

COMPARATIVE STUDY AND ANALYSIS OF FACE DETECTION ALGORITHMS

Prof. Rushabh Shah¹, Aalisha Sheth², Biyanta Shah³

Aayushi Shah⁴, Dr. Priyanka Sharma⁵

^{1,2,3,4} *Department of CSE, Institute of Technology, Nirma University, (India)*

⁵ *Professor, Raksha Shakti University, Gujarat, (India)*

ABSTRACT

Automatic facial detection of various features of the face is an important stage for face recognition. Recognizing different facial features such as nose, lips, eye corners, jaws in different illumination and inclinations is a difficult task. Face detection is the process of detecting the location and size of faces. In this paper, we will first explain and implement the Viola Jones face detector. We will then survey various techniques, which are used to detect and track faces such as Kanade-Lucas-Tomasi (KLT) and CAM-shift. Comparison of different algorithms and doing analysis of it will give researchers a better understanding of this field.

Keywords: CAM Shift algorithm, Eigen values, Face detection, Face Recognition, Image Processing, ROI region of interest, VIOLA JONES Face Detector

I INTRODUCTION

With the rapid increase of computational powers and availability of modern sensing, analysis technologies, computers are becoming more and more intelligent. Many research projects and commercial products have demonstrated the capability for a computer to interact with human in a natural way by looking at people through cameras, listening to people through microphones, understanding these inputs, and reacting to people [1]. One of the fundamental techniques that enable such natural human-computer interaction (HCI) is face detection. Face detection is the step stone to all facial analysis algorithms, including face alignment, face modeling face recognition, face verification, facial expression recognition, and many more.

Face detection is the process where the face is processed and matched bit wise with the underlying face image in the database. Face detection is also the psychological process with which we locate the faces in a visual scene. The ability to detect faces by a computer is affected by the orientation of the image and the visual color. Face detection is basically a special case of object class detection where objects of all sizes and at all locations are to be found in an image and that object belongs to a specific class. Face detection algorithms focus on the frontal detection of human faces [1]. It is similar to image detection in which the image is matched bit wise. Even if some facial feature is modified then the system will fail to detect the feature as similar to the previous one.

Face detection can also be applied in various fields especially in biometrics with face recognition, marketing and photography. Biometric based techniques are always more preferred for authorization than classical techniques like password access, PIN access, Smart cards, tokens etc. PIN and passwords are difficult to remember and can be stolen for misuse.

This is a general idea of face detection. Our paper will focus on the seminal Viola-Jones face detector, Kanade-Lucas-Tomasi and CAM-shift algorithm, face detection methods of these algorithms and their comparison. The flow of the paper goes as follows. Following the introduction, in section II we discuss the Viola-Jones face detector. In section III we cover the Kanade-Lucas-Tomasi face detection and tracking algorithm. In section IV we have the discussion on CAM-shift algorithm. Section V contains comparison between the algorithms discussed in the above sections. The paper ends with the acknowledgments and references.

II VIOLA JONES FACE DETECTOR

Viola Jones face detector is proposed by Paul Viola and Michael Jones in 2001.

The main characteristics of Viola-Jones are: [2]

- 1) *Robustness: It has a very high detection rate for frontal face imaging.*
- 2) *Real time: For real time applications we require processing speed of 2 frames per second which is workable in Viola-Jones.*

One of the limitations of this algorithm is that it only detects full view frontal upright images. That is, in order to be detected the entire face must point towards the camera and must not be tilted towards any side.

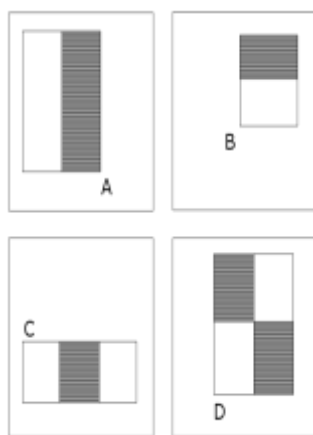


Fig 1: feature types

The features employed by the detection framework universally involves the sum of image pixels within rectangular areas. The value of any given feature is always simply the sum of the pixels within clear rectangles subtracted from the sum of the pixels within shaded rectangles.

Viola Jones works on the fact that all human faces share the similar type of information. There are certain properties which are common for the human faces.

- The nose bridge region is lighter than the eye region.
- Eyes are darker than the upper jaw region.

The value of dark or light is applied to each Harr- feature which helps us in detection of the facial features. The rectangle features that are obtained is by the following formula [2]:

Value: Σ (Pixels in dark area) – Σ (Pixels in brighter area)

Each of these features are related to a special location in these sub-windows. 'cascadeobjectdetector()' from the vision package of MATLAB is used to detect the face from the image. The function 'cascadeObjectDetector()' uses the basic Haar Features and Cascading to detect the facial organs. The following images are the results of face detection using Viola-Jones.

From the below results we can see that Viola-Jones aptly detects frontal face images, but fails to detect tilted, side or rotated faces. Also, all the features of the face are not detected everytime. They vary according to the orientation and contrast of the face and the background.

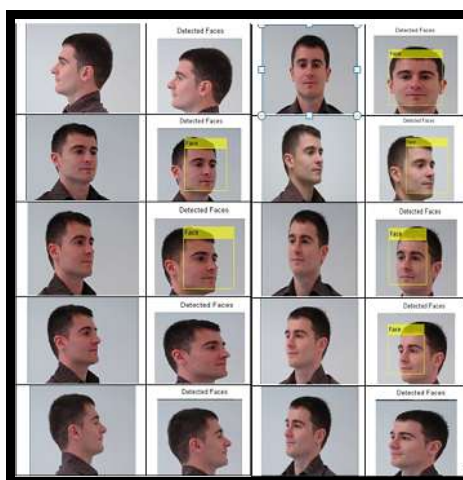


Fig 2: Results of Viola- Jones face detector [8]

III KANADE LUCAS TOMASI ALGORITHM

The KLT tracking algorithm [3] computes displacement of features or interest points between consecutive video frames when the image brightness constancy constraint is satisfied and image motion is fairly small. Assuming a local translational model between subsequent video frames, the displacement of a feature is computed using Newton's method to minimize the sum of squared distances within a tracking window around the feature position in the two images [4].

In KLT algorithm, the face is detected using the 'cascadeobjectdetector()' object. While it is possible to use the cascade object detector on every frame, it is computationally expensive. It may also fail to detect the face whenever the subject turns or tilts his head.

To solve this issue, we track a set of feature points across the video frames. Once the face is detected, it uses the same feature points to track down the image if at all it moves. The feature points are tracked using 'detectMinEigenFeatures()'. Once each point is tracked in the previous frame, the point tracker uses those values to find the corresponding point in the current frame.

The KLT algorithm is used to detect and track the face from a running video. The Below images show the features detected by the 'detectMinEigenFeatures()' function and these values are carried further to track the tilted, side face or rotated face.

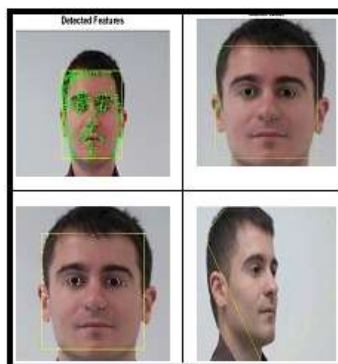


Fig 3: Results of Kanade-Lucas-Tomasi [8]

IV. CAM-SHIFT ALGORITHM

CAM-shift is the abbreviated form of Continuously Adaptive Mean shift algorithm. In CAM-shift algorithm, a back projection image is obtained from the color histogram model of each image frame, the size of searching window is adjusted adaptively on the object's size, and the center position of window is computed iteratively. By extending the searching window's size adaptively, the issue of the object losing instantaneously from the current window for motion acceleration is solved [5].

The face is detected using the cascade object. The cascade detector object uses the Viola-Jones algorithm which by default detects faces which are upright frontal images. Once the face is located the next step will be to identify a single feature which will help to locate the face in the video. Feature like texture or hue of the skin can be chosen. In our implementation we have chosen 'skin' to track down the features. The skin tone will not change as the face changes its orientation and thus is a reliable parameter to track down features of the face and detect it in the video.

We get the skin tone information by extracting the hue from the frame and converting it to HSV color space. With the skin tone as the selected parameter, we use the 'vision.HistogramBasedTracker' for tracking. The histogram-based tracker uses CAM-shift algorithm, which provides the capability to track an object using a histogram of pixel values [6]. Here the Hue channel pixels are extracted from the nose region of the detected face. This is the region of interest for us [7]. These are used to initialize the histogram.

There are two things we need to take care of while implementing CAM-shift algorithm. The first is that if the contrast between the skin tone of the face and the background is not enough then the face will not be detected properly. If the face is not detected for the first time then it will not be detected during the entire video.

The second point will be to ensure that the starting of the video consists of a human face. If the starting frame consists of text or other kinds of graphics, then the detector object will be unable to detect the face in the rest of the video because it will not receive the points, which will help initializing the histogram.



Fig 4: The first frame where face is detected

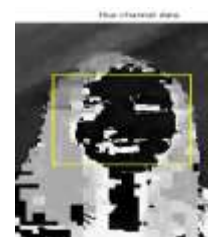


Fig 5: For skin tone, the frame is converted to HSV color

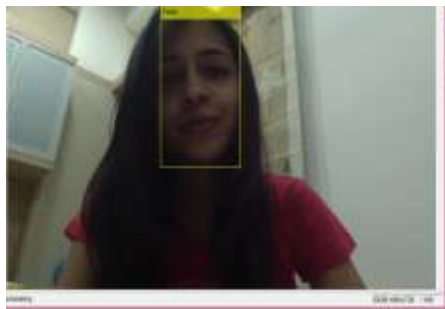


Fig 6: The face being detected in all the video frames it appears in

V COMPARISON

The Viola Jones algorithm uses a series of steps for detection of the faces, but the disadvantage here lies that only frontal faces are detected. The tilted and rotated faces are not detected; also if the color tone, size or brightness is not good enough the frontal faces don't get detected. The problem also lies that the facial organs are not always correctly identified using the 'CascadeObjectDetector()' . The below images show the issues discussed above for the Viola Jones Algorithm.



Fig 7: Flaws of Viola Jones Algorithm [8]

This Algorithm is best used when the frontal face is clearly visible and recognized, along with all factors appropriate.

Kanade-Lucas-Tomasi (KLT) and CAM-shift algorithms both are used for face detection and feature tracking from the input video files. As discussed in the above sections, both use Viola Jones for face detection. Kanade-Lucas-Tomasi calculates Eigen vectors that help in detecting the face throughout the video whereas CAM-shift uses the skin tone as a medium here and converts the detected face image into the Hue channel data and plots it on the Histogram.

One of the major flaws is that if in the initial video frame, a frontal face is not detected, then both algorithms fail to spot any face in the entire video.

In Kanade-Lucas-Tomasi, if we have a slightly dark image, or an image with less contrast then it will fail to detect Eigen values on the face, which will then not detect tilted or rotated faces ahead in the video stream.



Fig 8: Lack of Detected Eigen Values

In CAM-Shift algorithm, the face will not be detected only if there is very high contrast. Once the face is detected, the algorithm converts it into hue channel data, which transforms it into an inverted grayscale image. Thus if there is less contrast the inverted image which will be the same, thus failing to detect the face along the video as it goes ahead.

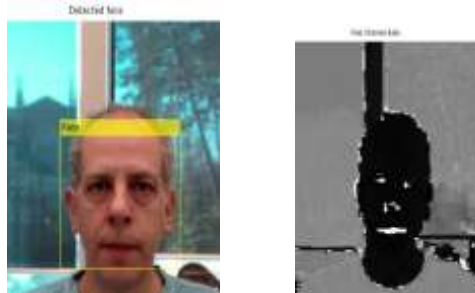


Fig 9: Very High Contrast thus no face detected [9]

VI CONCLUSION

In this paper we conclude that, Viola Jones cannot detect tilted, rotated or side faces. KLT and Cam-shift will not detect a face in the entire video if there is an absence of face in the first frame. KLT detects face based on the eigenvectors detected from the first frame and has difficulty in detecting a face having low contrast with the background. Cam – Shift detects face by using a feature to track them, that feature can be skin tone or contrast background or any feature of the face. Here, we have used skin tone as a feature and converted the first frame to hue channel data and then detects the face in the rest of the video.

VII ACKNOWLEDGEMENTS

We would like to thank our mentor, Assistant Professor Rushabh Shah who mentored us during our paper and guided us with his experience and useful comments.

REFERENCES

- [1] Cha Zhang and Zhengyou Zhang, June 2010, “ A Survey of Recent Advances in Face Detection”, Available at: <https://www.research.microsoft.com>.
- [2] Nikolay Degtyarev and Oleg Seredin, “Comparative Testing of Face Detection Algorithms”, Available at: <https://lda.tsu.tula.ru/papers/degtyarev-2010-icisp-ctfd.pdf>.
- [3] Sudipta N Sinha et al., May 2006 “GPU-Based Video Feature Tracking and Matching”, Available at: <https://cs.unc.edu>.
- [4] Carlo Tomasi and Takeo Kanade, April 1991, “Detection and Tracking of Point Features”, Available at: <https://www.csie.ntu.edu.tw>
- [5] Sun Kai and Liu Shi-rong, “Combined Algorithm with Modified Camshift and Kalman filter for Multi-Object Tracking”, Available at: <https://en.cnki.com.cn>
- [6] <https://in.mathworks.com/help/vision/examples/face-detection-and-tracking-using-camshift.html>. Last Viewed on January 23, 2015

[7] Shah, Rushabh, Priyanka Sharma, and Rutvi Shah. "Performance Analysis of Region of Interest Based Compression Method for Medical Images." Advanced Computing & Communication Technologies (ACCT), 2014 Fourth International Conference on. IEEE, 2014.

[8] <https://gtav.upc.edu/research-areas/face-database>. Last Viewed on January 20,2015 1

[9] https://d28rh4a8wq0iu5.cloudfront.net/images/recoded_videos Last viewed on January 19,2015