Review of Literature Face Detection via Viola Jones

Mr. Pushpendra Singh[1], Dr. Shabana Urooj[2]
[1,2]Department of Electrical Engineering
[1,2]Gautam Buddha University Greater, Noida, Uttar Pradesh (India)

Abstract— Face recognition has been one of the most studied topics in the last decades. Its research has increased in many applications in communications and automatic access control systems. In this paper, a new face recognition method based on face detection technique is introduced. The goal of the propose method is to detect faces that exist in the image and trying to locate each face in a previously prepared database using simple variance calculations and Euclidean distance of extracted facial features. The features under consideration are eyes, nose and mouse. Furthermore, a new method to extract facial features is developed based on feature location with respect to face dimensions. The proposed algorithm has been tested on various images and its performance is found to be good in most cases. Experimental results show that our method of detection and verification achieves very encouraging results with good accuracy and simple computations.

Key words: Viola jones face detection algorithm, Adaboost algorithm for face detection, Haar cascade classifier

I. Introduction

Face detection is a part of face identification. When we see at the person's face, can get the information such as the expression, gender, age and ethnicity. Face detection is useful in many applications such as surveillance system, human machine interaction, biometrics, gender classification etc. For human beings face detection is an easy task but face detection is quite a tough task for a computer. A digital image is made up of finite number of elements each of which has a particular location and value. These elements are known as pixel and picture element. These elements take participation to find out the face. Face detection method can be broadly classified into two categories: Appearance based approach and feature based approach. In the appearance based approach, the whole image is used as a input to the face detector. In feature based approach face detection is based on the features extracted from an image. Features can be i.e. skin color or edges and sometimes they have knowledge of the face geometry [1]. The appearance based approach which we used in this paper has the potential to identify the face from an image using haar cascade classifier. To getting the detailed knowledge of the face detection can read [1] [2].

II. Methods of Detection

The following methods are generally used to detect the faces from a still image or a video sequence.

A. Viola Jones Face Detection Algorithm:

The Viola–Jones object detection framework [3] is the first object detection framework to provide competitive object detection rates in real-time proposed in 2001 by Paul Viola and Michael Jones. Even though it can be trained to detect a variety of object classes, it was motivated mainly by the problem of face detection. This face detection framework is capable of processing images extremely rapidly while achieving high detection rates. There are three key of assistance.

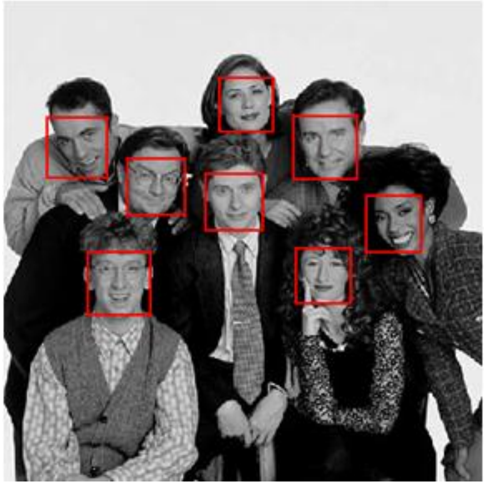

Fig. 1: Viola Jones face detection

Detection process by using viola – Jones algorithm.

The first is the introduction of a new image illustration called the Integral Image which allows the features used by our detector to be computed very quickly

The second is an easy and efficient classifier which is built using the AdaBoost learning algorithm to select a small number of critical visual features from a very large set of potential features.

The third contribution is a process for combining classifiers in a cascade which allows background regions of the image to be quickly discarded while spending more computation on promising face-like regions.

1) *Advantages:*

It is the most admired algorithms for face detection in real time.

The main advantage of this approach is uncompetitive detection speed while relatively high detection accuracy, comparable to much slower algorithms.

High accuracy. Viola Jones gives accurate face detection. Constructing a cascade of classifiers which totally reduces computation time while improving detection accuracy.

The Viola and Jones technique for face detection is an especially successful method as it has a very low false positive rate.

2) *Disadvantages:*

- Extremely long training time.
- Limited head poses.
- Not detect black Faces.

B. AdaBoost Algorithm for Face Detection

Boosting is an approach to machine learning based on the idea of creating a highly accurate prediction rule by combining many relatively weak and incorrect rules. The

AdaBoost algorithm was the first practical boosting algorithm, and one of the most widely used and studied, with applications in numerous field. Using boosting algorithm to train a classifier which is capable of processing images rapidly while having high detection rates. AdaBoost is a learning algorithm which produces a strong classifier by choosing visual features in a family of simple classifiers and combining them linearly. Although AdaBoost [4] is more resistant to over fitting than many machine learning algorithms, it is repeatedly sensitive to noisy data and outliers .AdaBoost is called adaptive because it uses multiple iterations to generate a single composite strong learner. AdaBoost creates the strong learner (a classifier that is well-correlated to the true classifier) by iteratively adding weak learners (a classifier that is only slightly correlated to the true classifier). Throughout each round of training, a new weak learner is added to the group and a weighting vector is adjusted to focus on examples that were misclassified in preceding rounds. The outcome is a classifier that has higher accuracy than the weak learners‘classifiers.

C. *What's Good About Adaboost:*

- Can be used with numerous different classifiers.
- Improves classification accuracy.
- Commonly used in many areas.
- Simple to implement.
- Not prone to overfitting.

D. *Advantages:*

No a prior knowledge**.** AdaBoost is an algorithm which only needs two inputs: a training dataset and a set of features (classification functions). There is no need to have any a priori knowledge about face structure.

1) *Adaptive algorithm*

At each stage of the learning, the positive and negative examples are tested by the current classifier. If an example is misclassified, i.e. it cannot clearly be assign in the good class. In order to increase the discriminant power of the classifier these misclassified examples are up-weighted for the next algorithm iterations.

The training errors theoretically converge exponentially towards 0. Given a finite set of positive and negative examples, the training error reaches 0 in a finite number of iterations.

Very simple to implement.

Do feature selection resulting in comparatively simple classifier Fast.

Simple and easy to Program.

No former knowledge needed about weak learner.

E. *Disadvantages:*

The result depends on the data and weak classifiers. The quality of the final detection depends highly on the consistence of the training set. Both the size of the sets and the interclass variability are important factors to take in account.

Quite slow training. At each iteration step, the algorithm tests all the features on all the examples which requires a computation time directly proportional to the size of the features and examples sets.

Weak classifiers too complex leads to overfitting. Weak classifiers too weak can lead to low margins, and can also lead to overfitting. Suboptimal solution. Sensitive to noisy data and outlier.

III. HAAR CASCADE CLASSIFIERS

The core basis for Haar classifier object detection is the Haar-like features. These features, rather than using the intensity values of a pixel, use the change in contrast values between adjacent rectangular groups of pixels. The contrast variances between the pixel groups are used to determine relative light and dark areas. Two or three adjacent groups with a relative contrast variance form a Haar-like feature. Haar-like features, as shown in Figure 1 are used to detect an image [5]. Haar features can easily be scaled by increasing or decreasing the size of the pixel group being examined. This allows features to be used to detect objects of various sizes.

A. *Integral Image*

The simple rectangular features of an image are calculated using an

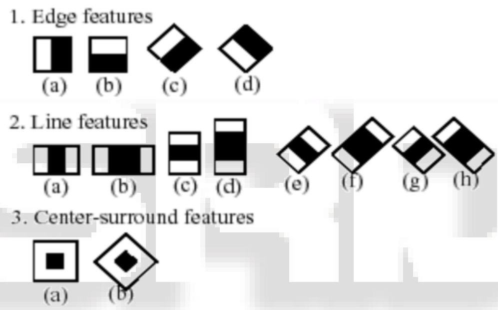

Fig. 2: Intermediate representation of an image, called the integralimage[6].

The integral image is an array containing the sums of the pixels' intensity values located directly to the left of a pixel and directly above the pixel at location (x, y) inclusive. So if A[x,y] is the original image and AI[x,y] is the integral image then the integral image is computed as shown in equation 1 and illustrated in Figure 2.

$$AI[x\ y]\ A\ x\ y\ (1) x\ x\ y\ y,\ (',\ ')'\ ,\ '=\leq\ \leq\Sigma$$

The features rotated by forty-five degrees, like the line feature shown in Figure 3, as introduced by Lienhart and Maydt, require another intermediate representation called the rotated integral image or rotated sum auxiliary image [7]. The rotated integral image is calculated by finding the sum of the pixels' intensity values that are located at a forty five degree angle to the left and above for the x value and below for the y value.

So if A[x,y] is the original image and AR[x,y] is the rotated integral image then the integral image is computed as shown in equation 2 an illustrated in Figure 3.

$$AR[x\ y]\ A\ x\ y\ (2) x\ x\ x\ x\ y\ y,\ (\ ',\ ')'\ ,\ '\ '$$

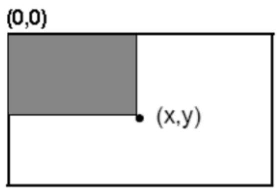

Fig. 3: Summed area of integral image

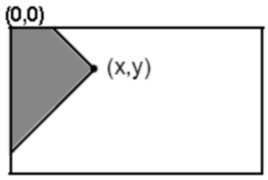

Fig. 4: Summed area of rotated integral image

It only takes two passes to compute both integral image arrays, one for each array.

Using the appropriate integral image and taking the difference between six to eight array elements forming two or three connected rectangles, a feature of any scale can be computed. Thus calculating a feature is extremely fast and efficient. It also means calculating features of various sizes requires the same effort as a feature of only two or three pixels. The detection of various sizes of the same object requires the same amount of effort and time as objects of similar sizes since scaling requires no additional effort [8].

B. Classifiers Cascaded

Although calculating a feature is extremely efficient and fast, calculating all 180,000 features contained within a 24 × 24 sub-image is impractical [Viola 2001, Wilson 2005].

Fortunately, only a tiny fraction of those features are needed to determine if a sub-image potentially contains the desired object [9]. In order to eliminate as many sub-images as possible, only a few of the features that define an object are used when analyzing sub-images. The goal is to eliminate a substantial amount, around 50%, of the sub-images that do not contain the object. This process continues, increasing the number of features used to analyze the sub-image at each stage.

The cascading of the classifiers allows only the sub-images with the highest probability to be analyzed for all Haar-features that distinguish an object. It also allows one to vary the accuracy of a classifier. One can increase both the false alarm rate and positive hit rate by decreasing the number of stages. The inverse of this is also true.

Viola and Jones were able to achieve a 95% accuracy rate for the detection of a human face using only 200 simple features [10]. Using a 2 GHz computer, a Haar classifier cascade could detect human faces at a rate of at least five frames per second.

IV. REGIONALIZED DETECTION

Since it is not possible to reduce the false positive rate of the classifier without reducing the positive hit rate, a method besides modifying the classifier training attribute is needed to increase accuracy [11]. The method proposed to is to limit the region of the image that is analyzed for the facial features. By reducing the area analyzed, accuracy will increase since less area exists to produce false positives. It also increases efficiency since fewer features need to be computed and the area of the integral images is smaller.

In order to regionalize the image, one must first determine the likely area where a facial feature might exist. The simplest method is to perform facial detection on the image first. The area containing the face will also contain facial features. However, the facial feature cascades often detect other facial features as illustrated. The best method to eliminate extra feature detection is to further regionalize the area for facial feature detection. It can be assumed that the eyes will be located near the top of the head, the nose will be located in the center area and the mouth will be located near the bottom.

The upper 5/8 of the face is analyzed for the eyes. This area eliminates all other facial features while still allowing a wide variance in the tilt angle. The center of the face, an area that is 5/8 by 5/8 of the face, was used to for detection of the nose. This area eliminates all but the upper lip of the mouth and lower eyelid. The lower half of the facial image was used to detect the mouth [12]. Since the facial detector used sometimes eliminates the lower lip the facial image was extended by an eighth for mouth detection only.

V. CONCLUSION & FUTURE WORK

In this paper, a new approach to face detection with Adaboost & feed forward neural network is presented. The method uses Adaptive & feed forward neural network for both finding feature points and extracting feature vectors. From the experimental results, it is seen that proposed method achieves better results compared to the graph matching and eigenface methods, which are known to be the most successive algorithms. In the proposed algorithm, since the facial features are compared locally, instead of using a general structure, it allows us to make a decision from the parts of the face. For example, when there are sunglasses, the algorithm compares faces in terms of mouth, nose and any other features rather than eyes. Moreover, having a simple matching procedure and low computational cost proposed method is faster than elastic graph matching methods. Proposed method is also robust to illumination changes as a property of Adaptive, which is the main problem with the eigenface approaches [13]. A new facial image can also be simply added by attaching new feature vectors to reference gallery while such an operation might be quite time consuming for systems that need training.

Although detection performance of the proposed method is satisfactory by any means, in future it would be further improved with some small modifications and/or additional preprocessing of face images. Such improvements can be summarized as

– A set of weights can be assigned to these feature points by counting the total times of a feature point occurs at those responses.
– When there is a video sequence as the input to the system, a frame giving the "most frontal" pose of a person should be selected to increase the performance of face detection algorithm.

– In order to further speed up the algorithm,q could be decreased with an acceptable level of decrease in detection performance.

VI. REFERENCES

[1] Ming-Husan Yang, David J.Kriegman, Narendra Ahuja, “Detecting Faces in Images: A Survey “,IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.24, pp. 34-58 January 2002.

[2] H. A. Rowley, S. Baluja, T. Kanade, “Neural Network-Based Face Detection”, IEEE Transactiono n Pattern Analysis and Machine Intelligence, Vol.20, pp. 39-51, 1998.

[3] Zhang ZhenQiu, Zhu Long, S.Z. Li, Zhang Hong Jiang, “Real-time Multi-View Face Detection”, Proceeding of the Fifth IEEE International Face Conference on Automatic Faceand Gesture Recognition, pp 142-147, 20-21 May 2002.

[4] Gavrila, D.M; Philomin, V.”Real-Time Object detection for Smart Vehicules”. International Conference on Computer Vision (ICCV99), Vol. 1, 20-25 September 1999.

[5] Rolf F. Molz, Paulo M. Engel, Fernando G. Moraes, Lionel Torres, Michel Robert,”System Prototyping Dedicated to Neural Network Real-Time Image Processing”, ACM/SIGDA Ninth International Symposium on Field Programmable Gate Arrays(FPGA 2001).

[6] Haisheng Wu, John Zelek,” A Multi-classifier Based Real-time Face Detection System”, Journal of IEEE Transaction on Robotics and Automation, 2003.

[7] Theocharis Theocharides, Gregory Link, Vijaykrishnan Narayanan, Mary Jane Irwin, “Embedded Hardware Face Detection”, 17th International Conference on VLSI Design, Mumbai,India, January 5-9, 2004.

[8] Fan Yang and Michel Paindavoine,”Prefiltering for Pattern Recognition Using Wavelet Transform and Face Recognition”,16th International Conference on Microelectronics, Tunisia 2004.

[9] Fan Yang and Michel Paindavoine,”Prefiltering for pattern Recognition Using Wavelet Transform and Neural Networks”, Adavances in imaging and Electron Physics, Vol. 127, 2003.

[10] Xiaoguang Li and shawki Areibi,”A Hardware/Software co-design Approach for Face Recognition”, 16th International Conference on Microelectronics, Tunisia 2004.

[11] Fan Yang and Michel Paindavoine,”Implementation of an RBF Neural Network on Embedded Systems: Real-Time Face Tracking and Identity Verification”, IEEE Transactions on Neural Networks, vol.14, pp. 1162-1175, September 2003.

[12] R. McCready, “Real-Time Face Detection on A Configurable Hardware System”, International Symposium on Field Programmable Gate Arrays, Montery, California, United States, 2000

[13] D. Gajski, N. Dutt, A. Wu, “High-Level Synthesis: Introduction to Chipand System Design”, Kluwer Academic Publishers,Boston, 1992.

[14] T. Kanade, “Picture Processing by Computer Complex and Recognition of Human Faces”.Technical Report, Kyoto University, Dept. of Information Science, 1973.

[15] S. Lin, S. Kung, and L. Lin, “Face Recognition/Detection by Probabilistic DecisionBased Neural Network,” IEEE Trans. Neural Networks, Vol.8, pp.114-132, 1997.

[16] R. Brunelli, T. Poggio, “Face Recognition: Features vs. Templates,” IEEE Trans. on PAMI, Vol. 12, No. 1, Jan. 1990. M. H. Yang, N. Ahuja, and D. Kriegman, “A survey on face “.

[17] S. Ranganath and K. Arun, “Face Recognition using Ransform Features and Neural Network,” *Pattern Recognition*, Vol.30,pp.1615-1622, 1997.

[18] S. Lawrence, C. Giles, A. Tsoi, and A. Back, “Face Recognition: A Convolutional Neural Network Approach,” *IEEE Trans. On p p . 9 8 - 1 1 3 ,1997.*

[19] M. Turk and A. Pentland. “Eigenfaces for Recognition.” Journal of Cognitive Science, pp.71-86, 1991.

[20] P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class Specific linear projection,” IEEE Trans. On PAMI, Vol.19, 1997.