Person tracking with non-overlapping multiple cameras

Sanjay Kumar Sonbhadra
Department of IT
IIIT, Allahabad
Prayagraj, India
rsi2017502@iiita.ac.in

Sonali Agarwal
Department of IT
IIIT, Allahabad
Prayagraj, India
sonali@iiita.ac.in

Mohammad Syafrullah
Program of master in CS
Universitas Budi Luhur
Jakarta, Indonesia
mohammad.syafrullah@budiluhur.ac.id

Krisna Adiyarta
Program of master in CS
Universitas Budi Luhur
Jakarta, Indonesia
krisna.adiyarta@budiluhur.ac.id

***Abstract*—Monitoring and tracking of any target in a surveillance system is an important task. When these targets are human then this problem comes under person identification and tracking. At present, large scale smart video surveillance system is an essential component for any commercial or public campus. Since field of view (FOV) of a camera is limited; for large area monitoring, multiple cameras are needed at different locations. This paper proposes a novel model for tracking a person under multiple non-overlapping cameras. It builds the reference signature of the person at the beginning of the tracking system to match with the upcoming signatures captured by other cameras within the specified area of observation with the help of trained support vector machine (SVM) between two cameras. For experiments, wide area re-identification dataset (WARD) and a real-time scenario have been used with color, shape and texture features for person's re-identification.**

***Index Terms*—Non-overlapping multiple cameras; Person detection; Person identification; Support vector machine; Video surveillance.**

I. Introduction

Since the birth of CCTV cameras in the early 1950s, video surveillance systems have been experiencing numerous continuous evolution. Due to the technological advancements, the surveillance systems became cheaper and ubiquitous, whereas the human resources required to supervise them is expensive. Meanwhile, the captured videos are often used merely as an archive, to verify a happening is known to have taken place. Large scale surveillance cameras are more useful real-time systems if they can distinguish events as they happen, and take action. In the presence of multiple cameras, massive data is generated and to handle and process these data is the most challenging task for any smart surveillance system.

In a smart video surveillance system, massive data processing is needed for extracting some useful information. If, the main objective of any surveillance system is to reduce effect of any financial or physical damage which can happen then the time required to take any decision is the most important factor. The decision time defines effectiveness of that video surveillance system.

Surveillance a large area, such as building, mall or campus to monitor the activities of any moving object (person, vehicle etc.), is very difficult task. More specifically, it is very hard to identify and track any suspicious person because field of view of single camera is limited and there are number of blind spots in case of multiple cameras. The team appointed for monitoring manually tries to find out any suspicious person in that target areas and try to track him or her by just watching the videos which are captured. This manual process takes a long time to perform any action. This paper proposes a fully automated person tracking mechanism in presence of multiple cameras. The objective of this research is to reduce the response time response time in identification of any suspicious activity.Tracking a person in multiple cameras of non-overlapping field is a challenging task due to bind spots, illumination change, pose change, view point change etc [1]. This research builds a relationship between every pair of cameras in presence of non-overlapping filed of view of cameras to track any person and take decision that person is suspicious.

Initially, the proposed model extracts the features such as color, shape and texture of a person when he/ she is observed by a camera. All this features together form a signature of detected person. At multi-camera level the model finds a relation between every two cameras on the basis of feature difference of same person between two cameras. Person re-identification is done on the basis of feature difference and trained support vector model between two cameras. Various distance measure methods are considered to compute the difference between two feature vectors. Predicated SVM score decides identification of any detected person. This system is useful for different real-time working places like coal mines, tunnels, forests etc. to track workers or visitor at several checkpoints. In case of any misshapen, this model can track and save lives of these people.

The rest of the paper is organized as follows: section 2 covers pre-existing research works whereas system description and proposed approach are discussed in section 3. Experimental results are discussed in section 4 and section 5 contains the concluding remarks.

II. Related work

The available literature on person tracking with the help of video surveillance mainly focused on person detection and its re-identification. Person detection is a special category of object detection task from video frames or images. Human beings are more complex than any other object in the term of

Fig. 1. Background image.

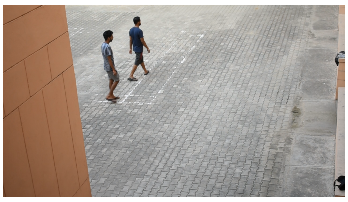

Fig. 2. Random captured image.

shape, color etc. because, humans wear clothes of various colors and styles. While walking pose of human beings changes frequently. It is also observed that mostly, the techniques of person detection depend on motion based or structure-based analysis.

In motion-based object detection techniques, Juhana et al. [2] proposed a model that extracts foregrounds from images which contain required object. They defined a background image similar to Fig. 1 and applied image subtraction between every image(Fig. 2) and background image to get a binary image similar to Fig. 3. Later, Blob analysis has been applied on binary images to get region of interest. This technique seems applicable for moving object detection, but it requires fixed video capturing device and stable illumination. This is a very simplest and easiest technique to detect any person from video. Major drawback of this technique is that it cannot categorized a moving object as person, vehicle or pet. It is also observed that this method is not applicable in densely populated area.

There are two types of background subtraction techniques as following:

a. Normal background subtraction: In this background is fix for all images.
b. Gaussian mixture model: In this background is updating during whole process and this update depend on definition of Gaussian model applied.

Structure of human body from front or back view is quite

Fig. 3. Binary image of foreground detected in Fig. 2.

Fig. 4. Person detection using HOG + SVM model.

similar and there are two structure-based person detection techniques as follows:

a. Haar cascade feature based.
b. Histogram of oriented gradients (HOG) and SVM based.

Haar Cascade feature based is a machine learning based object detection techniques, given by Paul et al. [3] inspire by face detection mechanism. In this approach, the proposed cascade function is trained with positive images (images which contain required object) and negative images (images which not contain required object). Authors defined haar features for objects, which are quite common in them. For human face, some haar features are region of nose and cheeks are brighter than reign of eyes, position and size of eyes, nose bridge and mouth etc.

Dalal et al. [4] used HOG + SVM model for person detection from any image. HOG is used as features and linear SVM classifier was trained on these HOG feature vectors of positive and negative sample images. For person detection in any image this model takes all possible sub images where human presents. A sample result is shown in Fig. 4 with SVM prediction values.

It is observed that the structure based techniques are advantageous over motion-based techniques because they are capable to differentiate human and non-human object and also able to detect any person which is not in motion. But these techniques take more processing time than motion-based techniques. Drawbacks of these techniques are detection of human from side view is poor and they also may give false detection as one of detected person is wrong as shown in Fig. 4.

Person detection methods operates independently on video frames which are captured during video surveillance and uses only those video frames in which person is detected. Mehmood et al. [5] used optical flow with feature matching and shape descriptors for person detection and tracking in single camera and extended this to multi-camera model. This method preprocesses the frames to reduce the effects of illumination change and noise reduction. Authors used optical flow to find region of interest in a video in which person is present and used region-based descriptor, i. e. minimum and maximum area for human availability in that video frame. Later, Chen et al. [6] use histogram of gradient (HOG) and support vector machine (SVM) for person detection.

Human wear clothes of different colors and shades. The way a person walks or stand is somehow different from another person. Structure of human body is mostly similar, but size and structure together can give some details about any person. Color, pose and structure gives different level of re-identification of any person. Color is one of the dominating feature which help us in differentiating between two persons while tracking a person. Chen et al. [6] added color features for object identification in multi-camera environment. There are different color models to define an object in several ways. RGB color space store color information in form of red, green and blue colors. HSV color space consist of hue (i.e. six primary color), saturation and value, and is closer to how human eye perceive an object. Chi et al. [7] take HSV color histogram over RGB color histogram as features for real time person tracking.

Afterwards, Martinel et al. [8] proposed a novel model to reduce the search space with the help of camera network topology. This model consideres color histogram of RGB and HSV and add shape and texture features for person identification in multi-camera environment. Shape and pose as alone are not very useful in person re-identification, but it can enhance the performance, when it is used together with color features. Authors used pyramid histogram of oriented gradients (PHOG) feature for storing information of shapes in any digital image. PHOG is computing HOG at different levels of same image, where level is number of equal rectangular sub-regions in which image is distributed. To store information about texture of any image haralick features [9] are used. Chen et al. [6] used number plate as additional feature for vehicle re-identification. This research also suggests to use face as additional feature for person re-identification, if human face can be detected and captured from CCTV cameras efficiently.

Similarity (or dissimilarity) measure and person re-identification done at multi-camera level. Li et al. [1] divided similarity measure techniques in two category: the first technique is RankSVM whereas another one is relative distance comparison (RDC). RankSVM method requires training and testing data samples, and gives ranks of target samples in order of similarity with test sample. Whereas, RDC methods consist of statistical computation between two or more set of features and give mathematical value as a similarity measure. Chi et al. [7] use chebyshev as distance measure to find dissimilarity between two features. Smaller the chebyshev value more similar the person.

III. Proposed approach

In this research, a novel model for person tracking over a multiple non-overlapping camera network is proposed. The whole process is divided into two parts: system configuration and system operation. It consists of following functionalities:

a. Person detection.
b. Feature extraction.
c. Feature difference.
d. Person re-identification.

A. System description

a) System configuration: There are multiple cameras in any surveillance system and all are somehow different from each other. The objective is to find a relation between every two cameras when they are capturing same persons. Fig. 5 shows the flow chart of system configuration. It is a cascaded approach, which first chooses a set of people, who are going to walk across whole surveillance area. We have given unique camera ids to each camera and person ids to each person. At single camera level the video is processed frame by frame to detect a person. When a person is detected, it extracts image of that person and region of interest as binary mask as shown in Fig. 6 and also note their person ids. All extracted images are grouped and masked according to camera ids. Feature extraction is performed on each image of every group and features of color, shape and texture are extracted. We compute feature difference between every two cameras in which one is test camera and another is target camera. The test camera is that from which the person needs to be tracked and target camera is that where images are registered with known person id. Since, this is configuration phase, so the person ids are known of persons in test cameras, but in operation phase person ids of persons is not known in test cameras. A binary SVM classifier is trained between every pair of test and target cameras. Each row in feature difference dataset is a training feature vector for SVM, and if that row belongs to same person ids in test and target camera then class of training feature vector is positive otherwise negative class. This trained SVM act as a relation between two cameras while identifying any person.

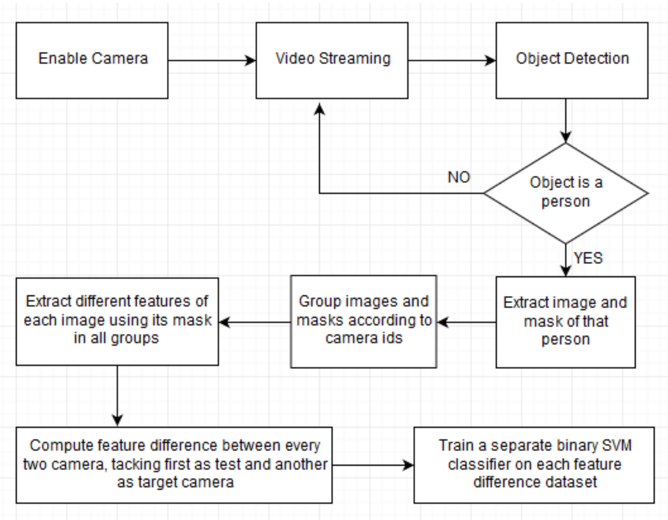

Fig. 5. Flow chart of system configuration.

Fig. 6. Image and its binary mask [10].

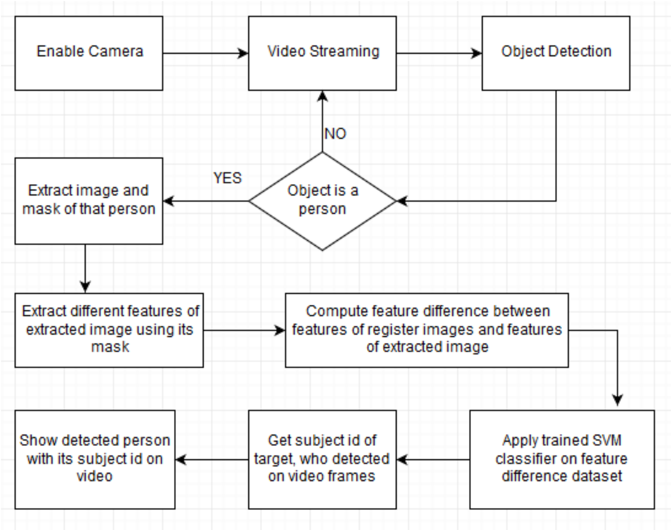

Fig. 7. Flow chart of system operation.

b) System operation: System operation is the second phase of proposed prototype. A trained SVM classifier is available between every pair of test and target cameras. The proposed model captures images and binary mask of each and every person at the starting of tracking system with unique person ids. These images are known as registered images and the camera use to capture these images is known as target camera. Feature extraction is performed on registered images with the help of their mask for further processing. Fig. 7 shows the flow chart of system operation. It is a cascaded approach, where initially person detection and feature extraction done at single camera level, then feature difference and person re-identification done at multi camera level.

At single camera level in each camera video is processed frame by frame to detect a person. When a person is detected, it extracts image and binary mask of that person. Here, the camera which detected person is known as test camera. The feature extraction is performed on extracted image with the help of its mask. The feature difference between features of registered images and extracted image is calculated to get feature difference dataset. Each row of feature difference dataset is used to trained SVM between test and target cameras to get SVM classes with predicted score for each row. For predicting person id of detected person on test camera highest SVM score of positive class is considered with corresponding person id from registered images of target camera. This person id is the predicated person id of detected person on video frames.

B. Person detection

Person detection is performed at single camera level in each camera to extract images and masks of detected persons for further processing. Generally, CCTV cameras are operating at 24 to 26 frames per sec (fps), and images of a person are not going under major changes in two continuous frames. So, without any significant loss of information choice of 8 to 10 fps is sufficient for person detection. The proposed methodology uses HOG + SVM model given by Dalal et al. [4] for person detection as shown in Fig. 4. It uses normal background subtraction model and detection boundaries to extract binary mask of detected person. Fig. 6 shows such one example of extracted image and its binary mask.

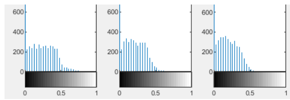

Fig. 8. Histogram of red, green and blue channels of given image in Fig. 6 respectively.

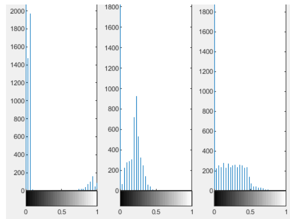

Fig. 9. Histogram of hue, saturation and value channels of given image in Fig. 6 respectively.

C. Feature extraction

Initially, pre-processing is performed on all extracted images and masks, where dimension normalization is applied so that all images and masks become of same dimension (i.e. of specific height and width). Later, enhancement on images is done, which includes noise reduction and edge sharpening is performed. Image histogram equalization is performed to all the images to achieve similar brightness for all the images. Finally, region of interest is chosen in the image for further processing.

a) Color features: The proposed method considers RGB and HSV color models for color feature of person. It computes histograms for both RGB and HSV color features. Each color model consists of three different channels. Each channel is 8 bit long which gives 256 shades of same color. Since, adjacent shades of color are not change drastically, all the 256 shades are grouped in 32 groups in continuous manner. Thus, there are 32 bins per channel. This approach gives three 1-D vector for RGB and HSV containing 32 elements each. as shown in Fig. 8 and Fig. 9. Different weights can be given to each channel according to their importance.

b) Shape features: For extracting shape information of a person, the proposed model uses pyramid histogram of oriented gradients (PHOG) feature of an image. PHOG features are computed at different levels for all three channels of HSV image. In this work, 9 bin per region per channel have been chosen. PHOG feature of image at level three (i.e. $2^3 \times 2^3 = 64$ regions) of Fig. 6 as shown in Fig. 10. This process gives three 1-D vector each consists of 765 elements. Different weights are assigned to each channel according to their importance.

c) Texture feature: Haralick features [9] are used to store information about texture of any image. it is observed

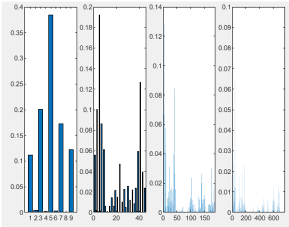

Fig. 10. pdf of PHOG at level 0,1,2 and 3 of hue channel of image in fig. 6 respectively.

Fig. 11. Images of 10 persons from 3 different cameras in WARD dataset [8].

that for rotation invariant feature, haralick features are more suitable. For this, gray level co-occurrence matrices (GLCMs) is calculated from gray-scale image of given RGB image. Haralick feature contains 14 statistical measures (e.g. correlation, entropy, variance etc). Thus, this phase generates one 1-D vector consists of 14 values.

D. Person identification

RGB, HSV, PHOG and Haralick features together form signature of a person. In this proposed model, distance matrix is computed for all respective 1-D feature vectors between test image and all target images. Chi-squared distance has been used to compute distance between RGB, HSV and PHOG features of images as in Eq. 1, where x and y are two feature vectors of same type, containing n elements each.

$$Chi - squared\ distance = \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i - y_i)^2}{(x_i + y_i)} \qquad (1)$$

Square of euclidean distance is used to compute the distance between Haralick features ((2)).

$$square\ of\ Euclidean\ distance = \sum_{i=1}^{n}(x_i - y_i)^2 \qquad (2)$$

After computing all the distance matrices, normalization is performed.

IV. Results and discussion

A. Dataset description

Wide area re-identification dataset (WARD) [10] captured in real-time environment has been used for experiments. It contains 4786 images of 70 different persons from 3 non-overlapping cameras. All the images have been captured at spatial resolution of 320 × 240 at 1 fps. Captured images size varies from 36 × 15 to 189 × 70 spatial dimension. Images with illumination and pose changes have been captured as shown in Fig. 11. It consists of binary mask of images which contains area of interest as shown in Fig. 6.

TABLE I
Feature difference dataset between test camera 1 and target camera 2 of WARD dataset.

rgb distance	hsv distance	phog distance	haralick distance	class
0.15321092	0.28384486	0.525006505	0.004234493	1
0.0805233	0.18712254	0.091253355	0.009831016	1
0.10536254	0.35250833	0.373341279	0	1
0.07497174	0.15953558	0.18716894	0.019855165	1
0.00490798	0.24960395	0.516470084	0.028932588	1
0.57295255	0.57718669	0.81904955	0.936046229	-1
0.5557917	0.48487088	0.549962541	0.970949694	-1
0.69871302	0.61433501	0.477288946	0.903180503	-1
0.66942374	0.61768988	0.598045711	0.910804392	-1
0.54032479	0.57260641	0.776959249	0.97417839	-1
0.6654321	0.57404439	0.390140833	0.238035163	-1
0.53553092	0.45256118	1	0.646295912	-1

TABLE II
Table of description of SVM models between two cameras of WARD dataset

Test Camera	Target Camera	Training Accuracy (%)	Testing Accuracy (%)
1	1	99.0171	97.8933
1 or 2	1 or 2	92.4171	90.5600
1 or 3	1 or 3	93.6286	90.2800
2	2	98.1714	97.5733
2 or 3	2 or 3	93.2914	90.6267
3	3	97.9429	96.6400

B. Feature difference dataset generation

Feature vectors from images and difference between feature vectors of two images are computed for further processing in this approach. Randomly, 10 distinct person ids have been selected and 5 images of each distinct person from test camera. Similarly 50 distinct person ids and 5 images of each distinct person have been selected from target camera randomly. Hence, total 50 (i.e. 10 × 5) images form test camera and 250 (i.e. 50 × 5) images from target camera have been randomly selected for experiments. Afterwards, feature difference between every pair of images, one from test camera and another from target camera is computed. This pre-processing step generates 12500 (i.e. 50 × 250) distinct vectors, which have features: RGB, HSV, PHOG, Haralick distances and labeled class. This generated data contains 10 of feature vectors of positive class (+1) and rest are of negative class (-1). This type of feature difference vector is generated between every pair of test and target cameras. Table I shows first few feature vectors of feature difference samples between test camera 1 and target camera 2, where camera 1 and 2 are camera number used in the WARD dataset.

C. Relationship between two cameras

A binary classifier SVM acts as a relationship between two cameras during the process of person re-identification. WARD dataset uses three different cameras (here name them camera

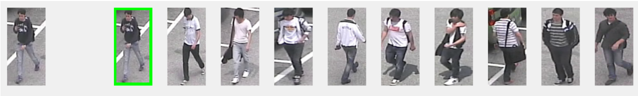

Fig. 12. Left most is test image. Top 10 matched images from left to right from target folder. Correct matched images have border around them.

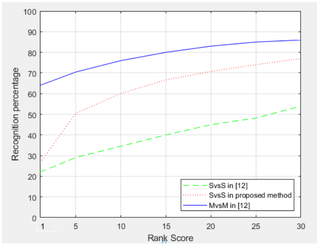

Fig. 13. CMC performance comparison curve between proposed method and existing approaches of [11].

1, 2 and 3) for capturing videos of peoples. Six different SVM classifiers have been used between every possible pair of cameras as shown in Table II. In ideal case, if there are n cameras, and one camera is used as target camera and other as test cameras, so total n different SVM need to be train between every possible pair of test and target cameras. Feature difference dataset as shown in Table I is used to train a SVM classifier between two cameras. MATLAB function 'fitcsvm' is used to create and train a binary SVM. There are 25000 instances in one feature difference dataset and in which 2500 instances (10%) are positive class instance. 17500 instances (70%) are used for training and 7500 instances (30%) for testing of SVM classifier. Radial basis function kernel function is used in these SVM models.

D. Person re-identification results in WARD

Every person re-identification dataset has its own challenges in terms of pose variation, illumination change etc. Martinel et al. [11] uses cumulative matching characteristic (CMC) curve to validate various approaches. In the proposed approach, 1000 images from WARD dataset have been randomly selected irrespective of person ids and camera ids as test images and for each test image a camera id as target camera id has been randomly selected and within that camera id 50 images of different person ids have been randomly selected as target images. Feature extraction is performed on all test images and target images, and feature difference operation is performed on each test image and their respective generated target images. Camera ids of test image and target images, and trained SVM classifier between these cameras are known. Feature difference dataset is used for this SVM to predict class of each feature vector. The target images are sorted in descending according to their SVM score of positive class. One of the result shown in Fig. 12.

There are two type of approach to build signature (collection of feature vectors) of a person. The first one is single shot in which single image is use to build signature of a person and the another is multiple shot in which multiple images are used. In the proposed approach single shot method is used to build signature of any person. Fig. 13 shows CMC curve as comparison of proposed approach and two approaches in [11] on WARD dataset. In this research, single shot vs single shot (SvsS) is used in which signature of test and target person build from single image. Martinel et al. [11] uses SvsS and multiple shot vs multiple shot (MvsM) in which multiple images are used to build signature of test and target persons.

From Fig. 13 it is evident that rank 30 score has 77% recognition rate in SvsS method of proposed approach while method of [11] has 55% recognition rate. The Proposed approach outperforms SvsS method of [11] as observed in CMC curve. This method does not track any person based on appearance on some previous frames with in same video, so multiple images of test person can not be captured in this, so MvsM method cannot applicable. MvsM method of [11] perform better than proposed approach because it randomly selects multiple images of same person from multiple cameras. Signature builds from multiple images of different cameras compensate pose, illumination, viewpoints variations in between cameras.

E. Person re-identification results in real scenario

To validate proposed approach, person re-identification task is performed on different videos, captured from two different stationary cameras (name them camera 1 and camera 2) which have non-overlapping field of view. There are five different subjects (persons) who are going to move across these two cameras. Fig. 14 shows that subjects from left to write namely person 1, 2, 3, 4 and 5. Three videos from camera 1 and two

Fig. 14. Five different subjects captured by camera 1 and 2 in real scenario.

Fig. 15. Detected person in a video frame.

Fig. 16. Foreground detection in video frame.

Fig. 17. Detected person and its binary mask.

Fig. 18. Person re-identification result 1 from camera 1 in real scenario.

videos from camera 2 have been captured. Camera 1 is used as target camera and video 1 of camera 1 is used for capturing target (registered) images of these five subjects. Fig. 15 shows that how a person is detected in a video frame, and Fig. 16 shows foreground detection of that video frame. Fig. 17 is extracted image and binary mask of detected person in video frame of Fig. 15.

Fig. 20. Person re-identification result 1 from camera 2 in real scenario.

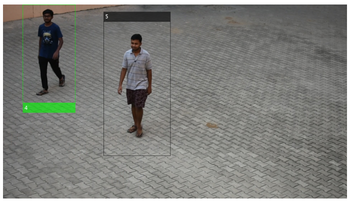

Fig. 21. Person re-identification result 2 from camera 2 in real scenario.

Images captured from video 1 of camera 1 and video 1 of camera 2 are used for generating feature difference dataset to train SVM between camera 1 and 2. Fig. 18 and Fig. 19 are person re-identification results on videos captured from camera 1 and Fig. 20 and Fig. 21 are person re-identification results on videos captured from camera 2. Same person in all video frames have same color rectangular boundary and same person id.

Fig. 19. Person re-identification result 2 from camera 1 in real scenario.

V. Conclusion

In this paper, multiple features (color, pose and texture) are used to build signature of a person, which enhances the rate of person re-identification whereas binary SVM is used as a relationship between two cameras in the process of person re-identification in non-overlapping field of view. CMC curve analysis has been done for person re-identification on WARD dataset for proposed approach. Experimental results show that 77% recognition rate is achieved as rank 30 for SvsS method. Many pre-existing approaches in this field miss the part of person detection from video frames. So, we create a small real-time scenario of two cameras of non-overlapping field of view to evaluate proposed approach and overall result is satisfactory. The false detection cannot be completely removed, but effect can be minimized by adding background subtraction to find percentage of area of detected region is in motion. We also get false re-identification; that can be treated as data drift and minimization of this can be a future work.

References

[1] L. Hou, W. Wan, K. Han, R. Muhammad, and M. Yang, "Human detection and tracking over camera networks: A review," in *2016 International Conference on Audio, Language and Image Processing (ICALIP)*, pp. 574–580, IEEE, 2016.

[2] T. Juhana and V. G. Anggraini, "Design and implementation of smart home surveillance system," in *2016 10th International Conference on Telecommunication Systems Services and Applications (TSSA)*, pp. 1–5, IEEE, 2016.

[3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, vol. 1, pp. I–I, IEEE, 2001.

[4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1, pp. 886–893, IEEE, 2005.

[5] M. O. Mehmood and A. Khawaja, "Multi-camera based human tracking with non-overlapping fields of view," in *2009 Fifth International Conference on Image and Graphics*, pp. 313–318, IEEE, 2009.

[6] Z. Chen, W. Liao, B. Xu, H. Liu, Q. Li, H. Li, C. Xiao, H. Zhang, Y. Li, W. Bao, *et al.*, "Object tracking over a multiple-camera network," in *2015 IEEE International Conference on Multimedia Big Data*, pp. 276–279, IEEE, 2015.

[7] J.-Y. Chi, M.-H. Sheu, and C.-C. Sun, "Non-overlapping multi-camera object tracking algorithm and system design," in *2017 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, pp. 409–410, IEEE, 2017.

[8] N. Martinel, G. L. Foresti, and C. Micheloni, "Person reidentification in a distributed camera network framework," *IEEE transactions on cybernetics*, vol. 47, no. 11, pp. 3530–3541, 2016.

[9] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, no. 6, pp. 610–621, 1973.

[10] R. S. Lab, "Person re-identification datasets.," Online; accessed January 3, 2020.

[11] N. Martinel and C. Micheloni, "Re-identify people in wide area camera network," in *2012 IEEE computer society conference on computer vision and pattern recognition workshops*, pp. 31–36, IEEE, 2012.