

A literature review on recent multi-object tracking methods based on HMM and particle filter

Kalyani Ahire¹, Prof.P.S Mohod²

Department of Computer Science & Engineering,, G.H.R.I.E.T.W.,RashtrasantukdojiMaharaj Nagpur University Nagpur, India

Abstract: Tracking of the object movement in the input frame of video is an important process for various real time application such as video-conferencing, human robotics or human computer interface or in the analysis of social interaction. The important step is to determine the path of the object. The various techniques to track the multiple moving objects in a input frame of video have been proposed. This paper gives a brief analysis of recent multi-object trackers algorithms based on markov model and particle filtering

Keywords: face tracking,HMM(Hidden markov model)MCMC(Markov Chain Monte Carlo), PF(Particle filtering), long term multi-object tracking.

I. Introduction

Object tracking has many practical applications (e.g. surveillance, HCI) and has long been studied in computervision. Generally there are three components to a tracking system: image representation (e.g. filter banks, subspaces, etc.), appearance model, and motion model.

The advanced multi-object tracking methods detect the object movement, orientation of object and the path of the objects. The probabilistic characteristics of the object tracking algorithms enable us to detect the unstable and unpredictable trajectory of the object in the video. The basic steps in the tracking of objects in video are: frames, adaptive background removal, and skinColor model, global appearance based model, candidate based feature extraction etc. In this paper, literature review of multi-object tracking methods are explained broadly.

II. Literature Review

A. Parameter in object tracking

There are various techniques for object tracking

[7], [8], [12 - 14], [23], [25], [27], [28]. When the ground truth, i.e. the objects to be tracked is manually segmented, one can compare and spot the correctly labeled object pixels (true positives – TP) and non-object pixels which are falsely labeled as object pixels (false positives – FP). False positive is also known as false alarm. *False reduction* is the decrease of the detection of false objects, which are represented by TP and FP. The ratio of false-positive detections that cannot be adapted to any ground-truth trajectories over the number of detections is known as the *false positive rate*.

B. Particle-Filter Based Multi-Object Tracker

The objects in multiple-object tracking methods are bigger when compared to the conventional point tracking methods. These are known as “extended” objects. One needs to create point measurements relative to extended object detections and apply one of the existing point-target tracking algorithms, for multiple-object tracking. Point-target tracking methods, JPDA (Joint Probabilistic Data Association), multi-dimensional assignment and the PHD (Probability Hypothesis Density) filter can be used for multi-object tracking. For tracking colored objects, zian et al proposed a particle filter for single colored object and manual initialization. Multi-object tracker based on color and PF with automatic object initialization/deletion was proposed.

C. Multi-Track Linking

Tracking a multiple faces in a video is an important process. Long-term occlusions are mainly responsible for problems during the data association process, resulting in “track-switch” or “track-lost” errors.

When occlusions occur, individual or multiple tracks become merged. After the occlusions, the merged tracks separate into individual tracks. To maintain the integrity of merging and splitting process, the “track linking method” assumes each track as a “tracklet” and links these tracklets. Earlier approaches used a local linking strategy that calculated the pairwise cost between tracklets in a repetitive manner. Track graph was introduced by Nillius et al and a Bayesian network interference algorithm as a global linking strategy. The global linking allows simultaneous matching of various tracklets, but it is computationally expensive.

Another division of data association techniques are sampling-based algorithms. Oh *et al* [15] introduced a framework to particulate the data association hypothesis and track a large number of objects, by a Markov Chain Monte Carlo (MCMC) approach. Khan *et al* [34] proposed a probabilistic model to affiliate merged and split components using a MCMC-based particle filter. Yu and Medioni [20] extended the work of Songhwai *et al* [16] to detect the appropriate temporal and spatial affiliation of segments with a Data-Driven MCMC sampling approach.

D. Long-Term Online Multiface Tracking Using Particle Filter and Hidden Markov Model (HMM)

Most of the multi-face detectors in the recent years are applicable only when the persons look towards the cameras, but this is not possible for all scenarios. The difficult head postures when they last for long time, it is tough to track the trajectories of the objects.

Many multiple face tracking methods have been proposed ([19, 21, 32, 28, 17]), which mainly concentrates on new features, better dynamics, multi-cue fusion mechanisms or adaptive models [7, 8, 9, 10]. The results are always based on short video sequences only.

The various multi-object trackers address the track termination and track initialization, especially in terms of performance evaluation. A high confidence threshold in the face detector may result in missing an early track initialization. The false tracks also occur because of low threshold false tracks.

Principled methods exist to integrate track creation and termination within the tracking framework, for example Reversible-Jump Markov Chain Monte Carlo (RJ-MCMC) [11], [12]. But to be effective, they require appropriate global scene likelihood models involving a fixed number of observations and these are difficult to build in multi-face tracking applications. Kalal *et al* [13] present an interesting approach for failure detection in visual object tracking that is based on the idea that a correctly tracked target can be tracked *backwards* in time. Unfortunately, the backward tracking greatly increases the overall computational complexity (by a factor linear in the backward depth). In a particle filter tracking framework, another solution is to directly model a failure state as a random variable within the probabilistic model [14]. However, this increases the complexity of the model and thus the inference, and it is difficult in practise to model the distribution of a failure state or failure parameters.

Dockstader *et al* [15] proposed to detect failure states in articulated human body tracking using a Hidden Markov Model (HMM). However, their method differs significantly from ours: they only use one type of observation (the state covariance estimate) which in our case proves to be insufficient for assessing tracking failure; their observation are quantised to use a standard discrete multinomial likelihood model, whereas our method learns these likelihoods in a discriminative fashion; and their HMM structure (number of states, connections) is specifically designed for their articulated body tracking application. In applications that are similar to ours the problem of deciding when to stop tracking a face is usually solved in a recursive manner. This means, assessing tracking failure is often left to the (sudden) drop of objective or likelihood measures which are not easy to control in practise.

In many scenarios of interest, the camera is fixed, and due to the application and the room configuration, people in front of the camera tend to behave similarly over long periods of time. However, most of the existing face tracking methods ignore this long-term information, as they concentrate on video clips that are often not longer than a minute.

Mikami *et al* [16] introduced the Memory-based Particle Filter where a history of past states (and appearances [17]) is maintained and used to sample new particles. However, they only addressed single, near-frontal face tracking, in high resolution videos and only evaluated the method on 30 to 60-second video clips. Finally, other works (e.g. [22]–[24]) tackle the problem of long-term *person* tracking by analysing the statistics of features from shorter tracks (tracklets), and by proposing methods to effectively associate them. The observations at every point in time are known while observation, and they mainly deal with tracking the position of the *full human body* as opposed to just faces. Another approach for multiple pedestrian tracking [25] associates smaller tracklets *on-line* and in a statistical sampling framework but no principled mechanism for starting and ending tracks is proposed.

VI. Conclusion

The multi-object tracking methods based on particle filter, Hidden Markov Model (HMM) and the appearance information of objects have been reviewed in this paper. The recent multi-object tracking methods show higher efficiency when compared to its predecessors, like computational efficiency calculated by observations.

References:

- [1] S. Duffner and J. M. Odobez, "Track Creation and Deletion Framework for Long-Term Online Multiface Tracking," *Image Processing*, IEEE Transactions on, vol. 22, pp. 272-285, 2013.
- [2] C. Cuevas and N. Garcia, "Efficient Moving Object Detection for Lightweight Applications on Smart Cameras," *Circuits and Systems for Video Technology*, IEEE Transactions on, vol. 23, pp. 1-14, 2013.
- [3] I. Ali and M. N. Dailey, "Multiple human tracking in high-density crowds," *Image and Vision Computing*, vol. 30, pp. 966-977, 2012.
- [4] C.-H. Kuo and R. Nevatia, "How does Person Identity Recognition Help Multi-Person Tracking?," 2011.
- [5] B. Benfold and I. Reid, "Stable multi-target tracking in real-time surveillance video," in *Proc. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 3457-3464.
- [6] P. Pan and D. Schonfeld, "Video Tracking Based on Sequential Particle Filtering on Graphs," *IEEE Transactions On Image Processing*, vol. 20, pp. 1641-1651, 2011.
- [7] S. Duffner and J. M. Odobez, "Exploiting long-term observations for track creation and deletion in online multi-face tracking," in *Automatic Face & Gesture Recognition and Workshops (FG 2011)*, 2011 IEEE International Conference on, 2011, pp. 525-530.
- [8] L. Cehovin, et al., "An adaptive coupled-layer visual model for robust visual tracking," in *IEEE International Conference on Computer Vision*, 2011, pp. 1363-1370.
- [9] A. Saffari, et al., "Online multi-class LPBoost," in *Computer Vision and Pattern Recognition (CVPR)*, 2010 IEEE Conference on, 2010, pp. 3570-3577.
- [10] B. Song, et al., "A stochastic graph evolution framework for robust multi-target tracking," presented at the Proceedings of the 11th European conference on Computer vision: Part I, Heraklion, Crete, Greece, 2010.
- [11] D. Mikami, et al., "Memory-based particle filter for tracking objects with large variation in pose and appearance," presented at the Proceedings of the 11th European conference on computer vision conference on Computer vision: Part III, Heraklion, Crete, Greece, 2010.
- [12] Z. Kalal, et al., "Forward-Backward Error: Automatic Detection of Tracking Failures," in *Pattern Recognition (ICPR)*, 2010 20th International Conference on, 2010, pp. 2756-2759.
- [13] W. Zheng, et al., "Tracking a large number of objects from multiple views," in *Computer Vision*, 2009 IEEE 12th International Conference on, 2009, pp. 1546-1553.
- [14] Q. Yu and G. Medioni, "Multiple-Target Tracking by Spatiotemporal Monte Carlo Markov Chain Data Association," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, pp. 2196-2210, 2009.
- [15] Y. Ming, et al., "Context-Aware Visual Tracking," *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, vol. 31, pp. 1195-1209, 2009.
- [16] D. Mikami, et al., "Memory-based Particle Filter for face pose tracking robust under complex dynamics," in *Computer Vision and Pattern Recognition*, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 999-1006.
- [17] Z. Kalal, et al., "Online learning of robust object detectors during unstable tracking," in *Proc. Int. Conf. Comput. Vis.*, 2009, pp. 1417-1424.
- [18] X. Junliang, et al., "Multi-object tracking through occlusions by local tracklets filtering and global tracklets association with detection responses," in *Computer Vision and Pattern Recognition*, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 1200-1207.
- [19] B. Babenko, et al., "Visual tracking with online Multiple Instance Learning," in *Computer Vision and Pattern Recognition*, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 983-990.
- [20] S. M. Bhandarkar and X. Luo, "Integrated detection and tracking of multiple faces using particle filtering and optical flow-based elastic matching," *Computer Vision and Image Understanding*, vol. 113, pp. 708-725, 2009.
- [21] J. Yao and J.-M. Odobez, "Multi-camera multi-person 3D space tracking with MCMC in surveillance scenarios," presented at the Proc. Eur. Conf. Comput. Vis., Workshop Multicamera Multimodal Sensor Fusion Algorithms Appl., Marseille, France, 2008.
- [22] E. Maggio, et al., "Efficient Multitarget Visual Tracking Using Random Finite Sets," *Circuits and Systems for Video Technology*, IEEE Transactions on, vol. 18, pp. 1016-1027, 2008.
- [23] Z. Li, et al., "Global data association for multi-object tracking using network flows," in *Computer Vision and Pattern Recognition*, 2008. CVPR 2008. IEEE Conference on, 2008, pp. 1-8.
- [24] C. Huang, et al., "Robust Object Tracking by Hierarchical Association of Detection Responses," presented at the Proceedings of the 10th European Conference on Computer Vision: Part II, Marseille, France, 2008.
- [25] C. Plagemann, et al., "Efficient failure detection on mobile robots using particle filters with Gaussian process proposals," presented at the Proceedings of the 20th international joint conference on Artificial intelligence, Hyderabad, India, 2007.
- [26] J. Hao, et al., "A Linear Programming Approach for Multiple Object Tracking," in *Computer Vision and Pattern Recognition*, 2007. CVPR '07. IEEE Conference on, 2007, pp. 1-8.
- [27] A. G. A. Perera, et al., "Multi-Object Tracking Through Simultaneous Long Occlusions and Split-Merge Conditions," in *Computer Vision and Pattern Recognition*, 2006 IEEE Computer Society Conference on, 2006, pp. 666-673.
- [28] P. Nillius, et al., "Multi-Target Tracking - Linking Identities using Bayesian Network Inference," presented at the Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2, 2006.
- [28] K. Zia, et al., "MCMC-based particle filtering for tracking a variable number of interacting targets," *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, vol. 27, pp. 1805-1819, 2005.
- [29] W. Ying, et al., "Tracking appearances with occlusions," in *Computer Vision and Pattern Recognition*, 2003. Proceedings. 2003 IEEE Computer Society Conference on, 2003, pp. 1-789-I-795 vol.1.
- [30] P. Perez, et al., "Color-Based Probabilistic Tracking," presented at the Proceedings of the 7th European Conference on Computer Vision-Part I, 2002.
- [31] C. Rasmussen and G. D. Hager, "Probabilistic data association methods for tracking complex visual objects," *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, vol. 23, pp. 560-576, 2001.
- [32] I. Haritaoglu, et al., "W4: real-time surveillance of people and their activities," *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, vol. 22, pp. 809-830, 2000.