INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT

ISSN 2278 – 0211 (Online)

# An Efficient Hierarchical Based Multi Face Detection and Gender Identification System

**Lavanya Mooram**
Student, Department of Information Technology
V. R. Siddhartha Engineering College, Kanuru, Vijayawada (A.P), India
**S. Sunita**
Assistant Professor, Department of Information Technology
V. R. Siddhartha Engineering College, Kanuru, Vijayawada (A.P), India

***Abstract:***
*Face detection and Gender classification is a very hot research topic in the fields of pattern recognition and computer vision. Face detection is based on identifying and locating a human face in the image, regardless of position, size, and condition. In this paper, we present an algorithm that comprises of two parts, Face detection and Gender classification which are based on haar features and Hierarchical clustering technique. We present experimental results of our algorithm for multi face images.*

***Keywords****: Haar features, Hierarchical clustering, Face detection, Gender classification*

## 1. Introduction

Face detection is a very hot research topic in the fields of pattern recognition and computer vision. Its applications are widely used in artificial intelligence, demographic data collection, identity authentication and human machine interaction etc. Face detection is based on identifying and locating a human face in the image, regardless of location, size, and condition. The goal of face detection is to determine whether any faces are present in the image or not and, if present, return the image location and extent of each face .Various challenges are associated with face detection. Various techniques have been developed for face detection in single and multi face images. Generally they are categorized into four categories. Those are knowledge based methods [1], feature based approaches [2], template matching methods [3] and appearance based methods [4].

Template based approach matches facial elements to preciously designed templates. This approach could be too complex due to extensive computation involved and is only effective when query and modal image have same scale, orientation and illumination properties. Appearance based method to refer any of the extracted characteristic of the image known as feature. However this technique requires good-quality image to extract the features correctly. Feature based face detection approach on the apparent properties of the face such as skin, color and face geometry are explored. It depends on feature deviation and analysis to gain the required knowledge about face [2].

Clustering is the method of grouping the data into sets so that data points within the set should have high similarity while those within different sets are dissimilar[5],[6]. Clustering plays a significant role in various fields including image processing, computational biology, mobile communication, machine learning, wireless sensor networks, medicine and economics. There has been a tremendous growth of using clustering techniques in gender classification. Identifying similar data to form various meaningful groups is extremely important in the field of image processing [7]. Among the various clustering techniques, the hierarchical based clustering technique is one of the powerful approach to detect the multiple faces in an image and to classify that faces. Hierarchical clustering algorithms can be further divided into agglomerative approaches and *divisive* approaches based on how the hierarchical dendrogram is formed. Agglomerative algorithms (bottom-up approach) initially regard each data object as an individual cluster, and at each step, merge the closest pair of clusters until all the groups are merged into one cluster. Divisive algorithms (top-down approach) starts with one cluster containing all the data objects, and at each step split a cluster until only singleton clusters of individual objects remain[8][9].

Gender classification can be done using various methods like gait, hand, face and iris. But gender classification using facial features has given major priority due to its increased attention. Gender classification refers to designate an image of a person into one of the categories of male or female [10]. Gender classification using face images has become major priority. By looking at ones' faces, it is difficult to identify who they are but also lot of information like their emotions, ages and genders are perceived. This is why gender

recognition by face has received much interest in computer vision research community over past two decades. The facial images may contain the variation in illumination, pose, background clutter, and partial occlusion. We consider all these variations in facial image and develop a reliable method to identify the gender. Gender identification methods can be divided into two main categories: geometry based and appearance based. The geometry based category focuses on extracting the geometry feature points from the facial images and describing the shape structure of the face. The appearance-based category can further be divided into two approaches: texture-based and statistical based. The former uses different texture descriptors to characterize a facial image about gender. The statistical-based approach aims at using different features which are quantified into probability to characterize a facial image about gender according to their visual traits [11].

Several face detection and gender classification methods are presented in literature. Simon J. D., et al [12] localize faces using a detector called standard sliding-window. Preprocess the face region by using Gabor filters and produces a feature vector. Additive sum of non-linear function of the data classifier is used. However, test on faces can be identified by commercial face detector. This leads to poor performance where gender classification is hardest. A. Jain, et al [13] developed gender classifiers with performance superior to existing gender classifiers. For feature vector ,Independent component analysis (ICA) is used to represent each image in low dimensional space. Different classifiers are studied in this lower dimensional space.. By using support vector machine (SVM) in ICA space they got 96% accuracy. Rodrigo Verschae et al [14] stated a framework for classifying face images using Adaboost and domain-partitioning based classifiers. They used a different features like LBP, wavelets, rectangular e.t.c., Bo Wu et al [15] stated two approaches :Support Vector Machines and Adaboost Learning methods to the problem of gender classification. SVM, gives better correct rate but requires more computational effort on the other hand adaboost are much faster with slightly worse performance. In this, a new Look Up Table(LUT) weak classifier based adaboost is presented for gender classification learning. Matthew Toews et al [16] stated a novel method for detecting, localizing, and classifying of faces from arbitrary viewpoints and in the presence of occlusion. Appearance based features are derived from local scale-invariant features. An appearance model is first learned for the object class, and to identify the model features Bayesian classifier is trained. Zhang et al [17] uses face profiles and ear images for gender classification to extract powerful features which are classified by support vector classification (SVC) with histogram intersection kernel , Hierarchical and discriminative bag of features technique is stated. An average classification accuracy of 97.65% is achieved.

In this paper, hardware architecture for face detection is proposed. Face Detection system creates a window is an integral part of the implementation of the Haar classification feature. And then Performs parallel clustering and classification to detect the faces in image. This paper is organized as follows: In Section II, proposed face detection and gender classification algorithms are explained. In Section III, experimental results and analysis are presented. Conclusion is presented in Section IV.

## 2. Proposed Algorithm

The proposed system consists of modules for face detection, gender classification system shown in Figure.1. The initial face detection module scans the captured image and detects the multiple human faces in an image. In Gender Classification module, for every detected face, Haar features are computed and minimum distance is calculated using Euclidean distance and based on distance measure hierarchical clustering is done and then classify the face is that face belongs to male or female and then count the number of males and females in an image. The block diagram shown in fig.1.

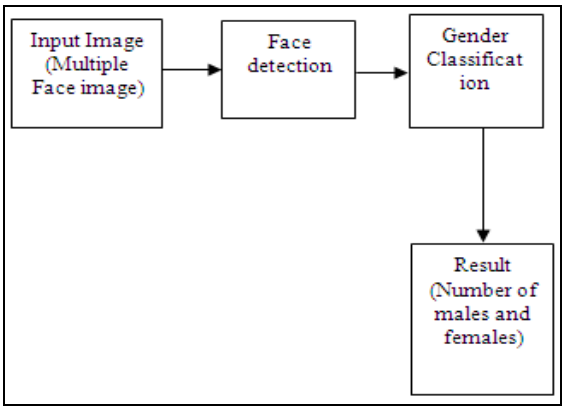


*Figure 1: Block Diagram for Gender Classification System*

### *2.1. Haar Based Face Detection*

Haar based face detection algorithm is a sub-window based algorithm with a dense or over complete feature set from which effective features are selected using Adaboost algorithm. Haar features are suitable for rigid objects and are capable of extracting the structural content of an object. Haar face detection involves two stages: training stage where the classifier is trained to detect faces and testing stage where the detection of face takes place in real-time. The Haar features computed are over complete. Thus the computation of these features as such from the image is time consuming. A faster computation of Haar features is possible using an intermediate representation for the image which we call the integral image.

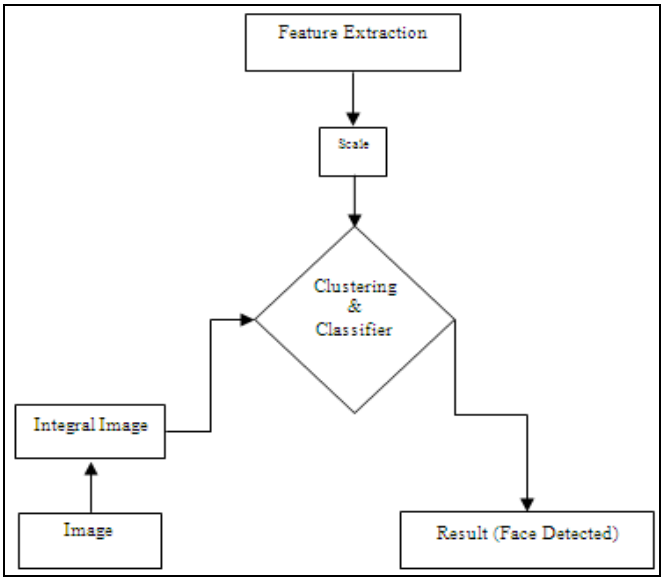


*Figure 2: Block Diagram for Face Detection*

2.1.1. Integral Image

The major step of this algorithm is to get the integral image representation of input image. This is done by assembly each pixel equivalent to the whole summation of all pixels above and to the leftward of the anxious pixel.

2.1.2. Feature Extraction

The proposed face detection technique uses a set of modest rectangular features to spot faces. Various types of rectangles are employed which are shown in below figure 3.

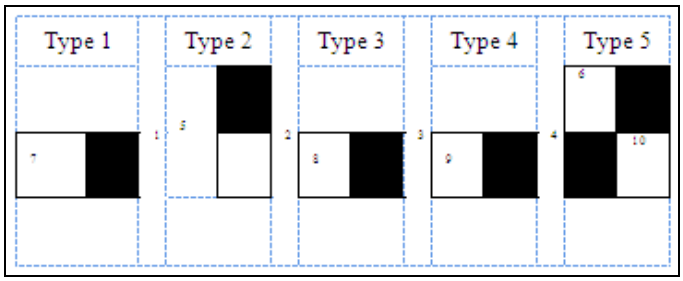


*Figure 3: Types of Haar Features*

2.1.3 Scaling

Scaling is required in most of the cases because size of face is variable in an image. In order to spot all the faces in an image the scaler module is used to resize the sub frame. It supports nearest neighbor, bilinear and poly-phase scaling algorithms. In all these algorithms nearest neighbor is used because it consumes less hardware resources

2.1.4 Classifier

Classifier uses the rectangle integral to calculate the value of a feature. If the image contain more than one face then general Classifier takes large amount of time to evaluate sub windows. So, Cascade architecture type Classifier is used. The cascaded classifier consists of stages each stage containing a strong classifier. The job of each stage is to tell whether a given sub window is definitely a face or not. 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 face is passed on to the next stage in the cascade. Figure shows the cascade architecture.

2.1.5. Haar Features

There are various positions, scales and kinds for haar features. There scheme and structure is simple. By addition of black area pixels and subtracts the sum of white area pixels in the rectangle we obtain the eigen values for an image. It allows very well grained review of files, it also rises the training time and can decrease the performance. For these reasons, we a select feature by creating a weak workbook for every haar feature. Specifically, to reduce the error of each haar feature in training we create a binary workbook on the basis of threshold. One is chosen as best workbook weak for that round for every round of increase(consistent to a special feature Haar). Finally, the effect of the increase is a powerful seed production is calculated as a combination of a weak linear threshold works. Haar based face detection method is both fast and efficient.

2.1.6. Gender Classification

**Step 1**: Feature Extraction: For every detected face haar features are calculated as explained in 2.1.5

**Step 2:** Hierarchical Clustering Algorithm:

Hierarchical clustering provides a hierarchy of partitions at multiple levels of granularity. Agglomerative clustering is one of the hierarchy clustering method which starts with separate cluster for each and point and continuously merge the two closest clusters until we get a single cluster.

Given a set of N items to be clustered, and an N*N distance (or similarity) matrix, the basic process of hierarchical clustering is this:

- Assign each feature as single cluster like $c_1$, $c_2$, $c_3$, .. $c_n$ where n is the no. of features extracted from an image.
- By using any similarity measure, find the distance matrix Dis.
- According to dis(a, b) = min dis(i, j) { i, is an object in cluster a and j in cluster b}
- Find the closest pair of clusters in the current clustering, say pair (a), (b),
- Merge clusters (a) and (b) into a single cluster to form a merged cluster. Store merged objects with its corresponding distance
- Update distance matrix, Dis, by deleting the rows and columns corresponding to clusters (a) and (b). Adding a new row and column corresponding to the merged cluster(a, b) and old cluster (k) is defined in this way:
- dis[(k), (a, b)] = min d[(k),(a)], d[(k),(b)].
- For other rows and columns copy the corresponding features from existing distance matrix.
- If all the extracted features formed into one cluster, stop the process. Else go to step 3 in an Hierarchical Clustering Algorithm.

**Step 3 Classification**: In the gender classification process, we use Haarcascade Classifier. For that frontal face and eye features of male and female faces are trained. For the test image we calculate the front face features and eye features of a detected face. And compare the feature values with trained male and female frontal face and eye features by calculating the Euclidean distance. Then we identified that face is belongs to male if it has minimum distance with male features . If it has minimum distance with female then we identified that face belongs to female. This process continues for all faces in an Image.

## 3. Implementation and Results

First the input image is converted into gray scale and its integral image representation is calculated. Then the input image is given to the modal. The modal detects the face by using haar features of adaptive boosting algorithm. We have used real time multiple faced image database for testing of algorithm. It includes 200 images. We have arbitrarily select 85% of images as training set and 15% as testing set. The algorithm is tested on JAVA. We have divided the images into two categories. Those are:

- Single frontal faces
- Near frontal multiple faces

### *3.1. Single Frontal faces*

*Figure 4*

| Image | Size | False Positive | False Negative | No. of faces detected | Gender Classification |
|---|---|---|---|---|---|
| 1 | 494X466 | 0 | 0 | 1 | Female |
| 2 | 494X466 | 0 | 0 | 1 | Male |

*Table 1*

*3.2. Multiple Frontal Faces*


*Figure 5*

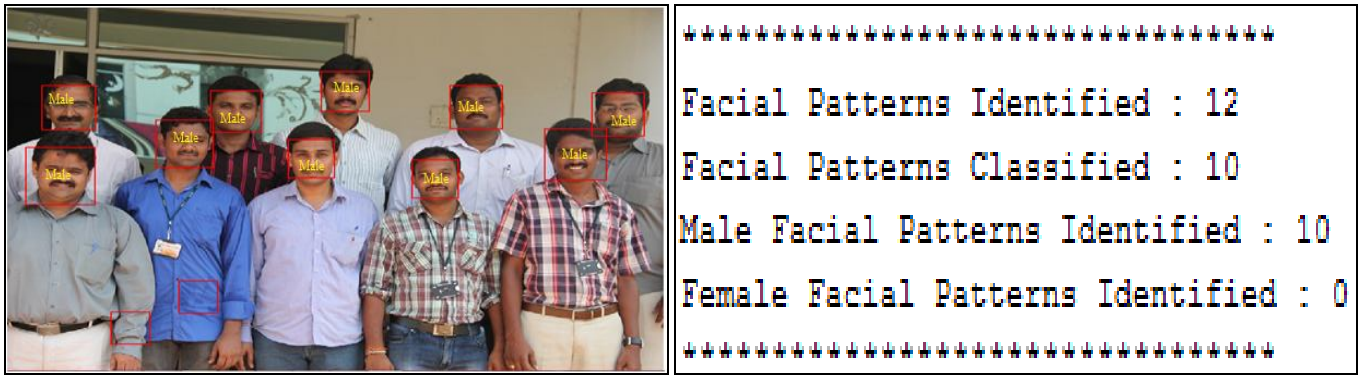


*Figure 6*


*Figure 7*

| Image | Size | Total Faces | False Positive | False Negative | No. of faces detected | No. of Males | No. of Females |
|---|---|---|---|---|---|---|---|
| 1 | **480X640** | **15** | **2** | **0** | **13** | **0** | **13** |
| 2 | **480X640** | **10** | **0** | **0** | **10** | **10** | **0** |
| 3 | **480X640** | **44** | **2** | **0** | **42** | **26** | **16** |

*Table 2*

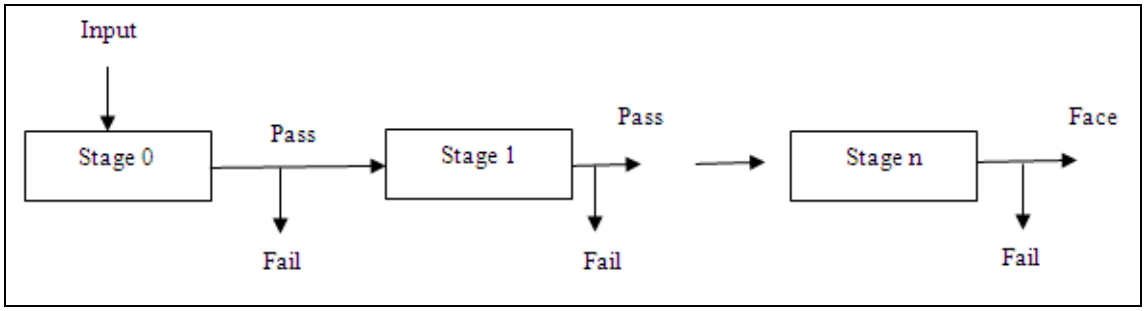


*Figure 8: Classifier*

**4. Conclusion and Future Work**

Face detection and Gender classification for different group images have been identified. Our future work is to improve for a given number of different images is to identify all the number of male or female in any given image.

**5. References**


1. G. Yang and T. S. Huang, "Human Face Detection in Complex Background," Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1994.
2. K.C. Yow and R. Cipolla, "A Probabilistic Framework for Perceptual Grouping of Features for Human Face Detection," Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, pp. 16-21, 1996.
3. Lanitis, C.J. Taylor, and T.F. Cootes, "An Automatic Face Identification System Using Flexible Appearance Models," Image and Vision Computing, vol. 13, no. 5, pp. 393-401, 1995.
4. K.-K. Sung and T. Poggio, "Example-Based Learning for View- Based Human Face Detection," IEEE Trans. Pattern Analysis andMachine Intelligence, vol. 20, no. 1, pp. 39-51, Jan. 1998.
5. Jiawei Han and M. Kamber, Data mining concepts and techniques, Morgan Kaufmann Publishers, 2007.
6. M H Dunhum, Data Mining- Introductory and Advanced Topics, Pearson Education,2007.
7. Roded Sharan, Ron Shamir, " A Clustering Algorithm with Applications to Gene Expression Analysis," Proc. Of American Association for Artificial Intelligence (ISMB-00), 307-316,2000.
8. Dubes, R. and Jain, A. Algorithms for Clustering Data. Prentice Hall, 1988.
9. Kaufman, L. and Rousseeuw, P.J. Finding Groups in Data: an Introduction to Cluster Analysis. John Wiley and Sons, 1990.
10. Ahmed Khan, Muhammad Naeem; Qureshi, Sheraz Ahmed and Riaz, Naveed. "Gender
11. Classification with Decision Trees", International Journal of Signal Processing, Image Processing & Pattern Recognition, 2013.
12. Shen-Ju Lin. "Gender Identification Using Feature Patch-Based Bayesian Classifier", Lecture Notes in Computer Science, 2011
13. Simon J. D., et al., "Gender Classification in Uncontrolled Settings Using Additive Logistic Models," in International Conference on Image Processing, 2011.
14. Jain, et al., "Gender identification using frontal facial images," in Multimedia and Expo, 2005, ICME 2005. IEEE International Conference on, 2005
15. Rodrigo Verschae , Javier Ruiz-del-Solar , Mauricio Correa, Gender classification of faces using adaboost, Image Analysis and Applications, 2006.
16. Bo Wu , Haizhou Ai , Chang Huang, LUT-based Adaboost for gender classification, Proceedings of the 4th international conference on Audio- and video-based biometric person authentication, . 2003.
17. Matthew Toews, Tal Arbel, Detection, Localization, and Sex Classification of Faces from Arbitrary Viewpoints and under Occlusion, IEEE Transactions on Pattern Analysis and Machine Intelligence, September 2009.
18. Guangpeng Zhang , Yunhong Wang, Hierarchical and discriminative bag of features for face profile and ear based gender classification, Proceedings of the 2011 International Joint Conference on Biometrics,2011