Boosting Multi-Hypothesis Tracking by means of Instance-specific Models

Michael Pätzold and Rubén Heras Evangelio and Thomas Sikora

Technische Universität Berlin

21. September 2012

Applications for Tracking in Crowded Environments

- Security Applications
  - Detection of Overcrowded Areas
  - Forensic Search
  - Protection of Restricted Areas
- Data Collection for Marketing Research:
  - Number of People entering Shopping Centers
    $\Rightarrow$ Conversion Rate Computation
  - Dwell Time in Service Areas
    $\Rightarrow$ Customer Satisfaction

flickr.com

Challenges in Crowded Environments

Challenges for Detector:

- Inter-Object-Occlusions
- Cluttered Background

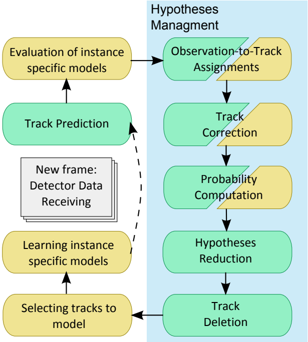
Challenges for Tracking:

- Track Initialization
- Manouvering targets

$\Rightarrow$ An effective tracking system should cope with all these challenges.

Tracking System Overview

- Standard MHT-Algorithm based on Head-Shoulder-Detector
- Usage of instance-specific appearance models
- Training of models for relieable objects only
- Incorporation of model information in MHT probality computation
- Application of instance-specific model information to guide tracks without detections



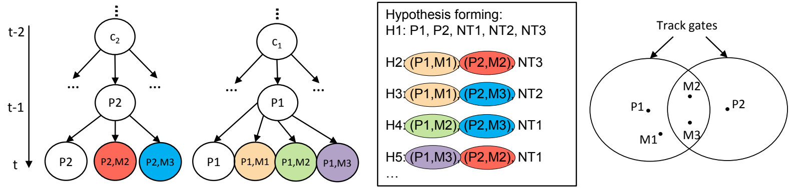
Multi-Hypothesis-Tracking

- Propagation of multiple Hypotheses from one time-step to the next one for the Task of Associating Observations to Tracks
- Each prior hypothesis generates a new set of hypotheses considering the current set of observations
- Individual track j is modelled by Kalman filter with state $\hat{x}_j$
- Computation of posterior probability of hypothesis i at time k [Reid1979]:

$$\begin{aligned} P_i^k &= \frac{1}{c} P_D^{N_{DT}} (1 - P_D)^{(N_{TGT} - N_{DT})} \beta_{FT}^{N_{FT}} \beta_{NT}^{N_{NT}} \\ &\times \left[\prod_{m=1}^{N_{DT}} \mathcal{N}(Z_m - H\hat{x}_j, S_j) \right] P_i^{k-1}, \end{aligned} \tag{1}$$

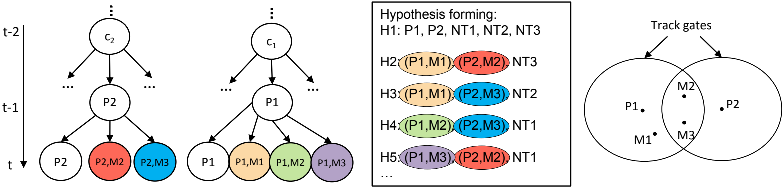
Tree-based MHT Track Managment

- Avoidance of storing redundant data [Blackman1999]
- Store each track state for a time-step in a particular node
- Spawn new child nodes for new possible tracks
- Track history: traverse tree branch from leaf to root
- Reference of all active tracks with common starting observation into the leafs of one track tree ⇒ Garbage Collection by Reference Counting



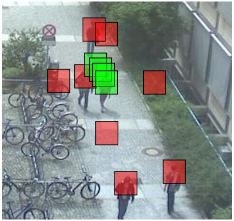
Selecting targets to learn

1. Merge duplicate tracks: $(\hat{x}_k - \hat{x}_j)'[S_j]^{-1}(\hat{x}_k - \hat{x}_j) \leq \sigma$
2. Collect all leaf nodes of all track trees
3. Reduce of number of nodes by replacing each node by its n_{LB}-th parent node
4. Count for each node the number of tracks ϵ that share this node
5. Keep node j if: $\epsilon_j > \frac{\kappa_{max}}{2}$



Learning: Boosting the Instance-specific Model

- Selection of training data similar to [2, 3]:
  - Negative samples: background and nearby tracks
  - Positive samples: history of own track (by ascending track tree)
- Feature pool **h**:
  - Set of haar-like features
  - Bins of normalized RGB-histogram
- Train instance-specific model c by applying AdaBoost to gathered data
- Store instance-specific model c in corresponding model node (used by all child nodes)



Evaluating the Models

- Selection of models to evaluate:
  - All models learned in previous time-step
  - Each model that exist for the tracks of the most probable hypothesis
- Evaluation of c at position x: $c(x) = \sum_{t=1}^{T} \alpha_t h_t(x), h_t \in \mathbf{h}$

Data association aided by the Models

- Up to now: dependance of Reid's MHT posterior probability on kinematics of measurements
- Extension: augmentation of Reid's equation by a term for object appearance similarity by using model c_j:

$$P_i^k = \frac{1}{c} P_D^{N_{DT}} (1 - P_D)^{(N_{TGT} - N_{DT})} \beta_{FT}^{N_{FT}} \beta_{NT}^{N_{NT}} \times \left[\prod_{m=1}^{N_{DT}} \rho \mathcal{N}(z_m - H\hat{x}_j, S_j) + (1 - \rho) \frac{c_j(z_m)}{\tau_{max}} \right] \times P_i^{k-1}, \tag{2}$$

- Normalization constant: τ_{max}
- Parameter for influence of kinematics or appearance: ρ

Unassigned Track Guiding

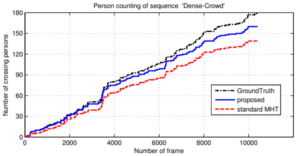
- Guiding of tracks by appearance in case of missing detector observations
- Minimal bounding box Ω including all child nodes of particular model c_j
- Probability map $P_{cj} = c_j(x) : \forall x \in \Omega$
- Computation of potential object positions $Z^j = \{z_1^j, z_2^j, \cdots, z_k^j\}$ by Mean-shift for P_{cj}
- Usage of measurements Z^j for each unassigned Kalman filter which shares model c_j
- Handling of measurement origin uncertainty by PDAF [1]

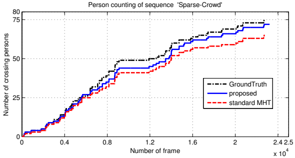
State Estimation by PDA-Filtering

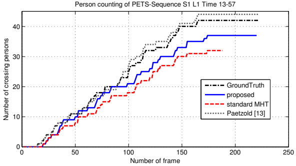
- Computation of association probability β_i^j for z_i^j using:
  - Likelihood Ratio: $\mathcal{L}_i = \frac{\mathcal{N}(z_i^j - H\hat{x}, P(k|k-1)) P_D \frac{c_j(z_i^j)}{\tau_{max}}}{\lambda}$
  - Normalized Confidence in instance-specific model c_j: $\frac{c_j(z_i^j)}{\tau_{max}}$
- State Estimation:
  $\hat{x}(k|k) = \hat{x}(k|k-1) + W(k) \sum_{i=1}^{m} \beta_i^j(k)(z_i^j - H\hat{x}(k|k-1))$
- Kalman Gain: $W(k)$
- Updated Covariance:
  $S(k|k) = \beta_0^j S(k|k-1) + (1-\beta_0^j) S^c(k|k) + \tilde{S}(k)$
- Standard Kalman error covariance: $S^c(k)$
- Measurement origin uncertainty: $\tilde{S}(k)$

Application: Person Counting:

- Standard application
- Counting of tracks crossing a fictive line
- Datasets: TU-Dataset, PETS
  - TUB-Dataset Sparse Crowd
  - TUB-Dataset Dense Crowd
  - PETS 2010: S1 L1 Time 13-57
- Outperforms standard MHT

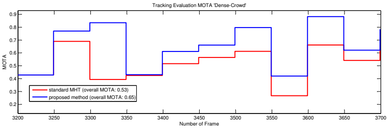


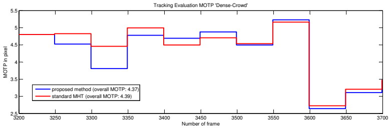




Tracking performance evaluation:

- Evaluation by CLEAR MOT metrics:
  - Object Configuration Error: MOTA
  - Average total Position Error: MOTP
  - Computation of measures for batches of 50 frames
- Dataset: most challenging part of the 'dense crowd'-sequence of TU-Dataset (passing 29 people)





Conclusions

- MHT-based tracking system with initialization of objects in crowded environment
- Relieable method for choosing objects for appearance model learning
- Efficient evaluation of trained models for different hypotheses
- Integration of instance-specific information into MHT-framework
- Guiding of tracks without detections by appearance model information

Literature

Y. Bar-Shalom, F. Daum, and J. Huang.
The probabilistic data association filter.
Control Systems Magazine, IEEE, 29(6):82 –100, 2009.

M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. Van Gool.
Robust tracking-by-detection using a detector confidence particle filter.
In *Proc. IEEE 12th Int Computer Vision Conf*, pages 1515–1522, 2009.

Cheng-Hao Kuo, Chang Huang, and R. Nevatia.
Multi-target tracking by on-line learned discriminative appearance models.
In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 685 –692, june 2010.