© 2018 IJSRST | Volume 4 | Issue 3 | Print ISSN : 2395-6011 | Online ISSN: 2395-602X

National Conference on Advances in Engineering and Applied Science (NCAEAS)
29th January 2018
Organized by : Anjuman College of Engineering and Technology (ACET) Nagpur, Maharashtra, India, In association with
International Journal of Scientific Research in Science and Technology

"A Review on Real Time Tracking and Face Recognition System"

Anam Darakhshan Saify Khan[1], Mohammad Nasiruddin[2]
[1]Student, Electronics and Communication, RTMNU, ACET, Nagpur, Maharashtra, India
[2]Associate Professor & HOD, Electronics and Communication, RTMNU, ACET, Nagpur, Maharashtra, India

ABSTRACT

This review summarizes the current state of research work made in the field of Real Time face recognition. In order to address all current face recognition researches, it include approaches that uses Blobs Acquisition and Segmentation as well as those uses Kalman Filter, Viola-Jones algorithm, Variation-ratio gain(VRG), Principle component analysis (PCA) Face recognition algorithm. This paper is a survey of various novel algorithm for different methods of feature extraction used by renowned personals. It gives a review on current research works where algorithm for face recognition typically extract facial features and compare them to a database to find the best match.

Keywords: Video Surveillance Cameras, Face Database, Face Recognition.

I. INTRODUCTION

Interest in face recognition, as a combination of pattern recognition and image analysis is still growing. Many papers are written and many real-world systems are being developed and distributed. As a non-invasive biometric method, face recognition is attractive for national security purposes as well as for smaller scale surveillance systems. However, in order to be able to claim that any face recognition system is efficient, robust and reliable, it must undergo rigorous testing and verification, preferably on real-world datasets.

The first thing need to create a face recognition system is a database of facial image of people that want to recognize also known as face gallery. Then perform a processing step known as feature extraction to store discriminative information about each face in compact feature vector. Following this we have to fit a model of the appearance of faces in the gallery so that we can determine between faces of different people in database the output of this stage is the classifier or a model that is used to recognize input images.

When we have input query image a face detection algorithm is used to find where the face is allocated in that image. We then crop, resize and normalize the face to match the size and pose of the image used in the training face gallery. Then performed the same feature extraction step that we did with the face gallery and run that through classifier or a model. The output is the label or an indicator to signify which person from the database the query image belongs to.

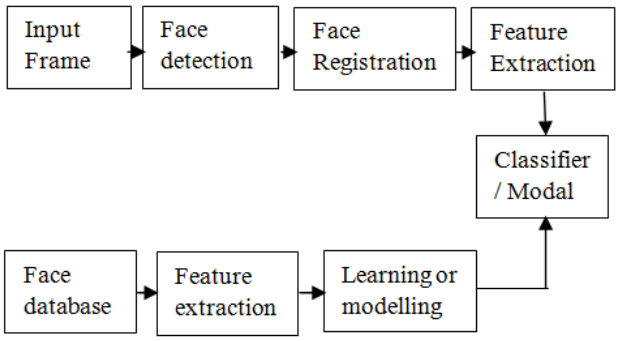

Figure 1. Face Recognition Workflow

II. VIDEO CAMERA SURVEILLANCE:

Real time people flow estimation can be very useful information for several applications like security or people management such as pedestrian traffic management or tourists flow estimation. The use of video camera to track and count people increase considerably in past few years due to the advancement of image processing algorithms and computers' technology. Several attempts have been made to track people but all those different ways can be classify in three categories of different complexity:

- ✓ Methods using region tracking features. To improve this methods some adding a classification scheme of pixel based on color or textures.
- ✓ Methods using 2D appearance of humans (using different models of humans)
- ✓ Methods using multiple cameras to make full 3D modelling.

The third category is more accurate than the two others because it rebuild precisely the scene (so it deals in a better way the occlusion problems) but it is also the most difficult with complex algorithms. Sometimes, this system required a complex camera set-up (calibration) and cannot operate in real-time because the 3D models are too slow. This is why most of the system used the other two categories.

In the paper [1], a novel approach is present to count number of people that pass through the view of an overhead mounted camera. Moving people are first detected as blobs and presented by binary masks, based on which possible multi-person blobs are further segmented into isolated persons according to their areas and locations. Each single person is tracked through consecutive frames using a correlation-based algorithm and a state diagram is proposed to count people entering and leaving the scene. Vision-based approaches are usually employed to accomplish this task, however, occlusion among people becomes a large obstacle when traffic is heavy. It solve the problem by placing a camera overhead to face downward, has been done by author Kim and Chen. Kim use convex hull approximation and velocity information to track and count people. Chen make use of HSI histogram information to track each person. They also use bounding-box algorithm to handle merge-splitting problems.

In this approach, blobs in current frame are firstly detected as groups of changed pixels using background subtraction. Define *New* blobs as blobs that just occur in current frame, and *Old* blobs as those also present in previous frame. Lost and Found algorithm is introduced to obtain the two types of blobs. As some of them may involve multiple persons, thus it propose a binary-level area-based segmentation scheme to ensure that each blob represents only one person. It then apply a correlation-based tracking method to establish correspondence between persons in two successive frames. Moreover, to facilitate bi-directional counting, the camera view is divided into three areas. Creative WebCam NX Ultra was used in our

experiments. It was placed 2.3 meters above the ground. The system has been tested on an Intel Pentium IV 2.4GHz PC with MatLab 6.5. The system achieved 100% accuracy at 3 fps. Neither false detection nor uncounted person events occurred in our experiment. One limitation of our system is that when more than four people join into a blob they are hard to be properly segmented.

In paper [2], a system is described a system for real time tracking of people in the video sequences. The input to the image is live or recorded video data acquired by a stationary camera in an environment where the primary moving objects are people. The output consist of trajectories which gives the Spatio-temporal coordinates of individual.

A system is developed for the interpretation of moving light displays (MLDs). Here, the objective is to segment out points from MLDs images of moving people into sets of corresponding to individual people and register different points on a person with different body parts. Individuals points are tracked across frames and points in each frame are clustered into individual objects based on their position and velocity. The motion field in each frame is obtained by matching intensity features in successive frames. The motion field is smoothed both temporally and spatially and then split into regions having the same quantized direction of motion. The posture of the person in each frame is estimated by matching gray scale edge the image with model edges. Figure shows main steps in people tracking system.

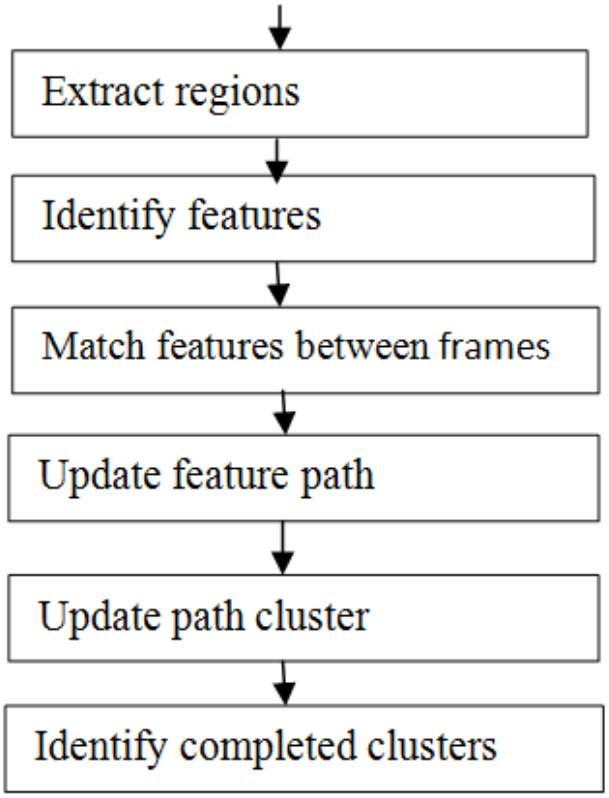

Figure 2. Steps in people tracking system

An application for counting people through camera is discussed in paper [3]; it performs the count distinction between input and output of people moving through the supervised area. The counter requires two steps: detection and tracking. The detection is based on finding people's heads through pre-processed image correlation with several circular patterns. Tracking is made through the application of a Kalman filter to determine the trajectory of the candidates. Finally, the system updates the counters based on the direction of the trajectories.

Here people counting is performed by extracting an appearance vector based on a color region of interest (ROI) and a probabilistic model, using the stereoscopic disparity map to resolve possible uncertainties. Two ROIs are defined at the top and bottom of the image. Next, the column histogram of

the optical flow is computed in those areas. The number of people crossing the area is obtained from the histogram values considering a minimum threshold. The count is obtained from the information of blobs crossing the ROIs.

For proper operation of system, a background image is dynamically updated to slowly capture small changes in scene illumination and introduce new static object. The background modelling has been divided into two parts; a comparison phase to obtain min/max value of each image pixel and the update phase. In comparison phase, the original image is compared pixel by pixel with one image containing the max value I_{max} and another with min value $I_{min.}$ If original image pixel i.e. I(m,n) is greater then corresponding pixel from $I_{max}(m,n)$ then pixel value of maximum image is modified with the image value see Eq.1 and Eq. 2 respectively.

$$vm(m, n) = [I\max(m, n) - I\min(m, n)]/2 \quad (1)$$

$$\text{Backupdate}(m, n) = \alpha \text{Back}(m, n) + (1-\alpha)\, vm(m, n) \quad (2)$$

Where parameter α determines the influence of the previous background value Back(m, n) and a new value $vm(m, n)$.

In the case of people tracking, the capture rate of the camera (in frames per second) should be sufficient to capture individually movements a number of consecutive times so that the estimation process may be consider linear in time. The aim of this technique is to obtain a good model to follow an object at each instant of time through an analysis of state variables.

Kalman filter used to estimate the orientation of the tool and the position of the center of the tool. Use of Kalman filter is combined with the information from a 2D stochastic model in order to identify the shape of person within image. The Kalman filter is a recursive procedure consisting of two main stages; prediction and correction. The first stage aim to estimate the motion, while the second is responsible for correcting the error in the motion. The key point is the feedback between the cracking and detection stages, allowing a more robust algorithm to be achieved, resolving temporal errors and partial occlusion that could occur in real image sequences such as the test videos used.

R. Venkatesh and A. Balaji Ganesh proposed research work and developed a real time system on moving object and face recognition using MATLAB. This paper [4]; implements a method to track and recognize the object in a surveillance area. They analyze usual pixel-approach. Camera system (webcam) acts as a sensor to track the object in surveillance area. Edge detection is an image segmentation process is implemented to have clear knowledge on real edges of real time video. Background separation algorithm provides clear knowledge about foreground and background. Video pre-processing such as frame separation, thresholding, binary operation, histogram equalization and edge detection of traffic video is done to track multiple objects and recognize it. Stepper motor may be used to orient the camera to any position to track and recognize object in surveillance. Contour let transform is used for feature extraction to recognize the object in a surveillance area and pattern matching also play an important role to recognize different objects in a video.

Two or more frames acquired at different time contain the information about relative motion between an imaging system and a scene. Therefore, the information about motions can be obtained through analysis and processing of frames acquired

at different time. Video sequence analysis methods can be classified into three methods: optical flow method, background difference method and adjacent frame difference method. Optical flow method reflects the frame variation caused by motion in a definite time interval. The motion field of frames is estimated to incorporate similar motion vectors into moving object solving transcendental equations is required in optical flow method. Background difference method is a technique for detecting the motion area by making the difference between the current frame and the background frame. An image is divided into foreground and background in this method. The background is modelled, and the current frame and the background model are compared pixel by pixel. Those pixels accordance with the background model are labeled as the background, while others are labeled as the foreground. In adjacent frame difference method, moving objects are extracted according to the differences among two or three continuous frames.

In order to acquire image using MATLAB, a video input object which represents the connection between MATLAB and the image acquisition device (in this case camera), must be created first. Video analysis normally requires certain video processing algorithm to prepare the image for further analysis. For this work, the video is analysed as it is viewed. To analyze color frames, it is necessary to first segment it. Edge detection is one of the most commonly used image segmentation methods in object detection. Since edges contain some of the most useful information in an image can be used to extract boundaries of each different object in an image.

Many edge detection algorithm have been developed which include Sobel, Prewitt, Roberts, Laplacian of a Gaussian, zero crossing and canny edge detectors, was used because of the fact that it is less susceptible to noises in comparison to other edge detection methods. The algorithm can be summarized as follows;

1. First the image was smoothed using a Gaussian filter with a specified standard deviation in order to reduce noises
2. Standard first-order edge detection was performed to find the edge locations and edge directions. An edge point will be the point whose strength is locally maximum in the direction of the gradient.
3. After this, non-maximum suppression will be applied. Non –maximum suppression is used to trace along the edge in the edge direction and suppress any pixel values (set to 0) that are not considered to be an edge. This will give a thin line in the output image.
4. Finally, Hysteresis thresholding was performed using two threshold values, T1 and T2 with T1<T2. Any result pixel with a value greater than T2 is categorized as a strong edge pixel, whist result pixel with value between the two threshold and adjacent to the strong edge pixel will be considered as weak edge pixel.

A primary objective of paper [5]; is the field trialing and ongoing development of a system for the robust detection and identification of persons of interest in a crowd. These people will often have non-frontal facial presentation, be photographed under various lighting conditions, and will exhibit natural expressions such images are typically acquired from CCTV cameras in public spaces as the subjects are not usually aware of camera placement. Other capabilities that are being trialed and developed include

1) Robust detection of background changes,
2) Tracking and identification of people by their appearance across multiple cameras,
3) Detecting suspicious events such as left luggage or the dangerous behavior of people,
4) Video summarization to produce brief video summaries of activity.

In this paper, the configuration of the trial system and some early results from commercial and NICTA research systems is presented. It also discuss the implementation and scalability challenges, as well as issues related to on-going real life trials in public spaces using existing surveillance hardware. The main capabilities that are currently offered by leading Intelligent Surveillance software vendors are demonstrated. Technology gaps are identified and opportunities for computer vision and pattern recognition research in the field of ICCTV are discussed. The main stages of processing in an intelligent visual surveillance system are: moving object detection and recognition, tracking, behavioral analysis and retrieval. These stages involve the topics of machine vision, pattern analysis, artificial intelligence and data management.

III. FACE RECOGNITION:

Face detection algorithms are usually divided into two general categories: (i) feature-based and (ii) learning-based methods. The algorithms from the first category are based on the assumption that face in the image can be detected based on some simple features, independent of ambient light, face rotation and pose. Thus, a simple method uses image projection to detect faces under the assumption that the background is uniform and with the vertical projection of the grey level image is determined the face position. Another feature-based face detection approach is based on a skin colour model determined by using the probability distribution in a colour space. The face is detected in image by applying a threshold on the modelled distribution. The algorithms from the second category are more robust but they need a greater computational effort. Learning-based methods use a number of training samples and benefit from statistical models and machine learning algorithms.

The detection of faces in an image is a subject often studied in computer vision literature. The algorithm which allowed face detection, imposing new standards in this area, was the Viola – Jones algorithm. In the paper [7], a practical implementation of a face detector based on Viola-Jones algorithm using Matlab cascade object detector is presented. Employing the system type object *vision Cascade Object Detector*, eight face detectors were developed using the *train Cascade Object Detector* function and tuning the number of cascade layer and the False Alarm Rate. For different tuning parameters, the performances of the face detectors were analysed.

The Viola – Jones algorithm is intended for real – time detection of faces from an image. Its real – time performance is obtained by using Haar type features, computed rapidly by using integral images, feature selection using the AdaBoost algorithm (Adaptive Boost) and face detection with attentional cascade.

The Viola-Jones face detector can run in real time because it is based on the following main ideas:
-Rapid computation of Haar-like features using the integral image;
- Classifier learning with AdaBoost to select the best feature;

- The attentional cascade structure which rejects the majority of the sub-windows in early layers of the detector, making the detection process extremely efficient.

Due to the simplicity of extracted features process and selection of the best features, Viola-Jones face detector is fast and robust, being reported many and various implementations for different applications.

Another paper written by Brian C. Lovell, Shaokang Chen and Ting Shan [8]; defines main approaches for face detection as feature based, image based, and template matching.

Feature based approaches attempt to utilize some priori knowledge of human face characteristics and detect those representative features such as edges, texture, colour or motion. It is shown in Fig. 3 where feature based, template matching & image based are main Face Detection Techniques. Edge detection is a necessary first step for edge representation. Two edge operators that are commonly used are the Sobel Operator and Marr-Hildreth operator. Edge features can be easily detected with a very short time but are not robust for face detection in complex environments. While texture based approach propose by detecting local facial features such as pupils, lips and eyebrows based on an observation that they are normally darker than the regions around them. Color feature based face detection is derived from the fact that the skin color of different humans (even from different races) cluster very closely. Various face detection techniques are present here:

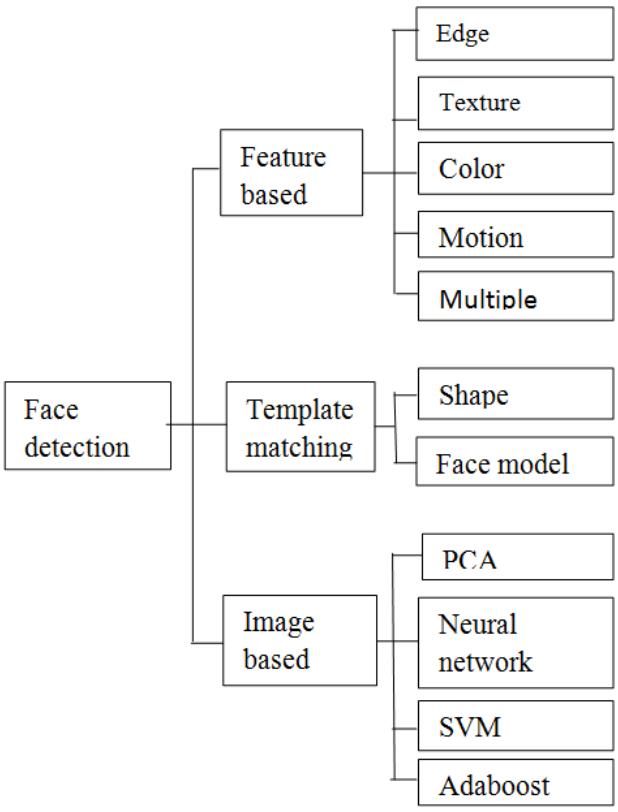

Figure 3. Face Detection Techniques

The Template matching approach is further divided into two classes: Feature searching and Face models. Feature searching techniques first detect the prominent facial features, such as eyes, nose, mouth, and then use knowledge of face geometry to verify the existence of a face by searching for less prominent facial features. Image-based approaches treat face detection as a two class pattern recognition problem and avoid using *a priori* face knowledge. It uses positive and negative samples to train a face/non-face classifier.

In summary, there are many varieties of face detection methods and to choose a suitable method is heavily application dependent. Figure 2 shows various face detection techniques and their categories. Generally speaking, feature-based

methods are often used in real-time systems when color, motion, or texture information is available. Template matching and image-based approach can attain superior detection performance than feature-based method, but most of the algorithms are computationally expensive and are difficult to apply in a real-time system.

IV. APPLICATION AND FUTURE SCOPE

Among the different biometric techniques, face recognition may not be most reliable and efficient. However one key advantage is that it does not require the cooperation of the test subject to work. Properly designed system installed in airport, multiplexes and other public places can identify individual among the crowd, without passerby even being aware of the system. Automated face recognition can be applied 'live' to search for a watch-list of 'interesting' people, or after the fact using surveillance footage of a crime to search through a database of suspects. Other biometric like fingerprints, iris scan and speech recognition cannot perform this kind of identification.

V. CONCLUSION

In this paper, we have given an introductory survey for the face recognition technology. Face recognition is a both challenging and important recognition technique. Among all the biometric techniques, face recognition approach possesses one great advantage, which is its user-friendliness (or non-intrusiveness). Different video camera surveillance and face recognition techniques have been studied to make system efficient and work on real time with more efficient result.

VI. REFERENCES

[1]. Hartono Septian, Ji Tao, and Yap-Peng Tan, "People counting by video segmentation and tracking",vol-1, June 2006, IEEE

[2]. Jakub Segen and Sarma (Gopal) Pingali "A Camera-Based System for Tracking People in Real Time" , Bell Laboratories, Lucent Technologies, Holmdel, NJ 07733,p-1015-4651, IEEE

[3]. Jorge Garcia, Alfredo Gardel, Ignacio Bravo, José Luis Lázaro, Miguel Martinez, and David Rodriguez "Directional People Counter Based on Head Tracking" IEEE transactions on Industrial electronics, vol. 60, no. 9, September 2013

[4]. R.Venkatesan and A.Balaji Ganesh, "Real Time Implementation on moving object Tracking and Recognition using MatLab", Velammal Engineering College, Chennai, India

[5]. P. Saraswathi, M. Prabha, M. Seema. "GSM based Smart Surveillance System using CCTV", vol-6, issue-06, june2016,VCET, Tamilnadu, India

[6]. Thiago Teixeira, Gershon Dublon, and Andreas Savvies', "A Survey of human sensing: methods for Detecting Presence, Count, Location, Track and Identity", ENALAB Technical Report 09-2010, Vol. 1, No. 1, September 2010

[7]. Elena Alionte Corneliu Lazar, Member, IEEE, "A Practical Implementation of Face Detection by using Matlab Cascade Object detector", 2015 19th International Conference on System Theory, Control and Computing (ICSTCC), October 14-16, Cheile Gradistei, Romania, 2015 IEEE.

[8]. Brian C. Lovell, Shaokang Chen and Ting Shan, "real time Face Detection and

Classification for ICCTV", Security and Survelliance Research Group, School of ITEE ,The University of Queensland, NICTA, Australia

[9]. Ankur Bansal, Mukesh Agarwal, Anima Sharma, Anindya Gupta, "A Review Paper on Facial Recognition", International Journal on Recent and Innovation Trends in Computing and Communication, vol 1, issue 4 , ISSN 2321-8169, APR 2013

[10]. Andrew W. Senior and Ruud M. Bolle, " Face Recognition and its Application", IBM T.J.Watson Research Center, Yorktown Heights, USA

[11]. Kui Jia and Shaogang Gong, "Multimodal Tensor Face for Simultaneous Super Resolution and Recognition", dept. of computer science, Queen Mary, University of London, London E1 4NS, UK, Tenth IEEE International Conference on Computer Vision, 1550-5499/05 © 2005 IEEE

[12]. Liyuan Li, Member, IEEE, Weimin Huang, Member, IEEE, Irene Yu-Hua Gu, Senior Member, IEEE, Ruijiang Luo, Member, IEEE, and Qi Tian, Senior Member, IEEE, " An Efficient technical approach to tracking multiple objects through Crowds for Real Time Intelligent CCTC Systems", IEEE Transactions on systems, man, and Cybernetics—part B: Cybernetics, vol. 38 no. 5, October 2008

[13]. Jae Young Choi, Yong Man Ro, Senior Member, IEEE, and Konstantinos N. (Kostas) Plataniotis, Senior Member, IEEE, "Color Face Recognition for degraded face image", IEEE transactions on systems, man, and Cybernetics—part B: Cybernetics, vol. 39, no. 5, October 2009

[14]. Ishita Gupta, Varsha Patil, Chaitali Kadam, Shreya Dumbre, "Face Detection and Recognition using Raspberry pie", 2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), 19-21 December 2016, AISSMS, Pune, India

[15]. Divyarajsinh N. Parmar, Brijesh B. Mehta, "Face Recognition methods and Application", Divyarajsinh N Parmar et al ,Int. J. Computer Technology & Applications, Vol. 4 (1),84-86 ISSN:2229-6093 Jan-Feb 2013

[16]. Matthew P. J. Ashby, "The Value of CCTV Surveillance Cameras as an Investigative Tool: An Empirical Analysis" Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK, Springer

[17]. Y M. Mustafah, A. W. Azman, A. Bigdeli, B. C. Lovell, "An Automated Face Recognition System for Intelligence Surveillance: Smart Camera Recognises Faces in Crowd", vol 1 , 2007 IEEE

[18]. Mislay Grgic, Kresimir Delac and Sonja Grgic, "SCface: Surveillance Camera Face Database" 30 October 2009, Springer science.

[19]. Anil K. Jain, Brendan Klare and Unsang Park, " Face Recognition: Some Challenges in Forensics", Michigan State University, East Lansing, MI, U.S.A

[20]. Zdenek Kalel, Krystina Mikolajczyk, Jiri Matas, "Face-TLD- Tracking-Learning-Detection Applied to Faces", Proceedings of 2010 IEEE 17th International Conference on Image Processing, September 26-29, 2010, Hong Kong, vol 1, ISSN 4244-7994 ©2010 IEEE

[21]. Prof. Mohd Nasiruddin, Faizan Ahmed, Ashar Quazi, Mubashara Mehrosh, Kahekashan Anjum," Computer Aided Design of Glaucoma Detection" International Journal of Engineering Science and Computing, March 2017, IJESC, Volume 7, Issue No.3