A survey of face detection algorithms

Ankit Srivastava, Suraj Mane, Aaditya Shah, Nirmit Shrivastava, Prof. Bhushan Thakare
Computer Department
Sinhgad Academy Of Engineering,
Pune, India.

Abstract:- This paper provides a brief insight of some famous and particularly important algorithms used for face detection. Face detection is a technology used by/ computer systems to detect faces in a given digital image. Automatic face detection is a very complex problem in image processing and many methods and algorithms have been proposed like Viola Jones, CNN (cascade neural networks), Eigenface etc. We have also tried to talk on each algorithm's efficiency and feasibility.

***Keywords*: Face detection, Haar features, Viola Jones.**

I. INTRODUCTION

Face detection is a process of detecting faces in a given scene. The algorithms proposed focuses on the frontal human faces. For humans, it's not a difficult task to do as they know how a face looks like for their brain has been collecting data since childhood. But for machines on the other hand, it's a lot difficult task. Also, the difficulty arises due to exaggerated facial expressions, visual variations in images, large area to find face in, etc. Machines work on the instructions given by us, it needs specific and clear instructions as what to do. Differentiating a face from a given image is a challenging task and in order to accomplish it we need to train machines. Given a digital image, it needs to find faces in it.

This paper is inspired from the work of Cha Zhang and Zhengyou Zhang. They have done incredible work by giving recent developments in face detection techniques. We have also tried to do the same by giving a survey on some important and recent techniques.

II. ALGORITHMS

I) Viola Jones

Viola Jones is a framework for detecting faces proposed by Paul Viola and Michael Jones in 2001. It achieves high detection rate while rapidly processing images. It gives a rate of 15 frames per second. The algorithm is implemented in OpenCV. It has four stages

1. Haar feature Selection
2. Creating an integral image
3. AdaBoost training
4. Cascading classifiers

Haar Features:

Every human face share some common properties, these properties can be matched using Haar like features. Few common features of human faces are:

- Eye region is darker than the nose bridge region.
- .Upper cheek region is brighter than the eye region.
- Location of eyes, mouth, nose-bridge etc.
- Value = Oriented gradients of pixel intensities.

These three features are calculated by the algorithm and are then searched in the image. Viola Jones algorithm uses a 24*24 window. It starts with one pixel per feature and matches it with the entire window. Value of each feature is calculated by subtracting the white region from the black region. Each feature gives one value. After that, two pixels are taken for each feature and is matched across whole window, it again gives one value. Same step is followed for other features also. This gives rise to about 16,000+ features each window. To optimize this number, Viola and Jones devised a method which is discussed in next section called integral image.

2. Integral image

Integral image allows very fast feature calculation. It was introduced so that the features can be calculated very rapidly on a large set of scales. It is used to represent images, and it can be evaluated using few operations per pixel. The integral image at location at x, y contains the sum of pixels above and left to x and y inclusive.

$$\text{I,I}(x, y) = \sum_{x' \leq x, y' \leq} i(x', y')$$

Where I,i(x,y) is the integral image and (x,y) is the original image.

3. AdaBoost training

A classifier is created for classifying the image using AdaBoost. This classifier is created using a small number of features from a large set of classifier using AdaBoost. All of the features calculated are not necessarily part of the face, so the unnecessary features are discarded using this classifier. All of the 160,000+ features are reduced to few hundreds by this classifier. Hence it's of prime importance that this classifier is efficient. It can be viewed as the feature selection. AdaBoost provides an effective learning algorithm and strong bound on generalization performance.

4. Cascading classifiers

Those sub-windows which are not rejected by the initial classifier are processed by a sequence of classifiers, each slightly more complex than the last. If any classifier rejects the sub-window, no further processing is performed. The structure of the cascaded detection process is essentially that of a degenerate decision tree, and as such is related to the work of Fleuret and Geman (2001) and Amit and Geman (1999). The complete face detection cascade has 38 classifiers, which total over 80,000 operations. Nevertheless the cascade structure results in extremely rapid average detection times. On a difficult dataset, containing 507 facesand75millionsub-windows, faces are detected using an average of 270 microprocessor instructions per sub-window. In comparison, this system is about 15 times faster than an implementation of the detection system constructed by Rowley et al. (1998). The reason for considering this algorithm first was because this algorithm was the first to process real time images, also it rejects false positive detection in very early stages. The computation time required for calculating the features is also low. Hence it provided a great means for applications extensively using face detection.

II) Robust Face Detection Using the Hausdorff Distance

This is a model based algorithm which works on the grey-scale still images. It uses Hausdorff distance, which makes it efficient and robust for real time applications. Hausdorff distance is used as a similarity measure between the image in which the object is to be detected and a sample face.

It works in two stages:

1) Coarse detection
2) Refinement

Each of which again uses segmentation and localisation.

Hausdorff Distance

Here I would like to explain what exactly Hausdorff distance is and how it works.

It basically is distance between two distinct point sets. For two finite point sets A and B the Hausdorff distance is given by

H(A, B) = max(h(A, B), h(B, A)), where

h(A, B) = max min(a-b) where a belongs to A and b belongs to B

here h(A, B) is called the directed Hausdorff distance from set A to B with some underlying norm (.) on the points of A and B.

A slightly different version is also present named as Modified Hausdorff distance (MHD). It is given by,

$$h_{\mathrm{mod}}(\mathcal{A},\mathcal{B}) = \frac{1}{|\mathcal{A}|}\sum_{a\in\mathcal{A}}\min_{b\in\mathcal{B}}\|a-b\|$$

It takes average of the single point distances.

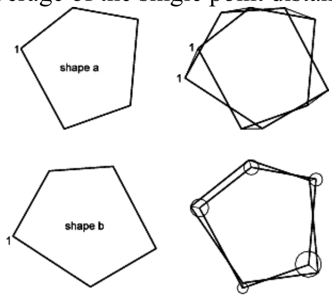

Fig. 1 Model fitting by scaling, translation and rotation

Coarse Detection: Here first of all area of interest (AOI) is determined for each incoming image before applying any method. This AOI is resampled to a fixed size.

- **Segmentation:** From the resampled image an edge intensity image is calculated with the sobel operator. Resulting in binary edge points.

- **Localisation:** With the binary representation A obtained by the segmentation step and a face model B, a localisation of the face in the image can be calculated.

Following image summarizes the two step process:

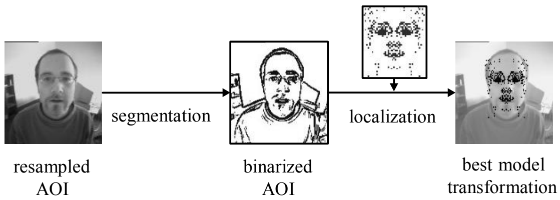

Fig. 2 Coarse Detection

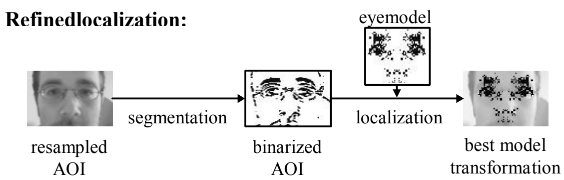

Fig. 3 Refinement

This concludes the Hausdorff Distance based algorithm for face detection which works on the edges of a grey scale images. This system is efficient for still images and can be used for detecting faces robustly which is insensitive to the brightness and exposure changes.

III) Convolution neural network cascade

This method was proposed in the paper Convolutional neural network cascade for face detection. It proposed a new method for detecting faces using CNNs in cascade. It can perform at a speed of 14 Fps on single core processor and at 100 Fps on a GPU. The computation is complex and hence requires high computational processing. Previous algorithms use hand crafted features; CNN learns the features by itself. It accepts large set of training data, this training is implemented on multiple cores in parallel. It uses a cascade of CNNs. There are 6 CNNs in cascade. The first 3 are used for classifying faces and non-faces and rest 3 are for bounding box calibration. The most confident detection window may not be well aligned to the face. As a result, without the calibration step, the next CNN in the cascade will have to evaluate more regions to maintain a good recall. The overall detection runtime increases significantly. This problem generally exists in object detection. It explicitly addresses this problem with CNNs in this work. Instead of training a CNN for bounding boxes regression as in R-CNN, they train a multi-class classification CNN for calibration. It's observed that a multi-class calibration CNN can be easily trained from limited amount of training data while a regression CNN for calibration requires more training data. It's believed that the discretization decreases the difficulty of the calibration problem so that it can achieve good calibration accuracy with simpler CNN structures.

Training:

For training the CNNs in cascade, it collects 5, 800 background images to get negative training samples and the faces in the Annotated Facial Landmarks in the Wild for positive training.

This helps it achieving the speed mentioned above. It shares the advantages of the CNNs. It rejects the false positive faces at low resolution quickly and carefully processes the challenging regions at higher resolution for accurate detection.

IV) Detecting faces using Eigenface

Method of using Eigenface was introduced by Sirovich and Kirby (1987) and used by Matthew Turk and Alex Pentland in face classification. When we provide input to the machine it contains a lot of noise i.e. light effects, poses, backgrounds etc. But all these inputs provided contains some pattern which appears in every image. These pattern are basic features which are present in every face image. Objects like mouth, nose, and eyes and their relative distance. These features or patterns are called as "eigenfaces" or principal components in general. We can extract these features using a mathematical method known as Principal Component Analysis (PCA). It is a mathematical method using which we can extract and use the principal components. We can also reconstruct the original image from these Eigenfaces. That means, if we have eigenfaces then we can have the image from which we got those eigenfaces. For this we need to combine the eigenfaces in particular proportion. Suppose we have an image in which the eigenface is present at a higher degree, then the share of that corresponding eigenface would be higher in the sum of all the eigenfaces. So for this purpose, each eigenface is given some weight according to its presence in the image. All the weighted eigenfaces are added to get the original image.

Working of the algorithm:

First of all a set of training images are given. From these training images eigenfaces are extracted. Then weight is calculated using the image and stored in the gallery. After that if a new image image comes while working, it's weights are calculated and compared with the weights present in the gallery. If a nearest match is present in the gallery then that image is said to be matched. For comparing Euclidian distance is used. More details on calculation of Eigen vectors and PCA can be found in (Pissarenko, 2002, pp. 70-72).

We have given just the rough idea about what is eigenface and how it works. Eigenface algorithm is mainly used for face recognition applications. Some of its advantages are:

- It simplifies the complex representation of the image.
- Its training is completely automatic.
- It can handle a large set of data.
- Once the data is calculated, it can be used in real time.

Some of their shortcomings are:

- It is very sensitive to scale and transformation, lightning.
- The most important Eigenfaces are about illumination encoding and they provide very less information about the actual face.
- Capturing expression changes is difficult.

III. CONCLUSION

We have tried to survey all the important and influential algorithms in most simple and understandable way. We have categorized each algorithm and provided their pros and cons. Characteristics of each algorithm is also given. We have provided the most abstract survey/review of each algorithm so as to enable one to get an idea about that algorithm.

IV. REFERENCES

1] Face Recognition: A Literature Survey, ”Technical Report CARTR-948, Center for Automation Research, University of Maryland, College Park, MD, 2000.

2] Face Recognition: A Literature Survey, W. ZHAO Sarnoff Corporation, R. CHELLAPPA University of Maryland, P. J. PHILLIPS National Institute of Standards and Technology AND A. ROSENFELD University of Maryland.

3] Robust Face Detection Using the Hausdorff Distance Oliver Jesorsky, Klaus J. Kirchberg, and Robert W. Frischholz BioID AG, Berlin, Germany.

4] Face Detection and Tracking using OpenCV S.V. Viraktamath*, Mukund Katti**, Aditya Khatawkar*** & Pavan Kulkarni*****Faculty, Department of Electronics and Communication Engineering, SDM College of Engineering & Technology, Dharwad, Karnataka, INDIA.

5] A Convolutional Neural Network Cascade for Face Detection Haoxiang Li†, Zhe Lin‡, Xiaohui Shen‡, Jonathan Brandt‡, Gang Hua† †Stevens Institute of Technology Hoboken, NJ 07030 {hli18, ghua}@stevens.edu ‡Adobe Research San Jose, CA 95110 {zlin, xshen, jbrandt}@adobe.com

6] Robust Real-Time Face Detection PAULVIOLA Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA viola@microsoft.com MICHAEL J. JONES Mitsubishi Electric Research Laboratory, 201 Broadway, Cambridge, MA 02139, USA mjones@merl.com

7] Ruiz-del-Solar, J and Navarrete, P.*Eigenspace-based face recognition: a comparative study of different approaches*, 2005

8] Turk, Matthew A and Pentland, Alex P. *Face recognition using eigenfaces*. Computer Vision and Pattern Recognition, 1991. Proceedings {CVPR'91.}, {IEEE} Computer Society Conference on 1991

9] M. Kirby; L. Sirovich (1990). "Application of the Karhunen-Loeve procedure for the characterization of human faces". IEEE Transactions on Pattern analysis and Machine Intelligence. 12 (1): 103–108.

10] M. Turk; A. Pentland (1991). "Face recognition using eigenfaces" (PDF). Proc. IEEE Conference on Computer Vision and Pattern Recognition. pp. 586–591.

11] A Survey of Recent Advances in Face Detection Cha Zhang and Zhengyou Zhang June 2010 Technical Report MSR-TR-2010-66 Microsoft Research Microsoft Corporation One Microsoft Way Redmond, WA 98052.

12] R. Lienhart and J. Maydt. An extended set of Haar-like features for rapid object detection. In Proc. of ICIP, 2002.

13] S. Li, L. Zhu, Z. Zhang, A. Blake, H. Zhang, and H. Shum. Statistical learningofmulti-view facedetection. In Proc.of ECCV, 2002.

14] Y.-Y. Lin and T.-L. Liu. Robust face detection with multiclass boosting. In Proc. of CVPR, 2005.

15] BEYMER, D. 1995. Vectorizing face images by interleavingshapeandtexturecomputations.MIT AI Lab memo 1537. Massachusetts Institute of Technology, Cambridge, MA.

16] BEVERIDGE, J. R., SHE, K., DRAPER, B. A., AND GIVENS, G. H. 2001. A nonparametric statisical comparison of principal component and linear discriminant subspaces for face recognition. In Proceedings, IEEE Conference on Computer Vision and Pattern Recognition. (An updated version can be found online at http://www.cs.colostate.edu/evalfacerec/news.html.)

17] BEYMER,D.J.ANDPOGGIO,T. 1995. Facerecognition from one example view. In Proceedings, International Conference on Computer Vision. 500–507.

18] BIEDERMAN,I. 1987. Recognition by components: A theory of human image understanding. Psych. Rev. 94, 115–147.

19] DARWIN, C. 1972. The Expression of the Emotions in ManandAnimals.JohnMurray,London,U.K.

20] BLANZ,V.AND VETTER,T. 1999. AMorphablemodel for the synthesis of 3D faces. In Proceedings, SIGGRAPH’99, 187–194.

21] BLACKBURN, D., BONE, M., AND PHILLIPS, P. J. 2001. Face recognition vendor test 2000. Tech. rep. http://www.frvt.org.

22] BRUCE, V. 1988. Recognizing faces, Lawrence Erlbaum Associates, London, U.K.

23] BRUCE,V.,BURTON,M.,ANDDENCH,N. 1994. What’s distinctive about a distinctive face? Quart. J. Exp. Psych. 47A, 119–141.