Robust Multiple Object Tracking by Detection with Interacting Markov Chain Monte Carlo

S.Santhoshkumar, S.Karthikeyan, B.S.Manjunath

Department of Electrical and Computer Engineering, University of California, Santa Barbara.
Robust Tracking by Detection

• Most of the existing tracking algorithms fail when applied on a new dataset.

• Solution: Scene priors could guide tracking and improve robustness.

• Two sources of scene priors
  • Local scene priors
  • Global scene priors
Scene Priors

• Local Information
  – Object specific information such as motion model, appearance.

• Global Scene Information
  – Interaction between multiple objects.
  – Entry/exit locations in the scene.
  – Trajectories of objects over the time.
Why Local and Global models?

• Single model is insufficient to capture all the motions characteristics of the object.

• Objects exhibit different motion patterns in different parts of the scene.

• Local object cues are insufficient during abrupt appearance and motion changes.

• Global model guides tracking by taking the entire scene into account.
Contributions

- Track multiple objects in a computationally efficient manner (parallelizable).
- Efficient algorithm to combine local and global information.
- Probabilistic approach that generalizes to different datasets.
Traditional Particle Filters

- Let $X_t = [x_t, y_t, s_t]$ be the object state and $Y_t$ be the measurement. Object tracking problem is formulated as a Bayesian MAP estimation:

$$\hat{X}_t = \arg \max_{X_t} p(X_t | Y_{1:t})$$

- In SIR particle filters, the posterior distribution is approximated by a set of “N” weighted particles:

$$p(X_t | Y_t) \approx p(Y_t | X_t) \sum_{r=1}^{N} \pi_{t-1}^{(r)} p(X_t | X_{t-1}^{(r)})$$
Limitations of Traditional Particle Filters

• Particle weight degeneracy.
  – Resampling helps only to an extent but it is a heuristic.
• Complexity varies exponentially with number of objects.
• Not easy to mix multiple samplers.
MCMC Sampling Algorithm

- Posterior is represented by a set of un-weighted particles.
- Complexity increases linearly with number of objects[1].
- Provides an efficient way to combine multiple samplers (local and global in our case).

Proposed Tracking by Detection (Notations)

Detections and Tracks

- Let $D_t=\{d_{jt}\}$ be the set of detections at time "t".
- $\{tr_{it}\}$ be the set of existing particle filter trackers.

Affinity Matrix

- Affinity for matching a detection $d_{jt}$ to the existing set of trackers $\{tr_{it}\}$ is given by:

$$S(i, j) = p_{size}(d_{jt}|tr_{it}) \cdot p_{pos}(d_{jt}|tr_{it}) \cdot p_{appearance}(d_{jt}|tr_{it})$$

- Use Hungarian algorithm to associate detection to trackers.
Local Particle Filters

Observation Model

\[ p_l(Y_t | X_t) = \underbrace{\mu(d, X_t)}_{\text{Detection Score}} \underbrace{p_{pos}(d | X_t)}_{\text{Position}} \]

\[ \mu(d, X_t) = \frac{\text{area}(B_d \cap B_x)}{\text{area}(B_d \cup B_x)} \]

Assumed to be Normal.

Metropolis Hastings Algorithm

Consists of two steps:

i. Proposal Step: A new state is proposed with proposal density:

\[ Q_l(X^*_t ; X_t) = p_l(X^*_t | X_t) \]

i. Acceptance Step: Proposed state is accepted with an acceptance ratio:

\[ \alpha_{\text{parallel}} = \min \left[ 1, \frac{p_l(Y_t | X^*_t)Q_l(X_t ; X^*_t)}{p_l(Y_t | X_t)Q_l(X^*_t ; X_t)} \right] \]
Global Particle Filters

Observation Model

\[ p_g(Y_t | X_{it}) = \Omega(d_i, X_{it}) \cdot \prod_{k \neq i} \psi(X_{it}, d_k) \cdot \phi(d_i, X_{it}) \]

\[ \psi(X_{it}, d_k) = 1 - \frac{1}{\beta} \exp\left(-\frac{\text{dist}(X_{it}, d_k)}{\sigma_i^2}\right) \]

- Similar to local particle filters, uses parallel Metropolis Hastings sampling algorithm for posterior estimation.
Interacting MCMC Particle Filters

For each object a set of two particle filters are used:

a. **Local Particle Filter:**
   i. Models the local motion of the object.
   ii. Does not take scene information into account.

b. **Global Particle Filter:**
   i. Models probable motions within the scene.
   ii. Takes *scene information* into account.
• In parallel mode, global particle filters act as parallel metropolis hastings sampler.

• At each time step "t", local particle filter operate in either parallel or interactive mode.

• In interactive mode, local particle filter communicates with the global particle filter and seeks the better state for object configuration $X_t$.

• Local particle filter accepts the state of global particle filter has its own with an interaction probability:

$$\alpha_{interacting} = \frac{p_g(Y_t|X_{it})}{p_l(Y_t|X_{it}) + p_g(Y_t|X_{it})}$$
Occlusion and Missing Detections

- Objects might not be detected due to lighting changes, illumination effects, occlusion and missing features.

- For a given tracker, if the object is neither detected nor associated:
  - The global particle filter proceeds with the prediction step and performs update using the other available detections and domain specific information.
  - The local particle filter operates completely in interactive mode.
  - Tracker pauses its operation after a certain number of continuously missed detection between frames (20 frames in our experiments).
Domain Priors

PETS-2009

- Kernel density estimator to model the probability density function on trajectories of the objects that moved over the scene for a period of time.

Melanosomes

- Velocity distribution of melanosomes modeled using Brownian motion model [2].

PETS-2009 Dataset

- S2-L2 sequence with 7 objects in the scene.
- We used HOG based pedestrian detector (Dalal et al. [3]).
- **Appearance likelihood:** We used normalized HSV based multi-dimensional color histogram with Bhattacharyya distance metric.
- Training data obtained from a different set of videos.

Melanosomes Dataset

• Melanosomes are organelles present in the Retinal Pigment Epithelium (RPE) layer of the retina.

• Melanosomes carry dark pigment and imaged using Bright-field microscopy (from mouse retina).

• We used thresholding followed by connected component analysis to generate detections.

• **Appearance likelihood**: We used histogram of Local Binary Patterns (LBP with radius 4 pixels) with Bhattacharyya distance metric.
## Experiments - PETS2009

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>304</td>
<td>278</td>
<td>17</td>
</tr>
</tbody>
</table>
Experiments - PETS2009


Experiments - Melanosomes

Conclusion And Future Work

• An efficient multiple object tracking by detection algorithm with IMCMC framework (300 milliseconds to track 50 objects per frame on a 2.4 GHz machine)

• Applicable for different domains (Pedestrians, Melanosomes)

• Extend pedestrian tracking to multiple cameras
Thank you!