18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

A Survey on Joint Tracking Using EM Based Techniques

Hua Lan*†, Xuezhi Wang‡, Yan Liang*†, Feng Yang*†, Zengfu Wang*†, Zhiyuan Shi*†
*School of Automation, Northwestern Polytechnical University, Xi'an, Shaanxi, 710072, P.R.China
†Key Laboratory of Information Fusion Technology, Ministry of Education, Xi'an, Shaanxi, 710072, P.R.China
‡School of Electrical & Computer Engineering, RMIT University, Melbourne 3000, Australia

***Abstract*—Many target tracking problems can actually be cast as joint tracking problems where the underlying target state may only be observed via the relationship with a latent variable. In the presence of uncertainties in both observations and latent variable, which encapsulates the target tracking into a variational problem, EM method provides an iterative procedure under Bayesian estimation framework to estimate/tracking the state of target in the process of minimise the latent variable uncertainty. In this paper, we describe this EM based Bayesian inference method in the joint tracking context. Some typical application examples in target tracking are presented.**

***Index Terms*—Target tracking, joint identification and estimation, expectation maximization, Bayesian inference**

I. INTRODUCTION

Target tracking is an information processing outcome which provides the trajectories of one or more moving objects based on sensor data and prior information of target motion, which is essential for many applications in the areas of defense, medical science, traffic control, navigation, and etc [1]. Particularly, the advances in multi-target tracking techniques have opened up numerous research venues as well as application areas [2].

Target tracking involves a wide range of topics including single sensor and multi-sensor tracking problems, in the later case, centralized or distributed structures are taken into account. Within each of these topics, tracking can be further categorized into single target and multi-target tracking problems, where uncertainties induced by false alarms and low sensor detection probability, data association between target and measurement, etc. must be addressed. While fully covered by the theories of estimation and optimization, the advanced target tracking techniques which reflect the cutting-edge research effort are under multi-disciplines. Since the modern target tracking are often required to deal with various uncertainties, nonlinear and high dimension issues simultaneously.

Roughly speaking, target tracking concerns two major problems: *what* and *where* the underlying target is. The former is related to target identity and the latter corresponds to target trajectory estimation. In target tracking context, parameter estimation refers in particular to estimating target kinematic state, while identification refers to resolving other unknown or random quantities, including the models of systems, data/track association states and latent variables which are related to target state under multiple hypothesis. For instance, in the case of tracking a maneuvering target in the presence of clutter, the parameter to be estimated is the state of target including position and velocity, while the identification signifies the measurement-to-track association and possible modes of maneuver. In general, the identification and estimation are highly related with each other like "chicken and egg". A poor identification of data association can deteriorate state estimation while the error yielded in the parameter estimation may result in a wrong target identification/labeling. As pointed out in [3], [4], jointly considering the identification risk and estimation error is more promising than separate identification and estimation. In the optimization theoretic parlance, this joint approach has the potential of arriving at a globally optimal solution, and its solution can always be obtained by an iterative algorithm. Consequently, the formulation of the joint identification and estimation problem which allows this problem to be solved iteratively is crucial. In our recent work [5], [6], [7], [8], we formulated target tracking problem as a joint tracking problem where the underlying target state x may only be observed via the relationship with a latent variable Θ. The object of target tracking is to estimate the target state x by using the observation y under the latent variable Θ, where the state x and latent variable Θ correspond to estimation and identification, respectively. We addressed such a problem using the expectation-maximization (EM) approach.

The EM algorithm proposed by A. Dempster et al [9] is an iterative inference method for the learning of parameters x in the presence of latent variables Θ. In general, it is an off-line, iterative, maximum likelihood method that is guaranteed to converge to a local maximum of the observed data likelihood function. The EM algorithm is based on the observation that the maximization of the complete data likelihood $P(\Theta, y|x)$, is usually easier than the maximization of the observed data likelihood function. The latent variables are unknown however. The most recent information about the latent variables is given by the posterior distribution $P(\Theta|x, y)$. Since the posterior over the latent variables Θ requires the knowledge of the parameters x and the maximization of the complete data likelihood requires the latent variables Θ, this result in a chicken-and-egg" problem. The EM algorithm uses an iterative procedure to circumvent this "chicken-and-egg" problem: after making an initial estimate of the target state x two steps including E-step and M-step are iterated, whereby the target state x and latent variable Θ are estimated and identified in the iterated loop, respectively. The most attractive of the EM algorithm is its convergence guarantee. Nevertheless, to apply the EM algorithm we must have knowledge of the posterior of the hidden variables given the observations, while both the E-step and M-step can still remain intractable when the dimensionality of the latent variables are too high [10].

In this paper, we consider the target tracking problems as a joint estimation and identification problem where the underlying target state may only be observed via the relationship with a latent variable. We show that the problem can be formulated and solved under a unified Bayesian estimation framework, though it is analytically intractable since the problem involves the *uncertainty*, *multi-mode*, *nonlinear* and *high-dimension*. The EM based Bayesian inference method is described in the joint tracking context to deal with the latent variable uncertainty. We derive the EM algorithm for joint estimation and identification problem in a way which leads to a better understanding and get more insights into the underlying joint tracking problem.

II. Bayesian Joint Tracking Problem

Consider the multisensor multitarget tracking system as:

$$x_k^i = f_{k-1}^{i,r}(x_{k-1}^i, a_{k-1}^i) + \Gamma_{k-1}^i v_{k-1}^i \quad (1)$$

$$y_k^{j,\mu,\tau} = h_k^{j,\tau}(x_k^i, b_k^{j,\tau}) + w_k^{j,\tau} \quad (2)$$

where x_k and y_k represent the system state and measurement. The state transition function f_k, measurement function h_k, and control matrix Γ_k are given. The continued-value a_k and b_k are unknown disturbed inputs in dynamic model and measurement model, respectively. The process noise v_k and measurement noise w_k are zero-mean white Gaussian noises with the known covariance Q_k and $R_k > 0$. The initial state x_0 is Gaussian distributed with known mean $\bar{x}_0$ and associated covariance Σ_0. The superscript $i \in \{1, 2, ..., t\}$, $r \in \{1, 2, ..., \iota\}$, $j \in \{1, 2, ..., s\}$, $\mu \in \{1, 2, ..., m_j\}$, and $\tau \in \{1, 2, ..., l_\mu\}$ are index of target, sensor, motion mode, measurement and propagation mode with $\{t, \iota, s, m_j, l_\mu\}$ are corresponding number. The superscript k is time instant.

Definition 2.1 Define $\mathcal{X}_k \triangleq \{x_k^1, \cdots, x_k^t\}$ as the states of targets with $i$the target state x_k^i, and $\mathcal{Y}_k \triangleq \{Y_k^1, \cdots, Y_k^s\}$ as the measurements of all sensors with jth sensor measurement set $Y_k^j \triangleq \{y_k^{j,1}, \cdots, y_k^{j,m^j}\}$, respectively.

Definition 2.2 Define the track-track association as $\alpha_k \triangleq \{i, j\}$, track-motion mode association as $\beta_k \triangleq \{i, r\}$, track-measurement association as $\gamma_k \triangleq \{i, \mu\}$, and measurement-propagation mode association as $\delta_k \triangleq \{i, \tau\}$, respectively. Let $\Theta_k \triangleq \{f_k, h_k, R_k, Q_k, \bar{x}_0, \Sigma_0, \alpha_k, \beta_k, \gamma_k, \delta_k, a_k, b_k\}$ as model dependent parameters.

The Bayesian inference is the foundational of modern target tracking, and its central task is the evaluation of the posterior distribution $p(\mathcal{X}|\mathcal{Y})$ of the state variables $\mathcal{X}$ given the observed data variables $\mathcal{Y}$, and the evaluation of expectations computed with respect to this distribution [11]. The Bayesian inference for multisensor multitarget tracking problem is computationally intractable in all but the simplest problem [12]. However, as described by John W. Tukey, *"An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate problem"*. Assuming the model parameter Θ_k are exactly known, and given the measurement $\mathcal{Y}$, the Bayesian inference is to obtain the probability density function of $\mathcal{X}$ conditioned on $\mathcal{Y}$ and Θ_k, i.e.,

$$p(\mathcal{X}_k|\mathcal{Y}_k) = p(\mathcal{X}_k|\mathcal{Y}_k, \Theta_k) \quad (3)$$

where the model parameters Θ_k are only partly known, that is, the association hypothesis α_k, β_k, γ_k, δ_k, and the inputs a_k and b_k are unknown. It turns out that the posterior distribution is in the form of hybrid, high dimensional integrations:

$$p(\mathcal{X}_k|\mathcal{Y}_k) = \sum_{\alpha_k}\sum_{\beta_k}\sum_{\gamma_k}\sum_{\delta_k}\int_{a_k}\int_{b_k} p(\mathcal{X}_k|\mathcal{Y}_k, \alpha_k, \beta_k, \gamma_k, \delta_k, a_k, b_k) da_k db_k \quad (4)$$

Remark 2.1: The key of target tracking is to solve the integration of hybrid system (4), which has the complex of uncertainty (continue variables a_k and b_k), multi-mode (discrete variables α_k, β_k, γ_k, and δ_k), high dimensionality of latent space Θ_k. Essentially, it is *NP-hard* problem. In the case of continuous variables, the required integrations may not have closed-form analytical solutions, while the dimensionality of the space and the complexity of the integrand may prohibit numerical integration. For discrete variables, the marginalization involve summing over all possible configurations of the hidden variables, there may be exponentially many hidden states so that exact calculation is prohibitively expensive in practice.

III. The EM framework for Joint Tracking

A. Expectation Maximization Algorithm

The EM algorithm has been widely used in the engineering and statistical literature as an iterative optimization procedure for computing maximum likelihood (or MAP) parameter estimates of *incomplete data problem*. Since formally introduced by Dempster et al. (1977) in three decades ago, the EM algorithm had received tremendous attention from researchers across different disciplines. The popularity of EM lies mainly in its ability to find the ML or MAP estimation for many popular statistics models in a widely acclaimed manner of simplicity and stability [13]. A good survey on the EM algorithm can be found in [14]. The common derivation of EM algorithm is based on Jensens' inequality. Here, we briefly describe the EM algorithm under the criterion of Maximum Likelihood Estimation (MLE).

Consider a probabilistic model in which we collectively denote all of the observed variables by $\mathcal{Y}$ and all of the hidden variables by Θ. The joint distribution $p(\mathcal{Y}, \Theta|\mathcal{X})$ is governed by a set of parameters denoted $\mathcal{X}$. Our goal is to maximize the log likelihood function given by

$$L(\mathcal{X}) = \log p(\mathcal{Y}|\mathcal{X}) = \log \sum_{\Theta} p(\mathcal{Y}, \Theta|\mathcal{X}) \tag{5}$$

A key observation is that the summation over the latent variables appears inside the logarithm, which results in complicated expressions for the maximum likelihood solution. The EM algorithm can be derived using *Jensens inequality*, i.e.,

$$l(\mathcal{X}|\mathcal{X}^{(n)}) \geq \sum_{\theta} p(\theta|\mathcal{Y}, \mathcal{X}^{(n)}) \log \frac{p(\mathcal{Y}|\theta, \mathcal{X}) p(\theta|\mathcal{X})}{p(\theta|\mathcal{Y}, \mathcal{X}^{(n)}) p(\mathcal{Y}|\mathcal{X}^{(n)})} \tag{6}$$

We emphasize that the object is to choose a values of $\mathcal{X}$ so that $L(\mathcal{X})$ is maximized. Directly optimizing of $L(\mathcal{X})$ is difficult, but optimizing the lower bound function $l(\mathcal{X}|\mathcal{X}^{(n)})$ is often significantly easier. Therefore, the EM algorithm calls for selecting $\mathcal{X}$ such that $l(\mathcal{X}|\mathcal{X}^{(n)})$ is maximized. That is,

$$\begin{aligned} \mathcal{X}^{(n+1)} &= \arg\max_{\mathcal{X}} \sum_{\Theta} p(\Theta|\mathcal{Y}, \mathcal{X}^{(n)}) \log p(\mathcal{Y}, \Theta|\mathcal{X}) \quad (7) \\ &= \arg\max_{\mathcal{X}} E_{\Theta|\mathcal{Y}, \mathcal{X}^{(n)}} (\log p(\mathcal{Y}, \Theta|\mathcal{X})) \end{aligned}$$

The conditional expectation $E_{\Theta|\mathcal{Y}, \mathcal{X}^{(n)}} (\log p(\mathcal{Y}, \Theta|\mathcal{X}))$ is also called Q function and written as $\mathcal{Q}(\mathcal{X}|\mathcal{X}^{(n)})$, which takes the expectation of complete-data likelihood with respect to the latent variables Θ. The EM algorithm completes the MLE by iteratively operating the following two steps:

- *E-step*: Determine the conditional expectation Q function;
- *M-step*: Maximize this expression with respect to $\mathcal{X}$.

Remark 3.1: The convergence of EM algorithm depends on filter initialisation, likelihood property and Q function.

Remark 3.2: The EM algorithm involves the two iterative steps E-step and M-step, which correspond exactly to the estimation and identification problem with the target tracking. Due to the inherent iterative convergence characteristics of EM algorithm, the closed feedback loop between state estimation $\mathcal{X}$ and identification Θ, which is desirable for dealing with the couple problem between estimation and identification.

Remark 3.3: The EM algorithm requires that $p(\Theta|\mathcal{X}, \mathcal{Y})$ is explicitly known, or at least we should be able to compute the conditional expectation of its sufficient statistics Q-function in (8). While $p(\Theta|\mathcal{X}, \mathcal{Y})$ is in general much easier to infer than $p(\mathcal{Y}|\mathcal{X})$, in many interesting problems, especially when the hidden variables are of high-dimensions, it is not possible and thus the EM algorithm is not applicable.

After it was formalized and generalized by A.P. Dempster et al, EM has received tremendous attention in many different research areas. Great efforts were made to simplify the implementation of EM in some situations, in particular, when there is no analytical solution available for the calculation of $\mathcal{Q}(\mathcal{X}|\mathcal{X}^{(n)})$ in the E-step or for the maximization of it in the M-step. Here, we briefly review several developments of the EM algorithm, including its convergence, initialization, extensions and properties.

1) Convergence of the EM Algorithm: The appealing property of EM algorithm is that the likelihood function monotonously increases with respect to the iteration number before reaching the local optimization value. However, the monotonicity of the EM algorithm guarantees that as EM iterates, its guesses won't get worse in terms of their likelihood, but the monotonicity alone cannot guarantee the convergence of the sequence $\{\mathcal{X}^{(n)}\}$ with n is the nth iteration. As stated in [13], there is no general convergence theorem for the EM algorithm: the convergence of the sequence $\{\mathcal{X}^{(n)}\}$ depends on the characteristics of likelihood function and Q-function, and also the starting point $\mathcal{X}^{(0)}$. Under certain regularity conditions, A.P. Dempster et al prove that $\{\mathcal{X}^{(n)}\}$ converges to a stationary point with a linear convergence rate. The detailed discussion on EM convergence is given by C. F. J. Wu [15], which is based on Zangwill's global convergence theorem. Consider the joint state estimation and mode identification of jump Markov linear systems, three EM schemes were presented and their convergence was given [16]. The convergence rate of the EM algorithm is addressed in [17]. However, the convergence of EM algorithm is still a open problem.

2) Initialization of the EM Algorithm: The EM algorithm is often sensitive to the choice of the initial parameter vector, efficient initialization is an important preliminary process for the future convergence of the algorithm to the best local maximum of the likelihood function [18]. All initialization strategies can be generally classified as deterministic or stochastic. Some popular deterministic initialization approaches choose starting values based on the solution obtained from hierarchical clustering, model-based Gaussian hierarchical clustering [19], or the multistage procedure based on finding the best local modes. A considerable disadvantage of all deterministic methods is their incapability to propose another starting point. The proposed starting value may lead to an incorrect solution or even no solution when the likelihood function is unbounded. Stochastic initialization strategies do not share this shortcoming as they normally allow restarting from another point of the parameter space. The general idea is to try different starting values of parameters and choose the one that yields the largest local maximum. Because of the need to repeat the initialization step several times, these procedures are typically more time consuming than deterministic approaches. Among the well-known stochastic initialization methods are the emEM [20] and RndEM [21] algorithms, share the common idea of trying different initial values of parameters and choosing the one that yields the largest local maximum.

3) Extensions of the EM Algorithm: The EM algorithm breaks down the potentially difficult problem of maximizing the likelihood function into two stages, the E-step and the M-step, each of which will often prove simpler to implement. Nevertheless, for complex models it may be the case that either the E-step or the M-step, or indeed both, remain intractable. This leads to several extensions of the EM algorithm as follows. More details see [22].

The *generalized EM (GEM)* [23] algorithm addresses the problem of an intractable M-step. Instead of aiming to maximize Q-function $\mathcal{Q}(\mathcal{X}|\mathcal{X}^{(n)})$ with respect to $\mathcal{X}$, it seeks instead to change the parameters in such a way as to increase its value. Again, because $\mathcal{Q}(\mathcal{X}|\mathcal{X}^{(n)})$ is a lower bound on the log likelihood function, each complete EM cycle of the GEM algorithm is guaranteed to increase the likelihood function.

The *expectation conditional maximization (ECM)* [24] algorithm replaces the M-step of EM with a sequence of simpler constrained or conditional maximization (CM)-steps, indexed by $s = 1, \cdot, S$. The advantage of the above strategy is that, in many cases, the CM-steps can be very simple (either analytical solutions or elementary numerical solutions are available) while $\mathcal{Q}(\mathcal{X}|\mathcal{X}^{(n)})$ itself is difficult to optimize directly over the whole parameter space.

The *expectation conditional maximization either (ECME)* [25] algorithm is an extension of ECM, which further partitions the CM-steps into two groups ψ_Q and ψ_L with $\psi_Q \bigcup \psi_L = \{1, \cdot, S\}$. While the CM-steps indexed by $s \in \psi_Q$ remain the same with ECM, the CM-steps indexed by $s \in \psi_L$ remain the same with EM. It has been shown theoretically and empirically that ECME typically has a greater speed of convergence than ECM, and enjoys the same stability as EM with typically higher efficiency than EM.

The *Monte Carlo EM (MCEM)* [26] algorithm addresses the problem of an intractable E-step. Instead of computing the intractable analytical solution of $\mathcal{Q}(\mathcal{X}|\mathcal{X}^{(n)})$, the MCEM algorithm approximate it via the Monte Carlo method. When implementing MCEM, maintaining a balance between efficiency and accuracy is important and relies on a smart choice of the sample size in the E-step. In general, the MCEM algorithm converges almost surely to the standard EM auxiliary function thanks to the law of large numbers.

The *parameter expanded EM (PX-EM)* [27] algorithm speeds EM by expanding the complete-data mode $p(\mathcal{Y}, \Theta|\mathcal{X})$ to a larger model $p(\mathcal{Y}, \Theta|\mathcal{X}, \alpha)$, and α is an auxiliary parameter whose value is fixed at α_0 in the original model. Then, the PX-EM algorithm can be considered to be an EM algorithm with respect to the expanded model. Actually, the PX-EM algorithm turns a low dimensional problem to a high dimensional one and it is developed for acceleration.

The *accelerated EM (AEM)* [28] algorithm is developed by appending a line search to each EM iteration. Formally, given starting value $\mathcal{X}_0$, in the $(n+1)$th iteration, the new estimation $\mathcal{X}^{(n+1)}$ is computed by $\mathcal{X}^{(n+1)} = \mathcal{X}^{(n)} + \alpha^{(n)} d^{(n)}$, where $d^{(n)}$ is a direction composed from the current direction and the previous directions, and $\alpha^{(n)}$ is a step size typically computed from a line maximization of the complete-data likelihood. The AEM algorithm usually converges much faster than EM, since conjugate direction method is considered to be one of the best general purpose optimization methods in terms of both stability and efficiency.

B. Applications of Target Tracking Base on EM

Consider the following discrete linear dynamic systems

$$x_k = F_{k-1}x_{k-1} + \Gamma_{k-1}v_{k-1} \quad (8)$$

$$y_k = H_k x_k + w_k \quad (9)$$

where x_k, v_k and w_k are Gaussian and mutually independent.

Compare the multisensor multitarget tracking system described by equations (1) - (2) with equations (8) - (9), the difference is the parameter Θ is partly known in multisensor multitarget tracking. The single sensor single target tracking without clutter is a special case when Θ is completely known, which is turned into the standard KF problem. In this sense, the KF can be regarded as the complete data estimation problem. In practice, the parameter Θ always cannot known completely, which makes the target tracking suffer from complex situation such as maneuvering target tracking, data association in the case of multiple target tracking or in the clutter environment and etc. Here, we review different applications of target tracking by using the EM algorithm according to different unknown parameter θ with $\theta \subset \Theta$. Actually, the EM algorithm has been used widely in target tracking, ranging from single sensor target tracking to cooperate tracking in sensor networks.

1) Single Sensor Multiple Target Tracking: Parameter estimation for linear dynamic systems $\left(\theta \triangleq \{F, H, R, Q\}\right)$: It is known from the theory that the KF is optimal in case that the model perfectly matches the real system, the entering noise is white and the covariances of the noise are exactly known. That is, the parameter θ is completely known. Shumway and Stoffer [29] first presented an EM algorithm for linear dynamical systems where the parameter θ is partly known (the measurement function h is known), this work was further modified by Ghahramani and Hinton [30], which presented a basic form of the EM algorithm with h unknown.

Nonlinear dynamic systems with model uncertainties $\left(\theta \triangleq \{h, R, Q\}\right)$: Consider the nonlinear state estimation problem with possibly non-Gaussian process noise in the presence of a certain class of measurement model uncertainty, Amin and Thia [31] proposed an EM-PF algorithm that casts the problem in a joint state estimation and model parameter identification framework. The E-step is implemented by a particle filter that is initialized by a Monte Carlo Markov chain algorithm, while the M-step estimates the parameters of the mixture of Gaussian, which is used to approximate the nonlinear observation equation. The EM-PF is used to solve a highly nonlinear bearing-only tracking problem and sensor registration problem in a multi-sensor fusion case.

Stochastic dynamic systems with unknown inputs $\left(\theta \triangleq \{a, b\}\right)$: Consider the joint estimation and identification problem of a class of discrete-time stochastic systems with unknown inputs in both the plant and sensors, Lan and Liang [5] proposed an EM based iterative optimization method. The system state is estimated by using the KS in E-step, and the analytical solution of unknown inputs is obtained in the M-step by maximizing the $\mathcal{Q}$-function. The proposed method is used to solve the maneuvering target tracking in electronic counter environments.

Maneuvering target tracking in clutter $\left(\theta \triangleq \{\beta, \gamma\}\right)$: The problem of maneuvering target tracking in a clutter free environment with unity detection probability is addressed in [32], where the EM algorithm is used to compute hard maneuver command assignments, the target state estimate is computed based on the measurements and the hard input control sequence estimate. Andrew and Vikram [33] extended this work to the case of maneuvering target tracking in clutter, where both uncertain origin of the measurements and the maneuvering command are uncertainty. The proposed scheme combines a hidden Markov model smoother (HMMs) and a Kalman smoother (KS), whereby the E-step computes the joint posterior probability density of association and control input by HMMs, and the M-step obtains the MAP sequence of target state based on a modified state-space model via the KS.

Multiple target tracking $\left(\theta \triangleq \{\gamma\}\right)$: The well-known probabilistic multiple-hypothesis tracker (PMHT) [34], [35], [36], [37] is a preferred multitarget target tracking and association algorithm derived from the application of the EM algorithm. A fundamental difference between the PMHT and other standard tracking approaches is that PMHT assumes that the assignment indices for each measurement are independent random variable. The PMHT algorithm forms an estimate of the unknown model states based on a collection of state observations with uncertain origin, and estimates the model states by maximizing the conditional expectation of the log likelihood with respect to the model to measurement assignments.

Multiple detection tracking systems in clutter $\left(\theta \triangleq \{\gamma, \delta\}\right)$: In most tracking systems which referred as single-detection systems are based on the common assumption that in every scan there is at most one measurement from each interested target. In fact, multiple measurements may be simultaneously generated by the same target via different measurement modes, and the association hypothesis among targets, measurements and measurement modes are unknown. Such systems are referred as multiple detection systems, including over-the-horizon radar (OTHR) [38], [39] etc. The difficulty in tracking a target for multiple detection systems arises from the uncertain origin of the measurements and the uncertainty of measurement mode. Pulford and Logothetis presented the expectation maximization data association (EMDA) [40] for fixed-interval Kalman smoothing conditioned on the MAP estimation of target-measurement-mode triple association sequence in the EM framework. However, the EMDA is suitable for off-line, batch computation. It is suggested that an approximate method is needed to decrease the computation cost. Lan and Liang proposed the joint multipath data association and state estimation (JMAE) [6] algorithm also based on the EM algorithm to obtain the approximate solutions, which carried out the identification of ionospheric mode and measurement association in the E-step, where the pseudo-measurement and a posterior probability of each propagation mode are derived. Meanwhile, the JMAE updates the state estimation in the M-Step, where path-conditional state estimates and multipath state fusion are implemented. Furthermore, Lan and Liang proposed a distributed EM based on consensus filtering (D-CEM) [7] to solve the approximate problem posed by JMAE. The DCEM regards the multiple detection systems as sensor networks where each sensor node corresponds to a measurement mode. In the E-step, each mode independently calculates local state estimate by using its associated measurement. A consensus filter is used to exchange its localized estimate with its neighbors and then fuse them. In the M-step, each mode uses the estimated global state to find the local optimal measurement in the nearest neighbor sense.

2) Multiple Sensors Multiple Target Tracking: Multisensor multitarget tracking systems introduce a major complication that is absent from single-sensor, single-target problems. In the traditional centralized setting for both measurement level fusion and track level fusion, multitarget tracking is difficult [41]. There is a combinatorial explosion in the space of possible multiple target trajectories due to the uncertainty of the track-track association α and track-measurement association γ at each timestep, i.e., $\theta \triangleq \{\alpha, \gamma\}$. Tracking is also complicated by the fact that, for many sensing modalities, targets in close proximity tend to interfere with sensing one another. Compensating for this problem often requires sensing in a higher-dimensional joint space, again increasing computational complexity. Due to the above challenges, multitarget tracking is still an open problem in centralized systems.

Consider the multisensor multitarget tracking in a centralized measurement level fusion, Molnar and Modestino proposed an iterative procedure for time-recursive multitarget/multisensor tracking based on EM algorithm [42]. More specifically, target updates at each time use an EM based approach that calculates the MAP estimate of the target states, under the assumption of appropriate motion models. The approach uses a Markov random field (MRF) model of the associations between observations and targets and allows for estimation of joint association probabilities without explicit enumeration. The advantage of this EM-based approach is that it provides a computationally efficient means for approaching the performance offered by theoretical optimum approaches that use explicit enumeration of the joint association probabilities. Frenkel and Feder [43] investigated the application of EM algorithm to the classical problem of multitarget tracking for a known number of targets, three different schemes, including EM-Newton, EM-Kalman and EM-HMM were proposed based on different optimization criteria. The EM-Newton was a second approximation of the recursive EM algorithm in maximum likelihood criteria, and the EM-Kalman algorithm used EM localization with Kalman tracking in MSE criteria, while the EM-HMM algorithm that used a discrete model for the parameters and a Viterbi search for the optimum parameter sequence in the MAP criteria. Sensor registration and data association are two fundamental problems in multisensor multitarget tracking systems, they actually affect each other. Li and Chen [44] presented a joint sensor association, registration and fusion algorithm based on the EM framework for multisensor multitarget tracking, i.e., $\theta \triangleq \{\alpha, \gamma, b\}$. More precisely, the target state is regarded as the missing data and estimated in the E-step via KS, and the optimal target-to-measurement association is chosen via the multi-dimensional assignment, thus the registration parameters is obtained in M-Step.

3) Distributed Target tracking in Sensor Networks: Distributed estimation and tracking is one of the most fundamental collaborative information processing problems in sensor networks [46]. Decentralized Kalman filtering involves state estimation using a set of local Kalman filters that communicate with all other nodes [47], [48], [49]. Control-theoretic consensus algorithms have proven to be effective tools for performing network-wide distributed computational tasks such as computing aggregate quantities and functions over networks [50]. Naturally, the distributed EM algorithm is developed for the distributed joint estimation and identification in sensor networks. Maybe the distributed EM algorithm was firstly proposed by Nowak for the joint density estimation and clustering in sensor networks [51], which is viewed as an application and adaptation of the incremental EM algorithm. It views the E-step and M-step both as the maximization of an "energy function" over distribution and parameters. Based on the partially increasing, the distributed EM algorithm constructs a path through the network, which passes through all nodes. The incremental based distributed EM algorithm uses the partially accumulated global sufficient statistics to estimate the parameters in each node. The convergence of the distributed EM algorithm is also investigated, under mild conditions, the distributed EM converges to a stationary point of the log likelihood function with a (at least) linear rate, potentially converging more rapidly than standard EM. This makes the distributed EM attractive for sensor network applications. However, the incremental based distributed EM algorithm is slow when the network becomes complex and demands a full network access in each updating step. Therefore, Gu [52] proposed a consensus based distributed EM algorithm to handle this difficulty through estimating the global sufficient statistics using local information and neighbors' local information. In the E-step, each sensor node independently calculates local sufficient statistics by using local observations. A consensus filter is used to diffuse local sufficient statistics to neighbors and estimate global sufficient statistics in each node. By using this consensus filter, each node can gradually diffuse its local information over the entire network and asymptotically the estimate of global sufficient statistics is obtained. In the M-step of this algorithm, each sensor node uses the estimated global sufficient statistics to update model parameters of the Gaussian mixtures. The convergence is proved and it is a stochastic approximation to the standard EM with probability one.

Consider the problem of distributed joint state estimation and identification for a class of stochastic systems with unknown inputs, Lan and Bishop [8] proposed a distributed EM algorithm to estimate the local state at each sensor node by using the local observations in the E-step, and three different consensus strategies developed to diffuse the local state estimation to each sensor's neighbours and to derive the global state at each node. In the M-step, each sensor identifies the local unknown inputs by using the global state estimate. Pereira [53] proposed a diffusion-based distributed EM algorithm (DB-DEM) for distributed estimation in unreliable sensor networks, where sensors may be subject to data failures and report only noise. The propagation of information across the network is embedded in the iterative update of the parameters, where a faster term for information diffusion is combined with a slower terms is controlled by assigning them appropriate time-varying step-size sequence.

IV. Conclusion

In this paper, an overview of the approaches for joint tracking using EM algorithm is presented. The discussion focus on the coupling relationship between estimation and identification problem of the target tracking. The interdependence between estimation and identification add additional difficulty for the solution of a target tracking problem. The joint tracking framework, which simultaneously takes both estimation error and identification risk of target state into account, can effectively solve this kind of target tracking problem using the EM algorithm. From another view point, this target tracking problem can be treated as the state estimation using *incomplete data*, that is, some unknown parameters needs to be identified before we can estimate the underlying target state. The EM algorithm is shown being an effective method to solve a range of joint estimation and identification problems in the paper.

Acknowledgments

This work was supported by the National Science Foundation of China (Grant61135001, 61374159, and 61374023), and the Fundamental Research Funds for the Central Universities (Grant G2015KY0801).

References

[1] Y. Bar-Shalom, X.R., Li, and T. Kirubarajan, *Estimation with Application to Tracking and Navigation: Theory Algorithms and Software*, New York: Wiley, 2001.

[2] D. Hall, and L. James, *Handbook of Multisensor Data Fusion*, CRC Press, 2001.

[3] X. Li, "Optimal Bayes joint decision and estimation," In *processing 2007 International Conference on Information Fusion*, Quebec, Canada, 2007.

[4] X. Li, "Joint tracking and classification based on bayes joint decision and estimation," In *Processings 2007 International Conference on Information Fusion*, Quebec, Canada, 2007.

[5] H. Lan, Y. Liang, F. Yang, Z. Wang, and Q. Pan, "Joint estimation and identification for stochastic systems with unknown inputs," *IET Control Theory and Applications*, 2013, 7(10), pp. 1377-1386.

[6] H. Lan, Y. Liang, F. Yang, Q. Pan and C. Guan, "An EM algorithm for multipath state estimation in OTHR target tracking," *IEEE Transactions on Signal Processing*, 2014, 62(11), pp. 2814-2826.

[7] H. Lan, Y. Liang, Z. Wang, F. Yang, and Q. Pan, " A distributed expectation-maximization algorithm for OTHR multipath target tracking," *Proceedings of the 17th International Conference on Information Fusion*, Spain, 2014, pp. 1-8.

[8] H. Lan, A. Bishop, and Q. Pan, "Distributed joint estimation and identification for sensor networks with unknown inputs," *The 2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, Singapore, 2014, pp. 1-6.

[9] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data," *Royal Statistical Society*, 1977, 39(1), pp. 1-38.

[10] D. Tzikas, A. Likas, and N. Galatsanos, "The variational approximation for bayesian inference," *IEEE Signal Processing Magazine*, 2008, 25(6), pp. 131-146.

[11] E. Cosme, J. Verron, P. Brasseur, J. Blum, and D. Auroux, "Smoothing problems in a bayesian framework and their linear Gaussian solutions," *American Meteorological Society*, 2012, 140, pp. 683-695.

[12] C. Bishop, *Pattern Recognition and Machine Learning*, New Work, NY, USA: Spring, 2006.

[13] M. Gupta, Y. Chen, "Theory and use of the EM algorithm," *Foundations and Trends in Signal Processing*, 2011, 4(3), pp. 223-296.

[14] G. McLachlan, and T. Krishnan, *The EM and Its Extension*, John Wiley & Sons, New York, 2008.

[15] C. Wu, "On the convergence properties of the EM algorithm," *The Annals of Statistics*, 1983, 11(1), pp. 95-103.

[16] A. Logothetis, and K. Krishnamurthy, "Expectation maximization algorithms for MAP estimation of jump markov linear systems," *IEEE Transactions on Signal Processing*, 1999, 47(8), pp. 2139-2156

[17] X. Meng, and D. Rubin, "On the global and componentwise rates of convergence of the EM agorithm," *Linear Algebra and its Applications*, 1994, 199, pp. 413-425.

[18] V. Melnykov, and I. Melnykov, "Initializing the EM algorithm in Gaussian mixture models with an unknown number of components," *Computational Statistics and Data Analysis*, 2012, 56, pp. 1381-1395.

[19] C. Fraley, "Algorithms for model-based Gaussian hierarchical clustering," *SIAM joural on Scientific Computing*, 1998, 20, pp. 270-281.

[20] C. Biernacki, G. Celeux, and G. Govaert, "Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models," *Computational Statistics and Data Analysis*, 2003, 41(3), pp. 561-575.

[21] R. Maitra, "Initializing partition-optimization algorithms," *IEEE Transactions on Computational Biology and Bioinformatics*, 2009, 6(1), pp. 144-157.

[22] A. Roche, "EM algorithm and variants: An informal tutorial," *IEEE Transactions on Computational Biology and Bioinformatics*, 2009, 6(1), pp. 144-157.

[23] J. Fessler, and A. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Transactions on Signal Processing*, 1994, 42(10), pp. 2664-2677.

[24] X. Meng, and D. Rubin, "Maximum likelihood estimation via the ecm algorithm: A general framework," *Biometrika*, 1993, 8(2), pp. 267-278.

[25] C. Liu, and D. Rubin, "The ecme algorithm: A simple extension of em and ecm with faster monotone convergence," *Biometrika*, 1993, 81(4), pp. 633-648.

[26] G. Wei, and M. Tanner, "A monte carlo implementation of the em algorithm and the poor mans data augmentation algorithms," *Journal of the American Statistical Association*, 1990, 85(411), pp. 699-704.

[27] C. Liu, D. Rubin, and Y. Wu, "Parameter expansion to accelerate em: The px-em algorithm," *Biometrika*, 1998, 85(4), pp. 755-770.

[28] M. Jamshidian, and R. I. Jennrich, " Conjugate gradient acceleration of the EM algorithm," *Journal of the American Statistical Association*, 1993, 88, pp. 221-228.

[29] R. Shumway, and D. Stoffer, " An approach to time series smoothing and forecasting using the EM algorithm," *Journal of the Time Series Analysis*, 1982, 3(4), pp. 253-264.

[30] Z. Ghahramani, and G. Hinton, " Parameter estimation for linear dynamical systems," *Technical Report CRG-TR-96-2*, 1996, University of Totronto, Dept. of Computer Science.

[31] A. Zia, T. Kirubarajan, J. Reilly, D. Yee, K. Punithakumar, and S. Shirani, "An EM algorithm for nonlinear state estimation with model uncertainties," *IEEE Transactions on Signal Processing*, 2008, 56(3), pp. 921-936.

[32] B. Scala, and G. Pulford, " Manoeuvring target tracking using the expectation-maximisation algorithm," *The Fourth International Conference on Control, Automation, Robotics and Vision*, Singapore 1996, pp. 2340-2344.

[33] A. Logothetis, V. Krishnamurthy, and J. Holst, "A bayesian EM algorithm for optimal tracking of a maneuvering target in clutter," *Signal Processing*, 2002, 82, pp. 473-590.

[34] R. Streit, and T. Luginbuhl, "Maximum likelihood method for probabilistic multi-hypothesis tracking," In *Processings of the SPIE*, 1994, 2235, pp. 394-405.

[35] P. Willett, Y. Ruan, and R. Streit, "PMHT: Problems and some solutions," *IEEE Transactions on Aerospace and Electronic Systems*, 2002, 38(3), pp. 738C754.

[36] R. Streit, *Possion Point Processes: Imaging, Tracking, and Sensing*. New Work, NY, USA: Spring, 2010.

[37] M. Wieneke, and W. Koch, "A PMHT approach for extended objects and object groups," *IEEE Transactions on Aerospace and Electronic Systems*, 2012, 48(3), pp. 2349-2370.

[38] G. Pulford, and R. Evans, "A multipath data association tracker for over-the-horizon radar," *IEEE Transactions on Aerospace and Electronic Systems*, 1998, 34(4), pp. 1165-1182.

[39] G. Pulford, "OTHR multipath tracking with uncertain coordinate registration," *IEEE Transactions on Aerospace and Electronic Systems*, 2004, 40(1), pp. 38-56.

[40] G. Pulford, and A. Logothetis, "An expectation-maximisation tracker for multiple observations of a single target in clutter," In *Proceedings of the 36th Conference on Decision and Control*, San Diego, California, 1997, pp. 4997-5003.

[41] D. Smith, and S. Singh, "Approaches to multisensor data fusion in target tracking: A survey," *IEEE Transactions on Knowledge and Data Engineering*, 2006, 18(12), pp. 1696-1710.

[42] K. Molnar, and J. Modestino, "Application of the em algorithm for the multitarget/multisensor tracking problem," *IEEE Transactions on Signal Processing*, 1998, 46(1), pp. 115-129.

[43] L. Frenkel, and M. Feder, "Recursive expectation-maximization (EM) algorithms for time-varying parameters with applications to multiple target tracking," *IEEE Transactions on Signal Processing*, 1998, 47(2), pp. 306-320.

[44] Z. Li, S. Chen, H. Leung, and E. Bosse, "Joint data association, registration, and fusion using EM-KF," *IEEE Transaction on Aerospace and Electronic Systems*, 2010, 46(2), pp. 496-507.

[45] J. Liu, M. Chu, and J. Reich, "Multitarget tracking in distributed sensor networks," *IEEE Signal Processing Magazine*, 2007, 24(3), pp. 36-46.

[46] I. Akyildiz, W. Su, and Y. Sankarasubramaniam, "A survey on sensor networks," *IEEE Communications magazine*, 2002, 40(8) pp. 102-114.

[47] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," *IEEE Conference on 46th Decision and Control*, 2007, pp. 5492-5498.

[48] R. Olfati-Saber, "Distributed Kalman filtering with embedded consensus filters," *IEEE Conference on 44th Decision and Control*, 2005, pp. 8179-8184.

[49] R. Olfati-Saber, and J. Shamma, "Consensus filters for sensor networks and distributed sensor fusion," *IEEE Conference on 44th European Control COnference*, 2005, pp. 6698-6703.

[50] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, 2007, 95(1),

[51] R. Nowak, "Distributed EM algorithms for density estimation and clustering in sensor networks," *IEEE Transactions on Signal Processing*, 2003, 51(8), pp. 2245-2253.

[52] D. Gu, "Distributed EM algorithm for Gaussian mixtures in sensor networks," *IEEE Transactions on Neural Networks*, 2008, 19(7), pp. 1154-1166.

[53] S. Pereira, R. Valcarce, and A. Zamora, "A diffusion-based EM algorithm for distributed estimation in unreliable sensor networks," *IEEE Signal Processing Letters*, 2013, 20(6), pp. 595-598.