}_{k-1}|\mathbf{Z}_{1:k-1})$ is known, the posterior at time k may be calculated, as in [18], by:

$$f_{k|k-1}(\mathbf{X}_k|\mathbf{Z}_{1:k-1}) = \int \Pi_{k|k-1}(\mathbf{X}_k|\mathbf{W}) f_{k-1|k-1}(\mathbf{W}|\mathbf{Z}_{1:k-1}) \delta\mathbf{W}, \tag{3}$$

$$f_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k}) = \frac{\nu_k(\mathbf{Z}_k|\mathbf{X}_k) f_{k|k-1}(\mathbf{X}_k|\mathbf{Z}_{1:k-1})}{\int \nu_k(\mathbf{Z}_k|\mathbf{W}) f_{k|k-1}(\mathbf{W}|\mathbf{Z}_{1:k-1}) \delta\mathbf{W}}, \tag{4}$$

where $\Pi_{k|k-1}(\mathbf{X}_k|\mathbf{X}_{k-1})$ is the multitarget Markov transition density and $\nu_k(\mathbf{Z}_k|\mathbf{X}_k)$ is the multitarget likelihood.

Propagating the full multitarget Bayes filter as outlined above is impractical due to the high computational cost. In order to overcome this, Mahler [17] proposed to propagate the first moment of $f_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k})$, known as the PHD, denoted $D_{k|k}(\mathbf{x}|\mathbf{Z}_{1:k})$ and commonly abbreviated to $D_{k|k}(\mathbf{x})$. The PHD has the useful property that $\int_S D_{k|k}(\mathbf{x})\mathrm{d}\mathbf{x}$ is equal to the expected number of targets in S. Hence, the PHD may be interpreted as a density of expected targets.

It is shown in [17] that by propagating only the PHD, equations (3)-(4) become:

$$D_{k|k-1}(\mathbf{x}) = \gamma_k(\mathbf{x}) + \int (p_S(\mathbf{w}) f_{k|k-1}(\mathbf{x}|\mathbf{w}) + b_{k|k-1}(\mathbf{x}|\mathbf{w})) D_{k-1|k-1}(\mathbf{w})\mathrm{d}\mathbf{w}, \tag{5}$$

$$D_{k|k}(\mathbf{x}) = F_k(Z_k|\mathbf{x}) D_{k|k-1}(\mathbf{x}), \tag{6}$$

where $f_{k|k-1}$ is the single target Markov transition density, $p_S(\mathbf{x})$ is the probability that a target with state $\mathbf{x}$ at time step $k-1$ will survive to time step k, $b_{k|k-1}$ describes the spawning of new targets from existing ones and γ_k describes the state-independent appearance of new targets. F_k is the multitarget likelihood as given in [17].

It should be noted that in approximating the multitarget Bayes filter by the propagation of its first moment, the PHD, it is necessary to assume that the multitarget density is approximately Poisson. In particular, this means that the variance of target number increases with the number of targets. The validity of this approximation will depend upon characteristics of the data such as the false alarm rate and the probability of missed detections.

B. SMC-PHD for Visual Tracking

The approach of using SMC-PHD for visual tracking was presented in [16] and [35]. These follow the approach of Vo et al. [32] using a weighted particle approximation to equations (5)-(6).

There are some additional complexities to performing visual tracking in this way. These are:

i) Clustering for state extraction - In order to extract individual targets, particles must be partitioned using a clustering method. [16] and [35] choose the expectation maximisation and k-means algorithms respectively. Both of these algorithms take the number of partitions as an input and it is well-known that the state extraction process can become unreliable when the number of partitions is estimated incorrectly [32]. In Pasha et al. [21], for example, "unreliability of clustering techniques for extracting state estimates" is given as the reason for eschewing the SMC implementation of the PHD filter.

ii) Data Association - Associating tracks between frames in order to give continuous tracks may be important, and is particularly so with regards to the tracking of bacterial motion, as presented in Section III. In [16], data association is performed externally to the PHD filter using a graph-based method. The graph-based method used does not permit clutter, and hence may give unreliable output in the cases where high clutter causes false detections to be output by the PHD filter.

iii) Birth Process - In the case where there is no prior localisation information about γ_k in equation (5) the birth density has to cover the whole of the state space, requiring a large number of particles [25]. [16] avoids this problem using the heuristic of only sampling particles in the vicinity of the measurements received.

The complexities described above are all taken into account in [16] and [35]. However, the methods used may be unreliable in certain circumstances, such as incorrect estimation of target number for clustering or declaration of false tracks for data association. This can be avoided by handling the complexities within the PHD filter as described below.

C. Simplified PHD Filter for Visual Tracking

1) GM-PHD Filter: In the case where the target dynamics and measurement model are both linear and Gaussian and the birth process may be expressed as a Gaussian sum, there is a closed form expression for the PHD recursion derived in [33].

In this case, the forward and measurement models may be expressed as:

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k + \mathbf{v}_k, \tag{7}$$

$$\mathbf{z}_{k+1} = H_{k+1} \mathbf{x}_{k+1} + \mathbf{w}_{k+1}, \tag{8}$$

where F_k and H_{k+1} are the matrices representing the noise-free motion and measurement model respectively, $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, Q_k)$ and $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, R_k)$ with Q_k and R_k the covariance matrices of process noise and measurement noise respectively which are assumed known.

The birth density γ_k is given by:

$$\gamma_k(\mathbf{x}) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(\mathbf{x}; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}), \tag{9}$$

where $J_{\gamma,k}$, $w_{\gamma,k}^{(i)}$, $m_{\gamma,k}^{(i)}$ and $P_{\gamma,k}^{(i)}$ are known model parameters. Note that any birth density may be approximated to any accuracy using such a Gaussian mixture [30]. It will be assumed for simplicity of presentation that there is no target spawning process but it should be noted that inclusion of a spawning process is a straightforward and well-known extension [17].

Under these assumptions, and with:

$$D_{k-1|k-1}(\mathbf{x}) = \sum_{i=1}^{J_{k-1}} w_{k-1}^{(i)} \mathcal{N}(\mathbf{x}; m_{k-1|k-1}^{(i)}, P_{k-1|k-1}^{(i)}), \tag{10}$$

the PHD prediction equation (5) becomes:

$$\begin{aligned}
D_{k|k-1}(\mathbf{x}) &= D_{S,k|k-1}(\mathbf{x}) + \gamma_k(\mathbf{x}), & (11)\\
D_{S,k|k-1}(\mathbf{x}) &= p_S \sum_{i=1}^{J_{k-1}} w_{k-1}^{(i)} \mathcal{N}(\mathbf{x}; m_{S,k|k-1}^{(i)}, P_{S,k|k-1}^{(i)}), & (12)\\
m_{S,k|k-1}^{(i)} &= F_{k-1} m_{k-1|k-1}^{(i)}, & (13)\\
P_{S,k|k-1}^{(i)} &= Q_{k-1} + F_{k-1} P_{k-1|k-1}^{(i)} F_{k-1}^T. & (14)
\end{aligned}$$

Similarly, with the predicted PHD denoted:

$$D_{k|k-1}(\mathbf{x}) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(\mathbf{x}; m_{k|k}^{(i)}, P_{k|k}^{(i)}), \tag{15}$$

the PHD update equation (6) becomes:

$$\begin{aligned}
D_{k|k}(\mathbf{x}) &= (1 - P_D) D_{k|k-1}(\mathbf{x}) + \sum_{\mathbf{z} \in Z_k} D_{D,k|k}(\mathbf{x}; \mathbf{z}), & (16)\\
D_{D,k|k}(\mathbf{x}; \mathbf{z}) &= \sum_{i=1}^{J_{k|k-1}} w_k^{(i)}(\mathbf{z}) \mathcal{N}(\mathbf{x}; m_{k|k}^{(i)}(\mathbf{z}), P_{k|k}^{(i)}), & (17)\\
w_k^{(i)}(\mathbf{z}) &= \frac{P_D w_{k|k-1}^{(i)} q_k^{(i)}(\mathbf{z})}{\kappa_k(\mathbf{z}) + P_D \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} q_k^{(j)}(\mathbf{z})}, & (18)\\
q_k^{(i)}(\mathbf{z}) &= \mathcal{N}(\mathbf{z}; H_k m_{k|k-1}^{(i)}, R_k + H_k P_{k|k-1}^{(i)} H_k^T), & (19)\\
m_{k|k}^{(i)}(\mathbf{z}) &= m_{k|k-1}^{(i)}(\mathbf{z}) + K_k^{(i)}(\mathbf{z} - H_k m_{k|k-1}^{(i)}), & (20)\\
P_{k|k}^{(i)} &= [\mathbb{I} - K_k^{(i)} H_k] P_{k|k-1}^{(i)}, & (21)\\
K_k^{(i)} &= P_{k|k-1}^{(i)} H_k^T (H_k P_{k|k-1}^{(i)} H_k^T + R_k)^{-1}, & (22)
\end{aligned}$$

where $\kappa_k(\mathbf{z})$ is the clutter density and P_D is the probability of detection, both assumed known.

The linear Gaussian assumptions in equations (7)-(8) are appropriate for a range of visual tracking scenarios, including the one presented in [35] where SMC-PHD is used. For the case where all noise is Gaussian, but the target

models are nonlinear, such as the visual tracking scenario presented in [16], Vo and Ma [33] present an adaptation to the GM-PHD filter. Their approach follows the method of single target tracking where extended Kalman filter (EKF) and unscented Kalman filter (UKF) methods are applied to tracking scenarios with nonlinear target models. EKF relies on a local linearisation of the target models, while UKF uses the unscented transform. Details of the single target EKF and UKF filters are given in [1] and [13] respectively. Full details of the EKF-GM-PHD and UKF-GM-PHD are given in [33]. This allows implementation of GM-PHD even when the dynamical and measurement models are nonlinear.

EKF-GM-PHD and UKF-GM-PHD may only be expected to work in cases with mildly nonlinear target models. In cases with severely nonlinear target models, or non-Gaussian noise, it may still be necessary to use the SMC-PHD filter. The textbook on single target tracking with nonlinear models by Ristic et al. [24] provides some examples where EKF and UKF will not be adequate.

Returning to the issues raised in Section II-B, the GM-PHD and its nonlinear approximations do not require troublesome clustering methods for state extraction. Individual Gaussian components with weights above a threshold provide the estimated target states. Furthermore, target birth is entirely captured by (9). Data association/track continuity can be handled within the GM-PHD filter. The method essentially involves assigning track labels to Gaussian components. These are maintained as they are propagated forward under the target motion model, as described in [20].

2) Improved SMC-PHD filter: The improved SMC-PHD filter was proposed recently in [25] in order to alleviate the issues with clustering and the birth process within the SMC-PHD filter. Similarly to the SMC-PHD filters presented for visual tracking in [16] and [35] it approximates the PHD by a set of weighted samples with:

$$D_{k|k}(\mathbf{x}|\mathbf{Z}_{1:k}) \approx \Sigma_{n=1}^{N_k} w_{k|k}^n \delta_{\mathbf{x}_{k|k}^n}(\mathbf{x}), \tag{23}$$

where $\{(w_k^n, \mathbf{x}_k^n)\}_{n=1}^{N_k}$ is the set of weights and samples at time k.

The standard prediction and update equations, from [32], for the random samples at time k given $\{(w_{k-1}^n, \mathbf{x}_{k-1}^n)\}_{n=1}^{N_{k-1}}$ are:

$$\begin{aligned}
\mathbf{x}_k^n &\sim \begin{cases} f_{k|k-1}(\cdot|\mathbf{x}_{k-1}^n, \mathbf{Z}_k), & n = 1, ..., N, \\ b(\cdot, Z_k), & n = N+1, ..., N+P, \end{cases} \\
w_{k|k-1}^n &= \begin{cases} w_{k-1}^n, & n = 1, ..., N, \\ \frac{1}{P}, & n = N+1, ..., N+P, \end{cases} \qquad (24) \\
w_k^n &= (1 - P_D(\mathbf{x}_k^n)) w_{k|k-1}^n + \\
&\quad \sum_{\mathbf{z} \in Z_k} \frac{\psi_{k,z}(\mathbf{x}_k^n) w_{k|k-1}^n}{\kappa_k(\mathbf{z}) + \sum_n \psi_{k,z}(\mathbf{x}_k^n) w_{k|k-1}^n}, \qquad (25)
\end{aligned}$$

where P is the number of particles assigned to newborn targets, P_D is the probability of detection, $\psi_{k,z}(\mathbf{x}) = P_D(\mathbf{x}) g(\mathbf{z}|\mathbf{x})$ where $g(\mathbf{z}|\mathbf{x})$ is the single target likelihood for measurement $\mathbf{z}$. $b(\cdot, Z_k)$ is the proposal density for newborn particles chosen in order to place such particles close to measurements, in a similar manner to the methods

used in [16] and [35]. Additionally, a resampling step is generally performed at the end of every time step as in the single-target particle filter [24].

In order to perform state extraction without explicitly clustering particles, the form of (25) is exploited. It can be seen that for each particle weight computed, there is a component due to each measurement, and an additional component relating to missed detections. The component for each measurement is considered in turn and separate state estimates are computed for each measurement as follows:

$$w_k^{n,j} = \frac{\psi_{k,z_{k,j}}(\mathbf{x}_k^n) w_{k|k-1}^n}{\kappa_k(\mathbf{z}_{k,j}) + \sum_n \psi_{k,z_{k,j}}(\mathbf{x}_k^n) w_{k|k-1}^n}, \tag{26}$$

$$\hat{\mathbf{x}}_{k,j} = \sum_{n=1}^{N+P} w_k^{n,j} \mathbf{x}_k^n, \tag{27}$$

$$P_{k,j} = \sum_{n=1}^{N+P} w_k^{n,j} (\mathbf{x}_k^n - \hat{\mathbf{x}}_{k,j})(\mathbf{x}_k^n - \hat{\mathbf{x}}_{k,j})^T, \tag{28}$$

$$W_{k,j} = \sum_{n=1}^{N+P} w_k^{n,j}, \tag{29}$$

for $j = 1, ..., m_k$. The weights due to the possibility of missed detection can be computed similarly by:

$$w_k^{n,0} = (1 - P_D(\mathbf{x}_k^n)) w_{k|k-1}^n. \tag{30}$$

A method for track continuity within the improved SMC-PHD filter was presented in [36]. This uses a similar method for track continuity to the one employed for GM-PHD by assigning track labels to individual particles, which are retained during propagation due to the target motion model. Thus, the improved SMC-PHD also avoids the issues raised in Section II-B.

III. Tracking Multiple Bacteria in Microscopic Video Footage

A. Context

Bacteria such as *Rhodobacter sphaeroides* are capable of constructing and living within surface associated multicellular communities known as biofilms. The formation of biofilms has huge implications for medical and industrial environments. Biofilms have been found to cause chronic infections, to grow on medical implants and to form plaque on teeth. In industry they have been found to block pipes, corrode metal, and to contaminate clean water distribution systems. For these reasons, bacterial behaviour near surfaces is a topic of current research interest [14] [29] [34].

Bacteria were first successfully tracked in microscopic visual data in 1971 by Berg [4] using a specially designed tracking microscope which adjusted its position and focus in real time to follow bacterial movement. Berg poetically described the results: "*The scene through the binocular is extraordinary. The bacterium being tracked seems to be stuck to the centre of the field, turning this way and that trying to free itself, while the other bacteria drift in and out of focus, then to and fro, in apparent synchrony*". In the same passage, Berg also described a telling incident: "*the organism was seen to collide with one which was not swimming; the microscope locked onto this bacterium and the*

other swam out of focus". This highlights the difficulty of multitarget tracking as opposed to single target tracking and is the reason why multitarget tracking methods are necessary for the problem. In Berg and Brown's widely-cited paper [5] presenting results from the tracking process, 59 trajectories are extracted from over 25 minutes of tracking time.

More recently, some analysis of bacterial behaviour has been carried out for 'tethered cells' [7], where cells are stuck to a microscopic slide cover by their flagella. This provides some information about characteristics of the cell motor but is an unnatural state in which to observe bacteria, which may influence results. Furthermore, parameters such as speed, distance and turn rate cannot be measured using this method.

Taboada et al. [31] appreciate that "*quantification of the different features of bacterial swimming requires the analysis of a large number of free-swimming cells*". In [31], Taboada et al. use a basic digital tracking method relying on cell overlaps between frames for track continuity. This method will encounter difficulties for crossing cell paths, just as Berg did, and as cells often move more than 200 times their own body length per second [2] a high frame rate is required. Consequently, Taboada et al. [31] are limited to obtaining 50 tracks.

It should be noted that theapplication of the PHD filter to microscopic bacterial data has been attempted before by Juang et al. [12]. However, the primary purpose for their implementation was for monitoring cell lineage. Problems were found in the course of the implementation with regard to consistently tracking moving cells. Juang et al. state, for example, that "cells that moved too quickly became untracked". The implementation that will be described below focuses on maintaining continuous tracks for cells in motion.

B. The Tracking Scenario

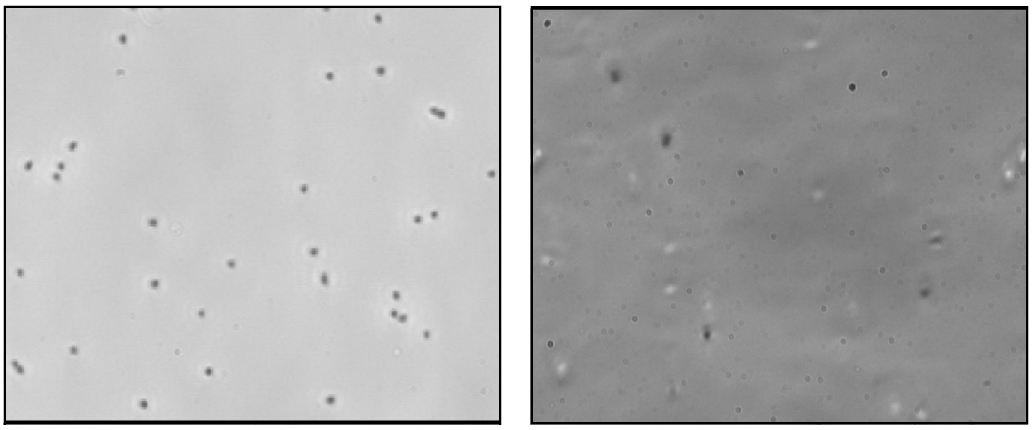
In order to record bacterial motion, cells were placed within a flat glass capillary which was sealed at both ends using silicone grease. Swimming was visualised using phase contrast microscopy at 40x magnification and videos were captured at 50 frames per second (fps). The species *Rhodobacter sphaeroides* was chosen as it is of research interest [22]. Videos were taken at, and away from, the surface of the capillary as differences in swimming behaviour of the bacteria between these regions are of particular interest with regard to biofilm formation.

Figure 1 shows frames from two different kinds of microscopic video. Figure 1(a) is taken at the surface of the capillary. The cells are quite clear, so that occlusion is the issue most likely to cause a missed detection of a cell. Figure 1(b) is taken away from the surface and deeper within the fluid. In this case the cells are more faint and difficult to detect, some are darker than the background while others are lighter due to differences in depth.

Linear, Gaussian models are sufficient to capture the target dynamics and measurement models, hence equations (7)-(8) are used with:

$$\mathbf{x}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k], \tag{31}$$

$$F_k = \begin{bmatrix} 1 & dt & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & dt \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad , \quad H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{32}$$



(a) (b)

Fig. 1. a) Frame from microscopic video taken at the surface. b) Frame from microscopic video taken away from the surface.

for all k where (x_k,y_k) is the (x, y) position of the object and dt is the sampling time. The covariances of the process and measurement noise are given by:

$$Q_k = \begin{bmatrix} \frac{dt^3}{3} & \frac{dt^2}{2} & 0 & 0 \\ \frac{dt^2}{2} & dt & 0 & 0 \\ 0 & 0 & \frac{dt^3}{3} & \frac{dt^2}{2} \\ 0 & 0 & \frac{dt^2}{2} & dt \end{bmatrix} q \ , \ R_k = \sigma_w^2 \mathbb{I}_2, \tag{33}$$

where the form of Q_k is as derived in [3] with q being the process noise scaling parameter and σ_w is the standard deviation of measurement noise.

The birth density, γ_k, will depend on the scenario but should be chosen to localise around the edges of the image in the case, as in fig. 1(a), where the cells are always clearly visible and must emerge from the sides. In the case, as in fig. 1(b), where cells may move into the field of view by changing depth the birth density must cover the whole image.

Occlusions may be taken into account by varying the value of the parameter P_D for probability of detection so that the probability of detection of a cell falls when it is in the vicinity of another cell. This helps maintain tracks through cell crossings.

Finally, it is well-known that some cells, including *Rhodobacter sphaeroides*, have a run-stop behaviour. This might be modelled as a jump Markov linear system, meaning that there are two different motion modes and switching between these two is a Markov process. Handling jump Markov linear systems within the PHD filter has been presented in [21], [23] and [37] for GM-PHD, SMC-PHD and the general case respectively. This method may be used to improve handling of rapid changes in behaviour while tracking but is omitted here for simplicity as the bacterial strain used did not display stopping behaviour.

C. Testing on Real Data

Testing the output of the tracker is a difficult task as no 'ground truth' is available. A comparison of the tracker output with a manual analysis was carried out on a selection of tracks. This process is difficult and time consuming. Taboada et al. [31] go as far as to suggest that "simple eye analysis would be extremely difficult or impossible". However, there is some merit in confirming correspondence between the tracker output and the tracks as identified by the human eye.

Manual Analysis: 100 tracks were randomly selected, 50 of which were from videos captured at the surface and 50 of which were captured away from the surface. A careful manual analysis was performed comparing the track output with the video in slow motion. Occlusions and faint tracks made identifying the track by eye particularly difficult and several viewings were sometimes necessary. 96 out of the 100 tracks chosen were deemed to match the video. A random sample of 20 of these are available for the reader's scrutiny, 10 from the surface and 10 away from the surface at the URLs:
`http://people.maths.ox.ac.uk/yatesc/ieee_movies/tracksurf1.avi` to
`http://people.maths.ox.ac.uk/yatesc/ieee_movies/tracksurf10.avi` and
`http://people.maths.ox.ac.uk/yatesc/ieee_movies/track1.avi` to
`http://people.maths.ox.ac.uk/yatesc/ieee_movies/track10.avi` for tracks at, and away from the surface respectively.

The reader may be more interested in the 4 incidents where the tracker output did not match up to the by-hand analysis. There were two at the surface and two away from the surface and the videos are available at the URLs:
`http://people.maths.ox.ac.uk/yatesc/ieee_movies/tracksurffail1.avi`,
`http://people.maths.ox.ac.uk/yatesc/ieee_movies/tracksurffail2.avi` and
`http://people.maths.ox.ac.uk/yatesc/ieee_movies/trackfail1.avi`,
`http://people.maths.ox.ac.uk/yatesc/ieee_movies/trackfail2.avi` respectively.

All four incidents contain a rare occurrence. One of the track failures at the surface includes a cell occluded three times in quick succession and the other an incident where two cells with similar trajectories are both simultaneously occluded by a third cell as they cross (see figs. 2(d)-2(f)). Away from the surface, one of the track failures is a case where a cell disappears off of the edge of the image at the same time as a new cell appears at a similar point. The other is of a large cell, which appears to be about to divide, spinning rapidly. A track is declared but its position is erratic.

In a tracking scenario with a large number of varied targets, some of which are faint and which frequently occlude each other, such incidents might be expected to give a low level of unreliable tracks, but the overall results give a confidence in the match between manual analysis and the tracker. It is also clear from some of the videos that the tracker allows detection of some cells which might have been too faint to detect by eye, and there were also incidents with occlusions where the tracker at first appeared to be incorrect but upon a more careful viewing was shown to be correct. In these cases the performance of the tracker may be better than that of manual analysis. For

self-containment purposes, we have included stills from two of the videos described above to give the reader an insight into what these events look like. The stills may be found in fig. 2.

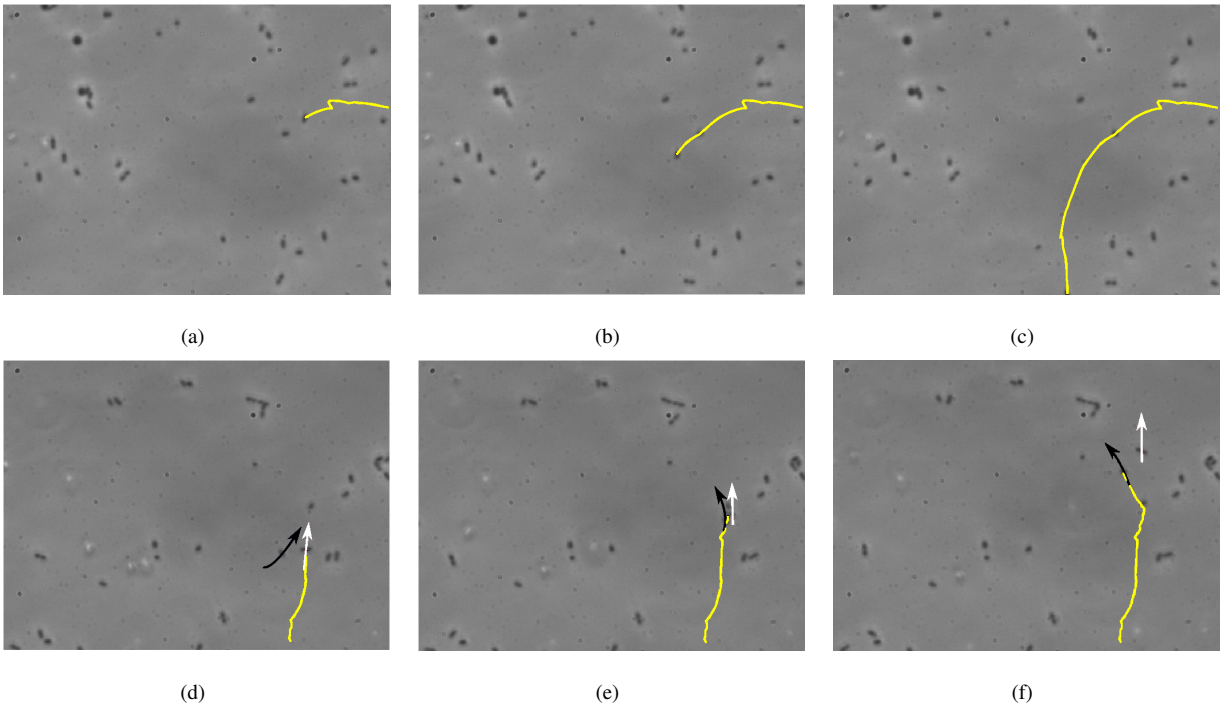


Fig. 2. Illustrative video frames to complement the above linked videos. Stills from videos illustrating a well-tracked cell (a)-(c) and a track which was lost (d)-(f) due to an occlusion with a moving cell and a stationary cell at the same time (an unusual event). (a) The track (yellow) is initiated almost from the moment the cell arrives on the screen. (b) The tracker manages to deal with an occlusion event as the tracked cell passes over a stationary cell. (c) The track continues until the cell leaves the screen. (d) A new cell is tracked as it appears on the screen and its motion path indicated with a white arrow through its centre. A cell which will occlude the tracked cell also has its motion path marked in a similar way with a black arrow. (e) The three cells occlude each other. (f) The path of the track is swapped to the cell marked with the black arrow as the cells separate.

D. Performance Comparison Using Simulated Data

It is desired to test the PHD filter against other recent methods from microscopic visual tracking. A recent review on the topic which gives a "good picture of the state-of-the-art" [26] for live-cell tracking is given by Jaqaman et al. [11] which broadly characterises tracking methods in terms of defining a cost between cell detections in subsequent frames and then performing a deterministic total cost minimisation. Examples of tracking methods based on cost minimisation may be found in the papers by Ouellette et al. [19] and Xie et al. [38] [39] for tracking in particle image velocimetry (PIV) and tracking *Escherichia coli* respectively. There is a fundamental difference between this class of methods and the PHD filter because the cost minimisation approach is deterministic as opposed to the statistical approach of the PHD filter as described above.

In order to obtain a reliable quantitative comparison between the PHD filter and alternative methods, it will be necessary to use simulated data. This is due to the absence of a known 'ground truth' in the real data.

Simulating images that accurately replicate experimentally-acquired data is extremely challenging. Different optical contrast-enhancement methods (ie. Differential Interference Contrast (DIC), Phase Contrast (PC) or Dark-Field (DF) microscopy), as well as chemical staining techniques can be used to visualise swimming bacteria, and therefore the appearance of each image will be dependent on many experimental factors. Furthermore, different bacterial species will have different shapes, sizes and refractive properties and the buffers that they require for swimming will potentially have different refractive indices. Therefore biological factors will also affect the appearance of the image data both in terms of image contrast and image noise. For these reasons, instead of simulating full images, the set of detections for each frame was simulated taking into account phenomena such as false alarms, occlusions and measurement noise.

The dynamics of the bacteria swimming at a surface were simulated using the motion model in equations (31)-(33). Code written for MATLAB is available upon request. The system process noise parameter, q, was set to 5×10^{-3}, as the tracks simulated using this level of process noise were observed to bear a very close resemblance to the manually verified experimental tracks for bacteria swimming at a surface. The simulated tracks were confined to a 'field of view' of size 640 px by 480 px, which is the field of view of the microscope camera used in this study. Tracks leaving this region were no longer considered as they are deemed to have swum out of view. The total simulation time was 90s and a time step of 1/50s was used as this is the between-frame interval used by the microscope camera. Simulated bacteria were spawned at random positions on the boundary of the field of view with a birth rate of 32 s^{-1}, equivalent to an average of approximately 20 bacteria in the field of view for any particular frame, which is comparable to the number in the experimental videos. Newly spawned bacteria were allocated an initial velocity with direction chosen from a uniform distribution on the half-circle facing into the field of view and magnitude (speed) drawn from a Gaussian distribution. The Gaussian distribution governing the initial speed is truncated so that only positive values are permitted. The mean and standard deviation are chosen to approximate the range of speeds observed in experimental data, and are given by 325 pxs^{-1} and 150 pxs^{-1}, respectively. In order to account for the fact that the number and location of bacteria are at steady state at the start of an experiment, an initial 'burn-in' time of several seconds was discarded from the start of each simulation. Tracks are therefore initially found distributed throughout the whole simulated field of view and subsequent tracks are spawned at the edges, as observed in the microscopy videos. False alarm measurements were uniformly distributed over the simulated field of view with the number of false alarms being approximated by a Poisson distribution with mean λ. Detections were also corrupted with measurement noise which was simulated by white, zero mean Gaussian noise with standard deviation $\sigma_w = 1$ px in each coordinate direction.

We note that the underlying motion model in equations (31)-(33) used in these simulations is equivalent to a random walk in velocity space. One drawback of this method is that the variance of the speed of the simulated bacteria is unbounded with time. Long tracks can exhibit physically unrealistic speeds as there is no explicitly defined maximum speed. Alternative process noise models exist, such as the Ornstein-Uhlenbeck process, which has been shown to be compatible with the motion exhibited by certain types of amoeba [6]. The use of the model is nevertheless deemed appropriate in the present work as, in practice, the field of view limits the duration of any

single track. This fact, coupled with the large value of the mean initial speed of tracks compared to the value of the variance of the process noise, means that no simulated tracks were observed exhibiting unrealistic speeds.

The performance of the GM-PHD filter was compared with two methods based on cost minimisation. The first is the nearest neighbour (NN) method as described in [19] which is used as a benchmark. The cost function in this case is the distance between the estimated position of the tracked cell and the position of the newly detected cell. This tracker is equivalent to the implementation in, for example, [9]. The second alternative method is the tracker from [39] denoted XIE hereafter. The cost function for this tracker is the sum of the distance between the predicted cell position and the new detection, and the difference between the predicted direction of motion of the cell and the observed direction of motion given the new detection. Information relating to image histograms is not included as it is not available in this simulated data. In order to permit equivalent computational complexity to the GM-PHD filter, the sliding window size for XIE was chosen as two. It should be noted that in practice simple heuristics are often used to improve filter performance but the results presented below consider only the performance of the well-known "vanilla" implementations of each filter.

1) Performance Comparison Using OSPA Metric: Evaluation in multitarget tracking may be performed using the optimal sub-pattern assignment (OSPA) metric [28], a consistent metric for the distance between two sets. In multitarget tracking performance evaluation, OSPA is used to measure the distance between the true set of tracks and the estimate returned by a given tracker. The OSPA metric takes two factors into account; the difference in localisation and the difference in cardinality (the number of targets). The relative weighting given to each of these components depends on a parameter, the 'cut-off' parameter, c. For the test here, the parameter c is chosen as $c =$ 10 pixels. This may be interpreted as errors in cardinality (e.g. targets not tracked) being penalised at the same rate as a localisation error of 10 pixels, the maximum localisation error considered. OSPA itself may be interpreted as an average error per target and hence has a maximum value of 10 pixels.

A significant problem with the NN and XIE trackers is the way in which false alarms are handled. In the NN implementation by Etgar et al. it is stated that "the bright peaks that were not used... initialize new trajectories". This means that any false alarms not assigned to existing trajectories always initialise new tracks. Similarly, in [27] on which the XIE algorithm is based, it is stated that "a track is either composed of sensor responses from a single world point or it is composed of points occurring from noise only". Therefore, it can be seen that both methods will produce track sets which contain a high level of tracks corresponding to noise. This results in noisy track sets and also means that cardinality error will be large whenever there are false alarms.

As an example of this problem, GM-PHD, NN and XIE trackers were evaluated on a simple scenario with an average of three cells in the field of view, no missed detections and some false alarms, uniformly distributed over the field of view where the number of false alarms in each frame has a Poisson distribution with mean λ. 4000 time step simulations were computed for a range of λ and the average OSPA distances are shown in fig. 3. This figure demonstrates the constistent strong performance of the GM-PHD filter while the other trackers have close to the highest possible OSPA distance even for low false alarm levels.

The results shown in fig. 3 highlight the inability of methods based on cost minimisation to handle the multitarget

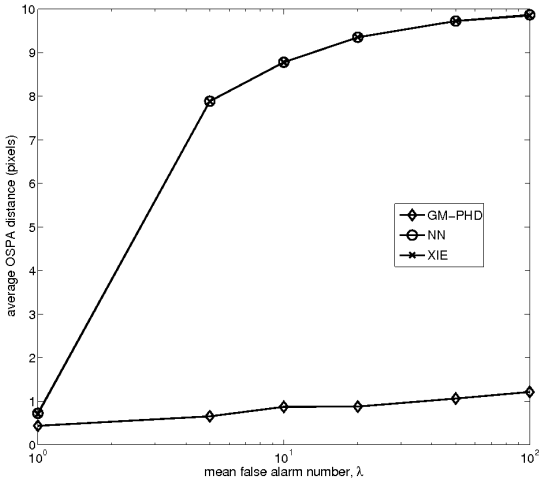


Fig. 3. Average OSPA distance for each tracker for a range of λ, the average number of false alarms.

aspect of the tracking problem, which is to jointly determine the number of targets and their locations. The PHD filter's statistical approach to handling varying target number and false alarms allows it to differentiate between false alarms and true targets directly, which results in less noisy track outputs.

2) Single Established Trajectory Performance Comparison: A further set of comparisons were implemented to determine the ability of the trackers to maintain continuous and accurate tracks on individually identified cell trajectories. In order to test performance under challenging tracking conditions the simulations included phenomena relevant to microscopic visual tracking which make the scenario more difficult. The phenomena included were uniformly distributed false alarms, spatially and temporally correlated false alarms and occlusions. Although these do not capture every aspect of microscopic visual data, they are broadly representative of the complications which make the problem difficult.

In all of the tests there were an average of 25 cells in the field of view unless otherwise stated. This is representative of the real data as shown in Section III-B. For each test, 100 individual trajectories were randomly selected. A track was initialised on the cell location in the first frame of its appearance and the average distance (in pixels) between the tracker's estimated location and the true location were measured. Distances above 10 pixels were considered lost tracks and penalised as 10 pixel errors, as in the OSPA metric with $c = 10$ pixels in Section III-D1.

Test 1: Uniform False Alarms - In this test, the data was corrupted by false alarms uniformly distributed in the field of view with the number of false alarms given by a Poisson distribution with mean λ. Simulations were carried out for a range of λ and results are shown in fig. 4. These results show that performance for GM-PHD and XIE are better than NN but broadly similar to each other for all but the highest levels of false alarms. The false alarm levels at which GM-PHD gains a significant advantage over XIE are unrealistically high for real data. Therefore, in realistic tracking conditions for uniformly distributed false alarms, the performance of GM-PHD and XIE is similar.

Test 2: Spatially and Temporally Correlated False Alarms - In order to represent the kind of false alarms seen

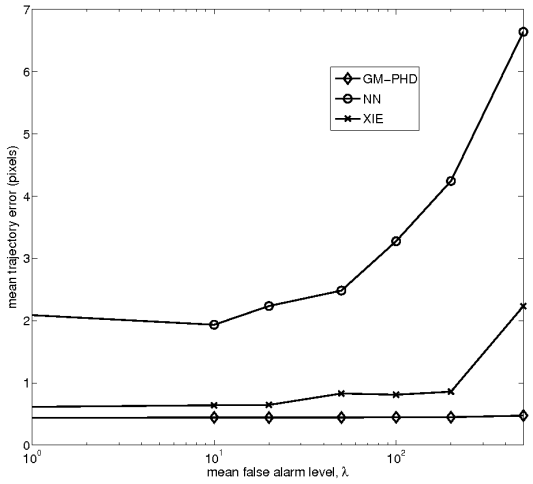


Fig. 4. Average trajectory localisation error (in pixels) for each tracker for a range of λ

in the real data, a number of false alarms were generated in the vicinity of each true cell, achieving spatial and temporal correlation in a manner similar to the false alarms observed the real data discussed in Section III-B as a result of the thresholding process. The number of false alarms generated by each cell was chosen to be Poisson distributed with mean γ. Simulations were carried out for a range of γ and the results are given in fig. 5. The results show that the advantage of the GM-PHD is greater in the case of correlated clutter than uniform clutter. An explanation for this is that the spatially correlated false alarms present conflicting options about which detection should be used to update the trajectory. Whereas the cost minimisation methods must choose one measurement, the PHD filter is able to propagate weighted hypotheses for each possible association. The results show that the statistical approach of the PHD filter is more suited to tracking a trajectory in data containing correlated false alarms.

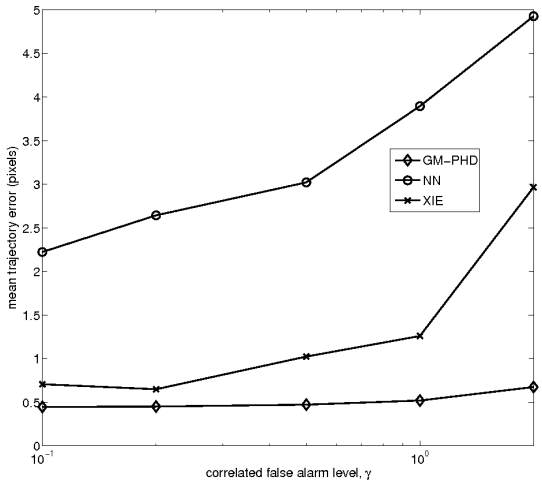


Fig. 5. Average trajectory localisation error (in pixels) for each tracker for a range of γ

Test 3: Occlusions - In this test, if the centroids of two simulated cells are within five pixels of each other, an occlusion is deemed to have occurred, and a single detection is recorded for the two cells at the midpoint of their two centroids. It is well known [8] that the PHD filter has a short memory with regard to targets for which there are missed detections (i.e. the weight of these targets drops very quickly). In order to alleviate this problem, the cardinalised PHD filter was proposed by Mahler [18]. However, a simpler approach to alleviating the problem is taken here. The existence of track labels permits a trajectory dependent target declaration threshold so that established targets with low weights can continue to be declared despite missed detections. Furthermore, occlusions can be taken into account explicitly by reducing the probability of detection, P_D for targets in the vicinity of other targets (and hence in danger of being occluded), as described in Section III-B.

In order to test the effect of occlusions on the different tracking algorithms, the cell density in the simulations was varied, as occlusions become more of a severe problem when there are more cells in the field of view. The results are given in fig. 6 which shows comparable performance between GM-PHD and XIE with significantly worse performance for NN.

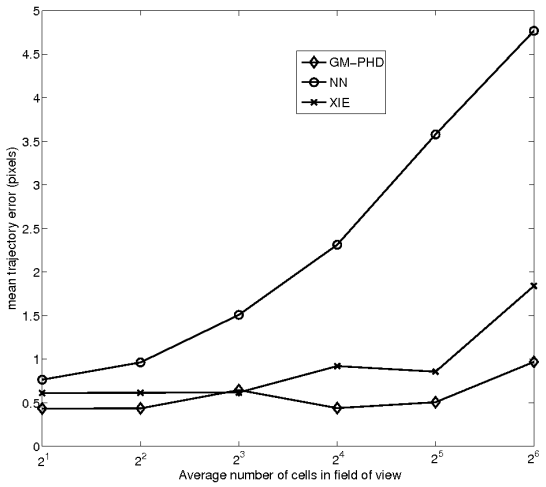
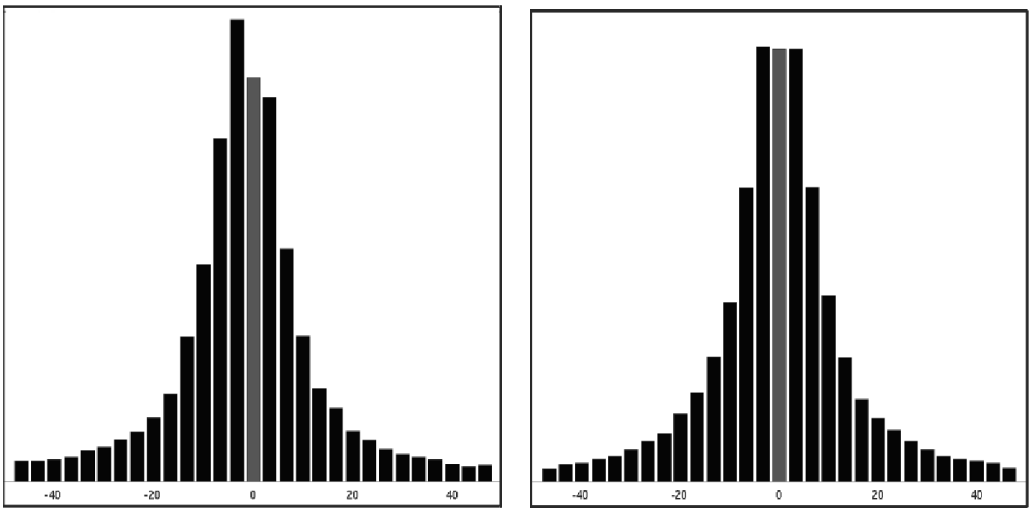


Fig. 6. Average trajectory localisation error (in pixels) for each tracker for a range of cell densities with occlusions

In summary, the performance of the GM-PHD filter is significantly better than traditional nearest neighbour based algorithms in a wide variety of tracking conditions. The tracking algorithm of Xie et al. is also able to outperform NN. The GM-PHD has a moderate performance advantage over XIE in tracking single trajectories, particularly in the presence of correlated false alarms. The real advantage of using GM-PHD is its ability to distinguish false alarms from true targets. In scenarios where there are a significant number of false alarms, the NN and XIE trackers are expected to output many tracks containing mostly or entirely false alarms whereas the GM-PHD does not, as fig. 3 clearly demonstrates.



(a) (b)

Fig. 7. Distribution of framewise angle changes: a) at the surface. b) away from the surface.

E. Results

An initial data set contained 78 minutes of video footage. Many of the tracks returned from this footage were immobile cells (which may have been, for example, dead or stuck to the surface). Tracks of immobile cells provide no useful information, so only tracks with a minimum level of movement were used. 4182 tracks were returned from 78 minutes of footage, which is nearly two orders of magnitude more than the 59 and 50 tracks returned in [5] and [31] respectively. The large number of tracks returned makes it possible to build up accurate probability distributions of parameters of interest from bacterial motion, an example of which is given below.

A bias towards spiralling in one orientation is consistent with fluid dynamical models for bacterial motion near a surface [29]. A direct comparison of this bias at the surface as compared with behaviour away from the surface was not previously possible. Figure 7 gives the distribution of angle changes (per frame) over the whole data set, showing a clear bias towards clockwise turns at the surface which is not present for cells away from the surface. This verifies the prediction from fluid dynamical models of bacterial motion and is an example of how the extraction of a large number of reliable tracks, free of noise related tracks, permits immediate in-depth, quantitative analysis of a range of features of bacterial motion.

Publications with further details on methods of analysis and biological impact are in preparation.

1) Computational Speed: Results were obtained using non-optimised MATLAB code on a machine with an AMD Phenom(tm) II X3 720 2.8GHz Processor. Performing the GM-PHD filter on 9000 frames of measurements with an average of 15 targets present took 543 seconds, giving a processing rate of about 16.5fps. The detection step may take longer depending on the image resolution and method used. It should be noted that computational complexity of GM-PHD is linear in both the number of measurements and the number of targets, so the speed will depend on both the number of cells present and the level of false alarms. Therefore, close to real time implementation of the

GM-PHD filter should be possible for 5 targets present at 50fps

A similar implementation on the same data set using the improved SMC-PHD filter with 2000 particles per target and 500 particles for new targets (chosen to enable direct comparison with the computational results presented in [16]) was carried out. This implementation took 6807 seconds, giving a rate of around 1.3 fps. The closest equivalent calculation in [16] is for 12 targets and the processing time given corresponds to around 0.7 fps. There are several factors which might explain the two-fold decrease in computation time as compared with the results in [16]. The improved SMC-PHD does not need clustering or data association, which reduces the computation. Furthermore, the presence of track labels within the PHD filter enables a gating process, as described in [36], which may have increased computational efficiency. Finally, some of the difference may be due to hardware, although the 3.2GHz processor used in [16] should be comparable to the 2.8GHz processor used here.

The difference in performance between the standard SMC-PHD filter and improved SMC-PHD gives a roughly two-fold increase in performance (ignoring hardware differences), but the striking difference is that gained by using GM-PHD which gives a 25-fold increase in speed by comparison with the computational results reported in [16]. This 25-fold decrease in computation may provide the most compelling practical argument for adopting GM-PHD for visual tracking.

IV. Summary and Future Work

The GM-PHD [33] and improved SMC-PHD [25] [36] filters have been presented with a view to implementation on visual tracking data. These implementations are shown to avoid the complications inherent in previous PHD filter implementations on visual data by Maggio et al. [16] and Wang et al. [35].

It is anticipated that in many visual tracking scenarios, it will be possible to use GM-PHD or one of its nonlinear extensions, which provide a significant computational advantage over SMC-PHD implementations.

The GM-PHD filter is applied to the problem of tracking bacteria in microscopic visual data. Using this technique, a significantly larger number of tracks is extracted than was previously possible for free-swimming bacteria. Using this large number of tracks, it is possible to build reliable distributions for parameters of interest for bacterial motion. This is being exploited to obtain potentially biologically significant results.

The output of the GM-PHD filter has been verified using a careful manual analysis showing good results. The GM-PHD filter has also been compared with alternative tracking methods using simulations. The simulated tests demonstrated the advantage of working within the finite set statistics framework in order to resolve the multitarget tracking problem in avoiding the declaration of false tracks whenever false alarms were present. Favourable performance is of the GM-PHD filter in comparison to other state-of-the-art tracking algorithms was also shown for the monitoring of a single trajectory.

Future work on multitarget tracking of bacteria could include implementing the method of Hozeinnezhad et al. [10] which avoids the detection step by working with image histograms within the finite set statistics framework.

ACKNOWLEDGEMENT

T. M. Wood would like to thank his academic supervisors David Allwright, Philip Bond and Irene Moroz, all with the University of Oxford, as well as Roger Benton and Stephen Long with Thales Underwater Systems and Glen Davidson with Thales Aerospace for guidance and useful discussions.

C.A.Y. would like to acknowledge the support and help of his supervisors Philip Maini, Ruth Baker and Radek Erban. He would also like to acknowledge support from the Engineering and Physical Sciences Research Council via the Doctoral Training Centre (University of Oxford) and a Scholarship from St. Catherine's College, (University of Oxford).

G. R. thanks his academic supervisors Philip Maini, Ruth Baker and Mark Leake for their support. He is also grateful to Alexander Fletcher for many helpful discussions and acknowledges financial support from the Engineering and Physical Sciences Research Council via the Doctoral Training Centre (University of Oxford).

D.A. thanks Prof. Judy Armitage from the Oxford centre for Integrative systems biology for her support.

REFERENCES

[1] B.D. Anderson and J.B. Moore, *Optimal Filtering*, Prentice-Hall, 1979

[2] J.P. Armitage and H.L. Packer, "Bacterial Motility and Chemotaxis", *In D.R. Soll and D. Wessels: Motion Analysis of Living Cells"*, pp. 1-24, 1998

[3] Y. Bar-Shalom and T.E. Fortmann, *Tracking and data association*, Academic Press, 1988

[4] H.C. Berg, "How To Track Bacteria", *Rev. Sci. Instrum.*, vol, 42, no. 6, 1971

[5] H.C. Berg and D.A. Brown, "Chemotaxis in Escherichia Coli analysed by Three-dimensional Tracking", *Nature*, vol. 239, pp. 500–504, 1972

[6] D. Campos, V. Méndez and I. Llopis, "Persistent random motion: uncovering cell migration dynamics", *Journal of theoretical biology*, vol 267, no. 4, pp. 526-534, 2010

[7] A.A. Chernova, J.P. Armitage, H.L. Packer and P.K. Maini, "Response kinetics of tethered bacteria to stepwise changes in nutrient concentration", *Biosystems*, vol 71, pp. 51-59, 2003

[8] O. Erdinc, P. Willett, and Y. Bar-Shalom, "Probability hypothesis density filter for multitarget multisensor tracking", *Proc. of the 7th International Conference on Information Fusion (FUSION)*, vol. 1, pp. 146-153, 2005

[9] L. Etgar, A. Nakhmani, A. Tannenbaum, E. Lifschitz, R. Tannenbaum, "Trajectory control of PbSe-γ-Fe_2O_3 nanoplatforms under viscous flow and an external magnetic field", *Nanotechnology*, vol. 21, no. 17, 2010

[10] R. Hoseinnezhad, B.-N. Vo, D. Suter and B.-T. Vo, "Multi-object filtering from image sequence without detection", *2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 1154-1157, 2010

[11] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S.L. Schmidt and G. Danuser, "Robust single-particle tracking in live-cell time-lapse sequences", *Nature Methods*, vol. 5, no. 8, pp. 695-702, 2008

[12] R. Juang, A. Levchenko and P. Burlina, "Tracking cell motion using GM-PHD", *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1154-1157, 2009

[13] S.J. Julier and J.K. Uhlmann, "Unscented filtering and nonlinear estimation", *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, 2004

[14] D.B. Kearns, "A field guide to bacterial swarming motility", *Nat. Rev. Microbiol*, vol. 8, pp. 634-644, 2010

[15] G. Li., L.K. Tam, and J.X. Tang, "Amplified effect of brownian motion in bacterial near-surface swimming.", *Proc. Natl. Acad. Sci USA*, vol 105, pp. 18355-18359, 2008

[16] E. Maggio, M. Taj and A. Cavallaro, "Efficient Multitarget Visual Tracking Using Random Finite Sets", *IEEE Trans. Circ. Syst. Vid.*, vol 18, no 8, pp. 1016–1027, 2008

[17] R.P.S. Mahler, "Multi-target bayes filtering via first-order multi-target moments", *IEEE Trans. Aero. Elec. Sys.*, vol 39, no. 4, pp. 1152-1178, 2003

[18] R.P.S. Mahler, *Statistical Multisource-Multitarget Information Fusion*, Artech House, 2007.

[19] N.T. Ouellette, H. Xu, E. Bodenschatz, "A quantitative study of three-dimensional Lagrangian particle tracking algorithms", *Experiments in Fluids*, vol. 40, pp. 301-313, 2006

[20] K. Panta, D.E. Clark and B.-N. Vo., "Data Association and Track Management for the Gaussian Mixture Probability Hypothesis Density Filter", *IEEE Trans. Aero. Elec. Sys.*, vol. 45, no3, pp. 1003-1016, 2009

[21] S. A. Pasha, B.-N. Vo, H. D. Tuan, and W.-K. Ma, "A Gaussian mixture phd filter for jump Markov system models", *IEEE Trans. Aero. Elec. Sys.*, vol. 45, no. 3, pp. 919-936, 2009.

[22] S.L. Porter, G.H. Wadhams, J.P. Armitage "Rhodobacter sphaeroides: complexity in chemotactic signalling", *Trends Microbiol.*, vol. 16 no. 6, 2008

[23] K. Punithankumar, T. Kirubarajan and A. Sinha, "Multiple-Model probability hypothesis density filter for tracking maneuvering targets", *IEEE Trans. Aero. Elec. Sys.*, vol 44, no 1, pp. 87–98, Jan 2008

[24] B. Ristic, S. Arulampalam and N. Gordon, *Beyond the Kalman Filter*, Artech House, 2004

[25] B. Ristic, D. Clark, B.-N. Vo, "Improved SMC implementation of the PHD filter", *Proc. 13th Int. Conf. Information Fusion, Edinburgh, 2010*

[26] M.J. Saxton, "Single-particle tracking: connecting the dots", *Nature Methods*, vol. 5, no. 8, 2008, 2008

[27] K. Shafique and M. Shah, "A noniterative greedy algorithm for multiframe point correspondance", IEEE. Trans. Pattern, Anal. Mech. Intell., vol. 27, no. 1, pp. 3962-3971, 2005

[28] D. Schuhmacher, B.-T. Vo and B.-N. Vo, "A consitent metric for performance evaluation of multi-object filters", *IEEE Trans. Signal Proces.*, vol. 56, no. 8, pp. 3447–3457, 2008

[29] H. Shum, E.A. Gaffney and D.J. Smith, "Modelling bacterial behaviour close to a no-slip plane boundary: the influence of bacterial geometry", *Proc, R. Soc. A*, vol 466, pp. 1725-1748, 2010

[30] H.W. Sorenson and D.L. Alspach, "Recursive Bayesian Estimation Using Gaussian Sums", *Automatica*, vol 7, pp 465–479, 1971

[31] B. Taboada, S. Poggio, L. Camarena, G. Corkidi, "Automatic Tracking and analysis system for free-swimming bacteria", *Proc. 25th Annual Conf. of the IEEE EMBS, 2003*

[32] B.-N. Vo, S. Singh and A. Doucet "Sequential Monte Carlo methods for multi-target filtering with random finite sets," *IEEE Trans. Aero. Elec. Sys.*, vol 41 pp. 1224-1245, 2005.

[33] B. Vo, and W. Ma "The Gaussian mixture probability hypothesis density," *IEEE Trans. Signal Proces.*, vol 54, pp. 40914104, 2006.

[34] N. Verstraeten, K. Braeken, B. Debkumari, M. Fauvart, J. Fransaer, J. Vermant, and J. Michiels, "Living on a surface: Swarming and biofilm formation", *Trends Microbiol.*, vol 16, pp. 496-506, 2008

[35] Y.-D. Wang, J.-K. Wu, A.A. Kassim and W. Huang, "Data-Driven Probability Hypothesis Density Filter for Visual Tracking", *IEEE Trans. Circ. Syst. Vid.*, vol 18, no 8, pp. 1085–1095, 2008

[36] T. Wood, D. Clark and B. Ristic, "Efficient resampling and basic track continuity for the SMC-PHD filter", *Proc. Cognitive Systems with Interactive Sensors 2010*

[37] T. Wood, "Interacting methods for manoeuvre handling in the GM-PHD filter", *IEEE Trans. Aero. Elec. Sys., In Press*

[38] J. Xie, S. Khan and M. Shah, "Automatic tracking of *Escherichia Coli* bacteria", *11th International Conference on Medical Imag Computing and Computer Assisted Information, MICCAI*. September 6-10, 2008, New York City.

[39] J. Xie, S. Khan and M. Shah, "Automatic Tracking of *Escherichia Coli* in phase-contrast microscopy", *IEEE Trans. Bio-Med. Eng.*

Trevor M. Wood was born in Greenwich, England in 1985. He received a B.A. degree in mathematics (first class honours) in 2006 and an MSc. degree in mathematical modelling and scientific computing (with distinction) in 2007, both from the University of Oxford. He is currently working towards a DPhil degree in sonar and radar tracking at OCIAM, University of Oxford.

David A. Wilkinson was born in Sevenoaks, England in 1983. He received an MBiochem. degree in 2006, and a DPhil. degree in 2011, both from the University of Oxford. He is currently employed as a research associate by the Oxford Centre for Integrative Systems Biology, University of Oxford

Christian A Yates was born in Manchester, England in 1985. He received a B.A. degree in mathematics (first class honours) in 2006 and an M.Sc. degree in mathematical modelling and scientific computing (with distinction) in 2007, both from the University of Oxford. He is currently working towards a D.Phil. degree in the modelling of cell migration at the Centre for Mathematical Biology, University of Oxford.

Gabriel Rosser was born in Bromley, England in 1984. He received an MChem. degree in chemistry (first class) in 2007 from the University of Oxford. He is currently working towards a DPhil. degree in bacterial taxis and population dispersion at the Centre for Mathematical Biology, University of Oxford.

REVIEW RESPONSES

We would once again like to thank the editor and the anonymous reviewers for their attention.

During the previous round of review, the issue was raised that some details of the numerical simulations were only referenced in an unpublished paper. This prevented the manuscript from being self-contained and also made it impossible for other researchers to reproduce the experiments. This error has now been corrected as full details of the simulation of bacterial motion are now included in Section III-D. Further details of changes made in response to the reviews are given below:

Review # 3:

The authors have augmented the section describing experiments on synthetic data (sec. III D) with results from a dataset generated using a model described in [25]. However the referenced work is unpublished and is neither available as a technical report. This leads to two problems: first as a reviewer I cannot assess the validity of the model and consequently of the results; second, as a researcher it would be impossible to reproduce and compare the author's results.

We have added full details of the model used in the simulation of the bacterial trajectories in Section III-D to address this issue.

Minor issues: -Sec III.D (1), first line: "evaluation in multitarget tracking is often performed using OSPA". I do not agree with this statement, OSPA is a fairly new (2008) and quite unknown evaluation metric. I suggest to rephrase this sentence.

This sentence has been reworded to take this into account.

-Fig.3: I feel that the comparison between GM-PHD, NN and XIE is somehow unfair. It is true that XIE and NN generate new tracks for every unassigned detection. However in practice naive heuristics are commonly used to filter out spurious trajectories. For example one might want to discard all the trajectories shorter than x frames. An interesting question to answer is whether GM-PHD is better than NN (or XIE) combined with simple heuristic.

We appreciate that in practice simple heuristics are used to improve tracker performance. However, many different engineers may have different favourite heuristics (indeed, we have our own heuristics that we occasionally use in conjunction with the GM-PHD filter) each giving a different overall performance. We think it is simpler and fairer to just evaluate the performance of the well-known "vanilla" algorithms.

As the point is a very relevant one, we have included a sentence to bring this issue to the attention of the reader.