Face Detection by Using Viola-Jones Algorithm

1. Shubham Bhausaheb Daule (2787739)
2. Shivani Phadtare (2787493)
3. Ramya Kakarla (2785312)

# Abstract – In digital image processing, face detection is one of the most important applications. The Viola-Jones algorithm is used to detect the face in an image. This algorithm is used to identifying and locating the human face irrespective of its size, situation, and surroundings. Face detection is a technique that detects the human face and ignoring anything else, like trees, bodies, and buildings. In this paper, the Viola-Jones algorithm is practically implemented by using MATLAB R2013a for detecting the human face in images. This algorithm is used to find out an automatic human face and it also calculates the accuracy of the system.

Index Terms - Viola-Jones, face detection, Haar-feature, Ad boost

**INTRODUCTION**

In the last few years, face detection has obtained crucial attention. In few areas, face detection gains superb function due to diverse reasons including verification of identities, huge range of industrial and regulation enforcement available for viable technologies. Face detection is one among many applications in digital photo processing. Its miles worried about the automated identification of a man or woman in a digital image. Face detection is used to figuring out and finding the human face no matter its length, position, and situation. Face detection is a clean challenge for the human mind; however, it is a completely tough task for laptop structures [1]. To come across the face effortlessly and correctly, the laptop device wishes some schooling elements to easily identify whether it is face or non-face. For detecting the face some threshold values are units based on those values a gadget can stumble on the human face. If the picture specifies the preferred threshold value, then the photograph is a face otherwise it is a non-face.

Face detection algorithms are divided into elements: (i) Feature based (ii) Learning based. The feature-based totally algorithms are based on the assertion that the face is detected based totally on a few simple capabilities, impartial of mild, face variation, and posture. The learning-based set of rules utilization numerous training fashions, advantage from statistical models and machine learning algorithms. (2)

Various complexities are linked with face detection algorithm that can be:

1. ***Quality of Image -***Face detection system is required a good quality image. Good quality images are collected under predictable circumstances. An image quality is necessary to extract the features from an image. When the calculations of features are not good then the robustness of the system will be lost.
2. ***Variations in Illumination -***Due to the lighting changes same faces images will be shown differently. The presence of an object can be changed due to variations in illumination.
3. ***Facial Expression -*** The person’s facial expressions are affected by the presence of faces.
4. ***Visual Angle -*** For different angles, the face images directly vary about the camera’s optical axis.

There are numerous algorithms thru which the face detection manner is achieved but, on this paper, the viola-jones set of rules is used for detecting the face from pix that is one of the maximum famous algorithms amongst all the face detection algorithms. Face detection additives discover or separate out human faces from the non-face objects present in a photograph. The image may be captured either in a controlled environment or in an uncontrolled environment [3] [5].

This paper is organized as: In section II the proposed face detection models is illustrated, in section III the viola-jones face detection algorithm is described, in section IV the viola jones face detection framework is explained, in V section the experimental results is shown and in VI section the performance evolution of the proposed algorithm is shown with the help of table.

**VIOLA-JONES FACE DETECTION ALGORITHM**

The viola-jones algorithm is used to detect the human face from an image. The system takes some face images or non-face images as input. After taking the input images the training phase will start in which the system detects the face. In the training phase, two types of sets are included that is positive image set or negative image set. In the positive image set, all the images are face images and in the negative image set, all the images are non-faces images. In the training phase, all the features are collected that are related to the face images, and all these features are stored in a file. After the training phase, the next phase is the testing phase. In the testing phase, all the stored features are applied to an input image and classified whether it is a face or not. If the image passes all the thresholds, then it is classified as face otherwise the image is classified as non-face [4] [13]. The Viola-Jones algorithm has four stages, and they are:

* + 1. *Haar feature selection*
    2. *An integral image*
    3. *Ad boost training*
    4. *Cascading classifiers*

1. ***Haar Feature Selection***

All human faces share some similar properties. These regularities may be matched using Haar Features [12]. These are:

* + - Location and size: eyes, mouth, bridge of nose
    - Value: oriented gradients of pixel intensities
    - The four features matched by this algorithm are then required in the image of a face. Rectangle features:

A picture containing logo

Description automatically generated

**Figure 1:** Different type of features

* + - Value = Σ (pixels in black area) - Σ (pixels in white area)
    - Various types of features like two, three and four-rectangles, Viola & Jones used two-rectangle features.
    - Difference in brightness between the white and black rectangles over a specific area.
    - Each feature is related to a special location in the sub-window [14].
  1. ***An Integral Image:***

In the subsequent step of the Viola-Jones face detection algorithm is rotate the input image into an integral image. This is completed by creation of every pixel equivalent to the total addition of all pixels above and to the left of the pixel. This is established in Figure 2.

|  |  |  |
| --- | --- | --- |
| 1 | 1 | 1 |
| 1 | 1 | 1 |
| 1 | 1 | 1 |

|  |  |  |
| --- | --- | --- |
| 1 | 2 | 3 |
| 2 | 4 | 6 |
| 3 | 6 | 9 |

**Figure 2:** The Integral Image

This makes the computation of the addition to the entire pixels within any specified rectangle using only four values [6] [7]. In the integral image, these values are the pixels that correspond with the corners of the rectangle in the input image. This is established in figure 3.

Graphical user interface

Description automatically generated with medium confidence

**Figure 3:** Sum Calculations

* 1. ***Adaboost Training:***

Viola Jones algorithm use a 24x24 window as the base window size to begin evaluating all the features in any given image. If we think about all feasible parameters of the Haar features like situation, degree, and type, then we need to calculate about 160,000+ features in any given window. By using this algorithm, we need to evaluate huge sets of features for every 24x24 sub-window in any new image. The basic idea is to eliminate a lot of features which are redundant and not useful. To select only those features that is very useful for us, which are done by Adaboost. Adaboost eliminate all the redundant features [11].

Adaboost is a machine learning algorithm which helps in judgment only the most excellent features between the entire those 160,000+ features. After these features are establish a weighted arrangement of all these features in used in evaluating and deciding any given window has face or not. These features are also called as weak classifiers.

A major component of the modified Adaboost algorithm is the determination of the most excellent feature and threshold. There seem to be no smart solution to this problem and Viola-Jones suggest a simple brute force method. This means that the determination of every latest weak classifier involves evaluating each feature on all the training examples to find the best performing feature. This is estimated to be the most time-consuming part of the training method [7].

* 1. ***Cascading Classifiers***

The cascaded classifier is collection of stages that contains a strong classifier. The work of every phase is to verify whether a particular sub-window is not a face or may be a face. When a sub-window is classified to be a non-face by a given phase it is discarded. A sub-window classified as a may be face is passed on to the next stage in the cascade. It follows that the additional stages a given sub-window passes, the higher chance that the sub-window really contains a face [15].

Diagram

Description automatically generated

**Figure 4:** The cascade classifier

**VIOLA-JONES FACE DETECTION FRAMEWORK**

The viola-jones face detection framework is the primary face detection structure to give competitive face detection charges in real-time planned in 2001. It was forced mainly by the difficulty of face detection, while it can be trained to identify a multiplicity of object classes. This algorithm is implemented in Open CV as cvHaarDetectObjects(). Viola Jones detector become well-known due to its open-source implementation in the Open CV library. To locate an object of an unidentified size is normally adopted to work this field that possesses a high competence and accuracy to locate the face region in an image [8].

1. Viola-jones face detection algorithm Face detection techniques can be divided into two main groups that are feature based approaches and image-based approaches. Image based approaches make use of linear subspace scheme, neural networks, and statistical approaches for face detection. Feature based approaches can be subdivided into low level analysis, feature analysis and active shape model. Face detection is restricted by trained scanning window classifiers. Viola-Jones Face Detection Algorithm is the first real-time face detection system [9].
2. Viola- Jones upper body detection Correct upper body detection improves the strength and reduces the difficult task of detecting upper bodies from images. The Viola-Jones algorithm is using the cascade object detector to detect people's upper body. This model detects the upper-body area, which is distinct as the head and shoulders area. This model uses Haar features to encode the fine points of the head and shoulder area.
3. Viola-jones face detection algorithm Eyes are detected based on the assumption that they are darker than additional part of the face, finding eye analogue segments searching small patches in the input image that are roughly as large as an eye and are darker than their neighborhoods. To throw away regions equivalent to eyebrows, the model uses the information that the middle part of an eye region is darker than other parts. Then a straightforward histogram examination of the region is complete for selecting eye regions [10].
4. Viola-Jones mouth detection algorithm This model is collected from weak classifiers, based on a result stump, which uses Haar features to encode mouth facts. Investigational outcomes prove that the algorithm is face image partition based on physical estimation of position of eyes, nose and mouth on face and can locate away the mouth region quickly [15].

**EXPERIMENT & RESULTS**

For the experiment we mainly using *‘FDetect = vision.CascadeObjectDetector;*’ This tool from the MATLAB. Here we are use this Image as example. In this paper we have created our own dataset. We have chosen face detection for particularly mouth, noise, eyes, face. For programming we have used MATLAB 2020a. The result of detection components of image are as follows:

A picture containing person, outdoor, person, lady

Description automatically generated

* **Face Detection**

Text

Description automatically generated

As you can see in the program, we are using bbx to detect and rectangle the face.

Graphical user interface, application

Description automatically generated

* **Nose Detection**

Graphical user interface, text, application

Description automatically generated

A picture containing text, person, person, screenshot

Description automatically generated

* **Mouth Detection**

Graphical user interface, text

Description automatically generated

A person with blonde hair

Description automatically generated with low confidence

* **Eyes Detection**

Graphical user interface, text

Description automatically generated with medium confidence

A person wearing glasses

Description automatically generated with low confidence

* **Multi Face Detection**

Also, the same program can work on multi face image. Program gave almost 98% accurate result.

A picture containing text, bunch, many, different

Description automatically generated

* **Web Cam based face Detection:**

For the web cam face detection, we were not able to detects faces with Viola-Jonas Algorithm. So, for the Webcam we have decide to use different method. We used the MultiObjectTrackerKLT function of multiple objects tracker using Kanade-Lucas-Tomasi (KLT) algorithm. Using the tracking ability tracker, we can track and detect faces in dark room and face with mask on.

*faceDetector=vision.CascadeObjectDetector ()*

*tracker = MultiObjectTrackerKLT()*

MultiObjectTrackerKLT properties:

* PointTracker- a vision Point Tracker object
* Bboxes- object bounding boxes.
* BoxIds- ids associated with each bounding box.
* Points- tracked points from all objects.
* PointIds- ids associated with each point.
* NextId- the next object will have this id.
* BoxScores- and indicator of whether or not an object is lost.

MultiObjectTrackerKLT methods:

* addDetections- add detected bounding boxes.
* track- track the objects.

Outputs for webcam-based face detection:

* Face detection under light:

A picture containing text, person, wall, indoor

Description automatically generated

* Face detection in dark:

A picture containing text, display, picture frame

Description automatically generated

* Face detection with Mask:

Graphical user interface, application

Description automatically generated

* Multi face detection:

A picture containing text, wall

Description automatically generated

**CONCLUSION**

In this paper, we have explored face detection in various situations. After this research we have concluded that, Viola-Jonas Algorithm works well while detection of mouth, eye and noise. Face detection is not accurate for noisy detection. Viola Jonas is very reliable if we put correct threshold value in it. Greater the threshold, higher the accuracy. Overall, Viola Jones method is by far widely used and efficient method for detecting a face. The same method can be used foe detecting the other objects.

**REFRENCES**

[1] L. Zhi-fang, Y. Zhi-sheng, A.K.Jain and W. Yun-qiong, 2003, “Face Detection And Facial Feature Extraction In Color Image”, Proc. The Fifth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA’03), pp.126-130, Xi’an, China.

[2] Prof.Dr. AbdulkadirErden, "Development of A Face Recognition System", July 2011.

[3] K. Seo, W. Kim, C. Oh and J. Lee, 2002, “Face Detection And Facial Feature Extraction Using Color Snake”, Proc. ISIE 2002 - 2002 IEEE International Symposium on Industrial Electronics, pp.457-462, L 'Aquila, Italy.

[4] J. Ruan and J. Yin, 2009, “Face Detection Based On Facial Features And Linear Support Vector Machines”, Proc. 2009 International Conference on Communication Software and Networks, pp.371-375, Macau, China.

[5] M. A. Berbar, H. M. Kelash and A. A. Kandeel, 2006, “Faces And Facial Features Detection In Color Images”, Proc. Geometric Modeling and Imaging― New Trends (GMAI'06), pp.209-214, London, UK.

[6] S. Kherchaoui and A. Houacine, 2010, “Face Detection Based On A Model Of The Skin Color With Constraints And Template Matching”, Proc. 2010 International Conference on Machine and Web Intelligence, pp. 469 - 472, Algiers, Algeria.

[7] Qiang-rong, Jiang, and Li Hua-lan. "Robust human face detection in complicated color images." Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on. IEEE, 2010.

[8] Das, Akanksha, Ravi Kant Kumar, and DakshinaRanjanKisku. "Heterogeneous Face Detection." Proceedings of the International Conference on Internet of things and Cloud Computing. ACM, 2016.

[9] Lander, Eric S., et al. "Initial sequencing and analysis of the human genome." Nature 409.6822 (2001): 860-921.

[10] C. Lin, 2005, “Face Detection By Color And Multilayer Feedforward Neural Network”, Proc. 2005 IEEE International Conference on Information Acquisition, pp.518-523, Hong Kong and Macau, China.

[11] S. Kherchaoui and A. Houacine, 2010, “Face Detection Based On A Model Of The Skin Color With Constraints And Template Matching”, Proc. 2010 International Conference on Machine and Web Intelligence, pp. 469 - 472, Algiers, Algeria.

[12] M. I. Razzak, M. K. Khan, K. Alghathbar and R. Yousaf, 2010, “Face Recognition Using Layered Linear Discriminant Analysis And Small Subspace”, Proc. 2010 10th IEEE International Conference on Computer and Information Technology (CIT 2010), pp.1407-1412, West Yorkshire, UK.

[13] D.N Pritha, L. Savitha and S.S. Shylaja, 2010, “Face Recognition By Feedforward Neural Network Using Laplacian Of Gaussian Filter And Singular Value Decomposition”, Proc. 2010 First International Conference on Integrated Intelligent Computing, pp.56-61, Bangalore, India.

[14] X. Wang, Q. Ruan and Y. Ming, 2010, “3D Face Recognition Using Corresponding Point Direction Measure And Depth Local Features”, Proc. ICSP 2010, pp.86-89, Beijing, China.

[15] Setu, Tania Akter, and Mijanur Rahman. "Human Face Detection and Segmentation of Facial Feature Region." Global Journal of Computer Science and Technology 16.1 (2016).