

Gender Detection from Frontal Face Images



Inspiring Excellence

A thesis submitted for the degree of
Bachelor of Science in Computer Science

Supervisor: Professor Md. Haider Ali

Co Supervisor: Annajiat Alim Rasel

Mirza Mohtashim Alam (16341021)

Swagatam Roy Rocky (12101053)

School of Engineering and Computer Science

Department of Computer Science & Engineering

DECLARATION

We, hereby declare that the thesis “Gender Recognition from Facial Images Detected using DSP Algorithm” is based on the results found by ourselves under the supervision of Professor Dr. Md. Haider Ali and co supervision of Annajiat Alim Rasel. It is carried out for the degree of Bachelor of Science in Computer Science. Materials of work used here found by other researchers are acknowledged by reference. This Thesis, neither in whole or in part, has been previously submitted for any degree.

Signature of Supervisor

Professor Dr. Md. Haider Ali

Signature of Co Supervisor

Annajiat Alim Rasel

Signature of Author

Mirza Mohtashim Alam

Signature of Author

Swagatam Roy Rocky

ACKNOWLEDGEMENT

We would first like to thank our thesis advisor Professor Dr. Md. Haider Ali of the Department of Computer Science & Engineering at BRAC University. The door to his office was always open whenever we ran into a trouble spot or had a question about our research or writing. He consistently allowed this paper to be our own work, but steered us in the right the direction whenever he thought we needed it.

We would also like to acknowledge our Co supervisor Annajiat Alim Rasel of the Department of Computer Science & Engineering at BRAC University as we are gratefully indebted to him for his valuable comments and support throughout this thesis.

Finally, we are grateful to the influences and support of faculties, friends from BRAC University and our family members, which has been a giant motivation in this thesis journey. This accomplishment would not have been possible without them. Thank you.

INDEX

DECLARATION	i
ACKNOWLEDGEMENT	ii
INDEX	iii
LIST OF FIGURES	v
LIST OF TABLES	vii
ABSTRACT	viii

Chapter 1: INTRODUCTION	1
1.1 Motivation	2
1.2 Aims and Objectives	3
1.3 Thesis Outline	4
Chapter 2: BACKGROUND STUDY AND RELATED WORK	5
2.1 Face Detection from Image	6
2.1.1 Finding Faces by Color	6
2.1.2 Edge-Orienting Matching	7
2.1.3 Hausdroff Distance	7
2.1.4 Viola and Jones	8
2.2 Pre Processing of Image	10
2.2.1 Local Pixel Grouping	10
2.2.2 Linear Filter	11
2.2.3 Median Filter	11
2.2.4 Adaptive Filter	12
2.2.5 Histogram Equalization	13

2.3 Classification	15
2.3.1 Gender Detection Classifier	15
2.3.2 Local Binary Pattern	15
2.3.3 Local Directional Pattern	16
2.3.4 Deformable Spatial Pyramid	17
Chapter 3: DESIGN APPROACH	19
3.1 Face Detection	21
3.1.1 Viola and Jones	21
3.2 Noise Reduction and Grayscale Conversion	23
3.3 Histogram Equalization	25
3.4 Data Analysis (Gender Detection)	27
3.4 Deformable Spatial Algorithm	27
Chapter 4: RESULTS AND DATA ANALYSIS	29
Chapter 5: CONCLUSION AND FUTURE PLAN	36

LIST OF FIGURES

Figure	Page
Figure 2.1: Image showing mixture models and Gaussian mixtures found from face.	6
Figure 2.2: Edge orientation vector field computed	7
Figure 2.3: Model fitting by scaling, translation and rotation	8
Figure 2.4: Example of a “non-face” on the surface of mars	8
Figure 2.5: Haar features that calculate positive and negative features from facial images	9
Figure 2.6: Linear filtering “salt and pepper” noise	11
Figure 2.7: Median filtering keeping border values unchanged to preserve edges	12
Figure 2.8: Comparison between Adaptive Filtering (Left) and Gaussian Filtering (Right)	12
Figure 2.9: Before Applying Histogram Equalization	14
Figure 2.10: Surrounding pixel intensity compared to middle pixel	15
Figure 2.11: Kirsch edge response masks in eight directions	16
Figure 2.12: The Graph Structure of DSP model	17
Figure 2.13: DSP matching method by dividing grid cells	18
Figure 3.1: All steps required for our proposed system	20
Figure 3.2: Detecting faces using Viola & Jones	22

Figure 3.3: Left: Real Image Right: Image after adaptive filtering	24
Figure 3.4: Before Applying Histogram Equalization	26
Figure 3.5: After applying Histogram Equalization	26
Figure 3.5: Recognized faces using DSP	31
Fig 4.1: Detection of female with short hair	32

LIST OF TABLES

Table	Page
Table 4.1: Accuracy of detection	30
Table 4.2: Accuracy with respect to preprocessing	31
Table 4.3: Experiment with short hair women	32
Table 4.4: Experiment with no hair men	33
Table 4.5: Success against training data (female)	33
Table 4.6: Success against training data (male)	34

ABSTRACT

In these modern days, gender recognition from facial image has been a crucial topic. To solve such delicate problem several handy approaches are being studied in Computer Vision. However, most of these approaches hardly achieve high accuracy and precision. Lighting, illumination, proper face area detection, noise, ethnicity and various facial expressions hinder the correctness of the research. Therefore, we propose a novel gender recognition system from facial image where we first detect faces from a scene using Haar Feature Based Cascade Classifier by Paul Viola and Michael Jones with the help of Adaboost technology. The face detection goal is achieved by OpenCV. After the detection of a face and noise is reduced using Histogram equalization. Finally, Deformable Spatial Pyramid (DSP) matching algorithm is used to match the processed facial image with the knowledge base containing classified male and female frontal face images. Our proposed system pulls out good accuracy in the experiment.

Chapter 1

INTRODUCTION

1. INTRODUCTION

The process of gender detection from facial images has become an important issue nowadays. This research will prove helpful in sectors such as surveillance system, criminology, security, field of psychology and biometric authentication [1]. A person's identity can be determined through his/her eyes even if the hairstyle has been changed, facial hair has been added or removed and the facial appearance of a person is altered. However, these tasks can be carried out easily by a human being but not by machines without any intelligence [2]. Moreover, the world is becoming more machine dependent nowadays. Thus, it has become an important topic in the field of image processing and computer vision. Demand of face detection, gesture detection, person recognition, motion capture and detection is increasing tremendously because of reliability in security and authentication process. Focusing on this matter, our target is to develop a system which detects all the faces from a scene and performs the gender identification process on the faces from the scene. There are two different approaches that can be considered. High level features such as distance between eyes, nose and mouth, measurement of different parts of human face can be an approach to detect the gender of a human being. This method can prove erroneous in cases because these high level features can provide inappropriate results or even unresolved [2]. For this reason, we have focused on the pixel level matching between images.

1.1 Motivation

According to a research: "High level Feature-based" approach uses a set of discriminative facial features such as nose, distance between eyes, mouth, eye brow, etc. which are extracted from facial image as classification attributes [1]. Even though this approach mostly works well, it requires images in a controlled situation with proper lighting, good details and so forth which is its biggest limitation. Most real life images are not captured in a manner suitable for this method. It also fails to work in situations where human eyes are not clearly understood, such as, an image of a male person who has prominent female attributes in his face or vice versa. Our primary motivation for this thesis has been to eliminate this limitation. This has been acknowledged by using "pixel-level matching" method which, irrelevant of the quality of the image, would correctly assume the gender of the people in the picture.

1.2 Aims and Objectives

The primary aim of this thesis is to detect the gender of a person from his/her facial image with decent accuracy. There has been a lot of work done in this field using various methods which all have their shortcomings. For example, the use of SVM on images and FERET database accumulates positive results for gender detection highlighted by Lapedriza *et al* [3]. They extracted facial features and classified it in two, external and internal, parts and stored data in FERET database. Also, Makinen *et al* [4] proposed a system using FERET database. However, FERET database requires images to be captured in controlled situations like consistent lighting. This is a problem because most real life images are not taken under controlled situations for the most part and is very impractical to do so. There are also methods specified on single features such as eye features [5] which works well as long as the subject has the required features. It ignores that some people are blind or does not have eye brows. There are other methods that acknowledge these faults but fail to provide satisfactory results.

This thesis has specifically targeted on the issue of establishing a new tactics which can help us to establish efficient operation of facial data extraction and gender classification techniques. The step by step procedure of this thesis is summarized here. This work marks the following issues:

- Firstly, we have used input images to detect faces using Viola & Jones algorithm for robust and real-time extraction of faces.
- Then, the image has been processed to reduce noise using Adaptive Filtering and for adjusting contrast, Histogram Equalization.
- Finally, we have used DSP (Deformable Spatial Pyramid) to produce extremely accurate results and efficient computation to reduce computational time.

1.3 Thesis Outline

The thesis is organized as follows:

- Chapter 1 is about the formal introduction of the thesis which includes our motivation for starting this thesis and our aims and objectives for it.
- Chapter 2 is the Literature review which consists of current knowledge and work done on this topic. Also, it includes all theoretical and methodological contributions for this thesis.
- Chapter 3 focuses on our proposed model for this thesis which the work flow we maintained throughout thoroughly.
- Chapter 4 shows the end results of our experiments and analysis of the data we acquired as a result.
- Chapter 5 is the conclusion which includes the summary of our thesis along with current limitations, future plans and further insight.

Chapter 2

BACKGROUND STUDY AND RELATED WORK

2. BACKGROUND STUDY AND RELATED WORK

2.1 Face Detection from Image

2.1.1 Finding Faces by Color

Since color is responsible for providing an efficient method which is computationally effective and robust in case of depth and partial occasions, it is still being used. It can work with a combination of various methods such as motion and appearance based face detection. Human being has skin which is shaped of a tight assembly in color space. That is why, when different races are considered, this method performs well [6].

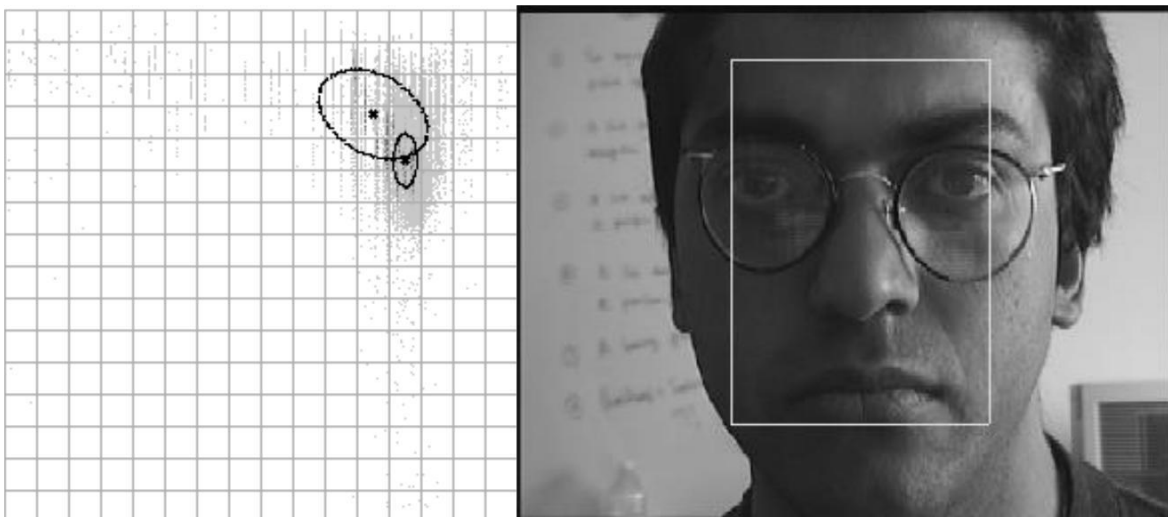


Figure 2.1: Image showing mixture models and Gaussian mixtures found from face [6].

2.1.2 Edge-Orientation Matching

This method is a local image feature to model objects for detection and is able to detect image in real time and so used in the field of “Real-time Face Detection”. Edge features are extracted from a 2D array of pixels which works as the basic feature for calculation. This is a gradient based method which requires convolving the image. The convolution of the image with the two filter masks gives two edge strength images. Sample of hand-labeled face images are used to create face model. The faces are cropped, aligned and scaled in the grey level domain. From this set of normalized face images an average face is computed. Vertically mirrored versions of each face in the set are adjusted to the average face. Finally the edge orientation vector field is calculated from the average face. For face detection, the model is shifted over the image, and at each image position the similarity between the model and the underlying image patch is calculated. The image is represented by its orientation field [7].

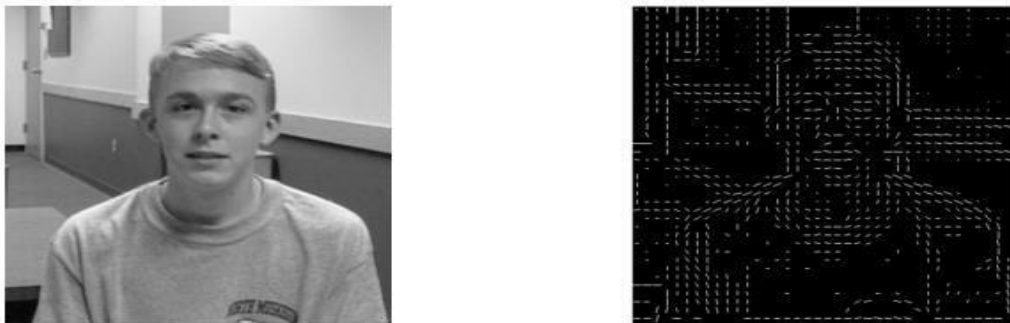


Figure 2.2: edge orientation vector field computed [7]

2.1.3 Hausdorff Distance

This method is considered as a robust scheme which uses Hausdorff distance for measuring similarity among general face model and probable instances of the object. Hausdorff distance is considered as a metric among sets of two points. Illumination and background is changed robustly by it [8].

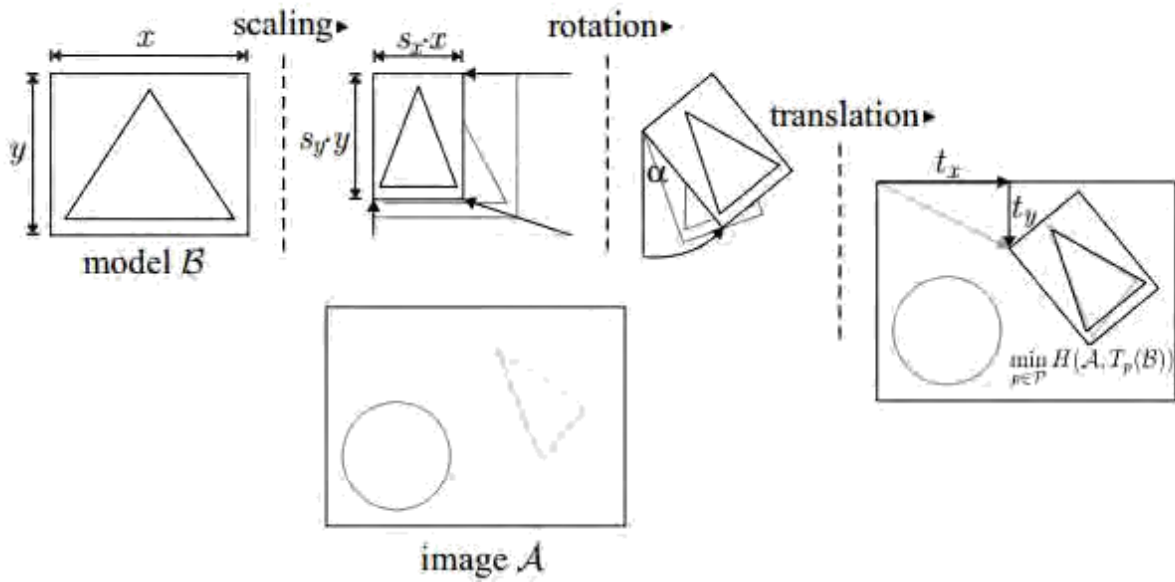


Figure 2.3: Model fitting by scaling, translation and rotation [8]

2.1.4 Viola & Jones

This algorithm by Viola and Jones has had the biggest impact on the field and the most popular currently. It can be used as both realtime and robust face detection. The main idea is to differentiate real faces (true images) from non faces (false images).



Figure 2.4: Example of a “non-face” on the surface of mars [9]

The algorithm has four stages:

- 1) Haar Feature Selection
- 2) Creating an Integral Image
- 3) Adaboost Training
- 4) Cascading Classifiers

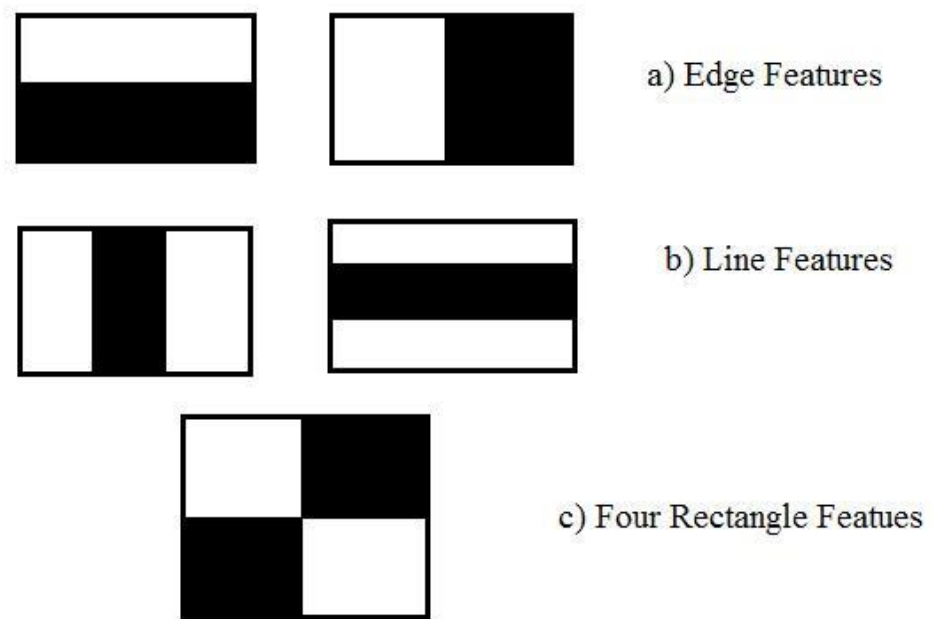


Figure 2.5: Haar-features that calculate positive and negative features from facial images [10].

2.2 Pre Processing of the Image

2.2.1 Local Pixel Grouping

Local Pixel Grouping is used for noise reduction because of its elementary nature. Principle Component Analysis (PCA) helps the system to be stable while noise is being cleared. It works generally well. Low variance components are discarded by PCA in the time of conserving Principal Components with Larger variance [11].

2.2.2 Linear Filter

Linear Filter is used for removal of noise using various filters like averaging filters or Gaussian filters, though it depends on the type of noise. For an example averaging filter is used for grainy type of noise in an image. Here, each pixel gets the average value of neighbor pixels.



Figure 2.6: Linear filtering “salt and pepper” noise [12]

2.2.3 Median Filter

Median Filter is a nonlinear method which removes noise from images while keeping the edges unharmed. It loops throughout the whole image by each pixel and replaces the value by neighboring pixel’s median value. The neighboring pixel pattern is called “Window”. It slides pixel by pixel inside the whole image. The median is calculated from the window in a numerical order by sorting the entire pixel values. Finally, the replacement has been done with the middle pixel value [13].

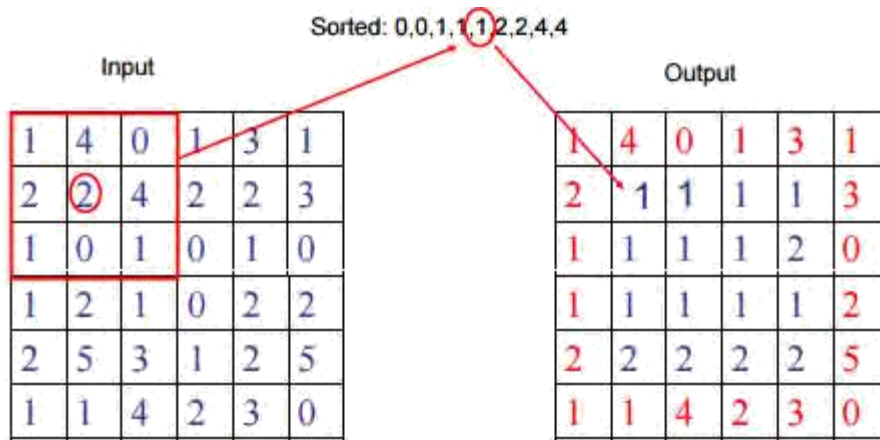


Figure 2.7: Median filtering keeping border values unchanged to preserve edges [13]

2.2.4 Adaptive Filter

This is considered as a better option than the previous methods discussed as it is consisted of all their attributes. It requires a filter called Wiener Filter. This method is better than linear filtering because of being selective in nature. The edges are kept intact, thus it generates more accuracy in results. Though it has good amount of accuracy, it requires more computation time. Wiener2 function is responsible for primary computations and implementation of the filter [14].

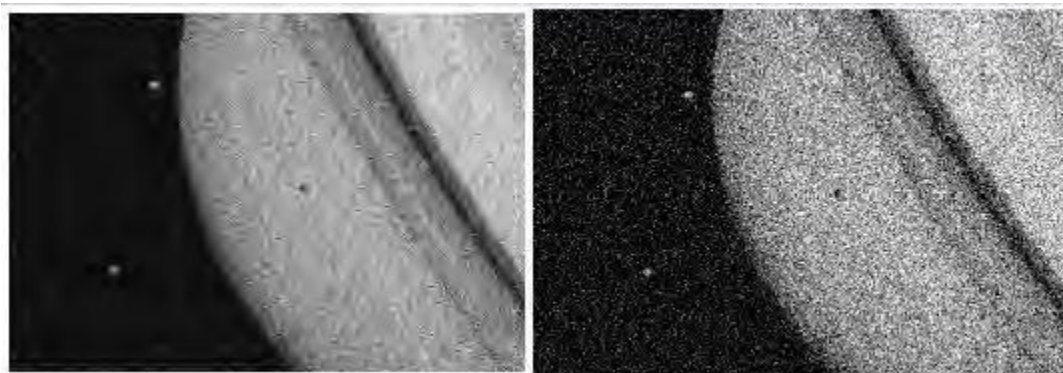


Figure 2.8: Comparison between Adaptive Filtering (Left) and Gaussian Filtering (Right) [14]

2.2.5 Histogram Equalization

Histogram equalization is a method where intensity of the images is stretched out so that the contrast is adjusted. The intensities are stretched across so that a better preprocessing has been done. Different image processing techniques require Histogram Equalization. We also have used this method as preprocessing technique which is described in work methodologies.

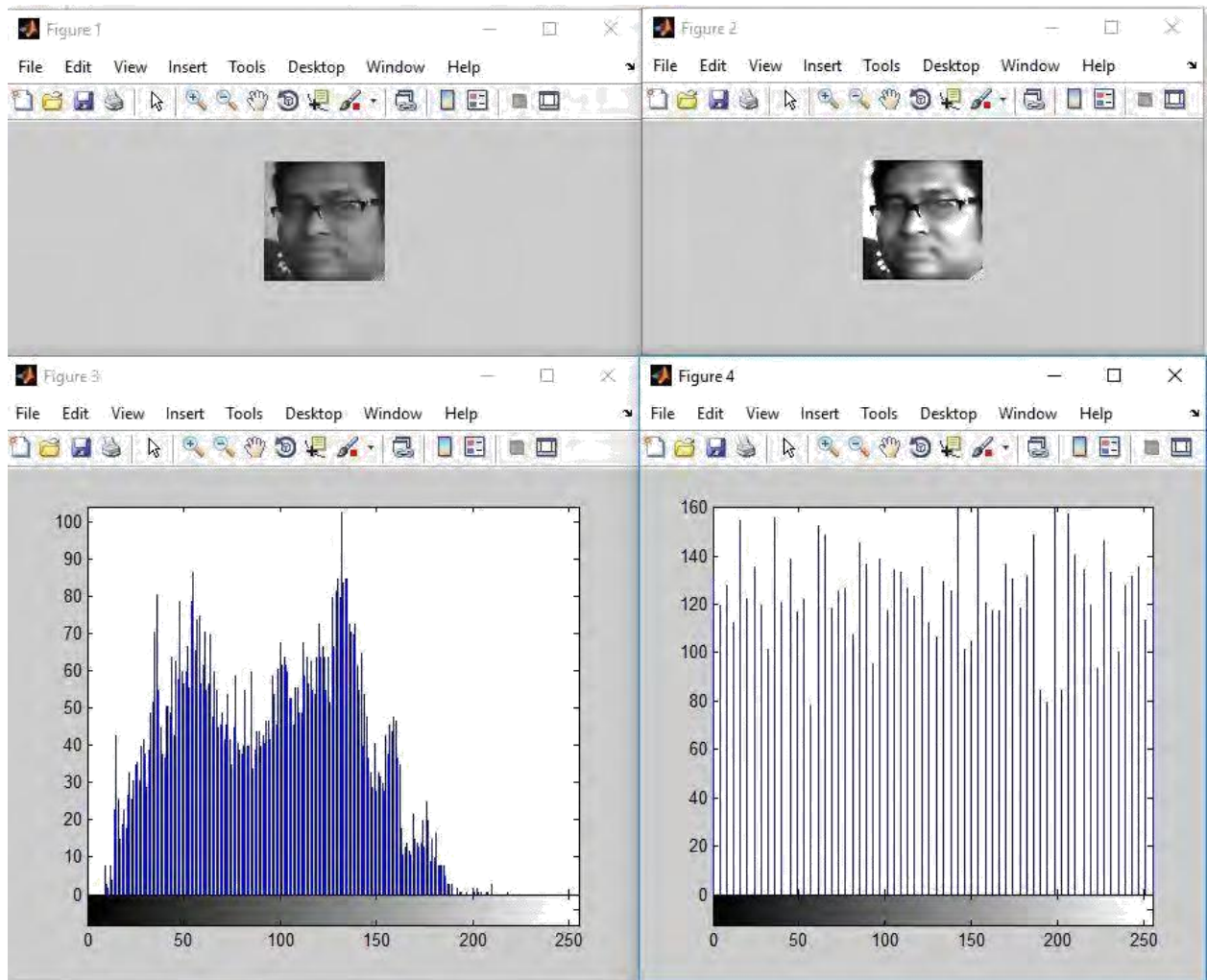


Figure 2.9: Before and after applying Histogram Equalization

2.3 Classification

2.3.1 Gender Detection Classifier

The Gender Detection Classifier stores images for classification purpose and makes classification based on similarity measure. A one dimensional array is responsible for storing the major four top most values of each module after Discrete Cosine Transformation has been done. Subsequently, Manhattan Distance is calculated between each point of the image to the training images by the classifier. The increment of hit or miss score depends on the whether the average distance reaches above or less than the predefined threshold. Resultant male or female depends on the hit or miss score. If hit score is greater for specific labels (Male, Female), then the image should be considered as that label [2].

2.3.2 Local Binary Pattern

This method checks points around a central point and tests the surrounding points if they are greater or less than the center point.

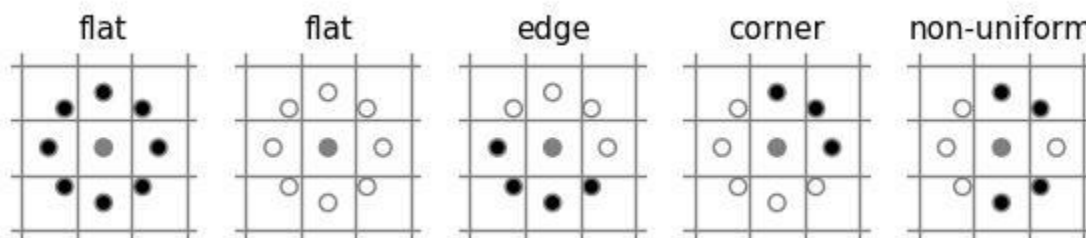


Figure 2.10: Surrounding pixel intensity compared to center pixel [15]

Points around the central point is checked and tested whether the points are greater or less. We can clearly view the above figure in order to clearly demonstrate this. Pixels which are represented by black or white dots are considered as less or more intense respectively compared to the central pixel. In case of being the surrounding pixels are all black or white, then image region is considered as flat and featureless. Corners or edges are the uniform patterns which are continuous groups of black or white pixels. On the other hand, non-uniform patterns are generated if pixels change to and fro between black and white pixels [15].

2.3.3 Local Directional Pattern

LDP is considered as a gray-scale texture pattern. It also portrays the spatial form of a local image texture. Corner response values are computed throughout the directions at the position at each pixel, generating code from the relative strength magnitude [16] [17]. For doing this, a particular pixel's eight directional edge response value is calculated by Kirsch masks in orientation (M0~M7) which are centered on self-position [17]. Figure below [17] shows the illustration.

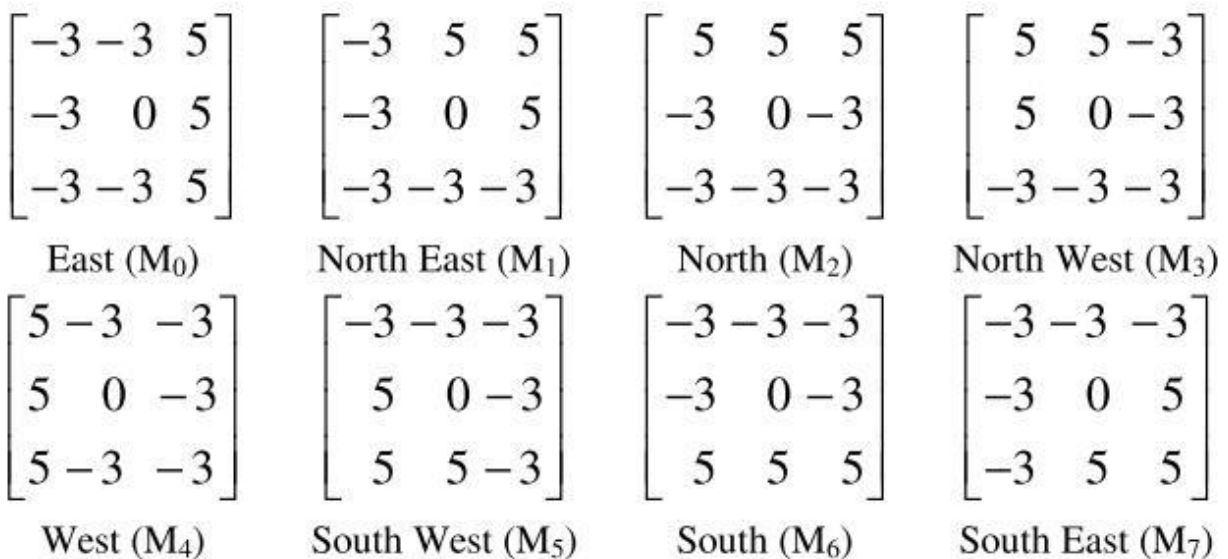


Figure 2.11: Kirsch edge response masks in eight directions [17]

2.3.4 Deformable Spatial Pyramid

A Deformable Spatial Pyramid popular for the calculation of dense pixel correspondences. It is a pyramid graph model which matches uniformity at various spatial degrees. Pixel level matching is improved by this method's regularization. Strict rigidity of the previous traditional pyramids has been overcome by the deformable aspect of the model [18].

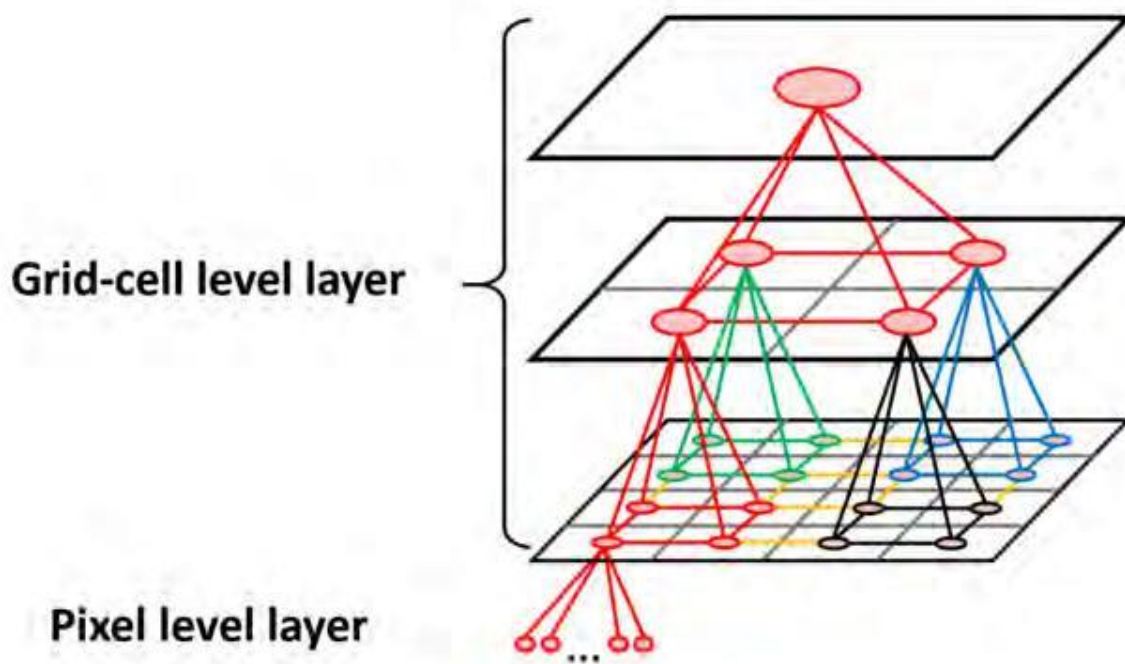


Figure 2.12: The Graph Structure of DSP model [19]

The entire image is divided into four rectangular grid cells. Further it is divided until it reaches to a specific pyramid level number. Finally, a pixel layer is added so that finest pixels have a width of one pixel. The pyramid is represented with a graph. Fixed scale or multi scale matching has been done on the basis of objective. Since the computation time is vast, effective computation is needed.[18]

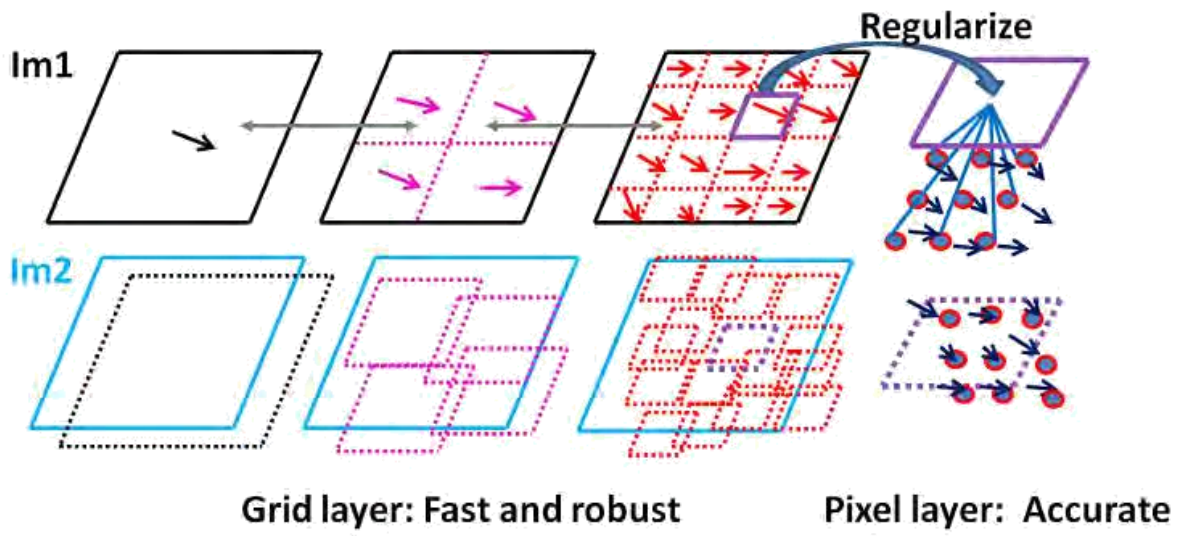


Figure 2.13: DSP matching method by dividing grid cells [18]

Chapter 3

DESIGN APPROACH

3. DESIGN APPROACH

In our design approach, first we have detected the faces from a scene from an input image. Detecting face region from a scene required Viola and Jones method [20], after that, noise reduction from all the faces has been done. From all the de-noised face images, we have applied histogram equalization technique to adjust image intensities in order to enhance the contrast [21]. Thus, it has required less calculation and processes for our research work. Finally all the preprocessed image have been passed through Deformable Spatial Pyramid (DSP) method for final detection. This algorithm performs a correspondence matching of the faces with the knowledge base.

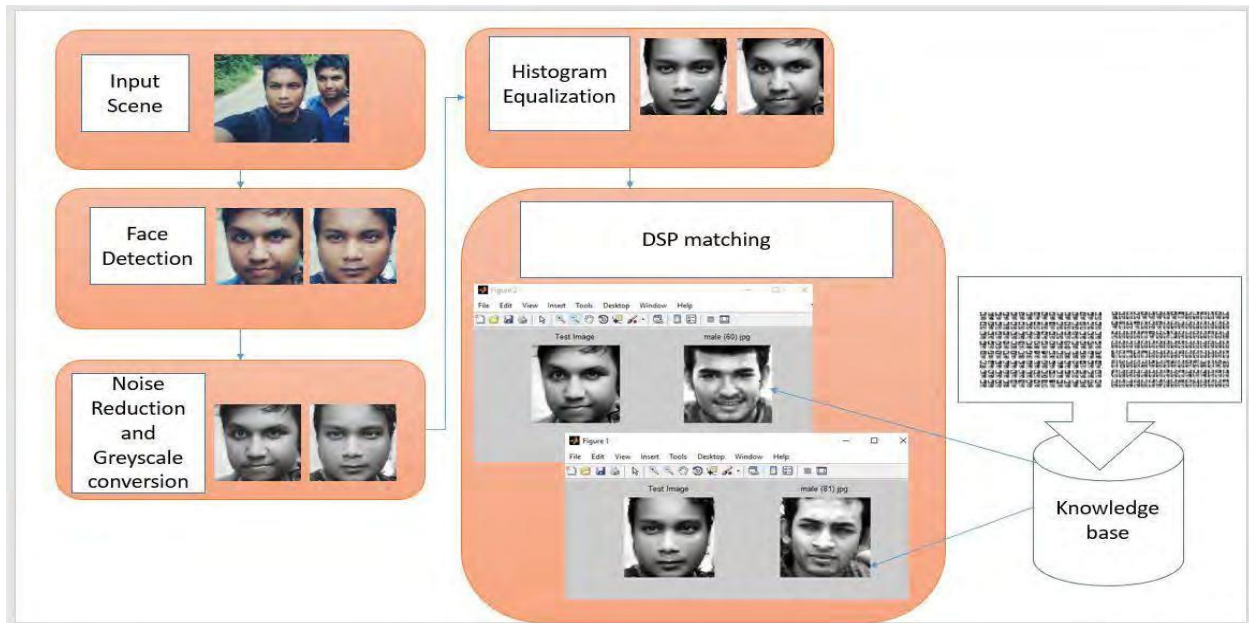


Figure 3.1: All steps required for our proposed system

3.1 Face Detection

Face detection from an input scene is a key process in our research. For the detection part Viola and Jones method has been used. A combination of positive and negative images is needed for training of the cascading function. All the highlighted features have been extracted from the image to match a certain pattern and every feature indicates a single value subtracted from the sum of the pixel under white rectangle from the sum of the pixels from black rectangle [20].

To avoid ending up calculating altogether more than 160,000 features in a single window, the integral value of the image is used which allows the calculation of only the sum of pixels in any given rectangle considering only the four cornering values [22]. It reduces the number of features and allows us for real-time detection. Moreover, for distinguishing between relevant features and irrelevant features, Adaboost technology [23] is used. This technology helps us finding out only the best features among all these features. Furthermore, after detection of the highlighting features, a weighted combination is used for final evaluation a face structure. Adaboost constructs a strong classifier by summing up all the weighted weak classifiers [24].



Figure 3.2: Detecting faces using Viola & Jones

3.2 Noise Reduction and Grayscale Conversion

After successfully detecting the face area from a scene, the region of interest (ROI) is detected. Thus, the detected image has been cropped into 90 x 90 sizes. Digital images are exposed to noise and noise is the result of faults in the image acquisition procedure. Therefore, pixel values do not reflect their true intensities as the real image. Various factors like film grain, CCD detector of the device, damaged film, and electronic transmission of image can introduce noise whereas reducing noise from an image has become a crucial factor for the accuracy of this research. To avoid unwanted result, adaptive filtering over linear filtering is chosen. Wiener Filter is used for this purpose. This approach produces better result than linear filtering in times because of being more selective. Besides, it keeps the edges and other high frequency parts of the image intact. Though, Wiener Filter method requires more computation process than the linear one [14]. Next the RGB image is converted into a grayscale format by eliminating hue and saturation while keeping the luminance intact. It is highly required for the DSP matching algorithm to perform faster.

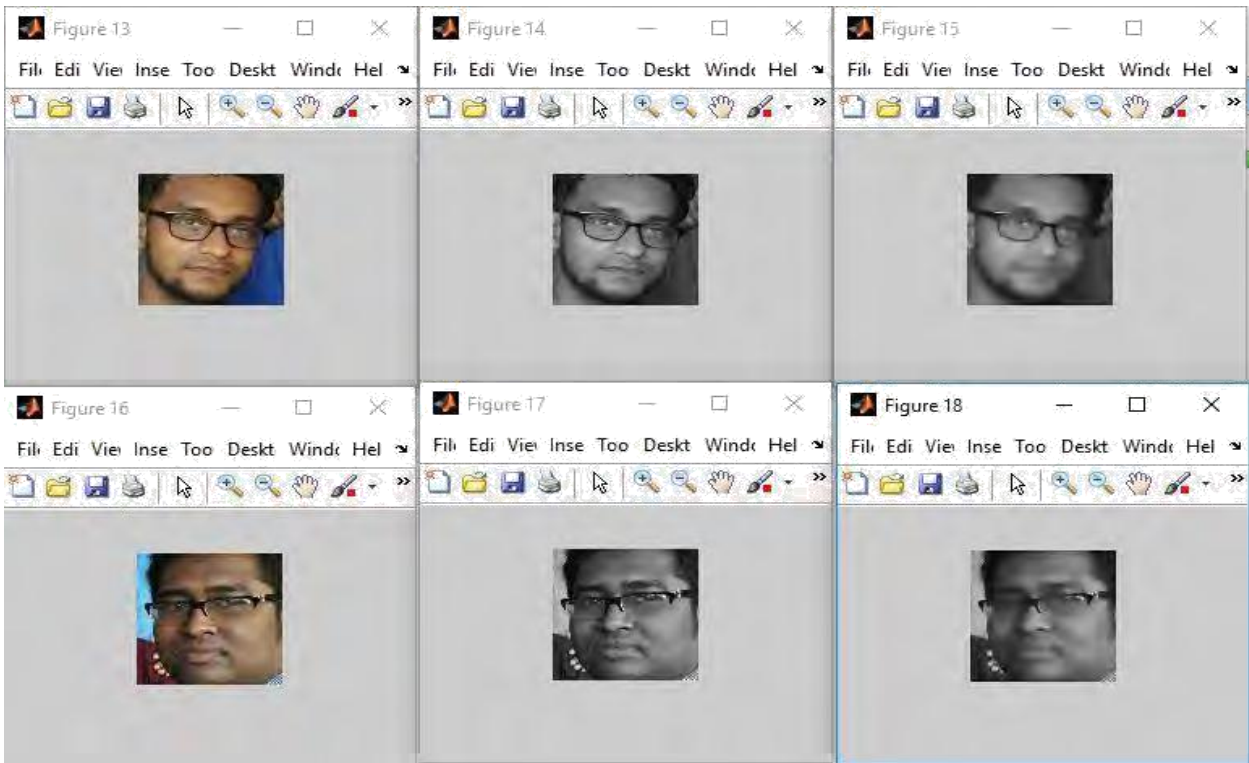


Figure 3.3: Image after Grey Scale conversion and adaptive filtering

3.3 Histogram Equalization

Histogram equalization is used for the enhancement of contrast of an input image. It works by stretching out the intensity range. This method is used in order to gain better accuracy as there are many images whose pixel might cluster around available range of intensities. As a result, a lot of under populated intensities might cause hindrance [25]. After applying Histogram equalization, intensity values are stretched over the intensity range. Let I be an image consisting of M X N matrix. Let the integer values of pixel intensities ranges from 0 to N-1, H denote a normalized histogram of image I and HE denote a histogram equalized image [26]. The normalized histogram H can be defined as,

$$H = \frac{\text{number of pixels with intensity } n}{\text{total number of pixels}} \quad (iii)$$

H = number of pixels with intensity n = total number of pixels. Where, n=0, 1, 2, 3...N-1. The histogram equalized image HE can be defined as,

$$HE_{ij} = \text{floor}((N - 1) \sum_{n=0}^{I_{ij}} H) \quad (iv)$$

Where the nearest integer is rounds down using floor function [26]. Below Figures show the state of the intensities along with a sample face image before and after processing with Histogram equalization method.

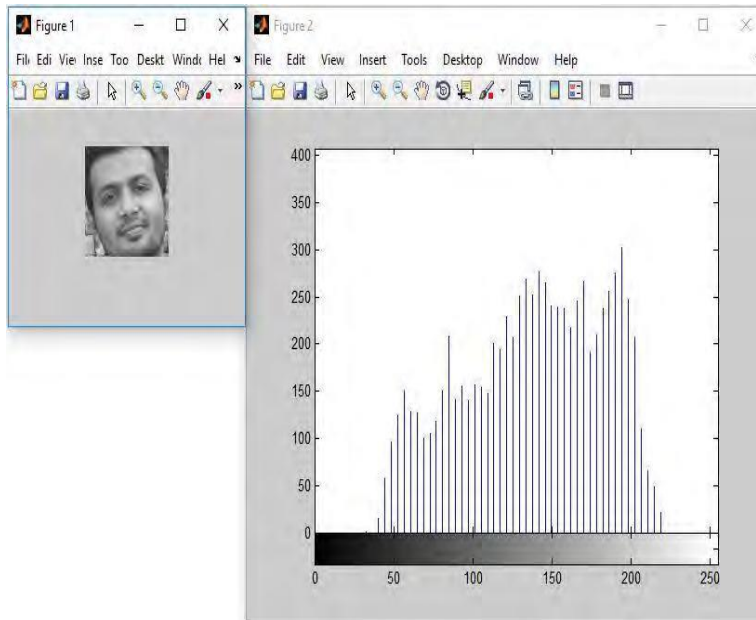


Figure 3.4: Before Applying Histogram Equalization

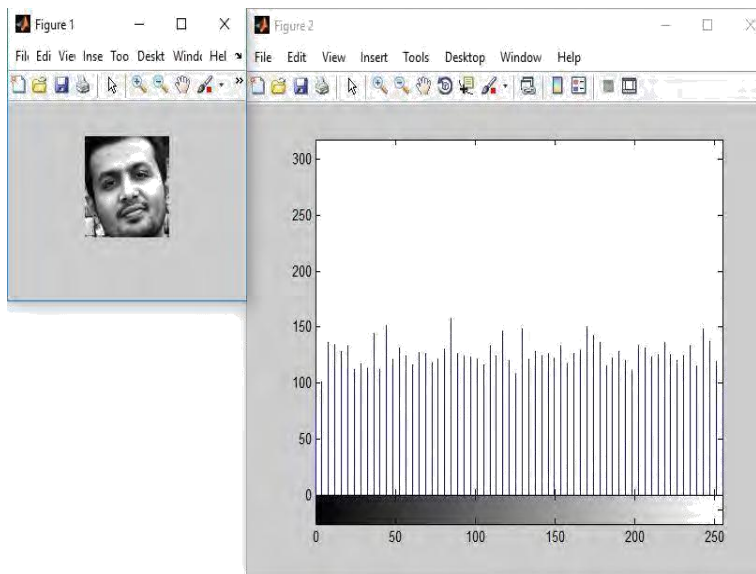


Figure 3.5: After applying Histogram Equalization

3.4 Data Analysis (Gender Detection)

3.4.1 Deformable Spatial Algorithm

For gender detection we have used DSP matching [18] algorithm in our system. Facial image is used to detect a correspondence matching with the given knowledge base. For convenience we have resized the input image [27] to a fixed size.

The pyramid model is firstly achieved through dividing image into four rectangular grids. Same division is also done inside the grids till three levels of pyramid is achieved. Then pixel layer is added at the bottom. Then a graph of size varying spatial nodes is defined. Edges span all levels and each grid is also defined as nodes. Edges spans through levels of pyramid and pixels are only connected to parent grid cell not among them. The SIFT descriptor [28] is used for computing the matching cost.

The DSP matching algorithm is formulated as an energy minimization problem. Loopy Belief propagation [29] is used to minimize the energy equation which is given by,

$$E(t, s) = \sum_i D_i(t_i, s_i) + \alpha \sum_{i,j \in \mathcal{N}} V_i(t_i, t_j) + \beta \sum_{i,j \in \mathcal{N}} W_{ij}(s_i, s_j) \quad (v)$$

At node i the translation node is defined in first image and is the data term which is known as the appearance matching cost for node i at translation. The scale smoothness term is defined with scale variable and is the smoothness term that regularizes the solution. Two constant weights α and β are also associated in the equation.

This algorithm survives different geometric variation and illumination variances as well. Multilevel pyramid helps it to succeed in case of different image resolution. Minimum scores among all scores returned by the DSP matching algorithm determines best match. Classified images in knowledge base are compared to our input via DSP [27]. Figure below represents some recognized male and female faces after dense correspondence matching using DSP algorithm.

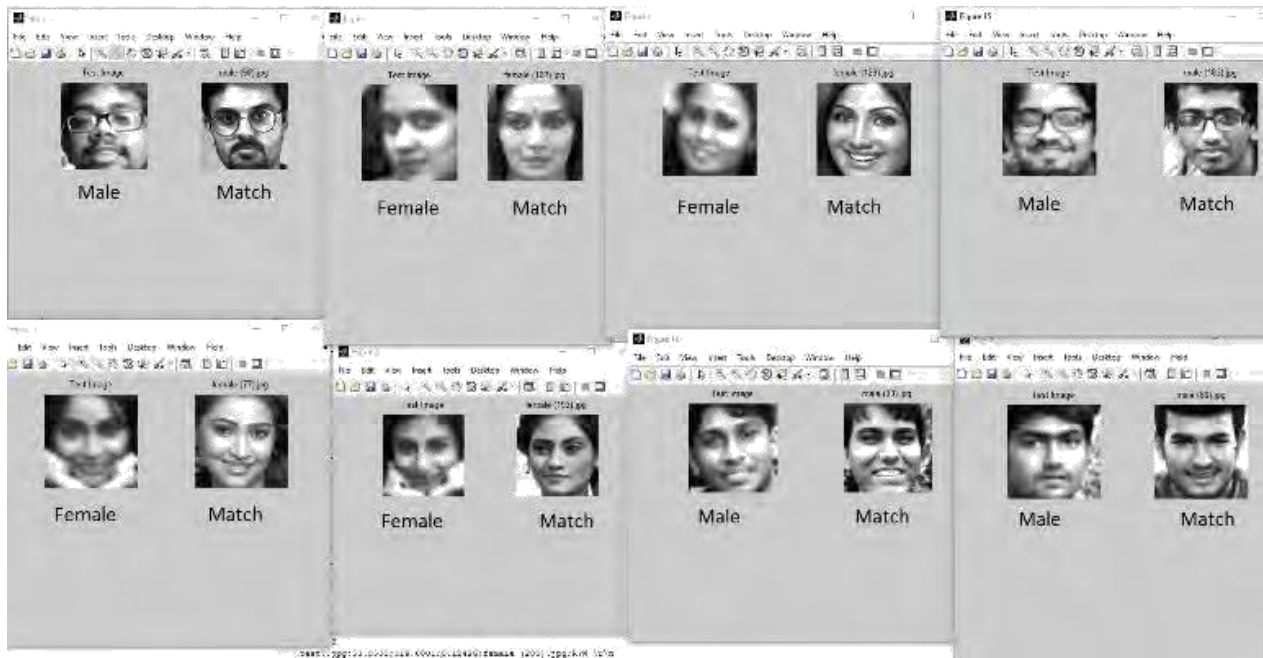


Fig 3.6: Recognized faces using DSP

Chapter 4

RESULT AND DATA ANALYSIS

4. RESULTS AND DATA ANALYSIS

In this paper we used the authors' [18] [30] publicly available code for DSP matching algorithm, vlfeat open source library [31] and opencv as well. For the detection part, Minhas Kamal's [32] publicly available code played a vital role. First, we tried to detect all the faces from a scene. For detection part we obtained good accuracy though it varies time to time. The result of detection part is shown in Table 4.1.

Table 4.1: Accuracy of detection

No of Images	Faces in the Scene	Detected Faces	M (Actual)	F(Actual)	Accuracy %(Face Detection)
1	68	64	38	30	94.12
2	8	8	3	5	100
3	20	19	14	5	95
4	16	11	7	4	68.75
5	56	51	23	28	91.07

After the detection part, we have tried to detect male and female from the detected faces. The recognition accuracy is 83.53% for male and 78% for female without any preprocessing. Performance increases (not every time) slightly after applying all the necessary preprocessing. We have acquired 87.05% accuracy for male and 80.52% accuracy for female, after doing preprocessing. The result is stated in Table 4.2.

Table 4.2: Accuracy of recognition with respect to preprocessing

	Total	Detected (Pre processing)	Detected (No Preprocessing)	Accuracy Without Pre- processing	Accuracy With Preprocess ing
Male	85	74	71	83.53	87.05
Female	77	62	60	78	80.52

Women with relatively short hair are very problematic for detection. So we have taken several images where women have relatively short hair. Some of the sample tested images of women with short hair have been provided in Figure below. Our detection has obtained 80% or less. Table 4.3 illustrates the detection and accuracy of women with relatively short hair. The increment of short hair women to the training set might increase the accuracy.

Table 4.3: Experiment with women with short hair

Sequence	No. of Images	Detected	True Detection	False Detection	Accuracy
1	10	9	8	1	80
2	5	3	3	0	60
3	8	6	4	2	50

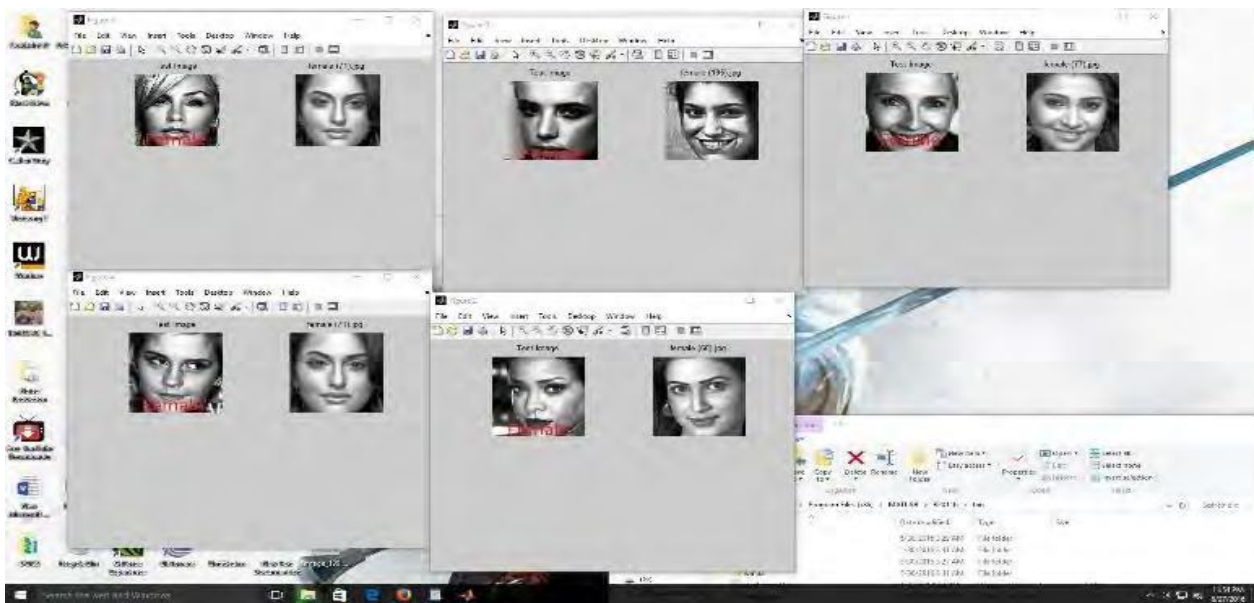


Fig 4.1 Detection of female with short hair.

our experiment with men who are shaved and bald headed.

Table 4.4: Experiment with bald headed men

Sequence	No. of Images	Detected	True Detection	False Detection	Accuracy
1	18	15	13	2	72.22
2	10	9	8	1	80
3	25	22	18	4	72

Table 4.5 demonstrates that if we increase the knowledge base which contains the training data, the success rate of true detection increases. We have taken frontal face image of 6 women and tested against the knowledge base. After elaborating the training data features rate of success drastically increases.

Table 4.5: Success against training data (female)

Test Image	Knowledge Base (Male)	Knowledge Base (Female)	True Detection	False Detection	Accuracy
6 Women	15	15	3	3	50
Same	60	60	4	2	66.67
Same	120	120	4	2	66.67
Same	240	240	5	1	83.33
Same	360	360	6	0	100

We have also tested it against input frontal face images of 15 men and found similar result. Increasing the knowledge base results into better success rate. Table 4.6 shows the experiment of 15 male input of success rate against the training data.

Table 4.6: Success against training data (male)

Test Image	Knowledge Base (Male)	Knowledge Base (Female)	True Detection	False Detection	Accuracy
15 Men	30	30	9	6	60
Same	90	90	10	5	66.67
Same	180	180	12	3	80
Same	250	250	12	3	80
Same	300	300	12	3	80
Same	350	350	13	2	86.67

Chapter 5

CONCLUSION AND FUTURE PLAN

5. CONCLUSION AND FUTURE PLAN

In our paper, we have proposed a gender detection system after detecting all the faces from a scene and have achieved decent accuracy. As we have used Viola and Jones algorithm for the face detection part, we attained some advantages. Face detection part is achieved by OpenCV. We have also used a publicly available dataset from Github [33]. It is extremely fast for feature computation and selects the features efficiently. We have not scaled the image rather than scaling the features. It is a generic detection scheme and possible to use for detection of other types of objects. For gender detection of facial image part, we have used DSP algorithm which shows grate accuracy. However, the accuracy sometimes gets worse if the test images varies greatly from the training images. But the main drawback is that it works slowly while calculating and matching. We have used the system in a supervised manner. It does not work robustly in case of rotation and scale variation. For this reason, we are hoping to solve this issue in our further work.

References:

- [1] Rahman, M. M., Rahman, S., Dey, E. K., & Shoyaib, M. (2015). A gender recognition approach with an embedded preprocessing. *International Journal of Information Technology and Computer Science (IJITCS)*, 7(7), 19.
- [2] Dey, E. K., Khan, M., & Ali, M. H. (2013). *Computer Vision Based Gender Detection from Facial Image*. LAP LAMBERT Academic Publishing.
- [3] Lapedriza, A., Maryn-Jimenez, M. J., & Vitria, J. (2006, August). Gender recognition in non controlled environments. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on* (Vol. 3, pp. 834-837). IEEE.
- [4] Makinen, E., & Raisamo, R. (2008). Evaluation of gender classification methods with automatically detected and aligned faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3), 541-547.

- [5] Alrashed, H. F., & Berbar, M. A. (2013). Facial gender recognition using eyes images. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(6), 2441-2445.
- [6] McKenna, S., Gong, S., & Raja, Y. (1997, July 11). Face Recognition in Dynamic Scenes. Retrieved October 21, 2017, from <http://vision.eecs.qmul.ac.uk/oldpages/research/papers/bmvc97/node2.html>.
- [7] Friba, B., & Kklbeck, C. (2001). Real-time face detection using edge-orientation matching. Edited by G. Goos, J. Hartmanis and J. van Leeuwen, 78.
- [8] Jesorsky, O., Kirchberg, K. J., & Frischholz, R. W. (2001, June). Robust face detection using the hausdorff distance. In *International Conference on Audio-and Video-Based Biometric Person Authentication* (pp. 90-95). Springer, Berlin, Heidelberg.
- [9] N. (2001, May 23). Unmasking the Face on Mars. Retrieved October 21, 2017, from https://science.nasa.gov/science-news/science-at-nasa/2001/ast24may_1/.
- [10] Face Detection using Haar Cascades. (n.d.). Retrieved October 21, 2017, from https://docs.opencv.org/master/d7/d8b/tutorial_py_face_detection.html#gsc.tab=0.
- [11] Priya, G. H., Venkatesh, B., Srivani, A., Sowmya, G., & Rajasekhar, N. (2010). Noise Removal in Image Using LPG-PCA (Local Pixel Grouping Principle Component Analysis) Algorithm. *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)* e-ISSN, 2278-2834.
- [12] Corso, J. (n.d.). Linear Filters and Image Processing. Retrieved October 21, 2017, from https://web.eecs.umich.edu/~jjcorso/t/598F14/files/lecture_0924_filtering.pdf.
- [13] Median Filtering. Retrieved October 21, 2017, from https://www.cs.auckland.ac.nz/courses/compsci373s1c/PatricesLectures/Image%20Filtering_2up.pdf
- [14] M. (n.d.). Documentation. Retrieved October 21, 2017, from <http://www.mathworks.com/help/images/noise-removal.html>.
- [15] Local Binary Pattern for texture classification¶. (n.d.). Retrieved October 21, 2017, from http://scikit-image.org/docs/dev/auto_examples/features_detection/plot_local_binary_pattern.html.
- [16] Ahmed, F. (2012). Gradient directional pattern: a robust feature descriptor for facial expression recognition. *Electronics letters*, 48(19), 1203-1204.

- [17] Jabid, T., Kabir, M. H., & Chae, O. (2010, January). Local directional pattern (LDP) for face recognition. In Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on (pp. 329-330). IEEE.
- [18] Kim, J., Liu, C., Sha, F., & Grauman, K. (2013). Deformable spatial pyramid matching for fast dense correspondences. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2307-2314).
- [19] Hur, J., Lim, H., Park, C., & Chul Ahn, S. (2015). Generalized deformable spatial pyramid: Geometry-preserving dense correspondence estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1392-1400).
- [20] Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-I). IEEE.
- [21] Vij, K., & Singh, Y. (2009). Enhancement of images using histogram processing techniques. *Int. J. Comp. Tech. Appl*, 2(2), 309-313.
- [22] Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision*, 57(2), 137-154.
- [23] Y. Freund & R. Schapire, Freund, Y., & Schapire, R. E. (1995, March). A decision-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory* (pp. 23-37). Springer, Berlin, Heidelberg.
- [24] Palla, K., & Kalaitzis, A. (n.d.). Robust Real-time Face Detection by Paul Viola and Michael Jones, 2002 Presentation by Kostantina Palla & Alfredo Kalaitzis School of Informatics University. Retrieved October 21, 2017, from <http://slideplayer.com/slide/8151084/>
- [25] Histogram Equalization¶. (n.d.). Retrieved October 21, 2017, from https://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/histogram_equalization/histogram_equalization.html.
- [26] Histogram Equalization.(n.d). Retrieved October 21, 2017, from http://www.math.uci.edu/icamp/courses/math77c/demos/hist_eq.pdf
- [27] S. N. Pinku, R. Biswas, M. Mostakim. A Novel Approach to Recognize the Speed Limit Traffic Sign Based on Deformable Spatial Pyramid. Unpublished.
- [28] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.

- [29] Felzenszwalb, P. F., & Huttenlocher, D. P. (2006). Efficient belief propagation for early vision. *International journal of computer vision*, 70(1), 41-54.
- [30] J Kim, J., Liu, C., Sha, F., & Grauman, K. (2013). Deformable Spatial Pyramid Matching for Fast Dense Correspondences. Retrieved April 17, 2017, from <http://vision.cs.utexas.edu/projects/dsp/> CVPR 2013.
- [31] VLFeat.org. (n.d.). Retrieved April 17, 2017, from <http://www.vlfeat.org/>.
- [32] M. (2016, October 08). MinhasKamal/GenderRecognizer. Retrieved April 17, 2017, from <https://github.com/MinhasKamal/GenderRecognizer>.
- [33] M. (2015, December 25). MinhasKamal/GenderRecognizer. Retrieved October 21, 2017, from <https://github.com/MinhasKamal/GenderRecognizer/tree/master/src/res/trainingData>.