Exclusive Association Sampling to Improve Bayesian Multi-Target Tracking

SHUN TAGUCHI AND KIYOSUMI KIDONO
Toyota Central R&D Labs., Inc., Nagakute 480-1192, Japan
Corresponding author: Shun Taguchi (s-taguchi@mosk.tytlabs.co.jp)

ABSTRACT Multi-target tracking has been studied for many years, yet it remains a challenging problem, particularly in terms of implementing data association when tracking targets over several time steps. To achieve robustness, probabilistic approaches have been proposed, including Bayesian multi-target tracking methods. However, these approaches involve high calculation costs, which are incompatible with real-time applications. We propose exclusive association sampling (EAS) to improve the efficiency of Bayesian multi-target tracking methods. Although EAS is a simple procedure composed of random sorting and sampling based on the observation probability, it can be employed to increase the efficiency of association candidate generation and the calculation of statistical values. In this study, we proposed two Bayesian multi-target tracking methods based on EAS: stochastic joint probabilistic data association (SJPDA) and Rao-Blackwellized particle filter with EAS (RBPF with EAS). Evaluation of these methods with simulated data shows that integrating EAS into these methods can enhance their speed and accuracy. Moreover, evaluation on open datasets used for pedestrian tracking on camera sequences shows that the proposed methods achieve significantly better performance on some important metrics compared with representative methods.

INDEX TERMS Multi-target tracking, joint probabilistic data association, Rao-Blackwellized particle filter.
FIGURE 1. How EAS works. We propose EAS, which can be applied to the data association problem to generate several candidates for association. EAS provides efficient samples of associations for calculating the weighted association for JPDA or for associating samples in the RBPF.

multi-target tracking, as a result they can improve the tracking efficiency (see Fig. 1).

We integrate the proposed method with two Bayesian filtering-based multi-target tracking methods: JPDA [1] and Rao-Blackwellized particle filtering (RBPF) [6]. The JPDA-based method, called stochastic JPDA (SJPDA), applies EAS to approximate joint probabilistic scores, and the simulation results show that this is achieved faster than with a recently developed improved JPDA variant [3]. The RBPF-based method, called RBPF with EAS, utilizes EAS to generate data associations for individual particles, and it achieves considerably higher accuracy than the conventional RBPF method [6]. We evaluate the proposed methods using open datasets on pedestrian tracking with a single camera. The results demonstrate that the proposed methods are considerably better than the original methods and that they achieve better performance on some important metrics than representative methods.

The remainder of this paper is organized as follows. In Section II, we review work related to this study. In Section III, we explain the data association problem. In Section IV, we explain EAS for improving Bayesian multi-target tracking. In Section V, we propose two Bayesian multi-target tracking methods using EAS. The experimental results are presented in Section VI. Finally, we conclude this paper in Section VII.

II. RELATED WORK
A. BAYESIAN MULTI-TARGET TRACKING IN EARLY TIME

The primitive approach to multi-target tracking involves the use of an optimization method to implement data associations within one-step observations. This association process can be modeled as an assignment problem and solved using a combinatorial optimization algorithm, such as the Hungarian method [7]. Each target is tracked using a Bayesian filter, e.g., a Kalman [8] or particle filter [9]. However, the one-step optimization approach is not robust because the association with the highest joint probability is not always correct. To achieve robustness, various probabilistic approaches that use multiple candidates for data associations or probability-weighted data associations have been proposed based on Bayesian filtering.

Multiple hypothesis tracking (MHT) [2], which is the most successful multiple-candidate approach, creates a tree of tracking hypotheses for each candidate target, with individual branches representing respective data association candidates. The solution is obtained by finding the most likely combination of tracks without conflicts. As MHT is essentially a breadth-first search algorithm, its performance strongly depends on its ability to prune branches within the search tree.

JPDA [1] is the most successful approach based on probability-weighted data association. It performs probabilistic data association by calculating joint probabilistic scores. However, naïve JPDA suffers from high combinatorial complexity because it considers all possible assignments of observations to targets when calculating the joint probabilistic scores. Correspondingly, as the number of targets increases, the technique becomes intractable in nearly all practical applications. To alleviate the computational burden, some heuristic approximations for JPDA have been proposed [10], [11]. Nevertheless, these methods sacrifice tracking accuracy to make the algorithms computationally tractable.

RBPF [6] is a multiple candidate-based approach in which candidates are generated for each data association set instead of for each candidate target, as in the case of MHT. However,
in situations with high numbers of cluttered detections, the RBPF sampling approach does not perform well because it does not consider data association exclusivity.

Certain types of sensors (such as radar and LIDAR) provide no information to identify objects beyond their positions. This has led to the development of tracking methods that do not distinguish individual objects, such as filtering methods based on random set theory [12], e.g., Gaussian mixture probability hypothesis density (GMPHD) filters [13], which estimate the states of multiple targets as a probability density. Because these methods do not provide IDs for individual targets, they are not suited to track individual targets.

B. RELATED STUDIES IN COMPUTER VISION

Multiple target tracking has also been extensively studied in the context of computer vision. Successful approaches are based on offline optimization [14]–[17], which follows a different strategy from the Bayesian filtering-based approaches described above. Sequences of frames are grouped together and the states and data associations of all targets are jointly inferred by optimizing a predefined objective. These approaches introduce respective assumptions to simplify conditional dependences and reduce complexity through the use of, for example, the min-cost flow [14]–[16] or k-shortest paths algorithm [17].

Improvement of optimization techniques and the usefulness of exploiting appearance models have led to a revisiting of Bayesian filtering methods, as exemplified by Kim et al. [4], who applied MHT to develop an appearance model. The Bayesian filtering framework can easily integrate different models, such as motion and appearance models, suggesting that a classical MHT implementation has come surprisingly close to the performance of state-of-the-art methods. Rezatofghi et al. [3] applied JPDA as an approximation, using the m-best solutions (JPDA_m), which calculates joint probabilistic scores from the m-best assignments instead of from all assignments. Furthermore, a recently developed integer linear programming algorithm [18] can quickly provide the m-best solutions and has been applied to significantly reduce the JPDA calculation time. Our SJPDA was inspired by JPDA_m [3] and demonstrates improved calculation speed by introducing a sampling strategy, rather than calculating the m-best solutions. Random-finite-set-based approaches have also been studied, as they are Bayesian filterings for multiple object tracking that can operate in real time without any special modifications. In particular, some GMPHD-based approaches have been proposed [19]–[22]. However, GMPHD-based approaches suffer from many ID switches, as they track probability density instead of individual objects. Almost all of these studies employ appearance information to identify individual objects. Some of these methods [20], [21] are evaluated in Sec. VI.

More recently, deep learning approaches using the appearance information have been intensively studied [23]–[25]. Milan et al. [23] proposed recurrent neural network based approach. Leal-Taixé et al. [24] employ a Siamese CNN to learn local appearance features to associate targets. Sadeghian et al. [25] propose an appearance model using the LSTM, which takes images in the tracklet step-by-step and predicts the similarity score. However, these approaches are highly dependent on appearance information; therefore they cannot be applied to sensors that do not provide appearance information about the target.

III. DATA ASSOCIATION PROBLEM

We start with a discussion of the data association problem, which is the main problem in multi-target tracking. Data association is the task of associating multiple observations with multiple estimated tracks. The data association problem is illustrated schematically in Fig. 2, where \( x_{i,k} \) is the state of the \( i \)-th track at time step \( k \), and \( y_{j,k} \) is the \( j \)-th observation at time step \( k \).

The connection matrix \( \theta_k \) between the \( M \) states \( x_k = \{x_{1,k}, \ldots, x_{M,k}\} \) and the \( N \) observations \( y_k = \{y_{1,k}, \ldots, y_{N,k}\} \) is unknown and must be estimated at each step. This process is called data association.

The primitive approach to connection estimation involves optimization of the joint observation probability \( p(y_k|x_k, \theta_k) \). The Hungarian method [7] is the most popular optimization method for tackling this problem. However, this one-best approach is not very robust because the association with the highest joint probability at each step is not always correct for the time-series observation. More robust probabilistic approaches introduce multiple candidates for data associations or probability-weighted data associations. Here, EAS is a technique for improving the efficiency of such probabilistic data associations.
k each time-step it is a module used to improve Bayesian multi-target tracking capability. EAS is not a stand-alone association algorithm; rather, weighted association samples according to the joint probability that is used for generating association samples. The main difference from the Hungarian method is the generation of weighted association samples according to the joint probability. EAS is not a stand-alone association algorithm; rather, it is an algorithm that is used for generating association samples. The main difference from the Hungarian method is the generation of weighted association samples according to the joint probability. EAS is not a stand-alone association algorithm; rather, it is a module used to improve Bayesian multi-target tracking methods such as JPDA and RBPF.

The EAS process applies the following procedure during each time-step $k$ (with $y_{j,k}$ written as $y_j$):

1) Randomly sort observations $\{y_1, \cdots, y_N\}$.
2) Associate the first observation $y_j$ with the target $x_i$ based on the probability $p(y_j|x_i)$.
3) Exclude the associated target from subsequent associations.
4) Repeat steps 2 and 3 for all observations.

Using this procedure, a sample $\theta$ of the associations is obtained. The procedure is illustrated schematically in Fig. 3.

EAS can approximate the joint probability $p(y|x, \theta)$ by introducing an appropriate weight $w_\theta$ for each sample. This joint probability $p(y|x, \theta)$ is written as follows:

$$p(y|x, \theta) = \prod_{i=1}^{M} \prod_{j=1}^{N} p(y_j|x_i)^{\theta_{ij}},$$

where $\theta_{ij}$ represents element $i,j$ of the association matrix $\theta$. If $x_i$ and $y_j$ are associated, then $\theta_{ij} = 1$. Otherwise, $\theta_{ij} = 0$. The sampling probability $q(\theta)$ is determined through the following procedure. First, an order of observations is generated from the $N!$ possible orders by random sorting (see Fig. 3(b)). Then, $N$-times sampling is executed. The sampling probability $q(\theta)$ is therefore:

$$q(\theta) = \frac{1}{N!} \prod_{j=1}^{N} (\frac{p(y_{d_j}|x_j)}{\sum_{i \in \chi_j} p(y_{d_j}|x_i)})^{\theta_{d_j}},$$

where $\chi_j$ is the candidate set of targets that are available for association, and $|\cdot|$ denotes the number of elements. $d_j$ is the $j$-th index of the random sorted observations. We introduce the weight $w_\theta$ to approximate the joint probability $p(y|x, \theta)$:

$$p(y|x, \theta) = w_\theta \cdot q(\theta),$$

where $\cdot$ denotes multiplication. This equation can be transformed as follows:

$$w_\theta = \frac{p(y|x, \theta)}{q(\theta)} = \frac{N! \prod_{i=1}^{M} \prod_{j=1}^{N} p(y_j|x_i)^{\theta_{ij}}}{\prod_{j=1}^{N} |\chi_j| \prod_{i=1}^{M} \left( \frac{p(y_{d_j}|x_j)}{\sum_{i \in \chi_j} p(y_{d_j}|x_i)} \right)^{\theta_{d_j}}},$$
where $\theta_{ij}$ is 1 at only one $i$ in $i = 1, \ldots, |X_j|$, otherwise 0. Therefore (4) can be transformed as follows:

$$w_\theta = N! \cdot \prod_{j=1}^{N} \sum_{x \in X_j} p(y_{dj} | x) \prod_{i=1}^{M} \frac{p(y_{dj} | x_i)^{\theta_{ij}}}{\prod_{i=1}^{M} p(y_{dj} | x_i)^{\theta_{ij}}}.$$  

(5)

When $\theta$ is fixed, an observation $j$ and its corresponding target $i$ are always included in $X_j$; therefore, the following relation holds:

$$\prod_{i=1}^{M} p(y_{dj} | x_i)^{\theta_{ij}} = \prod_{i=1}^{M} p(y_{dj} | x_i)^{\theta_{ij}}.$$  

(6)

Substituting (6) into (5), the suitable weight $w_\theta$ can be obtained as follows:

$$w_\theta = N! \cdot \prod_{j=1}^{N} \sum_{x \in X_j} p(y_{dj} | x).$$  

(7)

By introducing the weight $w_\theta$, EAS can approximate the joint probability $p(y | x, \theta)$.

The complete algorithm for EAS is presented in Algorithm 1. The computational complexity of EAS is $O(MN)$, however, because multiple sampling is generally necessary for applying multi-target tracking, the time complexity becomes $O(SMN)$, where $S$ is the number of samples. EAS can be calculated on an order close to that of the Hungarian method. When the number of targets and observations exceeds the number of samples, EAS is faster than the Hungarian method. In this algorithm, false detection and new targets are not considered. To address false detection and new targets, it is necessary to introduce corresponding dummy targets. Because the exclusion rule does not apply to a dummy target, we do not exclude it from the available set $X$, even if an observation has already been assigned to it.

**V. MULTI-TARGET TRACKING METHODS USING EAS**

In this section, we propose applications of EAS to two multi-target tracking methods: stochastic JPDA and RBPF with EAS.

**A. STOCHASTIC JPDA**

Stochastic JPDA (SJPDA) is a JPDA approximation method. The original JPDA is a multi-target tracking method using probability-weighted associations based on joint probabilities, which are calculated for each available association. The probabilistic distribution of the $i$-th target $p(x_i | y_{1:k})$, where $y_{1:k}$ denotes $[y_1, \ldots, y_k]$, is calculated by updating the previous distribution $p(x_i | y_{1:k-1})$ using the current observation $y_k$:

$$p(x_i | y_{1:k}) = \sum_{j=1}^{N} p(\theta_{ij}) \cdot p(x_i | y_{1:k-1}, y_{ij}).$$  

(8)

where $p(x_i | y_{1:k-1}, y_{ij})$ is the posterior distribution using an associated observation $y_{ij}$, which can be calculated using a Kalman filter. The association probability $p(\theta_{ij})$ is calculated as

$$p(\theta_{ij}) = E[\theta_{ij}] \approx \frac{\sum_{\theta^{(s)} \in \Theta_{EAS}} \theta^{(s)}_{ij} \cdot p(y_k | x_i, \theta^{(s)})}{\sum_{\theta^{(s)} \in \Theta_{EAS}} p(y_k | x_i, \theta^{(s)})},$$  

(9)

where $\Theta_{EAS}$ is the set of samples of the association given by EAS. By introducing EAS, we can approximate the association probability $p(\theta_{ij})$ using any number of samples of associations.

**B. RBPF WITH EAS**

Another proposed multi-target tracking method using EAS is the RBPF-based method. RBPF is a particle filter method that tracks states using multiple hypotheses of associations.

---

**Algorithm 1 Exclusive Association Sampling**

**Input:** Tracked states $x$ and observations $y$

**Output:** An association matrix $\theta$ and its weights $w_\theta$

1: $\theta \leftarrow 0^{M \times N}$
2: $\chi \leftarrow \{x_1, \ldots, x_M\}$
3: $w_\theta \leftarrow N!$
4: /* Process in Random Order */
5: for all $d_j \in RandomSort([1, \ldots, N])$ do
6: /* Sum the Probability of Candidates */
7: $p_\Sigma \leftarrow 0$
8: for all $x_i \in \chi$ do
9: $p_\Sigma \leftarrow p_\Sigma + p(y_{dj} | x_i)$
10: end for
11: /* Association Sampling */
12: Sampling $x_i$ from $\chi$ based on $\frac{p(y_{dj} | x_i)}{p_\Sigma}$
13: $\theta_{ij} \leftarrow 1$
14: $w_\theta \leftarrow w_\theta \cdot p_\Sigma$
15: Remove $x_i$ from $\chi$
16: end for
17: return $\theta, w_\theta$
Although RBPF-based multi-target tracking was previously proposed by Särkkä et al. [6], their sampling method does not consider the exclusivity of associations, resulting in some sampled associations being invalid as two or more observations are associated with one target. EAS efficiently removes such invalid associations.

In the EAS-based RBPF, the probability $p(x_k | y_{1:k})$ is represented as follows:

$$p(x_k | y_{1:k}) = \sum_{s=1}^{S} p(x_k^{(s)} | y_{1:k}), \quad (11)$$

where $S$ is the number of particles and $p(x_k^{(s)} | y_{1:k})$ is the $s$-th particle of the RBPF. Note that each particle contains all tracked states in the Kalman filter form. EAS is applied to update each particle as follows:

$$p(x_k^{(s)} | y_{1:k}) = \sum_{j=1}^{N} \theta_j^{(s)} p(x_k^{(s)} | y_{1:k-1}, y_{j:k}), \quad (12)$$

where $\theta_j^{(s)}$ is a sample of associations produced by EAS. This sample of associations is generated for each particle and is used to update all of the tracked states for that particle.

The operation of RBPF with EAS is illustrated schematically in Fig. 4, where $x_k^{(s)}$ represents the $s$-th particle at time-step $k$, which includes all tracked states. In the prediction step, tracked states in each particle are predicted by the system model. The data associations are then constructed through sampling, and each state is updated by the associated observation. The proposed method employs EAS to sample the associations and updates each tracked state based on the EAS-associated observations, with the probability of each particle updated using the weight produced by the EAS algorithm. Finally, resampling is performed to proceed to the next step.

VI. EXPERIMENTAL RESULTS

We evaluated the two proposed multi-target tracking methods first on simulated data and then on real camera sequences. Specifically, we used single camera sequences from open datasets [26]–[28] to evaluate the methods for pedestrian tracking.

A. EVALUATION ON SIMULATIONS

To evaluate the speed and accuracy of the proposed tracking methods, we applied them to a simulated scenario, in which three moving targets cross in front of each other. The simulated state space model is given as:

$$x_k = Fx_{k-1} + u_k,$$
$$y_k = Hx_k + v_k,$$

where $F$ is a Gaussian noise with covariance $Q = \text{diag}[0.1 \tau, 0.1 \tau]$, while the observation error is Gaussian noise with variance $R = 0.3$. The number of false detections was simulated using a Poisson distribution with mean $\lambda = 3$, and the value was simulated using a uniform distribution on $[0.0, 10.0]$. The targets were given the initial states $[1.0 \ 0.8]^T$, $[3.0 \ 0.1]^T$, and $[8.0 \ -0.7]^T$, where $(\cdot)^T$ denotes the matrix transpose operation. The simulation was performed 100 times for 100 steps.
In the evaluation, the initial i-th estimated state was $\hat{x}_{i,0} = [x_{i,0.0}, 0.0]^T$ with covariance $P_{i,0} = \text{diag}[1.0, 1.0]$. To evaluate the performance of the two proposed methods, we compared them with several existing methods, namely the Hungarian method, original JPDA [1], JPDA$_{m}$ [3], and the original RBPF [6]. The JPDA$_{m}$ method was applied as an online method, in which the branch and bound algorithm was employed to calculate the $m$-best solutions. However, if outputting the results online, the accuracy of both RBPF and RBPF with EAS would be reduced because of the loss of particle continuity. Therefore, we processed the results offline by particle smoothing, which traces particles backwards from the last maximum likelihood particle. All of the algorithms (JPDA$_{m}$, SJPDA, RBPF, and RBPF with EAS) used 100 candidates. The observation probability $p(y_j|\hat{x}_i)$ was modeled as follows:

$$p(y_j|\hat{x}_i) = \begin{cases} 1 & i = 0 \text{ (false detection)}, \\ \mathcal{N}(y_j, \hat{x}_i, P_j) & 0 < i \leq M \text{ (targets)}, \end{cases}$$

(14)

where $V$ is the volume of the detection area, which in this case was set to $V = 10$. Because false detections are introduced in this evaluation, but the number of targets does not change, we introduced only one dummy target for false detections $\hat{x}_0$. The evaluation was performed by a single core of a computer with an Intel Core i7-4770K CPU running at 3.50 GHz with 32 GB of RAM.

A sample of detections and tracking results are shown in Fig.5, where it can be observed that the Hungarian method and RBPF produced numerous missed tracks as a result of the large number of cluttered detections. The Hungarian method tended to switch targets, while the RBPF tracked noise. In contrast, JPDA, JPDA$_{m}$, and the two proposed methods were all able to track targets in the presence of cluttered detections.

We compared the run times and accuracy of the target positions; the results are listed in Table 1, from which it can be observed that both of the proposed methods achieved significantly improved performance.

| Method          | Run Time [ms] | Accuracy MAE | RMSE |
|-----------------|---------------|--------------|------|
| Hungarian       | 6.92          | 1.55         | 2.72 |
| JPDA [1]        | 5965.17       | 0.22         | 0.28 |
| JPDA$_{m}$ [3]  | 168.42        | 0.18         | 0.23 |
| SJPDA (ours)    | 61.81         | 0.17         | 0.26 |
| RBPF [6]        | 707.25        | 1.11         | 2.38 |
| RBPF+EAS (ours) | 718.53        | 0.47         | 1.04 |

Table 1. Evaluation results for simulations. SJPDA is faster than the JPDA$_{m}$ without any loss of accuracy. RBPF with EAS is significantly more accurate than the original RBPF with almost the same speed.

SJPDA was faster than and almost as accurate as JPDA$_{m}$. Although both methods are approximations of JPDA, they are more accurate than the original JPDA. This may be because these methods exclude improbable associations during the approximation process. Even though RBPF with EAS had a similar run time as conventional RBPF, its accuracy was considerably higher. These results suggest that applying exclusivity to the data associations improved the robustness of RBPF against cluttered detection. The accuracy of both proposed methods was also considerably higher than that of the Hungarian method. The JPDA and RBPF perform probability-weighted and multi-candidate-based tracking, which provides the robustness with crossed targets. The EAS can improve the efficiency of the JPDA and RBPF. This causes the increase in speed of the JPDA, and the improvement of the accuracy of the RBPF with the same number of candidates. Although JPDA was originally more accurate than the Hungarian method, it required more calculation time, however the EAS improved the calculation speed. The improved RBPF achieves higher accuracy than that of the Hungarian method. Overall, the results validate the proposed approach in terms of its ability to enhance the performance of Bayesian multi-target tracking methods.

B. EVALUATION ON OPEN DATASET OF PEDESTRIAN TRACKING WITH A SINGLE CAMERA

1) MOT16 DATASET

Next, we evaluated the proposed methods using real-world data by applying them to the MOT Challenge, a recently developed multiple object tracking (MOT) benchmark [26]. The MOT Challenge provides various video sequences and detections of pedestrians. Only trackers can be evaluated using these detections. We employed the MOT 16 datasets [29] and conducted an evaluation based on the popular CLEAR MOT performance metrics of the MOT Challenge [30]. These include multiple-object tracking accuracy (MOTA), which combines errors such as false positives (FP), false negatives (FN), and identity switches (ID Sw.) into a single number. Next, the multiple-object tracking precision (MOTP) is the overlapped ratio between the tracking result and ground truth rectangles. In addition, the mostly tracked (MT) and mostly lost (ML) scores [31] represent how many targets are tracked for more than 80% and less than 20% of their life-spans, respectively, based on ground-truth trajectories.

Neither SJPDA nor RBPF with EAS includes a mechanism for tracking a time-varying number of targets. We therefore introduced a special assignment to generate new targets and an existence probability that increases when an observation is assigned and decreases at the prediction step, as follows:

$$p(e_{i,k}|y_j) = \begin{cases} p_{e} \cdot \theta_{ij} & \theta_{ij} > 0, \\ p_{e}(e_{i,k-1}) - p_{d} & \theta_{ij} = 0, \end{cases}$$

(15)

where $p(e_{i,k})$ is the existence probability of the $i$-th target; $p_{d}$ is the decrease in the existence probability in one step and is set as $4.0\Delta t$, where $\Delta t$ is a sampling interval; and $p_{e}$ is the detection precision, which is set as 0.8. Targets that have an existence probability below 0.1 after prediction are removed. The observation probability is modeled to be dependent on the existence probability and the detection precision.
as follows:

\[
p(y_j|x_i) = \begin{cases} 
\frac{1}{V} & i = 0 \text{ (false detection)}, \\
p(e_i) \cdot p_p \cdot p(y_j, \hat{x}_i) & 0 < i \leq M \text{ (targets)}, \\
\frac{1}{V} & i = M + 1 \text{ (new target)},
\end{cases}
\]

where \( V \) is defined by the size of the camera image. The MOT challenge datasets contain low-confidence detections. We utilized detections with a higher confidence than the threshold value \( c = 0.0 \).

The MOT 16 evaluation results are listed in Table 2. In this study, the focus is only on position tracking and appearance is not considered. Representative trackers (as of 2020/06/04) that do not use appearance are shown at the top. The two proposed methods are compared with the original methods in the bottom half. SJPDA and RBPF with EAS are confirmed to have higher performance than the original methods. SJPDA is considerably faster than JPDA\(_m\) and performs better on the main metric, MOTA. Although this is partly due to the different parameter settings, SJPDA achieves a higher recall rate than JPDA\(_m\), although it has more false positives than JPDA\(_m\). In contrast, JPDA\(_m\) has a low false positive rate but...
FIGURE 6. Tracking results obtained using the SJPDA and RBPF with EAS on the MOT16 [29] datasets. The left panel presents the result of SJPDA, and the right panel displays the result of RBPF with EAS. The frame of the detection result is drawn in a different color for each tracking ID. To some extent, both methods can track pedestrians in image sequences without any appearance information.
### TABLE 2. Tracking performances evaluated using the MOT 16 [29]. The tracking performances of representative trackers that do not use the appearance are shown at the top. The performances of the two proposed methods and the conventional methods are listed in the lower half. The proposed methods achieved better performance than the conventional methods. Moreover, SJPDA is significantly faster than the representative methods and has comparable accuracy, and the MT and ML of RBPF with EAS are the best among the considered methods.

| Tracker       | Online | Hz  | MOT | MOTP | MT | ML | FP | FN | Recall | Precision | ID Sw. |
|---------------|--------|-----|-----|------|----|----|----|----|--------|-----------|------|
| EAMTT [32]    | yes    | 11.8| 33.8| 75.1 | 60| 373| 373| 11184 | 49.3   | 90.8      | 965  |
| LPFD [33]     | no     | 49.3| 35.7| 75.8 | 66| 385| 385| 5054  | 39.0   | 93.3      | 915  |
| GMPHDNIT [20] | yes    | 8.8 | 33.3| 76.8 | 42| 425| 425| 1750  | 36.1   | 97.4      | 3499 |
| CEM [34]      | no     | 0.3 | 33.2| 75.8 | 59| 413| 413| 6837  | 37.3   | 90.9      | 642  |
| JPDAm [3]     | no     | 22.2| 26.2| 76.3 | 51| 312| 312| 3689  | 31.0   | 93.3      | 365  |
| SJPDA(ours)   | yes    | 83.8| 35.0| 75.2 | 64| 379| 379| 10936 | 41.9   | 87.5      | 1538 |
| RBPF [6]      | no     | 9.1 | 33.8| 74.7 | 43| 386| 386| 7721  | 38.9   | 90.2      | 1555 |
| RBPF+EAS(ours)| no     | 7.6 | 33.9| 74.8 | 70| 344| 344| 15145 | 43.5   | 84.0      | 2264 |

* This method has not been submitted to MOT 16; we therefore implemented and submitted it.

### TABLE 3. Tracking performance evaluated using the PETS2009-S2L2 [27]. The tracking performances of representative trackers that do not use the appearance are shown at the top. The performances of the two proposed methods and the conventional methods are listed in the lower half.

| Tracker       | Online | Hz  | MOT | MOTP | MT | ML | FP | FN | Recall | Precision | ID Sw. |
|---------------|--------|-----|-----|------|----|----|----|----|--------|-----------|------|
| GMPHD_OGM [21]| yes    | -   | 45.7| 69.5 | 5  | 4  | 1060| 4247| 58.7   | 85.1      | 286  |
| AdTobKF [37]  | yes    | -   | 37.1| 69.1 | 2  | 8  | 778 | 5474| 46.8   | 86.1      | 225  |
| TENSOR [38]   | no     | -   | 43.7| 70.3 | 5  | 7  | 900 | 4484| 56.4   | 86.6      | 412  |
| EAMTT [32]    | yes    | -   | 38.7| 69.7 | 2  | 8  | 691 | 5388| 49.7   | 76.7      | 221  |
| JPDAm [3]     | no     | -   | 36.3| 65.9 | 3  | 10 | 959 | 5452| 47.0   | 83.5      | 142  |
| SJPDA(ours)   | yes    | 89.3| 36.8| 68.4 | 1  | 8  | 957 | 5316| 48.3   | 83.9      | 233  |
| RBPF [6]      | no     | 5.0 | 40.5| 68.4 | 2  | 4  | 1224| 4592| 55.4   | 82.3      | 304  |
| RBPF+EAS(ours)| no     | 4.8 | 44.5| 69.1 | 6  | 3  | 1351| 4019| 61.0   | 82.3      | 341  |

* We implemented and evaluated this method.

### TABLE 4. Tracking performance evaluated using the TUD-Crossing [28]. The tracking performances of representative trackers that do not use appearance are shown at the top. The performances of the two proposed methods and the conventional methods are presented in the lower half.

| Tracker       | Online | Hz  | MOT | MOTP | MT | ML | FP | FN | Recall | Precision | ID Sw. |
|---------------|--------|-----|-----|------|----|----|----|----|--------|-----------|------|
| GMPHD_OGM [21]| yes    | -   | 60.4| 73.1 | 5  | 2  | 109 | 311 | 71.8   | 87.9      | 16   |
| AdTobKF [37]  | yes    | -   | 53.0| 72.8 | 2  | 3  | 23  | 482 | 56.3   | 96.4      | 13   |
| TENSOR [38]   | no     | -   | 54.5| 73.0 | 3  | 2  | 64  | 397 | 64.0   | 91.7      | 40   |
| EAMTT [32]    | yes    | -   | 46.0| 72.9 | 3  | 8  | 110 | 436 | 54.6   | 85.8      | 27   |
| JPDAm [3]     | no     | -   | 60.9| 74.2 | 4  | 3  | 34  | 335 | 65.3   | 94.2      | 2    |
| SJPDA(ours)   | yes    | 40.8| 71.0| 74.2 | 7  | 1  | 55  | 252 | 77.1   | 93.9      | 13   |
| RBPF [6]      | no     | 2.1 | 44.4| 77.1 | 6  | 0  | 261 | 334 | 69.6   | 74.6      | 17   |
| RBPF+EAS(ours)| no     | 1.8 | 51.0| 72.9 | 10 | 0  | 331 | 188 | 82.9   | 73.4      | 21   |

* We implemented and evaluated this method.

S. Taguchi, K. Kidono: EAS to Improve Bayesian Multi-Target Tracking

a high number of missed detections. SJPDA is also faster than all the other methods and performs as well as the representative methods. RBPF with EAS also performs better than the original RBPF in terms of the MOTA and achieves a higher recall rate than the original RBPF. These results suggest that the exclusive association is effective for improving the performance on these metrics. Furthermore, the MT and ML of RBPF with EAS are the best among the considered methods. RBPF with EAS requires offline processing, and the computation time is not fast. However, it performs well for continuous tracking. These results confirm that EAS can improve Bayesian multi-target tracking for real scenes and that the proposed methods achieve performances comparable to or better than the representative methods. The result shows that although the proposed methods have heuristic procedures concerning false detections and new targets, they work well in a real situation with false alarms, miss-detection, and new targets. In particular, the proposed method is superior than a random-finite-set-based approach, GMPHDNIT [20], which suffers from many ID switches.

Images of the tracking results obtained using SJPDA and RBPF with EAS are presented in Fig. 7. Successful tracking of pedestrians is shown in the same color frame. To some extent, both methods can track pedestrians in image sequences without any appearance information. Because the proposed methods are based on Bayesian filtering, they can exploit higher-order information such as appearance information, like MHT-DAM [4]. To create an application that focuses on tracking images, such enhancements should be included to improve the performance.

2) PETS2009 AND TUD DATASET

We also evaluated the proposed methods using the PETS2009 [27] and TUD [28] benchmarks. These datasets are included...
in the 2DMOT15 [35] benchmark, which is a previous benchmark of MOT16 and provides detections to evaluate the pure tracking performance. 2DMOT15 contains the same sequences as MOT16; however, the data, PETS2009-S2L2 and TUD-Crossing, are not included in MOT16. We apply the proposed method by using the detection of 2DMOT15 and evaluate its performance using the ground-truth data provided in [36]. The results of PETS2009-S2L2 and TUD-Crossing are listed in Tables 3 and 4, respectively. Representative trackers (as of 2020/09/25) that do not use appearance are presented at the top. The two proposed methods are compared with the conventional methods in the bottom half. The performance results of competitors are also calculated from the raw results of 2DMOT15 provided in [26] and the ground-truth data provided in [36]. These results do not contain speed information; therefore, only the speed of the proposed methods and the RBPF are evaluated. The MOTA of SJPDA is more accurate than that of JPDA\textsubscript{m}, and most of the other metrics of SJPDA are also better than those of JPDA. Although the speed of JPDA\textsubscript{m} has not been evaluated in these datasets, SJPDA can be expected to be faster than JPDA\textsubscript{m} based on previous results. The RBPF with EAS achieves better performance than the original RBPF on most metrics. In addition, the proposed methods achieved as good a performance as the representative methods. SJPDA achieves the highest MOTA on PETS2009-S2L2. RBPF with EAS achieves the best MT and ML on both datasets. In these datasets, GMMPHD_OGM [21], which is a GMMPHD-based tracking method, achieves good performance. This may be due to the fact that this method uses information about overlaps of detection boxes to improve accuracy.

Images of the tracking results obtained using SJPDA and RBPF with EAS are presented in Fig. 7. Successful tracking of pedestrians is shown in the same color frame. Both methods can track pedestrians in image sequences without any appearance information. These results reveal that the SJPDA has more misdetections, and RBPF has more false positives, which may be responsible for the differences in accuracy in each dataset. In these experiments, we employed the same parameters for both methods to ensure a fair evaluation; however, in practice, it is better to choose the appropriate parameter for each method separately.

VII. CONCLUSION

We presented a novel sampling method to enhance probabilistic data association methods. The proposed method, called EAS, applies probabilistic and exclusive sampling of data associations. EAS can approximate joint probabilities through the application of easily obtained weights. Based on the EAS scheme, we proposed two multi-target tracking methods: SJPDA and RBPF with EAS. The simulation results demonstrate that SJPDA is faster than JPDA\textsubscript{m} [3], while RBPF with EAS achieves considerably higher accuracy than the conventional method [6]. Furthermore, the results of the

![Tracking results using SJPDA and RBPF with EAS on the (a) PETS2009 and (b) TUD datasets. The left panel presents the result of SJPDA, and the right panel displays the result of RBPF with EAS. The frame of the detection result is drawn in a different color for each tracking ID.](image-url)
evaluation using open datasets on pedestrian tracking show that the proposed methods using EAS achieve higher performance than the original methods. The proposed methods achieved significantly better performance than the representative methods on some important metrics.

REFERENCES

[1] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, “Sonar tracking of multiple targets using joint probabilistic data association,” IEEE J. Ocean. Eng., vol. JOE-8, no. 3, pp. 173–184, Jul. 1983.
[2] D. Reid, “An algorithm for tracking multiple targets,” IEEE Trans. Autom. Control, vol. AC-24, no. 6, pp. 843–854, Dec. 1979.
[3] S. H. Rezafatofighi, A. Milan, Z. Zhang, Q. Shi, A. Dick, and I. Reid, “Joint probabilistic data association revisited,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 3047–3055.
[4] C. Kim, F. Li, A. Ciptadi, and J. M. Rehg, “Multiple hypothesis tracking revisited: Blending in modern appearance mode,” presented at the ICCV, 2015.
[5] S. Maresca, P. Braca, J. Horstmann, and R. Grasso, “Maritime surveillance using multiple high-frequency surface-wave radars,” IEEE Trans. Geosci. Remote Sens., vol. 52, no. 8, pp. 5056–5071, Aug. 2014.
[6] S. Särkkä, A. Vehtari, and J. Lampinen, “Rao-Blackwellized particle filter for multiple target tracking,” Inf. Fusion, vol. 8, no. 1, pp. 2–15, Jan. 2007.
[7] B. N. Vo and W. K. Ma, “The Gaussian mixture probability hypothesis density filter,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4091–4104, Nov. 2006.
[8] L. Zhang, Y. Li, and R. Nevatia, “Global data association for multi-object tracking using network flows,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2008, pp. 1–8.
[9] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes, “Globally-optimal greedy algorithms for tracking a variable number of objects,” in Proc. CVPR, Jun. 2011, pp. 1201–1208.
[10] A. A. Butt and R. T. Collins, “Multi-target tracking by Lagrangian relaxation to min-cost network flow,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2013, pp. 1846–1853.
[11] J. Berclaz, F. Fleuret, E. Turetken, and P. Fua, “Multiple object tracking using K-Shortest paths optimization,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 9, pp. 1806–1819, Sep. 2011.
[12] Y. M. Song and M. Jeon, “Online multiple object tracking with the hierarchically adopted GM-PHD filter using motion and appearance,” in Proc. IEEE Int. Conf. Consum. Electron.-Asia (ICCE-Asia), Oct. 2016, pp. 1–4.
[13] Z. Fu, F. Angelini, J. Chambers, and S. M. Naqvi, “Multi-level cooperative fusion of GM-PHD filters for online multiple human tracking,” IEEE Trans. Multimedia, vol. 21, no. 9, pp. 2277–2291, Sep. 2019.
[14] A. Milan, S. H. Rezafatofighi, A. Dick, I. Reid, and K. Schindler, “Online multiple-target tracking using recurrent neural networks,” in Proc. AAAI, 2017, pp. 1–17.
[15] L. Leaf-Taixe, C. Canton-Ferrer, and K. Schindler, “Learning by tracking: Siamese CNN for robust target association,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2016, pp. 418–425.
[16] Y. Bar-Shalom, A. Alahi, and S. Savarese, “Tracking the untrackable: Learning to track multiple cues with long-term dependencies,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 300–311.
[17] S. Maresca, P. Braca, J. Horstmann, and R. Grasso, “Maritime surveillance using multiple high-frequency surface-wave radars,” IEEE Trans. Geosci. Remote Sens., vol. 52, no. 8, pp. 5056–5071, Aug. 2014.
[18] S. Särkkä, A. Vehtari, and J. Lampinen, “Rao-Blackwellized particle filter for multiple target tracking,” Inf. Fusion, vol. 8, no. 1, pp. 2–15, Jan. 2007.
[19] B. N. Vo and W. K. Ma, “The Gaussian mixture probability hypothesis density filter,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4091–4104, Nov. 2006.
[20] C. Kim, F. Li, A. Ciptadi, and J. M. Rehg, “Multiple hypothesis tracking revisited: Blending in modern appearance mode,” presented at the ICCV, 2015.
[21] A. A. Butt and R. T. Collins, “Multi-target tracking by Lagrangian relaxation to min-cost network flow,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2013, pp. 1846–1853.
[22] J. Berclaz, F. Fleuret, E. Turetken, and P. Fua, “Multiple object tracking using K-Shortest paths optimization,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 9, pp. 1806–1819, Sep. 2011.
[23] Y. M. Song and M. Jeon, “Online multiple object tracking with the hierarchically adopted GM-PHD filter using motion and appearance,” in Proc. IEEE Int. Conf. Consum. Electron.-Asia (ICCE-Asia), Oct. 2016, pp. 1–4.
[24] A. Milan, S. H. Rezafatofighi, A. Dick, I. Reid, and K. Schindler, “Online multiple-target tracking using recurrent neural networks,” in Proc. AAAI, 2017, pp. 1–17.
[25] L. Leaf-Taixe, C. Canton-Ferrer, and K. Schindler, “Learning by tracking: Siamese CNN for robust target association,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2016, pp. 418–425.