International Journal of Recent Engineering Research and Development (IJRERD)
ISSN: 2455-8761
www.ijrerd.com || Volume 03 – Issue 07 || July 2018 || PP. 117-122

Detection and Tracking of Human Faces with Image Processing and MATLAB

Nang Khin Su Yee[1], Hla Soe[2],
[1]*Technological University (Mandalay)*
Myanmar
[2]*Technological University (Kyaukse),*
Myanmar

Abstract: This paper describes how to detect and track the multiple human faces. Crimes are causing all over the world in daily life. Therefore, the detection and tracking the human faces are the most important role in nowadays. Although the different algorithms were used for detection and tracking, Viola-Jones algorithm was used in this research because it gives a very fast detection of human faces in video live. This algorithm is used in MATLAB R2014a to simulate for detection and tracking of multiple human faces and it can be also used on computer vision technique to reduce a lot of problem.
Keywords: Multiple faces, Detection, Tracking, Viola-Jones algorithm and MATLAB.

1. Introduction

Multiple human face detection and tracking systems are the most critical and fundamental step in video surveillance systems and then it can also be to improve and develop a system's performance in fields such as security, safety, human activity monitoring etc. Although the face recognition system can distinguish the different human faces and only indicates whether or not a face is present in an image, face tracking can determine the exact location of the face [1].

In this research, the Vioal – Jone algorithm was applied for detection and tracking the human face. There are three intermediate steps in this algorithm [6].The first is the development of a new image representation called "integral image". The second step deals with the construction of classifiers that helps us to segregate desired features from the set of large number of features using a technique called "AdaBoost".Third step deals with the cascading of different classifier [5][6].

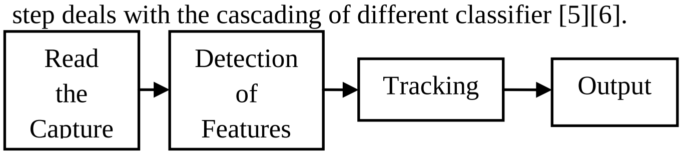


Figure 1. Block Diagram of Human Face Detection and Tracking

2. Viola – Jone Algorithm

Face detection system consists of mainly three key contributions:
- It introduced a new, efficient, method for features calculation, based on an integral image.
- It introduced a method for aggressive features selection, based on AdaBoost learning algorithm.
- It introduced an idea of combining classifiers in a cascade.

2.1. Integral Image

The first step of the Viola-Jones face detection Algorithm is to turn the input image into an integral image. Integral image, also known as a summed area table, is an algorithm for quickly and efficiently computing the sum of values in a rectangle subset of a pixel grid. In Figure 2, the integral image at location x; y contains the sum of the pixels from the original image above and to the left of x; y inclusive:

$$ii(x,y)=\sum_{x'\le x, y'\le y} i(x',y') \quad (1)$$

Where ii(x, y) an integral is image and i(x, y) is an original image.

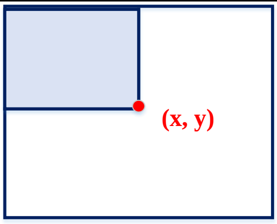


Figure 2: Integral Image

The integral image can be computed in a single pass over the original image by maintaining a cumulative row sum at each location x,y. Rectangle features can be calculated using only a few accesses to it see Figure 3. Two-rectangle features require six array references, three-rectangle features require eight array reference and four-rectangle features require nine array reference.

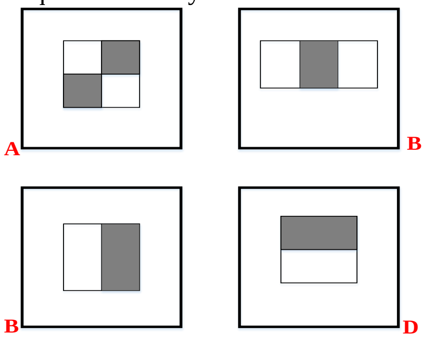


Figure 3: (A) Shows Four-rectangle Features, (B) Shows Three-Rectangle Feature, (C) and (D) Shows Two-rectangle Feature

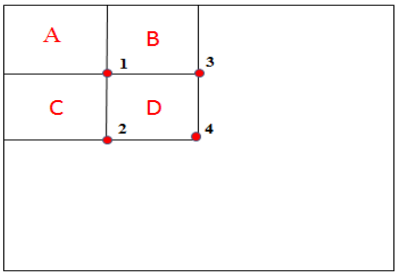


Figure 4: The sum of the pixels within rectangle D can be computed as 4 + 1 - (2 + 3), where 1 to 4 is values of the integral image

Integral image at location 1 is the sum of pixels in region A; at location 3 sum of pixels in region A+B, at location 2 C+A and at location 4 A+B+C+D Where ii(x,y) is an integral image and i(x',y') is the original image. Using the following formula,

$$r(x,y) = r(x,y-1) + i(x,y) \quad (2)$$
$$ii(x,y) = ii(x-1,y) + r(x,y) \quad (3)$$

Where r(x,y) is the cumulative row sum r(x,-1) = 0 and
ii (-1,y) = 0.

The detector, based on resolution is used 24x24. In others, every image frames divided into 24x24 sub-windows and features are extracted at all possible locations and scales for each such sub-window. This results in an exhaustive set of rectangle features which counts more than 160,000 features for a single sub-window.

2.2. AdaBoost Learning Algorithm

AdaBoost learning algorithm is used as a feature selection mechanism. AdaBoost is a learning classification function which when given a set of features and a training set of positive and negative images, to learn a classification function any number of machine learning approaches could be used. A classification result of a learning algorithm to improve is used by combining a collection of weak classifiers to form a strong classifier. The algorithm starts with equal weights for all examples. In each round, weights are updated so that the misclassified examples receive more weight. AdaBoost is used to train the classifiers as well as to select a small set of features. When used in its real form, AdaBoost learning function boosts the performance of a simple learning algorithm (sometimes called weak). AdaBoost has capability to achieve large margins rapidly which is one of the key features of this algorithm.

Each sub window consists of 160,000 features associated with it which is much larger than the number of pixels. Features can be computed very efficiently but computing the complete set is expensive. Minuscule number of these features can be combined to form an efficient classifier. Most importantly the challenge is to find these features. A weak learning algorithm is designed for development of a classifier that learns and adjusts by changing its threshold value. A single AdaBoost classifier consists of a weighted sum of many weak classifiers, where each weak classifier is a threshold on a single Haarlike rectangular feature. The weight associated with a given sample is adjusted based on whether or not the weak classifier correctly classifies the sample. A single weak classifier is defined as:

$$h(x, f, p, \theta) = \begin{cases} 1 & if \quad pf(x) < p\theta \\ 0, & otherwise \end{cases} \tag{4}$$

The weak learning algorithm was restricted to the set of classification functions, which of each was dependent on a single feature. A weak classifier h(x, f, p, θ) was then defined for a sample (i.e. 24x24 sub-window) by a feature f, a threshold θ and a polarity p indicating the direction of the inequality.Compute the strong classifier is :

$$H(x) = \begin{cases} 1, & \sum_{t=0}^{T} \alpha_t \quad h_t(x) \geq \gamma_t \\ 0, & otherwise \end{cases} \tag{5}$$

$$\alpha_t = \log \frac{1}{\beta_t} \gamma_t$$

where and t is chosen such that all positive training samples are correctly classified

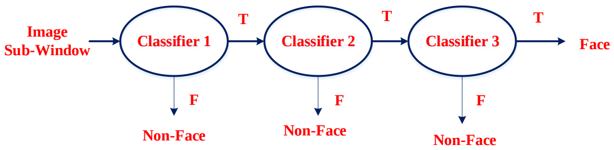


Figure 5: Computational Risk Minimization

2.3. Cascade Classifier

The cascaded classifier is composed of stages each containing a strong classifier from AdaBoost. The job of each stage is to determine whether a given sub-window is definitely not a face or maybe a face. When a sub-window is classified to be a non-face by a given stage it is immediately discarded. Conversely a sub-window classified as a may be face is passed on to the next stage in the cascade. It follows that the more stages a given sub-window passes, the higher the chance the sub-window contains a face.

The key features of this algorithm were increase performance detection while reducing the detection time. Simpler classifiers separate the positive and the negative from the image, hence are called weak classifiers or initial classifiers. Then additional features are added to other classifiers and hence are called complex or strong classifiers. Because strong classifiers complexity goes on increasing towards the final output. They take the output of the previous classifier as their input and perform either of the two operations, i.e. they either reject the window, or pass it on and trigger the next more complex classifier for further processing. This process is going on until the accurate detection is achieved. In other words, the sub window is rejected in the path. The sub window reaching the output has to pass all the intermediate classifiers in order to get detected.

AdaBoost classifier is used for constructing stages in the cascade. Its threshold is adjusted to minimize false negatives. In general, a lower threshold yields higher detection and higher false positive rates. For reducing computation time need to reduce the number of features. In this doing, decreases the accuracy. There is a tradeoff between the accuracy and speed. In order to overcome this problem, first stage classifier constructed with one feature to become strong classifier by reducing the threshold to minimize false negative. Adjustment of the threshold can be done to detect 100% of the faces with a false positive rate of 50% in figure 7. The second classifier goes through more difficult task then the first. The probability of face detection is increased by eliminating unwanted areas in early stages. More specific and detailed elimination is done in further stages for increasing the probability of face detection.

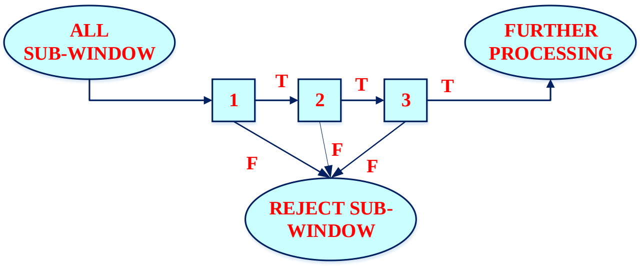


Figure 6: Detectors Cascade

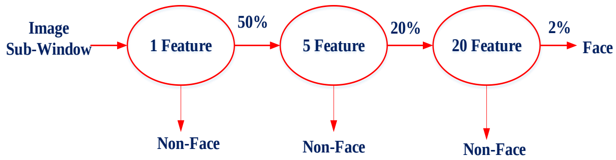


Figure 7: Every sub window is subjected to a series of classifiers

3. Algorithm of the Research

In this research, the 'Vision. Cascade Object Dector' called the Viola-Jones algorithm was used to detect human's faces. How to detect and track the human's face are shown in the following steps.

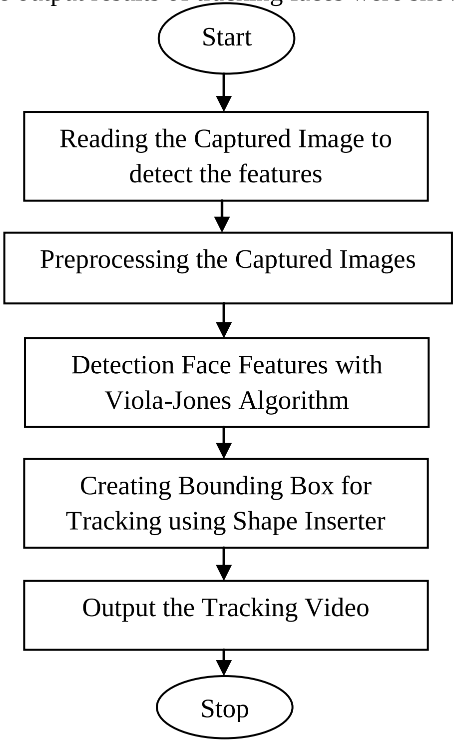
Step 1: Reading the captured input image to detect the human face.

Step 2: Pre-process the captured images to improve the image quality

Step 3: Imputing the captured image to Cascade-Object-Detector and then, it detects and tracks the people's face using Vision Cascade Object Detector.

Step 4: The bounding box for tracking the face was created.

Step 5: The output results of tracking faces were shown with the created bounding box.



la

Figure 8. Flow Chart for Real Time Face Detection and Tracking Algorithm

4. Tests and Results

This paper approaches for detection and tracking of human face with reducing computation time while achieving high detection and tracking accuracy. Real time human face detection and tracking algorithm developed better than the previous versions in terms of computation time and feature extraction. The code had been developed for real time detection and tracking of human' face can run in i7 processor dell laptop. The results are shown in the following figures.

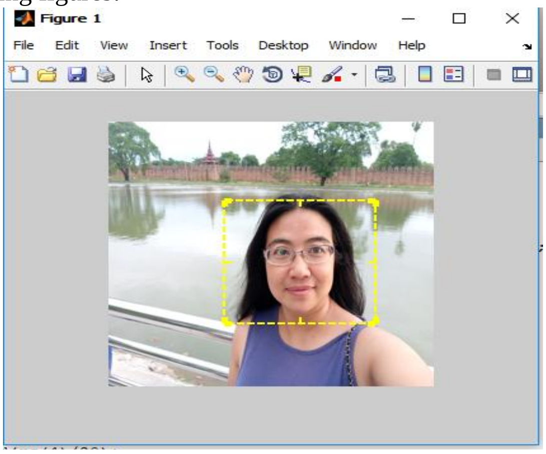


Figure 9: Shows Real Time Face Detection and Tracking of One of the Human's Face

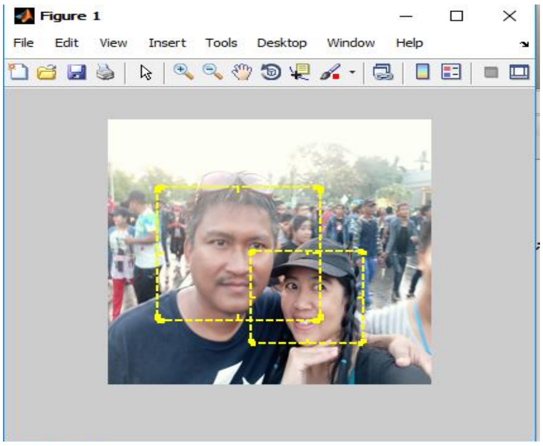


Figure 10: Shows Real Time Face Detection and Tracking of Two of the Human's Face

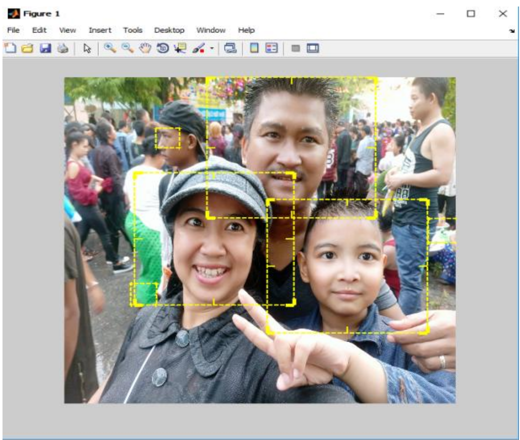


Figure 11: Shows Real Time Face Detection and Tracking of Multiple Face

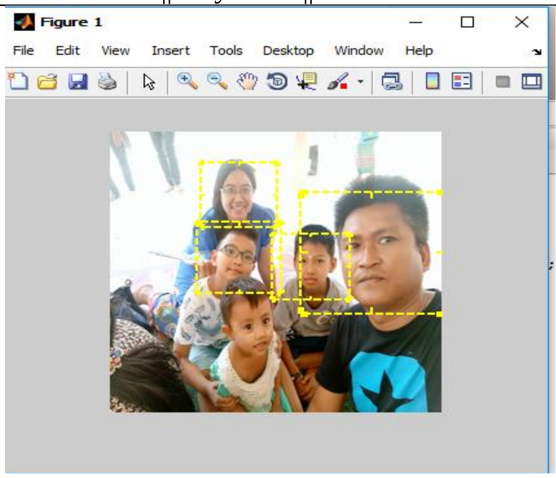


Figure 12: Shows Real Time Face Detection and Tracking of Multiple Face

5. Conclusion

In this paper, Viola-Jones face detection and tracking algorithm had been developed to detect and track. This method is a widely used method for detection of object. It mainly used broader applications such as computer vision region and image processing. In this paper, AdaBoost learning algorithm is used as a feature selection mechanism. Evaluating the entire set of features is still extremely expensive and cannot be performed by a real time application. This paper described experimental results on difficult human's face detection and tracking. Advantages of this method are that detection is very fast and extremely small of false positive. Main disadvantages of this algorithm is that it can only detect frontal, upright faces with some moderate deviation for in plane and out of plane rotations and training is slow.

References

[1]. J. Chatrath, P. Gupta, P. Ahuja, A. Goel and S. M.Arora, "Real Time Human Face Detection and Tracking", International Conference on Signal Processing and Integrated Networks (SPIN) C-4 Janak Puri NEW DELHI-58, INDIA, 2014.

[2]. C. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection", In International Conference Vision, 1998.

[3]. Robert E. Schapire, Yoav Freund, Peter Bartlett, and Wee Sun Lee. Boosting the margin: A new explanation for the effectiveness of voting methods. *In Proceedings of the Fourteenth International Conference on Machine Learning,* 1997.

[4]. Yoav Freund and Robert E. Schapire, " A decision-theoretic generalization of on-line learning and an application to boosting", In *Computational Learning Theory: Eurocolt 95*, pages 23-37. Springer-Verlag, 1995.

[5]. R. Makovetsky and G. Petrosyan, "Face Detection and Tracking with Web Camera", Fundamental of Computer Vision, COMP-558

[6]. Kaiqi Cen, "Study of Viola-Jones Real Time Face Detector", cenkaiqi @ gmail.com