

# HUMAN FACE DETECTION AND TRACKING

**Vrushang G. Patel**

UG Student, Dept. of ICT,  
L J Institute of Engineering & Technology, Ahmedabad, Gujarat State

**Dr. Anil Suthar**

Department of EC Engineering,  
L J Institute of Engineering & Technology, Ahmedabad, Gujarat State

## ABSTRACT

*Now day's security at the border area, airports, and smart cities is important task. Many CCTV is available for monitoring but in present scenario monitoring as well tracking & identification are real challenges. So, security can be updated with Internet of Things. We proposed to develop a design such systems by using face recognition with smart E-mail alerts & tracking system. We use for recognition system using python programming on ARM processor. Whenever any unknown or unauthorized person is recognized which is not in listed database send the E-mail alerts with tracked location of the same to administrator. We achieved excellent performance efficiency of the system by calculating the Face detection rate.*

**Key words:** Internet of Things, Face Recognition, Python Programming.

**Cite this Article:** Vrushang G. Patel and Dr. Anil Suthar, Human Face Detection and Tracking. *International Journal of Computer Engineering & Technology*, 9(4), 2018, pp. 187–195.

<http://www.iaeme.com/IJCET/issues.asp?JType=IJCET&VType=9&IType=4>

---

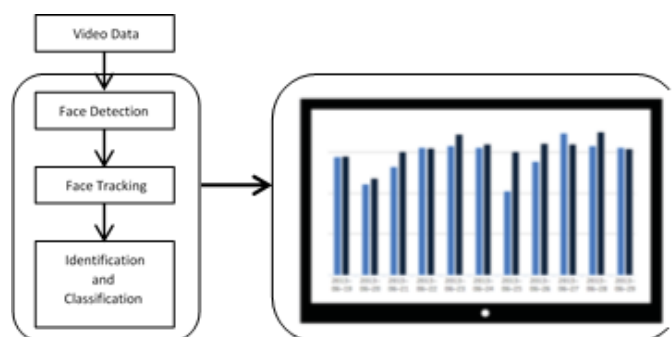
## 1. INTRODUCTION

Our brain having a unique feature that is can think only in images not in words. Once you may forget to keep your any things but you will never forget to bring a face with you. Everyone a unique face given by God, It is the most important part of our body. It can reflect many emotions of a human. Since long back, every human use nonliving thing for authentication and to get grant access in restricted areas like NASA, DRDO and ISRO etc. Basically biometric systems have two types as physiological characteristics (face, fingerprint, finger geometry, hand geometry, palm, iris, ear and voice) and behavioral characteristics (gait, signature and keystroke dynamics). Our behavioral may changes because of illness, fear, hunger etc. Now if we compare to biometrics systems with Face detection and recognition system, it is more expensive, exact, easy to understand and non-intrusive process. It has two categories as face detection (1:1) and face recognition (1: N). In the face detection we have to classify between face versus non face region while in recognition process we have to compare that single face image with multiple images from the input image.

Face recognition identifies persons on face images or video frames; it extracts features from an input face image and compares labeled faces in a database. This comparison is based on feature similarity metric and label of the most similar database entry is used to label the input image. In this case if the similarity value is below a certain threshold the input image is labeled as unknown.

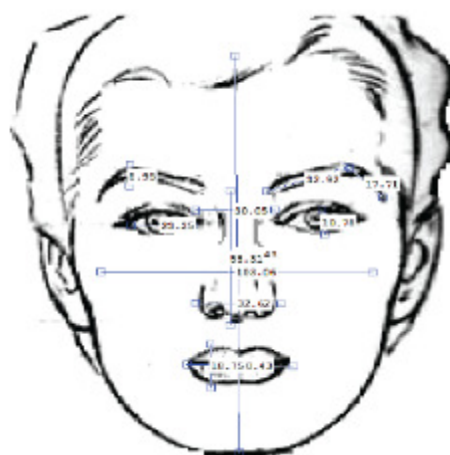
By using camera/web cam capturing an images have to come across some problems like pose (position of camera), presence of structural components (spectacles and beard), facial expression, occlusion (obstructed by someone), image orientation (variation in rotation), imaging condition (lightning and camera characteristics) etc. During 2006 a common feature like Face detection has been observed in digital cameras. In the year 2008 and 2010, automatic face detection and recognition system is placed in Europe and New Zealand respectively of security purposed for citizen.

In this paper we propose systems which extract identity of person and fining the director and location of unauthorized moving person from input video stream. Basic block diagram of real time person's classification system is presented in Figure. 1.



**Figure 1** Statics analysis of real time person classification

In this case inaccuracies at face position determination can lead to wrong decisions as well as quality of face detection step is critical to the final result of the entire system. The basic fundamental study which its mean idea is derived from Leonardo Davincci principle. For determining the geometry of the face, such distance as length and width of the face, length and width of the eyebrows, length and width of the eyes, distance between eyes, length and width of the nose, length and width of the mouth have been used. Figure 2 show face characteristics related to these distances.



**Figure 2** Face characteristics measurements

## 2. LITERATURE REVIEW

Face recognition method is based on information theory approach of coding and decoding the face image is discussed in [1]. In this method connection of two stages – Face detection using Haar Based Cascade classifier and recognition using Principle Component analysis. Various face detection and recognition methods have been evaluated [2] and also solution for image detection and recognition is proposed as an initial step for video surveillance. Implementation of face recognition using principal component analysis using 4 distance classifiers is proposed in [3]. A system that uses different distance measures for each image will perform better than a system that only uses one. The experiment show that PCA gave better results with Euclidian distance classifier and the squared Euclidian distance classifier than the City Block distance classifier, which gives better results than the squared Chebyshev distance classifier. A structural face construction and detection system is presented in [4]. System consists the different lightning, rotated facial image, skin color etc. [5]. The first approach applying AAM to age estimation, which extracts craniofacial growth and skin aging during childhood and adulthood. Different classifiers (including shortest-distance classifier, quadratic function and neural networks) are compared when AAM is employed as the feature representation. The approach also differentiated between 1) age-specific estimation, which is based on the assumption that the aging process is identical for everyone; and 2) appearance-specific estimation, which follows the assumption that people who look similar tend to have similar aging processes. Subsequently, a personalized age estimation used in the specialty of aging processes is then introduced to cluster similar faces before classification. In addition, [6] modeled the aging process with AAM based on a sequence of age-ascending face images for the same individual. Hence, different aging models can be learnt for different persons..

Applied principal component analysis (PCA) to represent each image as a feature vector in a low dimensional space; genetic algorithms (GA) were then employed to select a subset of features form the low dimensional representation that mostly encodes the gender information [7]. Four different classifiers were compared in this study: the Bayesian decision making, a neural network (NN), support vector machines (SVM) and a classifier based on linear discriminant analysis (LDA). [8] Focused on the edge information and exploited a neural network (NN) classifier for gender recognition. In particular, they computed the density histograms of the edge images, which were successively treated as input features for the NN. [9] Base their gender recognition system on a Gaussian Process Classifier (GPC). Facial images are first normalized to a standard dimensions and background and hair information was removed. Parameters for the GPC are learned using Expectation Maximization (EM) - Expectation Propagation (EP) algorithm. Finally GPC is used for classification.

## 3. BASIC METHODS FOR FACE RECOGNITION

In this section we understand about the basic system uses different techniques [10] for face recognition, age estimation, weight estimation and gender estimation namely;

- Viola Jones Algorithm:

The basic principal of algorithm is to detect the faces from the given input image. Before this there were so many images processing approach but all of them were time consuming due to making the entire image to the fix size and then run the image in the detector. Opposite of this is the viola Jones algorithm were the detector is rescale and whatever the size of image would be.

- Histogram of oriented gradients (HOG) Algorithm:

The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of

gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, scale-invariant feature transform descriptors, and shape contexts, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy. It use for gender estimation procedure.

- Local binary patterns (LBP):

LBP is one of the binary patterns which are used for feature extraction. In this the face image is firstly divided into small regions from which LBP features are extracted gives the output in histogram. LBP is used because there are micro patterns which are invariant of monotonic grey scale transformation. Combining all this gives the face image. LBP is widely used in much application due to its high tolerance against object recognition texture analysis and high discriminative power.

- Support vector machines:

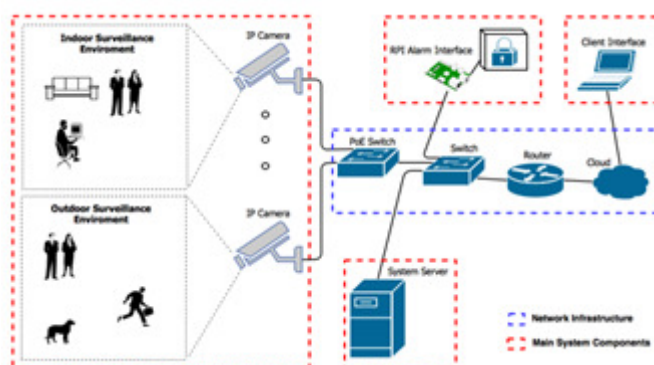
In machine learning, support vector machines (SVMs) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. In support vector machine is used to analyze the complex data and gives the result. SVM is very useful in finding patterns which are very useful and not complex.

- Convolution Neural Network:

In machine learning, a convolution neural network (CNN) is a type of feed-forward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. Individual cortical neurons respond to stimuli in a restricted region of space known as the receptive field. The receptive fields of different neurons partially overlap such that they tile the visual field. The response of an individual neuron to stimuli within its receptive field can be approximated mathematically by a convolution operation. Convolutional networks were inspired by biological processes and are variations of multilayer perceptron designed to use minimal amounts of preprocessing. They have wide applications in image and video recognition, recommender systems and natural language processing.

#### 4. OBJECT ORIENTED AND AGILE MODELS

We proposed system environment as per figure 3, in this model we added more than two IP/USB cameras mounted in different location and try to identify the person based on our proposed algorithm.



**Figure 3** Proposed Basic model of systems environment

#### 4.1. Working

Initially image is captured real time using the USB/IP camera. Open CV is used at face detection stage. Open CV (Open Source Computer Vision) is a library of programming functions mainly aimed at real-time computer vision. In simple language it is library used for Image Processing. It is mainly used to do all the operation related to Images. Now face detection and extraction algorithm will work i.e. viola Jones algorithm which uses Haar feature based cascade classifiers algorithm for face detection [11]. As long as a face is detected, a red bounding box is drawn on the face in the image.

#### 4.2. Methodology

Face Recognition is very important part of our research work. It is based on the Deep- Learning Concepts and DLIB implementation algorithm. It is able to give 95% accuracy on the one single photo of the user. It is able to recognize the people even after any changes of hair-style. There are many faces are stored in database for recognition. The algorithm generates the universal face encoding of that face which are stored in database [11]. Images has been identify with  $x^{(i)}$  and  $x^{(j)}$  of rows and columns, in our proposed work we use learning parameters of NN is define as encoding  $f(x^{(i)})$  is approximately 128. This algorithm matches that face encoding matrix to the live streaming face encoding features. We compute the distance  $d$  between their encoding  $f(x1)$  and  $f(x2)$ . If this function is less than threshold i.e. hyper parameter, it means that two images are same person, if not, they are two different persons.

$$d(x^{(1)}, x^{(2)}) = \|f(x^{(i)}) - f(x^{(j)})\|_2^2 \quad \text{eq.} \quad (1).$$

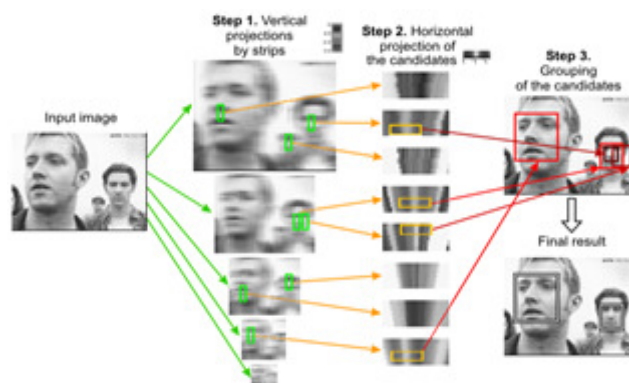
Distance function between two encoding of  $x1$  and  $x2$ , this is working for any two images  $x_i$  and  $x_j$ . If  $x^{(i)}$  and  $x^{(j)}$  are the same person then  $\|f(x^{(i)}) - f(x^{(j)})\|^2$  is small and if  $x^{(i)}$  and  $x^{(j)}$  are the different person then  $\|f(x^{(i)}) - f(x^{(j)})\|^2$  is large.

The encoding the universal face encoding means there are no another person's available whose face encoding matches with the each other. The calculation of loss functions is equal to

$$\sum_{i=1}^N [\|f(x^{(i)}) - f(x^{(j)})\|_2^2 - \|f(x^{(i)}) - f(x^{(j)})\|_2^2 + \alpha]_+ \quad \text{eq.} \quad (2)$$

#### 4.3. Procedure

We had several test variables that were varied in the experiments and thereby producing the 120 combinations each with unique test conditions. As per figure 4 if many person in group then start testing with many combinations.



**Figure 4** Experimental executions in group

At first, we trained the gender classifiers using automatically detected and, in some combinations, aligned faces. If an alignment method was used then it was used both during training and testing. Only successfully detected and aligned faces were used for training. We trained separate classifiers for each combination.

#### 4.4. Database of Images

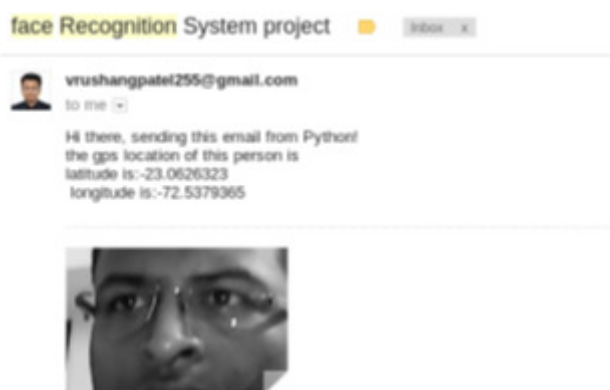
The Face recognition algorithm is working on concept of one-shot learning where the only one image of any person is enough for face recognition. The database of known people contain only one image of each person let's have a look. Here the database of known people is shown where it contain only photo of face in the gray scale or color image it does not matter as per shown figure 5.



**Figure 5** Database of trained persons

#### 4.5. E-mail Alerts

Also we developed smart surveillance system that is able to send E-mail alerts automatically when any person whose photo is not available in the database. The E-mail contains the photo of that person. E-mail is also containing information of latitude and longitude values so that user can get information regarding the geo-logical location of the person (figure 6).



**Figure 6** E-mail alerts indication

#### 4.6. Global Positioning System Tracking

The GPS Tracking is one of the most essential parts of the surveillance system. Along with this system model we implemented this feature which is able to give accurate location of the camera where the unknown person was recognized [12]. It is able to show the pointer Location in the satellite map so that the user can monitor the location of the unknown person. it can be very useful for army related surveillance systems. If any unknown person so which is not listed in



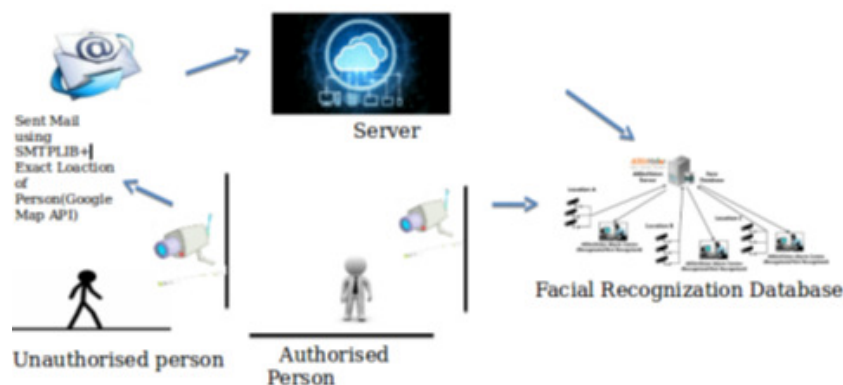
our system databases then gave accurate location in form of green pointer in Google-Map with all longitudes and latitudes. For reference purpose one result shown in figure 7. The Map has facility to auto zoom as well as switching views from map to satellite or vice-versa.



**Figure 7** Google map location of unknown identity

## 5. IMPLEMENTED FINAL MODEL

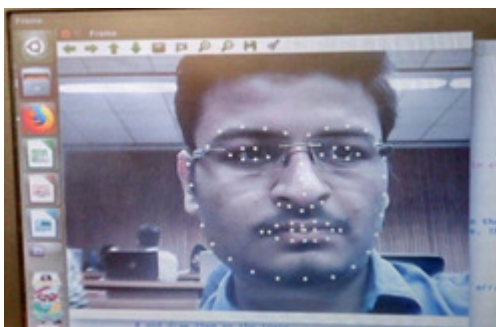
Here we show in figure 8 the working idea of the smart surveillance system if it will develop with server base database access and storage of millions of people's wide range face recognition [12]. The open source servers which are working on Linux base operating system can be used for the server base coding techniques. The person who is authorized will be able to cross any kind of security but person who has not any kind of access he/she will be identified in the system and will have to face security problems.



**Figure 7** Working idea of the smart surveillance system [12]

## 6. TESTING WITH VARIOUS SAMPLES

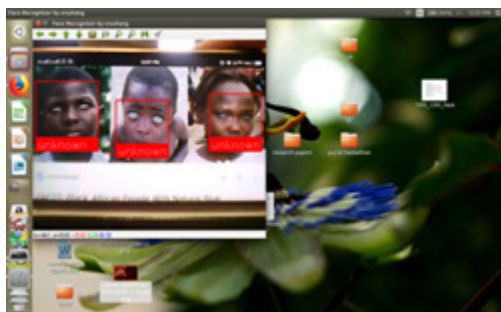
Figure 8.1 to 8.5 shows the various samples we tested and identified the persons.



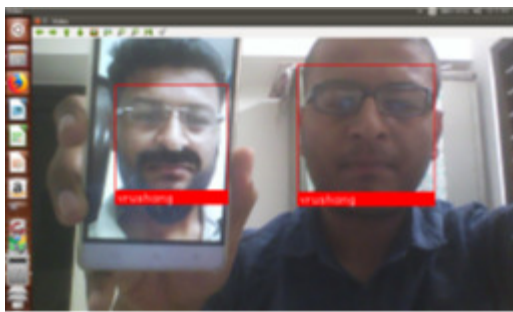
**Figure 8.1**



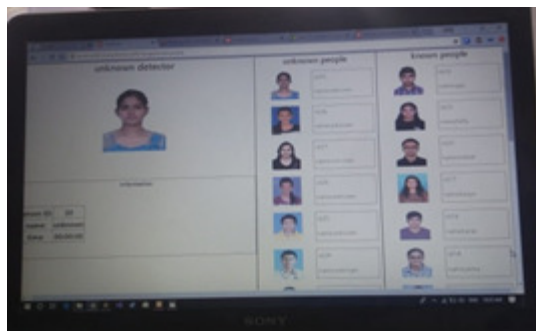
**Figure 8.2**



**Figure 8.3**



**Figure 8.4**



**Figure 8.5**

## 6. CONCLUSION

Face detection and recognition is currently a very active research area. Some of the more algorithms are still too computationally cheap to be applicable for real time processing. Face detection and tracking is being challenging for many researchers with real time Image sensor. With the advancement the real time face detection in remote monitoring is help for building much efficient application. Moreover such technology can be useful in tracking the lost object under dynamic environment. We tested more than 50 images with getting 98% of accurate detection of person. Further enhancement of this work can be extended with stereo depth analysis of face detection using two image sensor interfaced with High speed Processor. The future scope of this is to improve the database of public where the large public database is available.

## REFERENCES

- [1] Sarala A. Dabhade & Mrunal S. Bewoor (2012), "Real Time Face Detection and Recognition using Haar - based Cascade Classifier and Principal Component Analysis", International Journal of Computer Science and Management Research, Vol. 1, No. 1.
- [2] Faizan Ahmad, Aaima Najam & Zeeshan Ahmed (2013), "Image-based Face Detection and Recognition: State of the Art", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue. 6, No. 1.
- [3] Hussein Rady (2011), "Face Recognition using Principle Component Analysis with Different Distance Classifiers", IJCSNS International Journal of Computer Science and Network Security, Vol. 11, No. 10, Pp. 134–144.
- [4] S. Sankarakumar, Dr.A. Kumaravel & Dr.S.R. Suresh (2013), "Face Detection through Fuzzy Grammar", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, No. 2.



- [5] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 442–455, Apr. 2002.
- [6] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2234–2240, Dec. 2007.
- [7] Sun Z., Bebis G., Yuan X. and Louis S.J., "Genetic feature subset selection for gender classification: a comparison study", in *IEEE Proceedings on Applications of Computer Vision*, pag. 165-170, 2002.
- [8] Nakano M., Yasukata F. and Fukumi M., "Age and gender classification from face images using neural networks", in *Signal and Image Processing*, 2004.
- [9] Kim H.-C., Kim D., Ghahramani Z. and Bang S.Y., "Appearance based gender classification with Gaussian processes", in *Pattern Recognition Letters*, vol. 27, iss. 6, pag. 618-626, April 2006.
- [10] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. IEEE*, vol. 83, pp. 705–740, 1995.
- [11] Lucas D.Introna, H.Nissenbaum.: "Facial Recognition Technology, A survey of policy and implementation Issues", CCPR.
- [12] R.Chandana, Dr.S.A.K.Jilani, Mr.S.Javeed Hussain, "Smart Surveillance System using Thing Speak and Raspberry Pi", *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 4, Issue 7, July 2015.