ISSN XXXX XXXX © 2017 IJESC

IJESC

Research Article Volume 7 Issue No.6

A Novel Approach for Detection of Faces using Viola-Jones Algorithm

Supriti gupta[1], Dr.Anil Ahlawat[2]
Student[1], Professor & Dean[2]
Department of Computer Science & Engineering
Krishna Institute of engineering & Technology, Ghaziabad, UP, India

Abstract:
The novel approach for face detection using the Viola-Jones algorithm with the probability based training algorithm which detect the eyes, nose, mouth and other body parts (like skin) which enhance the speed for the detection of faces is presented. Human face detection and recognition has become a major field of interest in current research because there is no deterministic algorithm to find face in a given image. Further the algorithms that exist are very much specific to the kind of images they would take as input and detect faces. While this algorithm detect features and compute the number of faces in the group of images in the given database with 100% efficiency for frontal page and rotatory images up to 15^{O}. Then the faces are recognized by the advance approach i-e SURF (speed up robust feature) in which we used Manhattan algorithm for the measurement of distances between the feature points which improve the performance of the recognized faces. The result are compared with existing algorithm and found the efficiency of this novel approach is better than existing approaches.

Keywords: Introduction, Viola-Jones algorithm, probability adaboost algorithm, skin color model, SURF, Manhattan algorithm.

I. INTRODUCTION

Detection of the human face is an essential step in computer vision and many biometric applications [1]. Human face plays an important part in our social interaction, conveying people's identity. In the past few years, face recognition has received a significant attention and regarded the most successful applications in field of image analysis [2]. Face recognition is the most relevant applications of image processing. Face detection cannot be straight forward as it has lots of variations like appearance of an image, variations in poses, occlusion, image orientation, and illuminating condition and face expressions. One of the fast methods for detection of faces is being given by Viola and Jones [3] .one of the major contribution of their work is use of cascade classifier for the face detection. Presented the fundamentals of their face detection algorithm. This algorithm only detects frontal upright faces; however, a modified algorithm was presented in 2003 that detects profile and rotated views [4]. In this paper, we present a novel skin color model RGB-HS-CbCr for human face detection for the Viola and Jones algorithm including the detection of faces with the help of skin color [5].With the Probabilistic Weighting Adjusted Adaboost techniques is used for increasing the performance of the Viola and Jones algorithm Although their method was successful in face detection, it faces false alarm challenges, which may increase in the presence of a complex background. False positives in an application can be a source of errors and need additional post processing to remove them proposed method is explained in [6] as in this they reduce the number of false positive. Then for the face recognition, we present a novel scale- and rotation-invariant interest point detector and descriptor, coined SURF (Speeded Up Robust Features). This algorithm also give the approximation or it could outperform the scheme's which is being proposed previously with respect to robustness, repeatability and distinctive, yet this can be compared and computed much faster with the help of Manhattan algorithm we find the distance or segmentation between the feature points in SURF methodology.

The flow chart for face detection and recognition is given in (fig 1).

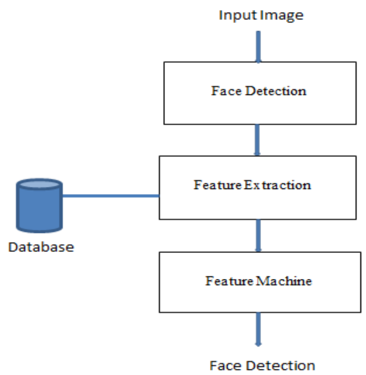

Figure.1. process of face detection and reorganization

II. VIOLA AND JONES ALGORITHM

The Viola–Jones algorithm is based on the object detection framework which is the first object detection framework which is used to provide the competitive object detection rates in the real-time proposed in 2001 by Paul Viola and Michael Jones. It train to detect many object but here we used for the detection of the faces. The problem to be solved is detection of faces in an image. It easy for human to do this, but a computer needs complete instructions and constraints. To make the task more manageable, Viola–Jones requires full view frontal upright faces. Thus in order to detect the faces, the entire face must focus on the camera and should not be tilted to either side.

The algorithm has four stages:

A. Haar feature selection

The Viola-Jones algorithm uses Haar-like features [3] that is, a scalar product between the image and some Haar-like templates. For the (x,y): Value = Σ (pixels in black area) - Σ (pixels in white area) To compensate the effect of different lighting conditions, all the images should be mean and variance normalized beforehand. there are different types of representation of haar feature that is in fig(2,3,4).

Figure.2. Haar Feature That Looks Similar To The Bridge Of The Nose Is Applied Onto The Face

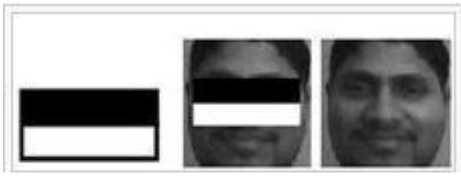

Figure.3. Haar Feature That Looks Similar To the Eye Region Which Is Darker Than the Upper Cheeks Is Applied Onto a Face

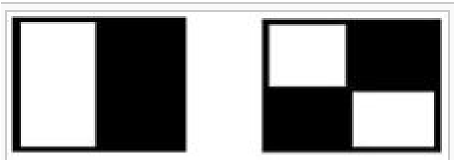

Figure.4. 3rd And 4th Kind Of Haar Feature

The representation of an image is called the integral image. Which evaluate the rectangular features of face at constant time, which have an advantage considerable speed over the more sophisticated alternative features. As each feature's rectangular area is to be considered always adjacent to at least one-other rectangle, it generally follows that any two-rectangle feature can be computed in the six array references, any three-rectangle feature in eight, and any four-rectangle feature in nine. At location (x,y) the integral image is consider, is sum of the pixels left and to the above of (x,y), inclusive.

B. Creating an integral images

Rectangle features can be computed very rapidly using an intermediate representation for the image which we call the integral image[7] The integral image at location contains the sum of the pixels above and to the left of , inclusive:

$$I\sum(x,y) = \sum_{\substack{x' \leq x \\ y' \leq y}} i(x'y')$$

We get the sum of integral image (Fig 5).

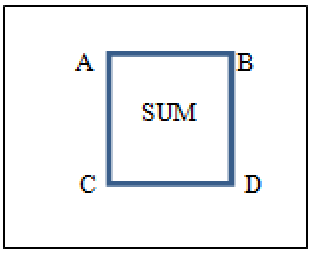

Figure.5. SUM= D-B-C+A

C. Adaboost training

The probability based algorithm we used for adaboost [6] which aims at an accurate algorithm to reduce false-positive detection rates. The probability was determined by computing both positive and negative classification errors for each weak classifier. The new weighting system used to give higher weights to the weak classifiers with the best positive classifications, which is used to reduce the false positives during detection of the faces.

D. Cascading feature

The cascade classifier is created by adaboost training [3] is used as a detector for the face detection process. The cascade consists of several classifier layers subordinately joined to each other (fig 6).

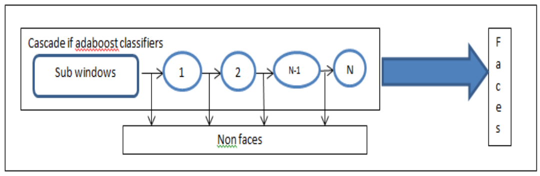

Figure.6. Cascade from simple to complex classifiers with N stages

The algorithm should concentrate on discarding non faces quickly and spend more time on probably face region. Based on haar cascade. To train and design a cascade we must choose: no. of stages in cascade(strong classifier)No. of features of each strong classifier. Threshold of each strong classifier

(the $\frac{1}{2}\sum_{t=1}^{T} \alpha_t$ *in definition*

Strong classifier definition:

$$h(x)=\begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2}\sum_{t=1}^{T} \alpha_t \\ 0 & otherwise \end{cases}$$

Where:

$$\alpha_{t=\log\frac{1}{\beta_t}}$$

$$\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$$

So, we get the false positive rate for an entire cascade is:

$$F = \prod_{i=1}^{K} f_i$$

Equation 1

Similarly, the detection rate is:

$$D = \prod_{i=1}^{k} d_i$$

Equation 2

III. FACE DETECTION USING (RGB-HS-CBCR SKIN COLOR MODEL)

Various colors spaces provide us various discriminality between skin pixels and non-skin pixels over various illumination conditions. Skin color models [8] that operate only on chrominance subspaces such as the Cb-Cr [9, 10, 11] and H-S [12] have been found to be effective in characterizing various human skin colors. While in this we used the concept of RGB-HS-CbCr concept of skin color model [5] for the detection of skin in the viola and jones algorithm [3]

IV. SURF (SPEED UP ROBUST FEATURE)

In 2006, three people, Bay, H., Tuytelaars, T. and Van Gool, L, "SURF: Speeded Up Robust Features" introduced a new algorithm called SURF. As name suggests, it is a speeded-up version of SIFT. Our fast detector(recognizer) and descriptor, called SURF (Speeded-Up Robust Features), was introduced in [13] For the recognition of faces here we used the concept of SURF(speed up robust feature) as it computed and compared much faster this is achieved by relying on integral images for image convolutions; by building on the strengths of the leading existing detectors and descriptors (in case, using a Hessian matrix-based measure for the detector, and a distribution-based descriptor); and by simplifying these methods to the essential. This leads to a combination of novel detection, description, and matching steps. A wide variety of detectors and descriptors have already been proposed in the literature (e.g. [14–19]). While this approach detector (recognizer) and descriptor is not only faster, but also more distinctive and equally repeatable.
The SURF generally works on main 3 concepts these are:

A. BASED ON FAST HESSIAN DETECTOR (INTEREST POINT DETECTION)

Interest point detection uses a very basic Hessian- matrix approximation. This lends itself to the use of integral images as made popular by Viola and Jones [3], which reduces the computation time drastically. Integral images in the more general framework of boxlets as proposed by Simard et al. [20]. We have already discussed about integral image in Viola and Jones [3]. Based on hassien matrix but use a very basic approximation – DOG (descriptor of Gaussian) and integral images (reduce the computation time).

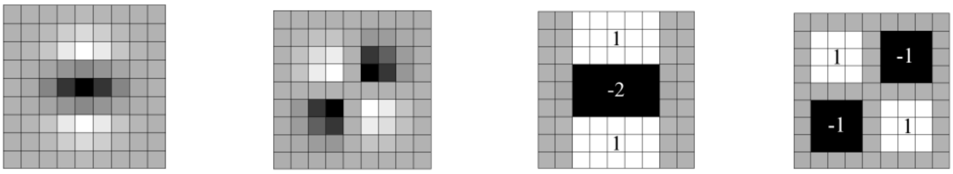

Figure.7. Left to right: the (discretised and cropped) Gaussian second order partial derivative in y- (Lyy) and xy-direction (Lxy), re- spectively; our approximation for the second order Gaussian partial derivative in y- (Dyy) and xy-direction (Dxy). The grey regions are equal to zero.

On the Hessian matrix we base our detector because of its good performance in accuracy. More precisely, we detect blob-like structures at locations where the determinant is maximum. In contrast to the Hessian- Laplace detector by Mikolajczyk and Schmid [21], we rely on the determinant of the Hessian also for the scale selection, as done by Lin-deberg [14]. Gaussians are optimal for scale-space analysis [19, 20],
Given a point x = (x; y) in an image I, the Hessian matrix H(x, σ) in x at a scale σ is defined as :

$$H(x,\ ) = \begin{bmatrix} l_{xx}(x,\sigma) & l_{xy}(x,\sigma) \\ l_{xy}(x,\sigma) & l_{yy}(x,\sigma) \end{bmatrix} \sigma$$

X(x,y), σ:scale(standard derivation of gaussian) Approximation LOG (laplacian of gaussian)with box filter – DOG

$$D_{xx}(x,\sigma) = \left(\frac{\partial^2}{\partial x^2} g(k\sigma) - \frac{\partial^2}{\partial x^2} g(\sigma)\right) \times I(x)$$

The 9 _ 9 box _lters in _gure 2 are approximations of a Gaussian with _ = 1:2 and represent the lowest scale (i.e. for the highest spatial resolution) for computing the blob response maps. We will denote them by Dxx, Dyy, and Dxy. In the rectangular regions the weights is applied to kept simple for computational efficiency. This yield:

$$\det(Happrox) = D_{xx}D_{yy} - (wD_{xy})^2$$

Where w (approximation Gaussian kernel) =0.9

The given (fig:8,9) shows the interest point detection which is done in the SURF algorithm which is used to detect the feature points of the faces then with the help of this algorithm we are going to match the faces.

Figure.8. image

Figure.9. Interest Point Detection of image

B. Local neighborhood descriptor

- Orientation descriptor: It is used for fixing a reducible orientation which is based on information from a circular region around the interest point. Based on haar wavelet
- Descriptor component : construct a square region centered around the interest point and orientation

The region is split up regularly in to smaller 4X4 square sub region.

V= $\left(\sum dx, \sum dy, \sum|dx|, \sum|dy|\right)$

C. Matching

The fast indexing is done through the sign of the laplacian for the underlying interest point.

The sign trace of hessian matrix.

Tr(H) or trace = $l_{xx} + l_{yy}$

OR

Tr(H)= $D_{xx} + D_{yy}$

also called the L1 distance [24]. The distance between a point x=(x1,x2…xn) and a point y=(y1,y2 ,…yn) is: $MD_{(x,y)} = \sum_{i=1}^{n} |x_i - y_i|$

Typically, the interest points are found at blob-type structures. From the reverse situation the sign of the Laplacian distinguishes bright blobs on dark backgrounds. This feature is available at no extra computational cost as it was already computed during the detection phase. In the matching stage, we only compare features if they have the same type of contrast, see figure 15. Hence, this minimal information allows for faster matching, without reducing the descriptor's performance. Note that this is also of advantage for more advanced indexing methods. E.g. for k-d trees, this extra information defines a meaningful hyper plane for splitting the data, as opposed to randomly choosing an element or using feature statistics.

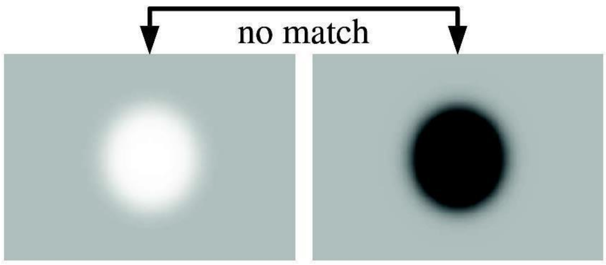

Figure.10. If the contrast between two interest points is different (dark on light background vs. light on dark background), the candidate is not considered a valuable match.

Experiments for camera calibration and object recognition highlighted SURF's potential in a wide range of computer vision applications. In the former, the accuracy of the interest points and the distinctiveness of the descriptor showed to be major factors for obtaining a more accurate 3D reconstruction, or even getting any 3D reconstruction at all in difficult cases. In the latter, the descriptor generalizes well enough to outperform its competitors in a simple object recognition task as well [26] while the implementation for object recognition is explain here [25] for the real time object recognition.

D. Manhattan algorithm

Manhattan distance is also called city block distance. It computes the distance that would be traveled to get from one data point to the other, if a grid-like path is followed. The Manhattan distance between two items is the sum of the differences of their corresponding components. Manhattan distance is

V. RESULT AND DISCUSSION

In this research paper have taken some Google images and some MIT/CMU database for the detection of faces. This (fig 11) contain the MIT/CMU database frontal face detection for the front face in which it is detected the feature like eye, nose, mouth (Fig 12) and we get the cropped image (Fig 13) of the front face from the whole images. The detection of the single images from the database is 100% as it detect the frontal face with the features.

Figure.11. Original image

Figure.12. Face detected with feature detected

Figure.13. the crop image after detection only face detection

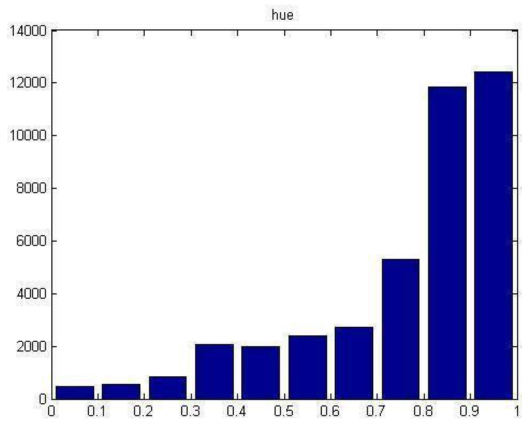

Figure.14. HUE histogram for the detected image

Then after that it also explains the multiple face detection in a given images how many faces are there in group of multi images Google image (fig 14). Where in this it detected that there are total 6 faces in the group of multiple face detection (fig 15). In this it also explains the detection of the rotated image about 15degree rotated image can be detected according to the detected algorithm (fig 16, 17).

Figure.15. original images of multi-image

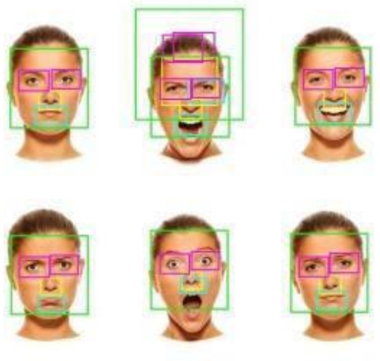

Figure.16. multiple face detection

Figure.17. original rotated image

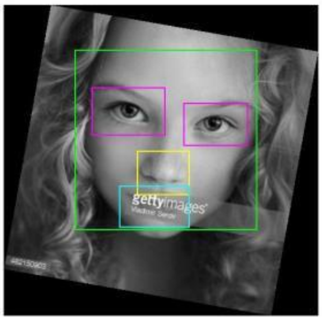

Figure. 18. detection of rotated image

While in this paper we work on the real time image detection method based on viola and jones algorithm for real time detection in which it capture the real time image form the webcam at the same time and detect the image as a face with the feature extraction fig(23, 24). While preprocessing of any image we get this result in the HSI (Hue saturation intensity). In this paper its show the preprocessing image for the web cam image which is after being used for the detection of the faces after the pre-processing of any image (Fig 19).

Figure.19. webcam image (real time)

Figure.20. pre-processed HSI image

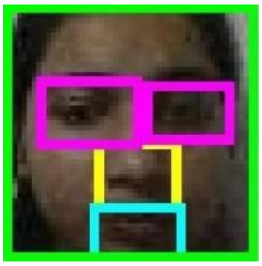

Figure.21. detected feature of webcam image

VI. CONCLUSION

This paper use concept of VIOLA and JONES face detector for face detection and for the feature detection eye, nose, mouth, and skin for the better recognition. It is robust, distinctive and less time consuming comparatively other algorithm. After detecting face it uses SURF ALGORITHM for face recognition. It approximates or even outperforms previously proposed schemes with respect to repeatability, distinctiveness, and robustness, yet can be computed and compared much faster. By using the concept of face detection and recognition algorithm together three improvements are done (a) dynamically build face detection (b) feature point detection and (c) fast matching algorithm. This comprises of various techniques used in digital image processing which put forwarded during the implementation of algorithms with high accuracy rate and improve the performance of the algorithms. These techniques are generally used for the dimensionality reduction and are influential in applications like security, surveillance etc.

VII. REFERENCES

[1]. R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain., "Face detection in color images," IEEE Trans. PAMI, vol. 24, no. 5, pp. 696-707, 2002.

[2]. W. Zhao., "Face recognition: a literature survey" et al (2000), Technical Report CAR-TR-948, University of Maryland, October 2000.

[3]. Viola, P. and Jones, M.,"Neural network-based face detection, rapid object detection using a Boosted Cascade of Simple Features", Proc., CVPR Vol.1, pp.511-518 (2001)

[4]. P. Viola and M. J. Jones., "Fast multi-view face detection," Mitsubish Electronic Research Laboratories TR2003-096, 2003.

[5]. Sayantan Thakur, Sayantanu Paul, Ankur Mondal, Swagatam Das, Ajith Abraham., "Face Detection Using Skin Tone Segmentation", "World Congress on Information and Communication Technologies"(2011).

[6]. Cesar Niyomugabo, Hyo-rim Choi, and Tae Yong Kim., "A Modified Adaboost Algorithm to Reduce False Positives in Face Detection," "Hindawi Publishing Corporation Mathematical Problems in Engineering", Article ID 5289413, 6 pages, 2016.

[7] Anonymous. Anonymous. In Anonymous, 2000.

[8]. M. J. Jones and J. M. Rehg., "Statistical color models with application to skin detection," Proc. CVPR'99, June 1999, pp. 274- 280.

[9]. S. Gundimada, L. Tao, and V. Asari., "Face Detection Technique based on Intensity and Skin Color Distribution," ICIP2004, pp.1413-1416, 2004.

[10]. S.L. Phung, A. Bouzerdoum, and D. Chai., "A Novel Skin Color Model in YCbCr Color Space and its Application to Human Face Detection," ICIP2002, pp. 289-292, 2002.

[11]. R.-L. Hsu, M. Abdel-Mottaleb, and A.K. Jain., "Face Detection in Color Images," IEEE Trans. PAMI, 24(5), pp.696-706, 2002.

[12]. K. Sabottka and I. Pitas., "Segmentation and Tracking of Faces in Color Images," AFGR'96, Killington, Vermont, pp. 236-241, 1996.

[13]. H. Bay, T. Tuytelaars, and L. Van Gool.," SURF: Speeded up robust features," In ECCV, 2006.

[14]. Lindeberg, T., "Feature detection with automatic scale selection," IJCV 30(2) (1998) 79 – 116.

[15]. Lowe, D.," Distinctive image features from scale-invariant keypoints, cascade filtering approach," IJCV 60 (2004) 91– 1104 http:// www. vision.ee. ethz. ch/ ˜ surf/ SURF: Speeded Up Robust Features 13

[16]. Mikolajczyk, K., Schmid, C.,"An affine invariant interest point detector. In: ECCV,"(2002) 128 – 142

[17]. Ke, Y., Sukthankar, R., "PCA-SIFT: A more distinctive representation for local image descriptors," In: CVPR (2). (2004) 506 – 513

[18]. Tuytelaars, T., Van Gool, L.," Wide baseline stereo based on local, affinely invariant regions," In: BMVC. (2000) 412 – 422

[19]. Matas, J., Chum, O., M., U., Pajdla, T., "Robust wide baseline stereo from maximally stable extremal regions," In: BMVC. (2002) 384 – 393

[20]. P. Simard, L. Bottou, P. Ha_ner, and Y. LeCun. Boxlets., "A fast convolution algorithm for signal processing and neural networks," In NIPS, 1998.

[21]. K. Mikolajczyk and C. Schmid., "Indexing based on scale invariant interest points," In ICCV, volume 1, pages 525 {531, 2001.

[22]. J.J. Koenderink., "The structure of images," Biological Cybernetics, 50:363 {370, 1984.

[23]. T. Lindeberg., "Scale-space for discrete signals," PAMI, 12(3):234-254, 1990.

[24]. ShubhenduTrivedi., "face Recognition using Eigen faces and Manhattan distance Classifiers: A tutorial", Feb 2009, face-recognition-using-eigen faces-and- distance- classifiers-a-tutorial

[25]. Fari Muhammad Abubakar.," IMPLEMENTATION OF SPEED UP ROBUST FEATURE FOR DETECTION AND TRACKING OF INANIMATE OBJECTS," International Journal of Computer Science and Mobile Computing, Vol.3 Issue.1, January- 2014, pg. 325-333.

[26]. Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008)., "Speeded-up robust features (SURF)," Computer vision and image understanding, 110(3), 346-359.

[27]. Supriti Gupta and Dr. Anil Ahlawat, "A Review of Face Detection and Recognition using Hybrid Approaches", International Journal of Computer & Mathematical Sciences (IJCMS); ISSN 2347 – 8527 Volume 6, 5 May 2017.