Computer Vision 2 – Lecture 9

Multi-Object Tracking II (02.06.2016)

Prof. Dr. Bastian Leibe, Dr. Jörg Stückler
leibe@vision.rwth-aachen.de, stueckler@vision.rwth-aachen.de

RWTH Aachen University, Computer Vision Group
http://www.vision.rwth-aachen.de

Content of the Lecture

- Single-Object Tracking
- Bayesian Filtering
  - Kalman Filters, EKF
  - Particle Filters
- Multi-Object Tracking
  - Introduction
  - MHT, (JPDAF)
  - Network Flow Optimization
- Visual Odometry
- Visual SLAM & 3D Reconstruction

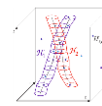
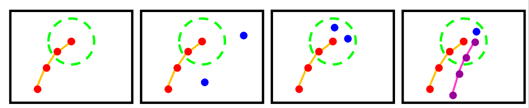


Image sources: Andreas Ess

Topics of This Lecture

- Recap: Track-Splitting Filter
  - Motivation
  - Ambiguities
- Multi-Hypothesis Tracking (MHT)
  - Basic idea
  - Hypothesis Generation
  - Assignment
  - Measurement Likelihood
  - Practical considerations

Recap: Motion Correspondence Ambiguities



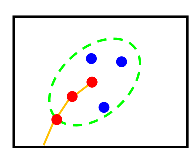
1. Predictions may not be supported by measurements
   - Have the objects ceased to exist, or are they simply occluded?
2. There may be unexpected measurements
   - Newly visible objects, or just noise?
3. More than one measurement may match a prediction
   - Which measurement is the correct one (what about the others)?
4. A measurement may match to multiple predictions
   - Which object shall the measurement be assigned to?

Let's Formalize This

- Multi-Object Tracking problem
  - We represent a track by a state vector $\mathbf{x}$, e.g.,
    $$\mathbf{x} = [x, y, v_x, v_y]^T$$
  - As the track evolves, we denote its state by the time index k:
    $$\mathbf{x}^{(k)} = \left[x^{(k)}, y^{(k)}, v_x^{(k)}, v_y^{(k)}\right]^T$$
  - At each time step, we get a set of observations (measurements)
    $$\mathbf{Y}^{(k)} = \left\{\mathbf{y}_1^{(k)}, \ldots, \mathbf{y}_{M_k}^{(k)}\right\}$$
  - We now need to make the data association between tracks $\left\{\mathbf{x}_1^{(k)}, \ldots, \mathbf{x}_{N_k}^{(k)}\right\}$ and observations $\left\{\mathbf{y}_1^{(k)}, \ldots, \mathbf{y}_{M_k}^{(k)}\right\}$:
    $z_l^{(k)} = j$ iff $\mathbf{y}_j^{(k)}$ is associated with $\mathbf{x}_l^{(k)}$

Mahalanobis Distance

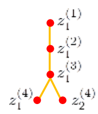
- Additional notation
  - Our KF state of track $\mathbf{x}_l$ is given by the prediction $\mathbf{x}_l^{(k)}$ and covariance $\Sigma_{p,l}^{(k)}$.
  - We define the innovation that measurement $\mathbf{y}_j$ brings to track $\mathbf{x}_l$ at time k as
    $$\mathbf{v}_{j,l}^{(k)} = (\mathbf{y}_j^{(k)} - \mathbf{x}_{p,l}^{(k)})$$



  - With this, we can write the observation likelihood shortly as
    $$p(\mathbf{y}_j^{(k)} | \mathbf{x}_l^{(k)}) \sim \exp\left\{-\frac{1}{2}\mathbf{v}_{j,l}^{(k)T}\Sigma_{p,l}^{(k)^{-1}}\mathbf{v}_{j,l}^{(k)}\right\}$$
  - We define the ellipsoidal gating or validation volume as
    $$V^{(k)}(\gamma) = \left\{\mathbf{y} | (\mathbf{y} - \mathbf{x}_{p,l}^{(k)})^T \Sigma_{p,l}^{(k)^{-1}} (\mathbf{y} - \mathbf{x}_{p,l}^{(k)}) \leq \gamma\right\}$$

Recap: Track-Splitting Filter

- Idea
  - Instead of assigning the measurement that is currently closest, as in the NN algorithm, select the *sequence* of measurements that minimizes the *total* Mahalanobis distance over some interval!
  - Form a track tree for the different association decisions
  - Modified log-likelihood provides the merit of a particular node in the track tree.
  - Cost of calculating this is low, since most terms are needed anyway for the Kalman filter.



- Problem
  - The track tree grows exponentially, may generate a very large number of possible tracks that need to be maintained.

Recap: Pruning Strategies

- In order to keep this feasible, need to apply pruning
  - Deleting unlikely tracks
    - May be accomplished by comparing the modified log-likelihood $\lambda(k)$, which has a χ^2 distribution with kn_z degrees of freedom, with a threshold α (set according to χ^2 distribution tables).
    - Problem for long tracks: modified log-likelihood gets dominated by old terms and responds very slowly to new ones.

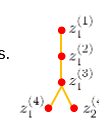
    $\Rightarrow$ Use sliding window or exponential decay term.
  - Merging track nodes
    - If the state estimates of two track nodes are similar, merge them.
    - E.g., if both tracks validate identical subsequent measurements.
  - Only keeping the most likely N tracks
    - Rank tracks based on their modified log-likelihood.

Topics of This Lecture

- Recap: Track-Splitting Filter
  - Motivation
  - Ambiguities
- Multi-Hypothesis Tracking (MHT)
  - Basic idea
  - Hypothesis Generation
  - Assignment
  - Measurement Likelihood
  - Practical considerations

Multi-Hypothesis Tracking (MHT)

- Ideas
  - Again associate sequences of measurements.
  - Evaluate the probabilities of all association hypotheses.
  - For each sequence of measurements (a hypothesized track), a standard KF yields the state estimate and covariance
- Differences to Track-Splitting Filter
  - Instead of forming a track tree, keep a set of hypotheses that generate child hypotheses based on the associations.
  - After each hypothesis generation step, merge and prune the current hypothesis set to keep the approach feasible.
  - Integrate track generation into the assignment process.



D. Reid, An Algorithm for Tracking Multiple Targets, IEEE Trans. Automatic Control, Vol. 24(6), pp. 843-854, 1979.

Target vs. Measurement Orientation

- Target-oriented approaches
  - Evaluate the probability that a measurement belongs to an established target.
- Measurement-oriented approaches
  - Evaluate the probability that an established target or a new target gave rise to a certain measurement sequence.
  - This makes it possible to include track initiation of new targets within the algorithmic framework.
- MHT
  - Measurement-oriented
  - Handles track initialization and termination

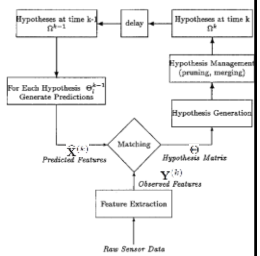
Challenge: Exponential Complexity

- Strategy
  - Generate all possible hypotheses and then depend on pruning these hypotheses to avoid the combinatorial explosion.

    $\Rightarrow$ Exhaustive search
  - Tree data structures are used to keep this search efficient
- Commonly used pruning techniques
  - Clustering to reduce the combinatorial complexity
  - Pruning of low-probability hypotheses
  - N-scan pruning
    - Select a single best hypothesis at frame K and prune all tracks that do not share the predecessor track at the $(K\text{-}N)^{\text{th}}$ frame.
  - Merging of similar hypotheses

Multi-Hypothesis Tracking (MHT)

- Ideas
  - Instead of forming a track tree, keep a set of hypotheses that generate child hypotheses based on the associations.
  - Enforce exclusion constraints between tracks and measurements in the assignment.
  - Integrate track generation into the assignment process.
  - After hypothesis generation, merge and prune the current hypothesis set.



D. Reid, An Algorithm for Tracking Multiple Targets, IEEE Trans. Automatic Control, Vol. 24(6), pp. 843-854, 1979.

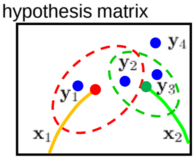
Hypothesis Generation

- Formalization
  - Set of hypotheses at time k: $\mathbf{\Omega}^{(\mathbf{k})} = \left\{\Omega_j^{(k)}\right\}$
  - This set is obtained from $\Omega^{(k\text{-}1)}$ and the latest set of measurements
    $$\mathbf{Y}^{(k)} = \left\{\mathbf{y}_1^{(k)}, \ldots, \mathbf{y}_{M_k}^{(k)}\right\}$$
  - The set $\Omega^{(k)}$ is generated from $\Omega^{(k\text{-}1)}$ by performing all feasible associations between the old hypotheses and the new measurements $\mathbf{Y}^{(k)}$.
- Feasible associations can be
  - A continuation of a previous track
  - A false alarm
  - A new target

Hypothesis Matrix

- Visualize feasible associations by a hypothesis matrix

$$\Theta = \begin{array}{c} \begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_{fa} & \mathbf{x}_{nt} \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \\ \mathbf{y}_4 \end{array}$$



- Interpretation
  - Columns represent tracked objects
  - Rows represent measurements
  - A non-zero element at matrix position (i,j) denotes that measurement $\mathbf{y}_i$ is contained in the validation region of track $\mathbf{x}_j$.
  - Extra column $\mathbf{x}_{fa}$ for association as *false alarm*.
  - Extra column $\mathbf{x}_{nt}$ for association as *new track*.

Assignments

- Turning feasible associations into assignments
  - For each feasible association, we generate a new hypothesis.
  - Let $\Omega_j^{(k)}$ be the j-th hypothesis at time k and $\Omega_{p(j)}^{(k-1)}$ be the parent hypothesis from which $\Omega_j^{(k)}$ was derived.
  - Let $Z_j^{(k)}$ denote the set of assignments that gives rise to $\Omega_j^{(k)}$.
  - Assignments are again best visualized in matrix form

Z_j	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_{fa}$	$\mathbf{x}_{nt}$
$\mathbf{y}_1$	0	0	1	0
$\mathbf{y}_2$	1	0	0	0
$\mathbf{y}_3$	0	1	0	0
$\mathbf{y}_4$	0	0	0	1

Assignments

Z_j	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_{fa}$	$\mathbf{x}_{nt}$
$\mathbf{y}_1$	0	0	1	0
$\mathbf{y}_2$	1	0	0	0
$\mathbf{y}_3$	0	1	0	0
$\mathbf{y}_4$	0	0	0	1

- Impose constraints
  - A measurement can originate from only one object.

  ⇒ Any row has only a single non-zero value.
  - An object can have at most one associated measurement per time step.

  ⇒ Any column has only a single non-zero value, except for $\mathbf{x}_{fa}$, $\mathbf{x}_{nt}$

Calculating Hypothesis Probabilities

- Probabilistic formulation
  - It is straightforward to enumerate all possible assignments.
  - However, we also need to calculate the probability of each child hypothesis.
  - This is done recursively:

$$\begin{aligned} p(\Omega_j^{(k)}|\mathbf{Y}^{(k)}) &= p(Z_j^{(k)}, \Omega_{p(j)}^{(k-1)}|\mathbf{Y}^{(k)}) \\ &\overset{Bayes}{=} \eta p(\mathbf{Y}^{(k)}|Z_j^{(k)}, \Omega_{p(j)}^{(k-1)}) p(Z_j^{(k)}, \Omega_{p(j)}^{(k-1)}) \\ &= \underset{\text{Normalization factor}}{\eta}\, \underbrace{p(\mathbf{Y}^{(k)}|Z_j^{(k)}, \Omega_{p(j)}^{(k-1)})}_{\text{Measurement likelihood}} \underbrace{p(Z_j^{(k)}|\Omega_{p(j)}^{(k-1)})}_{\text{Prob. of assignment set}} \underbrace{p(\Omega_{p(j)}^{(k-1)})}_{\text{Prob. of parent}} \end{aligned}$$

Measurement Likelihood

- Use KF prediction
  - Assume that a measurement $\mathbf{y}_i^{(k)}$ associated to a track $\mathbf{x}_j$ has a Gaussian pdf centered around the measurement prediction $\hat{\mathbf{x}}_j^{(k)}$ with innovation covariance $\widehat{\boldsymbol{\Sigma}}_j^{(k)}$.
  - Further assume that the pdf of a measurement belonging to a new track or false alarm is uniform in the observation volume W (the sensor's field-of-view) with probability W^{-1}.
  - Thus, the measurement likelihood can be expressed as

$$\begin{aligned} p\left(\mathbf{Y}^{(k)}|Z_j^{(k)},\Omega_{p(j)}^{(k-1)}\right) &= \prod_{i=1}^{M_k} \mathcal{N}\left(\mathbf{y}_i^{(k)};\hat{\mathbf{x}}_j,\widehat{\boldsymbol{\Sigma}}_j^{(k)}\right)^{\delta_i} W^{-(1-\delta_i)} \\ &= W^{-(N_{fal}+N_{new})} \prod_{i=1}^{M_k} \mathcal{N}\left(\mathbf{y}_i^{(k)};\hat{\mathbf{x}}_j,\widehat{\boldsymbol{\Sigma}}_j^{(k)}\right)^{\delta_i} \end{aligned}$$

Probability of an Assignment Set

$$p(Z_j^{(k)}|\Omega_{p(j)}^{(k-1)})$$

- Composed of three terms
  1. Probability of the number of tracks N_{det}, N_{fal}, N_{new}
     - Assumption 1: N_{det} follows a binomial distribution

$$p(N_{det}|\Omega_{p(j)}^{(k-1)}) = \binom{N}{N_{det}} p_{det}^{N_{det}}(1-p_{det})^{(N-N_{det})}$$

where N is the number of tracks in the parent hypothesis

     - Assumption 2: N_{fal} and N_{new} both follow a Poisson distribution with expected number of events $\lambda_{fal}W$ and $\lambda_{new}W$

$$\begin{aligned} p(N_{det},N_{fal},N_{new}|\Omega_{p(j)}^{(k-1)}) &= \binom{N}{N_{det}} p_{det}^{N_{det}}(1-p_{det})^{(N-N_{det})} \\ &\cdot \mu(N_{fal};\lambda_{fal}W)\cdot\mu(N_{new};\lambda_{new}W) \end{aligned}$$

Probability of an Assignment Set

2. Probability of a specific assignment of measurements
   - Such that $M_k = N_{det} + N_{fal} + N_{new}$ holds.
   - This is determined as 1 over the number of combinations

$$\binom{M_k}{N_{det}}\binom{M_k-N_{det}}{N_{fal}}\binom{M_k-N_{det}-N_{fal}}{N_{new}}$$

3. Probability of a specific assignment of tracks
   - Given that a track can be either *detected* or not *detected*.
   - This is determined as 1 over the number of assignments

$$\frac{N!}{(N-N_{det})!}\binom{N-N_{det}}{N_{det}}$$

⇒ When combining the different parts, many terms cancel out!

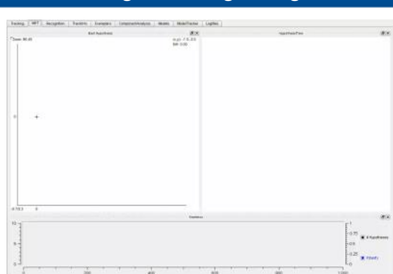
Measurement Likelihood

- Combining all the different parts
  - Nice property: many terms cancel out!
  - (Derivation left as exercise)

  ⇒ The final probability $p\left(\Omega_j^{(k)}|\mathbf{Y}^{(k)}\right)$ can be computed in a very simple form.
  - This was the main contribution by Reid and it is one of the reasons why the approach is still popular.

- Practical issues
  - Exponential complexity remains
  - Heuristic pruning strategies must be applied to contain the growth of the hypothesis set.
  - E.g., dividing hypotheses into spatially disjoint clusters.

Laser-based Leg Tracking using MHT



K. Arras, S. Grzonka, M. Luber, W. Burgard, Efficient People Tracking in Laser Range Data using a Multi-Hypothesis Leg-Tracker with Adaptive Occlusion Probabilities, ICRA'08.

Laser-based People Tracking using MHT

References and Further Reading

- A good tutorial on Data Association
  - I.J. Cox. A Review of Statistical Data Association Techniques for Motion Correspondence. In *International Journal of Computer Vision*, Vol. 10(1), pp. 53-66, 1993.
- Reid's original MHT paper
  - D. Reid, An Algorithm for Tracking Multiple Targets, IEEE Trans. Automatic Control, Vol. 24(6), pp. 843-854, 1979.