A COMPARATIVE STUDY ONMULTI PERSON TRACKING AND DETECTION

[1]S.SATHYANARAYANAN, [2]Dr. ANNASARO VIJENDRAN,
[1]Research Scholar,2Directors
[1]Computer science department, SNR SonsCollege, Coimbatore.
[2]Computer applications department, SNR SonsCollege, Coimbatore.

Abstract: Nowadays intelligent environment are most often mentioned as a purpose for doing work on visual person-tracking, quite making an intelligent atmosphere exposes many actual world problems in visual tracking that must be solved to make the technology useful. In the context of tracking in intelligent environments, previous researchers offered a person-tracking system that solves most of the real-world issues. With advances in technology, surveillance systems can become more automatic than manual systems whereperson tracking is although critical, is highly challenging. Tracking and recognizing objects and person movement from surveillance video motion, followed with the automatic summarization of its content has become a hot topic of study. Many researchers have contributed to the field of automatic video surveillance via detection and tracking algorithms. Previous research work is inadequate for comprehensive analysis of person tracking. The context of a surveillance domain could also be recognizedwith introduction of semantics. Such semantics may extend surveillance methods to participate in person tracking analysis principal to the area. This paper grants a survey on person tracking analysis is done with the aim of analyzing the capabilities of the state-of-art methodologies with specified focus on semantically enhanced evaluation.

Keywords: person tracking, video surveillance, Activity recognition, objecttracking.

1. INTRODUCTION

Multi-persontracking could be a vital topicin video surveillance context. Its determination will profit several applications. For instance, knowing the situation of various individuals over time will greatly facilitate the linguistics analysis of video, like group/interaction detection [1, 2], scene understanding [3] then on. On the opposite hand, the output of a multi-person huntsman will be fed to some higher level method like behavior cue extraction for action/event recognition [4]. However, multi-human tracking remains a difficult task, particularly in single camera tracking things, or in multi-camera cases with tiny overlap or high state of affairs, notably because of low image quality, device noise, and dimension loss because of projection of 3D objects in image planes, occlusions, clutter, unpredictable motions and look changes of individuals.

As task-specific object detectors become a lot of and a lot of reliable, one approach for multi-person tracking is to trust alone on the output of human detectors, that is named tracking-by-detection or detection-based tracking. Through this paradigm, person detection is performed first on the pictures. Then, the tracking step tries at associating the detections are like an equivalent person by assignment labels to the detection outputs. The most benefit is that discriminatively trained detectors are typically a lot of powerful at assessing the presence of humans in a picture compared to plain generative models. Another advantage is that no manual (re-)initialization is required since it's implicitly handled through the employment of the detector output at each frame. However, to achieve success, it's necessary to affect human detector inherent flaws: lost detections and false alarms, however additionally unprecise localization and size because of the presence of projected shadows or partial occlusion for example. Another lot of general challenge lies within the proven fact that individuals typically have similar appearances.

In apply, sadly, person tracking is very troublesome. Samples of the info we have a tendency to want to tackle are displayed in Fig. 1. The primary challenge evident in these pictures is that people's appearances vary wide, and folks amendment their look in several environments, that complicates person detection. Despite wonderful advances in detection [5, 6], it's still removed from trivial to observe individuals in a very style of poses, sporting a spread of article of clothing, and in littered environments packed with occlusions. A good field-of-view that truncates individuals,

ISSN 2229-5518

additionally as high distinction illumination, are further difficulties typically encountered in indoor environments.

Another challenge is that the complexness of the motion patterns of multiple individuals within the same scene. Tracking one person is sufficiently troublesome as they move willfully and erratically. Tracking multiple individuals, however, is difficult by their interactions; assumptive independence between targets' motions is meager. As in Fig. 1, individuals keep out of every other's personal house and ne'er occupy precisely the same house. On the opposite hand, individuals might opt to move along as a gaggle for a while. To model these interactions, we have a tendency to propose putting constraints between the targets' motions, partly removing the independence assumption. The rest of the paper is prepared as follows: the part 2 discusses related works on person tracking. part 3, describes the problems and the solutions. Part 4 concludes the work.

Fig 1: Typical examples of outdoor and indoor tracking scenarios for tracking multiple people in a wide variety of difficult situations

2. LITERATURE SURVEY

In this section, numerous strategies reviewed supported multi-person detection and tracking. To the contrary of generative strategies [7], detection-based trackers use a discriminative classifier to assess the presence of an object in an exceedingly scene, that is mostly a lot of sturdy, as progressive detectors provide excellent performance at detective work humans [8, 9]. The detector's output is employed to get target hypotheses in every frame, that then have to be compelled to be transitively joined to make trajectories with consistent identity labels.

Tracking-by detection will so be developed as a knowledge association downside, which typically depends on affinity models between detections in sequential frames supported motion constraints and intrinsic object descriptors like color [10]. The association downside is addressed by some approaches on a multi-frame basis [11-13]. Dependencies are typically sculptured victimisation graphs, and therefore the optimisation downside then consists find the simplest ways between all the detections in separate frames. The method are often applied on probably giant time windows, thus on overcome the scantness within the detection sets induced by uncomprehensible detections and conjointly to touch upon false alarms, however the quality of the optimisation will increase quickly. Moreover, thanks to the temporal vicinity of association thought of during this context, tracking-by-detection techniques will perform poorly in presence of semipermanent occlusions, i.e. severalsequent uncomprehensible detections. Or else, to scale back the computation and to increasingly increase the temporal vary for correspondences, stratified approaches are often thought of, within which low-level tracklets are initial produced therefore integrated at a higherlevel.

In [14], the lower level associate's pairs of detections in adjacent frames supported their similarity in position, size and look. The ensuing tracklets are then fed into a most A Posteriori (MAP) association downside resolved by the Hungarian algorithmic rule, and more refined at a better level to model scene exits and occluders. Because there are fewer tracklets than detections, the quality of the optimisation is reduced, however any wrong association created at the

Most visual tracking strategies specialize in tracking single object or multiple objects individually [6, 15]; they sometimes try and notice correct look models that distinguish one object with all alternative targets or backgrounds, and adopt meanshift [16] or particle filtering [17] like approach to on-line change target look models, and use updated models to endlessly track targets. On the opposite hand, most association primarily based strategies specialize in tracking multiple objects of a pre-known category at the same time [18, 19]. They sometimes associate detection responses created by a pre-trained detector into long tracks, and notice a world optimum resolution for all targets. Look models are typically pre-defined [20] or on-line learned to tell apart multiple targets globally [21]; additionally, linear motion models between tracklet pairs [22, 23] are typically adopted to constrain motion smoothness.

Although such approaches might acquire world optimized look and motion models, they're not essentially ready to differentiate tough pairs of targets, that means close ones with similar appearances, as look models for characteristic a selected try of targets could also be quite totally different with those used for characteristic all targets, and former motion models aren't stable for non-static cameras. However, our online CRF models think about each world and pairwise discriminative look and motion models. Note that CRF models also are adopted in [24]. Each [24] and this approach relax the idea that associations between tracklet pairs are freelance of every alternative. However, [24] centered on modeling association dependencies, whereas this approach aims at higher distinction between tough pairs of targets and so the meanings of edges in CRF are totally different. Additionally, [24] is an offline approach that integrates multiple cues on pre-labeled ground truth information, however our approach is a web learning technique that finds discriminative models mechanically while not prelabeled information .

A novel approach for multi-person tracking by detection in an exceedingly particle filtering framework is mentioned [25] with that objects are often half-track in occluded atmosphere. A detection-based three-level stratified association approach is introduced to robustly track multiple objects in huddled environments from one camera [26]. A replacement accommodative approach is planned to integrate multi cue in tracking multiple human driven by human detections [27]. Implementation of many object tracking algorithms is completed with totally different preprocessing strategies and their performances are evaluated for various video sequences [28]. Changed Background subtraction technique is planned to search out moving objects in an exceedingly video sequences [29]. A completely unique algorithm is developed for period detection and tracking of multiple moving objects, that consecutive integrate the entropy distinction technique with accommodative threshold and therefore the quick level set technique [30].

Person tracking has been extensively studied, primarily from one camera perspective [31, 32]. Previous work has conjointly restrained tracking persons across multiple cameras and therefore the associated hand-off downside [33-35]. Concerning the utilization of overlapping cameras, one among the most strategies for crucial abstraction positions is to geometrically rework pictures supported a preset ground plane homography [36-38]. During this case, the abstraction organization of the scene is calculable by projected metameric second objects on the bottom plane of the scene victimisation camera activity and same transformation. This ends up in aa pair of.5D-like approach, wherever the second pictures are combined into a second projection of a 3D scene. Trackingwill then be done supported the calculable ground plane positions.

In [39] the tactic of victimisation same projections of foreground blobs for detective work individuals is extended by victimisation multiple projection levels. Not solely are projections created on the bottom plane, however multiple height levels are outlined at that pictures are homogeneously remodeled and compared, so making a 3D stack of second projections giving far more detail on the chance of a person's position.

In [40] random models are wont to estimate a ground plane occupancy map that is employed to trace individuals. Camera activity is required to search out a standard ground plane map all told pictures, however the article segmentations are not remodeled onto this ground plane. As an alternative, the second segmentations of the separate camera pictures are directly wont to estimate the occupancy chance of sure prefixed ground plane locations. Rather than victimisation the bottom plane projection to trace, it's conjointly doable to trace within the second camera pictures and use inter-camera matching info to relate objects in numerous camera pictures [41]. During this case, object positions are determined in every individual camera image, half-track individually and so matched between camera's, supported look model and geometrical options. This sort of technique is particularly helpful for consistent labelling of individuals over multiple cameras and an extended amount of your time.

In [42], person positions are found by matching colours on epipolar lines all told cameras. Foreground pictures are projected onto horizontal planes within the 3D house in [43], detective work objects at ground plane locations wherever multiple foreground regions come across in multiple planes. Similarly, [44] uses pictures containing the quantity of foreground components higher than every pixel to make 3D detections at positions with the best accumulated score.

In [45], people's principal axis are matched across cameras. In [46], a Probabilistic Occupancy Map (POM) is bestowed for person detection. A generative model employing a discretized ground plane and stuck size regions of interest approximates the marginal chance of occupancy by accumulating all proof received from foreground pictures from each camera. Detections in [47] are generated employing a volume carving primarily based 3D scene reconstruction, projected onto the bottom plane. The same as [47, 48] proposes a model within which multiple volume carving primarily based scene configuration hypothesis are evaluated. Rather than determination hypothesis choice in 3D, the graph cut algorithmic rule is employed to label the pixels of every camera image as background or one among the individuals within the scene. In [49], an reiterative model is conferred labeling individual voxels of a volume reconstruction as either a part of an object, background or static occluder.

3. INFERENCE FROM EXISTING SOLUTION

The main drawbacks of person and object tracking methods are given below.

Table 1: research gap for object and person tracking methods

Authors	Method	Advantages	Disadvantages
Yao and Odobez [7]	Reversible Jump Markov Chain Monte Carlo (RJ-MCMC)	It is efficiently update tracksor initialize new tracks. It isfast and powerful human detector.	It is only handle partial occlusion, on the use of longer term constraints on the dynamics.
Pirsiavashet al., [12]	Greedy algorithms	Fast, simple, and scalable. It improves accuracy.	The computational cost is high.
Xing et al., [19]	dual-mode two-way Bayesian inference approach	It is more accurate.It obtains satisfactory tracking results on many typicalreal world sports videos.	Tracking multiple highlydynamic and interactive person detection is difficult.
Afef et al., [28]	Camshift and the Kalman filter	It has better precision, goodreliability and less execution time.	The accuracy of tracking is less and computational cost is high. In dynamic tracking it's not efficient.
Wanhyun Cho et al., [30]	Clausius Entropy theory	It is more reliable and robust and improve the detection and tracking performance	It has high time-consuming and it is not suitable for very large region search.
Santos et al., [44]	Kalman filters and normalized method	it does not require initial people segmentation and it has good robustness	It is not fast and execution time is high.
Fleuret et al., [46]	global optimization	very simple model and obtain very good performance	It is not suitable for large region and it has less reliability, high cost.
Liem et al., [47]	two-step method for the joint person localization and track assignment	very good real-time performance has been achieved	Tracking multiple person detection in dynamic way is difficult.

Solution to beat These Issues: despite the fact that a lot of analysis has worn out the area of person tracking Analysis, problems and challenges still prevail. Recognizing the background scene as indoors and out of doors isn't ample, it are often extended upon by more scene classification throughout pre-processing. Motion detection and tracking methodologies are often considerably increased victimization temporal segmentation and linguistics descriptions. The classifier chosen for activity recognition has impact on the sort of activity foreseen. Most of the algorithms don't with efficiency handle multiple object interaction recognition. There's active analysis ongoing during this space, with the increasing want for ways to search out the interaction between teams of people.

Activity recognition is presently in deep trouble events that happen continuously; it is often extended to distinct event recognition ways. There are often two totally different eventualities with relevancy distinct activities, with two persons distant from each other however concerned in an activity (e.g. waving to every alternative) and therefore the other one being, one person activity constant activity repeatedly with pauses. Linguistics descriptions play a significant role find the sort of activity.

4. CONCLUSION

This paper developed the multi-person tracking task as an association drawback between detections. One in all the areas which require more concentration is shaping an even format for linguistics descriptions of various activities. The methodologies during this survey are primarily classified with relevancy linguistics concerned, therefore dividing the entire method into high level and low level. This survey reviewed some algorithms within the domain of person tracking analysis as well as each the low-level and high-level techniques so as to induce a much better understanding of the state of art techniques. The association was expressed as a labeling method employing a Conditional Random Field framework.

REFERENCES

1. W. Ge, R. Collins, and B. Ruback. Automatically detecting the small group structure of a crowd. In IEEE Workshop on Application of Computer Vision (WACV), 2009.
2. T. Yu, S. Lim, K. Patwardhan, and N. Krahnstoever. Monitoring, recognizing and discovering social networks. In CVPR, 2009.
3. X. Song, X. Shao, H. Zhao, J. Cui, R. Shibasaki, and H. Zha. An online approach: learning-semantic-scene-by tracking and tracking-by-learning-semantic-scene. In CVPR, 2010.
4. C. Chen, A. Heili, and J. Odobez. Combined estimation of location and body pose in surveillance video. In AVSS, 2011.
5. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, 2005.
6. P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. PAMI, 32(9), Sept. 2010.
7. J. Yao and J.-M. Odobez. Multi-camera multi-person 3D space tracking with MCMC in surveillance scenarios. In ECCV workshop on Multi 0Camera and Multi-modal Sensor Fusion Algorithms and Applications (ECCV-M2SFA2), Marseille, Oct. 2008.
8. P. Dollar, S. Belongie, and P. Perona. The fastest pedestrian detector in the west. In BMVC, pages 68.1–68.11, 2010
9. A.Yilmaz, O. Javed, and M. Shah. Object tracking: A survey. ACM Comput. Surv., 38(4):13, 2006.
10. H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Globally-optimal greedy algorithms for tracking a variable number of objects. In CVPR, pages 1201–1208, 2011.
11. Z. Wu, T. H. Kunz, and M. Betke. Efficient track linking methods for track graphs using network-flow and set-cover techniques. In CVPR, pages 1185–1192, 2011
12. A. Andriyenko and K. Schindler. Multi-target tracking by continuous energy minimization. In CVPR, pages 1265–1272, 2011.
13. C. Huang, B. Wu, and R. Nevatia. Robust object tracking by hierarchical association of detection responses. In ECCV, pages 788–801, 2008.
14. Y. Li, H. Ai, T. Yamashita, S. Lao, and M. Kawade. Tracking in low frame rate video: A cascade particle filter with discriminative observers of different lifespans. In CVPR, 2007.
15. S. Wang, H. Lu, F. Yang, and M.-H. Yang. Superpixel tracking. In ICCV, 2011.
16. D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of nonrigid objects using mean shift. In CVPR, 2000.
17. M. Isard and A. Blake. Condensation – i99p density propagation for visual tracking. International Journal of Computer Vision, 29(1):5–28, 1998.
18. B. Song, T.-Y. Jeng, E. Staudt, and A. K. Roy-Chowdhury. A stochastic graph evolution framework for robust multi-target tracking. In ECCV, 2010.
19. J. Xing, H. Ai, L. Liu, and S. Lao. Multiple players tracking in sports video: a dual-mode two-way bayesian inference approach with progressive observation modeling. IEEE Transaction on Image Processing, 20(6):1652–1667, 2011
20. J. Xing, H. Ai, and S. Lao. Multi-object tracking through occlusions by local tracklets filtering and global tracklets association with detection responses. In CVPR, 2009
21. H. B. Shitrit, J. Berclaz, F. Fleuret, and P. Fua. Tracking multiple people under global appearance constraints. In ICCV, 2011
22. S. Stalder, H. Grabner, and L. V. Gool. Cascaded confidence filtering for improved tracking-by-detection. In ECCV, 2010.
23. M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. V. Gool. Online multi-person tracking-by-detection from a single, uncalibrated camera. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(9):1820–1833, 2011.
24. B. Yang, C. Huang, and R. Nevatia. Learning affinities and dependencies for multi-target tracking using a crf model. In CVPR, 2011.
25. Breitenstein .M.D, Reichlin .F, Leibe .B, KollerMeier .E &Gool .L .V. (2009), "Robust Tracking-by-Detection Using a Detector Confidence Particle Filter." IEEE International Conference on Computer Vision ICCV, 1515–1522
26. Huang .C, Wu .B &Nevatia .R.(2008), " Robust Object Tracking by Hierarchical Association of Detection Responses." International Conference on Computer Vision ECCV, 788–801.

27. Yang .M, Lv .F, Xu .W & Gong .Y.(2009), " Detection Driven Adaptive Multi-Cue Integration for Multiple Human Tracking." IEEE International Conference on Computer Vision ICCV,1554–1561.
28. Afef .S, Jammoussi&Ameni .Y.(2012), "Object Tracking System Using Camshift, Meanshift and Kalman Filter." World Academy of Science Engineering and Technology, 6:598-603.
29. Chinchkhede .D.W &Uke.N.J. (2012), "Image Segmentation in Video Sequences Using Modified Background." IEEE International Journal of Computer Science & Information Technology IJCSIT, 4: 93-104
30. Wanhyun Cho, Sunworl Kim &GukdongAhn.(2011), "Detection and Tracking of Multiple Moving Objects in Video Sequence Using Entropy Mask Method and Fast Level Set Method." IEEE International Conference on Computer Vision ICCV, 1-6.
31. I. Haritaoglu, D. Harwood, and L. S. Davis. W4: Real-Time Surveillance of People and Their Activities. PAMI, 22(8):809–830, Augustus 2000.
32. Mittal, A., Davis, L.: M 2 Tracker: a multi-view approach to segmenting and tracking people in a cluttered scene. IJCV 51(3), 189–203 (2003)
33. S. Khan et al. Human tracking in multiple cameras. In ICCV, 2001.
34. B. Prosser et al. Multi-camera matching using bi-directional cumulative brightness transfer functions. In BMVC, 2008.
35. W. Zajdel, Z. Zivkovic, and B. J. A. Kröse. Keeping Track of Humans: Have I Seen This Person Before? ICRA, pages 2081–2086, April 2005.
36. R. Eshel and Y. Moses. Homography based multiple camera detection and tracking of people in a dense crowd. In CVPR, pages 1–8, 2008.
37. W. Hu et al. Principal axis-based correspondence between multiple cameras for people tracking. PAMI, 28(4):663–671, 2006.
38. S. Khan and M. Shah. A multiview approach to tracking people in crowded scenes using a planar homography constraint,. In ECCV, 2006.
39. D. Arsic et al. Applyingmulti layerhomography for multi camera person tracking. In ICDSC, 2008.
40. F. Fleuret et al. Multicamera people tracking with a probabilistic occupancy map. PAMI, 30(2):267–282, 2008.
41. S. Calderara et al. Bayesian-competitive consistent labeling for people surveillance. PAMI, 30(2):354–360, 2008.
42. 16. Mittal, A., Davis, L.: M 2 Tracker: a multi-view approach to segmenting and tracking people in a cluttered scene. IJCV 51(3), 189–203 (2003)
43. Khan, S., Shah, M.: Tracking multiple occluding people by localizing on multiple scene planes. IEEE Trans. on PAMI 31(3), 505–519 (2009)
44. Santos, T.T., Morimoto, C.H.: Multiple camera people detection and tracking using support integration. PRL 32(1), 47–55 (2011)
45. Calderara, S., Cucchiara, R., Prati, A.: Bayesian-Competitive consistent labeling for people surveillance. IEEE Trans. on PAMI 30(2), 354–360 (2008)
46. Fleuret, F., et al.: Multicamera people tracking with a probabilistic occupancy map. IEEE Trans. on PAMI 30(2), 267–282 (2008)
47. Liem, M., Gavrila, D.M.: Multi-person localization and track assignment in overlapping camera views. In: Pattern Recognition, vol. 6835, pp. 173–183 (2011)
48. Huang, C.C., Wang, S.J.: A bayesian hierarchical framework for multitarget labeling and correspondence with ghost suppression over multicamera surveillance system. IEEE Trans. on ASE 9(1), 16 –30 (2012)
49. Guan, L., Franco, J.S., Pollefeys, M.: Multi-view occlusion reasoning forprobabilistic silhouette-based dynamic scene reconstruction. IJCV 90(3), 283–303 (2010)