Multi-target tracking and Bayesian estimation

Author: Qi Chen,

A B S T R A C T

This report gives a comprehensive and fundamental introduction to multi-target tracking (MTT) with an emphasis on Bayesian estimation. Target tracking estimates the target status, especially position information, with measurements from the observation station (possibly sensor or camera). MTT is overlapped and commonly divided into 1) dynamic motion modelling, 2) data association, 3) state estimation, 4) track maintenance and 5) multi-sensor data fusion. This report analyzes the role of each part except (4). The state estimation is given separately and detailly from the view of Bayesian analysis with the Kalman filter as an example. In addition, some simulations of simple MTT problems are given at the end to make the abstract clearer.

I declare that this work is my own and I acknowledge the contribution of others where appropriate.

I. Introduction

Multi-target tracking (MTT) has been extensively researched since the 1970s, derived from military applications, and is widely implemented in many fields, such as transportation and aerospace [1]. MTT systems estimate the target status, especially position information, with measurements from the observation station (possibly sensor or camera). When a stochastic system, i.e., is subjected to various noises, a simplest (single target) tracking system mainly considers state estimation. Suppose the worst-case scenario, if no prior knowledge is available, the collected data from the observations will output as the result. Otherwise, when priori knowledge, e.g., dynamic motion model and statistical information, is available, the result will be more accurate. Ideally, the more information considered, the more accurate result will be obtained. In addition, the measurement may not be always collected from the target, due to false alarms, clutters, etc. It makes the tracking problems different from a state estimation problem. Beyond single target tracking, MTT is more complex, which needs data association to identify multiple targets before state estimation. Moreover, modern tracking systems usually involve more than one observation station, so multi-sensor data fusion plays an important role in today's MTT problems as well. As seen above, MTT can be systematically divided into 1) dynamic motion modelling, 2) data association, 3) state estimation, 4) track maintenance and 5) multi-sensor data fusion.

The rest of this report is organized as follows: section II gives a brief and systematic introduction to the MTT system. Section III introduces the state estimation from a Bayesian view. Section IV gives some simulations of some simple MTT problems by MATLAB. Since the simulation result only serves to make the abstract clearer, the MATLAB codes won't be given in this report.

II. Multi-target tracking system

At an early stage, MTT systems were only designed for specific applications, i.e., no general MTT concept had been created. Some early approaches of the MTT problems are well summarized in the paper [2]. In 1986, Researchers from Massachusetts Institute of Technology and Tokyo Institute of Technology firstly gave a general framework of MTT without a priori identification [3]. The paper [3] considers the data association as a core challenge of the MTT problems. Fig 2.1 gives a general framework of the MTT problems. Different from a simple state estimation problem, the MTT system meets various problems, e.g., target identification and track maintenance. The functional element of track maintenance in Fig 2.1 is important in practice, because the number of targets is usually unknown, and possibly generating false alarms and clutters. However, it will complicate the issue if considering the track maintenance in this report. Due to space limitations, the track maintenance and detection won't be included in this report. For more details, readers should refer to [4][5].

A. Dynamic motion modelling

As told in section I, the dynamic motion model of the target provides evidence to state estimation. The efficiency of state estimation partly depends on the accuracy of the modelling of the motion process. In the survey [6], the recent achievements of dynamic models
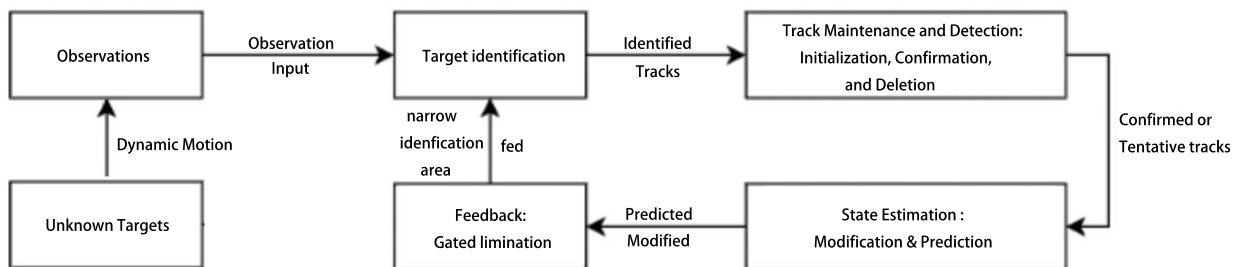

Figure 2.1 General multi-target tracking system

are well documented. However, for many cases of maneuvering target tracking, the accurate dynamic motion model is not available because the uncertainty of the target motion can't be well quantified. Moreover, a highly accurate model may be not applicable sometimes, because a more informative model brings more computation complexity. For real-time tracking, the computation complexity is extremely crucial. Therefore, a simple dynamic motion model is also valuable. In the following, two most used mathematical models of CV (constant velocity) motion and CA (constant acceleration) motion will be given as an example of building a space-state model. For more complex models, readers should refer to [7], which gives a detailed modelling process of coordinated turn (CT) motion and augmented coordinated turn (ACT) model. When a stochastic system, the most used recursive model is in the following form:

$$X(k+1) = \Phi X(k) + GW(k) \tag{1}$$

where X is the target state, Φ is a time-varying vector of X, W is the noise of motion, G is a time-varying vector of W. The equation is given as a discrete-time model, and the sampling interval is T. Assume that W is an Additive Gaussian White Noise (AWGN), then the following equation is satisfied

$$E[W(k)] = 0, E\left[W(k)W^T(j)\right] = Q(k)\delta_{kj} \tag{2}$$

where Q is the variance of W. The variable vectors take different values in CV and CA models. In CV model, the vectors are given as follows:

$$X = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}, \Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, G = \begin{bmatrix} T^2/2 \\ T \end{bmatrix} \tag{3}$$

where x is the target position, $\dot{x}$ is the target velocity, and T is the sampling interval. In CA model, the vectors are given as follows:

$$X = \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix}, \Phi = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix}, G = \begin{bmatrix} T^3/6 \\ T^2/2 \\ T \end{bmatrix} \tag{4}$$

where $\ddot{x}$ is the target acceleration. CV and CA dynamic motion models are the most used models, which are both linear and of significant reference value. While the uncertainty of maneuverings is not considered, its computation complexity is well-deserved. In section IV, some simulations based on CV motion targets are given. By the simulation, we can reasonably assume that a random motion target in a linear Gaussian system can be described with a CV model, whose variance of the process noise is extremely large.

B. Target identification

MTT is significantly different from a simple state estimation problem partly because of target identification. Target identification is one of the most challenging problems for multi-target tracking. Before state estimation, the observations need to be associated with the correct track, i.e., be identified, if more than one target is tracked. In sensor (possibly radar) tracking, Target identification is a pattern classification problem with a particular emphasis on position characteristic. By comparison, in image processing, the characteristics of the targets are more varied, therefore, the MTT problems are different in sensor tracking and image tracking. The paper [8] is a well-documented literature review of MTT problems in image tracking. In this report, we focus on specific situations in sensor tracking where a target can be treated as a "point target".

In 1967, T. M. Cover and P. E. Hart gave the Nearest Neighbor (NN) classification method in the paper [9] and proved its efficiency in a mathematical way. The NN algorithm is a decision-driving method. By enumerating all possible data, it finds the nearest detection and regards it as the correct unclassified target. The advantage of the NN algorithm is its low computation complexity. A Binary searching process is done recursively in the NN searching process. However, this algorithm brings insufficient information to the result that only one detection makes contributions to the result. In addition, the NN classification algorithm performs well only when using a sparse model. When the differentiation of the targets is not obvious, then other approaches are desired. In 1975, Bar-Shalom, Y.

and Tse, E. gave a probability data association (PDA) algorithm which enables all detections to make weighted contributions to the result based on their probability of association [10]. The author of [10] gave a further approach in [11], which was named as a joint probability data association (JPDA) algorithm. Due to limited space, the principles of these algorithms are not included in this report, readers should refer to [12] for more details. Instead, a simulation result of the NN algorithm is given in section IV, which offers a visual process of a decision-driving classification algorithm.

C. State Estimation

Target identification and State estimation are two core challenges in the MTT problems. After accurate target identification, the MTT problems become a collection of multiple state estimation problems if not considering false alarms and clutters. Besides (1), the recursive equation of dynamic motion, the observation process is also concerned in state estimation. A most used observation equation is:

$$Z(k) = HX(k) + GV(k) \tag{5}$$

where Z is the measurement, H is a time-varying vector for Z, V is the noise of measurement, G is a time-varying vector for V. Assume that V and W are independent and identically distributed. As seen in (1) and (5), many factors (possibly ambient noise and sensor accuracy) make the result inaccurate. State estimation proposes to modify the result with given evidence, i.e., priori knowledge. In section III, some details of state estimation will be given. This part only proposes to give the readers a complete view of the MTT problems and bring the measurement equation (5) into consideration.

D. Multi-sensor data fusion

A principled definition of information fusion is given in paper [13], as "Information fusion is the study of efficient methods for automatically or semi-automatically transforming information from different sources and different points in time into a representation that provides effective support for human

or automated decision making"[1]. For MMT problems, multi-sensor data fusion is highly like data association because the observation type is uniform for each observation station. Since data association has been introduced in subsection B, Multi-sensor data fusion won't be expanding in this report. Instead, a simple simulation will be given in section IV for a clearer understanding. For more details, readers should refer to paper [13] [14]. The paper [13] gives some approaches for complex data fusion problems from the view of Artificial Intelligence, e.g., Convolutional Neural Network (CNN), which is expanding in [14].

III. Bayesian estimation and Kalman filter

As a profound state estimation approach, Kalman filter (KF) follows the framework for Bayesian analysis, which "revolutionized the field of estimation …(and) open up many new theoretical and practical possibilities"[2] [15][16]. This section will introduce state estimation from a Bayesian view. It is mentioned that KF only has efficient performances for Linear Gaussian systems. For nonlinear non-Gaussian systems, the problem is more complex. Some well-performed methods derived from KF are given in paper [16][17].

A. Linear Gaussian model

It is well known that the sum of an infinite number of uniform distributed random variables is subjected to a Gaussian distribution. Ideally, the noise is regarded as a random variable, and the sum of noises is subjected to a Gaussian distribution. The advantage of Gaussian distribution is that it describes the probability distribution with only two parameters (mean and variance), which highly decreases the computation complexity. When a linear system, if the priori distribution is a Gaussian distribution, the posteriori distribution will be a Gaussian distribution as well. Therefore, the advantage of the Linear Gaussian system is that both the priori and posteriori distribution of the MTT systems can be described with only mean and variance.

---

[1]   *I have discussed and received written permission from Dr Yoshi Gotoh to include this quotation in this document.

[2]   *I have discussed and received written permission from Dr Yoshi Gotoh to include this quotation in this document.

B. A Bayesian view of State Estimation

The idea of Bayesian analysis was firstly proposed by Rev. Thomas Bayes in 1763 [18]. It answers a philosophical question that when new evidence is given, how to estimate the true with a balance of the former and new evidence. A comment is given in paper [19], that "Bayesian approach offers a unified and intuitive viewpoint particularly adaptable to handling modern-day control problems"[3]. Learning state estimation from a Bayesian view, therefore, lies a foundation of learning other state estimation approaches. When a real time MTT system consists of (1) and (5), essentially, state estimation looks for the probability density function (pdf) of $p(x_k \mid z_k, x_{k-1})$. According to Bayes' rule,

$$p(x_k \mid z_k, x_{k-1}) = \frac{p(z_k|x_k,x_{k-1})p(x_k|x_{k-1})}{p(z_k|x_{k-1})} \qquad (6)$$

The observation process is a Markov process, therefore,

$$p(z_k \mid x_{k-1}) = p(z_k) \qquad (7)$$

And $\frac{1}{p(z_k)}$ is a constant, indicated as η, then the equation (6) becomes

$$p(x_k \mid z_k, x_{k-1}) = \eta p(z_k \mid x_k)p(x_k \mid x_{k-1}) \qquad (8)$$

According to (8), the state estimation becomes a hidden Markov process. Further assume that it's a linear Gaussian system whose mean and variance is available, then $p(z_k \mid x_k)$ is involved in (1) and $p(x_k \mid x_{k-1})$ is involved in (5). The probability density function (pdf) of $p(x_k \mid z_k, x_{k-1})$ can be described with $p(z_k \mid x_k)$ and $p(x_k \mid x_{k-1})$, as seen in (8), which is one of the most used equations for both Bayesian filter and Kalman filter. For more mathematical process, readers should refer to paper [16]. The paper [16] gives a complete and strict mathematical process from Bayesian estimation to Kalman filter.

C. Kalman filter

A strictly mathematical derivation of the Kalman filter is given in paper [20]. Due to space limitation, this

---

section only gives a "prediction-update" procedure of the Kalman filter briefly, recursively processing $p(x_k \mid z_k, x_{k-1})$ at each time k. For linear Gaussian systems, the priori and posteriori pdf can be both described with only mean and variance. The procedure can be ordered to 1) state prediction, 2) observation prediction, 3) covariance prediction and Kalman gain calculation, 4) Update the target state, and 5) update the covariance. A simple approach to understand the Kalman filter is included in paper [21]. However, it's not mentioned in other documents that even without the specific gain (Kalman gain), the result will still be modified. Because if a production of two Gaussian distributions, e.g., $\mathcal{N}(\mu_1, \sigma_1^2)$, $\mathcal{N}(\mu_2, \sigma_2^2)$, is still a Gaussian distribution, e.g., $\mathcal{N}(\mu, \sigma^2)$, then the following equation is satisfied,

$$\frac{1}{\sigma^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \qquad (9)$$

where σ^2 is always smaller than σ_1^2 and σ_2^2. Therefore, the posterior covariance is always smaller than the prior covariance and the estimated state will become more and more accurate as time goes on, when using the Kalman filter.

IV. Some simulations of the MTT problems

Since learning MTT problems with only theory is possibly abstract to the beginners, this section gives a visual view of some MTT problems by MATLAB. As seen in section I, MTT problems can be divided into 5 parts. This section involves a simple simulation for each part of MTT problems, except track maintenance.

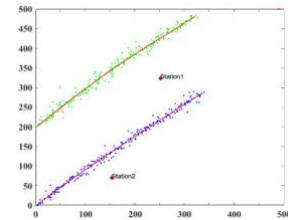## A. Multi-station and multi-target model

Figure 4.1 A Multi-observation and multi-target system

Two random observation stations are built to observe two CV targets from two random locations. The dynamic motion model follows the rules (1) and (3), when the

---

measurement model follows the rule (5). For each time k, each station will generate two observations for each target. Two approximately straight lines are observed in the figure 4.1, surrounding with the observations, that means the variance of the process noise is very small and the variance of the measurement noise is relatively bigger.
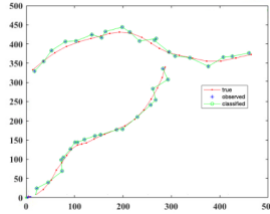
B.　NN classification used in a MTT problem

Figure 4.2 NN classification used in a MTT problem

In Fig 4.2, two CV targets, who are subjected to a significant process noise, are observed. The observations are classified well with the NN classification algorithm. Interestingly, these two trajectories are not like a line because the variance of the process noise is very large. From this view, we can have a reasonable scenario that a random motion target can be described as a CV model whose variance of the process noise is extremely large. In this simulation, the NN classification algorithm works efficiently. However, for non-sparse tracking problems, some advanced approaches are necessary.
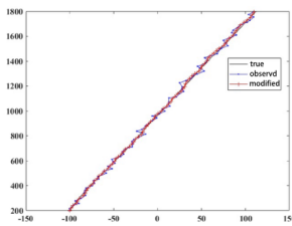
C.　Kalman filter

Figure 4.3 Kalman filter

A one-station and two-target model is built in this simulation. The true track is an approximately straight line, generated by a CV motion target. It is seen n Fig 4.3 that the observed track is crooked and inaccurate. The Kalman filter modifies the track significantly. In addition, the priori knowledge of the original state is not

available, so the first "predication" may be significantly different from the true, because it is a manmade value. The equation (9) proves that even the first "predication" is inaccurate, Kalman filter is still effective.
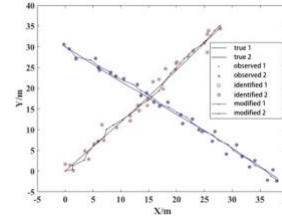
D.　Multi-sensor data fusion
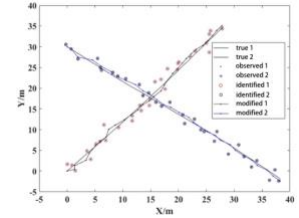

Figure 4.4 one-station track.　　　Figure 4.5 two-station track

In the last simulation, both the variance of the process noise and the variance of the measurement noise is set at a low value. However, in practice, the uncertainty caused by the noise seriously affects the state estimation. Therefore, multi-sensor data fusion is necessary in MTT problems, which reduces the uncertainty. Figure 4.4 shows the track without data fusion. Figure 4.5 shows a two-station track, integrated by the observations from two stations. In this simulation, the data is simply integrated by taking the mean value of two observations for each target. The two-station track is significantly better than the one-station track in this simulation.

IV. Conclusion

This report gives a comprehensive and brief introduction of MTT with a particular emphasis on state estimation. Linear Gaussian model is a special case of state estimation and Kalman filter is the optimal estimation when a linear Gaussian model. Some "mathematically friendly" explanations of Linear Gaussian model and Kalman filter are given in this report. In addition, the last section discusses some simulations of simple MTT problems, which make the abstract clearer.

References

[1]　Hunde, A.N. (2021) 'Tracking and fusion of multiple detections for multi-target multi-sensor tracking applications in urban traffic', *SAE International Journal of Connected and Automated Vehicles*, 4(1), pp. 49–64.

[2] Bar-Shalom, Y. (1978) 'Tracking methods in a multitarget environment', *IEEE transactions on automatic control*, 23(4), pp. 618–626.

[3] Mori, Shozo, C., Chee-Yee, Tse, Edison, Wishner, Richard P (1986) 'Tracking and classifying multiple targets without a priori identification', *IEEE transactions on automatic control*, AC-31(5), pp. 401–409.

[4] Vanfossen, B.E., Jones, J.D. and Spade, J.Z. (1987) 'Curriculum tracking and status maintenance', *Sociology of education*, 60(2), pp. 104–122.

[5] Nagarajan, V., Sharma, R.. and Chidambara, M.. (1984) 'An algorithm for tracking a maneuvering target in clutter', *IEEE transactions on aerospace and electronic systems*, AES-20(5), pp. 560–573.

[6] Rong Li, X. and Jilkov, V.. (2003) 'Survey of maneuvering target tracking. Part I. Dynamic models', *IEEE transactions on aerospace and electronic systems*, 39(4), pp. 1333–1364.

[7] Yuan, X., Lian, F. and Han, C. (2014) 'Models and algorithms for tracking target with coordinated turn motion', *Mathematical problems in engineering*, 2014, pp. 1–10.

[8] Luo, W., Xing, J., Milan, Anton, Zhang, X., Liu, W., Kim, Tae-Kyu (2021) 'Multiple object tracking: A literature review', *Artificial intelligence*, 293, p. 103448.

[9] Cover, T. and Hart, P. (1967) 'Nearest neighbor pattern classification', *IEEE transactions on information theory*, 13(1), pp. 21–27.

[10] Bar-Shalom, Y. and Tse, E. (1975) 'Tracking in a cluttered environment with probabilistic data association', *Automatica* (Oxford), 11(5), pp. 451–460.

[11] Fortmann, T., Bar-Shalom, Y. and Scheffe, M. (1983) 'Sonar tracking of multiple targets using joint probabilistic data association', *IEEE journal of oceanic engineering*, 8(3), pp. 173–184.

[12] Roecker, J.. (1994) 'A class of near optimal JPDA algorithms', *IEEE transactions on aerospace and electronic systems*, 30(2), pp. 504–510.

[13] Wan, E.. and Van Der Merwe, R. (2000) 'The unscented Kalman filter for nonlinear estimation', in Proceedings of the *IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium* (Cat. No.00EX373). IEEE, pp. 153–158.

[14] Kavzoglu, T. and Mather, P.M. (1999) 'Pruning artificial neural networks: An example using land cover classification of multi-sensor images', *International journal of remote sensing*, 20(14), pp. 2787–2803.

[15] 'Shirley Ann Jackson to Receive the National Medal of Science' (2016) *The Journal of blacks in higher education* [Preprint].

[16] Fang, H., Tian, N., Wang, Y., Zhou, M., Haile, M. (2017) 'Nonlinear bayesian estimation: from Kalman filtering to a broader horizon'.

[17] Wan, E.. and Van Der Merwe, R. (2000) 'The unscented Kalman filter for nonlinear estimation', in Proceedings of the *IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium* (Cat. No.00EX373). IEEE, pp. 153–158.

[18] Mr. Bayes and Mr. Price, (1763) 'An essay towards solving a problem in the doctrine of chances,'

Philosophical Transactions, vol. 53, pp. 370–418.

[19] Malakoff, D. (1999) 'Bayes offers a 'new' way to make sense of numbers', *Science* (American Association for the Advancement of Science), 286(5444), pp. 1460–1464.

[20] Meinhold, R.J. and Singpurwalla, N.D. (1983) 'Understanding the Kalman filter', *The American statistician*, 37(2), pp. 123–127.

[21] Arnold, T., Bertus, M.J. and Godbey, J. (2008) 'A simplified approach to understanding the Kalman filter technique', *The Engineering economist*, 53(2), pp. 140–155.