MULTI-TARGET TRACKING USING PREDEFINED REFERENCE SET IN NON-OVERLAPPING CAMERAS BY ADABOOST ALGORITHM

P. R. Magdum¹ and S. B. Jadhav²

¹,² E & TC Department, Bharati Vidyapeeth’s College of Engineering, Kolhapur

Abstract - Multi-target tracking is still recent approach which is used to find the same object across different camera views and also used to find the location and sizes of different object at different places [7]. Tracking and detection of moving objects are challenging research topic of many computer vision applications. Nowadays, the demand of surveillance camera is increasing rapidly which is useful for developing surveillance as well as monitoring purpose. Some previous methods are used for multi-target tracking that are color histogram, brightness transfer function (BTF) [11]. Many times it is not possible to cover complete area of interest by using single camera, such a cases there is need to use multi-target tracking system with non-overlapping field of views (FOV) [2]. In this paper we use method of feature extractions that is AdaBoost. The paper proposes the reference set based tracking in non-overlapping FOV’s due to overlapping FOV’s are having high cost. In this work widely used features HSV color histograms, Local Binary Pattern (LBP), Histogram of Gradient (HOG) are used to extract color, texture, shape of target. We use LBP, HOG, HSV color histogram features to determine person’s characteristics.

Keywords - Multi-target tracking, AdaBoost, feature extraction, LBP, HOG, HSV color histogram.

I. INTRODUCTION

The paper presents an object extraction system for multi-target tracking application that needs pixel wise extraction accuracy. Multi-target tracking is the research approach which is used to find the locations and sizes of multiple targets. Nowadays CCTV cameras are used for many purposes such as in public places, schools, airports, colleges etc. For security purpose the rate of use of CCTV cameras are increases [2]. The paper proposes to track the object robustly using multi-target tracking concept in non-overlapping field of views. There are two types of field of views that are overlapping FOV’s and non-overlapping FOV’s. The paper focused on non-overlapping FOV’s. The main goal of this project to find the appearance of same object across the multiple cameras output tracks. In computer vision automatic tracking of multiple targets is the recent topic which is used for industrial applications such as security purpose or monitoring purpose. The idea of multi-target tracking which is done by comparing both outputs that are output obtained from different cameras is compared with multi-cam dataset output. Using techniques for human detection, multi-target tracking system can play important role to capture location of people at public areas such as stores and travel sites and then produce congestion analysis to assist in the management of the people [3]. In such a way tracking system can monitor express ways and junctions of the road network. In some cases it is necessary to analyse the behaviour of people and vehicles also and check whether these behaviours are normal or abnormal. For example multi-target tracking system set in parking lots and supermarkets could track abnormal behaviour of theft which is useful to identify any criminals quickly and then contact the police immediately [1].

II. RELATED WORK

In this paper we propose the novel approach of multi-target tracking in non-overlapping cameras. With the help of tracking system we get the exact location and size of the object which are captured by the surveillance camera. Due to variation in illumination conditions, and camera imaging
characteristics, there might be significant appearance change of target across camera views. In multi-target tracking system there is need to track same object across multiple camera output tracks provided cameras are consist of non-overlapping regions [2]. Consequently the same target may appear in very different in two cameras which causes to error. Hence in most previous methods the appearance similarity is captured either using color histogram or based on brightness transfer function (BTF) [11]. BTF is used to find color difference between different cameras output. Our main aim is to implement multi-object tracking system in non-overlapping cameras. This project proposes reference set based appearance of the target to determine the similarity of multiple targets in different camera. The goal of multi-tracking across camera is to associate tracks in different cameras which contain the same object [2]. Our main goal is to implement the algorithm for multi-target tracking in non-overlapping cameras using predefined reference set.

III. CONTRIBUTION OF THE PAPER

The contributions of this paper are as follows
- The paper proposes the work for providing solution to track same target across multi-cam dataset [2].
- A novel reference set based appearance model to improve multi-target tracking in a network of non-overlapping camera.
- To track one target in one camera output video stream and register target features.
- On another camera i.e. multi-cam dataset output extracting features of target and comparing with pre-registered target features and to identify whether same target.
- This paper gives the detail information of art in personal identification at a distance and fusion of data from multiple cameras [8].
- We also produce detail discussion on future research direction in multi-target tracking system.

IV. PROPOSED WORK

The basic aim of multi-target tracking across non-overlapping cameras is to automatically recover the trajectories of all targets and keep their identities consistent while they travel from one camera to another camera [3]. Here we propose two methods for tracked target detection

4.1 AdaBoost Method:

Basically, AdaBoost is stands for ‘Adaptive Boosting’. It is good for things with soft edges like facial recognition. It is used for producing a strong classifier. The strong classifier is defined as threshold linear combination of weak classifier outputs. Adaboost is very simple to implement and it does feature selection resulting in relatively simple classifier. It is sensitive to noisy data and outliers. Adaboost improves classification accuracy and it can be used with many different classifiers [17]. It is called adaptive boosting because it uses multiple iterations to generate a single composite strong learner. Adaboost is an ensemble learning algorithm that can be used for classification. It contains many weak classifiers and minimizes overall error.

![Fig.1 Block diagram using AdaBoost method](image-url)
4.1.1 Camera 1 output:
In this block here first track moving object and take it as a camera 1 output. After this take its features and store it as a reference set.

4.1.2 Divide image with equal height:
After tracking object the whole body is dividing into two parts that are torso and legs. Torso is upper part of detection and legs are lower part of detection. Here whole body is divided into upper and lower part with equal height [2].

4.1.3 Register for further use:
All above obtained information is stored as a reference set which is used for comparison purpose. This is our register set which is use to comparison with different cameras outputs dataset.

4.1.4 Compare with register features:
In this block compare pre-recorded multiple cameras outputs dataset with the registered features to identify appearance of similar object. For comparison purpose here uses two methods that are as follows.

V. EXPERIMENT

5.1 Feature Extraction:
The features obtained from first block the features are extracted with using AdaBoost feature extraction method. AdaBoost is stands for ‘Adaptive Boosting’. AdaBoost is the method of feature extraction. The role of AdaBoost algorithm is used for selection purpose. It can be used in conjunction with many other types of learning algorithms to improve performance of the system [17]. It is referred as the best classifier. Here we take three kinds of appearance features that are HSV color histograms, Local binary pattern (LBP) and histogram of gradient (HOG). These appearance features are used to capture color, texture, and shape information of object. Here we are partitioning body image into torso and legs for each detection with equal height. For this total nine features are extracted that are Torso HSV, Torso LBP, Torso HOG, Body HSV, Body LBP, Body HOG, Legs HSV, Legs LBP, Legs HOG. For distance measurement the Bhattacharya coefficient is used [2]. To measure the similarity of two discrete or continuous probability distributions the Bhattacharya distance is proposed.

5.1.1 HISTOGRAM OF GRADIENT:
HOG is a feature descriptor in computer vision and image processing which is used for object detection. It is used to count occurrence of gradient orientation localize portions of an image. The image is divided into small connected regions called cells. A HOG direction is compiled. HOG has advantage over other descriptor, it operates on local cells and it is invariant to geometric and photometric.

5.1.2 HSV COLOR HISTOGRAM:
Histogram is a graphical representation of number of pixel in an image. X-axis represents tonal scale i.e. black at the left and white at the right. Y-axis represents number of pixel in an image in a certain area of tonal scale. A color histogram of an image represents the distribution of the composition of colors in the image. It shows different types of colors appeared and number of pixel in each type of the color appeared. HSV stands for the angle around the central axis corresponds to “hue”. The distance from the axis corresponds to “saturation”. The distance along the axis corresponds to “value”. Hue deals with the purity of the color. Saturation determines the amount of white light mixed with the original value. Value gives the intensity of color.

Principle of the formation of a color histogram:
We can simply count the number of pixel for each 256 scales in each of the 3 RGB color and plot them on 3 individual bar graphs. When we compute the pixel of different colors in an image if the color space is large then we can first divide the color space into certain number of small intervals.
Each of the intervals is called a bin. This process is called color quantization. Then by counting the number of pixel in each of the bins we get the color histogram of the image.

5.1.3 LOCAL BINARY PATTERN (LBP):

LBP is a type of visual descriptor used for classification in computer vision. LBP is a particular case of texture spectrum model. It has been found a powerful feature for texture classification. When LBP is combined with HOG it improves the detection performance.

VI. RESULTS

The following images show our result of experiment. Figure 1 is the original image which has been taken from dataset. Here we want to extract three features that are HOG which shown in figure 1a, HSV Color histogram which is shown as figure 1b and figure 1c indicate that RGB to HSV color transformed image, LBP features which is shown in figure 1d. Figure 2 shows the different image and its three individual features as per the figure 1. Figure 2 indicate another example of feature extraction.

![Fig.1](image1)

![Fig.2](image2)

6.1 TABLE ANALYSIS:

| Sr. No. | HOG features | HSV Color histogram | LBP features |
|---------|--------------|---------------------|--------------|
| 1       | 0.5051       | 0.00024             | 0.09803      |
| 2       | 0.5054       | 0.00275             | 0.32941      |
| 3       | 0.5108       | 0.09694             | 0.23137      |
| 4       | 0.5000       | 0.01362             | 0.39607      |
| 5       | 0.5052       | 0.00097             | 0.38823      |
| 6       | 0.4954       | 0.00105             | 0.35686      |
| 7       | 0.4955       | 0.00121             | 0.42352      |
| 8       | 0.4915       | 0.01241             | 0.41960      |
| 9       | 0.5108       | 0.01411             | 0.14901      |
| 10      | 0.5051       | 0.00681             | 0.21176      |
REFERENCES

[1] Weiming Hu, Tieniu Tan, Fellow, IEEE, Liang Wang, and Steve Maybank “A Survey on Visual Surveillance of Object Motion and Behaviors” IEEE Transaction on systems, MAN, and Cybernetics—Part C: Applications and reviews, VOL. 34, NO. 3, August 2004

[2] Xiaojing Chen, Student Member, IEEE, Le An, and Bir Bhanu, Fellow, IEEE “Multi-Target Tracking in Non-overlapping Cameras Using a Reference Set” IEEE SENSORS JOURNAL, VOL. XX, NO. XX, DECEMBER 20XX

[3] Xiaojing Chen, Zhen Qin, Le An, Member, IEEE, and Bir Bhanu, Fellow, IEEE “Multi-person Tracking by Online Learned Grouping Model with Non-linear Motion Context” IEEE Transactions on Circuits and Systems for Video Technology, VOL. XX, NO. XX, month year

[4] Xiaobai Liu, Liang Lin, Shuicheng Yan, Senior member, IEEE, Hai Jin, Senior member, IEEE, and Wenbing Tao “Integrating spatio-temporal context with multi-view representation for object recognition in visual surveillance,” IEEE Transaction on circuits and systems for video technology, vol. 21, no. 4, April 2011.

[5] Mario I. Chacon-Murgaui, Senior member, IEEE, and Sergio Gonzalez-Durante. “An adaptive neural-fuzzy approach for object detection in dynamic backgrounds for surveillance systems,” IEEE, Transactions on industrial electronics, vol. 59, no.8, August 2012.

[6] Xiaojing Chen, Zhen Qin, Le An, Member, IEEE, and Bir Bhanu, Fellow, IEEE “Multi-person tracking by online learned grouping model with non-linear motion context” IEEE Transactions on circuits and video technology, vol. XX, no. XX month year

[7] I. Saleemi, K. Shafique, and M. Shah, “Probabilistic modeling of scene dynamics for applications in visual surveillance,” IEEE Transactions on pattern analysis and machine intelligence, vol.31, no.8, pp. 1472-1485, Aug 2009.

[8] W. Hu, M. Hu, X. Zhou, T. Tan, J. Lou, and S. Maybank, “Principal axis-based correspondence between multiple cameras for people tracking,” IEEE Transactions on pattern analysis and machine intelligence, vol.28, no. 4, pp. 663-671, April 2006.

[9] X. Chen, Z. Qin, L. An, and B. Bhanu, “An online learned elementary grouping model for multi-target tracking,” in processing of IEEE Conference on computer vision and pattern recognition (CVPR), 2014.

[10] C.T. Chu, J. N. Hwang, K.M. Lan, and S. Z. Wang, “Tracking across multiple cameras with overlapping views based on brightness and tangent transfer functions”, in processing of ACM/IEEE International Conference on distributed smart cameras (ICDSC), 2011, pp. 1-6.

[11] T. D’Orazio, P. Mazzeo, and P. Spagnolo, “Color brightness transfer function evaluation for non overlapping multi camera tracking,” in processing of ACM/IEEE International Conference on Distributed smart Cameras (ICDSC), 2009, pp. 1-6.

[12] O. Yin, X. Tang, and J. Sun, “An associate-predict model for face recognition”, in processing of IEEE Conference on computer vision and pattern recognition (CVPR), 2011, pp.497-504.

[13] Seung-Hwan Bae, Student Member, IEEE, and KukJin Yoon, Member, IEEE, “Robust online multi-object tracking with data association and track management”, IEEE Transactions on image processing, vol.23, no.7, July 2014.

[14] Ying-Jia Yeh and Chiou-Ting Hsu, Member, IEEE, “Online selection of tracking feature using AdaBoost”, IEEE Transactions on circuits and systems for video technology, vol.19, no.3, March 2009.

[15] Saurav Subedi, Yimin D. Zhang, Senior Member, IEEE, Moeness G. Amin, Fellow, IEEE and Braham Himed, Fellow, IEEE, “Group sparsity based multi-target tracking in passive multi-static radar systems using Doppler-only measurements”, IEEE Transactions on signal processing, vol.64, no.14, July 15, 2016.

[16] Luis O. Jimenez-Rodriguez, Member, IEEE, Emmanuel Arzuaga-Cruz, student member, IEEE, and Miguel Velez-Reyes, Senior Member, IEEE, “Unsupervised linear feature-extraction methods and their effects in the classification of high-dimensional data”, IEEE Transactions on geosciences and remote sensing, vol. 45, no.2, February 2007.

[17] Shuangqing Wu and Hiroshi Nagahashi, “Parameterized AdaBoost: Introducing a parameter to speed up the training of real AdaBoost”, IEEE Signal processing letters, vol. 21, no.6, June 2014.

[18] Xiaojing Chen, and Bir Bhanu, Fellow, IEEE, “Integrating social grouping for multi-target tracking across cameras in a CRF model,” IEEE Transactions on circuits and system for video technology, vol. xx, month year.