Analysis of object tracking algorithms performance on event-based datasets

Alexandra-Claudia Olaru

Nergis Tomen, Ombretta Strafforello, Xin Liu

Delft University of Technology

Introduction

In event-based camera, each pixel adjusts and reacts to temporal brightness changes. The output is an asynchronous event [1], which is described by (x, y, t, p), with the location (x, y), timestamp t and polarity p (shows whether the intensity increased or decreased).

Object tracking = approximation of the trajectory of an object in the image plane over time [2]

An object is represented by multiple points, which are the events mapped into a 2D plane, with location (x, y) and timestamp t as value. The center of mass of each object is tracked in time.

Object tracking divided into 2 main steps:

- Object localization - cluster events and the resulted clusters represent the moving object
- Trajectory estimation - correct the detected location, using a non-linear filter, to consider possible errors

Research question

How well can single and multiple object tracking perform using event-based data?

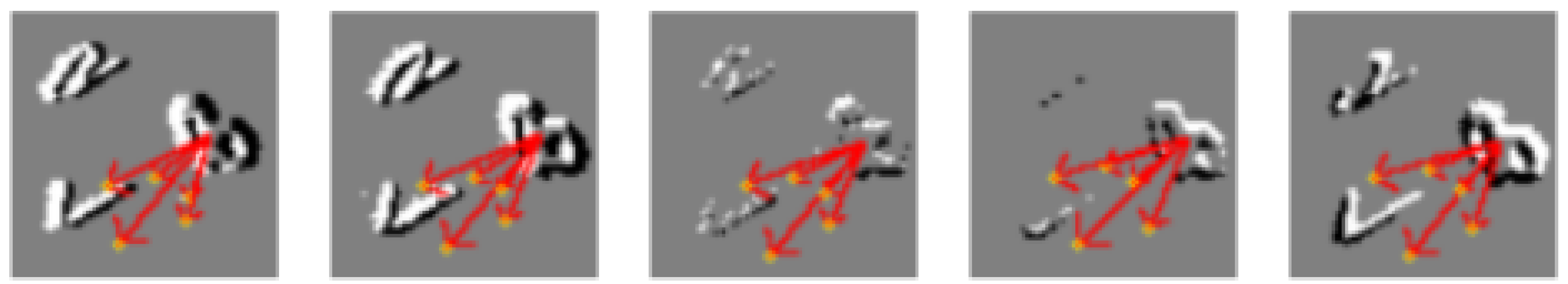


Figure 1:Sample particle movement for one of the objects in multi target tracking. In red the trajectory from the center of mass of the object to the particles. In yellow the position of the particles.

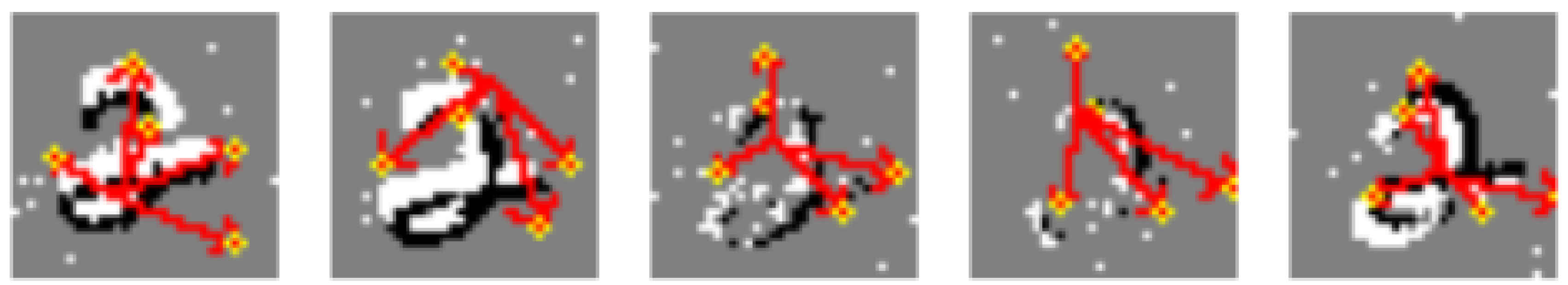


Figure 2:Sample particle movement for one object in the image plane. In red the trajectory from the center of mass of the object to the particles. In yellow the position of the particles.

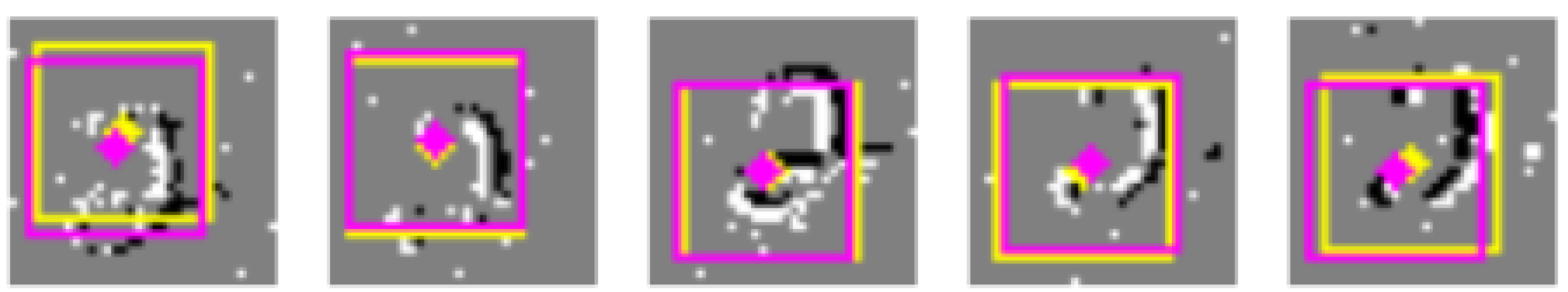


Figure 3:Visualization of object tracking over time. In pink current position and centroid. In yellow, previous position and centroid estimated using Particle filter.

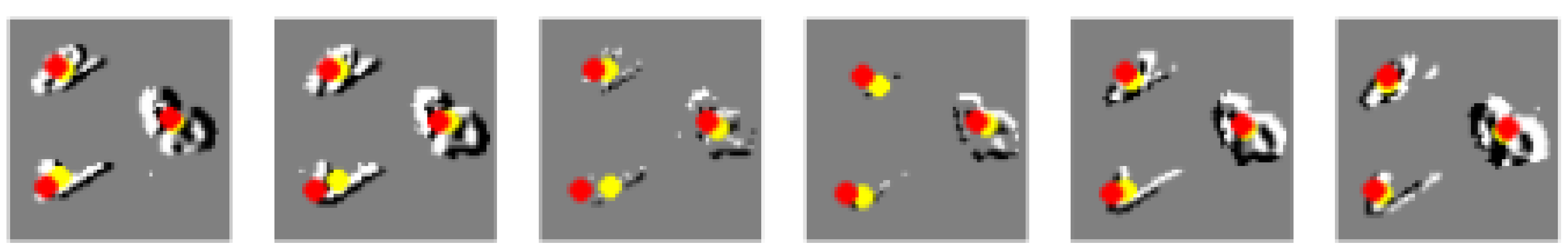


Figure 4:Difference between the centroid computed by the cluster algorithm (yellow dots) and the centroid estimated by the particle filter model (red dots).

Methods

Single object detection:

1. Detect object using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm
   - Points are grouped based on the distance measurement and a minimum number of points
2. Compute the centroids $(x_{centroid}, y_{centroid})$ of clusters using the formula:
   x

$$centroid = \frac{x_{min} + x_{max}}{2}, \qquad y_{centroid} = \frac{y_{min} + y_{max}}{2}$$

   where (x_{min}, y_{min}), (x_{max}, y_{max}) represents the position of the events with the minimum, respective the maximum value of coordinates in a cluster.
3. Apply an estimator to approximate the trajectory over time

Multi objects detection:

1. Detect object using Mean Shift Clustering algorithm [3]
   - Candidate of a centroid is the mean of the points from a specified area
   - The update for a new sample is done by obtaining the nearest centroid for a given sample
2. Compute the centroids $(x_{centroid}, y_{centroid})$ of clusters using the mean shift vector 1

$$m(x_i) = \frac{\sum_{x_j \in A(x_i)} Kernel(x_j - x_i) * x_j}{\sum_{x_j \in A(x_i)} Kernel(x_j - x_i)} \tag{1}$$

   where $A(x_i)$ is the neighborhood of samples in an area of a given radius from x_i.
3. Maintain order of clusters over time
4. Apply an estimator to approximate the trajectory over time

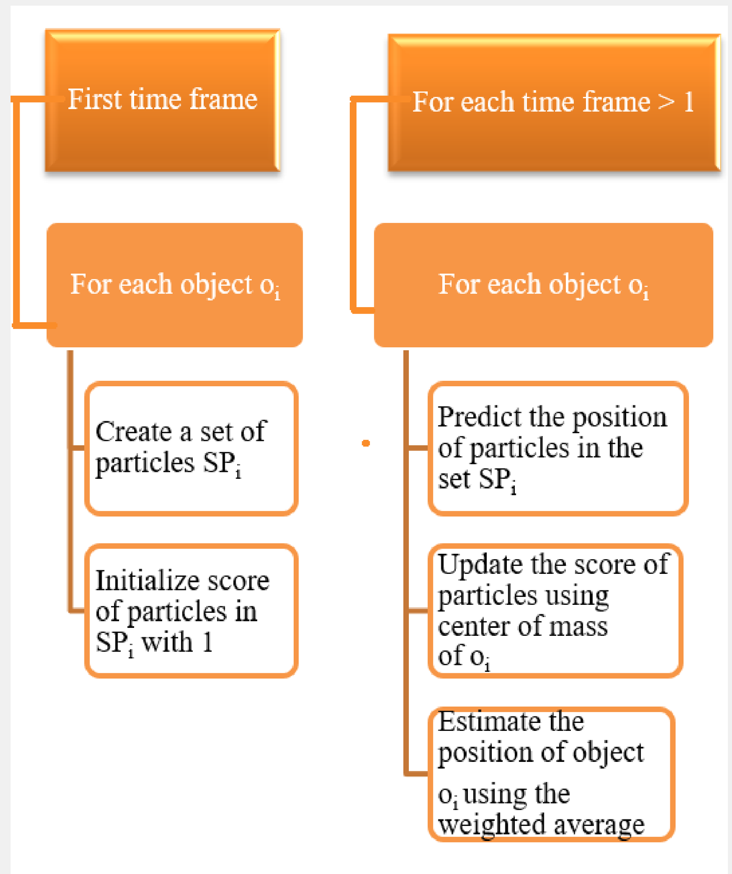


Figure 5:Flowchart of the tracking model

Particle filter [4]

1. Randomly draw N particles p_i^k from a uniform distribution, with equal score s_i^k
2. At each step k:
   - Predict the position of p_i^k using a predefined motion model

$$p_i^{k+1} = p_i^k + M(p_i) + \alpha \tag{2}$$

   - Update the score s_i^k, computing the probability $pr(s_i^{k1}|y_k)$, where y_k is the observation
   - Resample if necessary
   - Estimate the current position using the weighted average of particles coordinates

A particle is:

- Hypotheses over the current state of the system
- Characterized by location with coordinate (x, y) and a rotation value
- Linked to a score, s_i, which shows how well the particle represent the observation

Contact: a.c.olaru@student.tudelft.nl

Results

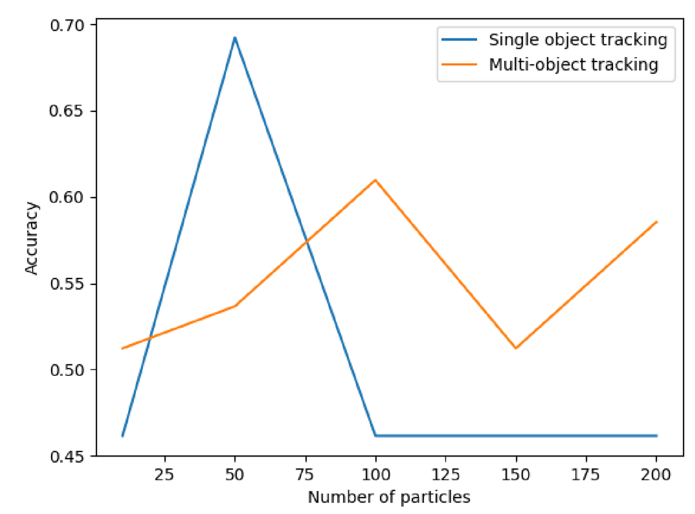
Clustering metrics

Clustering performance metric	Clustering Algorithm	Clustering average score
Silhouette score	Mean Shift	0.739
	DBSCAN	0.659
Calinski & Harabasz Index	Mean Shift	3263.1
	DBSCAN	1731.5
Davies-Bouldin Index	Mean Shift	0.343
	DBSCAN	0.455

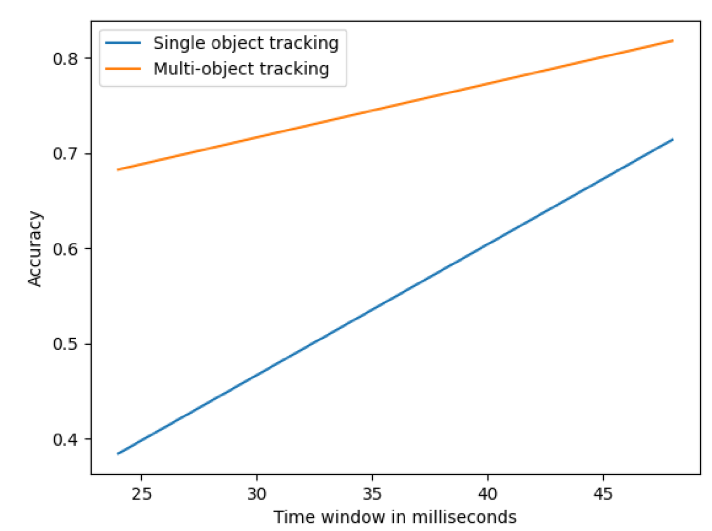
Figure 6:Clustering performance evaluation for multiple objects in the image plan.

Tracking accuracy

Variation of accuracy relative to the number of particles



Variation of accuracy for different time frame length



Conclusions

- Mean Shift provides better results for multi target tracking
- A non-linear estimator performs well on event-based data because it is a non-linear system, without a Gaussian distribution.

	Single object tracking	Multi-target tracking
Maximum accuracy for time frame of length 24ms	0.7	0.6
Particles number linked to the largest accuracy	50	100
Time frame length linked to best results	48ms	48ms

Limitations

- N-MNIST dataset [5] does not completely reflect the reality
- Results influenced by the features of the sensor
- The motion model can be improved using a more precise velocity.

References

[1] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis, *et al.*, ``Event-based vision: A survey,'' *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 1, pp. 154--180, 2020.

[2] D. Birant and A. Kut, ``St-dbscan: An algorithm for clustering spatial--temporal data,'' *Data & knowledge engineering*, vol. 60, no. 1, pp. 208--221, 2007.

[3] F. Barranco, C. Fermuller, and E. Ros, ``Real-time clustering and multi-target tracking using event-based sensors,'' in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5764--5769, IEEE, 2018.

[4] D. Weikersdorfer and J. Conradt, ``Event-based particle filtering for robot self-localization,'' in *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 866--870, 2012.

[5] G. Orchard, A. Jayawant, G. K. Cohen, and N. Thakor, ``Converting static image datasets to spiking neuromorphic datasets using saccades,'' *Frontiers in Neuroscience*, vol. 9, 2015.