Comparison Between Viola-Jones and KLT Algorithms and Error Correction of Viola-Jones algorithm

Mohammad Ashraful Islam[1] Md. Anin Naeem[2], Md. Nazmul Hasan[3]

[1] *Lecturer, Department of Computer Science and Engineering, Mawlana Bhashani Science & Technology University, Bangladesh*

[2] *Student, Department of Computer Science and Engineering, Mawlana Bhashani Science & Technology University, Bangladesh*

[1]*shovon.415@gmail.com*

ABSTRACT:

Abstract— Face detection is the first step of automated face recognition process in all pattern matching or artificial intelligence or robotic vision system. There are many face detection algorithms and they have their own strengths and limitations in terms of speed and accuracy in different situations. Viola-Jones and KLT are two well known algorithms used for real time face detection and they work differently. Viola-Jones algorithm is much faster even though it is proved to be erroneous in some specific circumstances. KLT is basically used for tracking by points feature in frames. In this paper, we have given an overview of these two algorithms and made comparative analysis amongst them. Later we have proposed a method that improves the accuracy of Viola-Jones algorithm by removing/minimizing false detection.

Keywords: Face Detection, Viola-Jones algorithm, KLT algorithm, Error Correction, Individual detection.

[1]INTRODUCTION

Automated face detection[2][8] and recognition in real time system has become an essence of time due to its extensive use in robotic vision, surveillance system, investigation/security system, facial recognition system [10], identity verification system, pattern recognition, computer vision, human computer interaction, image processing and may others. Fast and error prune face detection algorithm in real time system is the prime most requirements for real time face recognition.

Real time face detection[11] is very much complicated computer vision problem due to different alterable situation arises, such as: position of a face, angle of a face with respect to camera, distance between face and camera, lightness/darkness, skin color, facial expression, hair, bread, glasses etc. Paul Viola and Michael J. Jones in their "Robust Real-Time Face Detection" showed a new algorithm that had significant accuracy in robust real-time system [9.0]. It was capable enough to process 15 frames per second with its 700 MHz Pentium III machine. Later there were a few proposed improvements on boosted cascade framework but in this paper we are proposing a method based on individual detection of different face parts.

[2]CONCEPTS

Our primary target of this paper is to remove the errors while detecting faces from a frame using Viola-Jones algorithm. We can distinguish between true detection and false detection but still there may be some faces that are not detected. Accuracy of a face detection algorithm depends on these three things. In real time, we will crop the all detected faces and will store them temporarily as a new image. We will try to identify individual face detection parts- left and right eye, mouth, nose within these individual images. If we can't find all these elements in a detected face, then we will discard that detection as a false detection.

This idea is implimented based on CascadeObjectDetector of RightEye, LeftEye, Mouth, Nose and FrontalFaceCART where every face parts are detected individually. The code requires Image Processing Toolbox and Computer Vision System Toolbox. This experiment was done with 4GB ram and 2.83 GHz machine.

[3]VIOLA-JONES ALGORITHM

The first object detection framework is Viola-Jones framework. This algorithm was developed by Paul Viola and Michael Jones and this algorithm is already put into effect in OpenCV in the act of cvHaarDetectObjects() which is used for basically face detection.

The main task is to detect all faces from a picture or image which a human can do easily but a device like computer or a robot always needs some information and compulsion. To do this task, Viola-Jones needs appropriate front view against the camera and the faces can't be bending to another sideways. Even though it detects frontal face properly but this algorithm seems to be vulnerable when face is bend at least 45^o or more. It is the main error or defect of this algorithm.

[4]STEPS OF VIOLA-JONES ALGORITHM

Haar[4] components is the first step which follows some basic similarities such as dark region and bright region, eye region and nose region where nose is supposed to produce more light than the eyes and eyes are supposed to produce less light than the cheeks. Location and size of eyes, mouth and nose and pixel volume of individual component is calculated. Haar is a square type/shaped component. Haar uses square box to find the pixel volume. Haar uses three types of components. They are-two square, three square and four square components.

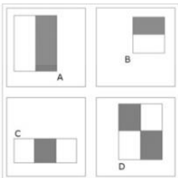

Figure:1. Haar Components

How the components are applied to a face is given below:

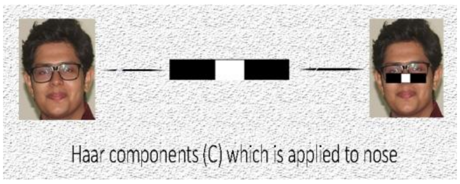

Figure:2. Nose

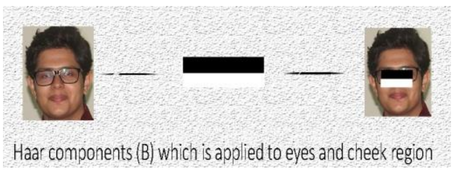

Figure:3. Eyes & Cheek

2

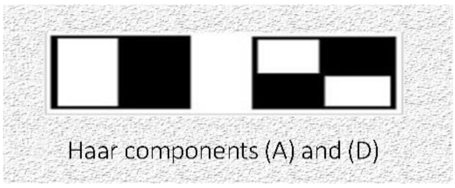

Figure:4. Others

Integral Image: Viola-Jones algorithm is so much faster because of integral image processing. Two square can execute six cluster source, any three square executes eight and four square can execute nine cluster source. And for this integral image, the process can avoid this section from counting the pixels which works with overhead and to the left side.

Adaboost training: This technique is used for finding the face and non-face from an image. It is the combination of weak classifier and density to create the strong classifier and it always discover individual square components and verge that always helps to differentiate the face and non-faces where non-faces are regarding as fault. After finding the fault, we try to calculate the best result and reducing the non-faces we get almost perfect score to finding the face.

Cascading: The last process is cascading which is used for boosting up the entire process to find the perfect outcome and this process is combining with some division which contains a strong classifier.

Example of Viola-Jones algorithm:

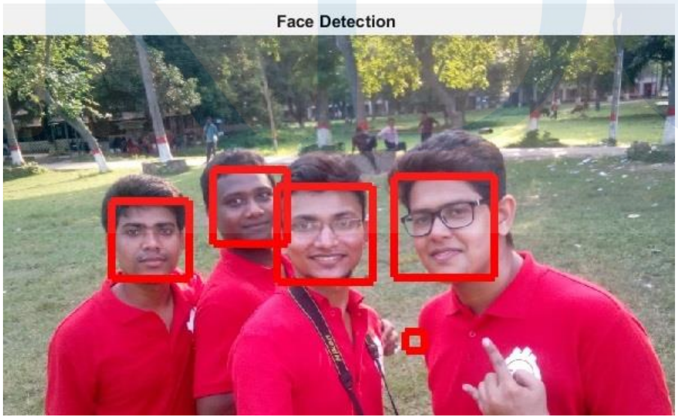

Figure:5. Detected faces(Viola-Jones algorithm)

[5]KLT ALGORITHM

Kanade-Lucas-Tomasi or KLT[1] algorithm is based on optical flow of point from frame to frame which is an access to component abstraction. This technique is planned because the common techniques are so much pricy and KLT is so much speedy than other techniques. This technique is basically used not for face detection but for face tracking continuously in a live streaming video through a camera or already in a saved video by calculating optical flow of a particular point from frame to frame. This technique can track multiple points from frame to frame where multiple points can be separated from each other. Moving objects can be tracked simply by tracking the particular points in a frame and finding the points in another frame. KLT algorithm is simply used for tracking the faces sequentially from frame to frame.

KLT algorithm measures the components or points of a frame and also detects the same points in another frame and then measures the movement of the points from previous frame to the present frame. This is a continuous process and KLT algorithm continuous to track the points until the end of the execution. This process is so much easier than the previous techniques. So the basic thing is detecting the selected points and measuring the movement.

Harris corners detection in the very first frame is the first work of this algorithm. After that KLT algorithm uses tracker to track the optical flow of the corners or points. This process is done by measuring the movement of the pixels. In every video, there is a moving object and this moving object has some specific points or pixels which can be easily detect by the tracker. KLT tracker can track each harries point and can also measure the movement of each pixels. Harries corners detection tracker is applied to every 10 to 15 frames to check the corners of every frames properly and this helps to find the points properly.

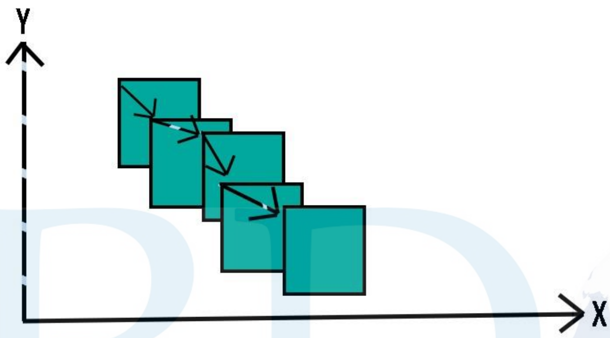

Figure:6. Movement of frames

Suppose (a,b) is the first corner point. The changeable vectors are m1, m2, m3, ……mn. So, the position of the first corner point in the next frame is: First corner point+ uprooted vector. By using Taylor series and hessian matrix we then find the position of the particular point in the next frame. That's how KLT algorithm works.

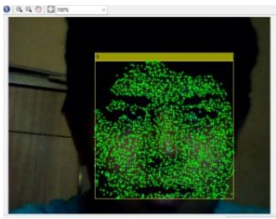

Figure:7. KLT algorithm(point tracking)

[6]COMPARISON BETWEEN VIOLA-JONES & KLT ALGORITHM

For finding the difference between Viola-Jones and KLT algorithm, we have to check these algorithms performance with some random test data. In all cases, both algorithms performed very well except one or two category. The comparison between these two algorithms is given below:

4

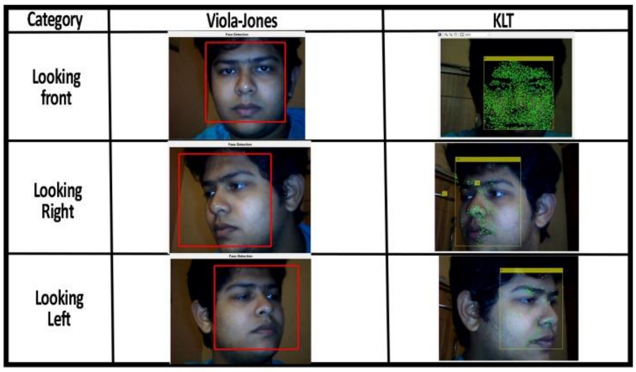

Figure:8. Comparison 01

Here, we can see that, looking front, right and left are performed very well by both algorithms, although Viola-Jones has some light issue and KLT has some glass issue.

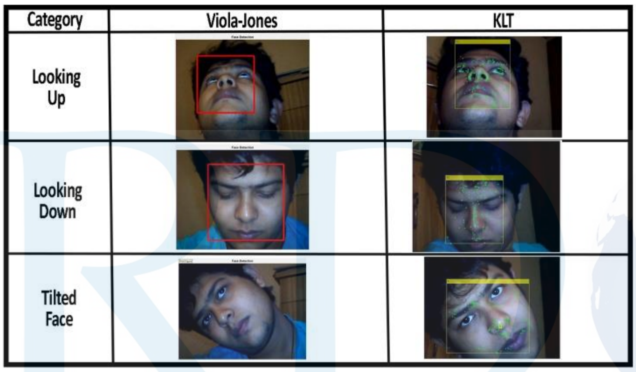

Figure:9. Comparison 02

In figure-09, we can see the big difference. Looking up and down are performed well by both algorithms but in tilted face detection, Viola-Jones algorithm failed to detect Where KLT detects tilted face properly. If the face tilted between 0^o to 45^o, Viola-Jones algorithm can detect that but if more than 45^o, it can't detect that face. It is the major difference between Viola-Jones and KLT algorithms. From figure-08 and figure-09, we can make the comparison [5] properly between these two algorithms [6] and can take the proper decision. Moreover, Viola-Jones is so much faster than KLT algorithms.

[7]ERROR CORRECTION OF VIOLA-JONES ALGORITHM

We have two main improvement scope with Voila-Jones algorithm.

1. Remove non-face detections/ false detections.
2. Detect tilted face.

We can say that, every face has five parts. They are- left eye, right eye, nose, mouth and full face in a whole. By detecting each part individually, we can detect faces easily. For doing this, we needed Image Processing Toolbox and Computer Vision System Toolbox. Computer Vision Toolbox has five detectors and they are – Face detector, Right eye detector, Left eye detector, Nose detector and Mouth detector. Each of them has a strong relation between them. Together they make a face. By detecting them together, we can detect a tilted face too.

So, we first detect a face by using cascade face detector which uses Viola-Jones method and then we cropped the face and saved it as an image. If there is a false detection, it will also be saved. Now, the saved images will be transferred to the individual face parts detectors. If the individual face parts are detected by the detectors of these images, then we can declare these images as different faces. The false detection will be discarded because detectors won't find any face parts there.

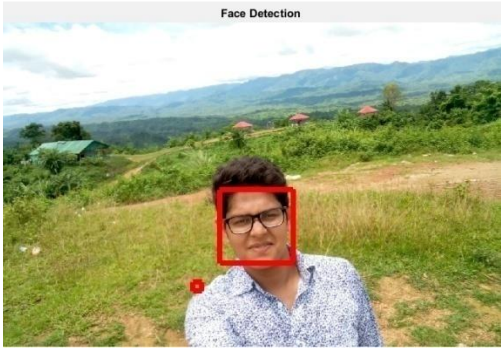

Figure:10. Face detection with error(Viola-Jones)

In this figure, we detected a face with a false detection using Viola-Jones method.

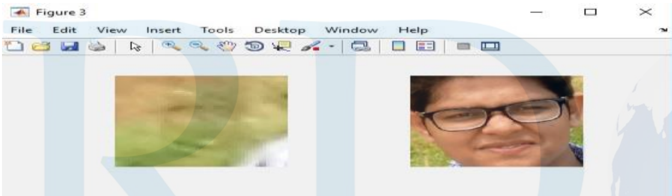

Figure:11. Cropped face with error.

Now we temporarily store these detected faces and transfer them to the individual face detector.

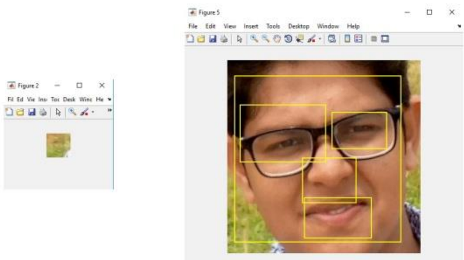

Figure:12. Individual Face Parts detected & false discarded

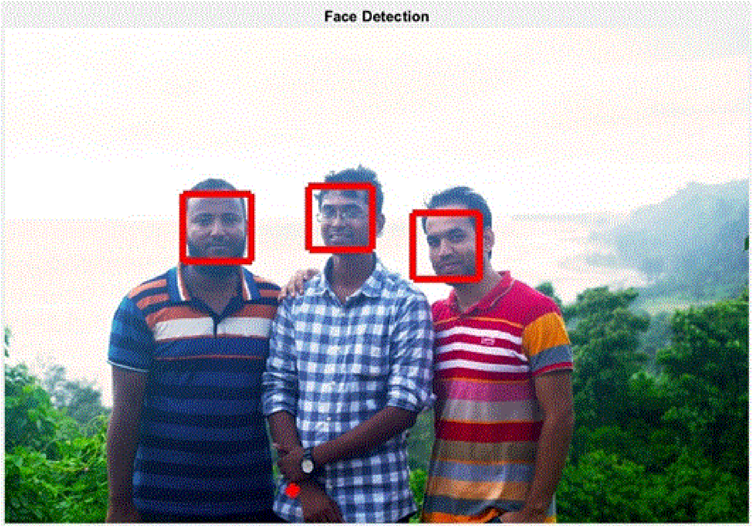

Figure:13. Face detection with false detection

Figure:14. Individual Face Parts detected & false discarded

Here, we can see that, the image with face is detected by the individual detectors and the false detection is discarded by the detectors. Thus, the errors can be corrected by using individual detectors. This individual face parts detector performs better than Viola-Jones and KLT algorithms but this detector's big lacking is-it's taking so much execution time for execution.

The full process is, we first build a detector which can detect individual face parts. Now pass an image through Viola-jones algorithm and detect faces. As a backend behind Voila-Jones, we crop detected face and send detected faces to the individual face parts detectors. If individual face detector fails to find individual face components, then we discard that face as false detection. The flow-chart of this individual detector is given below:

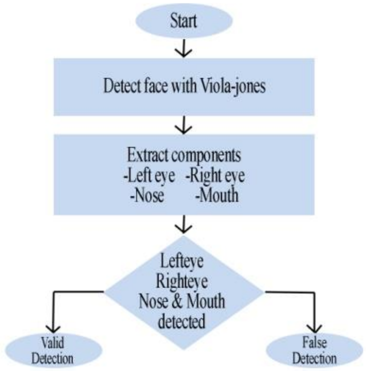

Figure 15: Flow-chart of the Individual Face Parts Detector

[8]RESULTS

Performance evaluation these three methods with some random test data are useful enough to show significant improvement of our method in terms of accuracy. We made the comparison by using some environmental variables. Here we used looking front, left, right, up, down and tilted environment. These environments helped to find the accuracy between these three methods. The Viola-Jones method has performed bad for tilted faces but the other methods are very useful here. Individual detector removes most of the problems of Viola-Jones methods. Here, Viola-Jones accuracy is 76.5%, KLT method's accuracy is 80.5% and Individual face parts detector's overall accuracy is 90.33%.

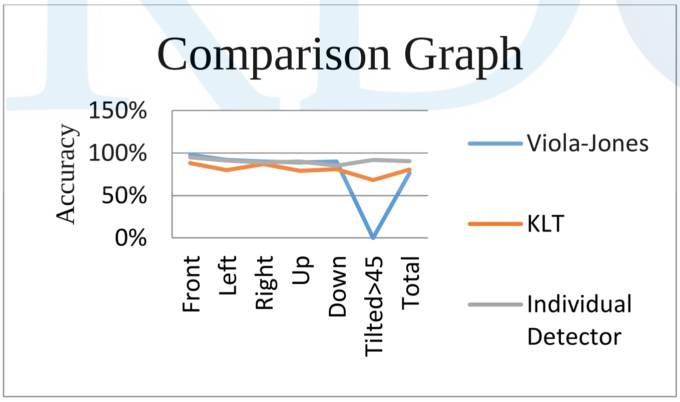

Figure:16. Comparison Chart

Now, a table is given here which shows how many faces are detected and how many faces are not detected and how many false detection are detected by these three methods.

No. of Images	Detection View	Viola-Jones	KLT	Individual face parts Detector
	No. of Faces Detected	51	47	73

	No. of Faces not Detected	30	34	8
	False Detection	12	9	1

Table:1. Comparison Table

We have worked with 50 images and this table shows the detection of faces and non faces and false detection for that 50 images. Here we see that, from 50 images Viola-Jones method detected 51 faces. This method couldn't detect 30 faces and there are 12 false detections. KLT method detects 47 faces and couldn't detect 34 faces and there are 9 false detections. And finally the Individual face parts detector detects 73 faces and couldn't detect 8 faces while we have got only single false detection.

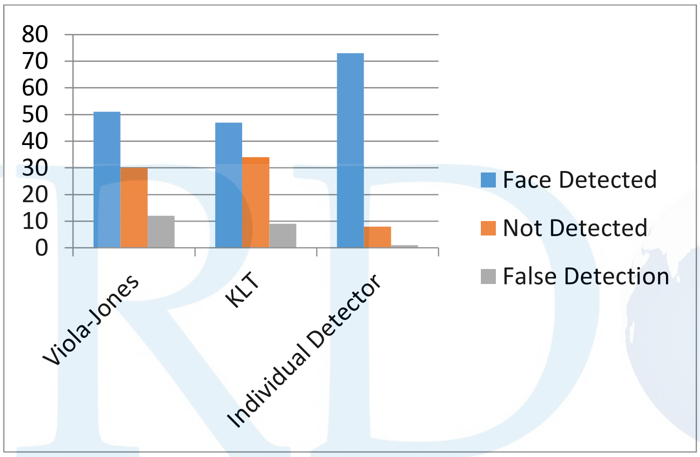

Figure 17: Graph based on Table 01

[9]CONCLUSION AND FUTURE WORK

From the comparison, we can see that Viola-Jones accuracy is 76.5% which can be improved by using the Individual face parts detector method which has the accuracy of 90.33%. This method removes most of the false detection of Viola-Jones. So this method will help the future researchers to build a proper method for face detection. There are some limitations too. This algorithm becomes very slow when there is 25/30 faces within a frame/image. One, improvement can be done using parallel processing for finding individual components of a face within detected.

REFERENCES

[1] J. Savitha and Dr. A.V. Senthil Kuman, “Face Tracking and Detection Using S-PCA & KLT Method”, 2014, International Journal of Advance Research in Computer Science and Management Studies, Volume 2.

[2] Deise Maia and Roque Trindade, “Face Detection and Recognition in Color Images under Matlab”, 2016, International Journal of Signal Processing, image Processing and Pattern Recognition.

[3] Yi-Shin Liu, Wai-Seng Ng and Chun-Wei Liu, “A Comparison of Different Face Recognition Algorithms”.

[4] Sander Soo, “Object detection using Haar-cascade Classifier”.

[5] Hemant Makwana & Taranpreet Singh, “Comparison of Different Algorithm for Face Recognition”, 2013, Global Journal of Computer Science and Technology Graphics & Vision(GJCSTGV).

[6] Prof. Rushabh Shah, Aalisha Sheth, Biyanta Shah, Aayushi Shah, Dr. Priyanka Sharma, “Comparative Study and Analysis of Face Detection Algorithms”, 2015, International Journal of Advanced Technology in Engineering and Science(IJATES).

[7] Monali Chaudari, Shanta sondur and Gauresh Vanjare, “A review on Face Detection and Study of Viola Jones method”, July 2015, International Journal Of Computer Trends and Technology(IJCTT).

[8] Mrs. Sunita Roy and Mr. Susanta Podder, “Face Detection and its applications”, 2013, International Journal of Research in Engineering & Advanced Technology (IJREAT).

[9.0] Paul Voila and Michael J. Jones, “Robust Real-Time Face Detection” , International Journal of Computer Vision 57(2), 137–154, 2004

[10] Yongzhong Lu, Jingli ZHOU and Shengsheng YU, “A Survey of Face Detection, Extraction and Recognition”, 2003, Computing and Informatics.

[11] Yi-Qing Wang , “An analysis of the Viola-Jones Face Detection Algorithm”, 2013, Image processing On Line.