Supplementary Material for "Online and Real-Time Tracking with the GM-PHD Filter using Group Management and Relative Motion Analysis"

Anonymous AVSS submission for Double Blind Review

Paper ID 15

In this supplementary material, we present the Gaussian mixture probability hypothesis density (GM-PHD) filtering theory.

1. The GM-PHD Filter

In section 1.1, we briefly introduce how to adopt the GM-PHD filter to our tracking work and its tracking process. In section 1.2, we present the parameter settings in the GM-PHD filter implementation.

1.1. Online Tracking Process

How to adopt the GM-PHD filter The original domain of the GM-PHD filter is radar/sonar data. Why the GM-PHD filter, rather than the particle PHD filter. clustering consume a lot of computing resource random probability removed.

This section gives how we revised the general tracking process of the GM-PHD filter. The PHD filter was originally devised for multi-sensor and multi-target tracking whose target has generally point data form such as (x, y). Dr. Mahler [1] defined targets as random finite set and approximated multi-target Bayes recursion by propagating the PHD filter. However, the PHD filter is still restricted to dealing with practical data. Then, Dr. Vo [2] implemented the PHD filter, using the Gaussian mixture model, named as the GM-PHD Filter.

In MOT of video data, the states and the observations imply the targets states and detection responses, respectively. States and Observations are represented as follows:

$$X_k = \{x_{1,k}, ..., x_{L_k,k}\}, \tag{1}$$

$$Z_k = \{z_{1,k}, ..., z_{N_k,k}\}, \tag{2}$$

where X_k and L_k denote a set of targets' states and the number of them at time k, respectively. A state vector x_k consists of (x, y, v_x, v_y), where x, y, v_x, and v_y indicate the x-axis center of the bounding box, the y-axis center of the bounding box, the x-axis velocity, and the y-axis velocity in y-axis, respectively. Likewise, Z_k and N_k denote a set of observations and the number of them at time k, respectively. An observation z_k consists of (x, y), where x and y indicate the x-axis and the y-axis center of the detection bounding box, respectively.

The tracking process of the GM-PHD filter is composed of four steps which are *Initialization*, *Prediction*, *Update*, and *Pruning* as follows.

Initialization Step:

$$\sum_{i=1}^{J_0} w_0^i \mathcal{N}(x; m_0^i, P_0^i), \tag{3}$$

where the Gaussian mixture models are initialized by using the initial observations from detection responses. Besides, when an observation is failed to find the association pair, i.e., updating target state, the observation initializes a new Gaussian model. We call that kind of initialization as *birth*. J_0 means the number of Gaussian models. Each Gaussian $\mathcal{N}$ represents each state model with weight w, mean vector m, input state vector x, and covariance matrix P. At this step, we set the initial velocities of mean vector to zeros. Each weight is set to the normalized confidence value of corresponding detection response. Also, how to set covariance matrix P is shown in subsection 1.2.

Prediction Step:

$$\sum_{i=1}^{J_{k-1}} w_{k-1}^i \mathcal{N}(x; m_{k-1}^i, P_{k-1}^i), \tag{4}$$

$$m_{k|k-1}^j = F m_{k-1}^j, \tag{5}$$

$$P_{k|k-1}^j = Q + F_{k-1} P_{k-1}^j (F)^T, \tag{6}$$

where we assume that we have already the Gaussian mixture of the target states at the previous frame $k-1$ as shown in (4). F is the state transition matrix and Q is the process noise covariance matrix. Those two matrices are constants in our tracker. Then, we can predict the state at time k using the Kalman filtering. In (5), $m_{k|k-1}^j$ is derived by using the velocity at time $k-1$. Covariance P is also predicted by the Kalman filtering method in (6).

Update Step:

$$\sum_{i=1}^{J_{k|k-1}} w_k^i(z)\mathcal{N}(x; m_{k|k}^i, P_{k|k}^i), \tag{7}$$

$$q_k^j(z) = \mathcal{N}(z; Hm_{k|k-1}^j, R + HP_{k|k-1}^j(H)^T), \tag{8}$$

$$w_k^i(z) = \frac{w_{k|k-1}^j q_k^j(z)}{\sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^l q_k^l(z)}, \tag{9}$$

$$m_{k|k}^j(z) = m_{k|k-1}^j + K_k^j(z - Hm_{k|k-1}^j), \tag{10}$$

$$P_{k|k}^j = [I - K_k{}^j H]P_{k|k-1}^j, \tag{11}$$

$$K_k^j = P_{k|k-1}^j(H)^T(HP_{k|k-1}^j(H)^T + R)^{-1}, \tag{12}$$

where the goal of update step is deriving (7). First, we should find an optimal observation z at time k to update a Gaussian model. The optimal z makes q_k into the maximum value in (8). In the perspective of application, the update step involves data association. Finding the optimal observations and updating the state models is equal to finding the association pairs. R is the observation noise covariance. H is the observation matrix to transit a state vector to an observation vector. Both matrices are constants in our application. After finding the optimal z, the Gaussian mixture is updated by (9)(10)(11)(12).

Pruning Step:

$$\tilde{X}_k = \{m_k^i : w_k^i \geq \theta_w, i = 1, ..., J_k\}, \tag{13}$$

$$\tilde{W}_k = \{w_k^j : m_k^j \in \tilde{X}_k, j = 1, ..., J_k\}, \tag{14}$$

$$\tilde{W}_k = \{\tilde{w_{k,1}}, ..., \tilde{w}_{k,\tilde{J}_k}\}, \tilde{J}_k = |\tilde{W}_k|, \tag{15}$$

$$w_k^i = \frac{\tilde{w}_k^i}{\sum_{l=1}^{\tilde{J}_k} \tilde{w}_k^l}, \tag{16}$$

$$X_k = \tilde{X}_k, \tag{17}$$

where the states with the weight under threshold θ_w are pruned as in (13). We experimentally set θ_w to 0.1. Then re-normalize the weights of the surviving states as in (16). Pruning step handles the false positives by the clutter.

1.2. Parameter Settings

The matrices F, Q, P, R, and H are used in *Prediction Step* and *Update Step*. Experimentally, We set the parameter matrices for the GM-PHD filter's tracking process as follows:

$$F = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, Q = \frac{1}{2}\begin{pmatrix} 5^2 & 0 & 0 & 0 \\ 0 & 10^2 & 0 & 0 \\ 0 & 0 & 5^2 & 0 \\ 0 & 0 & 0 & 10^2 \end{pmatrix},$$

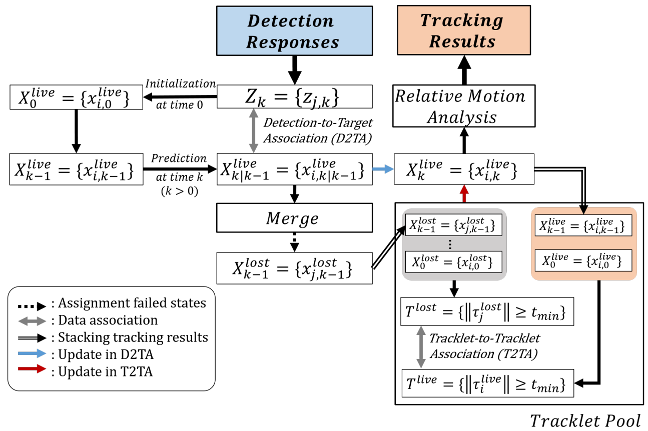

Figure 1. Flow chart of the proposed tracking framework with the two-stage association strategy (D2TA and T2TA) using group management and relative motion analysis (Merge and Relative Motion Analysis).

$$P = \begin{pmatrix} 5^2 & 0 & 0 & 0 \\ 0 & 10^2 & 0 & 0 \\ 0 & 0 & 5^2 & 0 \\ 0 & 0 & 0 & 10^2 \end{pmatrix}, R = \begin{pmatrix} 5^2 & 0 \\ 0 & 10^2 \end{pmatrix},$$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

References

[1] R. P. Mahler. Multitarget bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4), Oct. 2003.

[2] B.-N. Vo and W.-K. Ma. The gaussian mixture probability hypothesis density filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104, Nov. 2006.