IET Radar, Sonar & Navigation

Review Article

IET Journals
The Institution of Engineering and Technology

Survey of amplitude-aided multi-target tracking methods

ISSN 1751-8784
Received on 18th June 2018
Revised 30th August 2018
Accepted on 18th September 2018
E-First on 25th October 2018
doi: 10.1049/iet-rsn.2018.5064
www.ietdl.org

Seung-Hwan Bae[1] ✉

[1]*Department of Computer Science and Engineering, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon 22012, Republic of Korea*

✉ *E-mail: shbae@inu.ac.kr*

Abstract: An amplitude-aided multi-target tracking (MTT) exploits amplitude as well as spatial features for MTT. Compared to MTT using a spatial feature only, it can maintain its performance even in densely cluttered environment because more accurate association is achieved by using both features for likelihood evaluation between tracks and measurements. Therefore, the goal of this study is to review the state-of-the-art amplitude-aided MTT methods and compare each other extensively. For a fair comparison, a unified MTT framework is developed, and various methods are implemented and compared based on the same framework. On the challenging visual MTT datasets, the implemented methods are evaluated for several aspects such as accuracy, speed, sensitivity, and stability. Moreover, this study presents a summary of the extensive evaluation and guideline to select an appropriate method for amplitude-aided tracking.

1 Introduction

Multi-target tracking (MTT) is to find states (positions, velocities, or sizes) of multiple targets with unknown object quantities and behaviours with given a set of sensor measurements. It plays an important role in many research areas such as radar-based aircraft tracking, sonar-based sea-animal tracking, and image-based object tracking. In most cases, the measurement origin is also unknown since a sensor usually receives mixed signals reflected from both targets and random clutter. Therefore, to distinguish the returns of the actual targets from those of background clutter and other interference, a detection algorithm such as a constant false alarm rate algorithm [1] is applied, and then a refined measurement that contains spatial information such as range, bearing, and Doppler velocity is generated. To date, however, false positives and missing detections are unavoidable since the performance of detectors is still incomplete. Thus, to successfully solve the MTT problem, a data association process is required, which would be able to correctly associate measurements with the corresponding tracks.

For the association between tracks and measurements, many association methods have been developed during the last decades. Simple greedy association methods such as the nearest neighbourhood [2] and the strongest neighbour [3] association can be performed rapidly, but they do not properly consider the cases in which multiple measurements originating from multiple targets are scattered. In order to find optimal solutions for a joint track-to-measurement assignment problem through single-frame and multi-frame searches, joint probabilistic data association [2] and multiple hypothesis tracking (MHT) [4] have been proposed. However, they incur large computational costs as the number of possible joint assignments combinatorially increases due to the number of tracks and measurements. In an attempt to alleviate the association complexity, linear multi-target integrated probabilistic data association (LMIPDA) [5] has been developed.

In addition, sequential Monte Carlo (SMC) methods for MTT have been developed. They can be divided into an association-based tracking [6–9] and association free tracking [10–12]. The first one decomposes a complex MTT problem into data association and state estimation problems based on a divide-and-conquer approach. Therefore, a data association is key to approximate a posterior distribution for each target. On the other hand, the latter one estimates states of all targets in a Bayesian framework without data association. Thus, each sample is a hypothesis on the number of targets present and the states of those targets. Although the methods for MTT are applied, association failures occur often in practical scenarios when targets are closely spaced or clutter is densely distributed in the target vicinity. In these situations, exploiting only spatial information is not sufficient for discriminating between target and clutter measurements.

In practical sensors such as a radar and a sonar, amplitude feature as well as spatial feature such as range, bearing, and Doppler velocity can be measured. In addition, the signal amplitude from a target is usually stronger than that from false alarms (*or* clutters). Therefore, we can exploit the amplitude feature for identifying between targets and clutters. Based on this finding, the authors in [13–19] utilise the amplitude information for the better association. In these studies, the motion likelihood and also amplitude likelihood are used for associating tracks and measurements. For improving single-target tracking in a cluttered environment, the authors in [13, 15] present the probabilistic data association filter including amplitude feature, and the highest probabilistic data association [13], respectively. To extend it for MTT, amplitude has been incorporated into MHT [16] and Viterbi data association [14]. In addition, for tracking multiple targets with different signal-to-noise ratios (SNRs), amplitude was also incorporated into a framework of finite set statistics [17]. Recently, the authors in [18, 19] developed SNR estimation methods with amplitude and presented LMIPDA with amplitude information (LMIPDA-AI) for more robust association.

However, it is still challenging to determine which method is the most appropriate for each application since a complete survey for amplitude-aided MTT methods is not conducted. Although Ehrman and Blair [20] present some comparisons of MTT methods using target amplitude, the discussed methods can be applicable only with the assumption that target SNRs are known and fixed during tracking. However, in many practical applications, the target SNR information may not be available beforehand. In addition, even though knowledge of the (initial) SNR is known beforehand, the target SNR should be updated consistently during tracking since it fluctuates due to the following reasons: (i) the received signal of the target is faded by propagation and attenuation through a medium [15]; (ii) the returns of most targets are composed of the sum of reflected rays from individual scattering points [17, 21]; and (iii) the aspect angle of the target is changed by the target motion [22].

In this paper, we present MTT methods with amplitude which can be operated under unknown target SNR. We first introduce a SNR marginalisation method [17] which marginalises an amplitude likelihood over a range of possible SNRs. We then introduce two SNR estimation methods for estimating target SNRs with amplitude measurements. The first one estimates the values using the SMC [18] method. The other method estimates them based on a maximum posterior approach (MAP) [19]. Since these methods were implemented for different MTT frameworks, we implement each method based on a unified MTT framework and generate several MTT systems by combining different amplitude methods. In addition, we implement MTT system which is known SNRs during tracking to consider an ideal case. To investigate the effects of using both amplitude and spatial measurements for MTT, we also implement MTT systems with one of both measurements. For evaluating tracking systems, we use the stand metric, optimal subpattern assignment (OSPA) [23] and visual surveillance benchmark datasets for MTT. The continuations of this paper can be summarised as follows:

- A concise survey of modern MTT methods using amplitude under an known and unknown SNR environment.
- A unified framework based on SMC to implement various amplitude-aided tracking methods.
- Extensive evaluation of various methods for the several aspects on the challenging visual MTT datasets.
- A practical guide to select a suitable method for amplitude-added MTT.

2 Problem definition

A non-linear discrete-time dynamic motion is used to model the behaviour of a target τ as follows:

$$\boldsymbol{x}_k^\tau = f_k(\boldsymbol{x}_{k-1}^\tau) + \boldsymbol{v}_{k-1}, \quad k = 1, 2, \ldots \tag{1}$$

where $\boldsymbol{x}_k^\tau \in \mathbb{R}^4$ denotes states of a target τ at time instant k composed of positions $(x_{1,k}^\tau, x_{2,k}^\tau)$ and velocities $(x_{3,k}^\tau, x_{4,k}^\tau)$ along with x and y coordinates, respectively. f_k is a non-linear function of the state $\boldsymbol{x}_{k-1}$ and $\boldsymbol{v}_{k-1} \sim \mathcal{N}(0, Q_k)$ is white Gaussian system noise. The initial state $\boldsymbol{x}_0^\tau$ is assumed to be Gaussian $\mathcal{N}(m_0^\tau, P_0)$ with the covariance P_0, where $m_0 = E(\boldsymbol{x}_0^\tau)$ and $P_0 = \mathrm{cov}\{\boldsymbol{x}_0^\tau, \boldsymbol{x}_0^\tau\}$.

In general, the measurement set obtained at one scan is composed of many measurements originated from multiple targets and clutter [2, 21]. Let us denote a set of measurements at scan k as $\mathbb{Z}_k = \{\boldsymbol{z}_{i,k}\}_{i=1}^{m_k}$, where the measurement vector $\boldsymbol{z}_{i,k} = [r_{i,k}, \theta_{i,k}, a_{i,k}]^\mathrm{T} = [\boldsymbol{y}_{i,k}, a_{i,k}]^\mathrm{T}$ is composed of range $r_{i,k}$, bearing $b_{i,k}$, and amplitude $a_{i,k}$ elements. As in probabilistic data association filter (PDAF) [2], the gating technique is used to reduce matching combinations between tracks and measurements. Using the gating technique, m_k^τ validated measurement in the gate of the track τ is determined by

$$\mathbb{Z}_k^\tau = \left\{\boldsymbol{z}_{i,k}^\tau : (\boldsymbol{v}_{i,k}^\tau)^\mathrm{T} (\boldsymbol{S}_k^\tau)^{-1} (\boldsymbol{v}_{i,k}^\tau) \le \gamma\right\}, \quad i = 1, \ldots, m_k^\tau, \tag{2}$$

where γ is a gate threshold and m_k^τ is the number of measurements in the gate of the track τ; $\boldsymbol{v}_{i,k}^\tau = \boldsymbol{y}_{i,k}^\tau - \bar{\boldsymbol{y}}_{k|k-1}^\tau$, is a zero-mean Gaussian residual with covariance $\boldsymbol{S}_k^\tau$. Given the predicted sample measurements $\boldsymbol{y}_{k|k-1}^{n,\tau}$ with weights $w_{k-1}^{n,\tau}$ from (37), the predicted measurement $\bar{\boldsymbol{y}}_{k|k-1}^\tau$ and the innovation covariance $\boldsymbol{S}_k^\tau$ are empirically calculated as

$$\begin{aligned} \bar{\boldsymbol{y}}_{k|k-1}^\tau &= \sum_{n=1}^{N} w_{k-1}^{n,\tau} \boldsymbol{y}_{k|k-1}^{n,\tau}, \boldsymbol{y}_{k|k-1}^{n,\tau} = h_k(\boldsymbol{x}_{k|k-1}^{n,\tau}) \\ \boldsymbol{S}_k^\tau &= \sum_{n=1}^{N} w_{k-1}^{n,\tau} (\boldsymbol{y}_{k|k-1}^{n,\tau} - \bar{\boldsymbol{y}}_{k|k-1}^\tau)(\boldsymbol{y}_{k|k-1}^{n,\tau} - \bar{\boldsymbol{y}}_{k|k-1}^\tau)^\mathrm{T}. \end{aligned} \tag{3}$$

Given the gated measurements, amplitude thresholding is exploited to filter out false alarms with the threshold DT because the amplitude from a target is usually stronger than false alarms [17].

Furthermore, a target-originated measurement $\boldsymbol{z}_{i,k}^\tau$ is modelled by a non-linear measurement model as

$$\boldsymbol{y}_{i,k}^\tau = h_k(\boldsymbol{x}_k^\tau) + w_k = \begin{bmatrix} \sqrt{(x_{1,k}^\tau)^2 + (x_{2,k}^\tau)^2} \\ \tan^{-1}\left(\dfrac{x_{2,k}^\tau}{x_{1,k}^\tau}\right) \end{bmatrix} + \begin{bmatrix} w_{r,k} \\ w_{\theta,k} \end{bmatrix}, \tag{4}$$

where the range noise $w_{r,k} \sim \mathcal{N}(0, \sigma_r^2)$ and the bearing noise $w_{\theta,k} \sim \mathcal{N}(0, \sigma_\theta^2)$ are uncorrelated Gaussian noise sequence. Here, it is assumed that the spatial $\boldsymbol{y}_{i,k}^\tau$ and amplitude $a_{i,k}^\tau$ measurements are independent of each other.

3 Modelling of amplitude likelihood

Probabilistic amplitude models for a target and false alarm (*or* clutter) are presented. These models are modelled with amplitude and target SNR. To model these amplitude likelihood functions under unknown target SNR environment, a marginalised amplitude likelihood function within a certain SNR range is described. Then, SNR estimation methods, which can infer SNR with measured amplitude at each scan, are discussed.

3.1 Target and clutter amplitude likelihood models

Let us assume that amplitude is the output of a bandpass matched filter that has an envelope detector attached. In this case, the probability density of the amplitude a follows a Rayleigh distribution [For clarity, the set, matrix, vector, and scalar are denoted by blackboard bold font, upper case boldface, boldface, and standard italic types, e.g. $\mathbb{A}, \boldsymbol{A}, \boldsymbol{a}, a$], as described in [15, 21]. The amplitude probability density of false detections (*or* clutter) can be expressed as

$$p_c(a) = \frac{a}{\sigma^2} \cdot \exp\left(\frac{-a^2}{2\sigma^2}\right), \quad a \ge 0, \tag{5}$$

where σ^2 is the variance (*or* power) of the in-phase and quadrature components (x_s, y_s) [15] of the narrow band noise coming out of the matched filter, where each component is assumed to be Gaussian but independent of each other. Here, $x_s, y_s \sim \mathcal{N}(0, \sigma^2)$ and the amplitude a is defined as $a = \sqrt{x_s^2 + y_s^2}$. Note that the amplitude density function is the representation of the power (σ^2) of each component. However, as discussed in [24], in the narrow band filter, the amplitude density function of the noise is modelled with the average (*or* total) noise power σ_{noise}^2 rather than the power σ^2 of the component since the receiver bandwidth-to-centre frequency ratio is usually small. Thus, it can be reformulated with average noise power σ_{noise}^2, where $\sigma_{\text{noise}}^2 = 2\sigma^2$:

$$p_0(a) = \frac{2a}{\sigma_{\text{noise}}^2} \cdot \exp\left(\frac{-a^2}{\sigma_{\text{noise}}^2}\right), \quad a \ge 0, \tag{6}$$

where the background noise is normalised as in [15, 17]. This means that the variance of the noise (6) is $\sigma_{\text{noise}}^2 = 1$, and the expected noise power N_0 is unity as

$$N_0 = E[a^2] = \int_0^\infty a^2 p_0(a)\, da = 1. \tag{7}$$

Let us define the expected (or mean) SNR [The SNR is represented in log scale: $\mathrm{SNR(dB)} = 10\log_{10}(d)$] $d = S/N_0$, where S is the signal power and d can be treated as the expected target signal power because $N_0 = 1$. In addition, a slow Rayleigh fading amplitude-modulated narrowband signal is considered in the presence of narrowband noise. In this case, the signal returned

from the target is expressed as the sum of the transmitted signal and the narrow band noise. Also, as described above, the noise has the normal distribution with zero mean and unit variance (i.e. $\sigma^2_{\text{noise}} = 1$). Therefore, the amplitude density function of a target follows Rayleigh distribution with the variance $1 + d$ (*i.e.* the SNR)

$$p_1(a, d) = \frac{2a}{1+d} \cdot \exp\left(\frac{-a^2}{1+d}\right). \tag{8}$$

However, to evaluate the signal power S from the target amplitude distribution (8), the expected target SNR d is needed to estimate because

$$S = E\left[a^2\right] = \int_0^\infty a^2 p_1(a, d) da\,. \tag{9}$$

Let us next consider the case in which the amplitude a exceeds a detection threshold DT, i.e. $a \geq DT$. Then, the amplitude density of the target becomes

$$p_1^{DT}(a, d) = \frac{1}{P_D} p_1(a, d) = \frac{2a}{1+d} \cdot \exp\left(\frac{DT^2 - a^2}{1+d}\right), \quad a \geq DT, \tag{10}$$

where the target detection probability P_D used for normalisation is calculated as

$$P_D \quad = \int_{DT}^\infty p_1(a, d) da = \exp\left(\frac{-DT^2}{(1+d)}\right). \tag{11}$$

Otherwise the amplitude probability density of false alarms is expressed from (6)

$$p_0^{DT}(a) = \frac{1}{P_{FA}} p_0(a) = 2a \cdot \exp\left(DT^2 - a^2\right), \quad a \geq DT, \tag{12}$$

where the clutter detection probability P_{FA} is given as

$$P_{\text{FA}} \quad = \int_{DT}^\infty p_0(a) da = \exp\left(-DT^2\right). \tag{13}$$

When the target SNR d is known, the amplitude likelihoods of both the target and clutter can then be computed as

$$\textbf{Target}: \quad g_a^{DT}(a|d) = p_1^{DT}(a, d), \tag{14}$$

$$\textbf{Clutter}: \quad c_a^{DT} = p_0^{DT}(a)\,. \tag{15}$$

To exploit the amplitude in the general case without the assumption of the known SNR, a target SNR marginalisation method is presented in Section 3.2. Then, SNR estimation methods are provided to find d during tracking in Section 3.3. Once the likelihood functions $g_a^{DT}(a|d)$ and c_a^{DT} are evaluated, the amplitude feature can be used to discriminate different targets and/or clutter. In Section 4, a unified MTT framework is presented in order to incorporate and evaluate the presented SNR marginalisation and estimation methods.

3.2 Marginalisation of target amplitude likelihood

In an attempt to evaluate the target amplitude likelihood (14), Clark *et al.* [17] present the SNR marginalisation method. Let denote the density of the target SNR d defined within the boundary $[d_1, d_2]$ as $p(d)$. Then, the likelihood and detection probability can be defined as

$$\begin{aligned} g_a(a) &:= \int_{d_1}^{d_2} p(\delta) g_a(a|\delta)\, \mathrm{d}\delta \\ P_D(\tau) &:= \int_{d_1}^{d_2} p(\delta) P_D(\delta)\, \mathrm{d}\delta\,. \end{aligned} \tag{16}$$

To design the prior $p(d)$, an uninformative prior is applied using the Fisher information. In Bayesian statistics, the uninformative prior is proportional to the square root of the Fisher information: $p(d) \propto \sqrt{I(d)}$. Let us consider amplitude samples, $\{a_i\}_{i=1}^n$, which are drawn independent identically distributed from the target likelihood $g_a^{DT}(a|d)$ (14). Then, it can be represented as

$$g_a(a_1, a_2, \ldots, a_n, d) = \prod_{i=1}^n g_a(a_i, d) \tag{17}$$

According to the Fisher information theory [25], the variance of any unbiased estimator $\text{var}(\hat{d})$ is bounded by the inverse of $I(d)$

$$\text{var}(\hat{d}) \geq \frac{1}{I(d)}, \quad I(d) = E_d\left[\left(\frac{\partial(\ln(\prod_{i=1}^n g_a(a_i|d)))}{\partial d}\right)^2\right] \tag{18}$$

Since the target likelihood $g_a^{DT}(a|d)$ is a product of Rayleigh distributions from (17), the Fisher information $I(d)$ is

$$I(d) = \frac{n}{(1+d)^2} \tag{19}$$

Then, $p(d)$ for the single object likelihood with one measurement is represented with $I(d)$ (19)

$$p(d) \propto \frac{1}{(1+d)} \tag{20}$$

where $p(d)$ is normalised for the region $[d_1, d_2]$.

This prior provides a uniform distribution in the *dB* domain, which can be shown as follows: suppose that $u(\xi)$ is distributed uniformly in *dB* domain. Then, the integration of $u(\xi)$ over a range of dB values $[dB_1, dB_2]$ and conversion of it into d domain can be represented as

$$\int_{dB_1}^{dB_2} u(\xi) = \int_{\phi^{-1}(dB_1)}^{\phi^{-1}(dB_2)} u(\phi(\delta))\phi'(\delta)\, \mathrm{d}\delta \propto \int_{\phi^{-1}(dB_1)}^{\phi^{-1}(dB_2)} \phi'(\delta)\, \mathrm{d}\delta \tag{21}$$

where $\phi(d) = 10log_{10}(1+d)$ is the function which change d into *dB* values. Therefore, $p(d) = \phi'(d) \propto 1/(1+d)$ provides the required function in the d domain. In addition, by the prior $p(d)$ the analytic solution for the likelihood function g_a (16) marginalised over d can be derived as

$$g_a(a) = \frac{2\left(\exp\left(-a^2/(2(1+d_2))\right) - \exp\left(-a^2/(2(1+d_1))\right)\right)}{a(\ln(1+d_2) - \ln(1+d_1))}. \tag{22}$$

3.3 Unknown SNR estimation

In this section, SNR estimation methods with amplitude measurements are discussed. The prior to describe the behaviour of a target SNR is first introduced and then estimation methods using SMC [18] and MAP [19] methods are provided.

3.3.1 SNR prior model: To estimate the target SNR d, the prior knowledge of d is required. Therefore, the following two cases can be considered for modelling the prior $p(d)$:

- Case 1: no previous knowledge of the SNR is available.
- Case 2: an initial (*or* previously estimated) SNR is provided, but it should be updated due to its random fluctuation.

In the first case, let us consider that the SNR can be any value within any boundary $[d_1, d_2]$, and then the prior $p(d)$ can be modelled using the uniform distribution as follows:

$$p(d) = \begin{cases} \dfrac{1}{d_2 - d_1}, & d_1 \le d \le d_2, \\ 0, & \text{otherwise}. \end{cases} \tag{23}$$

In the second case, let us consider that the SNR is randomly fluctuated in the vicinity of the initial (*or* previously estimated) SNR $\hat{d}^{\tau}_{k-1}$. The prior $p(d)$ can be modelled by using the Gaussian random walk with the mean $\hat{d}^{\tau}_{k-1}$ and variance σ_d^2 as follows:

$$p(d) = \mathcal{N}\left(d; \hat{d}^{\tau}_{k-1}, \sigma_d^2\right), \quad \hat{d}^{\tau}_{k} \ge 0. \tag{24}$$

Once the motion of d is modelled with both prior (23) and (24), the unknown target SNR can be updated with the SMC and MAP methods discussed in the next section.

3.3.2 SMC-based SNR estimation: Suppose that unknown target SNR sequence d_k can be modelled based on the prior (24) as

$$d^{\tau}_{k+1} = d^{\tau}_{k} + \varphi_k, \quad k = 1, 2, \ldots \tag{25}$$

where φ_k is white Gaussian noise sequence represented by $\varphi_k \sim \mathcal{N}(0, \sigma_\varphi^2)$.

The unknown SNR d^{τ}_k of the target τ can be estimated recursively with the predication and update procedure of SMC as

$$\begin{aligned} &\textbf{Predict}: p(d^{\tau}_k | a^{\tau}_{1:k-1}) = \int p(d^{\tau}_k | d^{\tau}_{k-1}) p(d^{\tau}_{k-1} | a^{\tau}_{1:k-1}) d d^{\tau}_{k-1}, \\ &\textbf{Update}: p(d^{\tau}_k | a^{\tau}_{1:k}) \propto p(d^{\tau}_k | a^{\tau}_k) p(d^{\tau}_k | a^{\tau}_{1:k-1}). \end{aligned} \tag{26}$$

where the prior model of a target SNR (24) can be applied for prediction. The sample weight $w^{n,\tau}_k$ for the sample $d^{n,\tau}_k$ is evaluated with the amplitude likelihood function (14). Then, d^{τ}_k can be estimated with N samples and weights as

$$\hat{d}^{\tau}_k = \sum_{i=1}^{N} w^{n,\tau}_k d^{n,\tau}_k, \quad \sum_{i=1}^{N} w_k = 1, \quad k = 1, 2, \ldots \tag{27}$$

3.3.3 MAP-based SNR estimation: To estimate an unknown SNR more accurately, one can use several amplitude measurements. In other words, rather than inferring the target SNR with an instant amplitude measurement a^{τ}_k of target τ at scan k, it can be estimated with a set of amplitude measurements stacked during Δ scans.

Let us denote the stacked amplitude measurements from time $k - \Delta$ to time k as $a^{\tau}_{k-\Delta:k}$. The MAP problem of finding an optimal SNR with respect to the collection of amplitudes $a^{\tau}_{k-\Delta:k}$ can be modelled by

$$\begin{aligned} \hat{d}^{\tau}_k &= \underset{d}{\text{argmax}} \prod_{a^{\tau} \in a^{\tau}_{k-\Delta:k}} p(a^{\tau}, d), \quad d \ge 0, \\ &= \underset{d}{\text{argmax}} \prod_{a^{\tau} \in a^{\tau}_{k-\Delta:k}} p(a^{\tau} | d) p(d), \\ &= \underset{d}{\text{argmax}} \sum_{a^{\tau} \in a^{\tau}_{k-\Delta:k}} \log(p(a^{\tau} | d)) + \log(p(d)), \end{aligned} \tag{28}$$

where the first likelihood term $p(a^{\tau}|d)$ is given by (14).

To design the prior term $p(d)$, let us consider both cases mentioned in Section 3.3.1. In the first case, the prior $p(d)$ is given by (23), and then the MAP problem can be reformulated by substituting the prior of (28) with the distribution (23) as

$$\hat{d}^{\tau}_k = \underset{d}{\text{argmax}} \sum_{a^{\tau} \in a^{\tau}_{k-\Delta:k}} \log(p(a^{\tau} | d)) + \log(c), \quad c = \frac{1}{d_2 - d_1}, \tag{29}$$

As c is constant, the MAP problem (29) can be transformed into the maximum likelihood estimation

$$\hat{d}^{\tau}_k = \underset{d}{\text{argmax}} \sum_{a^{\tau} \in a^{\tau}_{k-\Delta:k}} \log(p(a^{\tau} | d)). \tag{30}$$

In the second case, the prior $p(d)$ using the Gaussian random walk with the mean $\hat{d}^{\tau}_{k-1}$ and variance σ_d^2 is represented by (24). In the similarity manner, by substituting the prior of (28) with (24), the following objective function can be derived:

$$\begin{aligned} \hat{d}^{\tau}_k &= \underset{d}{\text{argmax}} \sum_{a^{\tau} \in a^{\tau}_{k-\Delta:k}} \log(p(a^{\tau} | d)) + \log(c), \\ c &= \mathcal{N}\left(d; \hat{d}^{\tau}_{k-1}, \sigma_d^2\right). \end{aligned} \tag{31}$$

Optimising both objective functions (30) and (31) can be regarded as non-linear least-squares problems. To solve these problems, Bae *et al.* [19] use the Levenberg–Marquardt method [26].

4 MTT framework

In this section, we present a robust MTT system consisting of data association and track state update parts.

4.1 Data association with spatial and amplitude feature

We briefly discuss LMIPDA-AI and refer to [18] for further details. Let us consider the following two association events:

- $\chi^{\tau}_{0,k}$: None of the measurements are associated with track τ
- $\chi^{\tau}_{i,k}$: The measurement $\boldsymbol{z}^{\tau}_{i,k}$ is associated with track τ.

The posterior association probabilities for both events $\chi^{\tau}_{0,k}$ and $\chi^{\tau}_{i,k}$ can be evaluated by

$$\beta^{\tau}_{i,k} = \begin{cases} \dfrac{P\left(\chi^{\tau}_{0,k}, \chi^{\tau}_{k} | \mathbb{Z}^k, M_k\right)}{P\left(\chi^{\tau}_{k} | \mathbb{Z}^k, M_k\right)} = \dfrac{1 - P^{\tau}_D P^{\tau}_G}{1 - \Psi^{\tau}_k}, & i = 0, \\ \dfrac{P\left(\chi^{\tau}_{i,k}, \chi^{\tau}_{k} | \mathbb{Z}^k, M_k\right)}{P\left(\chi^{\tau}_{k} | \mathbb{Z}^k, M_k\right)} = \dfrac{P^{\tau}_D P^{\tau}_G}{1 - \Psi^{\tau}_k} \cdot \dfrac{\Lambda^{\tau}_{i,k}}{\Phi^{\tau}_{i,k}}, & i = 1, \ldots, m^{\tau}_k, \end{cases} \tag{32}$$

where P^{τ}_D and P^{τ}_G are the target detection and gate probability, respectively.

Then, using the spatial and amplitude features, the likelihood model $\Lambda^{\tau}_{i,k}$ can be acquired based on the assumption of independent spatial measurement $\boldsymbol{y}^{\tau}_{i,k}$ and amplitude measurement $a^{\tau}_{i,k}$. Consequently, $\Lambda^{\tau}_{i,k}$ is derived such that

$$\begin{aligned} \Lambda^{\tau}_{i,k} &\equiv p\left(\boldsymbol{z}^{\tau}_{i,k} | \chi^{\tau}_{k}, \mathbb{Z}^{k-1}\right) = p\left(\boldsymbol{y}^{\tau}_{i,k} | \chi^{\tau}_{k}, \mathbb{Y}^{k-1}\right) \\ &\quad \times p\left(a^{\tau}_{i,k} | \chi^{\tau}_{k}, \mathbb{A}^{k-1}\right) = \Lambda^{p,\tau}_{i,k} \cdot \Lambda^{a,\tau}_{i,k}, \end{aligned} \tag{33}$$

where the spatial likelihood function $\Lambda^{p,\tau}_{i,k}$ and the amplitude likelihood function $\Lambda^{a,\tau}_{i,k}$ become

$$\begin{aligned} \Lambda^{p,\tau}_{i,k} &= \mathcal{N}\left(\boldsymbol{v}^{\tau}_{i,k}, \boldsymbol{S}^{\tau}_{k}\right) \\ &= |2\pi \boldsymbol{S}^{\tau}_{k}|^{-1/2} \exp\left(-\frac{1}{2} (\boldsymbol{v}^{\tau}_{i,k})^{\mathrm{T}} (\boldsymbol{S}^{\tau}_{k})^{-1} \boldsymbol{v}^{\tau}_{i,k}\right), \\ \Lambda^{a,\tau}_{i,k} &= g^{DT}_a(a^{\tau}_{i,k} | \hat{d}^{\tau}_k), \end{aligned} \tag{34}$$

Note that the amplitude likelihood $\Lambda_{i,k}^{a,\tau}$ can be modelled in different ways. From the marginalised distribution (22), $\Lambda_{i,k}^{a,\tau}$ can be represented within a certain interval $[d_1, d_2]$. In addition, one can infer (10) by estimating the target SNR $\hat{d}_k^{\tau}$ from (27), (30) or (31). The innovation covariance $\boldsymbol{S}_k^{\tau}$ and the residual $\boldsymbol{v}_{i,k}^{\tau}$ of the track are computed using (3). $\Psi_k^{\tau} = P_D^{\tau} P_G^{\tau} \cdot \left(1 - \sum_{i=1}^{m_k^{\tau}} \Lambda_{i,k}^{\tau} / \Phi_{i,k}^{\tau}\right)$ is obtained using the target likelihood (33) and scatterer models. $\Phi_{i,k}^{\tau}$ is probability that measurement $\boldsymbol{z}_{i,k}^{\tau}$ is originated from scatterer [18]. The track existence probability is predicted and updated as

$$\begin{aligned} \textbf{Predict}: P\left(\chi_k^{\tau} | \mathbb{Z}^{k-1}\right) &= \alpha_{11} P\left(\chi_{k-1}^{\tau} | \mathbb{Z}^{k-1}\right) \\ &\quad + \alpha_{21}\left(1 - P(\chi_{k-1}^{\tau} | \mathbb{Z}^{k-1})\right), \\ \textbf{Update}: P\left(\chi_k^{\tau} | \mathbb{Z}^{k}\right) &= \frac{\left(1 - \Psi_k^{\tau}\right) \cdot P\left(\chi_k^{\tau} | \mathbb{Z}^{k-1}\right)}{1 - \Psi_k^{\tau} \cdot P\left(\chi_k^{\tau} | \mathbb{Z}^{k-1}\right)}, \end{aligned} \tag{35}$$

where transition probabilities $\alpha_{11} \equiv P(\chi_k^{\tau} | \chi_{k-1}^{\tau})$ and $\alpha_{21} \equiv P(\chi_k^{\tau} | \tilde{\chi}_{k-1}^{\tau})$ are set to 0.98 and 0.02.

4.2 Particle filtering for MTT

In particle filtering for MTT, the main issue is how to estimate posterior distributions of multiple targets' states with inaccurate detections. To handle this difficulty, the authors in [10, 11] present joint multitarget probability density for estimating target states and cardinality simultaneously. However, the computational complexity increases exponentially when the number of particles increases. To mitigate the curse-of-dimensionality problem, the authors in [6–9] solve the data association and state estimation problems separately. Based the Rao-Blackwellisation method, Särkkä *et al.* [7] used particle filtering and Kalman filtering for data association and state estimation, respectively. Ekmanm [6] presented data association method by combining probabilistic data association and nearest neighbour techniques. In [8], a game-theoretic framework is developed for deterministic association. Sample-based joint probabilistic data association [9] is provided for applying a ensemble square root filter for MTT. The reader is kindly referred to [27, 28] for more details of particle filtering for MTT.

In this study, we also present an association-based SMC method by using the data association with spatial and amplitude feature. Here, the states of each target can be modelled by linear or non-linear dynamic models.

Basically, the Monte Carlo filter (*or* particle filter) [29, 30] performs the following two-step recursion procedure:

$$\begin{aligned} \textbf{Predict}: p(\boldsymbol{x}_k^{\tau} | \boldsymbol{z}_{1:k-1}^{\tau}) &= \int p(\boldsymbol{x}_k^{\tau} | \boldsymbol{x}_{k-1}^{\tau}) p(\boldsymbol{x}_{k-1}^{\tau} | \boldsymbol{z}_{1:k-1}^{\tau}) \, \mathrm{d}\boldsymbol{x}_{k-1}^{\tau}, \\ \textbf{Update}: p(\boldsymbol{x}_k^{\tau} | \boldsymbol{z}_{1:k}^{\tau}) &\propto p(\boldsymbol{z}_k^{\tau} | \boldsymbol{x}_k^{\tau}) p(\boldsymbol{x}_k^{\tau} | \boldsymbol{z}_{1:k-1}^{\tau}) . \end{aligned} \tag{36}$$

The recursion requires a motion model $p(\boldsymbol{x}_k^{\tau} | \boldsymbol{x}_{k-1}^{\tau})$ and a likelihood model $p(\boldsymbol{z}_k^{\tau} | \boldsymbol{x}_k^{\tau})$. To deal with complicated distributions which are analytically intractable, the particle filter approximates the two steps using a set of weighted samples $\{\boldsymbol{x}_k^{n,\tau}, w_k^{n,\tau}\}_{n=1}^{N}$, where N is the number of particles.

When measurements $\boldsymbol{z}_{1:k}^{\tau}$ of the track τ up to scan k are provided, all states of the target up to scan k (i.e. trajectory) $\boldsymbol{x}_{1:k}^{\tau}$ can be updated well by (36). In most MTT scenarios, however, it is not easy to identify origin of measurements because the detection responses are often unreliable (e.g. false positive, missing and inaccurate detections) and the responses of other targets are present.

To update the states of a track with unreliable measurements, we need to select a measurement $\boldsymbol{z}_k^{\tau}$ corresponding to the track τ in the set $\mathbb{Z}_k$. In this study, we select $\boldsymbol{z}_k^{\tau}$ among $\mathbb{Z}_k^{\tau} = \{\boldsymbol{z}_{i,k}^{\tau}\}_{i=1}^{m_{a,k}^{\tau}}$ according to the posterior association probability $\beta_{i,k}^{\tau}$ (32). More specifically, when the number of particles is $N = 100$, 100 measurements are generated by random selection in the set $\mathbb{Z}_k^{\tau}$ according to $\beta_{i,k}^{\tau}$. This procedure is rather similar to resampling in particle filtering: a measurement with a high value $\beta_{i,k}^{\tau}$ is more likely to be selected in the set. By incorporating this measurement selection step, we extend the two-step recursion (36) of the conventional particle filtering to a three-step procedure as follows:

$$\begin{aligned} \textbf{Predict}: p(\boldsymbol{x}_k^{\tau} | \boldsymbol{z}_{1:k-1}^{\tau}) &= \int p(\boldsymbol{x}_k^{\tau} | \boldsymbol{x}_{k-1}^{\tau}) p(\boldsymbol{x}_{k-1}^{\tau} | \boldsymbol{z}_{1:k-1}^{\tau}) \, \mathrm{d}\boldsymbol{x}_{k-1}^{\tau}, \\ \textbf{Select}: \boldsymbol{z}_k^{\tau} &\sim \beta_{i,k}^{\tau}, \ \boldsymbol{z}_k^{\tau} \in \{\boldsymbol{z}_{i,k}^{\tau}\}_{i=1}^{m_{a,k}^{\tau}}, \\ \textbf{Update}: p(\boldsymbol{x}_k^{\tau} | \boldsymbol{z}_{1:k}^{\tau}) &\propto p(\boldsymbol{z}_k^{\tau} | \boldsymbol{x}_k^{\tau}) p(\boldsymbol{x}_k^{\tau} | \boldsymbol{z}_{1:k-1}^{\tau}) . \end{aligned} \tag{37}$$

where, for sample prediction, we exploit the dynamic system model (1). The sample weight is evaluated with the selected measurement using the likelihood model (4)

$$\begin{aligned} p(\boldsymbol{x}_k^{\tau} | \boldsymbol{x}_{k-1}^{n,\tau}) &= \mathcal{N}(\boldsymbol{x}_k^{\tau}; \boldsymbol{x}_{k|k-1}^{n,\tau}, \boldsymbol{Q}_k), \quad \boldsymbol{x}_{k|k-1}^{n,\tau} = \boldsymbol{f}_k(\boldsymbol{x}_{k-1}^{n,\tau}), \\ p(\boldsymbol{z}_k^{\tau} | \boldsymbol{x}_k^{n,\tau}) &= \mathcal{N}(\boldsymbol{z}_k^{\tau}; \boldsymbol{z}_{k|k-1}^{n,\tau}, \boldsymbol{R}_k), \quad \boldsymbol{z}_{k|k-1}^{n,\tau} = \boldsymbol{h}_k(\boldsymbol{x}_{k|k-1}^{n,\tau}) . \end{aligned} \tag{38}$$

By incorporating the above recursion (37), we present an overall algorithm to implement the amplitude-aided MTT methods discussed in Algorithm 1 (see Fig. 1).

5 Experimental results

In this section, we compare different MTT systems on the challenging visual surveillance datasets.

5.1 Implementation

Based on Algorithm 1 (Fig. 1), we have implemented and compared the MTT systems (TR1–TR6) using spatial and amplitude measurements in a different way. For this comparison, based on the same MTT framework using LMIPDA(-AI) and particle filtering mentioned in Section 4, we have implemented the following MTT systems by combining different methods discussed in Section 3:

- (TR1) with spatial measurement;
- (TR2) with amplitude measurement based on strongest neighbour data association [3];
- (TR3) with spatial and amplitude measurement based on the known SNR [15];
- (TR4) with spatial and amplitude measurement based on the marginalisation method [17];
- (TR5) with spatial and amplitude measurement based on the SMC method [18];
- (TR6) with spatial and amplitude measurement based on the MAP method [19].

Here, the system (TR1) only uses spatial measurement (i.e. range and angle) without using amplitude. (TR2) associates a measurement with the strongest amplitude in each track gate as done in [3]. Systems (TR3)–(TR6) use both spatial and amplitude measurements. (TR3) exploits known SNR d and evaluate the amplitude likelihood (14) with known d. On the other hand, in (TR4)–(TR6), the SNR d is assumed to be unknown. (TR4) models the marginalised likelihood function (22) with marginalised SNRs over a range $[dB_1, dB_2]$. $dB_1 = 0$ (dB) and $dB_2 = 30$ (dB) is set to determine the prior of SNR used for (20) and (23). For (TR5) and (TR6), we estimate SNRs $\hat{d}_k^{\tau}$ for each target. (TR5) estimates $\hat{d}_k^{\tau}$ using the SMC method (27). We generate 59 samples uniformly from 1 to 30 dB with interval 0.5 dB. In (TR6), we estimate it using the objective function (31) and the variance σ_d^2 is set to 5. We set $\Delta = 5$ when solving (31).

Input : A set of measurements: $\mathbb{Z}_k$ and a set of trackers $\{\Omega_{k-1}^{\tau}\}_{\tau=1}^{M_k}$, where each tracker is composed of $\Omega_k^{\tau} = \left\{\mathbf{x}_{k-1}^{n,\tau}, w_{k-1}^{n,\tau}, d_{k-1}^{\tau}\right\}_{n=1}^{N}$.

Output: Updated trackers $\{\Omega_k^{\tau}\}_{\tau=1}^{M_k}$, $\Omega_k^{\tau} = \left\{\mathbf{x}_k^{n,\tau}, w_k^{n,\tau}, d_k^{\tau}\right\}$

for *TR1* **to** *TR6* **do**
  // **Step1:** Select a set of validated measurements $\mathbf{Z}_k^{\tau}$;
  for $\tau \leftarrow 1$ **to** M_k **do**
    $\mathbf{Z}_k^{\tau} = \{\mathbf{z}_{i,k}^{\tau} : (\mathbf{v}_{i,k}^{\tau})^{\mathrm{T}}(\mathbf{S}_k^{\tau})(\mathbf{v}_{i,k}^{\tau})\} \leq \gamma, a_{i,k} \geq DT$ by (2)
  end
  // **Step2:** Evaluate likelihoods;
  for $\tau \leftarrow 1$ **to** M_k **do**
    if *TR4* **then**
      Marginalize an amplitude likelihood $g_a(a)$ by (22)
    end
    if *TR5* **then**
      Estimate a target SNR d_k^{τ} by (27)
    end
    if *TR6* **then**
      Estimate a target SNR d_k^{τ} by (28)
    end
    for $i \leftarrow 1$ **to** m_k^{τ} **do**
      Evaluate a likelihood $\Lambda_{i,k}^{\tau}$ by (33)
    end
  end
  // **Step3:** Evaluate track existence and association probabilities;
  for $\tau \leftarrow 1$ **to** M_k **do**
    for $i \leftarrow 1$ **to** m_k^{τ} **do**
      Evaluate a posterior association probability $\beta_{i,k}^{\tau}$ by (32)
    end
    Update a track existence probability $P\left(\chi_k^{\tau}|\mathbb{Z}^k\right)$ by (35)
  end
  // **Step 4:** Particle filtering in Sec. 4.2 ;
  for $\tau \leftarrow 1$ **to** M_k **do**
    for $n \leftarrow 1$ **to** N **do**
      Prediction: $\mathbf{x}_k^{n,\tau} \sim p\left(\mathbf{x}_k^{\tau}|\mathbf{x}_{k-1}^{n,\tau}\right)$
      Selection: $\mathbf{z}_k^{\tau} \sim \beta_{i,k}^{\tau},\ \mathbf{z}_k^{\tau} \in \mathbb{Z}_k^{\tau}$
      Update: $w_k^{n,\tau} \leftarrow w_{k-1}^{n,\tau} \cdot p\left(\mathbf{z}_k^{\tau}|\mathbf{x}_k^{n,\tau}\right)$
    end
    $w_k^{n,\tau} \leftarrow \frac{w_k^{n,\tau}}{\sum_{j=1}^{N} w_k^{j,\tau}}$;
    Resampling$(\{\mathbf{x}_k^{n,\tau}, w_k^{n,\tau}\}_{n=1}^{N})$;
  end
end

Fig. 1 *Algorithm 1: the overall algorithm for MTT systems (TR1)–(TR6)*

5.2 Evaluation metric

As a performance measure, the OSPA metric [23] is used. Given the true and estimated sets composed of states of multiple targets, we measure the *localisation distance* and *cardinality distance*. The localisation distance evaluates the state similarities between matched pairs in the true and estimated sets. On the other hand, the cardinality distance evaluates how well the number of existing tracks matches the number of true targets. As an overall performance measure, the *OSPA distance* representing the total error is calculated by summing both the localisation and cardinality distances. For all the distance metrics, a smaller distance indicates better results.

In the OSPA metric, the cut-off parameter is set to $c = 100$, which determines the relative weighting of penalties assigned to the cardinality and localisation errors. The order parameter then is set to $p = 1$ which determines the sensitivity of the metric to outliers.

5.3 Visual multi-object tracking dataset

To compare systems (TR1–TR6) in real MTT environment, we use the publicly available VS-PETS 2009 benchmark dataset [31]. In the dataset, PETS S2.L1 and PETS S2.L2 sequences for MTT evaluation are exploited. PETS S2.L1 and S2.L2 sequences consist of 795 and 436 frames and the resolution of each image is 768(pixels) × 576(pixels). About 23 and 74 objects exist for PETS S2.L1 and S2.L2, respectively. As shown in Figs. 2*a* and 3*a*, the trajectories of multiple objects are complicated. In particular, PETS S2.L2 sequence is very challenging because many objects are moving and interacting with each other. We allocate each object to an initial SNR within (5 dB, 20 dB), and the object SNRs fluctuate at each scan according to the Gaussian distribution (24) with the variance $\sigma_d = 10$.

5.4 Detection

We first obtain detection results using a HOG detector [32] which was pre-trained for detecting pedestrians. Measurements of objects are assumed to detect with $P_D = 0.95$ and some detections for the objects are removed according to P_D. From each detection, spatial locations (x and y positions) and its sizes (width and height) are obtained. For each object SNR at frame k, amplitude measurements are generated according to Rayleigh distribution (9).

For each sequence, we generate more clutters with various clutter density: $\lambda = 6.78 \times 10^{-5}$, $\lambda = 1.13 \times 10^{-4}$, and $\lambda = 1.58 \times 10^{-4}$ (measurements/frame/pixel^2) for PETS L1, and $\lambda = 4.52 \times 10^{-5}$, $\lambda = 9.04 \times 10^{-5}$ and $\lambda = 1.36 \times 10^{-4}$ (measurements/frame/pixel^2) for PETS L2. As a result, from 20 to 70 clutters are produced randomly at each frame.

5.5 Tracking parameters

For a fair comparison, all the systems (TR1)–(TR6) use the same detections and tracking parameters. As a proposal density used for predicting particles, we use the prior of sample motion $p(\boldsymbol{x}_k^{\tau}|\boldsymbol{x}_{k-1}^{n,\tau}) = \mathcal{N}(\boldsymbol{x}_k^{\tau}; f(\boldsymbol{x}_{k-1}^{n,\tau}), \boldsymbol{Q}_k^{\tau})$, where f is in a general non-linear transfer function, and the noise covariance is set to

$$\boldsymbol{Q}_k = \mathrm{diag}[0.05(\mathrm{pixel})^2, 5(\mathrm{pixel/frame})^2, 0.05(\mathrm{pixel})^2, 5(\mathrm{pixel/frame})^2]^{\mathrm{T}}$$

The thresholds of gate and amplitude are set to $\gamma = 15$ and $DT = 0.7$.

For initialising new tracks, a two-step track initialisation method [33] is used. The initial states of a track are computed by associated measurements during few recent scans. The initial covariance is

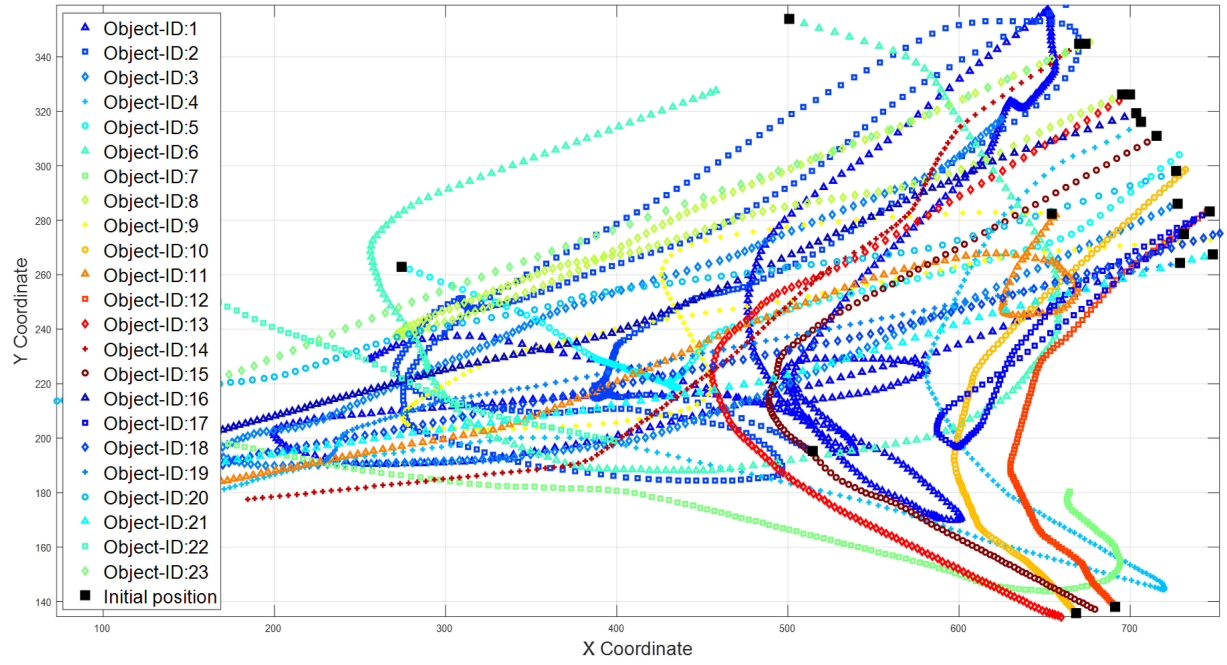
$$\boldsymbol{P}_k = \mathrm{diag}[1(\mathrm{pixel})^2, 25(\mathrm{pixel/frame})^2, 1(\mathrm{pixel})^2, 25(\mathrm{pixel/frame})^2]^{\mathrm{T}}$$

in consideration of maximum distances and velocities of moving objects in an image.

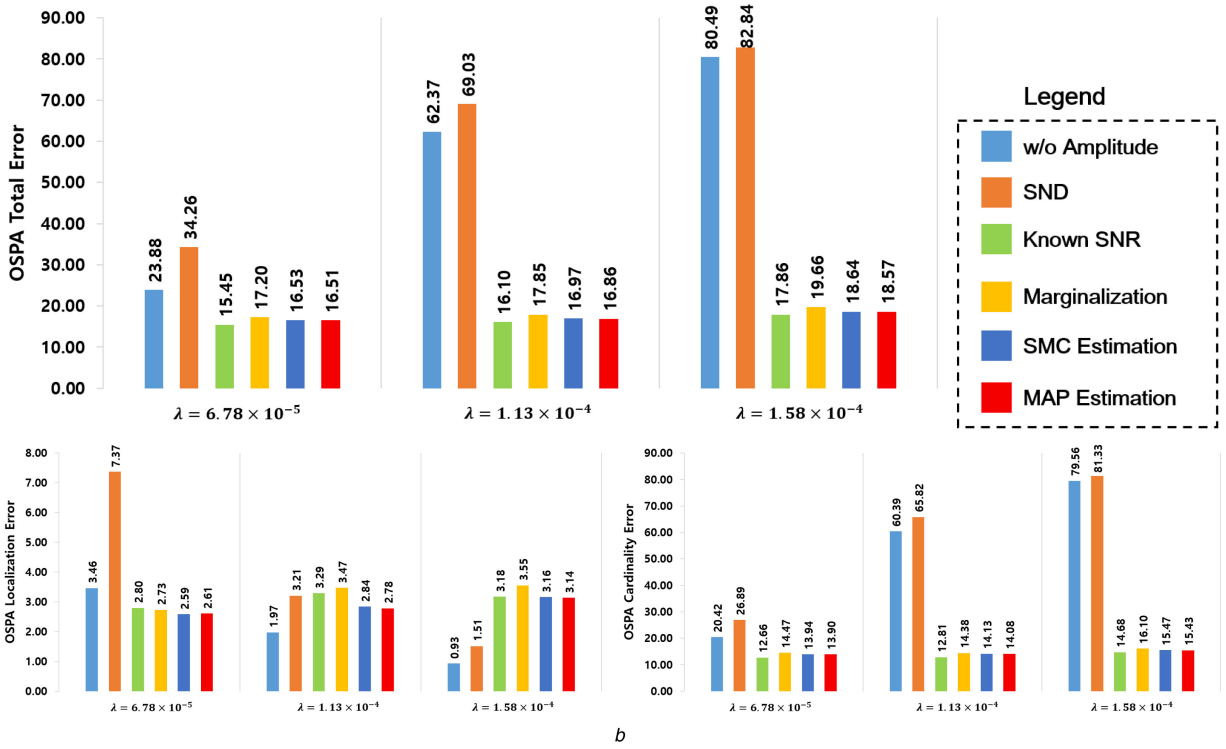
5.6 Accuracy comparison

5.6.1 PETS S2.L1 sequence: In Fig. 2, we compare the MTT systems (TR1)–(TR6). Fig. 2*a* shows the true trajectories of all the objects in this sequence. In Fig. 2*b*, we show the OSPA total, cardinality, and location error rates for the systems. The OSPA total errors of systems (TR1) and (TR2) using spatial or amplitude information only are much higher that those of (TR3)–(TR6) using both measurements. As clutter density λ increases, OSPA errors of (TR1) and (TR2) significantly increases. The main reason is that the clutter measurements in the vicinity of objects and/or with high amplitude are generated more as λ become higher. As a result, more clutter measurements are likely to associate with tracks in (TR1) and (TR2). However, we found that (TR3)–(TR6) can still maintain the good performance even for high λ since both spatial and amplitude likelihood $\Lambda_{i,k}^{p,\tau}$ and $\Lambda_{i,k}^{a,\tau}$ are applied for track-to-measurement association as in (32). Therefore, to associate a clutter measurement with a track in (TR3)–(TR6), the clutter has high spatial and amplitude likelihood at the same time.

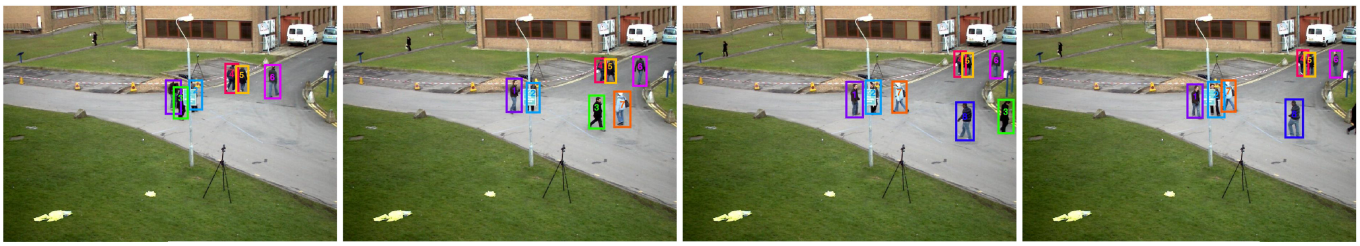
As expected, (TR3) with known SNR shows the better OSPA performance than (TR4)–(TR6) without the knowledge of SNR. In addition, (TR5) and (TR6) which estimate the object SNRs show the better rates than (TR4) using the marginalisation. This implies



a



b



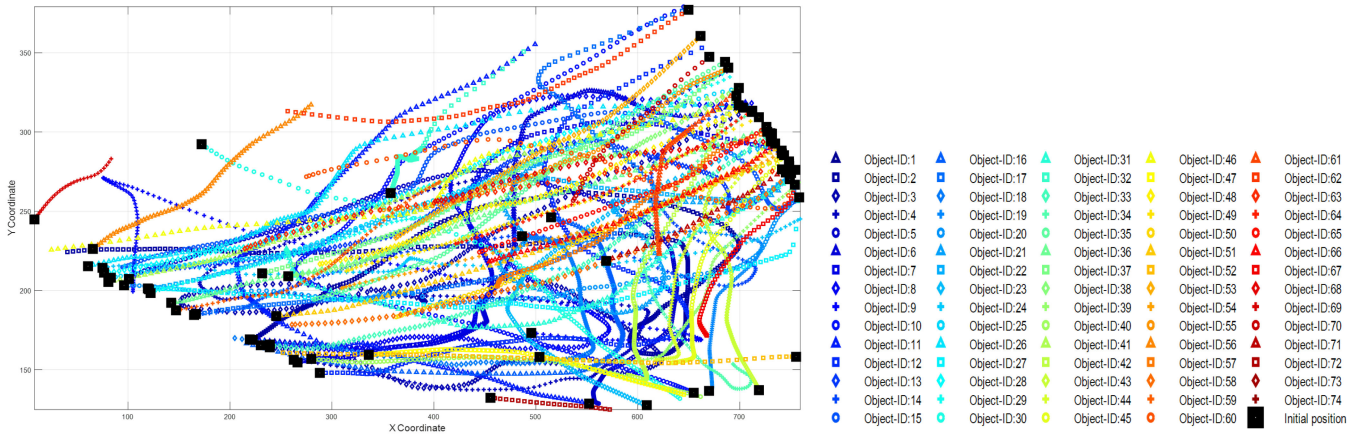
c

Fig. 2 *Comparisons with different MTT systems (TR1–TR6) for the PETS S2.L1 over 795 frames*

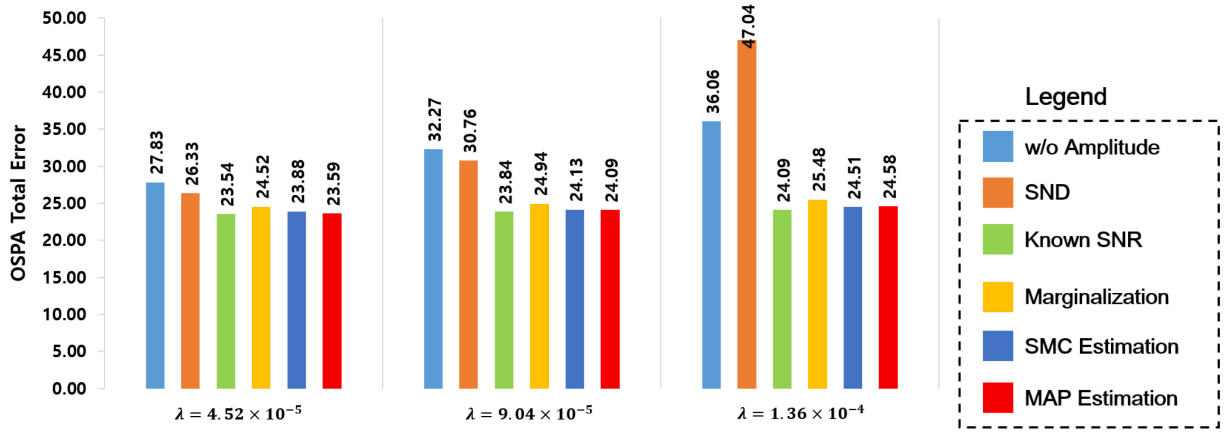
(a) True trajectories, ***(b)*** Comparison of different methods in terms of OSPA metrics, ***(c)*** PETS-L1 tracking results using the TR6. From left to right: frame 107, frame 128, frame 146, frame 149

that the amplitude likelihood function can be computed more accurately by using the estimated SNRs. Moreover, (TR5) and (TR6) show the almost similar performance.

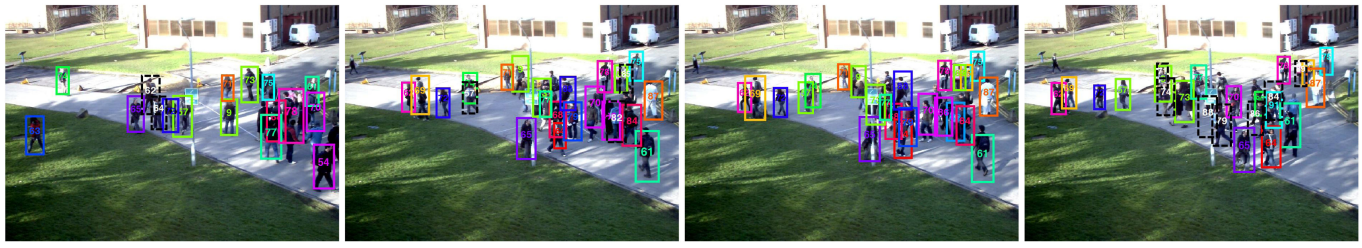
5.6.2 PETS S2.L2 sequence: Fig. 3 shows the tracking results of the (TR1)–(TR6) on the PEST S2.L2 sequence. As mentioned, this sequence is very challenging due to complex motions between



a



b



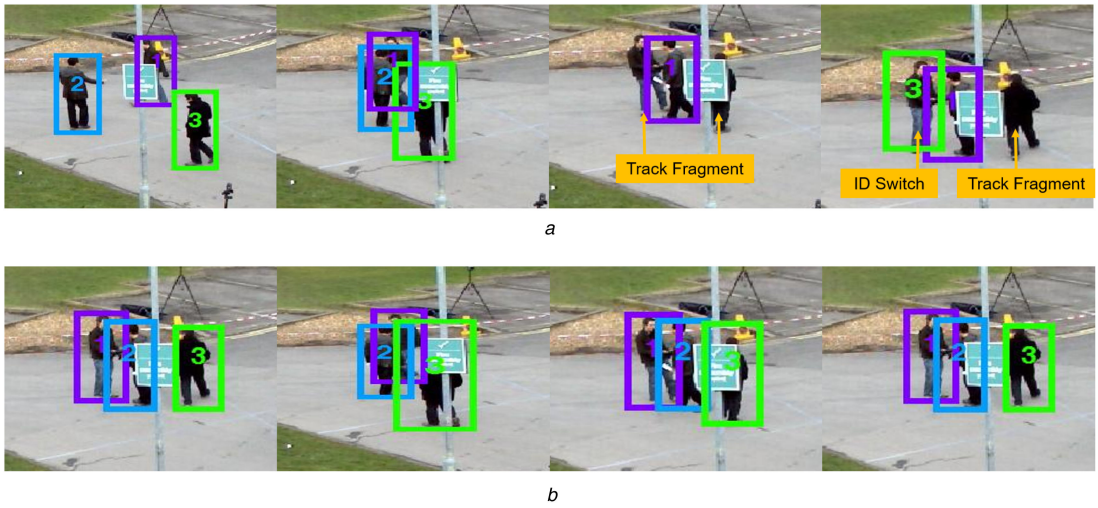
c

Fig. 3 *Comparisons with different MTT systems (TR1–TR6) for the PETS S2.L2 over 436 frames*

(a) True trajectories, ***(b)*** Comparison of different methods in terms of OSPA metrics, ***(c)*** PETS-L2 tracking results using the TR6. From left to right: frame 373, frame 403, frame 405, frame 416

many objects. Fig. 3*b* compares the performance of different systems. We also confirm that exploiting both spatial and amplitude measurements improve MTT performance when comparing (TR1) and (TR2) with (TR3)–(TR6). Even though (TR3) with known SNR shows the best rates for all clutter densities, (TR4)–(TR6) are comparable with the (TR3). Here, (TR5) and (TR6) based on SNR estimation methods show the better performance than (TR3). Since many objects exist, the cardinality errors of the (TR1) and (TR2) systems increase greatly as λ increases. However, the cardinality errors for (TR3)–(TR6) are not increased much for the large λ. The low cardinality error means that the number of generated tracks is close to the number of true objects.

5.6.3 Qualitative evaluation: For more comparison between the marginalisation and SNR estimation methods, we demonstrate the tracking results of (TR4) and (TR6) on the PETS S2.L1 sequence in Fig. 4. As can be seen, some track fragments and ID switch occur when using (TR4). On the other hand, (TR6) maintains the identities of objects correctly under occlusions. We provide the more tracking results of (TR6) in Figs. 2*c* and 3*c*. Many objects are successfully tracked even in complex scenes.



a

b

Fig. 4 *For PETS S2.L1 sequence, comparison results of (TR4) and (TR6). The selected frames are #18, #26, #39, #44*
(a) Tracking results using (TR4) on PETS S2.L1, ***(b)*** Tracking results using (TR5) on PETS S2.L1

Table 1 Computational complexity of the main algorithms for TR1–TR6. The scores are computed with $DT = 1$ and $\lambda = 1.58 \times 10^{-4}$ on PETS S2.L1 sequence

	(TR1)-w/o Amp.	(TR2)-SND	(TR3)-Known	(TR4)-Marg.	(TR5)-SMC	(TR6)-MAP
(steps 1 and 2) gating and likelihood	1.55×10^{-2}	9.27×10^{-3}	5.18×10^{-3}	4.99×10^{-3}	5.12×10^{-3}	1.56×10^{-2}
(step 3) data association	8.51×10^{-4}	4.60×10^{-4}	2.84×10^{-4}	2.66×10^{-4}	2.55×10^{-4}	2.36×10^{-4}
(step 4) particle filtering	3.36×10^{-2}	3.07×10^{-2}	2.76×10^{-2}	2.85×10^{-2}	2.78×10^{-2}	2.38×10^{-2}
overall complexity	4.99×10^{-2}	4.04×10^{-2}	3.30×10^{-2}	3.38×10^{-2}	3.32×10^{-2}	3.96×10^{-2}

5.7 Speed comparison

Table 1 compares the speed of the systems (TR1)–(TR6). As described in Algorithm 1 (Fig. 1), the computational time (second/frame) of the main steps in each system is also computed. As shown, the particle filtering step in all the systems is most computationally expensive. (TR6) requires the more time in step 1 due to the target SNR estimation. Interestingly, the algorithm complexity of (TR1) and (TR2) is lower than (TR4)–(TR6) because (TR4)–(TR6) contain the additional process of marginalising an amplitude function or estimating a target SNR as shown in Algorithm 1 (Fig. 1). However, they show the lower speed than (TR4)–(TR6). The main reason is that many false tracks are generated by the inaccurate association in (TR1) and (TR2).

5.8 Sensitivity and stability analysis

In this section, we present sensitive and stability analysis of MTT systems (TR1)–(TR6). To this end, on PETS S2. L1, we evaluate the variations of OSPA total errors (*or* distance) by changing important parameters, which might affect the overall MTT performance, at 795 frames based on 200 Monte Carlo runs. In this evaluation, $\lambda = 1.58 \times 10^{-4}$ is used to evaluate the sensitivity and stability of each system under heavy cluttered environment.

5.8.1 Amplitude threshold: In amplitude-aided tracking, the amplitude threshold DT can affect the overall MTT performance. The reason is that DT can be exploited for filtering out false alarm based on the assumption the amplitude from a target is stronger than false alarms. Moreover, DT is used for amplitude likelihood computation for the target (10) and clutter (12).

To show the effect of DT, we change the DT with [0:1:10] for all the systems (TR1)–(TR6). We then evaluate the OSPA total errors and tracking speed for each system in Figs. 5*a* and *b*. For all the systems, the tracking complexity is reduced by using high DT, but their performance can be degraded because target originated measurements can be filtered out. Except for (TR3), (TR6) achieves the lowest OSPA score of 18.44 for $DT = 1$. Remarkably, we found that the most systems (TR3)–(TR6) by using amplitude likelihoods produce the high performance when applying the low DT. Therefore, we recommend that exploiting the low DT between 0 and 1 is suitable for amplitude-aided MTT systems in order to obtain good tracking performance.

In addition, for $DT \geq 2$, (TR2), which associates a measurement with the highest amplitude to a track, shows the almost similar performance with other systems (TR4–TR6). When averaging the OSPA results for all DT, (TR5) shows the better rates except for ($TR3$), but the performance difference between (TR5) and (TR4)/(TR6) is ∼1.5%. The average speed of (TR4) and (TR5) is higher rather than ones of other systems. When DT is low, the speeds of (TR4) and (TR5) are much faster than those of (TR1) and (TR2) by reducing false track generation as described in Section 5.7.

5.8.2 Number of samples in a particle filter: All the systems (TR1)–(TR6) use the particle filter to approximate the posterior distribution of hidden states with a set of weighted samples. One key parameter of a particle filter is the sample number N. As shown in [34, 35], exploiting a large number of samples improves the particle diversity. This indicates that the overall tracking accuracy can be enhanced because the posterior distribution can be approximated better. However, using many samples increase the computational complexity. Therefore, many studies [36, 37] have been conducted to determine the number of samples to trade the accuracy and speed properly. In this paper, we have also investigated the stability and sensitivity of each system towards the number of samples.

Fig. 5*c* shows the OSPA errors for different N. When $N = 10$, the OSPA errors of all the systems are peaked. However, we found that their performance does not increase considerably after $N = 25$. This means that N is not a sensitive parameter for the MTT systems. In addition, they reach the stable performance when N exceeds to 25. We can also infer that posterior distributions of target states are affected more by the selected measurement rather than the predicted samples. Therefore, we found that improving data association between tracks and measurements is important to capture a desirable posterior distribution in particle-filter-based MTT.

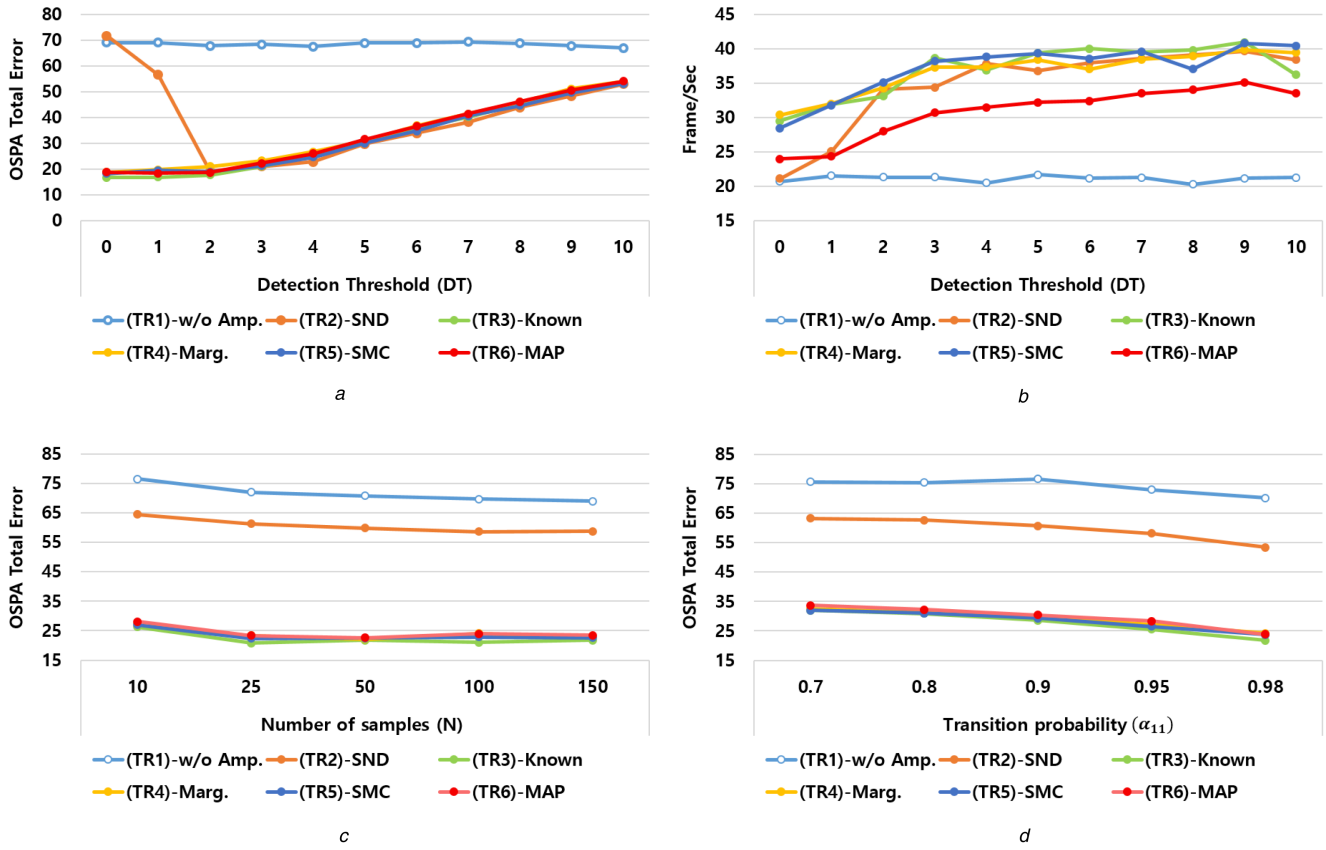


Fig. 5 *For PETS S2.L1 sequence, comparison of different MTT methods (TR1–TR6) by changing DT, N, and* α_{11}
(a) OSPA total error for *DT*, ***(b)*** Tracking speed, ***(c)*** OSPA total error for *N*, ***(d)*** OSPA total error for α_{11}

Table 2 On the PETS S2.L1, the averaged OSPA scores and standard deviation of (TR1)–(TR6) are computed for the hyperparameters *DT*, *N*, and α_{11}

	(TR1)-w/o Amp.	(TR2)-SND	(TR3)-Known	(TR4)-Marg.	(TR5)-SMC	(TR6)-MAP
detection threshold (*DT*)	68.48/0.77	39.93/16.64	32.17/13.81	33.67/12.98	32.36/13.00	33.19/13.42
number of samples (*N*)	71.58/2.98	60.55/2.39	22.32/2.25	23.86/1.91	23.47/1.98	24.29/2.19
transition probability (α_{11})	74.11/2.61	59.57/4.01	27.73/4.14	28.98/3.42	28.48/3.40	29.68/3.84
average	71.39/2.12	53.35/7.68	27.41/6.73	28.84/6.10	28.10/6.13	29.05/6.48

5.8.3 Transition probability: Fig. 5*d* shows the variation of OSPA errors for systems (TR1)–(TR6) according to the transition probability α_{11} used for the prediction of a track existence probability (35). As α_{11} decreases, the predicted track existence probability also decreases. Using the low transition probability α_{11} decreases the tracking performance as shown. All the systems show the best rate when $\alpha_{11} = 0.98$. These results mean that the performance of the systems is rather sensitive to the transition probability, but the stability can be maintained easily by using the high α_{11}.

5.8.4 System stability: We found that all the systems maintain the stability since they do not diverge during the 200 Monte Carlo runs. However, for a more comparison, in Table 2, we provide the averaged OSPA rates and standard deviation for the hyperparameters. The small standard deviation indicates that a system is not sensitive to the parameter, and keeps the high stability for the variation of the parameter. Even though (TR1) shows the lowest standard deviation among them, it is not difficult to conclude that (TR1) achieves the highest stability since its OSPA average score is too low. In overall, (TR3)–(TR6) show the higher stability than (TR1) and (TR2) since they show low OSPA error and standard deviation. This comparison implies that using both features also improves the system stability.

5.9 Summary

This section provides the summary of the experimental results and a guideline to select amplitude-aided MTT methods.

We have compared the performance of systems (TR1)–(TR6) in terms of several aspects. We first compare the tracking accuracy of the systems in Figs. 2 and 3. By comparing the mean accuracy (OSPA total error) of (TR1)–(TR6) in both sequences, we rank the systems in terms of the accuracy as follows:

$$\text{TR3} > \text{TR6} > \text{TR5} > \text{TR4} > \text{TR1} > \text{TR2} \tag{39}$$

We then compare the speed of the systems in Table 1, and the systems can be ranked for the speed as follows:

$$\text{TR3} > \text{TR5} > \text{TR4} > \text{TR6} > \text{TR2} > \text{TR1} \tag{40}$$

In Fig. 5, the sensitivity and stability are evaluated by changing the values of the detection threshold (*DT*), sample number (*N*), and transition probability (α_{11}). From the results, we can order the parameters to affect the system performance as follows:

$$DT > \alpha_{11} > N \tag{41}$$

In Table 2, we show the overall performance of each system for each parameter. By summing the scores of the average OSPA rate and standard deviation, we can order the system stability as follows:

$$\text{TR3} > \text{TR5} > \text{TR4} > \text{TR6} > \text{TR2} > \text{TR1} \quad (42)$$

To sum up, we can present the following key guidelines:

- Using both spatial and amplitude features improves the accuracy and also the speed of a MTT system.
- Using both features reduces the sensitivity to the hyperparameters and improves the stability under parameter change.
- When inferring the amplitude likelihood function, estimating a target SNR is more accurate than marginalising the function.
- Except for (TR3) using the known SNR, the system (TR5) produces the best trade-off value between the accuracy and speed.
- A track-to-measurement association is key when approximating the distribution of multiple targets' states using particle filtering.

6 Conclusion

This study has presented the review of amplitude-aided MTT methods. For implementing the methods in a SMC framework, we have designed a unified MTT framework based on amplitude-aided data association. The implemented MTT methods are compared extensively for the several aspects such as accuracy, speed, sensitivity, and stability on challenging visual MTT datasets. From the experimental comparisons, we have provided a practical guide to select a suitable method when developing amplitude-aided tracking systems in real applications.

7 Acknowledgment

This work was supported by the Incheon National University International Cooperative Research Grant in 2017.

8 References

[1] Gini, F., Greco, M.V., Farina, A.: 'Clairvoyant and adaptive signal detection in non-Gaussian clutter: a data-dependent threshold interpretation', *IEEE Trans. Signal Process.*, 1999, **47**, (6), pp. 1522–1531

[2] Bar-Shalom, Y., Li, X.R.: '*Multitarget-multisensor tracking: principles and techniques*' (YBS Publishing, Storrs, CT, 1995)

[3] Li, X.R.: 'Tracking in clutter with strongest neighbor measurements. Part I: theoretical analysis', *IEEE Trans. Autom. Control*, 1998, **43**, (11), pp. 1560–1578

[4] Reid, D.: 'An algorithm for tracking multiple targets', *IEEE Trans. Autom. Control*, 1979, **24**, (6), pp. 843–854

[5] Mušicki, D., La Scala, B.: 'Multi-target tracking in clutter without measurement assignment', *IEEE Trans. Aerosp. Electron. Syst.*, 2008, **44**, (3), pp. 877–896

[6] Ekmanm, M.: 'Particle filters and data association for multi-target tracking'. Proc. Int. Conf. Information Fusion, Cologne, Germany, 2008, pp. 1–8

[7] Särkkä, S., Vehtari, A., Lampinen, J.: 'Rao-blackwellized Monte Carlo data association for multiple target tracking', *Inf. Fusion*, 2007, **8**, (1), pp. 2–15

[8] Chavali, P., Nehorai, A.: 'Concurrent particle filtering and data association using game theory for tracking multiple maneuvering targets', *IEEE Trans. Signal Process.*, 2013, **61**, (20), pp. 4934–4948

[9] Jinana, R., Raveendran, T.: 'Particle filters for multiple target tracking'. Proc. Int. Conf. Emerging Trends in Engineering, Science and Technology, Kerala, India, 2016, pp. 980–987

[10] Kreucher, C., Kastella, K., Hero, A.O.: 'Multitarget tracking using the joint multitarget probability density', *IEEE Trans. Aerosp. Electron. Syst.*, 2005, **41**, (4), pp. 1396–1414

[11] Morelande, M.R., Kreucher, C.M., Kastella, K.: 'A Bayesian approach to multiple target detection and tracking', *IEEE Trans. Signal Process.*, 2007, **55**, (5), pp. 1589–1604

[12] Yi, W., Morelande, M.R., Kong, L., *et al.*: 'Computationally efficient particle filter for multitarget tracking using an independence approximation', *IEEE Trans. Signal Process*, 2013, **61**, (4), pp. 843–856

[13] Kim, D.S., Song, T.L., Mušicki, D.: 'Highest probability data association for multi-traget particle filtering with nonlinear measurements', *IEICE Trans. Commun.*, 2013, **E96-B**, (1), pp. 281–290

[14] La Scala, B.F.: 'Viterbi data association tracking using amplitude information'. Proc. Int. Conf. Information Fusion, Stockholm, Sweden, June–July 2004

[15] Lerro, D., Bar-Shalom, Y.: 'Automated tracking with target amplitude information'. American Control Conf., San Diego, CA, USA, 1990, pp. 2875–2880

[16] Van Keuk, G.: 'Multihypothesis tracking using incoherent signal-strength information', *IEEE Trans. Aerosp. Electron. Syst.*, 1996, **32**, (3), pp. 1164–1170

[17] Clark, D., Ristic, B., Vo, B.-N., *et al.*: 'Bayesian multi-object filtering with amplitude feature likelihood for unknown object SNR', *IEEE Trans. Signal Process.*, 2010, **58**, (1), pp. 26–37

[18] Bae, S.-H., Kim, D.Y., Yoon, J.H., *et al.*: 'Automated multi-target tracking with kinematic and non-kinematic information', *IET Radar Sonar Navig.*, 2012, **6**, (4), pp. 272–281

[19] Bae, S.-H., Park, J.Y., Yoon, K.-J.: 'Joint estimation of multi-target SNR and dynamic states in cluttered environment', *IET Radar Sonar Navig.*, 2017, **11**, (3), pp. 539–549

[20] Ehrman, L.M., Blair, W.D.: 'Comparison of methods for using target amplitude to improve measurement-to-track association in multi-target tracking'. Proc. Int. Conf. Information Fusion, Florence, Italy, 2006, pp. 1–8

[21] Skolnik, M.I.: '*Introduction to radar system*' (Mc-Graw-Hill, New York, 2002, 3rd edn.)

[22] Slocumb, B.J., Klusman, M.E.: 'A multiple model SNR/ RCS likelihood ratio score for radar-based feature-aid tracking'. Proc. SPIE. Signal and Data Processing of Small Targets, San Diego, CA, USA, August 2005

[23] Schuhmacher, D., Vo, B.-T., Vo, B.-N.: 'A consistent metric for performance evaluation of multi-object filters', *IEEE Trans. Signal Process.*, 2008, **56**, (8), pp. 3447–3457

[24] Field, E., Lewinstein, M.: 'Amplitude-probability distribution model for VLF/ELF atmospheric noise', *IEEE Trans. Commun.*, 1978, **26**, (1), pp. 83–87

[25] Cover, T.M., Thomas, J.A.: '*Elements of information theory*' (Wiley, New York, 2006, 2nd edn.)

[26] Nocedal, J., Wright, S.: '*Numerical optimization*' (Springer-Verlag, New York, 2006, 2nd edn.)

[27] Vo, B.-T., Mallick, M., Bar-shalom, Y., *et al.*: '*Multitarget tracking*' (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2015)

[28] Wang, X., Li, T., Sun, S., *et al.*: 'A survey of recent advances in particle filters and remaining challenges for multitarget tracking', *Sensors*, 2017, **17**, (12), pp. 1–21

[29] Arulampalam, M.S., Maskell, S., Gordon, N., *et al.*: 'A tutorial on particle filters for non-linear/non-Gaussian Bayesian filtering', *IEEE Trans. Signal Process.*, 2002, **50**, (2), pp. 174–188

[30] Ristic, B., Arulampalm, S., Grodon, N.: '*Beyond the Kalman filter: particle filter for tracking application*' (Artech House, London, 2004)

[31] 'VS-PETS dataset', Available at http://www.cvg.reading.ac.uk/PETS2009, accessed March 2009

[32] Dollár, P., Belongie, S.J., Perona, P.: 'The fastest pedestrian detector in the west'. British Machine Vision Conf., Aberystwyth, UK, September 2010, pp. 1–11

[33] Bar-Shalom, Y.: '*Multitarget-multisensor tracking: applications and advances*' (Artech House, Boston, MA, USA, 1990), pp. 25–42

[34] Zand, G., Taherkhani, M., Safabakhsh, R.: 'Exponential natural particle filter'. Computing Research Repository (CoRR), 2015, abs/1511.06603, Available at: https://arxiv.org/abs/1511.06603

[35] Li, T., Sun, S., Sattar, T.P., *et al.*: 'Fight sample degeneracy and impoverishment in particle filters: a review of intelligent approaches', *Expert Syst. Appl.*, 2014, **41**, (8), pp. 3944–3954

[36] Straka, O., Simandl, M.: 'A survey of sample size adaptation techniques for particle filters', *Proc. IFAC*, 2009, **42**, (10), pp. 1358–1363

[37] Elvira, V., Mguez, J., Djuric, P.M.: 'Adapting the number of particles in sequential Monte Carlo methods through an online scheme for convergence assessment', *IEEE Trans. Signal Process.*, 2017, **65**, (7), pp. 1781–1794