International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-4, February 2020

Real Time Face Recognition and Tracking System with Multi-Camera Arrangement

Kantilal P Rane

Abstract*: Object tracking and face reorganization has received tremendous attention in the video processing community due to its various applications in video surveillance, traffic monitoring and so on. A single camera is not capable to scan 3d view of specified space. So, we use multiple cameras, placed in different sections of the area with overlapping region in field of view (FOV). Every camera will capture the video scene of itself FOV. The system is able to track human successfully by setting up correspondence between objects captured in multiple cameras. Thus, it saves the hectic job of manual tracking. There is a search window available for each object that gives the object's trajectory. Tracking of object will be given by continuation of this process. For monitoring objects in areas like car parking, banks, hotels etc for security purpose, this system is best. Over the last couple of years, many algorithms and results have been presented for the problem of object tracking and recently the focus has been concentrated on real time person tracking with multiple cameras. Secondly, face detection is one of the best ways of identification. The main applications of automated face recognition are of biometric authentications and surveillances. Face recognition systems has became popular in biometric field as it is non intrusive and does not require the human interference. Up to that, there is no solution or technique that provides robust methods. This paper presents the detection of the face of the person, recognize and do tracking with use of multiple web cameras. Generally in daily camera security systems, cameras have been continuously remain on and large data storage is required in the system. In this real time object tracking system, Infrared sensors are used which indicates presence of person or object. Cameras will turn on only when object is detected by sensor after then the face recognition is carried out. It has capability of high speed processing and achieved low computational requirements. In similar areas, efficiency, accuracy, and speed of identification are the main tackled issues.

***Keywords*: HOG, multiple cameras, multiple sensors, and RGB- GRAY.**

I. INTRODUCTION

Nowadays, there is growing demand of security that leads with great importance to investigate the military applications based video surveillance system. More number of surveillance cameras is installed in the areas such as banks, stations and borders.

Revised Manuscript Received on February 28, 2020.
***** Correspondence Author
Kantilal P Rane*, Professor, KCES's COEIT, Jalgaon, Email: kantiprane@rediffmail.com

The available data makes it impracticable to guarantee watchful monitoring by human operators for long periods of time because of routine and fatigue. As a result, in the case of event suspicious activities, video feeds are mainly used for forensic measurements. To make helpful to the human operators for identification of important or suspected persons in videos, multi camera object tracking and face detection system can be used. For moving object detection and tracking; fast, efficient and robust methods are required.

II. LITERATURE SURVEY

CCTV is the first generation in visual surveillance system which was developed in 1940's for the military use in Germany for observing the launch of V2 rockets. They used black and white monitor to view launch of rocket V2 [1]. Video surveillance was used in stores, homes and banks to secure their properties. In 1980's, this traditional CCTV system are used analog video cameras which are connected by coaxial cables to surveillance system for monitoring by operators through the cameras connected to videotape recorders. They used VCR for recording of footage and used multiplexer that allowed multiple cameras to be viewed on single monitor. In 1990's VCR were replaced with DVR with the data archived on hard drives. Then in mid 2000's, they have network connections so the video data can be stored on servers. Video servers allowed real time video capture for remotely viewing facility on smart phone, PC or Laptop. In later 2000's, an IP network system were used where the data is continuously being transmitted over the network improving quality of image and resolution. It was reported that the United Kingdom has more number of cameras per person than other country in the world [2]. Initial stage required for video analysis is moving object detection [3]. The next step in the video analysis is object tracking [4]. The final step of an "intelligent" visual surveillance system is to analyze the information obtained of the video and to identify actions in a scene.

In 1960's, initially semi-automated recognition system for facial recognition was developed to locate the various features like nose, ears, mouth and eyes on the images. In 1970s, twenty first specific markers are used by Goldstein and Harmon [5] such as hair color and thickness of lips for specific recognition. In 1988, standard linear algebra technique was developed by Kirby and Sirovich [6] for face recognition. In 2007, 2006 FRVT results was reported by NIST [7] that initially demonstrated automated face recognition system experimented as well as or better than a human faces taken with variant lighting conditions.

They also showed an improvement across vendors from the FRVT 2002 results [8]. However, the best performing systems still only achieved a false reject rate (FRR) of 1 in a 100 measured at a false accept rate of 1 in one thousand. This translates unable to accurately identify 1% of available database but falsely identify 0.1%. These best of results were used for controlled illumination. The main reason for poor performance is that faces have a large variations and repetitions of many face images. Their lightening conditions and expressions are also different [9]. The faces can also be obscured by hair, jewelry, etc., and its appearance changed by make-up. This has important application in forensics and crime detection, for example photo and video match for identification of missing persons.

Major issues have become found in authentication and identification in today's digital world. Face detection has importance in authentication and identification [9]. With the use of this, people identify each other and so it is natural way that can teach computers for same work. Face recognition systems are becoming very much popular in biometric authentication as they are non-intrusive and do not really require the human's cooperation, though the recognition accurateness is quite not high for large scale application.

As shown in below Fig 1 is the flow of Traditional system [10].The details are specified as follows.

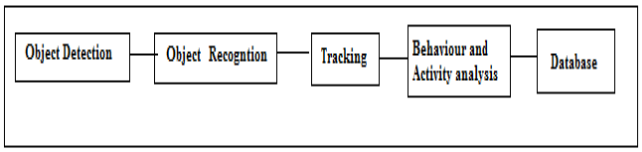

Fig 1: Traditional surveillance system

A. **Object detection**

In this separation of the foreground pixels with the background pixels is done. Two main aspects for object detection are 1) background subtraction process and 2) temporal difference process. First one consists of the subtraction of two consecutive frames done by thresholding. The second one is based on the subtraction of a background model and the current image by a labeling process. The temporal difference technique has good performance in dynamic environment due to its adaptive information. After extracting all the relevant object pixels, it was shown unsatisfactory results. While, background subtraction has better performance to extract object information but it is sensitive to various changes in the environment.

B. **Object Representation**

After detecting an object from frames, there are different types of object shape representations like primitive geometric shapes, point and skeletal model.

C. **Object Tracking**

Tracking techniques is divided into two models as 3-D Models and 2-D models with or without explicit shape models [11].

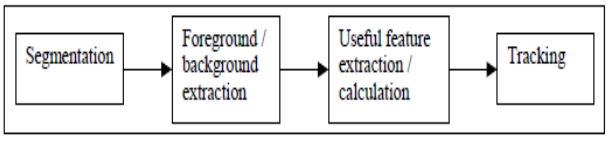

Fig 2: Object Tracking Processing

1) Segmentation: It is the method of identifying components of the image which includes operations such as boundary detection and thresholding. Boundary detection detects edges in the image [11, 12]. Thresholding is method of reducing grey levels of the image.
2) *Foreground extraction*: It is separation process of the foreground and background image. Foreground extraction method made use of difference images in order to find objects that are moving and objects that are not moving. The result of the subtraction is a difference image.
3) *Background extraction*: As soon as foreground is extracted, subtraction operation can be used to extract the background.
4) *Feature Extraction*: Features are selected based on object representation. Colour features are generally used for histogram based.
5) *Object Tracking*: It is the process of finding path followed by object in each frame of video sequence. Tracking can be accomplished by methods like Point Tracking, Kernel Tracking, and Silhouette Tracking.

Due to its major role in tracking, it uses the object representation e.g. color is used for histogram based representations and edges are used for contour based representation [13]. Usually many tracking algorithms are combining these two features.

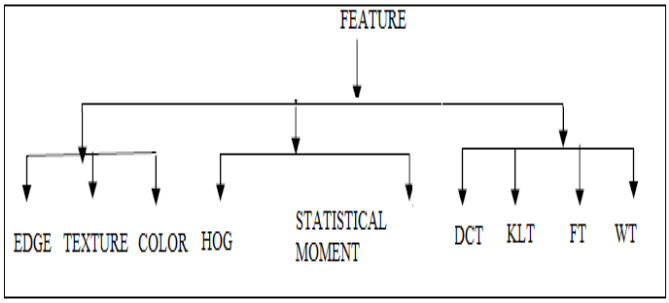

Fig 3: Types of Features

III. PROBLEM STATEMENT

With the existing system following problems are there:

D. A single camera shows limited view of the person that may enter in to the FOV of camera.
E. Object carrying some goods will be hided.
F. For large, area human intervention is needed for object tracking.
G. Large memory requirement for continuous on set up.

IV. METHODOLOGY

The face is one of the important objects in an image or video. An important ability of human face recognition, surveillance systems, video-conferencing is to detect the exact location of faces in image and then facial features extraction is performed. The proposed method includes proximity sensors. Logitech web cameras, PIC controller, MAX 232, RS 232, and MATLAB tool. Given below (Fig 4) is the experimental set up for the proposed system.

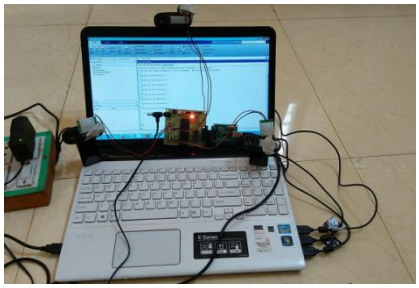

Fig 4: System Arrangement for Multi-camera Object Tracking

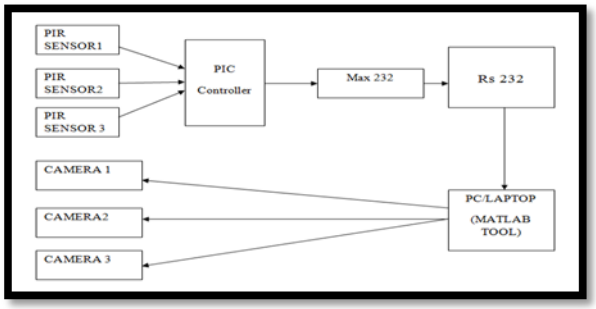

Fig 5: System Flow for Multi-camera Object Tracking

As shown in Fig 4, three sensors are used to sense human. Depending on sensed output of proximity sensor, particular camera related with that sensor will turn on. That means if first sensor is sensing human presence then camera first will turn on. Similar is happened with sensor and camera number 2 and 3. Sensors output is connected to PIC controller, and through the serial cable connector, it is provided to operating system. Proximity sensors are interfaced with microcontroller that is embedded with hi-tech C language programming. MAX 232 is used for the conversion of voltage level. This is required to make compatible TTL device with serial port of PC. This chip contains charge pumps which pumps the voltage to the desired level. Primarily Max 232 is used in serial communication. To sense the presence of an object, proximity IR sensors are used. Communication RS 232 is standard for serial communication of data. PIC controller is used to carry out number of tasks. It is normally used to interface peripheral with computer system.

The real time videos are obtained from the web camera. These video signals are given as an input to the CPU or Laptop where the following functions are performed.

1. As soon sensor sense presence of object it turns ON web camera.
2. Then the video acquired from the web camera are converted into frames.
3. The motions in the video are captured.
4. Once the motion is detected, template matching is performed to detect the object.
5. After the detection of the object, the camera placed on the respective sensor is turned on and the person is tracked by system by noticing output in window containing option 1 to 6.
    1. Option 1: If object sensed by Sensor No 1
    2. Option 2: If object sensed by Sensor No 2
    3. Option 3: If object sensed by Sensor No 1 and 2
    4. Option 4: If object sensed by Sensor No 3
    5. Option 5: If object sensed by Sensor No 1and 3
    6. Option 6: If object sensed by Sensor No 2 and 3
    7. Option 7: If object sensed by Sensor No 1, 2, 3

Face detection is the analysis of the behaviour method of detecting and extracting the facial features in a gray-level image and it is divided into two. Initially, by testing all valley regions in an image, possible regions of human eye are detected. Genetic algorithm is used to select a pair of eye of human for creating a possible face candidate. The fitness value of each object is checked out based on Eigenface's projections. All possible face areas are normalized for illumination for improvement of the level of detection reliability. Performing the number of iterations, all the face images with high fitness value are chosen for again verification. At this stage, the face similarity is measured and presence of differential facial features is checked and verified for each face of object. The facial features are obtained by evaluating the normalized face regions. Eyebrow, nose and mouth corner are the various facial features extracted.

For object face detection and face identification, HOG is used. It is shape descriptor. HOG features are calculated by taking orientation histogram of edge level in local region. Steps for calculating HOG are as follows.

1. Fixed window is chosen for face recognition.
2. Normalize Gamma and colour correlation for each block.
3. Compute the Gradient for all the blocks.
4. Collect weighted votes for each gradient orientation for each spatial block.
5. Collect HOG for all the blocks in the image.

V. FLOW CHART OF PROPOSED SYSTEM

The main aim of implementing Real time face tracking system with multiple cameras is to sense the person with proximity sensor. Here we are using multiple sensors with particular cameras. If person is detected by sensor particular camera which was connected to sensor output is turned ON. And video of that person is captured. Still image of that person is separated out and cropping of face shape is done. After face shape cropping, feature extraction is carried out. Using Histogram of gradient technique, frames are converted into block string. Finally already stored training images of persons are comparing with image captured from cameras. If both the faces match, we can say that suspect is detected. Face detection of suspect is done accurately. It will display the message "Suspect is detected". If already stored image faces views and face captured by our camera doesn't match, then message will display "Suspect is not detected".

VI. EXPERIMENTAL RESULTS

The proposed method is experimented using MATLAB. It has performed mainly in two phases.

Face visualization, Feature extraction and Classifier algorithm. Training set images are converted into feature cells by taking into consideration face features of each image. To generate training images dataset, we capture images in different views. After conversion of feature cell using HOG techniques following results are obtained. Simulation is done by using MATLAB.

Comparison of real time captured image and training images is done and then face identification of person is carried out.

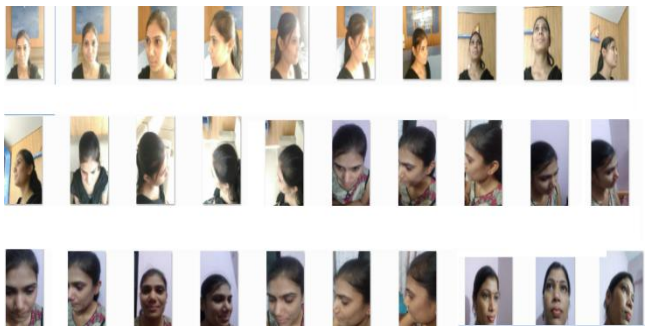

Fig 7: Multiple Images to Create Training Images Samples

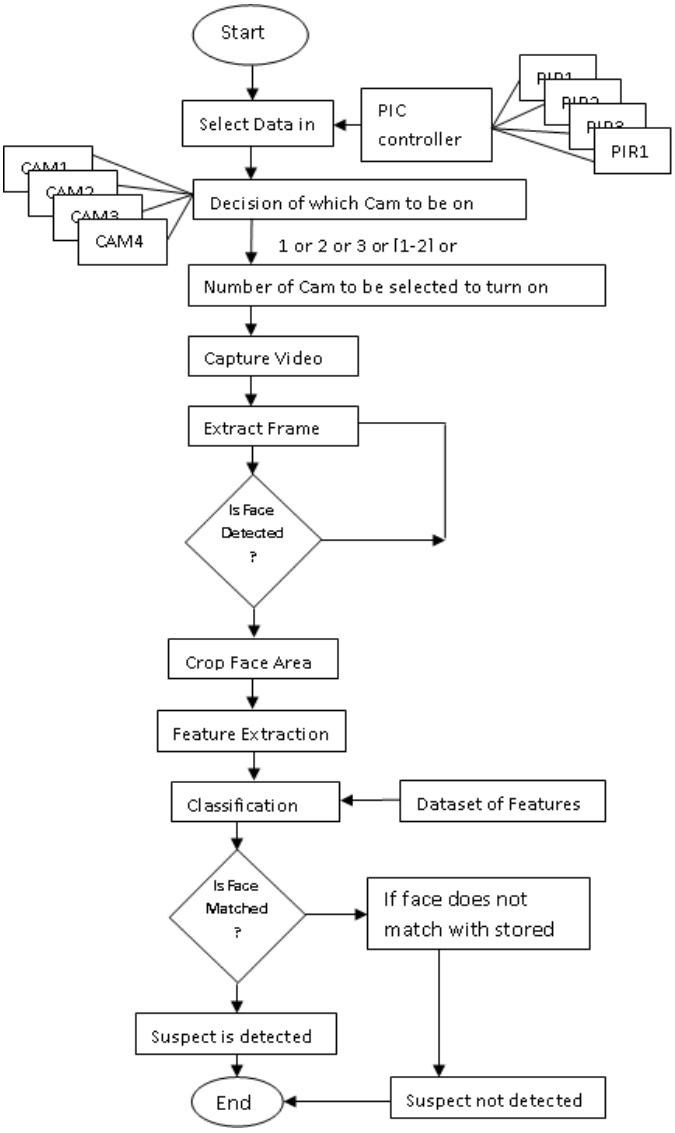

Fig 6: Flowchart of Proposed System

A. Initialization

After running the MATLAB code, cameras will initialize and editor window display ready command.

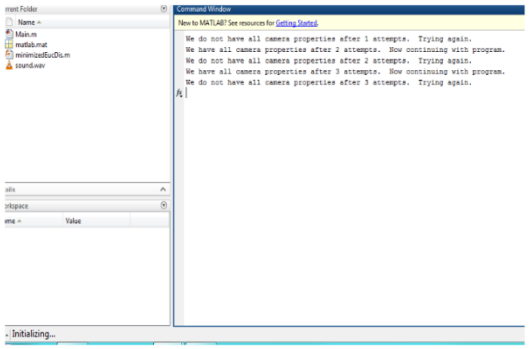

Fig. 8: Initialization of Cameras

B. Be Ready

After 3 attempts of computing the program, busy command is displayed at left bottom corner.

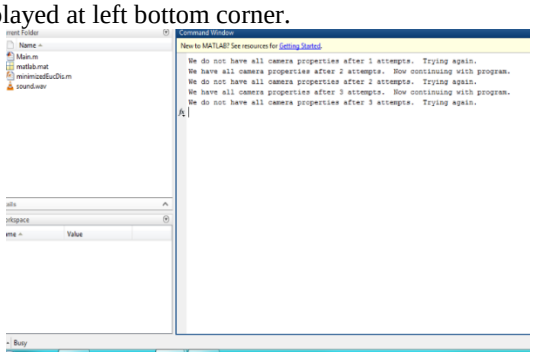

Fig 9: Be ready screen to Capture Video

C. Video Capture

With the help of video player, video is captured. Then from captured video, data is compared with training data sets.

D: Crop

Extract/crop face from video. Then captured face is compared with training data set.

Fig.10: Cropped and Matched Face

E: Final Output Window

If cropped face is matched with stored training set, then the message will appear that means object is pointed out.

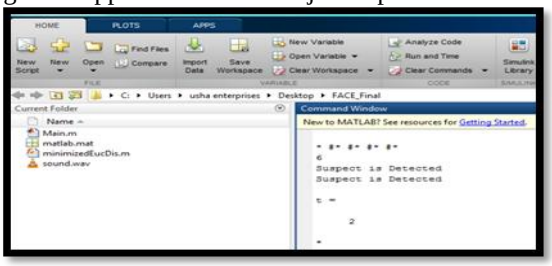

Fig.11: Output Window

F: False Reorganization

In false case, false reorganization is displayed then process exits with message displayed showing object is not pointed out.

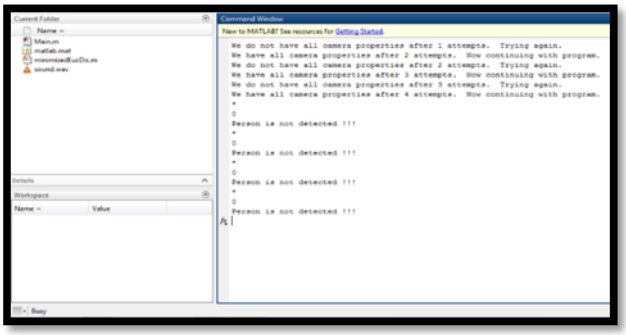

Fig12: False Recognition

VII. CONCLUSION

In this work, with combination of multiple sensors and cameras, we have done object tracking and face recognition. Human intervention not necessary and real time object tracking and face detection is done. It's easy to use with less complexity. Camera will only turn ON when object is sensed by particular sensor. This reduces requirement of large energy. Due to options provided, we can easily tracks objects. With the use of HOG, we can able to visualize a face in orientation from +90 to -90 and detection rate varies in few seconds depending on the reading of first video frame. Later it gives highest detection rate with the possibly accurate orientation (degree) of a face. Similar detection process may be used for incoming video with good detection rate of max 5ms delay. As object appears in FOV of particular camera then face of that object is detected and then it is matched against the face images (Wanted Criminals) available in database. However, using HOG descriptor, large database is required where the templates of the faces are needed to be stored for better accuracy. The entire system including of software as well as hardware can easily be turned into a completely wireless one. For better results, HD cameras can be used with high ranged sensors from long distances.

REFERENCES

1. Ka Ki Ng, "Background Subtraction and Object Tracking With Applications in Visual Surveillance", Purdue University, West Lafayette, Indiana, 1-134, December 2011.
2. T. Reeve, "How many cameras in the uk? Only 1.85 million, claims acpo lead on cctv," Web archive: http://www.securitynewsdesk.com/ 2011/03/01/how-many-cctv-camer-as-in- t he-uk/, Accessed: December 2019.
3. M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 4, pp. 657–662, April 2006.
4. Li and Chellappa, "A generic approach to simultaneous tracking and verification in video," IEEE Transactions on Image Processing, vol. 11, pp. 530–544, May 2002
5. A.J Goldstein, L. Harmon, A. Lesk, "Identification of Human Faces', Proceedings Of The IEEE,,Vol59, issue 5, pp. 748 -760,1971.
6. Sirovich and Kirby, "Low Dimensional Procedure for characteristics of Human Faces," Journal of Optical Society of America, vol. 4, pp 519-525,1987.
7. P.J. Phillips, W.T. Scruggs, A. J. O'Toole, P.J. Flynn, K.W. Bowyer, C.L. Schott, M. Sharpe, FRVT 2006 and ICE 2006 Large-Scale Results, 29 March 2007, NISTIR 7408, National Institute of Standards and Technology.
8. P.J. Phillips, P. Grother, R.J. Micheals, D.M. Blackburn, E. Tabassi, M. Bone, Face Recognition Vendor Test 2002.
9. Andrew W. Senior, Ruud M. Bolle, "Face Recognition and its Applications" Achievements and Challenges in Fingerprint Recognition, Chapter 04,101-115.
10. M. Valera and S. Velastin, "Intelligent distributed surveillance systems: a review," IEE Proc. of Vis. Image Signal Processing, Vol. 152, No. 2, 192–204 April 2005.
11. Ferryman, J.M., Maybank, S.J., and Worrall, A.D. 2000. Visual surveillance for moving vehicles. Int. J. Comp. Vis., 37(2):187–197.
12. R. Gonzalez and R. Woods – Digital image processing – second edition – Prentice Hall
13. L. G. Shapiro and R. M. Haralick – Computer and robot vision, volume 1 and 2, 1st edition – Prentice Hall.

AUTHORS PROFILE

Dr. Kantilal P Rane is presently working as a Professor and Principal in KCES's College of Engineering & Information Technology, Jalgaon. He has completed his PhD in 2013 from KBC North Maharashtra University, Jalgaon. He published 57 International papers and attended 24 National and International conferences. His area of research includes Networking, IoT, and Image Processing.