World Applied Sciences Journal 29 (Data Mining and Soft Computing Techniques): 140-145, 2014
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2014.29.dmsct.26

A Survey on Face Detection and Tracking

[1]R. Jagathishwaran, [2]K.S. Ravichandran and [3]Premaladha Jayaraman

[1]Computer Science and Engineering, School of Computing, India
[2]Associative Dean of ICT, School of Computing, India
[3]Research Scholar, School of Computing, India

Abstract: Face detection and Tracking are important research areas in the field of computer vision and image processing. Face detection is a computer technology that helps to determine the locations and size of the human faces. Face detection techniques are used in cameras for auto focus. Face detection and tracking are the two processes done by using various approaches. It is applied mainly in surveillance. The main purpose of these processes are detect and track the face even in poor viewing conditions in surveillance application. In this paper various techniques used for people detection and tracking like adaptive color based particle filter, fuzzy based particle filter algorithm and so on are discussed. Comparisons between the various approaches are illustrated, Performance measures in terms of number of particles used, root mean square error values etc have been reported. Drawbacks for the techniques like tracker facing the problems while detection and tracking has been explained. Reasons why fuzzy based particle filter is best among all the approaches have been produced.

Key words: Face Detection · Face Tracking · Fuzzy based particle filter · Particle filter · Kalman filter

INTRODUCTION

In unconstrained environment, face detection and tracking needs robust tracking and segmentation are needed to provide the normalized face. Two broad approaches are used namely, motion based approach and model based approach. While joining all the motions at a time robust technique are needed for motion based approach. Model based approach needs more semantic knowledge and computationally requires more cost due to scaling, rotation, translation and deformation. Both the model and motion based approaches are combined in a closed loop, motion based tracker reduces the searching space in model based face detection and later it aids the tracking. This has been described in Face tracking and poses representation [1]. Computer vision application requires the task, object tracking. Color based tracking methods are proposed by using mean shift, Anticipation of following reference location is calculated by a kalman filter, described in [2]. To present a target to track video sequences, a particle filter is present. This filter uses simple linear dynamical model and a likelihood model based on color histograms which describes in adaptive color based particle filter [3]. For resolving the stereo vagueness in face detection and tracking a new fuzzy based algorithm is used. More than one fuzzy system is used to remove the unwanted regions detected by the face detector is also described [4].

Related Works: In most of the tracking approaches kalman mean shift [2] has been used for tracking. It will find the target in the next frame which uses the Bhattacharyya coefficient [5]. Lots of processes have been done in stereo vision for finding the distance. Darrell *et al* [6] system will detect and track more than one person. Whole process has done using skin, face detector and disparity map. Grest and Koch [7] person position has been estimated by using a particle filter and color histogram is created for the face and the real position is computed by using the stereo vision. Moreno *et al* [8] delivered a system which is capable of detecting and tracking single face using kalman shift and mix the color and stereo information. Harville [9] and Munoz-Salinas *et al* [10] delivered a system which

Corresponding Author: R. Jagathishwaran, Computer Science and Engineering, School of Computing, India.

detects and tracks multiple people by using the plan-view maps. Soft computing techniques are also applied in the computer vision as described in Kil-jae and Bien [11] and in Bloch [12]. Dealing with uncertainty and vagueness, fuzzy logic [13] is used. Nowadays particle filters are used for tracking algorithms, because they compute the dynamic system through the observation has developed in Gordon and Salmand [14]. Another concept for solving stereo vagueness problem and uncertainty issue is fuzzy logic based particle filter algorithm. In Vermaak *et al* [15] the number of particles generated are computed by using the Fuzzy System. For Surveillance applications, it may require large set of variables so the system may face the difficulties to understand all the rules. For solving this problem hierarchical structures are used [16]. Solana *et al* [17] proposed a segmenting algorithm for finding the motion objects in H.264 type videos. It uses fuzzy techniques to depict the location and to measure the detected regions. Rubia E.O *et al* [18] proposed a prediction mechanism towards successive binary images. In Thomas lukasiewicz *et al* [19] covered the problem of uncertainness and unclearness in patterns and logics. David mercier [20] used a mathematical belief function for reducing the data from sources. Hirai *et al* [21] proposed a tracking system for back of human and shoulder. Leonid sigal *et al* [22] described a method for segmenting the skin on the sequences of video and it do segmentation on varying lighting conditions on tracking. In Yi sun *et al* [23] a system is proposed that tracks multiple locations in grouping scenes. Michael isard and Andrew Blake [24] solved the problem of tracking in dense environment. It used the condensation procedure which performed better in run time. Paul viola [25] used three approaches for detecting face namely integral image representation, adaboost and cascade classifier. In Raja Tanveer Iqbal [26] specific components are used to improve the performances of detection in various lighting conditions and poses. Papageorgiou *et al* [27] give a learnable framework for object detection based on the wavelet transform. Francois fleuret [28] escribed about region of interest(ROI) detection course to fine. It accepts all grayscale images and measures the executions in terms of false positives. In [29], shows the work of choosing appropriate features to make use in system learning and delivered in search problem. [30] suggested a algorithm for finding the head of a person by using the ellipse whose positions are updated as per the head movement. Micheal Harville [31] shows the model for foreground segmentation and background subtraction. Model gives more strength without loss of real time performances. In Ming-Hsuan Yang [32] describes various approaches for detecting faces. In [33], the object is detected by using Haar feature extraction and cascade classifier. Henry Schneiderman [34] a thesis submitted to demonstrate how to detect 3-d objects. Thomas Kailath [35] author used Bhattacharya distance for measuring the convergence and reducing the error rates. Jaco Vermaak [36] presents Monte Carlo techniques for many target tracking. Zia khan [37] proposed a particle filter with mcmc filter carries on moving targets. Gayathri a.patil [38] used fuzzy classifier and skin color for detecting the human face. kun peng *et al* [39] proposed a algorithm for detecting eye using pattern matching techniques.

Approaches Used for Face Detection and Tracking: In face tracking and pose estimation, [1] faces are detected by using simple shape model, color and texture. Photometric representations are used to model the internal structure of faces. Eigen faces are used to detect the faces. Tracking process is done by matching the scale and local estimates. Matching faces does not accurately track the face in every new frame. This problem is resolved by motion grouping. In [2], faces are detected by computing the target location and for tracking, kalman filters are used. First, the pixel location x_i of the destination is displaced at zero. Let b: r^2 _{1...n} be the measure applied at target pixel. Probability of color u is computed by employing a convex and monotonic function k: [0, 8] _ R. Robustness of the estimation has been improved after the weightage increases. Then target candidate are computed by denoting the pixel location of the target candidate centered at y in the current frame and the probability of color u in the target candidate is computed by

$$P_u(y) = c_n \sum_{i=1}^{n} K\left(y - \frac{x_i}{h} \right)^2 [b(x_i) - u]$$

For decreasing the value of distance, Bhattacharyya co-efficient is increased [2]. For tracking there are two shifters, first calculates each x and y co-ordinates and the second for changing velocity. By using the mean shift optimization, tracking process run on every new frame come after by a kalman filters which gives the determined position. In adaptive color based particle filter [3] a method is proposed to track the target in video sequences. It presents a particle filter which uses both the simple linear dynamic model and likelihood model based on colour histogram. Dynamic model predicts the state of target by $St_{+1=}$ A * S_t + W_t where A is the

deterministic component and W_t is a multivariate Gaussian. After predicting the state then next frame should be a target. In next frame likelihood model has been calculated. By using the color histogram these are done. Particle filter algorithm is comprises five steps. First, re-sampling is done for avoiding degeneracy. Then based on the dynamic model particles are propagated. On the basis of likelihood model, weights $\Pi^{(n)}$ of each and every particles are updated. In the new frame the posterior state is computed. Finally adapting the target color distribution for increasing the dependability and robustness. In [4] for detecting face detector is used which is given by the open source computer vision viola Jones algorithm is used for detecting the face. It works on grey level images and gives the output in the form of rectangular regions. Two tests are conducted to remove the false positives from the detected face. First test is to check whether the pixels satisfy three conditions such as 1, it should be a part of foreground. 2, it should have stereo information.3, it should not contain occluded pixels. Then the three values are fuzzified using appropriate linguistic variable, then defuzzified value gives visible person as output. If the defuzzified value greater than $á_1$ then it is passed to second test. In the second test, the objective is as same as to remove the false positives. It takes three input average differences, depth and standard deviation and these values are fuzzified and similar depth value is computed by the defuzzified values obtained. Two fuzzy systems use the mandami inference method. For tracking it uses the fuzzy based particle filter algorithm and five linguistic variables involved here are fuzzy system region information, fuzzy system face information, fuzzy system particle to position distance information, fuzzy system torso information. The fuzzy system confidence is computed by the defuzzified value of FSRI and FSFI. These fuzzy systems are constructed by using hierarchical fuzzy system [5]. In [17] video is given as input and it is fully decoded for extracting the vector points. Secondly the data which are extracted are marched by h.264 data. Noisy data are moved out by the guaranteed vectors computed. [18] uses fuzzy inference system for processing the crisp set to produce the output image. Fuzzy system is developed by using the if then rules which defined by the expert, genetic programming is used at end of the stage for optimization. [19] makes use of fuzzy description which combines programs and logic of fuzzy and finally got the truth values as outputs. In [40], projected model break down the input and output signals of noise. For identifying the noise fuzzification, maximize model is developed. In [21] author used robust sensing and recognition system. It gives strong tracking as it takes texture input from the dress and shoulder. In [22] skin colors are estimated by markov model and informed at run time. At each and every frame several inputs are applied like grading, transformation and translation. In [23], for combining the inputs it combines the Monte corlo and Dezert Smarandache theory. In [25], integral image helps the detector to find the region of interest very quickly and adaboost will choose the relevant features for detecting face. Cscade detector will quickly remove the false positives. [26] proposed a fuzzy method for detecting the area of interest and it choose the face based on visual aspect and move it over to the geometric classifier. [27] uses Support Vector Machine for separating the various types of object and it provides very low false positives. [28] got the fast detection while applying the spatial arrangement and testing the shape and performance in course to fine.

Performance Analysis: In face tracking and pose estimation [1] head tracker tracks 60 frames from various lighting conditions and twelve people were captured, manually cropped and smoothened to 64*64 pixels. Using kalman filter and temporal zero crossing a real time motion tracker has been implemented. The combination of model based and the motion based representation gives more robustness for the closed loop system. In mean shift and optimal prediction [2] proposed system are applied to multiple video sequences to compute the operational time and cost complexity. Tracker successfully tracks even in the presence of occlusion and the ellipse size (hx,hy) = (55,39) is obtained. System work 30 frames per second at 600MHZ pc which has been implemented in java. In [3] the algorithm is executed in matlab. Performance of the system detection mainly depends on the size of the rectangle and the number of particles used. 100 particles is enough for all experiments which has been proved, filter is capable of detecting face in a cluttered environment and changes in direction. Overall tracker performs well in 211 frames has proved. Algorithm is more robust even if frames having the similar colors. In [4] the performance of the proposed system has been tested in Intel core i5 2.67GHZ processor. At a time, trackers are capable of tracking 4 persons in real time. It has been tested in real time situations where two or more people move freely in the video. After testing the algorithm with various videos it requires 50 particles to finish the process. To measure performance between the various approaches, three measures are taken namely:

Table 1: Performance Measures

	Fuzzy based particle filter algorithm	Adaptive color based particle filter algorithm	Mean shift and optimal prediction
RMSE position	8.85px	35.99px	58.49px
RMSE rectangle size	4.88px	61.39px	220.87px
Processing time per cycle and person	22.64ms	12.62ms	17.65ms

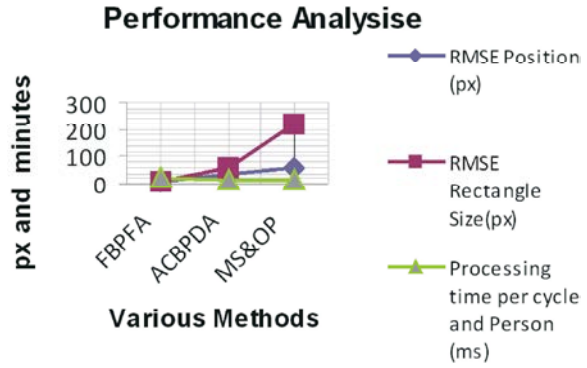

Fig. 1: Performance Analysis

1. Root mean square error position; 2. root mean square error rectangular size and 3.processing time taken per cycle and person and tabulated in table 1.

In the above Table.1 px stands for pixels, ms stands for milliseconds. Drawbacks of mean shift and optimal prediction are: It detects only one person at a time. It detects the face based on the skin color. So it losses its target by covering the neck portion which is not the area of interest. Drawbacks of adaptive color based particle algorithm is, it is not capable of differentiating the foreground and the background if both belong to similar colour. Finally, the target is lost due to this reason. But fuzzy based particle filter algorithm works well. It detects more than one people at a time. By using the stereo information, it differentiates the foreground and background. Tracker will not confuse if both objects belongs to similar colour. Hereby, the best algorithm for tracking is fuzzy based particle filter. It performs better than the other approaches as described above. The figure 1 shows the root mean square error position, root mean square rectangular size and processing time per cycle values for various approaches.

In [17] the approaches are tested in various sequences of videos. It takes 20, 26 and 38 ms as the execution time for 320*240,640*240 and 640*480 frame sizes. [40] After the data set applied, data are portioned in to two halves. First set of data are chosen for training and second set of data have used for testing. The projected approach got the least error in the context of performance measure like mean, median. In [22] getting good accuracy of 24 percentage when testing it in 17 out of 21 trials. In [26] compared to other method, accuracy rates increases up to 70% and false positive rates touches zero.

CONCLUSION

This paper presents the discussion about various face detection and tracking approaches, performance measures and drawbacks. Tracking faces is the main process used in different fields like surveillance application, computer vision and image processing. Particle filter algorithm is frequently used for tracking process even though it performs well it faces some difficulties like, if both objects are very close to each other with similar colour then the tracker fails. So, fuzzy based particle filter algorithm is suggested for solving this type of problem. Fuzzy system is constructed by hierarchical structures. It provides good performance than particle filters and requires more processing time.

REFERENCES

1. Stephen McKenna, shaogang Gong and J.J. Collins, 1997. Face Tracking and Pose Representation, Machine Vision Lab, Department of Computer science Queen mary and Westfield College, Mile End Rd. London.
2. Comaniciu, D., V. Ramesh and P. Meer, 2000. Real-Time Tracking of Non-Rigid Objects using Mean Shift, To appear, IEEE Conf on Comp. Vis. and Pat. Rec., Hilton Head Island, South Carolina.
3. Nummiaro, K., E. Koller-Meier and L.V. Gool, 2003. An adaptive color-based particle filter. Image and Vision computing, 21: 99-110.
4. Rui Paul, Eugenio Aguirre, Miguel Garcia-Silvente and Rafael Munoz-Salinas, 2012. A new fuzzy based algorithm for solving stereo vagueness in detecting and tracking people, International Journal of Approximate Reasoning, 53: 693-708.
5. Aherne, F., N. Thacker and P. Rockett, 1997. The Bhattacharyya metric as an absolute similarity measure for frequency coded data, Kybernetica, 32: 1-7.

6. Darrell, T., G. Gordon, M. Harville and J. Woodfill, 2000. Integrated person tracking using stereo, color and pattern detection, International Journal of Computer Vision, 37: 175-185.
7. Grest, D. and R. Koch, 2004. Realtime multi-camera person tracking for immersive environments, IEEE Sixth Workshop on Multimedia Signal Processing, pp: 387-390.
8. Moreno, F., A. Tarrida, J. Andrade-Cetto and A. Sanfeliu, 2002. 3d real-time head tracking fusing color histograms and stereovision, in: International Conference on Pattern Recognition, pp: 368-371.
9. Harville, M., 2002. Stereo person tracking with adaptive plan-view templates of height and occupancy statistics, Image and Vision Computing 2*:* 127-142.
10. Muñoz-Salinas, R., E. Aguirre and M. Garcfa-Silvente, 2007 People detection and tracking using stereo vision and color, Image and Vision Computing, 25: 995-1007.
11. Kil-jae, L. and Z. Bien, 1997. A model-based machine vision system using fuzzy logic, International Journal of Approximate Reasoning, 16*:* 119-135.
12. Bloch, I., 2008. Defining belief functions using mathematical morphology - application to image fusion under imprecision*,* International Journal of Approximate Reasoning, 48: 437-465.
13. Yager, R. and D. Filev, 1994. Essentials of Fuzzy Modeling and Control, John Wiley & Sons, Inc.
14. Gordon, N. and D. Salmand, 1995. Bayesian state estimation for tracking and guidance using the bottstrap filter, Journal of Guidance, Control and Dynamics, 18: 1434-1443.
15. Zhenjiang, C. and L. Zongli, 2008. Fuzzy particle filter used for tracking of leukocytes, Proceedings of the 2008 International Symposium on Intelligent Information Technology ApplicationWorkshops, (page. 562-565, Year of Publication: 2008).
16. Torra, V., 2002. A review of the construction of hierarchical fuzzy systems, International Journal of Intelligent Systems, 17: 531-543.
17. Solana-Cipres, C., G. Fernandez-Escribano, L. Rodriguez-Benitez, J. Moreno-Garcia and L. Jimenez-Linares, 2009. Real-time moving object segmentation in h.264 compressed domain based on approximate reasoning, International Journal of Approximate Reasoning, 51: 99-114.
18. Rúbia E.O. Schultz, Tania M. Centeno, Gilles Selleron and Myriam R. Delgado, 2009. A soft computing-based approach to spatio-temporal prediction, International Journal of Approximate Reasoning, 50: 3-20.
19. Lukasiewicz, T. and U. Straccia, 2009. Description logic programs under probabilistic uncertainty and fuzzy vagueness, International Journal of Approximate Reasoning, 50: 837-853.
20. David Mercier, Eric lefevre and Francois delmotte, 2012. Belief functions contextual discounting and canonical decompositions, International Journal of Approximate Reasoning, 53: 146-158.
21. Hirai, N. and H. Mizoguchi, 2003. Visual tracking of human back and shoulder for person following robot*,* IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 1: 527-532.
22. Sigal, L., S. Sclaroff and V. Athitsos, 2004. Skin color-based video segmentation under time-varying illumination, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26: 862-877.
23. Sun, Y. and L. Bentabet, 2010. A particle filtering and dsmt based approach for conflict resolving in case of target tracking with multiple cues, Journal of Mathematical Imaging and Vision, 36: 159-167.
24. Isard, M. and A. Blake, 1998. Condensation-conditional density propagation for visual trackings, International Journal of Computer Vision, 29: 5-28.
25. Paul viola and Michael Jones, 2004. Robust Real-Time Face Detection*,* International Journal of Computer Vision, 57(2): 137-154.
26. Iqbal, R., C. Barbu and F. Petry, 2006. Fuzzy component based object detection, International Journal of Approximate Reasoning, 45: 546-563.
27. Papageorgiou, C., M. Oren and T. Poggio, 1998. A general frame work for object detection. In International Conference on Computer Vision.
28. Fleuret, F. and D. Geman, 2001. Coarse-to-fine face detection. International Journal of Computer Vision, 41: 85-107.
29. Pat Langley, 1994. Selection of Relevant Features in Machine Learning, Proceeding of the AAAI Fall Symposium on Relevance.
30. Stan Birchfield, Elliptical Head Tracking Using Intensity Gradients and Color Histograms, Stan Birchfield Computer Science Department Stanford University Stanford, CA 94305
31. Michael Harville, Gaile Gordon, John Woodfill, 2001. Foreground Segmentation Using Adaptive Mixture Models in Color and Depth*,* IEEE .
32. Ming-Hsuan Yang, Member, IEEE, David J. Kriegman, Senior Member, IEEE and Narendra Ahuja, Fellow, IEEE,2002. Detecting Faces in Images: A Survey, IEEE transactions on pattern analysis and machine intelligence, pp: 24.

33. Rainer Lienhart and Jochen Maydt, 0000. An Extended Set of Haar-like Features for Rapid Object Detection, Intel Labs, Intel Corporation, Santa Clara, CA 95052, USA.
34. Henry Schneiderman, 2000. a Statistical Approach to 3D Object Detection Applied to Faces and Cars, Robotics Institute Carnegie Mellon University Pittsburgh, PA 15213 May 10.
35. Thomas Kailath, 1967. The Diver,gence and Bhattacharyya Distance Measures in Signal Selection, IEEE transactions on communication technology, 15: 1.
36. Jaco vermaak and simon J. godsill, 2005. Monte Carlo Filtering for Multi-Target Tracking and Data Association, IEEE transactions on aerospace and electronic systems, 41: 1.
37. Zia khan and tucker balch, 2005. Mcmc-based particle filtering for tracking a variable number of interacting targets, IEEE transactions on pattern analysis and machine intelligence, 27: 11.
38. Gayatri A. Patil and K.V. Joshi, 2013. face detection using skin color segmentation, International Journal Of Advance Research In Science And Engineering IJARSE, 2: 6.
39. Kun Peng, Liming Chen, Su Ruan and Georgy Kukharev, 2005. Robust and Efficient Algorithm for Eye Detection on Gray Intensity Face, Springer-Verlag Berl, pp: 302-308.
40. Hong, S., H. Lee and E. Kim, 2009. A new probabilistic fuzzy model: fuzzification-maximization (fm) approach, International Journal of Approximate Reasoning, 50: 1129-1147.