Chapter 2
Preliminaries

Modern radar systems include radar signal processors, data collectors, and radar data processors, besides traditional antennas, transmitters, receivers, and monitors. Signal processors are used mainly for target signal detection and for restraining the irrelevant signals of noise wave disturbances from earth or ocean, multi-path effects, atmospheric environment, universal noises, and man-made disturbances. The video output signals after signal processing are compared with certain detection threshold. If the signals are beyond the detection threshold, we "discover" targets. Then, we send the signals to data collectors, to measure the distance, angle, and radial velocity of the target, and for some radar, we can even measure the dispersion, polarization, and geometry of the target. The output of data collectors is the approximation of target observation. The observation from tradition radar is called observation data or measurement data. The observation data output by data collectors will be dealt with by data processors for all kinds of relevant processing. Radar data processing is used for computations such as association, tracking, filtering, smoothing, and prediction, after radar gets the location and motional parameters such as radial distance, radial velocity, and azimuth and pitch angle of targets. The processing of radar measurement data can effectively restrain the random errors introduced during the target measuring, precisely estimate the parameters relevant to location and motion of targets, and get stable tracks of targets.

From the perspective of information processing, if radar signal processing is the first-time processing, radar data processing is the second-time processing, and situation evaluation and threat judgment is the third-time processing. Some references called the interception determination, interception instruction computation, and interception means selection and kill probability computation the third-time processing. And some references in the early Soviet Union called data fusion the third-time processing. In the recent years, with the development of hardware, algorithms, and computer performance, there is qualitative leap in the information processing capability of radar, which enables the connectivity between "the physical, data, and cognitive layers" in radar. It is why the authors propose the group-target tracking algorithm.

W. Geng et al., *Group-target Tracking*, DOI 10.1007/978-981-10-1888-6_2

Target tracking is an interdisciplinary branch of science integrating information processing technology, cybernetics, and modern mathematics. There are many foundational theories, and we cannot elaborate in detail. This chapter will only expatiate on the content relevant to the research of this book.

2.1 Common Glossary

Next, we concisely introduce the basic glossary relevant to the target tracking. Please note that what we give in the following is glossary for multi-target, and the glossary for group-target will be given in the remainder chapters.

(1) Measurements

Measurements, also called measures, observations, or observation data, are the measurement data relevant to the state of targets and are contaminated by noises. Modern radar measurement is unnecessarily the original data output by receivers, but is the observation data output by data collectors after signal processing. The types of measurement include target radial distance, azimuth and pitch angle, radial velocity, the differences in time, frequency, and phase between the arrivals of signal to two sensors, and signal frequency and range transmit by targets.

Measurements come from possibly correct measurements of targets, false measurements of noise waves from earth/ocean, cloud, and rain, mistaken measurements of unexpected targets and baits in the target environment, and also false-alarm signal due to radar noises. Furthermore, missing detection of target signals may occur. So there are uncertainties in measurements.

(2) Tracks

The estimation of target state, according to the set of measurements from one target, goes through computation to give the track of the continuous motion of targets. A complete track also includes a track number, track quality, possible tracks, tentative tracks, validated tracks, terminated tracks, and response time of track initiation. The number of track can be regarded as the reference to all parameters related to the track. Track quality indicates the reliability of tracks, through which we can timely and correctly initiate tracks to build up new target archives and can also timely and correctly terminate tracks to remove redundant target' archives. Possible tracks consist of single measurement spots. Tentative tracks consist of two or more observation data and have comparatively low track quality. They can be target tracks or false tracks caused by random disturbances. Tracks, after initiation, will be transformed into tentative tracks or terminated tracks, so tentative tracks are also called temporary tracks. Validated tracks are those that have stable outputs or track quality of some standard. They are also called reliable tracks or stable tracks and usually are seen as real target tracks. Terminated tracks are those that have track quality below some fixed value or consist of isolated observation data. And the corresponding process is called track termination or track ending. The main task for track termination is to timely delete false tracks and keep real tracks. Response time of track initiation is the time period from the entering of target into radar power area to the construction of the track, usually taking radar scan time as its unit. Quick track initiation usually takes 3–4 radar scan cycles, while low track initiation usually takes 8–10 radar scan cycles.

(3) Data association

In a general sense, data association means the buildup the connection between the measurement data at one time and the measurement data or tracks at other time. It is the procedure to determine whether there measurement data come from the same target, or to pair correct observation data and tracks.

The objects of data association include track initiation, track maintenance, and track association. The association between measurement and measurement, or observation data and observation data, is the procedure of generating stable tracking according to the observation data obtained at some time. Track maintenance, also called track updates or track preservation, means the association between measurement and track, or observation data and track. Track association, also called track integration, means the association between track and track. Mathematically, data association can be divided into deterministic model and probability model. For the deterministic model, the source of measurement is deterministic, and the truth that it is unnecessarily correct is ignored. For the probability model, according to the occurrence probability of each true/false event, correct the approximation of target state through the probability. They two are corresponding to different data association algorithms, which will be discussed in detail in the following sections.

(4) Tracking

Data association and tracking are the two basic issues in radar data processing. There is only tracking issue for single targets, while there are both association and tracking issues for multi-targets.

Tracking means the processing of measurements derived from targets, to maintain the approximation of the current state of targets. The typical state of targets includes dimension information such as location, velocity, and acceleration; "signal characteristic" information such as strength, frequency, pulse width, and pulse recurrence frequency of transmit signals; and constants or slowly varying parameters such as coupling coefficient and transmission velocity of electromagnetic wave or sound waves.

Multi-target tracking means the concurrent processing of measurements from multi-targets, to maintain the approximation of the current state of multi-targets.

(5) Association gates

Association gates, also called association gates of tracking gates, take the predicted location of target under tracking as center and determine the area where the measurements of this target possibly occur. The dimension of the area is subject to the probability of correctly receiving target returns. Association gates have shapes such cuboid, cube, sphere, and ellipsoid. The principle of determining the shape and dimension of association gates is to make true measurements fall inside association gates most likely and in the same time make the number of irrelevant observation data as little as possible.

(6) Motional models

Motional models are the assumptions about the motional rules of targets. We can only obtain the state equations under some assumptions. Maneuver represents the uncertain changes in target motion, such as sudden acceleration and turning, which is one of the difficulties in target tracking research.

(7) State approximation

State approximation is the smoothing of the past motional state of targets, such as location, velocity, and acceleration, the filtering of the current motional state, and the prediction of the future motional state of targets.

2.2 Kalman Filtering

Literally, filtering is to filter waves. For example, in an automatic control system, the automatic control of system is realized through the feedback of system outputs. The system outputs usually include some disturbing signals and noises. When distilling some quantity of feedback as control quantity, there always exist random errors. Thus, we must appropriately filter in order to decrease control errors.

From the prospective of classical filtering, useful signals and noises are distributed within different frequency bands (sometimes overlaps possibly happen). Thus, we can use some classical filtering networks with certain frequency selection features to eliminate noises as many as possible, and keep useful signals with little aberration. However, the signals and noises we meet with sometimes are random, and their characteristics can be only described statistically. For example, the flying motion of missiles is random. The location and velocity are naturally random; meanwhile, measuring equipment has random errors as well. Therefore, we cannot adopt general classical filters to separate useful signals from measuring results, but can only compute the optimal estimation of useful signals through statistically estimation methods. From the statistical prospective, the closer the outputs of a filter are to practically useful signals, the better the performance of the filter is.

During the course of statistical filter development, early Wiener filter performs static processing of the time-independent statistical characteristics of signals. In this procedure, the statistical characteristics of useful signals and useless noises can be connected with their frequency-domain characteristics. Thus, Wiener filter and classical filters come down in one continuous line in conception. Wiener filter is applied widely during the Second World War. However, the demerits of Wiener filter are as follows: (1) The all-historical observation data must be used, which leads to great storage and computation cost; (2) when new observation data are obtained, we cannot perform recursion but have to recalculate; and (3) it can hardly be applied to filtering for unstable processes.

To overcome the demerits of Wiener filter, Kalman proposed a filtering method with recursion function in the early 60s, which is called Kalman filtering. Different from Wiener filter, Kalman filtering processes time-varying statistical characteristics, but not consider from the angles of frequency domain or time domain. In the last half-century, Kalman filtering has been widely applied in many domains.

2.2.1 Filter Description

According to the diverse principles in parameter estimation, there are four basic methods including maximum likelihood estimation, maximum a posterior estimation, least square estimation, and minimum mean square error estimation. All these estimation methods are used for time-independent parameters. In the nature, many parameters are a time function, which leads to the estimation problem of time-varying parameters. The estimation of time-varying parameters is also called state estimation. Since unknown parameters are a time function in state estimation, we must consider the evolvement of unknown parameters and observation data with time, during the course of processing measurement data.

State variable method is a valuable method for describing dynamic systems. Using this method, the input and output relationships of systems can be discussed in time domain through the state transition model and the output observation model. Inputs can be described by dynamic models consisting of certain time functions and stochastic processes representing unpredictable variables or noises. Outputs are a function of states, usually disturbed by random observation errors, and can be described by measurement equations. For physical clarity, we use both figures and mathematical expressions to describe the process.

As shown in Fig. 2.1, the state vector is denoted by $X(k)$, the measurement vector is denoted by $Y(k)$, and $Z^{-1}I$ represents storage unit. The mathematical description of state equation and measurement equation is as follows:

(1) State equation

$$X(k+1) = \Phi(k+1,k)X(k) + G(k)W(k) + U(k) \tag{2.1}$$

where $X(k) \in R^{n\times 1}$ is target state vector, $W(k) \in R^{p\times 1}$ is known inputs or control signals, $U(k) \in R^{n\times 1}$ is the white Gaussian noise sequence with zero mean, $Q(k)$ is covariance matrix, and $\Phi(k+1,k) \in R^{n\times n}$ and $G(k) \in R^{n\times p}$ are state transition

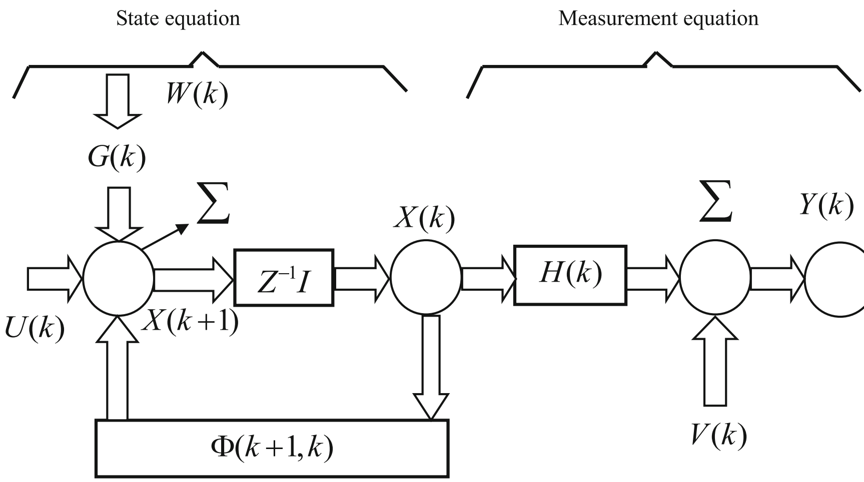

Fig. 2.1 The information flowchart of linearly dynamic discrete systems

matrix and input matrix, respectively. The association function of $U(k)$, white Gaussian noise, is:

$$E\left(U(i)U^T(k)\right) = \begin{cases} Q(k), & i = k \\ 0, & i \neq k \end{cases} \tag{2.2}$$

$\Phi(k+1,k)$, as the transition matrix from time k to time $k+1$, has the following characters.

(a) Multiplication law

$$\Phi(n,k)\Phi(k,m) = \Phi(n,m) \tag{2.3}$$

(b) Inversion law

$$\Phi(n,m)^{-1} = \Phi(n,m) \tag{2.4}$$

where k, m, n are integers.

(2) Measurement equation

$$Y(k) = H(k)X(k) + V(k) \tag{2.5}$$

$Y(k) \in R^{m\times 1}$ is measurement vector, $V(k) \in R^{m\times 1}$ are irrelevant measurement noises, and its covariance matrix is $R(k)$. $H(k) \in R^{m\times n}$ is observation matrix.

The association function of $V(k)$, white noises with zero mean, is:

$$E\big(V(i)V^T(k)\big) = \begin{cases} R(k), & i = k \\ 0, & i \neq k \end{cases} \tag{2.6}$$

Process noise and observation noise are independent from each other. So we get:

$$E\big(U(i)V^T(k)\big) = 0 \quad \forall\, i, k \tag{2.7}$$

The measurement as in Eq. 2.5 identifies the relationship between measurable system output $Y(k)$ and state $X(k)$. The unknown state can be computed by arraying the state equation and observation equation, with a predefined optimization manner. Also, we assume that the initial state $X(0)$ is a Gaussian sequence.

In short, there is three prior information for implementing Kalman filtering discrete-time dynamic systems. First, initial state $X(0)$ is a Gaussian sequence. Second, initial state is independent to process noises and measurement noises. Third, process noises and measurement noises are independent from each other. On this condition, the linear character of state equation and measurement equation can keep the Gaussian character of state and measurement. According to the known measurements at time i and before time i, some kind of estimation of state $X(k)$ at time k is denoted by $\widehat{X}(k, i)$. According to the time indicated by state estimation, Kalman filtering can be described as follows.

When $k = i$, it is called filtering, and $\widehat{X}(k, i)$ is the filtered value of state $X(k)$ at time k;
When $k < i$, it is called smoothing, and $\widehat{X}(k, i)$ is the smoothed value of state $X(k)$ at time k;
When $k > i$, it is called prediction, and $\widehat{X}(k, i)$ is the predicted value of state $X(k)$ at time k.

2.2.2 Filtering Models

The Kalman filtering, under the law of linear mean square error estimation, is very suitable for recursive algorithms to smooth the past and current state of targets and to predict the future motional state of targets, including parameters such as, velocity and acceleration.

The least mean square error estimation of state vector is defined as follows:

$$\widehat{X}(k/k) = E[X(k)/Y_k] \tag{2.8}$$

where $Y_k = \{y(k), k = 1, 2, 3, \ldots, m\}$. The covariance matrix of state errors companying with Eq. 2.8 is:

$$\begin{aligned} P(k/k) &= E\Big[\Big(X(k) - \widehat{X}(k/k)\Big) \cdot \Big(X(k) - \widehat{X}^T(k/k)\Big)/Y_k\Big] \\ &= E\big[\widetilde{X}(k/k) \cdot \widetilde{X}^T(k/k)/Y_k\big] \end{aligned} \tag{2.9}$$

where $\widetilde{X}(k/k)$ is the filtering error. If the error joints with the state equation, we can get the one-step prediction of state:

$$\begin{aligned} \widehat{X}(k+1/k) &= E[X(k+1)/Y(k)] \\ &= E[\Phi(k+1,k)X(k) + G(k)W(k) + U(k)/Y(k)] \\ &= \Phi(k+1,k)\widehat{X}(k/k) + G(k)W(k) \end{aligned}$$

The error of the one-step predicted value is:

$$\widetilde{X}(k+1/k) = X(k+1) - \widehat{X}(k+1/k) = \Phi(k+1,k)\widetilde{X}(k/k) + U(k) \tag{2.10}$$

The covariance for computing the error of the one-step predicted value is:

$$\begin{aligned} P(k+1/k) &= E\Big[\widetilde{X}(k+1) \cdot \Big(X(k+1) - \widehat{X}^T(k+1/k)\Big)/Y_k\Big] \\ &= \Phi(k)P(k/k)\Phi^T(k) + Q(k) \end{aligned} \tag{2.11}$$

Please note that the covariance for computing the error of the one-step predicted value is a symmetric matrix and can be used to measure the uncertainty of predictions. The smaller the $P(k+1/k)$ is, the more precise the prediction is.

Taking the expectation of measurement equation at time $k+1$ under the condition of Y_k, we can similarly get the one-step prediction of measurements:

$$\widehat{Y}(k+1/k) = E[Y(k+1)/Y_k] = H(k+1)\widehat{X}(k+1/k) \tag{2.12}$$

The error of the one-step predicted measurement is:

$$v(k+1) = \widetilde{Y}(k+1/k) = Y(k+1) - \widehat{Y}(k+1/k) = H(k+1)\widetilde{X}(k+1/k) + U(k) \tag{2.13}$$

As this is a measurement of the new information generated during the process of prediction, the prediction error is called "new information" and is also called measurement residual in some references.

The covariance of the new information is:

$$\begin{aligned} S(k+1) &= E\big[\widetilde{Y}(k+1/k) \cdot \widetilde{Y}^T(k+1/k)/Y_k\big] \\ &= H(k+1)P(k+1/k)H^T(k+1) + R(k+1) \end{aligned} \tag{2.14}$$

The covariance of the new information is also a symmetric matrix and can be used to measure the uncertainty of the new information. The smaller the covariance is, the more precise the measurement value is.

The covariance between state and measurement is:

$$P_{XY}(k+1/k) = E\left[\left(\widetilde{X}(k+1/k)\cdot\widetilde{Y}^T(k+1/k)\right)/Y_k\right] = P(k+1/k)H^T(k+1) \tag{2.15}$$

Filter gains are:

$$K(k+1) = P_{XY}(k+1/k)S^{-1}(k+1) = P(k+1/k)H^T(k+1)S^{-1}(k+1) \tag{2.16}$$

So the equation for computing the estimation update of state at time $k+1$ is:

$$\widehat{X}(k+1/k+1) = \widehat{X}(k+1/k) + K(k+1)v(k+1) \tag{2.17}$$

The equation tells that the estimation $\widehat{X}(k+1/k+1)$ at time $k+1$ is the sum of the predicted value $\widehat{X}(k+1/k)$ at this time and a correction item. And this correction item is relevant to gains and the new information.

Furthermore, the updating equation of covariance is:

$$\begin{aligned} P(k+1/k+1) &= P(k+1/k) - P(k+1/k)H^T(k+1)S(k+1)H(k+1)P(k+1/k) \\ &= [I - K(k+1)H(k+1)]P(k+1/k) \\ &= P(k+1/k) - K(k+1)S(k+1)K^T(k+1) \\ &= [I - K(k+1)H(k+1)]P(k+1/k)[I - K(k+1)H(k+1)]^T \\ &\quad - K(k+1)R(k+1)K^T(k+1) \end{aligned} \tag{2.18}$$

where I is the unit matrix with the same dimension as the covariance matrix. Equation 2.18 can guarantee the symmetry and positive definiteness of the covariance matrix P.

Another expression for filtering gain is:

$$\begin{aligned} K(k+1) &= P(k+1/k+1)H^T(k+1)R^{-1}(k+1) \\ &= \big[P(k+1/k)H^T(k+1) - P(k+1/k)H^T(k+1)S^{-1}(k+1) \\ &\qquad H(k+1)P(k+1/k)H^T(k+1)\big]R^{-1}(k+1) \\ &= P(k+1/k+1)H^T(k+1)S^{-1}(k+1) \\ &\qquad \big[S(k+1) - H(k+1)P(k+1/k)H^T(k+1)\big]R^{-1}(k+1) \end{aligned} \tag{2.19}$$

2.2.3 Summary of Filtering and Prediction Models

According to the above-mentioned analysis, the filtering basic equation and one-step prediction equation are as follows, respectively.

Filtering estimation equation is:

$$\widehat{X}(k/k) = \widehat{X}(k/k-1) + K(k) \cdot \left[Y(k) - H(k)\widehat{X}(k/k-1)\right] \tag{2.20}$$

State prediction equation is:

$$\widehat{X}(k/k-1) = \Phi(k/k-1)\widehat{X}(k-1/k-1) \tag{2.21}$$

The filtering gain matrix is:

$$K(k) = P(k/k-1)H^T(k) \cdot \left[H(k)P(k/k-1) \cdot H^T(k) + R(k)\right]^{-1} \tag{2.22}$$

The covariance of prediction errors is:

$$\begin{aligned} P(k/k-1) = \Phi(k/k-1)P(k-1/k-1) \cdot \Phi^T(k/k-1) \\ + G(k-1) \cdot Q(k-1)G^T(k-1) \end{aligned} \tag{2.23}$$

The covariance of filtering estimation is:

$$P(k/k) = [I - K(k)H(k)]P(k/k-1) \tag{2.24}$$

The predicted value of measurements is:

$$\widehat{Y}(k) = H(k)\widehat{X}(k/k-1) \tag{2.25}$$

The new information (residual) vector is:

$$v(k) = Y(k) - H(k)\widehat{X}(k/k-1) \tag{2.26}$$

The covariance matrix of residual vector is:

$$S(k) = H(k)P(k/k-1)H^T(k) + R(k) \tag{2.27}$$

The one-step equation is:

$$\widehat{X}(k+1/k) = \Phi(k+1/k)\widehat{X}(k/k-1) + K(k)\left[Y(k) - H(k)\widehat{X}(k/k-1)\right] \tag{2.28}$$

The one-step predicted gain matrix is:

$$K(k) = \Phi(k+1/k)P(k/k-1) \cdot \left[H(k)P(k/k-1)H^T(k) + R(k)\right]^{-1} \quad (2.29)$$

The one-step covariance matrix is:

$$P(k+1/k) = [\Phi(k+1,k) - K(k)H(k)] \cdot P(k/k-1) + G(k)Q(k)G^T(k) \quad (2.30)$$

2.3 Target Motional Models

2.3.1 Problem Description

(1) Problem posing

Target tracking generally includes target maneuver model identification, maneuver detection, estimation of maneuver quantity, and tracking algorithms. When the target maneuver model is inconsistent with the practical motional situation, mistaken track association or loss of targets can happen. Thus, the target motional model is a key issue in target tracking domain.

The kinetic model of target motion can generally be categorized into two classes—evenly variable target model and maneuver target model. Evenly variable motion means even motion, evenly accelerated motion, and evenly decelerated motion. Maneuver means that there are sudden changes in direction and acceleration during the process of target motion. The statistical characteristics of systematical dynamical noises due to stochastic factors, such as atmospheric disturbance, and observation noises can be correctly described in advance. But maneuver is stochastic and sudden and the direction and quantity of maneuver is unknown, so that we can hardly build up the kinetic model of the maneuvering target timely and correctly. Kalman filtering has good performance in non-maneuvering target tracking, but bad precision in maneuvering target tracking, or even divergence of filters. Therefore, the key is to precisely build up a maneuvering target model to support the problem of maneuvering target tracking. This is why the maneuvering target model is emphasized all the time.

The maneuver model, as one of the basic elements in maneuvering target tracking, is a key but also a challenge, especially in aspects of maneuver detection of sudden changes in velocity and direction of targets and the estimation of maneuver quantity. The maneuvering target model needs to be consistent with the practical maneuver and to be easily dealt with mathematically.

The maneuver detection of target motion is a kind of judgment mechanism, to deal with the problem of the inconsistency between the existing motional model and practical situation upon the occurrence of target maneuver. This inconsistency essentially means the changes in the statistical characteristics of new information sequence caused by maneuver. Thus, the maneuvering target detector can be seen as a filter taking inputs from the new information sequence of Kalman filtering. Maneuver detection method mainly includes the method based on the statistical characteristics of new information sequence, the method based on the extrapolation errors, the geometric method based on the target motional tracks, and the methods of knowledge repository or neural networks emerging in recent years. No matter which method we adopt, we are faced with the following problems in practical applications:

First, maneuver detection threshold. The threshold value is subject to false-alarm probability and detection probability. Thus, maneuver detection confidence needs to be considered.

Second, the estimation lags of maneuver detection. The maneuvering target detection, as a method of a posterior processing, necessarily has lags. Thus, we need to choose an appropriate detection data window length, to timely track the practical maneuver situation of targets.

Third, the transient error problem during the transition between filters for maneuver and non-maneuver.

Fourth, the estimation of maneuver quantity. Currently, the least square method, self-adaptive algorithm, and the geometric method for detector and target interception figure are mainly in use.

Currently, the comparatively effective method is the multi-model method, in maneuvering target tracking. But this method is faced with two difficulties. First one is the design of target maneuver characteristic detector. To guarantee that multiple models can completely cover all the possible maneuvers of targets, we need to comprehensively consider the influence of the signal-to-noise ratio change due to maneuver on the capability of maneuvering target detection, the influence of tracking errors due to maneuver on the capability of tracking, and the influence of filter-model parameter revising due to maneuver quantity on tracking precision. Second one is the robustness. On the one hand, maneuver detection needs to judge on the change in new information sequence mean. Whether it is caused by target maneuver or exterior disturbance, namely the source of abnormal values, needs to be determined. The difficulty for a tracking system lies in how to compromise between the judgment of abnormal values and the precision of system state estimation. Thus, the maneuver detection result without eliminating abnormal values is unreliable, before the processing of target maneuver detection. On the other hand, maneuver detection needs to judge on the loss of target observation data, due to target maneuver, system missing alarms, or exterior disturbances. The maneuvering target detection, under this circumstance, needs to comprehensively consider signal processing.

(2) Review of target maneuver models

In 1970, Singer regarded the acceleration caused by maneuver and atmospheric overfall as a disturbance to even motion and proposed a first-moment stationary Markov process to describe the maneuver acceleration. But this method is suitable for the situation when target has inconsiderable maneuver, but not for the situation of considerable maneuver. The important contribution of the Singer model lies in that it provides a new idea of treating the maneuver as a disturbance to system equations, to the research of maneuver models. In 1973, Mcaulay applied the statistical detection theory to maneuvering target Kalman filtering, constructed the maneuvering target detector, and for the first time comprehensively investigated the maneuvering target tracking problem. Mcaulay took advantage of the new information sequence characteristic of optimal filters, treated the maneuvering target detection problem as a binary hypothesis testing problem, and used two different target models to perform Kalman filtering under conditions of maneuver and non-maneuver, respectively. Thorp introduced a binary random variable to describe maneuver and non-maneuver, directly constructed two sets of Kalman filters, and denoted the final estimation value by weighted linear combination of the filtering results from the two sets.

In order to solve the problem of simply categorizing target motional state as maneuver and non-maneuver, Moose categorized the pattern of maneuvering targets into half-Markov process and Singer first-moment self-regressive process and constructed the correlated Gaussian noise model with random switch mean. He first divided maneuver acceleration into N possible maneuvering acceleration instructions, used a filter set consisting of N augmented Kalman filters to perform filtering, then computed the N possible filtered values of state at some time and the occurrence probability of each filtered value of state, and finally set the filtered value of state at this moment to be the weighted value of the probability. Although this method is more advanced compared to the Thorp method, it needs to choose N acceleration values in advance to approximate the acceleration value of the actual maneuvering target. Obviously, the complexity and computation cost increase with N. Chan regarded maneuver acceleration as an unknown constant and proposed a model to estimate the maneuvering target acceleration by using least square estimation. To degrade the complexity of the model, Chan further simplified the algorithm. Maneuvering detector was no more a detector based on new information sequence, but a detector based on the difference between the extrapolated filtering value and observation value. Bar-Shalom proposed a dimension-varying filtering method. A location–velocity 2-dimensional Kalman filter was adopted when targets had no maneuver, while a location–velocity–acceleration 3-dimensional augmented Kalman filter was adopted when targets had maneuver. In 1990, Bar-Shalom proposed an interacting multiple model tracking algorithm.

To improve the demerits of the existing maneuver models in expensive computational cost or computation delays, Wen proposed a type of increment model to perform detection and estimation of target maneuver characteristics, when using Kalman filter to track targets. The characteristic of the model is the introduction of a vector recording acceleration to the system equation. Take the known maneuver quantity at the last time point as a reference; assume the changes in acceleration to be a set of augmented maneuver; then predict the set of state estimation at the next time point by using the set of augmented values and reference values; and compute the mostly possible state estimation based on Bayesian equation. The mostly obvious feature of this method is the simplicity.

A Chinese researcher Zhou Hongren proposed a current statistical model of maneuvering targets. For each concrete tactical occasion, people care only the "current" probability density of maneuver acceleration, namely the current probability of target maneuver. When targets currently maneuver with acceleration, its value range of the acceleration in the next sampling time point is limited and can be within the neighboring range of the "current" acceleration. Therefore, when describing the probability density of maneuver acceleration, it is completely unnecessary to consider all the possibilities of maneuver acceleration values. Additionally, they are many other maneuvering target models. Here, we do not discuss in detail any more.

2.3.2 *Basic Motional Model*

According the above-mentioned analysis, no matter we use sampling-interval decreasing method or multiple models, they are essentially based on evenly variable motional models. Group-target tracking aims at obtaining the holistic equivalent measurement, so that equivalent measurement represents the mean motional characteristics of group-targets. Under the condition of dense multi-targets, there is no considerable maneuver unless for formation stunt shows.

The group-target tracking in this book is based on evenly variable motional models, so this section will only discuss the classical evenly variable motional model. For the even motion of evenly variable motion and the transformation between maneuver and non-maneuver models, the reader is referred to relevant references.

In a right-angle coordinate system, set the state equation to be $X(k) = \Phi(k, k-1)X(k-1) + G(k-1)W(k-1)$. The evenly variable motional model is a second-moment model. Its 1-dimensional, 2-dimensional, and 3-dimensional state equations are as follows:

(1) 1-dimensional state equation

$$\begin{bmatrix} x(k) \\ \dot{x}(k) \\ \ddot{x}(k) \end{bmatrix} = \begin{bmatrix} 1 & T & \frac{1}{2}T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(k-1) \\ \dot{x}(k-1) \\ \ddot{x}(k-1) \end{bmatrix} + \begin{bmatrix} \frac{1}{2}T^2 \\ T \\ 1 \end{bmatrix} w_x \tag{2.31}$$

(2) 2-dimensional state equation

$$\begin{bmatrix} x(k) \\ \ddot{x}(k) \\ \dddot{x}(k) \\ y(k) \\ \dot{y}(k) \\ \ddot{y}(k) \end{bmatrix} = \begin{bmatrix} 1 & T & \frac{1}{2}T^2 & 0 & 0 & 0 \\ 0 & 1 & T & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T & \frac{1}{2}T^2 \\ 0 & 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(k-1) \\ \ddot{x}(k-1) \\ \dddot{x}(k-1) \\ y(k-1) \\ \dot{y}(k-1) \\ \ddot{y}(k-1) \end{bmatrix} + \begin{bmatrix} \frac{1}{2}T^2 & 0 \\ T & 0 \\ 1 & 0 \\ 0 & \frac{1}{2}T^2 \\ 0 & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} \tag{2.32}$$

(3) 3-dimensional state equation

$$\begin{bmatrix} x(k) \\ \dot{x}(k) \\ \ddot{x}(k) \\ y(k) \\ \ddot{y}(k) \\ \ddot{y}(k) \\ z(k) \\ \dot{z}(k) \\ \ddot{z}(k) \end{bmatrix} = \begin{bmatrix} 1 & T & \frac{1}{2}T^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & T & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T & \frac{1}{2}T^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & T & \frac{1}{2}T^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(k-1) \\ \dot{x}(k-1) \\ \ddot{x}(k-1) \\ y(k-1) \\ \ddot{y}(k-1) \\ \ddot{y}(k-1) \\ z(k-1) \\ \dot{z}(k-1) \\ \ddot{z}(k-1) \end{bmatrix} + \begin{bmatrix} \frac{1}{2}T^2 & 0 & 0 \\ T & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2}T^2 & 0 \\ 0 & T & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2}T^2 \\ 0 & 0 & T \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_x \\ w_y \\ w_z \end{bmatrix} \tag{2.33}$$

where $x(k), \dot{x}(k)$, $\ddot{x}(k), y(k), \dot{y}(k)$, $\ddot{y}(k), z(k), \dot{z}(k)$, and $\ddot{z}(k)$ represent the location, velocity, and acceleration of motional targets in directions of x, y, and z, respectively. T is the sampling cycle. w_x, w_y, w_z are independent Gaussian white noises with zero mean and variance $\sigma_{w_x}, \sigma_{w_y}, \sigma_{w_z}$. The sampling cycle and acceleration variance have considerable influence on the performance of filters. We can choose sampling cycle, but the acceleration variance is an unknown quantity which is difficult to determine.

2.4 Basic Algorithms for Multi-Target Tracking

2.4.1 Review

Data association and track maintenance are the primary problems to be solved in multi-target tracking, currently including maximum likelihood method and Bayesian methods [67, 68, 71]. The maximum likelihood method includes mainly artificial plotting method, track bifurcation method, joint likelihood method, 0–1 integer programming method, and association method. Bayesian methods include Singer's nearest-neighboring method, the all-neighboring optimal method of Singer, Sea, and Housewright, the probabilistic data association method proposed of Shalom, Jaffer, and Tse, Shalom's joint probabilistic data association method, Reid's multiple hypotheses method, Blom's multiple model method, and the interacting multiple model of Blom and Shalom [1, 28, 33, 35, 37]. Recently, L. Hong proposed a interacting multi-velocity multiple model method based on the wavelet transform [46−53].

For the moment, Bayesian algorithms, based on Bayesian rules, still play the leading role and can be categorized into two classes. (1) The optimal Bayesian algorithms process the set of all the validated measurements before current time point, give the probability of each measurement sequence, and synthesize. This class includes the optimal Bayesian algorithm and multiple hypotheses methods and has considerable complexity and computation cost. (2) The second-optimal Bayesian algorithms process only the set of all the validated measurements at current time point. This class includes the nearest-neighboring algorithm, probabilistic data association algorithm, and joint probabilistic data association algorithm and has comparatively low computation cost. The aim of track association (or track identification) is not only to build up an archive of the targets being tracked, but also to complete tasks such as navigation or attack by utilizing target track information. Therefore, although the optimal Bayesian algorithms have many merits compared to the second-optimal Bayesian algorithms, its complexity and high computation cost limit the instantaneity, which is unbearable in realistic applications. In a sense, the association algorithms with good instantaneity and appropriate computation cost have the most powerful vitality [1].

Since group-target tracking is a kind of dense multi-target tracking algorithm based on Bayesian algorithms, next we will emphasize on the nearest-neighboring method, probability data association algorithm, and joint probability data association algorithm.

2.4.2 The Nearest-Neighboring Association Algorithm

In 1971, Singer et al. proposed a tracking filter which had fixed memory demand and could work under condition of multi-returns [37]. This filtering algorithm is the earliest, simplest, and most effective method so far. It takes the statistically "nearest-neighboring" observation data falling inside tracking gates, compared to the predicted location of the targets under tracking, as correlated observation data (although in 1973 Singer and Sea extended the work and developed a class of optimal tracking filters based on the relevant characters of prior statistical estimation [35], which still fell inside the domain of the nearest-neighboring filters). The tracking gates, predicted centers of tracks and the observation data in this cycle, and the relationships among them are shown in Fig. 2.2.

Assume that tracks have been built up before time $k-1$, and the new measurements at time k are $Y_i(k),\ i = 1, 2, \ldots, N$. The difference vector between the ith measurement and tracks is defined as the difference between measurements and predicted values. Namely, the filter residual is:

$$v_i(k) = Y_i(k) - H(k)\widehat{X}(k/k-1) \tag{2.34}$$

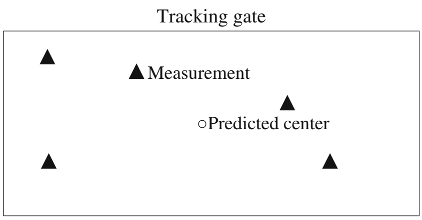

Fig. 2.2 Demonstration of the relationships among tracking gates, predicted centers, and measurements

According to Eq. 2.27, $S(k)$ is the covariance matrix of the residual vector. Then, the statistical distance is:

$$
\begin{aligned}
g_i(k) &= \left\| Y_i(k) - H(k)\widehat{X}(k/k-1) \right\|^2 \\
&= \left[Y_i(k) - H(k)\widehat{X}(k/k-1) \right]^T S^{-1}(k) \cdot \left[Y_i(k) - H(k)\widehat{X}(k/k-1) \right]
\end{aligned} \tag{2.35}
$$

$g_0 = \min(g_i)$ is the observation data correlated with tracks, where $i = 1, 2, \ldots, N$.

The nearest-neighboring data association algorithm is mainly applicable to tracking single target or sparse multi-targets. The merits lie in low computation cost and implementation easiness, and the demerit lies in the easy occurrence of association mistakes when returns have comparatively high density [2].

Generally, the applications of the nearest-neighboring filtering algorithm should obey the principles as follows: (1) There is only one measurement within the association gate of some track, and then, the measurement is naturally correlated to the track. (2) Some measurement falls only within one association gate of tracks, and then, the measurement is correlated to the track. (3) There are multiple measurements within the association gate of some track, and then, the measurement which is nearest to the predicted location is correlated to the track. (4) Some measurement falls within multiple association gates of tracks, and then, the measurement is correlated to the nearest track. Reference [2] believes that the solution to these four principles is not sole and is relevant to the successive sequence of measurements. And the solution considers not the situation when there is no correct measurement. In this case, the recent returns that are screened out are unnecessarily true measurements, which can lead to association mistakes.

2.4.3 The Probabilistic Data Association Algorithm

During the process of target tracking, there are real target returns and other returns caused by receiver noises and all kinds of disturbing among the returns that radar detects. Likewise, there are real target observation data and false observation data caused by noise wave residual, receiver noises, and disturbances among the observation data output by radar collectors. There false observation data are all called false measurements (or false observation data). Since true measurements and false measurements are mixed together, we need to figure out which measurement or which set of measurements are derived from real targets when updating filter states. The nearest-neighboring algorithm believes that all returns are effective and the returns nearest to predicted locations are real target returns, and considers not the possibility that all returns are all false measurements, which makes easy occurrence of association mistakes. In 1972, Bar-Shalom and Jaffer proposed a probabilistic data association algorithm for tracking single targets, by utilizing all returns within tracking gates [19]. In this algorithm, all returns within tracking gates are regarded as valid returns; each return can derive from real targets but with different probabilities. This method comprehensively considers all returns within tracking gates, computes the weighted coefficients of each return probability and the weighted sum according to substantial relevant situations, and uses equivalent measurement after weighted summation to update filters.

As we know, data association needs primarily to build up the tracking gates of targets and validates candidate returns based on the tracking gates. According to Kalman filter equations, filtering residual is $v_i = Y(k) - H(k)\widehat{X}(k/k-1)$, and its covariance is $S(k)$. Assume that real measurements at time k obey Gauss distribution, and then, we get:

$$p[Y(k)/Y^{k-1}] = N[Y(k); \widehat{Y}(k/k-1), S(k)] \tag{2.36}$$

The tracking gate is defined as:

$$V_k(\gamma) = \left\{Y(k) : \left[Y_i(k) - H(k)\widehat{X}(k/k-1)\right]^T S^{-1}(k) \cdot \left[Y_i(k) - H(k)\widehat{X}(k/k-1)\right] \le \gamma\right\} \tag{2.37}$$

In Eq. 2.37, γ is a threshold relevant to radar measuring precision and dynamic errors of filters.

The measurements fall inside tracking gates at time k are denoted by:

$$Y(k) = \{Y_i(k)\}_{i=1}^{m_k} \tag{2.38}$$

In Eq. 2.38, m_k is the number of measurements within tracking gates, and it is a random variable. The accumulative measurement is:

$$Y^k = \{Y(j)\}_{j=1}^{k} \tag{2.39}$$

To estimate the target state according to the measurements at the current time point, the measurements before time k can be computed, by assuming that the state vector $X(k)$ at time k based on the condition Y^{k-1} satisfies the second-optimal Bayesian rules.

$$p\left[X(k)/Y^{k-1}\right] = N\left[X(k); \widehat{X}(k/k-1), P(k/k-1)\right] \tag{2.40}$$

Thus, we can realize recursive computation. Considering the possibility that no measurement is true, we define:

$$\begin{cases} \theta_i(k) = \{Y_i(k) - \text{Measurement derived from targets}\} \\ \theta_0(k) = \{\text{Measurement at time } k \text{ that derived not from targets}\} \\ i = 1, 2, \ldots, m_k \end{cases} \tag{2.41}$$

The conditional probability under the condition of Y^k is:

$$\beta_i = p\{\theta_i(k)/Y^k\} \quad i = 1, 2, \ldots, m_k \tag{2.42}$$

These events are mutually exclusive and complete. Then, we get:

$$\sum_{i=0}^{m_k} \beta_i(k) = 1 \tag{2.43}$$

According to the all-probability theorem, we get the conditional mean at time k is:

$$\widehat{X}(k/k) = E\left[X(k)/Y^k\right] = \sum_{i=0}^{m_k} E\left[X(k)/\theta_i(k), Y^k\right] p\{\theta_i(k)/Y^k\} = \sum_{i=0}^{m_k} \beta_i(k) X_i(k/k) \tag{2.44}$$

In Eq. 2.44, $\widehat{X}_i(k/k)$ is the corrected state estimation under condition of event $\theta_i(k)$.

$$\widehat{X}_i(k) = \widehat{X}(k/k-1) + K(k)v_i(k) \quad i = 1, 2, \ldots, m_k \tag{2.45}$$

In Eq. 2.45, $v_i(k)$ and $K(k)$ are the residual and gains output by standard Kalman filter, respectively. Under the condition of $\theta_i(k)$ $(i \neq 0)$, measurements are derived from targets and are correct measurements. Considering the possibility that no measurement is correct, we substitute measurements with predicted values, namely the state estimation at this moment, is:

$$\widehat{X}_0(k/k) = \widehat{X}(k/k-1) \tag{2.46}$$

We substitute Eqs. 2.45 and 2.46 into Eq. 2.44 and get the state correction equation of probabilistic data association as follows:

$$\widehat{X}(k/k) = \widehat{X}(k/k-1) + K(k)v(k) \tag{2.47}$$

In Eq. 2.47, $v(k) = \sum_{i=1}^{m_k} \beta_i(k) v_i(k)$. Please note that $v(k)$ is formally linear, but essentially nonlinear, because of its dependence on $\beta_i(k)$. The corresponding error estimation covariance is:

$$P(k/k) = \beta_0(k)P(k/k-1) + [1 - \beta_0(k)]P^c(k/k) + \widetilde{P}(k) \tag{2.48}$$

In Eq. 2.48,

$$\widetilde{P}(k) = K(k)\left[\sum_{i=1}^{m_k} \beta_i(k)v_i(k)v_i^T - v(k)v^T(k)\right]K^T(k) \tag{2.49}$$

$$P^c(k/k) = [I - K(k)H(k)]p(k/k-1) \tag{2.50}$$

Next, we give the computation method of $\beta_i(k)$ [1, 21, 24–36, 68]. The assumptions made for probabilistic data association algorithm are as follows.

First, tracks are already initiated. Second, there is at most one true measurement among all the measurements at every time point, and the occurrence probability of this even is P_D. Third, correct measurements obey Gauss distribution. Fourth, false measurements all obey uniform distribution.

True measurements are:

$$P\left[Y_i(k)/\theta_i(k), Y^{k-1}\right] = f\left[Y_i(k)/Y^{k-1}\right] = P_G^{-1}N[v_i(k), 0, S(k)] \quad i = 1, 2, \ldots, m_k \tag{2.51}$$

In Eq. 2.51

$$N[v_i(k), 0, S(k)] = |2\pi S(k)|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}v_i^T(k)S^{-1}(k)v_i(k)\right] \tag{2.52}$$

$$P\left[Y(k); \theta i(k), m_k, Y^{k-1}\right] = \begin{cases} V_k^{-m_k}P_G^{-1}|2\pi S(k)|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}v_i^T(k)S^{-1}(k)v_i(k)\right] \\ V_k^{-m_k} \end{cases} \tag{2.53}$$

P_G is the probability that correct measurements fall inside tracking gates. For simplicity, it is set to be 1.

False measurements that fall inside tracking gates are independent from each other and obey uniform distribution.

$$P[Y_i(k); \theta_j(k), Y^{k-1}] = V_k^{-1}, \quad i \neq j \tag{2.54}$$

In Eq. 2.54, V_k is the volume of association gates.

According to Eq. 2.42, we get

$$\beta_i(k) = p\left\{\theta_i(k)/Y^k\right\} = P\left\{\theta_i(k)/Y(k), m_k, Y^{k-1}\right\} \quad i = 1, 2, \ldots, m_k \tag{2.55}$$

According to Bayesian theorem and multiplication theorem, we can get:

$$\beta_i(k) = \frac{1}{c}p\left[Y(k)/\theta_i(k), m_k, Y^{k-1}\right]P\left\{\theta_i(k)/m_k, Y^{k-1}\right\} \quad i = 1, 2, \ldots, m_k \tag{2.56}$$

In Eq. 2.56

$$c = \sum_{i=0}^{mk} p\left[Y(k)/\theta_i(k), m_k, Y^{k-1}\right]P\left\{\theta_i(k)/m_k, Y^{k-1}\right\} \tag{2.57}$$

where c is a normalization constant. Since $i = 0$, all possible events are considered.

When $i = 0$, all measurements are false measurements. According to Hypothesis 4, if the valid measurements before time k and m_k measurements at time k are all derived from noise waves, the probability density of $Y(k)$ is:

$$P\left[Y(k);\theta_0(k),m_k,Y^{k-1}\right] = \prod_{i=1}^{m_k} P\left[Y_i(k);\theta_0(k),m_k,Y^{k-1}\right] = V_k^{-m_k} \tag{2.58}$$

When $i = 1,2,\ldots,m_k$, according to Hypothesis 3, if there is certainly one measurement, among the valid measurements before time k and m_k measurements at time k, that is derived from targets, the probability density of $Y_i(k)$ is:

$$\begin{aligned} p\left[Y_i(k)/\theta_i(k),m_k,Y^{k-1}\right] &= P_G^{-1}N\left[Y_i\pi(k);\widehat{Y}(k/k-1),S(k)\right] \\ &= P_G^{-1}N[v_i(k);0,S(k)] = P_G^{-1}(2\pi)^{-\frac{1}{2}}\cdot|S(k)|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}v_i^T(k)S^{-1}(k)v_i(k)\right] \end{aligned} \tag{2.59}$$

Then, we get the association probability density of $Y(k)$:

$$\begin{aligned} P\left[Y(k);\theta_i(k),m_k,Y^{k-1}\right] &= P\left[Y_i(k);\theta_0(k),Y^{k-1}\right]\prod_{i=1}^{m_k} P\left[Y_i(k);\theta_0(k),m_k,Y^{k-1}\right] \\ &= V_k^{-m_k}P_G^{-1}|2\pi S(k)|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}v_i^T(k)S^{-1}(k)v_i(k)\right] \end{aligned} \tag{2.60}$$

Integrating Eqs. 2.55 and 2.57, we can get:

$$P\left[Y(k);\theta i(k),m_k,Y^{k-1}\right] = \begin{cases} V_k^{-m_k}P_G^{-1}|2\pi S(k)|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}v_i^T(k)S^{-1}(k)v_i(k)\right] \\ V_k^{-m_k} \end{cases} \tag{2.61}$$

Substituting Eq. 2.58 into Eq. 2.53, we can get:

$$\beta_i(k) = \frac{V_kP_G^{-1}|2\pi S(k)|^{-\frac{1}{2}}r_i(m_k)\exp\left[-\frac{1}{2}v_i^TS^{-1}(k)v_i\right]}{r_0(m_k)+\sum_{i=1}^{mk}V_kP_G^{-1}|2\pi S(k)|^{-\frac{1}{2}}r_i(m_k)\exp\left[-\frac{1}{2}v_i^TS^{-1}(k)v_i\right]}, \quad i=1,2,\ldots,m_k \tag{2.62}$$

$$\beta_0(k) = \frac{r_0(m_k)}{r_0(m_k)+\sum_{i=1}^{mk}V_kP_G^{-1}|2\pi S(k)|^{-\frac{1}{2}}r_i(m_k)\exp\left[-\frac{1}{2}v_i^TS^{-1}(k)v_i\right]} \tag{2.63}$$

In Eqs. 2.61 and 2.62,

$$\gamma_i(m_k) = P[\theta_i(k)/m_k, Y^{k-1}], \quad i = 1, 2, \ldots, m_k \tag{2.64}$$

It is the prior probability of event $\theta_i(k)$ that is correlated with the number of valid measurements. According to reference [2], we directly get:

$$\begin{aligned}\gamma_i(m_k) &= P\left[\theta_i(k)/m^t = m_k, Y^{k-1}\right] = P[\theta_i(k)/m^t = m_k] \\ &= \begin{cases} \frac{1}{m_k} P_D P_G \left[P_D P_G + (1 - P_D P_G)\frac{\mu_f(m_k)}{\mu_f(m_k - 1)}\right]^{-1}, & i = 1, 2, \ldots, m_k \\ (1 - P_D P_G)\frac{\mu_f(m_k)}{\mu_f(m_k - 1)} \left[P_D P_G + (1 - P_D P_G)\frac{\mu_f(m_k)}{\mu_f(m_k - 1)}\right]^{-1}, & i = 0 \end{cases}\end{aligned} \tag{2.65}$$

There are two types of models that can be used for the computation of $\mu_f(m_k)$.

One is parameter model. Assume that m^f obeys Poisson distribution with parameter λV_k, then

$$\mu_f(m_k) = P(m^f = m_k) = \mathrm{e}^{-\lambda V_k} \frac{(\lambda V_k)^{m_k}}{m_k} \tag{2.66}$$

In Eq. 2.66, λV_k is the number of false measurements within association gates.

The other is nonparameter model. Assuming that m^f obeys uniform distribution, then

$$\mu_f(m_k) = P(m^f = m_k) = \frac{1}{N}, \quad m_k = 1, 2, \ldots, N - 1 \tag{2.67}$$

Using parameter models, we simplify Eq. 2.60 as

$$\begin{aligned}P\left[\theta_i(k)/m^t = m_k, Y^{k-1}\right] &= P[\theta_i(k)/m^t = m_k] \\ &= \begin{cases} \frac{P_D P_G}{P_D P_G m_k + (1 - P_D P_G)\lambda V_k}, & i = 1, 2, \ldots, m_k \\ \frac{(1 - P_D P_G)\lambda V}{P_D P_G m_k + (1 - P_D P_G)\lambda V_k}, & i = 0 \end{cases}\end{aligned} \tag{2.68}$$

Substituting Eq. 2.68 into Eqs. 2.62 and 2.63, we get association the following matrices.

$$\beta_i(k) = \frac{e_i(k)}{b(k) + \sum_{i=1}^{m_k} e_i(k)} \quad i = 1, 2, \ldots, m_k \tag{2.69}$$

$$\beta_0(k) = \frac{b(k)}{b(k) + \sum_{i=1}^{m_k} e_i(k)} \quad i = 1, 2, \ldots, m_k \tag{2.70}$$

In Eq. 2.70,

$$e_i(k) = \exp\left[-\frac{1}{2}v_i^T S^{-1}(k)v_i(k)\right], \quad i = 1, 2, \ldots, m_k \tag{2.71}$$

$$b(k) = \lambda|2\pi S(k)|^{\frac{1}{2}}(1 - P_D P_G)/P_D, \quad i = 1, 2, \ldots, m_k \tag{2.72}$$

2.4.4 The Joint Probabilistic Data Association Algorithm

When measurements fall within the intersecting areas of different target tracking gates, measurements can be derived from any target, and then, we need to compute the association probability [22, 25, 28] of the association between each measurement and all kinds of possible source targets. In order to get the association probability, we do not set an independent tracking gate for each target and make the consistency between tracking gates and the whole surveillance area. The aim is to make equal the probability density function of each false measurement within the whole surveillance area and thus to get the conditional probability of association events. But this method does not ignore those events that should be ignored and consequently increases extra computation burden. Therefore, Bar-Shalom introduced the conception of cluster matrix. The structure of cluster matrix is:

$$\Omega = \left[\omega_{jt}\right] = \begin{bmatrix} \omega_{10} & \cdots & \omega_{1T} \\ \vdots & \ddots & \vdots \\ \omega_{m_k 0} & \cdots & \omega_{m_k T} \end{bmatrix} \tag{2.73}$$

In Eq. 2.73, ω_{jt} is a binary variable. When $\omega_{jt} = 1$, measurement $j(j = 1, 2, \ldots, m_k)$ falls inside the validation gate of target $t(t = 0, 1, \ldots, T)$. Meanwhile, when $\omega_{jt} = 0$, measurement j does not fall inside the validation gate of target t. When $t = 0$, there is no targets, and then the column elements ω_{j0} that Ω is corresponding to is all 1. This is because each measurement can be derived from noise wave or false alarm. Therefore, the purpose of ignoring extra events and decreasing computational burden can be achieved. Since the measurement that falls inside each target tracking gate is regarded as valid measurement, the above-mentioned methods correspondingly obtain the cluster matrix (or called validation matrix) including m_k measurements and T targets. “Cluster” is defined as the maximum set of intersecting tracking gates. Targets are categorized into different groups according to different “clusters” [22]. There is always a matrix with dual-value elements to correlate to each group like this. This matrix is called the cluster matrix.

For any given multi-target problem, once the validation matrix Ω, which reflects the association status between valid returns and targets or noise waves, is given, we can get the association matrix representing all association events by splitting the validation matrix. The two basic assumptions for the splitting are as follows:

Assumption 1 Each measurement has an only source, which means that each measurement is derived from either targets, or noise waves or false alarms. That is to say, we do not consider the situation when one target has multiple returns.

Assumption 2 For any given target, there is only one measurement that is derived from it. If one target is correlated with multiple measurements, only one of the measurements will be selected and set to be true, and others are false. This is consistent with the assumptions made for PDA algorithm.

The basic idea of JPDA algorithm lies in that all valid measurements that fall inside the tracking gates of different targets t can be derived from target t, but the values of association probability of each measurement are different. Thus, we first define association event and then define joint association event.

The association event between the ith measurement and target t is defined as:

$$\begin{aligned}\theta_{it}(k) = \{&\text{Valid measurement } Y_i(k) \text{ is derived from targets}\},\\ &i = 1, 2, \ldots, m_k; \quad t = 0, 1, 2, \ldots, T\end{aligned} \tag{2.74}$$

When t $= 0$, $\theta_{i0}(k)$ denotes the event that measurement $Y_i(k)$ is derived from false measurements. According to the assumptions for PDA algorithm, the event that measurements are associated with target t has the following characters:

(1) Mutual exclusivity:

$$\theta_{it}(k) \bigcap \theta_{jt}(k) = \emptyset, \quad i \neq j \tag{2.75}$$

(2) Completeness:

$$\bigcup_{i=0}^{m_k} P[\theta_{it}(k)/Y^k] = 1 \tag{2.76}$$

The association probability of association event is:

$$\beta_{it}(k) = P[\theta_{it}(k)/Y^k], \quad i = 0, 1, 2, \ldots, m_k; \quad t = 0, 1, 2, \ldots, T \tag{2.77}$$

$\beta_{it}(k)$ denotes the probability value when the ith measurement is associated with target t. According to all-probability equation, we can get:

$$\widehat{X}^t(k/k) = E[X^t(k)/Y^k] = \sum_{i=0}^{m_k} E[X^t(k)/\theta_{it}(k), Y^k] P[\theta_{it}(k)/Y^k] = \sum_{i=0}^{m_k} \beta_{it}(k) \widehat{X}_i^t(k/k) \tag{2.78}$$

In Eq. 2.78,

$$\sum_{i=0}^{m_k} X_i(k) = E[X^t(k)/\theta_{it}(k), Y^k], \quad i = 0, 1, 2, \ldots, m_k \tag{2.79}$$

denotes the state estimation of the ith measurement at time k when we perform Kalman filtering on it and target t. $\widehat{X}_0^t(k/k)$ denotes the situation when there is no measurement is derived from targets at time k. We must use the predicted value $\widehat{X}^t(k/k-1)$ as measurement values, namely $\widehat{X}_0^t(k/k) = \widehat{X}^t(k/k-1)$.

Next, we define joint events and the association probability of joint events, according to the definition of association events.

Let $\theta(k) = \{\theta_j(k)\}_{j=1}^{\theta_k}$ be the set of all possible association events at time k, and θ_k is the number of association events in $\theta(k)$. Then,

$$\theta_j(k) = \bigcap_{i=0}^{m_k} \theta_{it_i}^j(k), \quad i = 0, 1, 2, \ldots, m_k; \ \ j = 0, 1, 2, \ldots, \theta_k; \ \ t_i = 0, 1, 2, \ldots, T \tag{2.80}$$

$\theta_j(k)$ denotes the jth association event, $\theta_{it}^j(k)$ denotes the event that measurement i is derived from target t_i in the jth association event, and $\theta_{i0}^j(k)$ denotes the event that measurement i is derived from false measurements in the jth association event. Thus, according to Eq. 2.74, we get:

$$\theta_{it}(k) = \bigcup_{j=1}^{\theta_k} \theta_{it_i}^j, \quad i = 0, 1, 2, \ldots, m_k; \ \ j = 0, 1, 2, \ldots, \theta_k; \ \ t_i = 0, 1, 2, \ldots, T \tag{2.81}$$

There are two known assumptions for joint probabilistic event data association. First, every measurement has a sole source. Namely, any measurement is derived from either targets, or noise waves or false alarms. In other words, the situation when one target has multiple returns and they are unidentified is not considered. Second, for any given target, at most one measurement is derived from it. If one target generates multiple measurements, only one of the measurements will be selected and set to be true, and others are false. An event is called a feasible event when the two assumptions are satisfied.

A joint event $\theta_j(k)$ is denoted by the association matrix as:

$$\Omega[\theta_j(k)] = \{\omega_{it}^j[\theta_j(k)]\}, \quad i = 0, 1, 2, \ldots, m_k; \ \ j = 0, 1, 2, \ldots, \theta_k \tag{2.82}$$

In Eq. 2.82,

$$\omega_{it}^{j}\left[\theta_j(k)\right] = \begin{cases} 1, & \theta_{it}^{j} \subset \theta_j(k) \\ 0, & \text{else} \end{cases} \tag{2.83}$$

denotes in the jth association event, whether measurement j is derived from target t. According to the above-mentioned assumptions, the association matrix satisfies:

$$\sum_{t=0}^{T} \omega_{it}^{j}[\theta_j(k)] = 1, \quad i = 0, 1, 2, \ldots, m_k \tag{2.84}$$

$$\sum_{i=0}^{m_k} \omega_{it}^{j}[\theta_j(k)] \leq 1, \quad t = 0, 1, 2, \ldots, T \tag{2.85}$$

For simplicity, we define two binary variables.

One is measurement association indicator:

$$\tau_i[\theta_j(k)] = \sum_{t=0}^{T} \omega_{it}^{j}[\theta_j(k)] = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \end{cases}, \quad i = 0, 1, 2, \ldots, m_k \tag{2.86}$$

denotes whether measurement i is associated with a real target in joint event $\theta_j(k)$. Let

$$\tau[\theta_j(k)] = \left\{\tau_1[\theta_j(k)], \tau_2[\theta_j(k)], \ldots, \tau_{m_k}[\theta_j(k)]\right\} \tag{2.87}$$

Then, $\tau[\theta_j(k)]$ reflects whether any measurement is associated with real targets in some joint event.

The other is a target detection indicator.

$$\begin{aligned} \phi_t\left[\theta_j(k)\right] &= \sum_{i=0}^{m_k} \omega_{it}^{j}[\theta_j(k)] \leq 1 \\ &= \begin{cases} 1, & \text{there is measurement } i \text{ that is associated to real targets} \\ 0, & \text{there is no measurement } i \text{ that is associated to real targets} \end{cases}, \\ t &= 0, 1, 2, \ldots, T \end{aligned} \tag{2.88}$$

denotes whether any measurement is associated with a target t in joint event $\theta_j(k)$, namely whether target t will be detected or not. Similarly, let:

$$\phi[\theta_j(k)] = \{\phi_1[\theta_j(k)], \phi_2[\theta_j(k)], \ldots, \phi_T[\theta_j(k)]\} \tag{2.89}$$

Then, $\phi[\theta_j(k)]$ denotes the event whether any target is detected in joint event $\theta_j(k)$. Let $\varphi[\theta_j(k)]$ denote the number of false measurements in joint event $\theta_j(k)$, then

$$\varphi[\theta_j(k)] = \sum_{i=1}^{m_k} \{1 - \tau_i[\theta_j(k)]\} \tag{2.90}$$

According to the above-mentioned discussion, once we get the validation matrix of the association situations between valid measurements and targets or noise waves, the feasible matrix of all feasible joint events can be obtained by splitting the validation matrix. According to the two assumptions, the splitting of validation matrix should obey the following two principles.

(1) There can exist only one nonzero element in each row of the validation matrix. This is the only nonzero element in this row of the validation matrix, which means each measurement has a sole source. Thus, Assumption 1 is satisfied.
(2) In the feasible matrix, except the first column, there can exist at most one nonzero element in each column, which means at most one measurement is derived from each target. Thus, Assumption 2 is satisfied.

The feasible matrix has a one-to-one correspondence with feasible joint events. Generally, we obtain the feasible matrix by splitting the validation matrix and then determine the feasible joint events. With the increase of target number and the number of valid measurements, the number of feasible matrices quickly expands exponentially. Meanwhile, the bigger the intersecting areas of association gates are and the bigger the number of intersecting areas is, the bigger the number of feasible matrices will be.

The computation of joint event association probability is as follows.

According to Bayesian rules, the conditional probability of the joint events of all measurements at time k is:

$$\begin{aligned} P\left[\theta_j(k)/Y^k\right] &= P\left[\theta_j(k)/Y(k), Y^{k-1}\right] = \frac{1}{c} P\left[Y(k)/\theta_j(k), Y^{k-1}\right] P\left[\theta_j(k)/Y^{k-1}\right] \\ &= \frac{1}{c} P\left[Y(k)/\theta_j(k), Y^{k-1}\right] P\left[\theta_j(k)\right] \end{aligned} \tag{2.91}$$

In Eq. 2.91,

$$c = \sum_{j=1}^{\theta_k} P\left[Y(k)/\theta_j(k), Y^{k-1}\right] P\left[\theta_j(k)\right] \tag{2.92}$$

Similar to the computation given in Eq. 2.60, we can get:

$$P\left[Y(k)/\theta_j(k), Y^{k-1}\right] = \prod_{i=1}^{m_k} P\left[Y_i(k)/\theta_j(k), Y^{k-1}\right] = \prod_{i=1}^{m_k} P\left[Y_i(k)/\theta_{it}^j(k), Y^{k-1}\right] \tag{2.93}$$

Since false measurements are assumed to obey uniform distribution, true measurements associated with targets are assumed to obey Gauss distribution and association gates are assumed to be corresponding to the whole surveillance area, namely $P_G = 1$, we can get:

$$P\left[Y_i(k)/\theta_{it}^j(k), Y^{k-1}\right] = \begin{cases} N_t[Y_i(k)] = N\left[Y_i(k); \widehat{Y}_{t_i}(k/k-1), S_{t_i}(k)\right], & \tau_i\left[\theta_j(k)\right] = 1 \\ V^{-1}, & \tau_i\left[\theta_j(k)\right] = 0 \end{cases} \tag{2.94}$$

In Eq. 2.94, $\widehat{Y}_{t_i}(k/k-1)$ is the predicted value of target t_i and $S_{t_i}(k)$ is the covariance of the corresponding new information. Substituting Eq. 2.94 into Eq. 2.93, we can get:

$$P\left[Y(k)/\theta_j(k), Y^{k-1}\right] = V^{-\varphi[\theta_j(k)]} \prod_{i=1}^{m_k} N_{t_i}[Y_i(k)]^{\tau_i[\theta_j(k)]} \tag{2.95}$$

As we know, if $\theta_j(k)$ is determined, target detection indicator $\phi[\theta_j(k)]$ and the number of false measurements $\varphi[\theta_j(k)]$ are also determined. Thus,

$$P[\theta_j(k)] = P\left\{\theta_j(k)/\phi[\theta_j(k)], \varphi[\theta_j(k)]\right\} P\left\{\phi[\theta_j(k)], \varphi[\theta_j(k)]\right\} \tag{2.96}$$

In fact, if the number of false measurements is determined, joint event $\theta_j(k)$ is also solely determined according to $\phi[\theta_j(k)]$. Then, the number of event including $\varphi[\theta_j(k)]$ false measurements should be $\mathrm{C}_{m_k}^{\varphi[\theta_j(k)]}$. The rest $m_k - \varphi[\theta_j(k)]$ true measurements can have $\{m_k - \varphi[\theta_j(k)]\}!$ possible joint events. Therefore, we can get:

$$P\left\{\theta_j(k)/\phi[\theta_j(k)], \varphi[\theta_j(k)]\right\} = \frac{1}{\mathrm{C}_{m_k}^{\varphi[\theta_j(k)]}\left\{m_k - \varphi[\theta_j(k)]\right\}!} = \frac{\varphi\left[\theta_j(k)\right]!}{m_k!} \tag{2.97}$$

Again on the account of

$$P\left\{\phi\left[\theta_j(k)\right], \varphi\left[\theta_j(k)\right]\right\} = \mu_f\left\{\varphi\left[\theta_j(k)\right]\right\} \prod_{t=1}^{T} \left(P_D^t\right)^{\phi[\theta_j(k)]} \left(1 - P_D^t\right)^{1-\phi[\theta_j(k)}} \tag{2.98}$$

In Eq. 2.98, P_D^t denotes the detection probability of target t. $\mu_f\{\varphi[\theta_j(k)]\}$ denotes the distribution function of prior probability of the false measurement number.

Substituting Eqs. 2.97 and 2.98 into Eq. 2.96, we can get the prior probability of joint event $\theta_j(k)$ as:

$$P[\theta_j(k)] = \frac{\varphi[\theta_j(k)]!}{m_k!}\mu_f\{\varphi[\theta_j(k)]\}\prod_{t=1}^{T}(P_D^t)^{\phi[\theta_j(k)]}(1-P_D^t)^{1-\phi[\theta_j(k)} \tag{2.99}$$

Likewise, substituting Eqs. 2.93 and 2.99 into Eq. 2.91, we can get the posterior probability of joint event $\theta_j(k)$ as:

$$\begin{aligned} & P[\theta_j(k)/Y^k] \\ & = \frac{\varphi[\theta_j(k)]!}{m_k!}\mu_f\{\varphi[\theta_j(k)]\}V^{-\varphi[\theta_j(k)]}\prod_{i=1}^{m_k}N_{t_i}[Y_i(k)]^{\tau_i[\theta_j(k)]}\prod_{t=1}^{T}(P_D^t)^{\phi[\theta_j(k)]}(1-P_D^t)^{1-\phi[\theta_j(k)} \end{aligned} \tag{2.100}$$

According to the probabilistic data association algorithm, the conditional probabilities of the probability distribution function $\mu_f\{\varphi[\theta_j(k)]\}$ with Poisson-distributed parameters and uniform-distributed nonparameters and the corresponding joint events are:

$$\mu_f\{\varphi[\theta_j(k)]\} = \mathrm{e}^{-\lambda V}\frac{(\lambda V)^{\varphi[\theta_j(k)]}}{\varphi[\theta_j(k)]!} \tag{2.101}$$

Substituting Eq. 2.101 into Eq. 2.100, we can get:

$$P[\theta_j(k)/Y^k] = \frac{\lambda^{\varphi[\theta_j(k)]}}{c'}\prod_{i=1}^{m_k}N_{t_i}[Y_i(k)]^{\tau_i[\theta_j(k)]}\prod_{t=1}^{T}(P_D^t)^{\phi[\theta_j(k)]}(1-P_D^t)^{1-\phi[\theta_j(k)} \tag{2.102}$$

In Eq. 2.102, c' is a new normalization constant.

$$P[\theta_j(k)/Y^k] = \frac{1}{c}\frac{\varphi[\theta_j(k)]!}{V^{\varphi[\theta_j(k)]}}\prod_{i=1}^{m_k}N_{t_i}[Y_i(k)]^{\tau_i[\theta_j(k)]}\prod_{t=1}^{T}(P_D^t)^{\phi[\theta_j(k)]}(1-P_D^t)^{1-\phi[\theta_j(k)} \tag{2.103}$$

Thus, we get the probability when the ith measurement is associated with targets as follows:

$$\begin{aligned} \beta_{it}(k) &= P\left[\theta_{it}(k)/Y^k\right] \\ &= P\left[\bigcup_{j=1}^{\theta_k}\theta_{it_i}^j/Y^k\right] = \sum_{j=1}^{\theta_k}P\left[\theta_i(k)/Y^k\right]\omega_{it}^j\left[\theta_j(k)\right], \\ i &= 0,1,2,\ldots,m_k;\quad t = 0,1,2,\ldots,T \end{aligned} \tag{2.104}$$

The association probability when there is no measurement is derived from targets is:

$$\beta_{0t}(k) = 1 - \sum_{i=1}^{m_k} \beta_{it}(k) \tag{2.105}$$

After obtaining the joint association probability, we can naturally give its filtering equations. Here, we do not give unnecessary details again.

2.4.5 Generalized Probabilistic Data Association Algorithm

Because the computation of the joint probabilistic data association algorithm is comparatively big, researchers propose multiple improved algorithms of JPDA, mainly including the precise nearest-neighboring data association ENNPDA, the coupling probabilistic data association CPDA, the JPDA algorithm with avoidance of track aggregation, joint synthesized probabilistic data association JIPDA, and comprehensive IJPDA. All these improvements make some assumption based on the feasible rules of JPDA algorithm and are simplified algorithms on the cost of precision and cannot essentially solve the computation cost problem. The feasible rule of JPDA is that one measurement can belong to one target and one target can at most possess one measurement.

In "T. Kirubarajan, Bar-Shalom, K.R. Pattipati. "Multiassignment for Tracking a large Number of Overlapping Objects." IEEE Trans. on Aerospace and Electronic Systems. Vol. 37, No. 1, Jan. 2001, pp. 2–20", the viewpoint of the non-one-to-one corresponding between measurements and tracks was also presented. They implement the multi–multi-correspondence between measurements and tracks through the repeated usages of one-to-one distributions. This is also based on the feasible rule of JPDA, and the repeated usages of one-to-one distributions lead to even bigger computation burden.

The generalized probabilistic data association breaks the limitation of the feasible rule of JPDA, under the condition of non-one-to-one correspondence between measurements and targets. It is another feasibility assumption of joint events, namely the two events of targets and measurements constitute a joint event, so that it is an algorithm under condition of repeated usages of both measurements and targets. Compared to JPDA algorithm, the computation load is decreased as the precision is improved.

(1) The idea of the algorithm

Corresponding to the conception of JPDA joint association event, we define the joint association event in the new algorithm. In order to be different from that in JPDA algorithm, the joint association event is called joint event. With known number of target T and number of measurements m_k, a joint event consists of the

Fig. 2.3 The matrix consisting of the cluster-probability statistical distance between measurements and targets

$$\begin{array}{c|ccc} & \multicolumn{3}{c}{\text{Target}} \\ \text{Measurements} & 0 & 1 & 2 \\ \hline 0 & f_{00} & f_{01} & f_{02} \\ 1 & f_{10} & f_{11} & f_{12} \\ 2 & f_{20} & f_{21} & f_{22} \end{array}$$

following two events and these two events represent the feasible rule of the new algorithm.

Event ①: Each target has measurements whose number can be several or 0;
Event ②: Each measurement has target sources, whose number can be several or 0.

Here, 0 targets mean no targets. Namely, all measurements are derived from false targets such as disturbances and noise waves, or new targets, except targets under tracking. 0 measurements mean no detection of targets.

Figure 2.3 shows the matrix consisting of the cluster-probability statistical distance between measurements and targets. f_{it} is the statistical distance between measurement i and target t. θ is assumed to be the association event between measurement i and target t.

The events that satisfy events ① and ② are:

$$\begin{aligned} &\theta_{00}\theta_{11}\theta_{22}, \quad \theta_{20}\theta_{21}\theta_{22}, \quad \theta_{00}\theta_{11}\theta_{02}, \quad \cdots \\ &\theta_{11}\theta_{22}\theta_{00}, \quad \theta_{11}\theta_{21}\theta_{01}, \quad \theta_{12}\theta_{22}\theta_{02}, \quad \theta_{12}\theta_{20}\theta_{02}, \quad \cdots \end{aligned}$$

According to the definition, joint events are the combination of the two types of events. We can see that the events that satisfy event ① are on a basis of targets, while the events that satisfy event ② are on a basis of measurements. Thus, a joint event set can be seen as the set of the two types of events. Then utilizing Bayesian rules, we can get the mutual belonging probability β_{it} between measurement i and target t (corresponding to the marginal probability in JPDA algorithm). Although the set of joint events can be partitioned as sets of other types of events, the method discussed before is the simplest, most efficient, and optimal partition.

Figure 2.4 further describes the idea of GPDA algorithm.

The processing on a basis of targets is to deal with the cluster-probability statistical matrix under assumption of "each target has measurements." The problem of "each measurement has its target source" is not considered. In contrary, the processing on a basis of measurements is to deal with the cluster-probability statistical matrix under the assumption of "each measurement has its target source." Similarly, the problem of "each target has measurements" is not considered.

The difference between GPDA and JPDA lies in whether repeated usage of targets and measurements is allowed. The processing on a basis of targets is

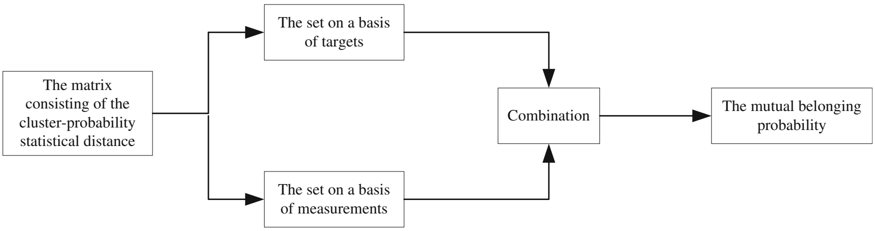

Fig. 2.4 The basic idea of GPDA algorithm

corresponding to the problem of repeated usage of measurements. Namely, "each target has measurements" (these measurements can be the same measurement) solves the problem of "one measurement is associated to multiple targets." Likewise, the processing on a basis of measurements is corresponding to the problem of repeated usage of targets. Namely, it solves the problem of "one target is associated to multiple measurements."

Combining the repeated usage of measurements and targets, we unify "one measurement is corresponding to multiple targets" and "one target is corresponding to multiple measurements" and implement the multi–multi-correspondence between measurements and targets.

(2) Relevant definitions

Like JPDA algorithm, GPDA algorithm is proposed to deal with multi-target tracking problem. The definition of the measurement set is as same as that in JPDA, and so are the set expressions. The validation set at time k is:

$$Y(k) = \{y_i(k)\}_{i=1}^{m_k} \tag{2.106}$$

where m_k is the number of measurements within validation area. Then, the accumulative set of measurements is:

$$Y^k = \{Y(j)\}_{j=1}^{k} \tag{2.107}$$

Similarly, we do not set an independent validation gate for each target and make the consistency between tracking gates and the whole surveillance area. Assume that we track T targets, then m_k is the number of measurements within validation area at time k. We give the following definitions.

Definition 1 $F = [f_{it}]$, $i = 0, 1, \ldots, m_k$, $t = 0, 1, \ldots, T$, where F is the matrix consisting of the cluster-probability statistical distance between measurements and targets. f_{it} is the statistical distance between measurement i and target t, which is also called probability density function.

Definition 2 Under the condition of $F = [f_{it}]$, Θ is the joint events that satisfy new feasibility rules, Θ_t is events that satisfy the assumption ①, and Θ_i is events that satisfy the assumption ②.

$$\Theta = \Theta_t \cup \Theta_i \tag{2.108}$$

(3) Deduction about mutual belonging probability

First construct matrix F, and assume that the state variable of target t obeys normalization distribution with mean of $\widehat{X}_t(k/k-1)$ and covariance of $P_t(k/k-1)$. Namely,

$$P\left[X_t(k)/Y_t^{k-1}\right] = N\left[X_t(k); \widehat{X}_t(k/k-1), P_t(k/k-1)\right] \tag{2.109}$$

Then, the probability density function when measurement $i\,(i \neq 0)$ is corresponding to target $t\,(t \neq 0)$ (Y. Bar-Shalom, T.E. Fortmann. Tracking and Association. Orlando, FL: Academic Press, 1998) is:

$$\begin{aligned} f_{it} &= p\left[Y_{it}/m_k, Y_t^k\right] = P_G^{-1} N\pi\left[Y_{it}(k); \widehat{Y}_t(k/k-1), S_t(k)\right] \\ &= P_G^{-1}|2\pi S_t(k)|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} v_{it}^T(k) S_t^{-1}(k) v_{it}(k)\right] \end{aligned} \tag{2.110}$$

where $v_{it}(k) = Y_{it}(k) - Y_{it}(k/k-1)$.

The probability density function when 0 measurement is corresponding to target $t\,(t \neq 0)$ means that there is no measurements to be corresponding to the target. Namely, the probability density function of no-detection event is $p_L = f_{0t} = (nV)^{-1}(1 - p_D p_G)$, where n is a proportional coefficient, V is the volume of wave gates, p_D is the detection probability, and p_G is the gate probability.

The probability density function when measurement $i\,(i \neq 0)$ is corresponding to 0 targets means that this measurement does not belong to any interested target within spatial surveillance area. Namely, the probability density function of the event when this measurement belongs to false targets is:

$$p_F = f_{i0} = \lambda \tag{2.111}$$

where false measurements within wave gates are regarded to obey Poisson distribution and λ is the density of false measurements, namely the number of noise waves within a unit volume.

The association between 0 measurement and 0 target means nothing, and its probability density function is:

$$f_{00} = 0 \tag{2.112}$$

According to the depiction, the matrix $F = [f_{it}]$ consisting of the cluster-probability statistical distance between measurements and targets is:

$$F = [f_{it}] = \begin{array}{cc} & \begin{array}{cccccc} 0 & 1 & 2 & \cdots & T \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ m_k \end{array} & \begin{bmatrix} f_{00} & f_{01} & f_{02} & \cdots & f_{0T} \\ f_{10} & f_{11} & f_{12} & \cdots & f_{1T} \\ f_{20} & f_{21} & f_{22} & \cdots & f_{2T} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f_{m_k0} & f_{m_k1} & f_{m_k2} & \cdots & f_{m_kT} \end{bmatrix} \end{array} \tag{2.113}$$

First, normalizing the matrix given in Eq. 2.113 on a basis of targets, we get the matrix $E_t = [\varepsilon_{it}]$ as:

$$\varepsilon_{it} = \frac{f_{it}}{c_t} \tag{2.114}$$

where c_t is the normalization constant on target t.

$$c_t = \sum_{i=0}^{m_k} f_{it} \tag{2.115}$$

Then, normalizing the matrix on a basis of measurements, we get $E_i = \left[\varepsilon'_{it}\right]$ as:

$$\varepsilon'_{it} = \frac{f_{it}}{c_i} \tag{2.116}$$

where c_t is the normalization constant on target t.

$$c_i = \sum_{t=0}^{T} f_{it} \tag{2.117}$$

The normalized matrix on a basis of targets and the normalized matrix on a basis of measurements are as follows:

$$\begin{array}{cc} & \begin{array}{cccccc} 0 & 1 & 2 & \cdots & T \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ m_k \end{array} & \begin{bmatrix} \varepsilon_{00} & \varepsilon_{01} & \varepsilon_{02} & \cdots & \varepsilon_{0T} \\ \varepsilon_{10} & \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1T} \\ \varepsilon_{20} & \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2T} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \varepsilon_{m_k0} & \varepsilon_{m_k1} & \varepsilon_{m_k2} & \cdots & \varepsilon_{m_kT} \end{bmatrix} \end{array} \tag{2.118}$$

$$\begin{array}{cc}
 & \begin{matrix} 0 & 1 & 2 & \cdots & T \end{matrix} \\
\begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ m_k \end{matrix} &
\begin{bmatrix}
\varepsilon'_{00} & \varepsilon'_{01} & \varepsilon'_{02} & \cdots & \varepsilon'_{0T} \\
\varepsilon'_{10} & \varepsilon'_{11} & \varepsilon'_{12} & \cdots & \varepsilon'_{1T} \\
\varepsilon'_{20} & \varepsilon'_{21} & \varepsilon'_{22} & \cdots & \varepsilon'_{2T} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\varepsilon'_{m_k0} & \varepsilon'_{m_k1} & \varepsilon'_{m_k2} & \cdots & \varepsilon'_{m_kT}
\end{bmatrix}
\end{array} \tag{2.119}$$

Based on E_i, E_t, we give the equation for computing mutual belonging probability.

According to Definition 2, we obviously get:

$$\sum_{f \in F_{it}} p\{\theta_{it}/\Theta, Y^k\} = \sum_{f_1 \in F_{1_{it}}} p\{\theta_{it}/\Theta_t, Y^k\} \cup \sum_{f_2 \in F_{2_{it}}} p\{\theta_{it}/\Theta_i, Y^k\} \tag{2.120}$$

where the meaning of $\Theta, \Theta_t, \Theta_i$ are as defined in Definition 2, F_{it} is the set of all the joint events f that include θ_{it}, $F_{1_{it}}$ is the set of all the events f_1 that include θ_{it} and satisfy assumption ①, and $F_{2_{it}}$ is the set of all the events f_2 that include θ_{it} and satisfy assumption ②.

According to the conditional Bayesian equation for multiple events, we can get:

$$\begin{aligned}
\sum_{f_1 \in F_{1_{it}}} p\{\theta_{it}/\Theta_t, Y^k\} &= \sum_{f_1 \in F_{1_{it}}} p\{\theta_{it}/Y^k\} \cdot p\{\Theta_t/\theta_{it}, Y^k\} \\
&= \sum_{f_1 \in F_{1_{it}}} p\{\theta_{it}/Y(k), Y^{k-1}\} \cdot p\{\Theta_t/\theta_{it}, Y(k), Y^{k-1}\} \\
&= p\left\{\theta_{it}/Y(k), \widehat{X}_t(k/k-1), P_t(k/k-1)\right\} \\
&\quad \cdot \sum_{f_1 \in F_{1_{it}}} p\left\{\Theta_t/\theta_{it}, Y(k), Y^{k-1}, \widehat{X}_t(k/k-1), P_t(k/k-1)\right\}
\end{aligned} \tag{2.121}$$

Because normalization has been performed before, we pass over the normalization coefficient in this equation. The first item to the right of the last equal mark in Eq. 2.121 is only the prior probability, and its computation equation is:

$$p\left\{\theta_{it}/Y(k), \widehat{X}_t(k/k-1), P_t(k/k-1)\right\} = \varepsilon_{it} \tag{2.122}$$

Because:

$$p\left\{\Theta_t/\theta_{it}, Y(k), Y^{k-1}, \widehat{X}_t(k/k-1), P_t(k/k-1)\right\} = p\left\{\theta_{it}/\Theta_t, Y^k\right\} = \prod_{\substack{r=0,t_r=0\\ r\neq i, t_r\neq t}}^{m_k,T} \varepsilon_{rt_r} \tag{2.123}$$

So that,

$$\sum_{f_1\in F_{1_{it}}} p\left\{\Theta_t/\theta_{it}, Y(k), Y^{k-1}, \widehat{X}_t(k/k-1), P_t(k/k-1)\right\} = \sum_{f_1\in F_{1_{it}}} \left(\prod_{\substack{r=0,t_r=0\\ r\neq i,t_r\neq t}}^{m_k,T} \varepsilon_{rt_r}\right) = \prod_{\substack{t_r=0\\ t_r\neq t}}^{T} \sum_{\substack{r=0\\ r\neq t}}^{m_k} \varepsilon_{rt_r}$$

Then substituting Eqs. 2.122 and 2.124 into Eq. 2.121, we can get:

$$\sum_{f_1\in F_{1_{it}}} p\left\{\theta_{it}/\Theta_t, Y^k\right\} = \varepsilon_{it}\cdot \prod_{\substack{t_r=0\\ t_r\neq t}}^{T} \sum_{\substack{r=0\\ r\neq t}}^{m_k} \varepsilon_{rt_r} \tag{2.125}$$

where $i = 0, 1, \ldots, m_k,\ t = 0, 1, \ldots, T$.

Likewise we can get:

$$\sum_{f_2\in F_{1_{it}}} p\left\{\theta_{it}/\Theta_t, Y^k\right\} = \varepsilon'_{it}\cdot \prod_{\substack{r=0\\ r\neq t}}^{m_k} \sum_{\substack{t_r=0\\ t_r\neq t}}^{T} \varepsilon'_{rt_r} \tag{2.126}$$

Substituting Eqs. 2.125 and 2.126 into Eq. 2.120, we can get:

$$\sum_{f\in F_{it}} p\left\{\theta_{it}/\Theta_t, Y^k\right\} = \varepsilon_{it}\cdot \prod_{\substack{t_r=0\\ t_r\neq t}}^{T} \sum_{\substack{r=0\\ r\neq t}}^{m_k} \cdot\varepsilon_{rt_r} + \varepsilon'_{it}\cdot \prod_{\substack{r=0\\ r\neq t}}^{m_k} \sum_{\substack{t_r=0\\ t_r\neq t}}^{T} \cdot\varepsilon'_{rt_r} \tag{2.127}$$

To guarantee the completeness of the probability of target t, we need to again perform normalization to the same target on the basis of Eq. 2.127 to compute the mutual belonging probability. Then, the mutual belonging probability of target t is:

$$\beta_{it} = \frac{1}{c}\sum_{f\in F_{it}} p\{\theta_{it}/\Theta_t, Y^k\} = \frac{1}{c}\left(\varepsilon_{it}\cdot\prod_{\substack{t_r=0\\t_r\neq t}}^{T}\sum_{\substack{r=0\\r\neq t}}^{m_k}\cdot\varepsilon_{rt_r} + \varepsilon'_{it}\cdot\prod_{\substack{r=0\\r\neq t}}^{m_k}\sum_{\substack{t_r=0\\t_r\neq t}}^{T}\cdot\varepsilon'_{rt_r}\right) \tag{2.128}$$

where c is the normalization coefficient for guaranteeing the completeness of measurements within target validation gates.

2.5 Summary

Any new algorithm cannot come out of nothing, so is the group-target tracking algorithm. This chapter first gives the common glossary of multi-target tracking, to provide a reference for the definitions of relevant glossary of multi-target. Then, this chapter gives the deduction procedure about relevant equations of Kalman filtering which is needed by group-target tracking. As for the problem of target maneuver detection and maneuver judgment, based on the depiction on the basic issues of target maneuver detection, this chapter introduces basic mobility models and their three kinds of expressions. Finally, this chapter gives the basic deduction procedures about the nearest-neighboring algorithm, the probabilistic data association algorithm, the joint probabilistic data association algorithm, and the generalized probabilistic data association algorithm, which have important enlightening meaning to group-target tracking, and then introduces the essential method and idea of their implementation. This founds the understanding of group-target and tracking algorithms.

http://www.springer.com/978-981-10-1886-2

Group-target Tracking
Geng, W.-d.; Wang, Y.-q.; Dong, Z.-h.
2017, XIII, 168 p. 40 illus., 31 illus. in color., Hardcover
ISBN: 978-981-10-1886-2