International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 3 Issue 4, April 2014

Face Detection in Real Time Based on HOG

Rekha N, Dr.M.Z.Kurian

***Abstract*— Naturally before recognizing a face, it must be located in the image. In some cooperative systems, face detection is obviated by constraining the user. Most systems use a combination of skin-tone and face texture to determine the location of a face and use an image pyramid to allow faces of varying sizes to be detected. Increasingly, systems are being developed to detect faces that are not full frontal. This paper proposes a design to detect faces in real time using hog descriptor. Taking HOG of the image and calculate the weights which contributes for the facial features. Positive weights with facial features like eyes, nose, and mouth are marked completely to visualize a face. This algorithm continuously detects the face from $+90^0$ to -90^0 rotations even for occluded faces with high detection rate.**

***Index Terms*— Face detection, Face recognition, Facial feature, HOG descriptor, Face visualization.**

I. Introduction

Human has special skill in analyzing and interpreting faces, and so face analyzing has an important role in man machine relationship and different research areas has been opened in this way. Face detection is considered to be the first task performed while processing scenes for varied purposes and its results are important for subsequent steps of automated human face recognition. Therefore the whole process should work predictably and quite reliably. The increased need of security in the country, both in Unconstrained and constrained areas and the rapid developments in the field of computer vision have paved much progress in face detection and face recognition system. These systems designs can be employed in surveillance and monitoring, biometrics, traffic assistance, health care, etc. Face detection is differentiating the face from any other objects (inter-class variability). Face recognition is differentiating one's face from the other (intra-class variability). Face detection and recognition poses challenging task as detecting any other objects and face has various facial features and color varies dynamically. Processing of a face in real time with occlusions, background structure and camera position adds to the existing challenges.

Manuscript received April, 2014.

***Rekha N** is currently pursuing M.Tech Digital Electronics in dept of E&C from Sri Siddhartha Institute of Technology, Tumkur, India.*

***Dr.M.Z.Kurian**, is currently working as a Dean & HOD in dept of E&C at Sri Siddhartha Institute of Technology, Tumkur. India.*

Face detection involves segmentation, extraction and verification of facial features from an unknown background. Face detection involves two approaches feature based where face knowledge is implicit into the system through mapping and training schemes and image based approach without feature derivation and analysis. Image approaches are mostly done by window scanning technique for detecting faces. A human detection technique involves two types part based and sub-window based approaches. Sub-window based approaches can be based on different types and combinations of features, such as histograms of oriented gradients (HOG) [1], covariance matrices and multi-level versions of HOG [2]. Part-based approaches split the body into several parts that are detected separately and, finally, the results are combined. The Haar based face detection system for person detection and implementation of face visualization models using HOG descriptor for frontal and non frontal faces has been implemented in this work for real time application. The motivation for using Haar face detection and HOG descriptor is that it is an easily trainable system for any object.

The organization of this paper is as follows. Section II presents a related work. Section III describes face detection based on haar classifier section IV describes face visualization models based on HOG section V describes advantages section VI applications and section VII concludes the paper.

II. RELATED WORK

There is a great diversity in the way facial appearance is interpreted for recognition by an automatic system. Currently a number of different systems are under development, and which is most appropriate may depend on the application domain. A major difference in approaches is whether to represent the appearance of the face, or the geometry. Brunelli and Poggio [3] have compared these two approaches, but ultimately most systems today use a combination of both appearance and geometry. Geometry is difficult to measure with any accuracy, particularly from a single still image, but provides more robustness against disguises and aging. Appearance information is readily obtained from a face image, but is more subject to superficial variation, particularly from pose and expression changes. In practice for most purposes, even appearance-based systems must estimate some geometrical parameters in order to derive a shape free representation that is independent of expression and pose artifacts. This is achieved by finding facial landmarks and warping the face to a canonical neutral pose and expression. Facial features are also important for geometric approaches and for anchoring local representations [4].

Moving object detection and tracking algorithm play an important role in the intelligent video surveillance system. This method combines the interframe difference method with background subtraction which makes use of color, texture information and dual threshold to detect moving targets. This method includes background subtraction and optical flow methods. Background subtraction method to detect moving object process involves first setting a threshold value for the variation in gray pixel and second threshold value is set for the moving target in whole image. If two threshold values met then the detection of target is achieved with some adaptability to change in background.

Optical flow method uses the instantaneous velocity which is generated in the continuous movements of the pixels in the moving object. This algorithm has poor anti-noise performance and calculation is complicated with large hardware setup. Interframe subtraction used to find difference between two adjacent frames to detect moving target. It has poor connectivity and having strong environmental adaptability to some extent [5].

Facial features classification using neural network is learning process with the neural network which compares the whole image to find the location of the face. A retinal connected neural network examines small windows of an image, and decides whether each window contains a face. The system arbitrates between multiple networks to improve performance over a single network. It first applies a set of neural network-based filters to an image, and then uses an arbitrator to combine the outputs. Filter receives an input of 20x20 pixel region of the image, and generates an output ranging from 1 to -1, signifying the presence or absence of a face. The filters examine each location in the image at several scales, looking for locations that might contain a face. The arbitrator then merges detections from individual filters and eliminates overlapping detections. The filtering algorithm has a preprocessing step, applied to a window of the image. Histogram equalization is performed, which non-linearly maps the intensity values to expand the range of intensities in the window passed through a neural network, which decides whether the window contains a face [6].

Face Detection & Smiling Face Identification Using Adaboost & Neural Network Classifier has two popular methods of detecting faces are presented [7], Adaboost and ANN, analyzing, evaluating ones' advantages and disadvantages. From the study, Adaboost (cascade of boosted Adaboost) has the fastest performance time; however the correctness rate is not high (because detection results depend on weak a) classifiers or Haar-like features); and it is proved by the experiments. Also boosting of intensity difference between pixels in the grey scale face images are used as simple feature for smile detection.

An intermediate system using a boosting algorithm and neural net-work to train a classifier which is capable of processing images rapidly while having high detection rates, Adaboost is an aggressive learning algorithm which produces a strong classifier by choosing visual features in a family of simple classifiers and combining them linearly. This is the model of combining AB and ANN for detecting faces. In this model, ABs have a role to quickly reject non face images; then ANNs continue filtering false negative images to achieve better results. The final result is face/non face [8].

Face detection using bag of facial features implemented and tested face detection method by using the SVM classifier. Elastic Bunch Graph Matching Method to extract the most important parts in the face (eyes, nose, etc) and from them we can obtain HOG descriptors without using the entire image, by reducing number of operations. In this paper, we explore the representational power of HOG descriptors for face detection with Bag of features [9]. We propose a simple but powerful approach to detect faces:(1) extract HOG descriptors using a regular grid, (2) vector quantization into different code words each descriptor, (3) apply a support vector machine to learn a model for classifying an image as face or non-face based on codeword histograms [10].

Real Time Face Detection and Recognition using Haar Based Cascade Classifier and Principal Component Analysis The experiment showed that using proposed face detection is performed by using Haar based cascade classifier, Combination of Log Gabor features and sliding window based feature selection method, Principal Component Analysis and Euclidean-Based distance measure we can achieve very high recognition accuracy(74-79%) and low equal rates(0.3 to o.4% Equal error rates) using real time database. In the future we are going to investigate the possibilities of using decomposed Log-Gabor feature vectors and multiple PCA spaces in order to have the possibility of using this method with an unlimited number of training images. Because the results of all compared methods showed that the accuracy of face recognition is very affected by the lighting conditions, in the future we are going to investigate different lighting normalization methods and test them with the Log-Gabor PCA face recognition method [11].

An Efficient Face Detection and Recognition System in this work, an approach for face detection and recognition system based on Haar wavelet and BICA is developed. For the face detection system, Haar based features capture the structural properties of the object and invariant to illumination, skin color and slight occlusions. The statistical approach of BICA partitions the image into sub-blocks and thus further reduces the dimensionality than the traditional ICA. The developed face detection and recognition system performs well and provide good recognition rate. This paper proposes an integrated system for effective face detection and recognition by combining Haar based face detection with BICA. This proposed scheme provides an efficient solution for tailgating problem. Also the proposed scheme detects the face, recognize the face, checks for the availability of similar face in the database [12].

Spoofing is the real concern with regard to the security of the biometric system. In this paper they have discussed face liveness detection, in which spoofing can be controlled in a smart way. In this paper, a novel approach of liveness detection based on skin elasticity is proposed. In which a set of face images is used for liveness detection based on their correlation coefficient and their discriminant analysis. In this method, firstly user is asked to do some simple mouth movement activities like chewing, forehead movement etc simultaneously. At the same time a set of face images is captured. Then after applying pre-processing techniques, feature extraction is done using correlation coefficient and image extension feature. Using some discriminant analysis method, images are discriminate and skin elasticity is calculated. Then output is compared with the stored database. If output is less than the stored value then image captured is a fake image else it is a real image. Since age factor plays a significant role in skin elasticity so threshold value can be according to ages. On the other hand, age can be used as a soft biometric for classifying the face database. Since user intervention is less it will provide a very good non intrusive solution against fake faces. On other way it is a user friendly method, so its acceptance criteria among users will be more. Also as it is a software base liveness detection method so it will be easily applicable on all pre-established face database [13].

III. FACE DETECTION BASED ON HAAR CLASSIFIER

A. Face detection

Given an arbitrary image, the goal of face detection is to determine whether or not there are any faces in the image and, if present, return the image location and extent of the face.

Face detection scans an image pyramid to detect faces regardless of scale and location, and uses a filtering hierarchy procedure to filter out locations that do not represent faces with successively more accurate face classifiers.

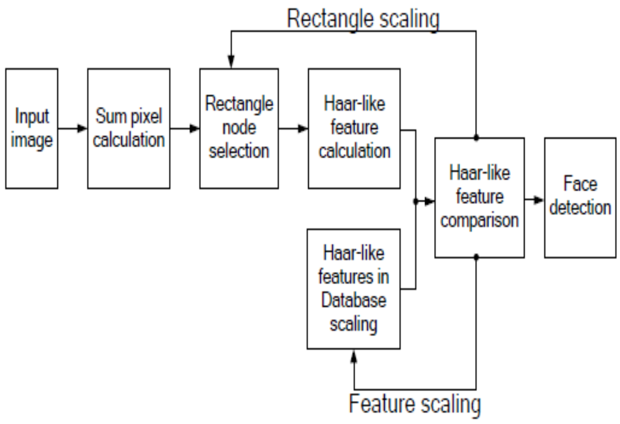


Fig 1: Flow diagram of face detection system

Face detection is the first step of any fully automatic system that analyzes the information contained in faces. It is mainly concerned with detecting faces in a given image. Numerous representations have been proposed for face detection, including pixel-based, parts-based, local edge features, haar wavelets and haar-like features. The system with haar-like features have high accuracy to detect the faces in different poses is shown in fig 1.

B. Haar classifier

There are three simple rectangular haar like features are used as shown in fig 2. The value of a two rectangle filter is the difference between the sums of the pixels within two rectangular regions. A three rectangle computes the difference between the sums of two outside rectangle with the center rectangle.

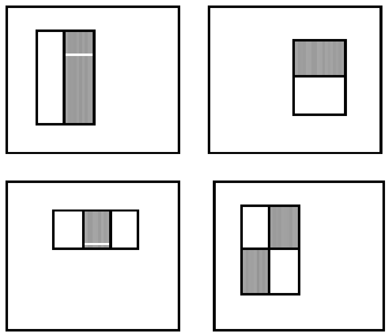


Fig 2: Haar like features used for face detection

A four rectangle filter computes the difference between the diagonal rectangle pairs. Each rectangle feature is binary threshold function constructed from a threshold, and a rectangle filter which is linear function of the image. Applying the haar feature basis to an input window and evaluated by passing it over several classifiers. Simple rectangular features are represented by integral image and features are calculated as shown in equation 1.

$$Feature = w1 \times RecSum(r1) + w2 \times RecSum(r2) \quad (1)$$

A faster computation of Haar features is possible using an intermediate representation for the image which we call the integral image. The integral image at location x, y contains the sum of the pixels above and to the left of x, y, inclusive is described in equation 2:

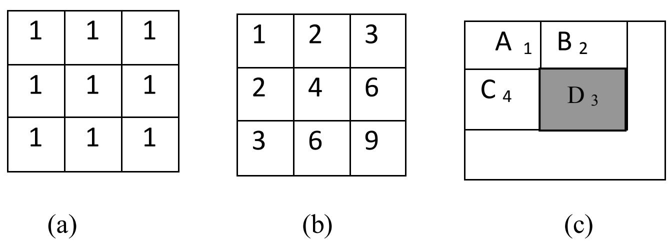


Figure 3: (a) the input image (b) the computation of integral image for the given input image (c) the computation of rectangular block from the integral image.

$$Integral(x, y) = \sum_{x' \leq x \;\; y' \leq y} Image\,(x', y') \quad (2)$$

Where *Integral(x, y)* is a pixel of the integral image at (x, y) and image (x', y') is a pixel of the image at (x', y'). Using the integral image any rectangular sum can be computed in four array references as shown in Figure 3(b). For computing the rectangular feature of block D, the value is given by equation3.

*SumD=*Integral3*(x, y)-Integral4(x, y)-Integral2(x,y)+ Integral1(x, y)* *(3)*

The Haar features are extracted using detector windows of various sizes. The cascading of classifier allows highest probability sub images to analyze for all haar features to distinguish an object is shown in fig 4.

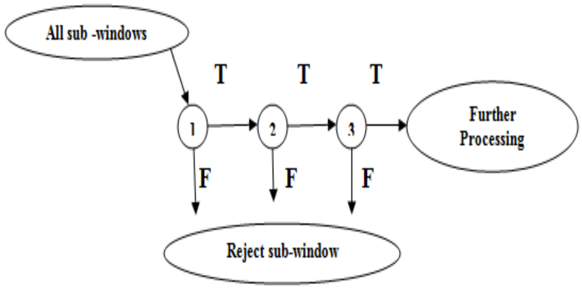


Fig 4: Cascaded haar classifier

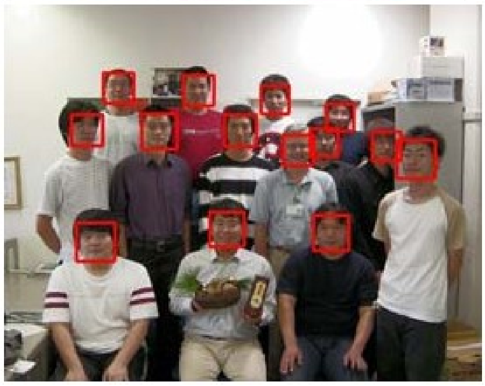
The number of features evaluated when scanning real images is necessarily a probabilistic process. Any given sub-window will progress down through the cascade, one classifier at a time, until it is decided that the window is negative or, in rare circumstances, the window succeeds in each test and is labeled positive. The expected behavior of this process is determined by the distribution of image windows in a typical test set. The key measure of each classifier is its "positive rate", the proportion of windows which are labeled as potentially containing the object of interest. The expected number of features which are evaluated in equation 4 is:

$$N = n_0 + \sum_{i=1}^{K}\left(n_i \prod_{j<i} p_j\right) \tag{4}$$

Where N is the expected number of features evaluated, K is the number of classifiers, Pi is the positive rate of the i^{th} classifier, and n_i are the number of features in the i^{th} classifier.

This cascaded classifier evaluates the input image and check for the presence of a rectangular window in a given image. Here each window is subdivided into several sub windows contributes for the presence of face or non face when passing it over several stages of classifiers. Each classifier evaluates the input window for the facial features which contributes for the face.

C. Results



IV. FACE VISUALIZATION MODELS BASED ON HOG

A sampling method is defined for every pixel and its neighbors are retrieved in a predefined manner to form a low level vector. The sampling is done for all trained images for every pixel in the image and therefore each pixel will have an associated with low level feature vector. HOG features are local descriptors, and human features are constituted by computing local direction of the gradient. Proposed descriptors could describe well the edge information of human; also the method is robust to illumination variations and small offset. The gradient of the pixel of (x, y) in an image can be denoted in equation 5 and 6 as:

$$G_x (x, y) = H(x+1, y) - H(x-1, y) \quad (5)$$
$$G_x (x, y) = H(x, y+1) - H(x, y-1) \quad (6)$$

Where $G_x (x, y)$ denotes the horizontal direction gradient of input image pixel, $G_y (x, y)$ denotes the vertical direction gradient and $H(x, y)$ denotes the pixel values. Then the gradient magnitude and direction of (x, y) can be represented in equation 7 as:

$$G(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)}$$
$$\alpha(x,y) = \tan^{-1}\left(\frac{G_y(x,y)}{G_x(x,y)}\right) \quad (7)$$

A vector quantization algorithm is used to construct the coding tree up to 5 levels; it can be extended to any number of levels if more features contributing for face. Each tree has several nodes all are connected to root node of a tree. An input image containing face, facial features will be marked up by HOG descriptor forming a feature facial cell. Each cell is grouped into blocks and blocks will be assigned with positive weights if gradient of an image contributes for the facial features. All these blocks are clustered together and corrected face visualization model is matched with the input face image. Like this so 13 models are generated using HOG which help us to visualize face in any direction from +90 degree to -90 degree. The resulting face visualization models generated using hog descriptor shown in figures from fig 5 – fig 9.

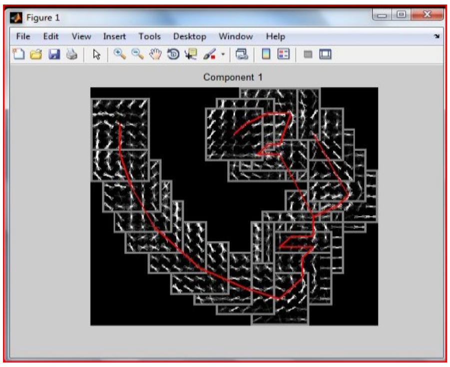


Fig 5: Simulated result of rotation +90^{0}.

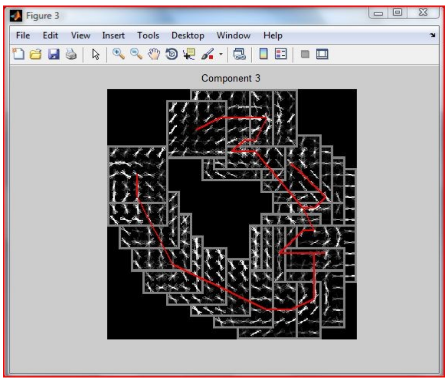


Fig 6: Simulated result of rotation 30^{0}.

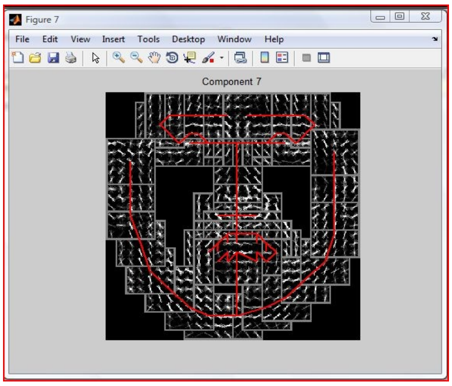


Fig 7: Simulated result of rotation 0^{0}.

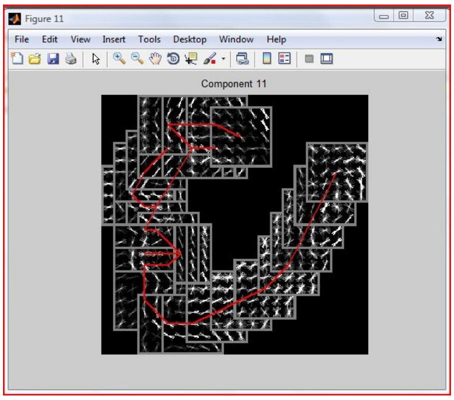


Fig 8: Simulated result of rotation -30^{0}.

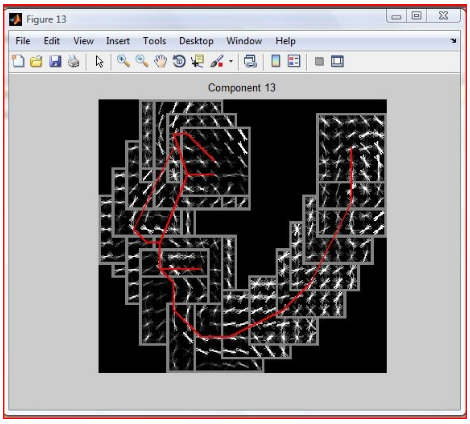


Fig 9: Simulated result of rotation -90^{0}.

A. Human feature extraction based on HOG

Challenges associated with face detection:

- Pose: frontal, 45degree, profile, upside down.
- Presence or absence of structural components: beards, mustaches, glasses, scarf.
- Facial expression.
- Occlusion.
- Image orientations.
- Imaging conditions like lightning and camera characteristics.

Hog descriptor implemented for an image is shown with an example in fig 10 and its steps to detect humans is as follows:

- Hog descriptor is based on dominant edge orientations.
- Edge detection applied.
- Image divided into cells.
- Histogram of edge orientations compiled.

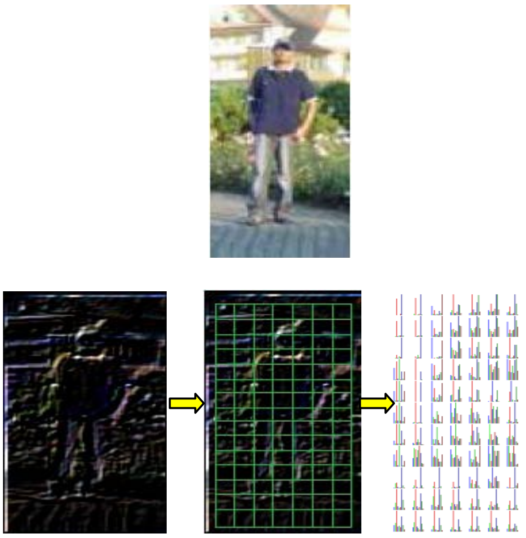


Fig 10: An example for Hog descriptor implementation

B. HOG algorithm implementation

Histogram of Oriented Gradient (HOG), a method of intensive descriptors that is used for local overlapped images, constitutes features by calculating the local direction of gradient. At present, the approach combining HOG with Support Vector Machine (SVM) has been widely applied to image recognition and achieved a great success especially in human detection.

The advantage of HOG feature is that it is based on histogram of oriented gradient. It can not only describe the feature of face contours, but also be not sensitive to light and small offset. Obtain the human facial features by combining the features of all blocks in line.

Face visualization models helps to develop a face detection algorithm based on HOG descriptor flow diagram is shown in fig 11 and designed specification are:

Frame Rate = 25.0000, Gain = 29, Gamma = 1, Hue = 0, Saturation = 13, block size for non maximal faces = 0.3, threshold = 0.65, Box cachesize =100000, detection time =8.6705secs.

Take the input image of 256*256 as an example shown in fig 11 shows the procedure of extracting depth image's HOG features, we calculate the HOG feature as follows:

1) Input an image which is the video frame captured continuously by web camera.

2) Gradient calculation: use the [-1, 0, 1] and [-1, 0, 1] median filter to perform filtering, calculate the vertical gradient and horizontal gradient of the image, and then calculate the gradient direction and gradient magnitude of each pixel.

3) Divide the inputting image into average small cells (including 256*256 pixels) and combine four cells into a small block, one block is constituted by a cell of 2*2.

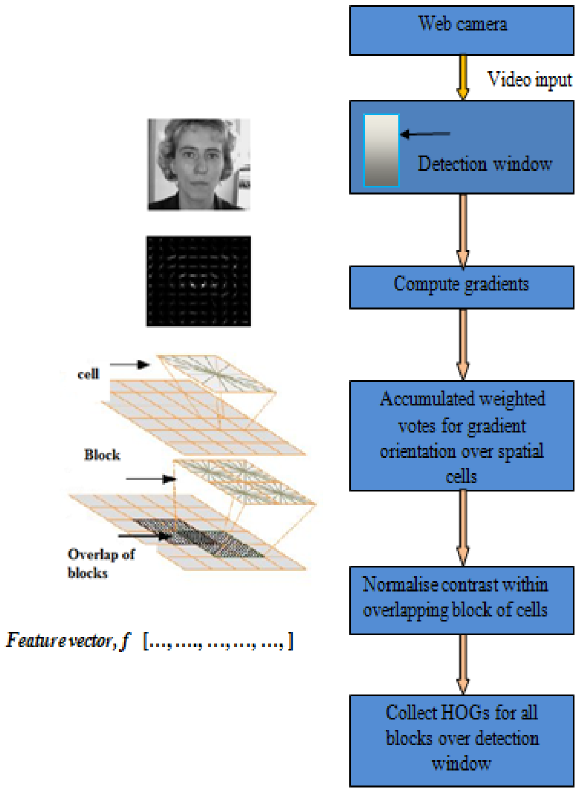


Fig 11: Face detection using HOG Descriptor

4) The selection of direction channel: divide 0^0 -180^0 or 0^0-360^0 into n channels averagely. In this paper, we divide +90^0 to -90^0 into thirteen equal parts, that is, thirteen channels in total. So there are 4*13=52 features in each block.

5) The acquisition of the histogram: get the statistics of each pixel in each cell of their histogram of orientated gradient. The abscissa of the histogram represents the thirteen direction channels selected in step 3, and the ordinate represents the summation of the gradient, belonging to a certain direction channel. Thus, we get a set of vectors.

6) The process of normalization: normalize the vectors in blocks in which pixels correspond with the vectors. Block normalization corrects local contrast variations and histograms for the cells of each block are normalized.

HOG method uses 6 basic parameters recalling: number of orientation bins, range of orientations to be considered, cell size, block size, overlap and normalization rule. In this paper, hog descriptor implemented by the $L2$-norm suggested in equation 8.

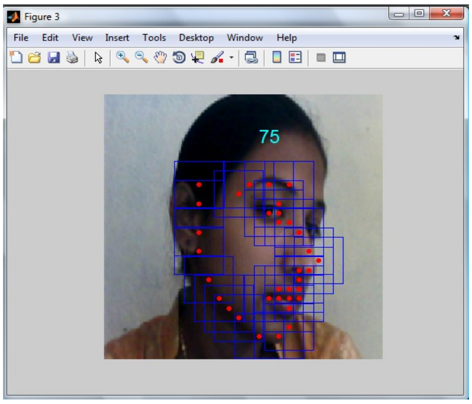
$$f = \frac{v}{\sqrt{||v||_2^2 + e^2}} \qquad (8)$$

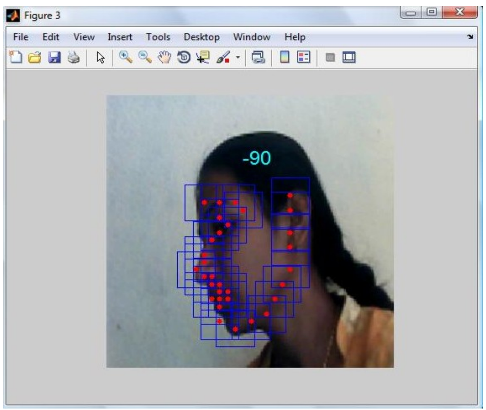
7) Form HOG features: combine all the vectors processed above and then form a set of vectors, which are the HOG features.

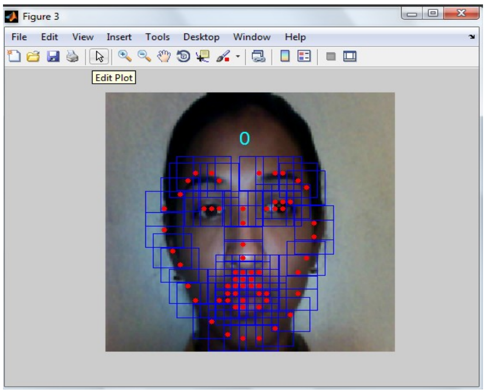
These steps ensure that little information is lost during the encoding process. Overall encoding focuses on capturing relevant fine grained features and adding the required degree $+90^0$ to -90^0 of invariance at each step to detect face or non face.

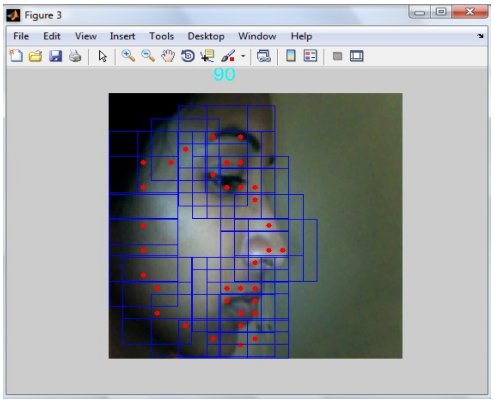
C. *Results*

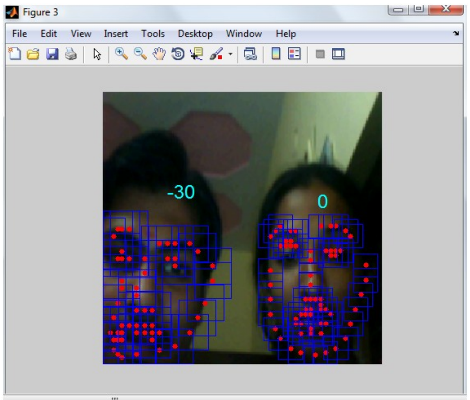
Face detection in real time implemented using hog descriptor and results are shown in fig below. First video frame is read and corresponding face visualization model will be set for the given image. Second checking the face orientations from $+90^0$ to -90^0 and if threshold is > 0.65 then detector return it as face with an face orientation angle, if it is below threshold then it detect it as non face. Later it gives highest detection rate returned with the correct orientation (degree) of a face. Similarly detection will be continued for incoming video frames with highest detection rate with a max delay of 5ms is shown in fig 12.











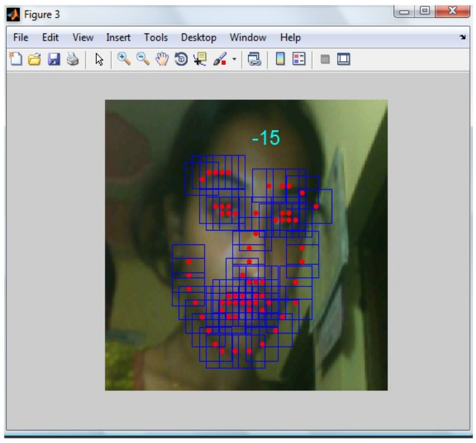


Fig 12: Face detection results based on HOG

V. ADVANTAGES

- Work stably with any skin color.
- Accept low-quality images (e.g. pixilated, blurred, noised, darkened, etc.)
- Detect occluded faces with any orientation.
- Determine the position of the particular facial features (e.g. eye pupils, nose, lips, etc.)
- Detect several faces on one image.

VI. APPLICATIONS

Face Detection technology may be used as a core component (basis) for a great number of applications with a wide sphere of usage.

- Smart captcha
- Webcam based energy/power saver.
- Time tracking service.
- Outdoor surveillance camera service.
- Video chat service.

VII. CONCLUSION

Face detection developed based on HOG descriptor is working well in real time. Face visualization models implemented based on HOG descriptor is able to visualize a face in any orientation from $+90^0$ to -90^0 and detection rate varies in few seconds depending on the faces that contained in the video frame. In future this method is helpful in face recognition systems.

REFERENCES

[1] N. Dalal, B. Triggs, *"Histograms of oriented gradients for human detection"*, CVPR, pp. 886-893, Vol. 1, 2005.

[2] S. Maji, A. Berg, J. Malik, *"Classification using intersection kernel support vector machines is efficient",* CVPR, 2008.

[3] Roberto Brunelli and Tomaso Poggio. *"Face Recognition: Features versus Templates"*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):1042–1052, October 1993.

[4] Andrew W. Senior and Ruud M. Bolle *"Face recognition and its applications"*, IBM T.J.Watson Research Center.

[5] Mengxin Li, Jingjing Fan, Ying Zhang, Rui Zhang, Weijing Xu, Dingding Hou *"Moving Object Detection and Tracking Algorithm"* TELKOMNIKA, Vol. 11, No. 10, October 2013, pp. 5539 ~ 5544.

[6] H. A. Rowley, S. Baluja, and T. Kanade, *"Neural Network-Based Face Detection,"* IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, Jan. 1998.

[7] S. Baluja and H. A. Rowley, *"Boosting set identification per-formance,"Int. J. Comput. Vis.*, vol. 71, no. 1, pp. 111–119, 2007.

[8] Sheenamol Yoosaf, Anish M P, *"Face Detection & Smiling Face Identification Using Adaboost & Neural Network Classifier",* International Journal of Scientific & Engineering Research, Volume 4, Issue 8, August 2013

[9]Dalal, N.Triggs, B: *"Histograms of Oriented Gradients for Human Detection, IEEE Computer Society* Conference on Computer Vision and Pattern Recognition", 2005.

[10] L. R. Cerna, G. Cámara-Chávez, D. Menott, *"Face Detection: Histogram of Oriented Gradients and Bag of Feature Method",* 2010.

[11] Sarala A. Dabhade, Prof. Mrunal S. Bewoor,*" 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 Issue 1 Aug 2012.

[12]Vaidehi V1, Annis Fathima A2, Teena Mary Treesa2, Rajasekar M2, Balamurali P3, Girish Chandra M3*"An Efficient Face Detection and Recognition System"* proceedings of international multi conference of engineers and computer scientists 2011 vol I.

[13] Dr. Chander Kant Nitin Sharma *"Fake Face Detection Based on Skin Elasticity",* International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 5, May 2013

