

# A Review: Face Detection Methods And Algorithms

<sup>1</sup>Neetu Saini, <sup>2</sup>Sukhwinder Kaur, <sup>3</sup>Hari Singh

<sup>1,2</sup> M. Tech. Scholar (ECE), DAV Institute of Engineering and Technology, Jalandhar (India)

<sup>3</sup> Assistant Professor (ECE), DAV Institute of Engineering and Technology, Jalandhar (India)

## ABSTRACT:

Face detection which is the task of localizing faces in an input image is a fundamental part of any face processing system. The aim of this paper is to present a review on various methods and algorithms used for face detection i.e. Haar cascade, adaboost, template matching etc. This paper also includes the algorithms of eye blink detection. Finally it includes some of applications of face detection

**Key Words:** Face Detection, eye detection, eye blink detection.

## 1. INTRODUCTION

Face detection is a computer technology that determines the locations and sizes of human faces in arbitrary (digital) images. It detects facial features and ignores anything else, such as buildings, trees and bodies. Human face perception is currently an active research area in the computer vision community. Human face localization and detection is often the first step in applications such as video surveillance, human computer interface, face recognition and image database management. Locating and tracking human faces is a prerequisite for face recognition and/or facial expressions analysis, although it is often assumed that a normalized face image is available.

In order to locate a human face, the system needs to capture an image using a camera and a frame-grabber to process the image, search the image for important features and then use these features to determine the location of the face. For detecting face there are various algorithms and methods including skin colour based, haar like features, adaboost and cascade classifier. Colour is an important feature of human faces. Using skin-colour as a feature for tracking a face has several advantages. Color processing is much faster than processing other facial features [20].

**Localization v/s detection:** Face localization: Find one and only one face assuming that it is shown in an image or video.

Face detection: Find all visible faces in an image or video.

**True-positive:** also called hit or detection; a correctly detected face.

**False-positive:** also called miss- or false-detection detecting a face where there is none actually.

**False-negative:** when missing a visible face

**True-negative:** describing non-face regions correctly as non-face region



Fig 1: Description of faces and non-faces [1].

## 2. METHODS OF FACE DETECTION

### 2.1 Methods using a skin colour model

- Normalize RGB colours and intensity values in the image.
- Mark pixels that match an established skin colour model.
- Remove regions (e.g. due to being 'small') that are unlikely to represent faces.
- Confirm a face appearance by verifying common features of a human face.

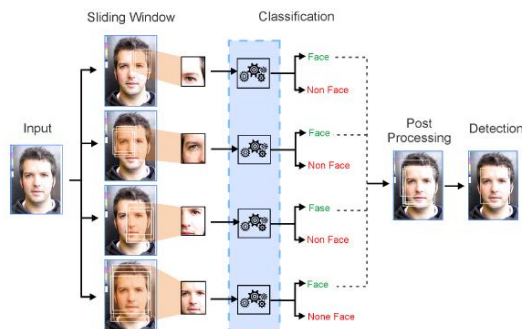
### 2.2 Feature-based methods

- Make explicit use of local face features (e.g. eyes, nose, mouth) and of geometric relationships between those.
- There should be only one face in the search domain (face localization) for avoiding geometric confusion.
- Method require good quality images.
- Potentially robust to most of the geometric and photometric deformations.
- Method is computationally expensive.

### 2.3 Appearance-based methods

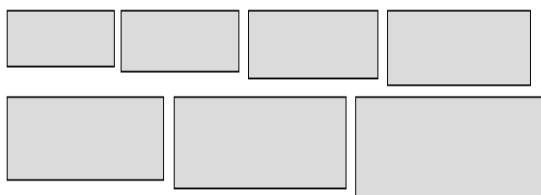
- Face detection as two-class pattern recognition.

- Apply a statistical learning method and use a training data set to built a face/no-face classifier.
- Applicable to low resolution images.
- Receive recently considerable attention.
- More successful than feature-based approaches.



**Fig 2: Concepts of Appearance-Based Methods [1].**

**Sliding Window:** The basic idea of appearance-based methods is the application of a sliding window. A sliding window scans the input image at different locations and resolutions. (Rotations are considered later.) The aim of this sliding search is to find the location of a face at some resolution. We do not change the image resolution; instead we change the size of the sliding window in a new search iteration. The window increases by a factor such as  $\Delta z > 1.1$ . The figure below illustrates increases by the lower bound factor  $\Delta z = 1.1$ . A sliding window is square, and there are rectangular Haar-like wavelets in the sliding window. More than this 10% increase in width and height of the sliding window ensures time efficiency. On the other hand, this may decrease the accuracy of detection. The figure illustrates resulting increases for rectangular wavelets:

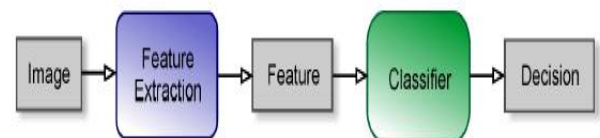


**Fig 3: Rectangular Haar-like wavelets in the sliding window [12].**

So a trade-off should be considered. Sliding window explores locations either “exhaustively” from top left corner of the image to the bottom right, with a scaling factor of  $\Delta z$  in x- and y-direction, or randomly, or in zigzag or spiral scan order, or by using some priority information (e.g. skin colour, local features), or any combination of above.

### Classification:

Each placed window is passed to a classifier for detecting either a face or no-face situation. This classification is the core and the most important part of the face detection system. After feature extraction, features need to be compared and matched by the classifier with desired features. The classifier then decides whether there is a face in the region or not (for more details, see below).



**Fig 4: Concept of classifier [15].**

### Post-processing:

After classification, usually there are multiple overlapping detections at different locations and sizes around a visible face. The goal of post-processing is to return a single detection per face. Methods applied for post-processing are usually heuristic (e.g. by applying mean calculations, or more advanced statistical methods).

### 2.4.TEMPLATE MATCHING

Template matching method that finds the similarity between the input images and the template images (training images). Template matching method can use the correlation between the input images and stored standard patterns in the whole face features, to determine the presence of a whole face features [22]. This method can be used for both face detection and face locations. In this method, a standard face (such as frontal) can be used. The advantages of this method are that it is very simple to implement the algorithm, and it is easily to determine the face locations such as nose, eyes, mouth etc based on the correlation values. It can be apply on the various variations of the images such as pose, scale, and shape. Sub-templates, Multi resolutions, and Multi-scales have been proposed to achieve the shape and the scale invariance and localization method based on a shape template of a frontal view face [17]. A Sobel filter is used to extract the edges.

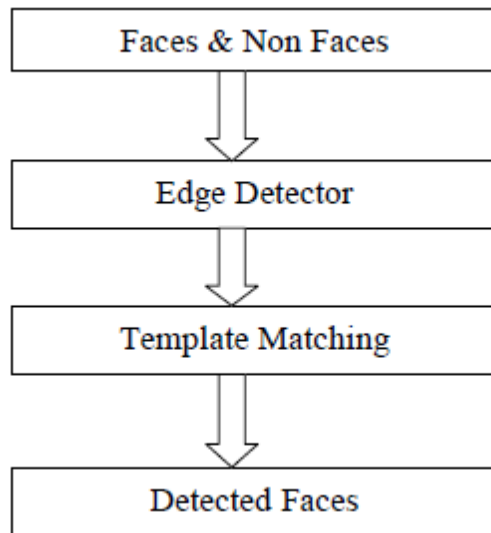


Fig 5:Block diagram of template matching method [22]

### 3. ALGORITHMS OF FACE DETECTION

#### 3.1 Haar like feature:

Haar-like wavelets are binary rectangular representations of 2D waves. A common visual representation is by black (for value 'minus one') and white (for value 'plus one') rectangles. The figure below shows a cut through a binary wavelet between  $x = 0$  to  $x = 1$ . The square above the 0-1-interval shows the corresponding Haar-like wavelet in common black-white representation. The rectangular masks used for visual object detection are rectangles tessellated by black and white smaller rectangles. Those masks are designed in correlation to visual recognition tasks to be solved, and known as Haar-like wavelets. By convolution with a given image they produce Haar-like features.[11],[12].

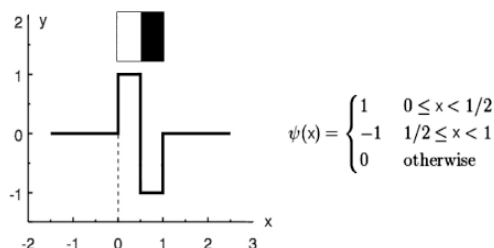


Fig 6: Representation of Haar-like wavelets [1],[12].

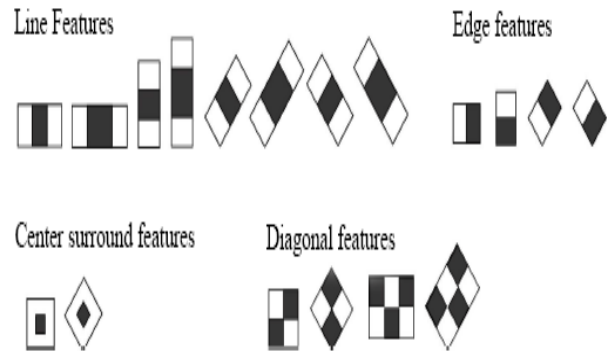


Fig 7: Feature prototypes of simple Haar-like . Black areas have negative and white areas positive weights [12].

Calculated features should be able to highlight important value distributions in objects of interest (e.g. in a face). For example, looking at the two faces below, for such a frontal upright view we may expect that faces have the following features:-  
 - Eye regions are darker compared to the bridge of the nose.

- Eye regions are darker compared to the cheeks.
- The iris region is darker compared to the sclera.



Fig 7:Distribution of eye region [1].

A Haar-like feature is determined by the convolution with the defining mask, having values -1 or +1 in its rectangular regions. Thus, this is simply done by subtracting the average of the pixel values in the black rectangles from the average of the pixel values in the white rectangles. If the difference is above at threshold, that feature (or wavelet) is said to be present. Used thresholds are specified during a training process.

#### 3.2. Integral images:

To determine the presence or absence of hundreds of Haar-like features at pixel locations and for several scales efficiently, Viola and Jones used integral images. In general, "integration" means adding small units together[17]. In this case, the small units are the pixel values. For pixel  $p = (x, y)$ , the integral value:

$$I(p) = \sum_{1 \leq i \leq x \wedge 1 \leq j \leq y} P(i, j)$$

[19].

is the sum of all the pixel values  $P(q)$ , where pixel  $q = (i, j)$  is not below and not on the right of  $p$ . For the mean, simply divide  $I(p)$  by  $x.y$ . See the figure on the left. TL stands for "top left".

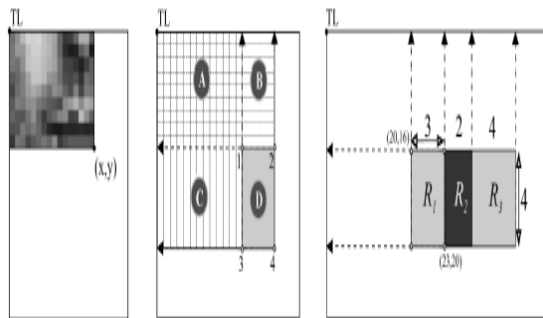


Fig 8: The Integral Image representation [17].

The integral image  $I$  is calculated as preprocessing. In the feature calculation process, sums need to be determined in rectangular areas such as for area D below. Assume that  $p_1$  is the lower right pixel in region A,  $p_2$  the lower right pixel in region B,  $p_3$  the lower right pixel in region C, and  $p_4$  the lower right pixel in region D. The sum of all pixel values in D equals

$$I(D) = I(p_4) + I(p_1) - I(p_2) - I(p_3)$$

### 3.3. AdaBoost:

Adaboost is an algorithm for constructing a "strong" classifier as linear combination.

$$H(x) = \text{sign} \left( \sum_{t=1}^T \alpha_t \cdot h_{a(t)}(x) \right)$$

Adaboost, short for Adaptive Boosting, is a machine learning algorithm, formulated by Yoav Freund and Robert Schapire[18]. It is a meta-algorithm, and can be used in conjunction with many other learning algorithms to improve their performance. Adaboost is adaptive in the sense that subsequent classifiers built are tweaked in favour of those instances misclassified by previous classifiers. Adaboost is sensitive to noisy data and outliers. In some problems, however, it can be less susceptible to the over fitting problem than most learning algorithms. The classifiers it uses can be weak (i.e., display a substantial error rate), but as long as their performance is slightly better than random (i.e. their error rate is smaller than 0.5 for

binary classification), they will improve the final model. Even classifiers with an error rate higher than would be expected from a random classifier will be useful, since they will have negative coefficients in the final linear combination of classifiers and hence behave like their inverses[18].

Adaboost generates and calls a new weak classifier in each of a series of rounds  $t = 1, \dots, T$ . For each call, a distribution of weights  $D_t$  is updated that indicates the importance of examples in the data set for the classification. On each round, the weights of each incorrectly classified example are increased, and the weights of each correctly classified example are decreased, so the new classifier focuses on the examples which have so far eluded correct classification.

### 3.4. Cascaded Weak Classifiers:

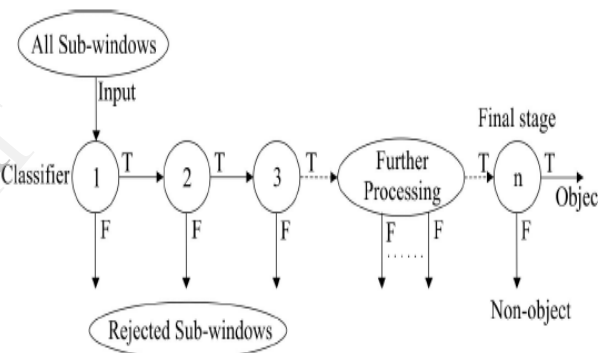


Fig 9: Cascade classifier [15].

Each of the weak classifiers has the task to detect a face [15][17]. They are performed in a cascade. A search window (sliding window) of  $24 \times 24$  pixels contains more than 180,000 different rectangular sub-windows of different size (isothetic or in 45 degree rotation). Only a small number of weighted Haar-like wavelets (usually less than 100) is sufficient to detect a desired object in an image, such as a face. The selection of such Haar-like wavelets can use available a-prior knowledge (e.g, the expected size of a face). A strong classifier is generated by boosting from a selected set of weak classifiers.

### 4.BLINK DETECTION

There are a number of image processing algorithms for eye blink detection [4],[13],[5]. A brief overview of three of these algorithms is provided.

**4.1 Contour Extraction:** In this technique, a set of 16 landmarks are created at regular intervals to outline the contour of the eye. Eight points are used to represent each eye.[6] The distance between the



highest and lowest landmark is denoted by  $d1$ , and the distance between the centroids of the two eyes is denoted by  $d2$ . Now  $d1/d2$  is computed, and assigned to a variable  $D$ . Now the value of  $D$  is used to distinguish between open eye and closed eye. Generally, a value of  $D$  equal to 0.158 implies open eye and a value equal to 0.016 implies closed eye. These values have been experimentally derived by [3].

#### 4.2 Gabor Filter:

The Gabor filter is used to extract arcs of the eye. Here, the eye region is first extracted and then the filter is applied to obtain the arcs of the eye. Then connected component labelling method is used to detect the top-bottom arcs. The distance between the arcs is measured to determine the blinking [4],[8].

#### 4.3. Median Blur Filtering:

In this method, the image of the eye is first threshold and then a median blur filter is applied to it. The resultant image obtained after applying the filtering shows a clear difference between the open and the closed eye, and hence helps in identifying eye blinks [6].

#### Applications:

Face detection is used in biometrics, often as a part of (or together with) a facial recognition system. It is also used in video surveillance, human computer interface and image database management. Some recent digital cameras use face detection for autofocus. Face detection is also useful for selecting regions of interest in photo slideshows that use a pan-and-scale Ken Burns effect [23].

Face detection is gaining the interest of marketers. A webcam can be integrated into a television and detect any face that walks by. The system then calculates the race, gender, and age range of the face. Once the information is collected, a series of advertisements can be played that is specific toward the detected race/gender/age. Face detection is also being researched in the area of energy conservation.

#### DISCUSSION

Different methods and algorithms of face detection have been reviewed in this paper. The choice of a face detection method in any study should be based on the particular demands of the application. None of the current methods is the universal best for all applications. Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real-time face detector. A Haar-like feature considers adjacent rectangular regions at a specific location in a

detection window, sums up the pixel intensities in each region and calculates the difference between these sums[11],[12]. This difference is then used to categorize subsections of an image. The key advantage of a Haar-like feature over most other features is its calculation speed.

In order to be successful a face detection algorithm must possess two key features, accuracy and speed. There is generally a trade-off between the two. Through the use of a new image representation, termed integral images, Viola and Jones describe a means for fast feature evaluation, and this proves to be an effective means to speed up the classification task of the system.

Adaboost, short for Adaptive Boosting, is a machine learning algorithm. Adaboost algorithms take training data and define weak classifier function for each sample of training data. It can be less susceptible to the over fitting problem than most learning algorithms. Bad feature of adaptive boosting is its sensitivity to noisy data and outliers.

The weak classifiers have the task to detect a face. They are performed in a cascade. A search window (sliding window) of  $24 \times 24$  pixels contains more than 180,000 different rectangular sub-windows of different size.

#### References:

- [1] Parris, J. Wilber, M.; Heflin, B. "Face and eye detection on hard datasets", Biometrics (IJCB), 2011 International Joint Conference on, Washington, DC, 11-13 Oct. 2011, pp. 1 – 10.
- [2] Takahashi, K., Mitsukura, Y. "Eye blink detection using monocular system and its applications "2012, Grad. Sch. of Sci. & Technol., Keio Univ., Yokohama, Japan. 9-13 Sept. 2012, pp 743 – 747.
- [3] Sanjay Kr. Singh, D. S. Chauhan, Mayank Vatsa, Richa Singh "A robust skin color based face detection algorithm" 2003. Tamkang Journal of Science and Engineering, Vol. 6, No. 4, pp. 227-234.
- [4] Udayashankar, A. ; Kowshik, A.R. ; Chandramouli, S. ; Prashanth, H.S. " Assistance for the Paralyzed Using Eye Blink Detection " 2012, Digital Home (ICDH), 2012 Fourth International Conference. Pp 104 – 108.
- [5] Michael Chau and Margrit Betke, . "Real Time Eye Tracking and Blink Detection with USB Cameras"2005. Boston University, USA. May 12,2005.
- [6] Liting Wang ; Xiaoqing Ding Chi Fang ; Changsong Liu and Kongqiao Wang "Eye blink detection based on eye contour extraction".2009, Proc. SPIE 7245, Image Processing: Algorithms and Systems VII, 72450R ,February 10, 2009
- [7] M. Takagi, K. Mohri, M. Katoh and S. Yoshino."Magnet-Displacement Sensor Using Magneto-Inductive Elements for Sensing Eyelid Movement". IEEE Translation Journal On Magnetics In Japan, Vol. 9,No.2, pp 78-83.

- [8] Kohei Aai and Ronny Mardiyanto , 2011. "Comparative Study on Blink Detection and Gaze Estimation Methods for HCI, in Particular, Gabor Filter Utilized Blink Detection Method". Proceedings of 18<sup>th</sup> International Conference on Information Technology: New Generations. Las Vegas, USA, pp. 441-446.
- [10] Cihan Topal, Ömer Nezihi Gerek and Atakan Doğan, 2008. "A HeadMounted Sensor-Based Eye Tracking Device: Eye Touch System". Proceedings of the 2008 Symposium on Eye tracking research & applications. Savannah, GA, USA, pp. 87-90
- [11] Lae-Kyoung Lee, Su-Yong An , Se-Young Oh, Aug 2011. "Efficient. Face Detection and Tracking with extended camshift and haar-like features". International Conference on Mechatronics and Automation (ICMA), 2011. Pohang, South Korea, pp. 507-513.
- [12] Duan-Sheng Chen , Zheng-Kai Liu, Aug 2011. "Generalized Haar-like features for Fast Face Detection". Conference on Machine Learning and Cybernetics, 2007. Hong Kong, pp. 2131-2135
- [13] Bhaskar, T.N., Foo Tun Keat , Ranganath, S., Venkatesh, Y.V., 2003." Blink Detection and Eye Tracking for Eye Localization". Conference on Convergent Technologies for Asia-Pacific Region, 10 oct 2003.
- [14] Cai, J., Goshtasby, A. and Yu, C., "Detecting Human Faces in Color Images," Proceedings of International Workshop on Multi-Media Database Management Systems, pp. 124-131 (1998).
- [15] Viola, P. and Jones, M., "Rapid Object Detection using a Boosted Cascade of Simple Features," Proc. of the Conf. On Computer Vision and Pattern Recognition (CVPR), Hawaii, USA, December 9-14, 2001, Vol. 1, pp. 511-518.
- [16] Adolf, F., "How-to build a cascade of boosted classifiers based on Haar-like features," *OpenCV's Rapid Object*.
- [17] P. Viola and M. Jones "Robust real time object detection" Proceedings of International Journal of Computer Vision, pp.137-154.2004.
- [18] Pu Han Jian-Ming Liao, "Face detection based on adaboost" Apperceiving Computing and Intelligence Analysis, 2009. ICACIA 2009. International Conference on, 23-25 Oct. 2009, pp. 337 – 340.
- [19] Ogunbona, Wanqing Li "face detection using generalised integral image features" 2009. Image processing (ICIP), 16th IEEE International Conference. ., Sydney, NSW, Australia 7-10 Nov. 2009, Pp 1229 – 1232.
- [20] Paul Viola & Micheal Jones. "Robust real-time object detection". Second International Workshop on Statistical Learning and Computational Theories of Vision Modeling, Learning, Computing and Sampling, July 2001.
- [21] Akhil Gupta, Akash Rathi, Dr. Y. Radhika "hands free pc control of mouse cursor using eye movement" International Journal of Scientific and Research Publications, Volume 2, Issue 4, April 2012.
- [22] Smita Tripathi, Varsha Sharma. "Face Detection using Combined Skin Color Detector and Template Matching Method". International Journal of Computer Applications (0975 – 8887) Volume 26– No.7, July 2011.
- [23] H Drewes, "Eye Gaze Tracking for Human Computer Interaction", A Dissertation submitted in the partial fulfilment of the Ph. D. Degree, 2010. www.ndltd.org
- [24] K Lin, J Huang, J Chen, and C Zhou, "Real-time Eye Detection in Video Stream", Fourth International Conference on Natural Computation, 2008, pp 193-197.
- [25] A Phirani, S K Patel, A Das, R Jain, V Joshi, and H Singh, "Face Detection using MATLAB", proceedings of the 2010 International Conference on Optoelectronics and Image Processing (ICOIP 2010), November 11-12, 2010, China, Volume-I, pp 545 – 548.
- [26] W Chen, T Sun, X Yang, and L Wang, "Face Detection Based on Half Face Template", The 9th International Conference on Electronic Measurement and Instruments (ICEMI'09), 2009, pp 4-54 to 4-58.
- [27] K Aai, and R Mardiyanto, "Comparative Study on Blink Detection and Gaze Estimation Methods for HCI, in Particular, Gabor Filter Utilized Blink Detection Method", 2011 8th International Conference on Information Technology: New generation, 2011, pp 441-446.
- [28] R L Hsu, M A Mottaleb, and A K Jain, "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, May 2002, Vol 24, Issue 5, pp 696-706.
- [29] S Saravanakumar, and N Selvaraju, "Eye Tracking and Blink Detection for Human Computer Interface", International Journal of Computer Application, Vol. 2, No. 2, May 2010, pp 7-9.
- [30] K Lin, J Huang, J Chen, and C Zhou, "Real-time Eye Detection in Video Stream", Fourth International Conference on Natural Computation, 2008, pp 193-197.