Data Association

CS 585 Lecture

Based on the book
“Data Association for Multi-Object Visual Tracking,”
Margrit Betke & Zheng Wu, 2016,
www.morganclaypool.com
Global Nearest Neighbor Standard Filter (GNNSF)

- Considers all possible measurement-to-track assignments within appropriate gating regions in each frame and generates the most likely assignment hypothesis.
- Solves a 2D binary assignment problem.
- The assignments are irrevocable and cannot be modified by future data.
- Advantage of GNNSF: Simple
- Disadvantage of GNNSF: Helpful future data will be overlooked
GNNSF in Computer Vision
Figure 2.1: Example of data association problem with Global Nearest Neighbor Standard Filter (GNNSF). The scene contains three pedestrians walking towards each other. At time $t$, two measurements $z_1$ and $z_2$ were returned by the pedestrian detector, while one pedestrian who is occluded by the street sign was missed. Detections falling into the gating region of an object, shown as a shaded region, are potential assignment candidates. The gating region is typically a hyper-ellipsoid if the residual measurement vector $\tilde{z}$ is Gaussian distributed. For the GNNSF method, the cost of pairing of a measurement $z$ and a predicted position $\hat{x}(t|t-1)$ is defined as the Mahalanobis distance between them. The cost to identify a missed detection or a false alarm also needs to be properly modeled with prior knowledge of detection and false alarm rates. GNNSF models the 2D assignment combinatorial problem as a bipartite graph where nodes on the left column represent the objects and nodes on the right column represent the detections. Possible pairings are expressed as edges between nodes with their associated costs. The node with a zero index, also known as a “dummy node,” is used as the placeholder for missing detections or false alarms. The optimal solution to this 2D bipartite assignment problem gives the best set of matchings such that, 1) every detection and object are assigned, and 2) the total cost is a minimum.
GNNSF

1. Predict the measurements and their covariances to estimate the validation gates.

2. Compute the cost for all possible measurement-to-track assignments within the each gating region.

3. Formulate the 2D assignment problem and obtain a global optimal solution as the best assignment hypothesis.

4. Perform tracking by updating the state of each object and its covariance from the assignment result.
GNNSF

1. Predict the measurements and their covariances to estimate the validation gates. **Use Kalman Filter**
2. Compute the cost for all possible measurement-to-track assignments within the each gating region.
3. Formulate the 2D assignment problem and obtain a global optimal solution as the best assignment hypothesis.
4. Perform tracking by updating the state of each object and its covariance from the assignment result. **Use Kalman Filter**
1. Predict the measurements and their covariances to estimate the validation gates. Use Kalman Filter

2. Compute the cost for all possible measurement-to-track assignments within the each gating region.

3. Formulate the 2D assignment problem and obtain a global optimal solution as the best assignment hypothesis.

4. Perform tracking by updating the state of each object and its covariance from the assignment result. Use Kalman Filter
GNNSF

Minimize cost function

\[
\min_{x_{i,j}} \sum c_{i,j} x_{i,j}
\]

s.t.

\[
\sum_{i:i>0} x_{i,j} = 1
\]

\[
\sum_{j:j>0} x_{i,j} = 1
\]

\[
x_{i,j} \in \{0, 1\}
\]

\[x_{i,j} = \text{binary variable to assign the } i\text{th object to the } j\text{th measurement, and } c_{i,j} \text{ is the corresponding cost.}\]

Add a “dummy object" and a “dummy measurement," both indexed by zero, to represent missing detections and false alarms, respectively.

The constraints enforce the assignment solution to be exhaustive and exclusive, except for the dummy object/measurement.
Solution to GNNSF

= Solutions to the Bipartite Matching Problem

Hungarian Method  $O(n^3)$ [Optimal]

Greedy Method $O(n^2)$ [Suboptimal]
MULTIPLE HYPOTHESES TRACKING (MHT)

Two main ideas:

1. Hypothesize all possible data associations over time
2. Use measurements that are received in later frames to resolve ambiguities in the current frame

Reid [1979]
Properties of Multiple Hypotheses Tracking

MHT

• propagates the current hypotheses in anticipation of subsequent data for better estimation
• provides a principled formulation to handle the complete life cycle of tracks including birth, growth and termination
• has exponential computational cost for the exponentially growing number of hypotheses
• needs heuristic techniques to enable a real-time performance
• most widely used algorithm in 40 years since its introduction for radar & 20 years since its introduction for video processing
Key Computation of MHT

• Evaluate the probability of a new set $\theta_{\ell}(t)$ of assignment hypotheses at time $t$ using Bayes rule.

• Set $\theta_{\ell}(t) = \ell$th hypothesis of a joint cumulative event (set of association histories) at $t$

• $Z(t)$ is the measurement at time $t$

$$P(\Theta_{\ell}(t) | Z(t))$$
Key Computation of MHT

• Evaluate the probability of a new set $\theta_{l}^{(t)}$ of assignment hypotheses at time $t$ using Bayes rule.

• Set $\theta_{l}^{(t)}$ is made up of the parent event $\theta_{m(l)}^{(t-1)}$ through $t-1$ and the offspring event (= current association event) $\theta(t)$

• $Z(t)$ is the measurement at time $t$

$$P(\Theta_{l}^{(t)} | Z^{(t)}) = P(\Theta_{m(l)}^{(t-1)}, \theta(t) | Z^{(t)})$$
Key Computation of MHT

- Set $\theta_{\ell}^{(t)}$ is made up of the parent event $\theta_{m(\ell)}^{(t-1)}$ through t-1 and the offspring event (= current association event) $\theta(t)$
- $P(Z(t) \mid .)$ is the likelihood of the measurements $Z(t)$ at current time t
- $P(\theta(t) \mid .)$ is the probability of a current association hypothesis
- $P(\Theta_{m(l)}^{(t-1)} \mid Z^{(t-1)})$ is the prior probability from the previous step

\[
P(\Theta_{l}^{(t)} \mid Z^{(t)}) = P(\Theta_{m(l)}^{(t-1)}, \theta(t) \mid Z^{(t)})
\]

\[
\propto P(Z(t) \mid \Theta_{m(l)}^{(t-1)}, \theta(t), Z^{(t-1)}) \times P(\theta(t) \mid \Theta_{m(l)}^{(t-1)}, Z^{(t)})
\]

\[
\times P(\Theta_{m(l)}^{(t-1)} \mid Z^{(t-1)})
\]
Algorithm for Computing Hypotheses Set

At each time step:

• Carry over $\theta_{m(\ell)}^{(t-1)}$

• Expand the hypothesis by enumerating all possible measurement-to-track pairings with the associated probability according to

$$P(\Theta_l^{(t)} | Z^{(t)}) = P(\Theta_m^{(t-1)} | \theta(t), Z^{(t)})$$

$$\propto P(Z(t) | \Theta_m^{(t-1)}, \theta(t), Z^{(t-1)}) \times P(\theta(t) | \Theta_m^{(t-1)}, Z^{(t)}) \times P(\Theta_m^{(t-1)} | Z^{(t-1)})$$
How to Compute $P(\Theta_l^{(t)} | Z^{(t)})$

We assume:

$N_d =$ number of current measurements associated with previously known objects, follows a binomial distribution with detection rate $P_D$

$N_{m(l)} =$ number of previously known objects

$N_f =$ number of false alarms, follows Poisson distribution with the density $\lambda_f$

$N_a =$ number of measurements associated with newly arriving objects, Poisson distribution with the density $\lambda_n$

$$P(\Theta_l^{(t)} | Z^{(t)}) \propto P_D^{N_d} (1 - P_D)^{(N_{m(l)} - N_d)} \lambda_f^{N_f} \lambda_n^{N_a} \left[ \prod_{j=1}^{N_d} \mathcal{N}(z_j; \hat{z}_l(t|t-1), S_l(t)) \right]$$

$$\times \quad P(\Theta_m^{(t-1)} | Z^{(t-1)}).$$

Uncertainty in data association

$S =$ Covariance matrix of innovation term in Kalman filter
Pruning the Hypotheses

First: Gating
Then: Note that many inconsequential, low-probability hypotheses do not need to be generated. Select one of two approaches:

1. Hypothesis-oriented MHT = “m-best algorithm” = only the m best hypotheses formed on the current measurements remain
2. Track-oriented MHT = Create a track graph data structure
MHT Track Tree

Figure 2.2: The data structure *track tree* for “track-oriented multiple hypotheses tracking.” Left: An example with measurements in three consecutive frames, where $z_{t,i}$ represents the $i$th measurement at time $t$. Right: A track tree whose root is the first measurement $z_{1,1}$ in the first frame. Squares represent actual measurements. Green circles represent missed detections. Red circles represent track terminations. A second track tree is generated for the first frame, which has the same structure and labels except that its root is labeled $z_{1,2}$. 