Face Detection: Survey

Shubham Sinha[1], Rahul Kumar[2], Shital Nalawade[3], Deepak Ray[4]

[1, 2, 3]Department of Electronics and Telecommunication, BharatiVidyapeeth Deemed University College of Engineering, Pune

[4]Assistant Professor, Department of Electronics and Telecommunication, Bharati Vidyapeeth Deemed University College of Engineering, Pune

Abstract: ***Though this paper we represent the different method of face detection to recognize or detect the faces in the images as well as real time. Actually the Face detection is a computer technology which is now a days used in many application to identify the object as well as face. Human face detection has always been an important problem for face, expression and gesture recognition. If the proper lights is not been provided then sometimes the error is also been defined.to erase these kinds of error some of the algorithm has been introduced.Many techniques is been used to detect the faces and the purpose of this paper is to provide the better knowledge about the method such than no error should be caused.Also the research has been done in many fields like Image recognition, object detection and face detection.***

1. Introduction

Due to the advancement of computer technology .Now a days a more efficient and friendly method for human computer Interaction is being developed which does not depend on the traditional devices. For ex: Keyboards, joysticks
Also due to the decrease in the price of the component of computer vision system and the increase of the performance of the computer system

Coupling of computer system & the computer vision system is embedded together. And these deployment is been seen in the desktop and embedded system

After the growth of the hardware also the expanding of the research has been seen in the field of image processing.
Going through many research papers & books it is been seen that psychophysicist, neuroscientists and engineers has been working in this field for very long time. Many applications has been used the research results

First stage of any face processing system is to detect the location of the faces in the image but this went to be a challenging task because of the variability in scale ,location, orientation (upright rotated), and pose (frontal, profile) occlusion, facial expression and lightning Condition also change the overall appearance of faces.

2. Different Types of Method used in Face Detection

- Knowledge-based methods
- Feature invariant methods
- Template matching methods
- Appearance-based methods

Method	Working	When to use
Knowledge-based method	Rules based	At a time of strictness
Feature invariant method	Grouping of edges. mixture of Gaussian . integration of skin color,size,shape.	At a place of low light.and for a blur image
Template matching	Shape template Active Shape model	When we need to take the record.
Appearance-based method	Joint statistics of local appearance and position Higher order statistics with HMM Eigenvector decomposition and clustering Gaussian distribution nd multilayer perceptron	For the detection of face,object etc.

2.1 Knowledge-Based Methods

These rules are based on method which encode human knowledge of what constitute a typical face. Usually the rules capture the relationship b/w facial features. These methods are designed mainly for face localization. This methods are developed based on the rules derived from the researcher knowledge of human faces. It is easy to come up with simple rules to describe the faces & their relationship .For example a face often appears in the image with the two eyes that are symmetric to each other, a nose & mouth. And Relation can be represented by Relative distance and positions

Working:
- From input image the facial feature is been extracted.
- And Identification of faces are based on the Coded rules.
- A verification process is usually applied to reduce false detection.

Approach and their Representative works-:
- Texture Space Gray-Level Dependence matrix of face pattern
- Skin Color-Mixture of Gaussian
- Multiple Features-Integration of skin color, size and shape
- Facial Features-Grouping of edges.

2.2Feature Invariant Methods

A feature-based approach to face recognition in which the features are derived from the intensity data without assuming any knowledge of the face structure is presented. The feature extraction model is biologically motivated, and the locations of the features often correspond to salient facial features such as the eyes, nose, etc. Topological graphs are used to represent relations between features, and a simple deterministic graph-matching scheme that exploits the basic structure is used to recognize familiar faces from a database. Each of the stages in the system can be fully implemented in parallel to achieve real-time recognition.

2.3 Template Matching

Template matching is a technique in digital image processing for finding small parts of an image which match a template image. It can be used in manufacturing as a part of quality control, a way to navigate a mobile robot, or as a way to detect edges in images.

In template matching, a standard face pattern is manually predefined or parameterized by a function.n. Given an input image, the correlation values with the standard patterns are computed for the face contour, eyes, nose, and mouth independently. The existence of a face is determined based on the correlation values. This approach has the advantage of being simple to implement. However, it has proven to be inadequate for face detection since it cannot effectively deal with variation in scale, pose, and shape. Multiresolution, multiscale, subtemplates, and deformable templates have subsequently been proposed to achieve scale and shape invariance.

For templates without strong features, or for when the bulk of the template image constitutes the matching image, a template-based approach may be effective. As aforementioned, since template-based template matching may potentially require sampling of a large number of points, it is possible to reduce the number of sampling points by reducing the resolution of the search and template images by the same factor and performing the operation on the resultant downsized images providing a search window of data points within the search image so that the template does not have to search every viable data point, or a combination of both

Approach and their Representative works-:

- Predefined face templates-shape template.
- Deformable Templates-Active shape Model

2.4 Appearance-Based Methods

Appearance based method treats the image as one of the two sub classes such as face or non- face.it rely on technique of statistical and machine learning.

The learned characteristics are in the form of distribution model or discriminant function that are consequently used in face detection. Many appearance based method can be understood in probabilistic framework. Appearance based method avoid difficulty in modeling 3d structure of faces by considering possible faces appearance under various conditions.

The model is been learned from a set of training images which should capture the representative variability of facial appearance. These learned models are then used for detection.

Disadvantages of this method is that with the large variation brought about by changes in facial appearance, lightning and expression make the manifold or face /non-face boundaries highly complex

Neural network and kernel based method

Nonlinear classification for face detection may be performed using neural networks or kernel-based methods.

Neural methods: a classifier may be trained directly using preprocessed and normalized face and non-face training sub windows.

- The input to the system of Sung and Poggio is derived from the six face and six non-face clusters. More specifically, it is a vector of 2 × 6 = 12 distances in the PCA subspaces and 2 × 6 = 12 distances from the PCA subspaces.
- The 24 dimensional feature vector provides a good representation for classifying face and non-face patterns.
- In both systems, the neural networks are trained by back-propagation algorithms.

Kernel SVM classifiers perform nonlinear classification for face detection using face and non-face examples. Although such methods are able to learn nonlinear boundaries, a large number of support vectors may be needed to capture a highly nonlinear boundary. For this reason, fast real time performance has so far been a difficulty with SVM classifiers thus trained.

AdaBoost-based Methods

An AdaBoost classifier is a meta-estimator that begins by fitting a classifier on the original dataset and then fits additional copies of the classifier on the same dataset but where the weights of incorrectly classified instances are adjusted such that subsequent classifiers focus more on difficult cases.

For AdaBoost learning, a complex nonlinear strong classifier $H_M(x)$ is constructed as a linear combination of M simpler, easily constructible weak classifiers in the following form

$$H_M(x) = \frac{\sum_{m=1}^{M} \alpha_m h_m(x)}{\sum_{m=1}^{M} \alpha_m} \quad (1)$$

Where x is a pattern to be classified, $h_m(x) \in \{-1.+1\}$ are the M weak classifier ,$\alpha_m \geq 0$ are the combinig coefficients in R, and $\sum_{m=1}^{M} \alpha_m$ is the normalizing factor. In the discrete version, $h_m(x)$ takes a discret value in$\{-1,+1\}$, whereas in the real version , the output of $h_m(x)$ is a number in R.$H_M(x)$ is real-valued, but the prediction of class label for x is obtained as $y(x)=sign[H_M(x)]$ and the normalized confidence score is $[H_M(x)]$.

The AdaBoost learning procedure is aimed at learning a sequence of best weak classifiers $h_m(x)$ and the best combining weights α_m.

A set of N labeled training examples $\{(x_1, y_1), \ldots, (x_N, y_N)\}$ is assumed available, where $y_i \in \{+1, -1\}$ is the class label for the example $x_i \in R^n$. Adistribution $[w_1, \ldots, w_N]$ of the training examples, where $w_i$is associated with a training example *(xi, yi)*, is computed and updated during the learning to represent the distribution of the training examples.

After iteration *m*, harder-to-classify examples (x_i, y_i) are given larger weights $w_i^{(m)}$, so that at iteration $m + 1$, more emphasis is placed on these examples.

AdaBoost assumes that a procedure is available for learning a weak classifier $h_m(x)$ from the training examples, given the distribution $[w_i^{(m)}]$.

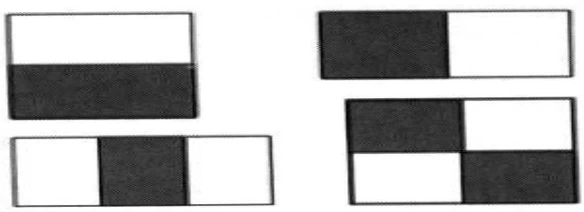
Haar-like features

Viola and Jones propose four basic types of scalar features for face detection as shown in figure. Such a block feature is located in a subregion of a subwindow and varies in shape (aspect ratio), size and location inside the subwindow.

For a subwindow of size 20 × 20, there can be tens of thousands of such features for varying shapes, sizes and locations. Feature *k*, taking a scalar value $z_k(x) \in R$, can be considered a transform from the n-dimensional space to the real line. These scalar numbers form an over complete feature set for the intrinsically low- dimensional face pattern.

Recently, extended sets of such features have been proposed for dealing with out-of-plan head rotation and for in-plane head rotation.

These Haar-like features are interesting for two reasons:
- Powerful face/non-face classifiers can be constructed based on these features
- They can be computed efficiently using the summed-area table or integral image technique.



Four types of rectangular Haar wavelet-like features. A feature is a scalar calculated by summing up the pixels in the white region and subtracting those in the dark region.

Constructing weak classifiers

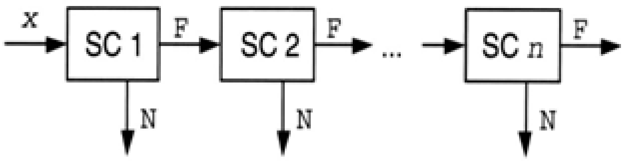
The AdaBoost learning procedure is aimed at learning a sequence of best weak classifiers to combine $h_m(x)$ and the combining weights α_m. It solves the following three fundamental problems:
- Learning effective features from a large feature set
- Constructing weak classifiers, each of which is based on one of the selected features
- Boosting the weak classifiers to construct a strong classifier
- AdaBoost assumes that a "weak learner" procedure is available.
- The task of the procedure is to select the most significant feature from a set of candidate features, given the current strong classifier learned thus far, and then construct the best weak classifier and combine it into the existing strong classifier.
- In the case of discrete AdaBoost, the simplest type of weak classifiers is a "stump". A stump is a single-node decision tree. When the feature is real-valued, a stump may be constructed by thresholding the value of the selected feature at a certain threshold value; when the feature is discrete-valued, it may be obtained according to the discrete label of the feature.
- A more general decision tree (with more than one node) composed of several stumps leads to a more sophisticated weak classifier.

Boosted strong classifier
- AdaBoost learns a sequence of weak classifiers h_m and boosts them into a strong one H_M effectively by minimizing the upper bound on classification error achieved by HM. The bound can be derived as the following exponential loss function:

$$J(H_M) = \sum_i e^{-y_i H_M(x_i)} = \sum_i e^{-y_i \sum_{m=1}^{M} a_m h_m(x)}$$

Where i is the index for training examples.

Cascade of Strong Classifiers
- A boosted strong classifier effectively eliminates a large portion of non-face subwindows while maintaining a high detection rate. Nonetheless, a single strong classifier may not meet the requirement of an extremely low false alarm rate (e.g. 10^{-6} or even lower). A solution is to arbitrate between several detectors (strong classifier), for example, using the "AND" operation.



A cascade of n strong classifiers (SC). The input is a sub window x. It is sent to the next SC for further classification only if it has passed all the previous SCs as the face (F) pattern; otherwise it exists as non-face (N). x is finally considered to be a face when it passes all the n SCs.

Approach and their Representative works
- Eigen face-Eigenvector decomposition and clustering
- Distribution-based-Gaussian distribution and multilayer perceptron
- Neural Network-Ensembles of neural networks and arbitration schemes
- Support Vector Machine-SVM with polynomial kernel
- Naïve Bayes Classifier-Joint statistics of local appearance and position
- Hidden Markov Model(HMM)-Higher order statistics with HMM

Important Stages

Feature detection stages each image is searched for location that are likely to match well in other images

Feature descriptor stageeach region around detected keypoint location into a more compact and stable descriptor that can be matched against other descriptor
Feature matching stage efficiently searches for likely matching candidates in other images
Feature tracking stage alternative to the third stage that only searches a small neighbourhood around each detected feature and as therefore more suitable for video processing

3. Future Work

Now after doing the survey on the whole topic of face detection from various research paper and internet .I m going to make the face detection software with the help of opencv library and we are going to use Appearance based method in the making of face detection software and by using a AdaBoost method we were going to make a Haar of frontal face.And for the matching of the face we are going to use viola-jones method

4. Conclusion

Through this paper we are presented an approach for face detection whichtell us about the different method of detecting faces.These different method minimizes computation time while achieving high detection accuracy. The algorithm which is been was used to construct a face detection system which is approximately very much faster than any previous approach. It can be used in the highly efficient detectors for other objects, such as pedestrians or automobiles, can also be constructed in this way.

This paper brings together new algorithms, representations, and insights which are quite generic and has many application in computer vision and image processing. In order to achieve true scale invariance, almost all face detection systems must operate on multiple image scales. The integral image, by eliminating the need to compute a multi-scale image pyramid, reduces the initial image processing required for face detection . Although significant progress has been made in the last two decades, there is still work to be done, and we believe that a robust face detection system should be effective under full variation in: .lighting conditions, . orientation, pose, and partial occlusion, . facial expression, and . presence of glasses, facial hair, and a variety of hair styles. Face detection is a challenging and interesting problem in and of itself. However, it can also be seen as a one of the few attempts at solving one of the grand challenges of computer vision, the recognition of object classes. The class of faces admits a great deal of shape, color, and albedo variability due to differences in individuals, non- rigidity, facial hair, glasses, and makeup. Images are formed under variable lighting and 3D pose and may have cluttered backgrounds. Hence, face detection research confronts the full range of challenges found in general purpose, object class recognition. However, the class of faces also has very apparent regularities that are exploited by many heuristic or model-based methods or are readily "learned" in data-driven methods. One expects some regularities when defining classes in general, but they may not be so apparent. Finally, though faces have tremendous within-class variability, face detection remains a two class recognition problem (face versus non face).

References

[1] The Viola Jones algorithm for face detection P. Viola, M. J. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision 57(2), 137–154, 2004
[2] Color-to-Grayscale: Does the Method Matter in Image Recognition? Christopher Kanan*, Garrison W. Cottrell
[3] https://pythonprogramming.net/haar-cascade-object-detection-python-opencv-tutorial/
[4] http://opencv.org/