Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/100628

A Comparative study of a Face Components based Model of Ethnic Classification using Gabor Filters

Hajra Momin[1] *and Jules-Raymond Tapamo*[2,*]

[1] School of Mathematics, Statistics, and Computer Science, University of KwaZulu-Natal, Westville Campus Durban 4001, South Africa
[2] School of Engineering, University of KwaZulu-Natal, Howard College Campus Durban 4001, South Africa

Received: 2 Jan. 2016, Revised: 8 Jul. 2016, Accepted: 10 Jul. 2016
Published online: 1 Nov. 2016

Abstract: The determination of ethnicity of an individual can be very useful in face recognition and person identification. In this paper, we propose a model of ethnic classification of a person from face images. The proposed model detects facial landmarks such as eyes, nose and mouth in a face image and then applies Gabor filters to each component to extract key facial features. K-Means, Naive Bayesian(NB), Multilayer Perceptron(MLP), SVM(Support Vector Machines) are then used to classify the human face image into different ethnic groups. Classification is performed in 2-class (Asian and Non- Asian), 3-class (Asian, White and Black) or 4-class (Asian, Indian, White and Black). The results show that the mouth and the nose outperform the eyes in the characterization of ethnicity, and the classification is improved when all these components are used. It is also important to note that although few face components are used, the proposed model is comparable or even outperforms some existing models.

Keywords: Gabor Filters, Face Component, Ethnic Classification, Support Vector Machines, Multilayer Perceptron, Naive Bayesian, K-Means.

1 Introduction

The human face is a key biometric that provides demographic information, such as ethnicity, age, and gender of a person. Conversely, ethnicity and gender also play an important role in face-related applications. Humans can easily recognize ethnicity through facial appearance. Currently, the human face is commonly used to determine the ethnicity of a person. A successful database indexing algorithm can significantly boost the performance of face identification. This can greatly improve the response time of face based person identification. The face recognition response time could be greatly improved if we divide the database of face images according to races and then search the images in the appropriate race database instead of searching the person in the whole single database. In brief, this will improve the search speed and efficiency of some biometric retrieval systems.

When Humans look at each other they process the face in a variety of ways to categorize it. By looking at somebody's face humans can easily recognize which race they belong to, identify their gender, and estimate their age. People are more accurate at categorizing faces of their own race than faces of other races. Face processing is a challenging task, mostly because of the inherent variability of the image formation process in terms of image quality, photometry, geometry, occlusion, change, and disguise. In [1,26], these challenges are discussed in detail.

The most critical characteristics in determining ethnicity are thought to be the facial organs such as eyes, nose and mouth [10]. Component based face identification and recognition gained momentum in the past two decades. Human face components such as eyes nose and mouth are the most important features in the human face. In [24] the connected component labeling algorithm is used for detecting face in digital images. Huang *et al.*[12] proposed a Support Vector Machine based face recognition system which decomposes the face into a set of components that are interconnected by a flexible geometrical model. They used a 3D morphable model to generate 3D face models from only two input

* Corresponding author e-mail: tapamoj@ukzn.ac.za

images. Heisele *et al.*[9] presented a component-based face recognition method and compared it to two global face methods, in relation to the robustness against the pose changes. But to our knowledge no work on race classification using face components has been introduced in the literature.

In this paper, we present an ethnic classification based on Gabor filter using four different classifiers. The following combinations of races have been considered for classification;

1. Two race groups: Asian and Non Asian
2. Three race groups: Asian, White and Black
3. Four race groups:Asian, Indian, White and Black

The remainder of this paper is organized as follows. In section 2, background and related works done on ethnic classification are presented. Section 3 develops the proposed face feature computation model and discusses the classifiers used to explore the effectiveness of the proposed model. Section 4 presents experimental results and outlines the findings followed by a conclusion and feature works in section 5.

2 Related work

Automatic face-based ethnic classification has the potential to boost the performance of face based person identification. During the last decade, research on face-based ethnic classification has emerged, and grown rapidly. Ou *et al.*[22] classified a frontal face image into Asian and non Asian. They used principal component analysis (PCA) for feature generation and independent component analysis (ICA) for feature extraction. SVM was then combined to some new classifiers to improve the classification rate. Their system achieved only 82.5% accuracy using a database containing 750 images. Lu and Jain[15] used the Linear Discriminant Analysis(LDA) scheme for ethnic classification. They classified faces into Asian and Non Asian classes. An overall accuracy of 96.3% was reported with 132 Asian and 131 non-Asian faces. Manesh *et al.*[17] also classified faces into Asian and Non-Asian using the appearance based method to determine the confidence of different facial region using SVM. They reported a 0.0261% error rate with faces normalized using the eye and the mouth positions.

Hosoi *et al.*[10] make use of Gabor Wavelets transformation and retina sampling to extract key facial features and then used Support Vector Machines for ethnic classification. They reported an overall accuracy of 94 % for ethnic estimation. The classification involved Asian, European and African face images.

Zhang *et al.*[36] explored the ethnic discriminatory index of both 2D and 3D face features. They used Multi-scale Multi-ratio Local Binary Pattern method which is a multi-modal method for ethnic classification. They claimed an accuracy of 99% of their system. However, their method has a poor performance for face expression variation.

Lyle *et al.*[16] used periocular region images using gray scale pixel intensities and periocular texture computed by local binary patterns as features for gender and ethnic classification and SVM as the classifier. They achieved 91% accuracy for the ethnic classification. Zhuang et al.[35] used the fusion of multi-view gait for the ethnic determination. They used Gait Energy Image (GEI) to analyze the recognition power of gait for ethnicity. However, the highest classification rate they achieved was 84%. Guatta *et al.*[8] used hybrid classification architecture for ethnicity classification of human faces. The hybrid approach consists of an ensemble of Radial Basis Functions network and inductive decision trees. The average accuracy of their system was 94% for ethnic classification.

Heisle *et al.*[9] used 3D models of faces to find ethnicity. They classified images into Asian and Non Asian. A range of pixel intensity was used as features and they employed two SVMs, one for each modality to infer ethnicity. The final decision was made by integrating the two SVM results. They reported an accuracy of 98%.

Yang and Ai [34] used Local Binary Patterns and Haar wavelets as features and the classifier they used was Adaboost to classify face images into Asian and Non Asian; an accuracy of 97% was reported. Guo *et al.*[7] used MORPH II database to investigate how gender and age affect the ethnic classification. They used biologically inspired features based on Gabor filters and Support Vector Machines is used as classifier. The classification involved Asian, European, and African faces images.

In [6], two types of features were used, Linear Discriminant Analysis based algebraic features and elastic model based geometric features. They classified images into three minority Chinese groups. An accuracy of 79% was reported using algebraic features and 90.95% with geometric features with K-Nearest Neighbours (k-NN) and C5.0 classifiers. Shakhnarovich *et al.*[29] propose a real time face detection and classification into gender and ethnicity. They used three types of rectangular filters to extract features. SVM and a boosted classifier were used to classify images the into two classes: Asian and Non Asian. They reported a 22.6% error rate. In [20], Muhammad *et al.* investigated and compared the performance of local descriptors for race classification from face images. They used two types of local descriptors in their study: Local Binary Patterns and Weber Local Descriptors (WLD). The accuracies obtained using LBP were 98.42 % for Asian, 95.56% for Black , 93.65% for Hispanic, 100% for Middle East and 98.18% for White. The accuracies obtained using WLD were 97.74% for Asian, 96.89% for Black, 92.06% for Hispanic, 98.33% for Middle East, and 99.53% for White.

Most existing methods typically use the whole face and complex models for ethnic classification. In this article, a method that extracts features from face landmarks and uses different classifiers to demonstrate the performance of the proposed model.

3 Materials and Methods

The aim in this paper is to use Gabor filters to characterize the human face for ethnic classification. After the features have been extracted from the face, different classifiers are used to evaluate the effectiveness of Gabor filters when applied to the face components. Figure 1 shows the structure of the face image classification process. The following sections discuss the feature extraction method proposed and the classifiers used.

3.1 Features Extraction

3.1.1 Gabor Filters

Gabor filters have been successfully applied to many computer vision applications. They have specifically been used for texture segmentation due to their appealing simplicity and ability to localize the joint spatial frequency. Some of the applications in biometrics are face recognition [3] and iris recognition [32].

The 2D Gabor filter can be represented as a complex sinusoidal signal modulated by a Gaussian kernel function [19] defined as

$$g(x,y) = s(x,y)w_r(x,y) \tag{1}$$

where $s(x,y)$ is a complex sinusoidal, known as the carrier defined as

$$s(x,y) = e^{j(2\pi(\mu_0 x + v_0 y + P))} \tag{2}$$

with parameters (μ_0, v_0) and P defining the spatial frequency of the phase sinusoid, respectively. In polar coordinates the spatial frequency are expressed as magnitude F_0 and direction θ_0 as

$$F_0 = \sqrt{\mu_0^2 + v_0^2} \tag{3}$$

$$\theta_0 = \tan^{-1}\begin{pmatrix}\mu_0 \\ v_0\end{pmatrix} \tag{4}$$

We then have the complex sinusoid rewritten as

$$s(x,y) = e^{j(2\pi F_0(x\cos\theta_0 + y\sin\theta_0) + P))} \tag{5}$$

The function $w_r(x,y)$ is the envelope, which is defined as

$$w_r(x,y) = Ke^{-\pi(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2)} \tag{6}$$

where (x_0,y_0) is the peak function, a, b are scaling parameters of the Gaussian, and the subscript r stands for a rotation operation.

$$(x-x_o)_r = (x-x_0)\cos\theta + (y-y_0)\sin\theta \tag{7}$$

and

$$(y-y_o)_r = -(x-x_0)\sin\theta + (y-y_0)\cos\theta \tag{8}$$

Considering $x_0 = y_0 = 0$ and $a = b = \sigma$, 2-D Fourier transform of the Gabor function is

$$g(x,y) = Ke^{(-\pi(a^2(x-x_0)_r^2 + (y-y)_r^2))}e^{j(2\pi F_0(x\cos\theta_0 + y\sin\theta_0) + P)} \tag{9}$$

where K , (a,b), θ and (x_0,y_0) are the scale of the magnitude, scale of the two axes, rotation angle and location of the peak of the Gaussian envelope respectively.

In our experiments we used 5 different frequencies and 8 orientations to extract texture features from the face components. The filter's parameters were the following:

– $K = 1$, $P = 0$, $\sigma = 1$,
– $F_0 \in \{0.13, 0.22, 0.31, 0.40, 0.49\}$
– and $\theta \in \{0, \pi/8, 2\pi/8, 3\pi/8, 4\pi/8, 5\pi/8, 6\pi/8, 7\pi/8\}$.

The Gabor filter bank composed of 40 channels was then created. The filters set obtained were then applied to the input facial images, by convolving the face image with each Gabor filter from this set. The resulting Gabor responses were then concatenated into a feature vector. The process is describe in the following section.

3.1.2 Face features computation

This Section describes how features were extracted from the face. The face feature extraction is initialised by the detection of facial landmark components(eye, mouth, nose) using the algorithm described in [18]. In our experiment, Gabor filter with 5 different frequencies and 8 orientations were used. For each component, the mean value for 5 frequencies and 8 orientations were computed and stored in feature vectors, $(\mu_{1\_Le}\dots\mu_{40\_Le})$ for the Left eye, $(\mu_{1\_Re}\dots\mu_{40\_Re})$ for the Right eye, $(\mu_{1\_No}\dots\mu_{40\_No})$ for the Nose, and $(\mu_{1\_Mo}\dots\mu_{40\_Mo})$ for the mouth. Similarly, for each component, standard deviations were computed for 5 different frequencies and 8 orientations and stored in vectors, $(\sigma_{1\_Le}\dots\sigma_{40\_Le})$ for the Left eye, $(\sigma_{1\_Re}\dots\sigma_{40\_Re})$ for the Right eye, $(\sigma_{1\_No}\dots\sigma_{40\_No})$ for the Nose, and $(\sigma_{1\_Mo}\dots\sigma_{40\_Mo})$ for the Mouth. In total there are 160 means, because there are 4 landmarks and each has 40 means. For each face, the means' vector is then represented as

$$\begin{aligned}\mu_{face} = (&\mu_{1\_Le}\dots\mu_{40\_Le}, \mu_{1\_Re}\dots\mu_{40\_Re}, \\ &\mu_{1\_No}\dots\mu_{40\_No}, \mu_{1\_Mo}\dots\mu_{40\_Mo})\end{aligned} \tag{10}$$

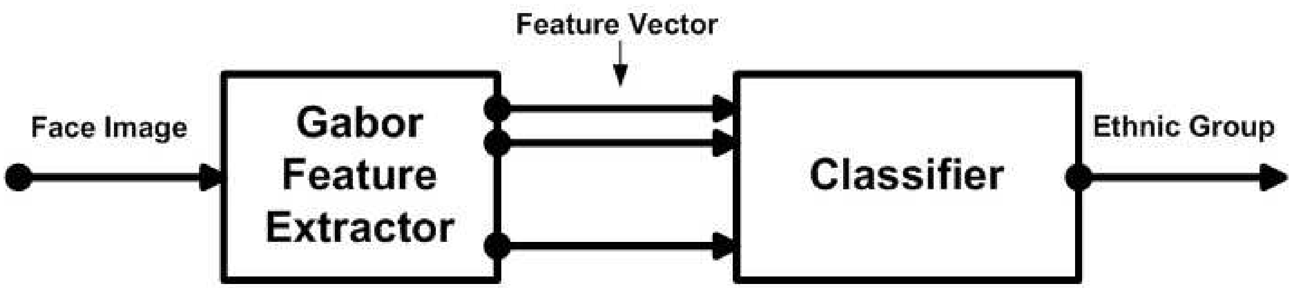


Fig. 1: Face based Ethnic classification flow diagram.

Similarly, the standard deviations' vector has 160 components and is represented as

$$\Sigma_{face} = (\sigma_{1\_Le} \dots \sigma_{40\_Le}, \sigma_{1\_Re} \dots \sigma_{40\_Re}, \sigma_{1\_No} \dots \sigma_{40\_No}, \sigma_{1\_Mo} \dots \sigma_{40\_Mo}) \tag{11}$$

A face was then represented by a feature vector of 320 components as

$$V_{face} = (\mu_{face}, \Sigma_{face}) \tag{12}$$

It is also important to note that, each face component was characterized by a vector of 80 features (40 means and 40 standard deviations) as

$$V_{LeftEye} = (\mu_{1\_Le} \dots \mu_{40\_Le}, \sigma_{1\_Le} \dots \sigma_{40\_Le}) \tag{13}$$

$$V_{RightEye} = (\mu_{1\_Re} \dots \mu_{40\_Re}, \sigma_{1\_Re} \dots \sigma_{40\_Re}) \tag{14}$$

$$V_{Nose} = (\mu_{1\_No} \dots \mu_{40\_No}, \sigma_{1\_No} \dots \sigma_{40\_No}) \tag{15}$$

$$V_{Mouth} = (\mu_{1\_Mo} \dots \mu_{40\_Mo}, \sigma_{1\_Mo} \dots \sigma_{40\_Mo}) \tag{16}$$

Figure 2 shows the framework of Gabor Feature extraction process.

3.2 Classification

Four classifiers, namely K-means, Naive Bayesian, Multi-layer Perceptron and Support Vector Machines were used to classify features extracted from the faces.

3.2.1 K-means Clustering

K-means [31] is a non-hierarchical clustering technique that follows a simple procedure to classify a given data set in a certain number of (k) clusters; only k must be known a priori. The K-Means clustering algorithm classifies input data points into multiple classes based on their inherent distance from each other (typically the Euclidian distance between the cluster centers and the candidate vector).

The aim of the k-means algorithm is to minimize the objective function that samples the closeness between the data points and the cluster centers, and is calculated as follows:
given $\{x_1, \dots, x_N\} \subset R^d$ (where d is the number of components of a feature vector) and $k \in \{1, 2, \dots, N\}$, the objective is to minimize the following function:

$$W(\omega_1, \omega_2, \dots, \omega_k; \mu_1, \dots, \mu_k) = \sum_{i=1}^{k} \sum_{x_j \in \omega_i} \|x_j - \mu_i\|^2 \tag{17}$$

where the clustering $\{\omega_1, \omega_2, \dots, \omega_k\}$ is a partition of $\{x_1, \dots, x_N\}$ and $\mu_1 \dots, \mu_k$ are the representatives of the clusters $\omega_1, \dots, \omega_k$.

The k-means clustering algorithm has been used by Chitade [2] and Samma [27] for image segmentation. Gutta and Wechsher[8] used k-means clustering algorithm for the ethnic classification of face images . In our experiment we used k-means algorithm with $k = 2$ for classifying the face images into two groups as Asian and Non Asian; with $k = 3$ for Asian, White and Black, and $k = 4$ for classifying images into 4 classes as Asian, Indian, White and Black.

3.2.2 Naive Bayesian

Bayesian classification provides practical learning algorithms and prior knowledge and observed data can be combined. It calculates explicit probabilities for hypothesis and it is robust to noisy input data. The Problem of classifying face images can be solved by using the Bayesian classifier. If C_i represents the i^{th} ethnic class; we can then have n ethnic classes $C_1, C_2 \dots, C_n$. Supposing that there are 3 ethic classes as Asian, White and Black, then $n = 3$. Each face image is represented by a feature vector $X = [X_1, \dots, X_j]$. If the class of a face with a feature X has been found, then the probability that an image belongs to that particular class C_i is given by the posterior probability $P(C_i|X)$ of that class C_i given the

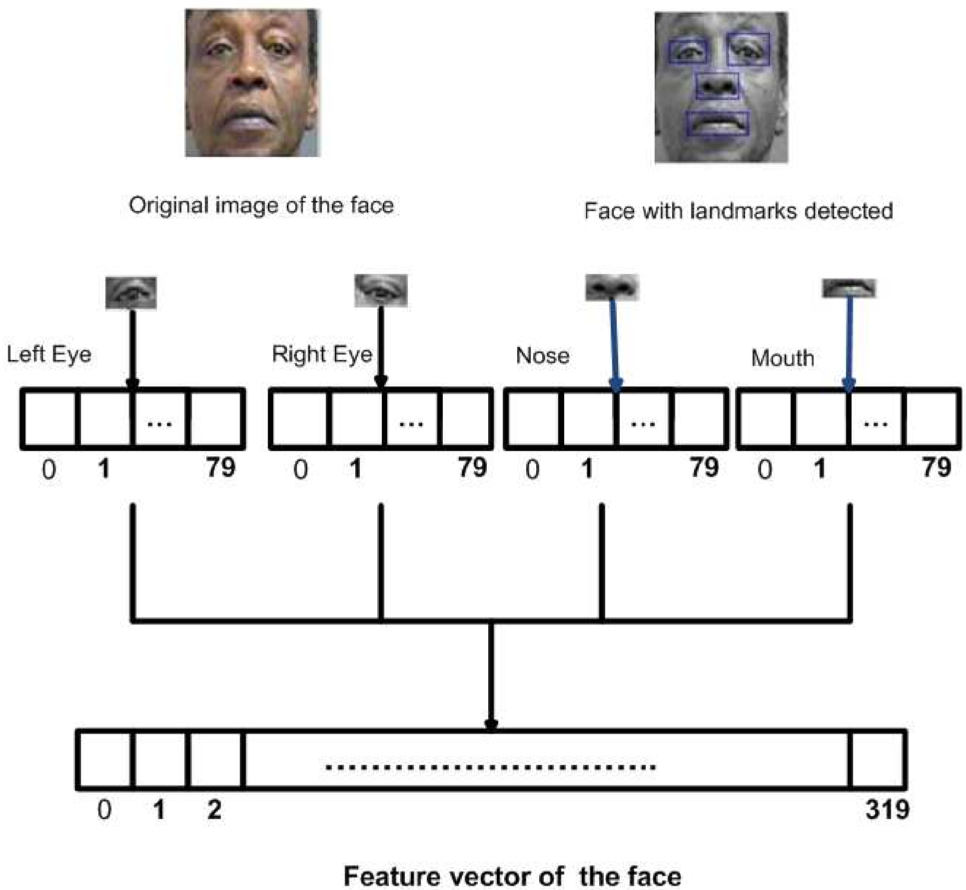


Fig. 2: Framework of Gabor Feature extraction process

feature vector X. One can write the Bayes theorem in the form [14,23]:

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \tag{18}$$

where

–$P(C_i|X)$ is the posterior probability of X,
–$P(X|C_i)$ is likelihood probability density function (pdf) for feature vector X given the image belongs to class C_i,
–and $P(C_i)$ is prior probability of the class C_i

And $P(X)$ is given by:

$$P(X) = \sum_{i=1}^{n} P(C_i)P(X|C_i) \tag{19}$$

which ensures that the sum of the predicted occurrence probabilities:

$$\sum_{i=1}^{n} P(C_i|X) = 1 \tag{20}$$

The prior probability $P(C_i)$ is simply taken as

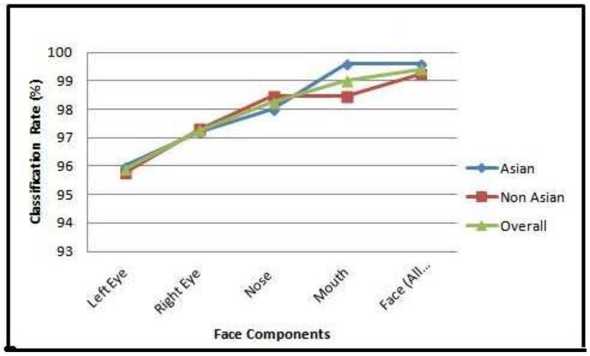
$$R_i / \sum_{j=1}^{N} R_j \tag{21}$$

Where R_i is the number of images of ethnic type C_i . To implement Equations 18 and 19, we need pdfs $p(X|C_i)$ for all race types C_i. One might get pdfs by parametric forms as for example Gaussian, lognormal or Gamma. In our method, however, we assumed that each class C_i has a Gaussian distribution, with mean μ_{C_i} and standard deviation σ_{C_i}, defined as

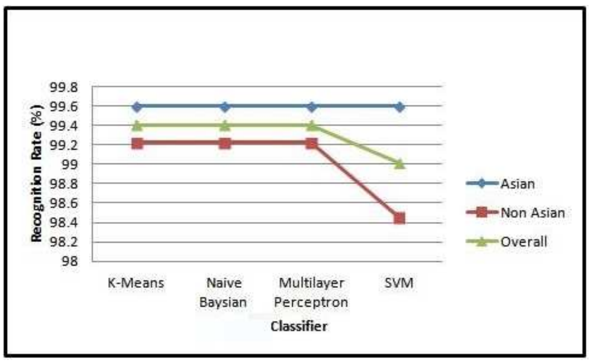
$$p(X|C_i) = \prod_{j=1}^{n} P(x_j|C_i)p(C_i) \tag{22}$$

Where $X = (x_1, \ldots, x_p)$ and

$$p(x_j|C_i) = \frac{1}{\sqrt{2\pi\sigma_{C_i}}} e^{-\frac{(x_j-\mu_{C_i})^2}{2\sigma_{C_i}}} \tag{23}$$

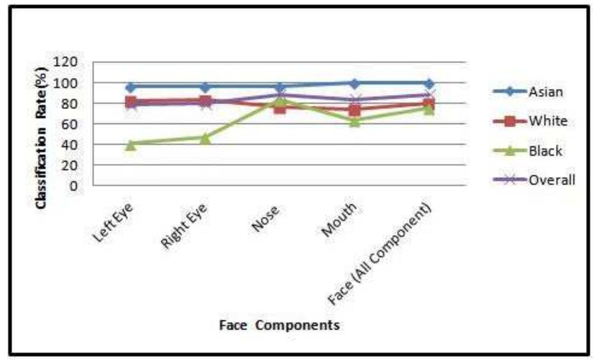


(a) Face components vs Classification rate in %

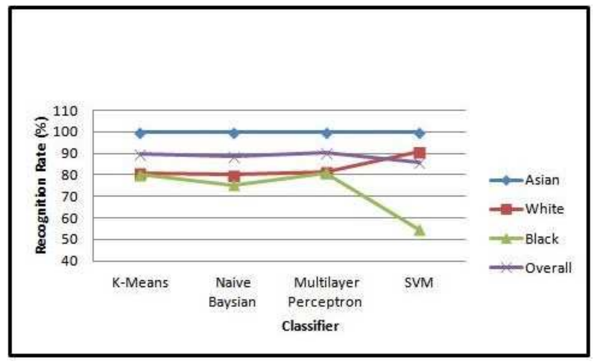


(b) Classifiers vs Classification rate in %

Fig. 3: 2-class (Asian and Non-Asian) Classification using 4 Classifiers and 4 face components.



(a) Face components vs Classification rate in %



(b) Classifiers vs Classification rate in %

Fig. 4: 3-class (Asian and Non-Asian) Classification using 4 Classifiers and 4 face components.

After getting the pdfs for all race types, Now one can use Equation 18 to calculate occurrence probability of any race type at known indicator variables. If the posterior occurrence probability is highest for a vector X into a race type C_i then vector X belongs to C_i race.

3.2.3 Multilayer Perceptron(MLP)

A multilayer perceptron(MLP) has already been used in face recognition. For instance, Tamura et al. [30] used multilayered neural network to classify gender from face images. The face images used were with multiple resolutions (from 32-by-32 to 16-by-16 and 8-by-8 pixels). The experiment conducted on 30 test images showed that their network is able to determine gender from face images of 8-by-8 pixels.

MLP is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate output. An MLP consists of multiple layers of simple, two state, sigmoid processing elements(nodes) or neurons that interact using weighted connection. An MLP network has an input layer, and any number of intermediate layers or hidden layers and an output layer. All neurons in each layer are fully connected to neurons in adjacent layers while there is no interconnection within a layer. In an MLP, all nodes and layers are arranged in a feed-forward manner. Weights measure the degree of correlation between the activity levels of neurons that they connect [28].

Let w_{ij} be the connection between a node i in the input layer and a node j in the hidden layer. The output y_{ku} of the j^{th} node in the hidden layer corresponding to the input vector X_u is given by:

$$y_{ju} = g(\sum_{i=1}^{d} w_{ij} x_{ui}); \quad j = 1, \ldots, h \tag{24}$$

where x_{ui} is the i^{th} component of the input vector X_u and $g(x)$ is the sigmoid function defined as

$$g(x) = \frac{1}{1+e^{-x}}. \tag{25}$$

Similarly , the k^{th} output y_{ui} of the output layer is

$$y_{ku} = g(\sum_{j=1}^{h} w_{jk} x_{ju}); \quad k = 1, \ldots, m \tag{26}$$

The MLP used will have a single hidden layer, and we will adopt the technique that assumes that if the input layer has *Nbrinputnodes* nodes and there are *Nbrclasses* classes then the number of nodes in the hidden layer *NbrNodesHidden* is defined as

$$NbrNodesHidden = \frac{Nbrinputnodes + NbrClasses}{2} \quad (27)$$

Three structures of the MLP depending of the number of classes considered:

(a)The first MLP is modeled to classify items from data structured in two groups : Asian and Non Asian. complete feature vector will have input layer with 320 attributes (nodes), hidden layer with 161 nodes, and output layer with 2 nodes (Asian and Non Asian). We have considered analysing the contribution of some face components in ethnicity. Therefore as each component feature vector has 80 components, the MLP in this case has an input layer with 80 nodes, hidden layer with 41 nodes and output layer with 2 nodes.

(b)When the population structures into three groups(Asian, White and Black) the input and hidden layers are the same as in (a) and the output layer will contain 3 nodes

(c)For 4-class classification the MLP has a hidden layers with 162 nodes for 320 nodes in input layer, and 42 hidden nodes for 80 nodes in input layer, and the output layer has 4 nodes (Asian, Indian, White and Black).

3.2.4 Support Vector Machine (SVM)

Support Vector Machines (SVM) is a relatively new machine learning method originally used for binary classification[4,33]. It has since also been used for multiclass classification [5,11]. It is one of the most used nonparametric supervised classifier available today. It provides a novel means of classification using the principle of structural risk minimization. The decision surface is a weighted combination of elements of the training set. These elements are called support vectors and characterize the boundary between the two classes. The goal of SVM is to produce a model which predicts the class to which a data instance belongs given its attributes.

Classification is done using training and testing data which consist of some data instances. The input to the SVM algorithm is a set (x_i, y_i) of labeled training data, where x_i is the data and y_i = -1 or 1 is the label.

$$x_i \in R^n, y_i \in \{-1, 1\} (i = 1, \ldots, l) \quad (28)$$

When we assume a definite and separable hyperplane exists, we can express this hyperplane as

$$(w.x) + b = 0 \quad (29)$$

By applying to the Quadratic Programming problem solving, we can express the decision function as

$$f(x) = \sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b \quad (30)$$

Here, $K(x, x_i)$ is a kernel function, and it is used to determine to which class the input data x belongs through $f(x)$.

3.3 General algorithm of the proposed method

Given a face image F. The process to classify it in a particular ethnicity is summarised in *Algorithm* 1.

Algorithm 1 Automatic Ethnic classification from face image

Require: F ▷ Given face image
Ensure: *Decision* ▷ classification of the face in one of the ethnic groups considered
1: (L, R, N, M) = **LME**(F) ▷ This function (LME) extracts the landmarks (LeftEye[L], RightEye[R], Nose[N],Mouth[M]) from the face (F), and saves them as 4 images
2: **Extract** face features from (L,R,N,M) and **Save** them in V_{face} using equations 9 and 12-16
3: $Decision = \textbf{Classify}(V_{face})$ ▷ Classification of a face (F) represented by the vector V_{face}

Given a face image of size $m \times n$. The running time of the extraction of landmarks is $O(mm)$, as all pixels are visited in the process; feature extraction consumes $O(pr)$, where $p \times r$ is the size of the biggest landmark; and the running time to classify each face after feature extraction is done in constant time; in summary, the running time of face classification is $O(mn) + O(pr) + O(1) = O(nm)$, as $m \gg p$ and $n \gg r$.

4 Experimental Results and discussion

4.1 Data Set

The Data set used for experimentation is a fusion of 3 different face data sets. We have used 511 images with different facial expression, illumination and orientation. The Asian group is composed of images from Asian face database [21] and Indian face database [13]. Most of the Asian faces are of Korean, Indian and Chinese origin. The Indian face database contains Indian faces with different illumination, expression and orientation. 52 images from the Indian face database and 199 images from the Asian face database were considered for experimentation. The Non Asian group is composed of images from MORPH database [25]. The MORPH database contains White and Black face images with different ages, illumination and expression. We have taken 260 images from this database. The $\frac{2}{3}rd$ of the images are considered for training and $\frac{1}{3}rd$ of images were used for testing.

4.2 Performance Evaluation

To evaluate our experiments, the performance metric used is the *accuracy* of each classification, which is the percentage of items correctly classified and defined as

$$Accuracy = \frac{Number\ of\ correctly\ classified\ Images}{Total\ Number\ of\ images} \quad (31)$$

4.3 Results and discussions

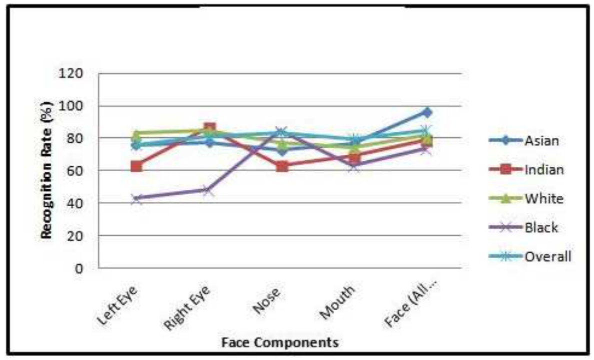
In our experiments we considered different classification algorithms, such as K-means clustering algorithm, naive bayesian, Multilayer perceptron and Support Vector Machines. Once the feature vectors are created using using Gabor filters, they are fed into each of the chosen classifiers. First images are classified into two ethnic groups as Asian and Non Asian. Table 1 gives ethnic classification results for each of the separate face component and for each classifier. The nose and the mouth give better results compares to the left eye and the right eye.

In the case of a 2-class classification involving Asian and Non Asian, it can be observed from Table 1 and Figure 3 that eyes are poor in characterizing ethnicity as the success rate is consistently below 95% for Asian and Non-Asian classification. On the contrary, the mouth and the nose seem to be indicators of ethnicity. Asian has consistently been the most identifiable ethnic group irrespective of the face component or the classifier used.

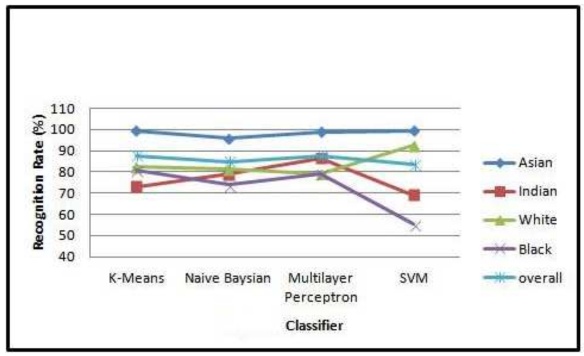
The results in Table 2 and Figure 4 show that when using a 3-class classifier where White, Asian, and Black are involved, there are many false negatives for Black identification. Asian remains the ethnic group which is consistently well classified, while White achieves a better accuracy than Black, but still outperformed by the Asian classification.

Based on Table 3 and Figure 5, the 4-class classification involving 4-ethnic groups (Asian, Black, Indian and White), Asian is still the most accurately classified ethnic group with Black achieving the lowest success rate. When used individually, all the face components are poor in characterizing ethnicity.

In General, the Nose and the Mouth are better indicators of the ethnic group compared to the eyes. This prompts us to make the hypothesis that the mouth and the nose could be further investigated for the characterization of face in ethnic classification. MLP gives better results compared to other machine learning algorithms used. Table 4 summarizes performances of some of the researches done in ethnic classification.



(a) Face components vs Classification rate in %



(b) Classifiers vs Classification rate in %

Fig. 5: 4-class (Asian and Non-Asian) Classification using 4 Classifiers and 4 face components.

5 Conclusion and Future works

We have presented a method for automatic ethnic classification based on texture analysis of face components. The face images are taken in complex contexts. The complexity of the images is presented by different illumination condition, facial expression, and different orientation. The proposed method was tested on images from 3 different face databases that were the Indian face database [13], MORPH database [25] and Asian face database [21]. We carried out experiments using 511 images with different complexities. Three cases were considered: (1) two different ethnic groups (Asian and Non Asian),(2) 3 ethnic groups (Asian, White and Black), and (3) 4 ethnic groups (Asian, Indian, White and Black). Different scenarios of use of the face components (left eye, Right Eye, Nose and Mouth) for ethnic classification were investigated. To extract features, banks

Table 1: Ethnic Classification for Asian and Non Asian

	Classification Rate (%)							
Face Component	k-means		Naive Bayesian		MLP		SVM	
	Asian	Non Asian	Asian	Non Asian	Asian	Non Asian	Asian	NonAsian
Left Eye	98.8	95.76	96.01	95.76	96.81	95.53	98.01	96.53
Right Eye	100	97.69	97.21	97.3	99.2	97.69	99.6	97.69
Nose	99.6	97.69	98	98.46	98.46	97.69	98.4	98.07
Mouth	99.6	97.69	99.6	98.46	99.2	98.46	99.6	98.46
All Components	**99.6**	99.23	99.6	99.23	**99.6**	99.23	99.6	98.84

Table 2: Ethnic Classification for Asian, White and Black

	Classification Rate (%)											
Face Component	k-means			Naive Bayesian			MLP			SVM		
	Asian	White	Black	Asian	White	Black	Asian	White	Black	Asian	White	Black
Left Eye	98	80.76	46.92	96.42	82.3	40.76	97.21	83.07	48.46	98.80	86.15	40.76
Right Eye	100	80	55.38	98.82	83.07	46.92	99.2	86.92	50.76	98.80	90	43.07
Nose	99.6	83.84	83.07	96.82	76.92	83.07	98.4	85.38	76.92	98.4	86.15	83.07
Mouth	99.6	76.92	63.84	99.6	73.84	63.84	99.2	88.46	60	99.6	82.30	66.15
All Component	**99.6**	80.76	80	**99.6**	80	75.38	**99.6**	81.53	80.76	**99.6**	86.92	76.15

Table 3: Ethnic Classification for Asian, Indian, White and Black

	Classification Rate (%)															
Face Component	k-means				Naive Bayesian				MLP				SVM			
	Asian	Indian	White	Black	Asian	Indian	White	Black	Asian	Indian	White	Black	Asian	Indian	White	Black
Left Eye	100	48.07	80	44.61	95.97	63.46	83.07	43.07	98.99	63.46	89.23	36.92	99.49	57.69	84.61	40
Right Eye	99.49	80.76	80	56.15	97.98	86.53	84.61	48.46	98.49	92.3	90.76	38.46	99.49	82.69	88.46	46.15
Nose	99.49	42.3	85.38	83.84	91.95	63.46	77.69	84.61	96.48	50	86.15	86.41	97.48	40.38	86.15	83.07
Mouth	98.99	50	75.38	64.61	96.98	69.23	74.61	63.07	97.48	65.38	89.23	58.46	98.99	65.38	83.84	65.38
All Components	**99.49**	73.07	82.3	80.76	**95.97**	78.84	81.53	73.84	**98.99**	86.53	79.23	79.23	**99.49**	78.84	88.46	76.15

of multichannel 2D Gabor filters were used to capture the texture information of the face components. K-means clustering algorithm, Naive Bayesian, Multilayer perceptron and SVM were used to classify images into ethnic classes.

Extensive experiments demonstrate that our method has achieved a high degree of accuracy: for Asian and Non-Asian:99.60%; When considering the Asian, White and Black grouping the accuracy dropped to 90.21%, and the 4-class classification with Asian, Indian, White and Black only achieved 87.67%. From our experiment we observed that mostly the nose and the mouth give better results compared to the eyes. In future works we would like to extend this work by adding more features like color and/or geometric features to further improve the accuracy of ethnic classification using Gabor filters. It will also be interesting to probe a bit more the use of the mouth and the nose for ethnic classification.

References

[1] R. Chellappa, C. L. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83:705–740, 1983.

[2] A. Z. Chitade and S. Katiyar. Color based image segmentation using k-means clustering. *International Journal of Engineering science and technology*, 2(10):5319–5325, 2010.

Table 4: Comparative performance of some ethnic classification systems

Reference	Features	Classifier	Race Groups	Database	Recognition
[22]	PCA	SVM	Asian, Non Asian	FERET	82.5%
[15]	range and pixel intensity	SVM	Asian, Non Asian,	376 subjects, 1240 scans	98%
[17]	Gabor	SVM	Asian, Non Asian	CAS-PEAL, FERET	96%
[10]	Gabor wavelet transform with retina sampling	SVM	Asian, European, African	HOIP dataset	Asian-96% European - 93% African-94%
[36]	MM-LBP	AdaBoost	Asian, White	FRGC v2.0	99.5%
[16]	LBP	SVM	Asian , Non Asian	FRGC Face dataset	91%
[35]	MPCA	Weighted sum rule and majority vote rule	Asian, American	Author's Database	84%
[8]	Grayscale Pixel intensities	Neural networks using SVM and DT's	Caucasian, South Asian, East Asian , African	FERET	92%
[34]	LBP,Haar like features	Adaboost	Asian, Non Asian	FERET,PIE	97%
[6]	LDA-PCA, Geometric feature	KNN, C5.0	Tibetan, Uighur, Zhuang	Author's Database	79% with algebriac features and 90.95% with Geometric features
[29]	Rectangular Features	Boosted Classifier and SVM	Asian, Non Asian	Faces from WWW	22.6% error rate
Proposed Method	Gabor Features	K-Means, NB, MLP and SVM	Asian,Indian, American, African	MORPH II database, Indian Face database, Asian Face database	Asian, Non Asian 99.60% Asian, White, Black 90.21% Asian, Indian, White, Black 87.67%

[3] K. Chung, S. Kee, and S. Kim. Face recognition using independent component analysis of gabor filter responses. In *in IAPR Workshop Machine Vision Applications*, pages 331–334, 2000.

[4] C. Cortes and V. Vapnik. Support vector machines. *Machine Learning*, 20:273–297, 1995.

[5] K Crammer and Y. Singer. On the algorithmic implementation of multiclass kernel-based vector machines. *Journal of Machine Learning Research*, 2:265–292, 2001.

[6] X. D Duan, C. R. Wang, Z. J. Li, J. Wu, and H. L. Zhang. Ethnic feature extraction and recognition of human faces. In *in Proceedings of the 2nd International Conference on Advanced Computer Control (ICACC)*, pages 125–130, 2010.

[7] G. Guo and G. Mu. A study of large scale ethnicity estimation with gender and age variations. In *in proceedings of the IEEE Computer Society Conference on computer Vision and pattern recognition Workshops (CVPRW)*, pages 79–86, 2010.

[8] S. Gutta and H. Wechsler. Gender and ethnic classification of face images. In *in Proceeding of the third IEEE international conference on Automatic face and Gesture Recognition*, pages 194–199, 1998.

[9] B. Heisele, P. Ho, J. Wu, and T. Poggio. Face recognition: component-based versus global approaches. *Computer Vision and Image Understanding*, 91:6–23, 2003.

[10] S. Hosoi, E. Takikawa, and M. Kawad. Ethnicity estimation with facial images. In *in Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, pages 195–200, 2004.

[11] C. W. Hsu and C.J. Lin. Comparison of methods for muticlass support vector machines. *IEEE Transactions on Neural Networks*, 13(2):415–425, 2002.

[12] J. Huang, V. Blanz, and B. Heisele. Face recognition using component-based svm classification and morphable models. *SVM 2002*, LNCS 2388:334–341, 2002.

[13] V. Jain and A. Mukherjee. The indian face database. *http://vis-www.cs.umass.edu/*, 2002.

[14] K. Jayech and M. A. Mahjoub. Clustering and bayesian network for image of faces classification. *International Journal of Advanced Computer Science and Applications*, pages 35–44, 2011.

[15] X. Lu and A. K. Jain. Ethnicity identification from face images. In *In proceeding of SPIE International Symposium on Defense and security: Biometric Technology for human identification*, pages 114–123, 2004.
[16] J.R. Lyle, P. E. Miller, S. J. Pundlik, and D. L. Woodard. Soft biometric classification using periocular region features. In *in Proceeding of the fourth IEEE International conference on Biometrics:Theory Applications and Systems(BTAS)*, pages 1–7, 2010.
[17] F. S. Manesh, M. Ghahramani, and Y. P. Tan. Facial part displacement effect on template based gender and ethnicity classification. In *in Proc of the 11th international conference on control automation Robotocs and vision (ICARCV)*, pages 1644–1649, 2010.
[18] H. Momin and J. R. Tapamo. Automatic detection of face and facial landmarks for face recognition. *Communications in Computer and Information Science Volume*, 260:244–253, 2011.
[19] J.R. Movellan. Tutorial on gabor filters. Technical report, UCSD MAPLab, 2005.
[20] G. Muhammad, M. Hussain, F. Alenezy, G. Bebis, A.N Mirza, and H. Aboalsamh. Race classification from face images using local descriptors. *International Journal on Artificial Intelligence Tools*, 21(5):1–24, 2012.
[21] Intelligent multimedia Lab. Asian face image database pf01. Technical report, Pohang University of Science and Technology, San 31, Hyoja-Dong, Nam-Gu, Pohang, 790-784, Korea.
[22] Y. Ou, X. Wu, H. Qian, and Y. Xu. A real time race classification system. In *in Proceedings of the 2005 IEEE International Conference on Information Acquisition*, pages 378–383, 2005.
[23] H. Rahimizadeh, M.H. Marhaban, R.M. Kamil, and N.B. Ismail. Color image segmentation based on bayesian theorem and kernel density estimaton. *European Journal of Scientific Research*, 26(3), 2009.
[24] A. Rakhmadi, M. S. M. Rahim, A. Bade, H. Haron, and I. M. Amin. Loop back connected component labelling algorithm and its implementation in detecting face. *World Academy of Science, Engineering and Technology*, 64, 2010.
[25] K. Ricanek Jr and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. In *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition(FGR'06)*, pages 341–345, 2006.
[26] A. Samal and P. Lyengar. Automatic recognition and analysis of human faces and facial expressions: A survey. *Pattern Recognition*, 25:65–77, 1992.
[27] A.S. Samma and R.A. Salam. Adaptation of k-means algorithm for image segmentation. *World Academy of Science, Engineering and Technology*, pages 58–62, 2009.
[28] K. P. Sankar and S. Mitra. Multilayer perceptron, fuzzy sets, and classification. *IEEE Transactions on Neural Networks*, 3(5):683–697, 1992.
[29] G. Shakhnarovich, P. A. Viola, and B. Moghaddam. A unified learning framework for real time face detection and classification. In *in Proceedings of the fifth IEEE International conference on automatic face and gesture Recognition*, pages 14–21, 2002.
[30] S. Tamura, H. Kawai, and H. Mitsumoto. Male/female identification from 8x6 very low resolution face images by neural network. *Pattern Recognition*, 29(2):331–335, 1996.
[31] M. W. Trosset. Representing clusters: K-means clustering, self-organizing, maps, and multidimensional scaling. Technical Report 08-03, Department of Statistics, Indiana University, Bloomington, 2008.
[32] C. C. Tsai, J. S. Taur, and C. W. Tao. Iris recognition using gabor filter optimized by the practicle swarm technique. In *in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC 2008)*, pages 921–926, 2008.
[33] V. Vapnik. *THe Nature of Statistical Learning*. Springer-Verlag, 1995.
[34] Zhiguang Yang and Haizhou Ai. Demographic classification with local binary patterns. In *Proceedings of the 2007 international conference on Advances in Biometrics*, ICB'07, pages 464–473. Springer-Verlag, 2007.
[35] D. Zhang, Y. Wang, and B. Bhanu. Ethnicity classification based on gait using multi-view fusion. In *in Proceedings of the IEEE conference on computer Vision and pattern Recognition*, pages 108–115, 2010.
[36] G. Zhang and Y. Wang. Multimodal 2d and 3d facial ethnicity classification. In *in Fifth International Conference on Image and Graphics(ICIG'09)*, pages 928–932, 2009.

Hajra Momin is PhD candidate in the School of Mathematics, Statistics and Computer Science at the University of KwaZulu-Natal. Her research interests are in the areas of Image Processing and Computer Vision. She is also a part-time Lecturer in Computer Programming and several other computing related courses.

Jules R Tapamo is Professor of Computer Science and Engineering at the School of Engineering , University of KwaZulu-Natal, South Africa. His research interests are Surface Characterization, Image Processing, Computer Vision, Machine Learning, Biometrics, Intelligent Monitoring, Activity recognition and Formal Methods. He is a member of the IEEE Computer Society, IEEE Signal Processing Society, IEEE Geoscience and Remote Sensing Society and the ACM.