Gender Classification of Consumer Face Images using Gabor Filters

Khurram Zeeshan Haider, Tabassam Nawaz, Hafiz Adnan Habib, Muazzam Maqsood, Tauqeer Ul Amin*

47080, University of Engineering and Technology, Taxila, Pakistan

50700, University of Gujrat, Gujrat, Pakistan*

Abstract:

Gender classification of consumer face images has been investigated using Gabor Filter method. Consumer images when taken into analysis such kind of classification make the task difficult for their diversity in appearance, pose, gesture and illumination conditions. The first step of this study is normalization of input image and next steps are extracting feature vector using Gabor filter which is further used as input for support vector machine SVM. SVM has proved to be finest supervised learning methodology for gender classification under applied experimental setup. Three face image databases [MUCT, AR and IMM] were utilized for training and testing. These databases were found precisely close to consumer face images for their variety of image acquisition procedure like use of different camera distance, subject pose, background and illumination.

Key words:

Consumer research; Gender classification; Image normalization; Gabor filters; SVM;

1. Introduction:

Male and female perception of purchasing either self visit or online shopping has discernable variation according to earlier studies [1]. Consumer behaviour is dependent on the culture, way of life, societal norms, customs, mores, traditions and daily life of a person intended to shop. The success or fall down depends upon not only his/her manner but the gender. Sex, gender identity, gender mind-set are unique limitations for consumer behaviour. [2]

There is a major difference between male and female towards purchasing even of organic food products [3]. According to the study carried out by a renowned car agent "Autogenie" which collected around 6000 new vehicles purchased in 2013 says men preferred sedans and Ute's more and women bought small cars and SUVs.

In consumer analysis gender identification can be the first step. A gender classification system uses face portion in a given image, makes computation/analyse it pixel by pixel and indications gender of person. A well structured gender classification system always increases the performance of lots of other computer applications like face detection and smart Human Computer Interface HCI. Even face reflection sequences can be classified with high level of confidence. [4]

Consumer faces are usually obtained from surveillance cameras at educational institutes, shopping centres, airports (local or international), residential colonies, enterprise bodies and other public areas. The surveillance cameras produce non-regular images for its non-uniformity in capturing speed, distance from the object, lens quality, medium of data communication, camera angle, brightness and face angle of the subject.

[5] has performed a successful experimentation to increase accuracy of face recognition system by separate training of male and female face images.

Humans can distinguish the gender by just looking at someone's face but how this can be solved with an automated system? To answer this question let us recall a face so we have to compute face features i.e. eyebrows, hair, skin and overall shape of a face. This task becomes more critical for consumer faces which are captured in real-world and real-time environment.

Consumer images come in variation of pose, illumination and ethnicity so the image is put into normalization phase first then Gabor filter is applied which is believed to be similar to human visualization system [6]. Filtered image is then sliced and histogram bins are obtained to calculate feature vector which is elaborated in detail in later sections. Training and testing is performed using Support Vector Machine and all experimentation has been performed using MUCT, AR and IMM face datasets. Accuracy achieved for MUCT, AR and IMM is up to 97.2%, 98% and 96.8% respectively.

2. Materials and methods

2.1 Datasets:

MUCT, AR and IMM have 3755, 4000+ and 240 number of images respectively. Out of which 80 male and 80 female images (for MUCT, AR and IMM) were used in training of SVM. The proposed scheme was tested for 250 randomly selected faces other than training images for MUCT and AR datasets. While 160 remaining images of IMM dataset were brought into testing.

To elaborate gender detection capability of proposed method, experimentation was carried out on 3 chosen

Manuscript received February 5, 2016
Manuscript revised February 20, 2016

datasets for their diversity in pose, female percentage and number of illumination/lightning sets among renowned 6 described below:

Table 1: Human face databases and their attributes

Database name	# of subjects	Image color	Background	Female subjects	light sets	Images
MUCT [7]	276	Full color	Blue shaded	51%	10	3755
AR [8]	126	Full color	White	41%	04	4000 +
IMM [9]	40	Mixed	Green	20%	02	240
BioId [7]	27	Not colored	Office	40%	01	1521
PUT [10]	200	Full color	unlike	11%	01	10 K
xM2VTS [11]	295	Full color	Blue	47%	01	1180

Our consideration is consumer face images so we need diversity in colour, illumination and percentage of female. MUCT, AR and IMM were chosen for experimentation purpose. xM2VTS was not selected for its single illumination set and same background. We did not choose PUT for its low percentage of female subjects i.e. 11%. BioId has good female subject percentage up to 40 but it has repetition of same background with single lightning set.

MUCT has 51% female subjects, 11 numbers of lightning sets and shaded background. AR and IMM have 4 and 2 lightning sets with 41% and 20% of female subjects respectively which is represented in Figure 1:

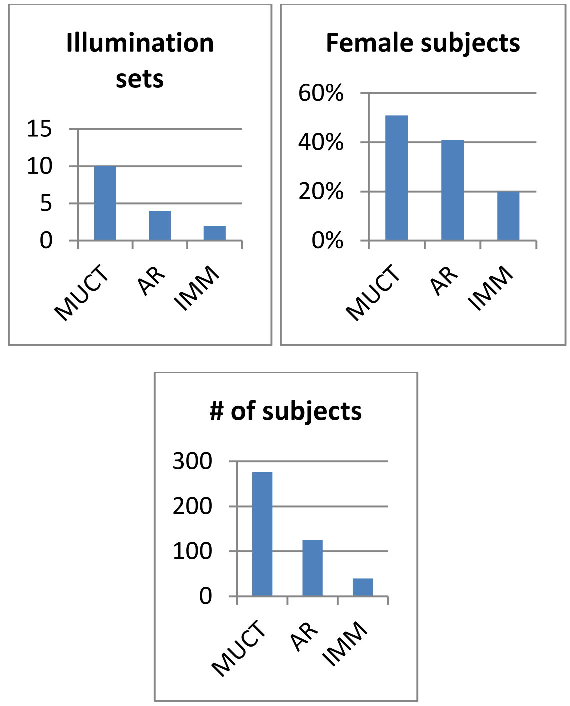

Figure 1: Representation of number of illumination sets, percentage of female subjects, total subjects in MUCT, AR, IMM which are used for experimentation.

Some consumer face examples are shown below which vary in lighting, background, age and ethnicity etc.:

Figure 2: Arbitrary captured consumer faces (up); MUCT, AR and IMM shown at 1st, 2nd and 3rd rows respectively (last three rows)

For the identification of consumer gender, three databases i.e. MUST, AR and IMM have been selected for their resemblance with real-time acquired/captured faces. Few raw images from the databases used, are shown in Figure 2.

2.2 Proposed Algorithm:

Idea is to pre-process the input image as shown in detailed in Figure 3. The normalized image is passed to the feature extraction section where orientation, scale and band are set then Gabor filter is applied and finally edge histogram is calculated to represent feature vector. Training and testing has been performed separately for both gender images and feature vectors are stored in feature repository. In the last classification has been performed using Support Vector Machine in training and testing. Detail of this mentioned methodology is described in later sections along-with results and discussion.

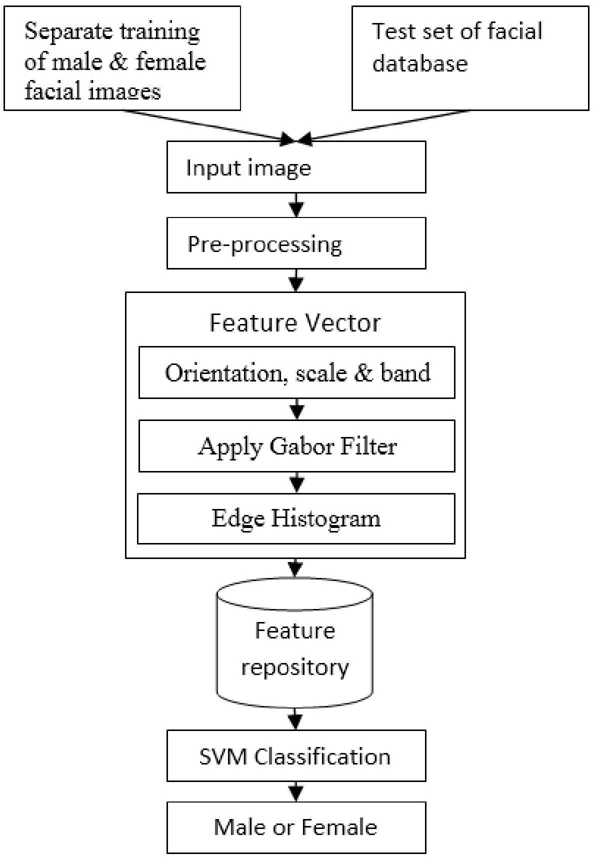

Figure 3: Block diagram of proposed method

For both training and testing the input image is pre-processed. Pre-processing scheme is elaborated in the following figure 4:

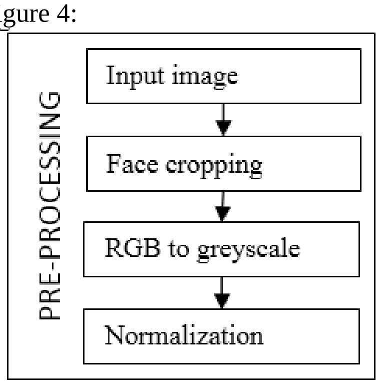

Figure 4: Pre-processing steps

Pre-processing steps include:
Face cropping done with viola and Jones
RGB to greyscale
Illumination normalization
Detected image using Viola and Jones method is cropped, scaled, converted into greyscale and then normalized. Greyscale images contain only intensity information so working with them becomes easy and timesaving Figure 5.

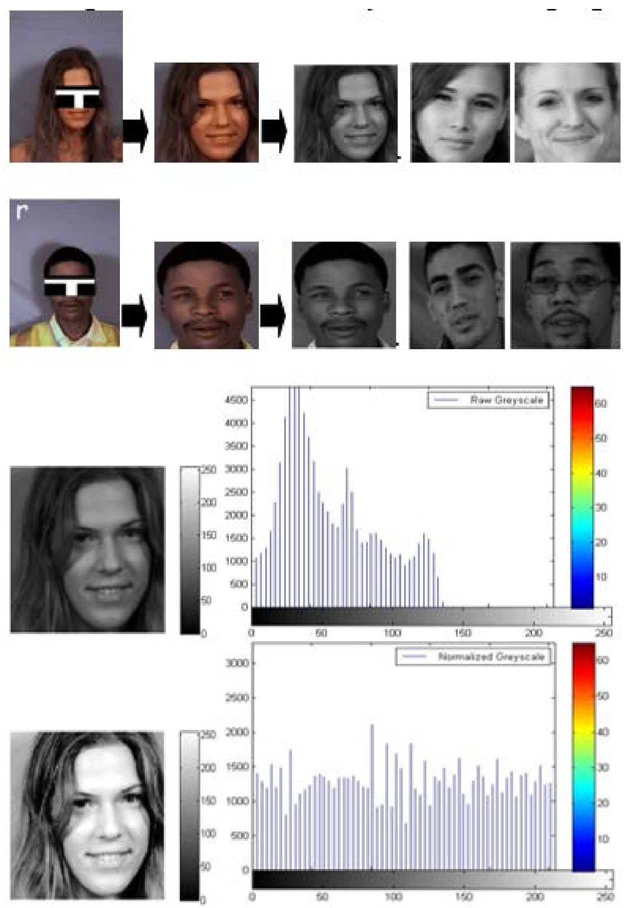

Figure 5: Upper row (female), 2nd row (male) →Viola & Jones classifiers (Left), Detected and cropped face (middle) and conversion colour to greyscale (right),
Last row →Original greyscale (left) and normalized one (right) [on the scale 0 to 255] and their respective histograms

After pre-processing the image is ready for feature vector extraction. While training feature vector is labelled as:

- 'm' for male &
- 'f' for female

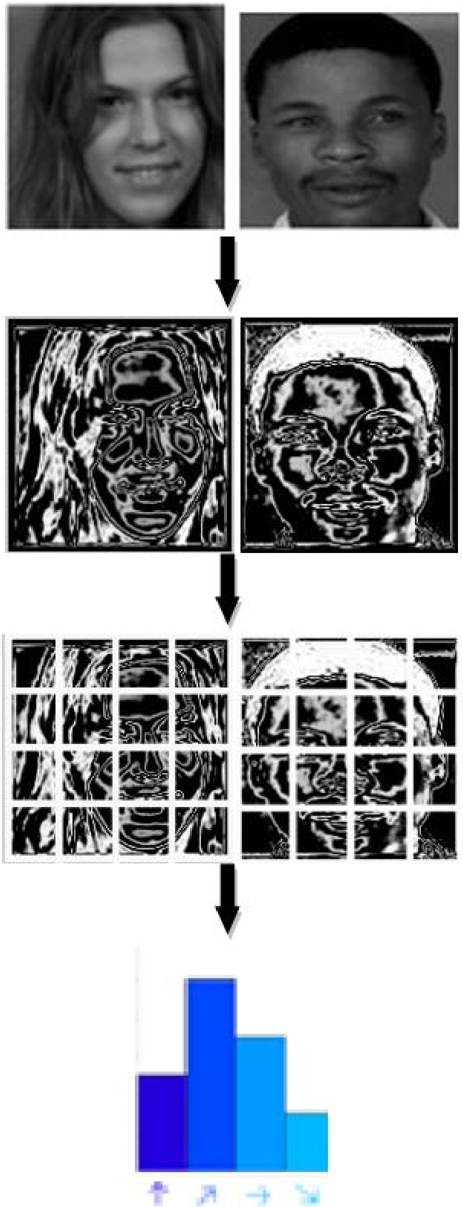

Figure 6: Left to right: Face detection + Normalization, Gabor filter application, slicing 4x4 sub-images (here) and 4-bin edge orientation histogram generic representation

2.2.1 Face Detection and Cropping

The first step is to identification of face portion in a given image. Here are some encouraging and adverse factors of different practices of “face detection” existing so far:

Table 2: Human “face detection” practices and their encouraging as well as adverse factors

Face detection	Encouraging factors	Adverse factors
Using colour [12]	Easy computation for face segmentation using colour matching	All skin colours cannot be matched Under changing lightning conditions performance slows down
movement of face and Blink eye [13]	Computation for video is easy because of its moving frames nature	Scheme be unsuccessful when background is in motion too
Hybrid [14]	If one not succeeds other factor may yield correct outcome.	Face detection may not be done correctly when background resembles the face colour.
Geometric model: Edge-Orientation Matching [15] Hausdorff Distance [16,17]	It was revolutionary technology at 2000 era when these methods out class in accuracy.	Varying lightning decreases the accuracy.
Viola and Jones method [18] (Weak classifier cascades)	Most excellent in results and popular now a days.	single face can be detected more than one time

Comparing different practices pulled us to use Viola and Jones method for its accuracy. Feature types employed by Viola and Jones are shown in the figure below:

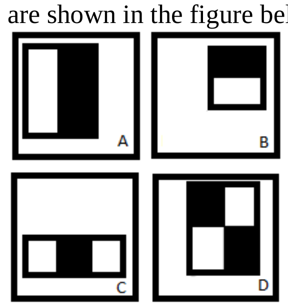

Figure 7: A & B edges, C vertical line and D diagonal features

In figure 7, A and D are used to locate vertical edges and diagonal respectively while C resembles to nose region and B resembles eyes. Benefits of this scheme are:

- Fast feature computation
- Efficient use of feature choice
- Scale as well as location invariant
- We do not scale image instead features could be

2.2.2 Image Normalization

Greyscale image enhancement and normalization can be understood with this simple example; the technique is also referred as Contrast stretching. Consider we have an image whose:
Previous minimum intensity = 45
Previous maximum intensity = 200
Required minimum intensity = 0
Required maximum intensity = 255
To obtain normalized image subtract 45 from each of pixel. It will make the new range from 0 to 155. Then multiply each of pixel intensity with 255/155making the new range 0 to 255. Previous pixel intensity histogram is shown in the last row (left) along-with normalized histogram (right) Figure 5.

2.2.3 Gabor Filter

Gabor filter has been applied to the normalized images. An input gray-scale image is densely filtered by a series of Gabor filters at different scaling and tilt angle theta ϴ. Therefore,
The filters appear in 4 angles and 16 scales (16x4 = 64 maps) that are in order of 8 bands. Band 1 is used with filter size of 7x7 with Lambda λ=3.5, Bandwidth σ=2.8 and theta ϴ=900 Figure 6. Equation applied to our input image is:

$$G_{x,y} = \exp\left(\frac{-(x\cos\theta + y\sin\theta)^2 + \gamma^2(-x\sin\theta + y\cos\theta)^2}{2\sigma^2}\right) \times \cos\left(2\Pi\frac{1}{\lambda}(x\cos\theta + y\sin\theta) + \phi\right) \quad (1)$$

2.2.4Slicing and SOBEL edge detection

The filtered images then sliced into 4x4 The edge detectors are based on simple 2x2 templates do not characterize edges well [19]. Therefore, following 3x3 template SOBEL edge detectors were used:

$$S_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} S_2 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$S_3 = \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} S_4 = \begin{bmatrix} -1 & 2 & 2 \\ -1 & -1 & 2 \\ -1 & -1 & -1 \end{bmatrix}$$

Figure 8: SOBEL filters S1 to S4 detect horizontal, vertical, 45O and 135O degree edges respectively

SOBEL filters S1 to S4 detect horizontal, vertical, 45O and 135O degree edges respectively which are shown in Figure 8. Here comes an important step of finding feature vector of the input image. The above image is sliced into 4x4, 6x6 and then 12x12 sub-images and the local edge histogram of each slice has been computed. For example for 6x6 sub-images we obtain 36 slices. For each slice we have used 4 bins. So edge histogram of 16 slices uses 64 bins.

2.2.5 Feature Vector Calculation

The histograms of all clipped regions have to be represented over a single vector which is obtained by appending all (column) values in a single cell array. A(:) is used in MATLAB as an structure array preserving its shape as before and returns its output in comma separated list. The acquired vector is to serve as support vector as an input of support vector machine SVM. Image obtained as output of SOBEL filtering is sliced as 1x1, 4x4 and 6x6 and then histogram is taken whose output is appended over a range of 1 to 4 for four types of edges. This gives the feature vector of length when sliced as 4x4 is 4x4x4=64 units and for 6x6 slices it is 6x6x4= 144 units. Extracted feature vectors are stored in feature repository and are used for training as well as testing with SVM.

2.2.6 SVM

SVM [20] falls under category of supervised learning. Gender classification is a task of separating both classes according to their features.
In proposed methodology training has been performed separately for both genders i.e. male and female. If D is a dataset for training, it is considered as a set of n points of the form:

$$D = \left\{(x_i, y_i) \mid \mathrm{x_i} \in \Re^P, y_i \in \{-1,1\}\right\}_{i=1}^{n} \quad (2)$$

Where y_i = {1,-1} and x_i is a P-dimensional Real vector.

3. Results and Discussion

Precision or specificity tells the ability of a system to identify only those images which are relevant amongst all images included in classification. Precision, accuracy and correctness of output in the field of gender classification mainly depend upon the database used for training and testing. Consumers' images in real-time acquisition environment change in sharpness, brightness, age and ethnicity that is why in the proposed experimental procedure MUCT, AR and IMM databases have been chosen for their varying conditions.

In the proposed method, SVM is applied on the every image present in the repository first for the female face images and then for the male face images. The class association for both male and female information is stored for training. The usability of this approach is that, after determining the edge histogram as feature vector we only need to classify using support vector machine. Overall, the classification accuracy is measured in the terms of the precision by applying the formula:

$$P_{[MUCT]} = \frac{N_{A[MUCT]}}{N_{R[MUCT]}} = \frac{243}{250} = 0.972 \qquad (3)$$

$$P_{[AR]} = \frac{N_{A[AR]}}{N_{R[AR]}} = \frac{245}{250} = 0.980 \qquad (4)$$

$$P_{[IMM]} = \frac{N_{A[IMM]}}{N_{R[IMM]}} = \frac{155}{160} = 0.968 \qquad (5)$$

Where

P = Precision, N_A =Accurate images found, N_R = All tested images

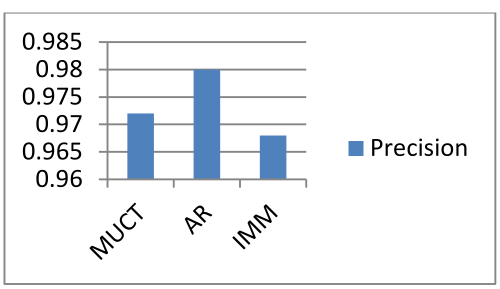

Figure 9: experimental results in terms of precision in MUCT, AR, IMM.

Proposed procedure With IMM as dataset shows 5 images were not recognized correctly among 160 and yielded precision .968. IMM images are different in expressions and pose also among 240 images, there are 40 female subjects whose percentage is 20% only. $P_{[MUCT]}$ & $P_{[AR]}$ are 0.972 and 0.980 respectively, MUCT has more illumination sets and uneven background which is close representation of consumer images as compared to AR and others.

For comparison purpose of state of art practices, following is the table along with graph of classification technique plus its feature extraction technique and its corresponding percentage accuracy. Accuracy of the presented method is among most excellent outcomes so far:

Table 3: Feature Extraction Technique+ Classifier Used with percentage accuracy

Feature Extraction Technique+ Classifier Used	*%age Accuracy*
2D PCA+SVM	90
ASM&LBP+SVM	93
DWT+NMM	90
G-LBP+SVM	87.1
LBPH+NN Euclidean	96.8
Scale-Invariant feature transform +Bayesian	81.1
Proposed [IR]	98

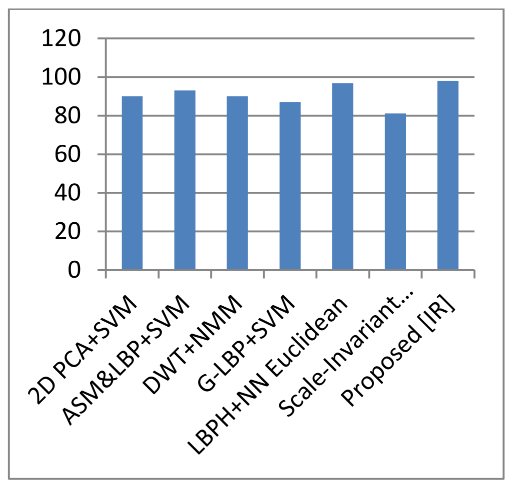

Figure 10: Percentage accuracy (Y-axis) of Feature Extraction Technique+ Classifier Used

Let's discuss favourable and unfavourable factors of different feature extraction techniques along with their classification mechanism presented in Figure 12. Two dimensional Principal Component Analysis 2D PCA for feature extraction and Support Vector Machine SVM for classification was used by [21]. Two favourable factors of this method were:

I. Multi-view face classification
II. Removal of redundant features

But for higher occlusion conditions the accuracy lowers to 86% otherwise it is 90%.

Active Shape Models belong to statistical methods in which an object iteratively deforms to become in shape of a new object. This method was first presented in 1995 byTim Cootes and Chris Taylor. This technique along with the use of Local Binary Pattern LBP is stated by [22]. They have used SVM for classification. The result accuracy is 93% for expression detection and it is very robust to external condition but if image contains composite background, it has adverse effect on feature extraction but our method is robust to composite background with the use of Viola and Jones techniques and is acceptable for real-time images of consumers.

Another feature extraction methodology is Discrete Wavelet Transform DWT which is used with Neural Network Model NNM for classification [23]. Accuracy is achieved up to 90% and Euclidean distance performance is better than L1 distance measure meanwhile accuracy is dependent upon the combination of more features. Our method uses simple but strong feature extraction technique by finding histogram of an image filtered through Gabor as well as high accuracy is attained.

Gradient Local Binary Pattern G-LBP and SVM were used for feature vector measure and classification respectively by [24]. It is robust to low quality images and illumination variations and achieved accuracy up to 87.1%. This method is close to ours but our results are more accurate and reliable because of the use of Gabor filters.

A similar method LBP-Histogram used with Neural Network Euclidean [25] achieved impressive accuracy rate ranging from 95.5% to 98.1%. For the problem under consideration SVM is a better choice for its coverage of Global Minima instead of Local Minima and time for training and testing with larger dataset.

Scale-Invariant Feature Transform with Bayesian [26] used internet images for feature extraction and classification but achieved accuracy rate just up to 81.1%. We have taken care of scale (proper use of Gabor band) as well as orientation denoted by theta and bandwidth, hence we able to achieve better outcome in terms of robustness and accuracy.

4. Conclusion

The presented method addresses consumer face images for its gender recognition. Pre-processing methods are discussed here for extracting face, converting it in greyscale and then using histogram equalization technique to normalize it. Gabor filtering along with slicing and histogram technique is proposed for determining feature vectors which are further used for training and testing of subject images. Results show the precision of the proposed method [MUCT, AR and IMM] 0.972, .98 and 96.8 which is quite promising and acceptable. Proposed method has proved itself as a better local plus global feature extraction as everything of the face including hair as well as expressions are taken into account which is fine for real-time consumers' images having variations in age, pose, expressions and illumination as presented in our datasets i.e. MUCT, AR and IMM.

5. Authors Contributions

Khurram Zeeshan Haider projected the research, performed the experiments, analyzed the data and set-up the figures. Dr. Tabassam Nawaz helped in scripting the paper. During the experimentation and paper development, suggestions and critical reviews were provided by Dr. Tabassam Nawaz, Dr. Hafiz Adnan Habib, Muazzam Maqsood and Muhammad Tauqeer Ul Amin. Requests for resources and study materials should be addressed to Khurram Zeeshan Haider. All authors have read and approved this manuscript.

6. Conflict of Interest

The authors declare no conflict of interest.

References

[1] Kolsaker, Ailsa, and Claire Payne. "Engendering trust in e-commerce: a study of gender-based concerns." Marketing Intelligence & Planning 20.4 (2002): 206-214.

[2] Atkins, Kelly Green, and Youn-Kyung Kim. "Smart Apparel Shopping A Multidimensional and Gender-Neutral Measure." Clothing and Textiles Research Journal (2015): 0887302X15602097.

[3] Wee, Chiew Shi, et al. "Consumers perception, purchase intention and actual purchase behavior of organic food products." Review of Integrative Business and Economics Research 3.2 (2014): 378

[4] Smith, Daniel F., Arnold Wiliem, and Brian C. Lovell. "Face Recognition on Consumer Devices: Reflections on Replay Attacks." Information Forensics and Security, IEEE Transactions on 10.4 (2015): 736-745

[5] Saatci, Yunus, and Christopher Town. "Cascaded classification of gender and facial expression using active appearance models." Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on. IEEE, 2006

[6] A. Ramachandran, "Modified Gabor Filter based Vehicle Verification", International Journal of Computer Applications® (IJCA) (0975 – 8887) National Conference cum Workshop on Bioinformatics and Computational Biology, NCWBCB- 2014

[7] Milborrow, Stephen, John Morkel, and Fred Nicolls. "The MUCT landmarked face database." Pattern Recognition Association of South Africa 201.0 (2010).

[8] Martinez, A. R., and R. Benavente. "The AR face database, 1998." Computer Vision Center, Technical Report 3 (2007).

[9] Nordstrøm, Michael M., et al. The IMM face database-an annotated dataset of 240 face images. Technical University of Denmark, DTU Informatics, Building 321, 2004.

[10] Kasinski, Andrzej, Andrzej Florek, and Adam Schmidt. "The PUT face database." Image Processing and Communications 13.3-4 (2008): 59-64.

[11] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, XM2VTS: The Extended M2VTS Database. AVPBA, 1999.
[12] Störring, Moritz. Computer vision and human skin colour. Diss. Computer Vision & Media Technology Laboratory, Faculty of Engineering and Science, Aalborg University, 2004.
[13] Królak, Aleksandra, and Paweł Strumiłło. "Eye-blink detection system for human–computer interaction." Universal Access in the Information Society 11.4 (2012): 409-419.
[14] Ban, Yuseok, et al. "Face detection based on skin color likelihood." Pattern Recognition 47.4 (2014): 1573-1585.
[15] Fröba, Bernhard, and Christian Külbeck. "Real-time face detection using edge-orientation matching." Audio-and Video-Based Biometric Person Authentication. Springer Berlin Heidelberg, 2001.
[16] Jesorsky, Oliver, Klaus J. Kirchberg, and Robert W. Frischholz. "Robust face detection using the hausdorff distance." Audio-and video-based biometric person authentication. Springer Berlin Heidelberg, 2001.
[17] Kim, JunYoung, et al. "Recognition of face orientation by divided hausdorff distance." Telecommunications and Signal Processing (TSP), 2015 38th International Conference on. IEEE, 2015.
[18] Mena, Alexandre Paz, Margarita Bachiller Mayoral, and Estela Díaz-Lópe. "Comparative Study of the Features Used by Algorithms Based on Viola and Jones Face Detection Algorithm." Bioinspired Computation in Artificial Systems. Springer International Publishing, 2015. 175-183.
[19] Smola, A., S.V. Vishwanathan, and Z. Yu, Sequential data segmentation. 2006, Google Patents.
[20] Weston, Jason. "Support Vector Machine." Tutorial, http://www. cs. columbia. edu/~ kathy/ cs4701/documents/jason_svm_tutorial. pdf, accessed 10 (2014).
[21] Rai, Preeti, and Pritee Khanna. "A gender classification system robust to occlusion using Gabor features based (2D) 2 PCA." Journal of Visual Communication and Image Representation 25.5 (2014): 1118-1129.
[22] Shbib, Reda, and Shikun Zhou. "Facial Expression Analysis using Active Shape Model." International Journal of Signal Processing, Image Processing and Pattern Recognition 8.1 (2015): 9-22.
[23] Wadkar, Pallavi D., and Megha Wankhade. "Face Recognition Using Discrete Wavelet Transforms." International Journal of Advanced Engineering Technology 3 (2012): 1-3.
[24] Huynh, Tri, Rui Min, and Jean-Luc Dugelay. "An efficient LBP-based descriptor for facial depth images applied to gender recognition using RGB-D face data." Computer Vision-ACCV 2012 Workshops. Springer Berlin Heidelberg, 2013.
[25] Bargal, Sarah Adel, et al. "Image-based ear biometric smartphone app for patient identification in field settings." Proc. International Conf. on Computer Vision Theory and Applications. 2015.
[26] M. Demirkus, M. Toews, J. J. Clark, and T. Arbel, "Gender classification from unconstrained video sequences," in Computer Vision and Pattern Recognition Workshops(CVPRW), IEEE Computer Society Conference on 2010, pp. 55–62, , 2010.

Khurram Zeeshan Haider completed his MS (Software Engineering) in 2011 from University of Engineering and Technology, Taxila, Pakistan. His research areas include Digital Image Processing, Machine Vision, Contextual Image Classification, video based Behavior Classification, Machine Learning, Software Development and Software Quality Engineering. He is PhD scholar at University of engineering and Technology, Taxila.

Tabassam Nawaz received his MS Computer Engineering in 2005 from CASE (Center for Advanced Studies in Engineering), Islamabad, Pakistan and subsequently he completed his Ph.D in 2008. He is the Chairman of Software Engineering Department at University of Engineering and Technology Taxila, Pakistan. He has published number of research papers in different Journals and Conferences. His research areas include Software Engineering, Programming, Data Structure, Computer Graphics, Computer Networks and Digital Image Processing.

Hafiz Adnan Habib completed his MS (Electrical Engineering) in 2004 and PhD (Electrical Engineering) in 2007 from University of Engineering and Technology, Taxila, Pakistan. He is author of number of research papers published in different journals and Conferences. His areas of interest include Video Tracking, Digital Image Processing, Computer Vision, Machine Learning and Artificial Intelligence. He is Chairman of Computer Science Department at University of Engineering and Technology, Taxila, Pakistan.

Muazzam Maqsood is currently doing his Ph.D. in Software Engineering from University of Engineering and Technology, Taxila. He has completed his MS degree in 2013 from University of Engineering and Technology, Taxila. His Major areas of interests are Machine Learning, Speech Processing, Recommender System and Image Processing.

Muhammad Tauqeer Ul Amin completed his MS (Software Engineering) from University of Gothenberg, Sweden in 2012. He is currently working as Assistant Professor in the Department of Software Engineering at University of Gujrat, Pakistan. His areas of interest include Empirical Software Engineering, Software Process Improvement, Requirement engineering, Software Project Management and Software Quality Engineering.