Age Invariant Face Recognition using Feature Level Fused Morphometry of Lip-Nose and Periocular Region Features (February 2016)

Amal S. O. Ali, Sarat C. Dass, Aamir S. Malik, Azrina Aziz & Vijanth S. Asirvadam

ABSTRACT

The Aging variation1 poses an interesting challenge for the task of automatic face recognition. Most face recognition studies that have addressed this problem focused on age estimation or aging simulation. Designing an appropriate feature representation and an effective matching framework for age invariant face recognition remains an open problem.

In this research study, a novel age-invariant face recognition framework that is built based on two face biometric traits is proposed. The first trait is a set of anthropometric measurements acquired from the lip-nose complex. Lip-nose complex measurements have been known in several physiological studies to be discriminative among different ethnicities and among different genders within the same ethnicity.

Keywords: anthropometric measurements, craniofacial growth, FG-NET, lip-nose complex, MORPH.

Classification: DDC Code: 621.317 LCC Code: TK3511

Language: English

London Journals Press

LJP Copyright ID: 975834
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 23 | Issue 1 | Compilation 1.0

Age Invariant Face Recognition using Feature Level Fused Morphometry of Lip-Nose and Periocular Region Features (February 2016)

Amal S. O. Ali[α], Sarat C. Dass[σ], Aamir S. Malik[ρ], Azrina Aziz[CO] & Vijanth S. Asirvadam[¥]

ABSTRACT

The Aging variation[1] poses an interesting challenge for the task of automatic face recognition. Most face recognition studies that have addressed this problem focused on age estimation or aging simulation. Designing an appropriate feature representation and an effective matching framework for age invariant face recognition remains an open problem.

In this research study, a novel age-invariant face recognition framework that is built based on two face biometric traits is proposed. The first trait is a set of anthropometric measurements acquired from the lip-nose complex.

Lip-nose complex measurements have been known in several physiological studies to be discriminative among different ethnicities and among different genders within the same ethnicity.

The second trait is based on extracting features of the periocular region using two robust descriptors, namely local binary patterns rotation invariant descriptor (LBPV) and GIST descriptors. The periocular area is considered as the most discriminative face area and is known to preserve its stability with aging. The two biometric face traits were combined at the feature level after being normalized separately using the Z-score rule and projected into a principle component analysis (PCA) subspace.

Eight algorithms were derived as part of the proposed framework for performing age invariant face verification and identification.

Furthermore, the proposed framework was used to produce demographic information, namely: age group, gender, and ethnicity from aging faces.

Experimental results show that the proposed framework reported an EER of 6.51% over the MORPH album 2 database (compared to 16.49% reported by Mahalingam et al), which is the largest public face aging database and 7.22% over the FG-NET database (compared to 24.08% reported by Mahalingam et al).

The proposed framework achieved an identification accuracy of more than 95% (compared to 66.40% reported by Park et al) over the MORPH album 2 database, which is the largest public face aging database and 93% over the FG-NET database (compared to 38.10% reported by Geng et al).

[1]This paragraph of the first footnote will contain the date on which you submitted your paper for review. It will also contain support information, including sponsor and financial support acknowledgment. For example, "This work was supported in part by the U.S. Department of Commerce under Grant BS123456".

The next few paragraphs should contain the authors' current affiliations, including current address and e-mail. For example, F. A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail: author@ boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

London Journal of Research in Computer Science and Technology

Index Terms: anthropometric measurements, craniofacial growth, FG-NET, lip-nose complex, MORPH.

I. INTRODUCTION

Human perception studies reveal that attributes derived from one's appearance, such as one's emotional state, attractiveness, perceived age, etc. tends to significantly influence interpersonal behavior [1].

Hence, for many decades, human faces have been closely studied in computer vision and psychophysics with the objective of characterizing the many factors that induce appearance variations and subsequently unearthing information pertaining to an individual from his/her varying facial appearances. Facial aging is known as a complex process that varies with both the shape and the texture of the facial area. Shape variations include craniofacial, whereas texture variations include skin coloration, lines and/or wrinkles. Shape and texture are both considered as the common facial aging patterns. Since the aging process occurs throughout different ages, it can be classified into the above-mentioned aging patterns. As a result, an age-invariant facial recognition method should account for these aging patterns [2].

The applications of age-invariant facial recognition are lost children investigations, human-computer interaction, and passport photo verification. These applications possess two basic attributes: 1) significant age difference between probe and gallery photos (photos acquired during enrollment and also authentication phases), and 2) failure to acquire the person's facial photo in order to update the template [3].

Furthermore, real-world conditions continue to be challenging, mainly due to the great amount of changes in the process of acquiring faces. For example, criminal offense inspections and society safety organizations must frequently fit a probe photo with registered photos in a database, which may depict significant differences in facial characteristics due to the presence of different age groups. The effects of aging on the performance of facial recognition systems have not been thoroughly studied. Thus, there is a need to develop facial recognition algorithms, which are generally invariant towards aging [4].

Forensic scientists proved that human face aging strongly depends on ethnicity and gender [5].

Although human faces have the same general manner in aging, each ethnic and gender group has distinct characteristics. Therefore, it is insufficient to assume that similar faces age in similar ways for each and every individual.

Several physiological studies [6] [7] [8] have proved that the arithmetic range of the lip-nose complex measurements is different for different ethnicities, and thus can be adopted for ethnicity classification. The same studies have also proven that such measurements have a different arithmetic range of different genders within the same ethnicity. Moreover, some of the studies [7] [8] concluded that there is a possibility that the growth rate for such measurements is unique for each individual. Inspired by the aforementioned physiological facts, the lip-nose complex measurements were used as the first facial trait in the proposed system.

The Periocular region has the densest and the most complex biomedical features on the human face, e.g. contour, eyelids, eyeball, eyebrow, etc., which could all vary in shape, size and color.

Biologically and genetically speaking, more complex structure means more "coding processing" occurs with fetal development, and therefore more proteins and genes are involved in the determination of appearance. Moreover, the Periocular region goes through little changes over time because the shape and location of the eyes remain largely unchanged [9]. As it is stable across ages and has strong discriminative power, using the Periocular region for age invariant face recognition is rational and wise. In the proposed model, texture features extracted from the Periocular region are used as the second facial biometric trait.

The remainder of this paper is organized as follows: Section 2 surveys previous research

efforts related to age-invariant face recognition and the Morphometry of the Lip-Nose Complex. The proposed system is discussed in details in Section 3. Section 4 is dedicated to the experiments and results. The conclusion and future work are presented in Section 5.

II. RELATED WORK

When Face recognition across age has not been explored much in the past in spite of its importance in real-world applications. A detailed survey of the effects of aging on face verification tasks can be found in [1, 2]. Earlier approaches [3, 4] perform recognition by transforming one image to have the same age as the other, or by transforming both the images to reduce the aging effects. Table I summarizes a number of age-invariant face recognition studies.

Chellapa et al. [3, 5] proposed a craniofacial growth model that characterized the shape variations in human faces across age variations.

The authors observed that the growth parameter k for different face features across age can be adapted in the model to characterize the face growth. Ramanthan and Chellapa [4] also proposed a two-step approach for modeling aging in adults, which comprised of a shape and texture variation model. The formulation of shape variations is performed by constructing physical models which characterizes the functionalities of the face muscles.

Table 1: Published Age-Invariant Face Recognition Studies

Reference	Approach /Method	Description of the Proposed System
Ramanathan et al.[5]	Generative/ Mathematical Modelling.	Shape (Craniofacial) growth modelling up to age 18.
Geng et al. [7]	Non-generative/ Subspace method.	Learn aging pattern on concatenated PCA coefficients of shape and texture across a series of ages.
Park et al.[6]	Generative/ 3D modelling.	Learn aging pattern based on PCA coefficients in separate 3D shape and texture spaces from the given 2D database.
Ling et al.[8], [9]	Non-generative/ Feature-based method.	GOP+SVM
Mahalingam et al. [11]	Non-generative/Featu re-based method.	HLBP + AdaBoost classifier.

Park et al. [6] designed an aging simulation technique that learns the aging patterns of shape and the texture based on PCA coefficients. A 3D morphable model is used to model the aging variations from a set of 2D face images. To model the aging variations from two-dimensional images, a morphable model that is three dimensional is employed. The prototype technique is applied to the 3D face data. Although the prototype technique is able to extract average patterns, many details crucial for age perception such as wrinkles and pigments are neglected. Park et al. fitted their 3D morphable model to a set of face images by fitting an active appearance model (AAM) and extracting a 3D model from the AAM.

Aging is performed by calculating a set of weights between an input face and exemplar faces in the

London Journal of Research in Computer Science and Technology

same age group. These weights are then used to build an aged face as the weighted sum of the resulting faces at the anticipated age. After the appearance is predicted, they used commercial face recognition software to evaluate the recognition performance and observed that aging prediction model improves the performance of face recognition algorithms. The shortcomings in three-dimensional face recognition include the significant size of the three-dimensional models which usually demands high computation cost in matching, as well as the expensive price of three-dimensional imaging sensors. Building a model for craniofacial growth and another one for modelling texture variations with the possibility of the need to transform the 2D models to the 3D space increases the computational complexity and consequently the total processing time.

Geng et al. [7] learned a subspace of aging pattern based on the assumption that similar faces age in similar ways. Their face representation is composed of face texture and the 2D shape represented by the coordinates of the feature points as in the active appearance models.

The aging pattern is developed which is defined as the sequence of face images sorted in time order, by constructing a subspace. The aging pattern of the query image is determined by projecting it in the subspace. The position of the face image in aging pattern which is sorted in the time order indicates the age. The above method requires prior information like the actual age of the individual, the face feature points to model the aging pattern in the face images.

The shortcomings of the above mentioned approaches are that the information about the age of the probe image is required in order to perform the age transformation. This information is usually not available in real-world applications.

Also, the accuracy in age transformation relies on the accuracy of the aging model. Such inaccuracies may result in inappropriate age transformations causing instabilities to these approaches. Moreover, Construction of face models is difficult and sometimes they do not represent the aging process very well, especially when the training sample size is limited. Further, the face aging process is very complex and, consequently, in order to construct the aging model, strong parametric assumptions are needed, which are often unrealistic in real-world face recognition scenarios.

Discriminative approaches proposed in the past [8, 9] follow a non-generative approach to perform face recognition across age progression.

Ramanthan and Chellappa [5] proposed a discriminative approach for face verification across age progression. The authors adapted the probabilistic eigenspace framework and a Bayesian model to learn the differences between intra-personal pairs and extra-personal pairs. To avoid the challenges caused by illumination variations, they use the half face that has better illumination (termed as PointFive faces) and use the symmetry property of the face to construct a face image.

Ling et al. [9] also used a non-generative discriminative approach for face verification of age separated images. They propose a face representation called gradient orientation pyramid, in which a Gaussian pyramid is generated for each face image and the gradient orientation is computed for each pixel at all levels of the pyramid. SVM is then used to classify the image pairs as intra-personal or extra-personal.

Bereta et al. [10] compared seven local descriptors commonly used in face recognition, e.g. LBP and WLD, and classify images by calculating a distance between feature vectors containing local texture. The Gabor coded MBLBP feature combined with the Euclidean distance brings the best recognition accuracy. Bereta et al. conducted experiments using MB-LBP to check the robustness of the features against ageing using FGNET database and with MB-LBP and Euclidean distance as similarity metric obtained 70% recognition accuracy. They also evaluated MB-LBP method along with Gabor features using Euclidean and cosine distances as similarity measurements and achieved 76% and 75% recognition rates respectively.

Mahalingam et al. [11] proposed a hierarchical local binary pattern (HLBP) feature descriptor used for face representation across age.

The representation by HLBP across minimal age, illumination, and expression variations combined with its hierarchical computation provides a discriminative representation of the face image.

The proposed face descriptor is combined with an AdaBoost classification framework to model the face verification task as a two-class problem. HLBP for each image is constructed by computing the uniform LBP at every level of the image pyramid and concatenating them together to form the HLBP descriptor. A dense sampling using LBP local descriptors allows the extraction of discriminatory information such as the distribution of the edge direction in the face image which is an age invariant feature.

A non-generative discriminative approach for the task of face recognition of age separated images is proposed in this research study. The framework proposed in this research study differs from the above mentioned non-generative approaches in the face representation and the classification framework. A discriminative feature based face representation coupled with a classification framework is proposed. The proposed framework has been applied to two face aging databases, which include both internal and external variations in the face images.

III. THE PROPOSED SYSTEM

A block diagram of the proposed age invariant face recognition framework is illustrated in Fig. 1.

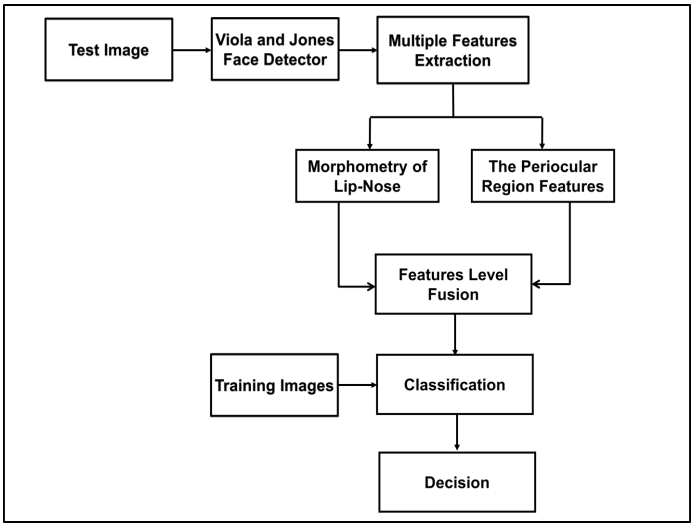


Fig. 1: Magnetization A Block Diagram of the Proposed Age Invariant Face Recognition Framework

Assume a framework where the test image may be derived from an image or a sequence of images containing human faces. The face area is detected using the real time implementation of the Viola and Jones face detector [109]. Such a detector is effective for real time applications. The training patterns of the Viola and Jones face and eye localizer are from a set of images with predefined coordinates (i.e., annotation). Based on these coordinates, the face feature vector of size 12x12 is extracted from the original face image using the landmark template (LMT) features [110]. The LMT features consist of multi-scale Gabor features (at 6 scales and 12 orientations) applied to 13 landmark facial points that are locally searched within the bounding box starting from the tracked state [111].

To detect a face in an input image, each of the possible sub-images is processed to see if it represents a face. The input images are preprocessed before inputting to the face detector.

A confidence score is produced by the face detector, indicating the system's confidence on this detection result. If the score is below some threshold, then no face is detected.

At the preprocessing stage, all face images in FG-NET and MORPH databases were properly

normalized and pre-processed. The pre-processing stage comprised of converting the color input images into 8-bit greyscale images and normalizing the face images photometrically by eliminating their means and scaling their pixels to unit variance [40]. Finally, pose correction was performed to non-frontal face images using the same method adopted by Kambhamettu et al. [41], which uses Active Appearance Model (AAM) technique as proposed by Cootes et al. [42]. A number of preprocessed sample images from the FGNET and MORPH databases are illustrated in Fig. 2 where the images in the top row are the original images and the images in the bottom row are the preprocessed images.

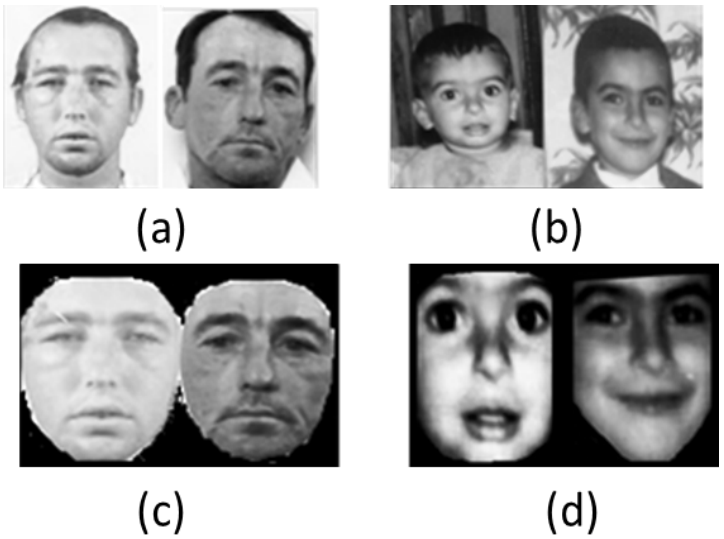


Fig. 2: Images Taken from the FG-NET and MORPH Databases in Panels (a) and (b) respectively Original Images are in the top row and Preprocessed Images are in the bottom row

In the proposed system, only 11 facial landmarks were considered and automatically localized using a facial landmark detector introduced recently by Zhou et al. [34], known as Exemplar-based Graph Matching (EGM). For the training stage, localizing the related landmarks was performed manually only for the MORPH database since the FGNET database was provided with annotations for a set of 68 landmarks, which included the 11 landmarks considered in the proposed system. For the MORPH database, about 1132 images were manually annotated and used for training the system. Fig. 3 illustrates examples of the automatic landmark localization using the EGM detector on sample image from the MORPH database.

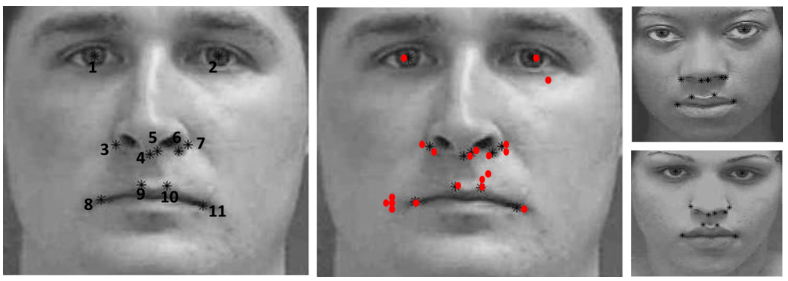


Fig. 3: Example of Response Map and Candidate Generations for a Face Image Collected from the MORPH Database. (a) A testing image labeled with 11 landmarks. (b) A 1/7 portion of all the candidates. The Black marks indicate the candidates that are transformed from exemplars, while red ones are the peaks of the response maps. (c) The top two exemplars found by the RANSAC step

Once the set of face landmarks is localized, a multilevel feature extraction is performed such that two set of face features are extracted from the test image and later fused. The first set of face features is a set of anthropometric measurements known as the lip-nose complex measurements.

Since angular measurements and tangential measurements on faces cannot be estimated accurately using photogrammetry of face images, a set of linear projective measurements acquired from the nose-lip complex are taken into account. Thus, five lip-nose complex measurements are considered in the proposed framework as follows:

1. Vertical length of the lip: bottom of the columella up to the highest point of the Cupid's bow. This specific parameter was basically identified alongside the bottom of columella up to the tubercle.
2. Cupid's bow width (philtrum peak to peak): absolute breadth for oral cavity (commissural to commissural length).
3. Columella width: the portion between the nostrils.
4. Nose area width: Distance from one alar to another.
5. Total width of the mouth: Distance from one oral commissure to another.

Fig. 4 depicts the lip-nose complex measurements considered in the proposed framework.

For face recognition problems, the face measurements extracted from face images lose their absolute scales because face images can be easily enlarged. Since the absolute scale information is lost in the face images, a relative scale has to be adopted to normalize the face measurements [5]. Moreover, ratios of distance, area, and angles are usually measured to compensate for the varying size of images [14].

The distance between the centers of two eyes is usually not affected by expressions and thus is relatively stable. Hence, this distance is used in this work as the reference scale. After calculating the set of lip-nose complex measurements, they are scaled through dividing each of the measurements by the distance between the centers of two eyes. Fig. 5 illustrates the notations related to the coordinates of the face landmarks considered in the proposed framework.

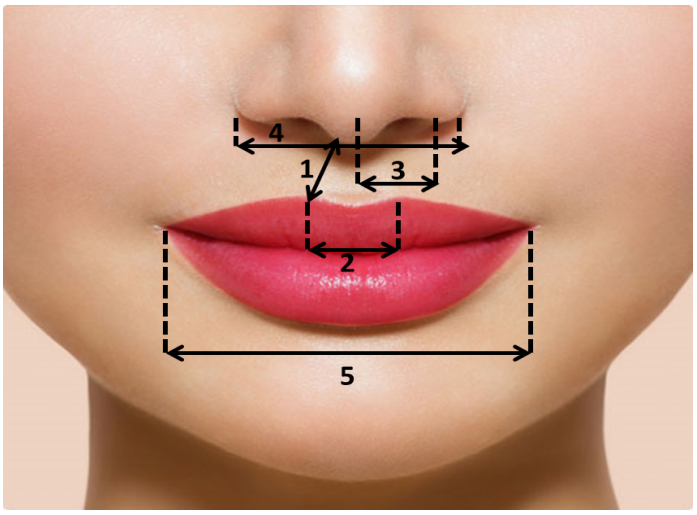


Fig. 4: The Lip-Nose Complex Measurements used in the Proposed Framework

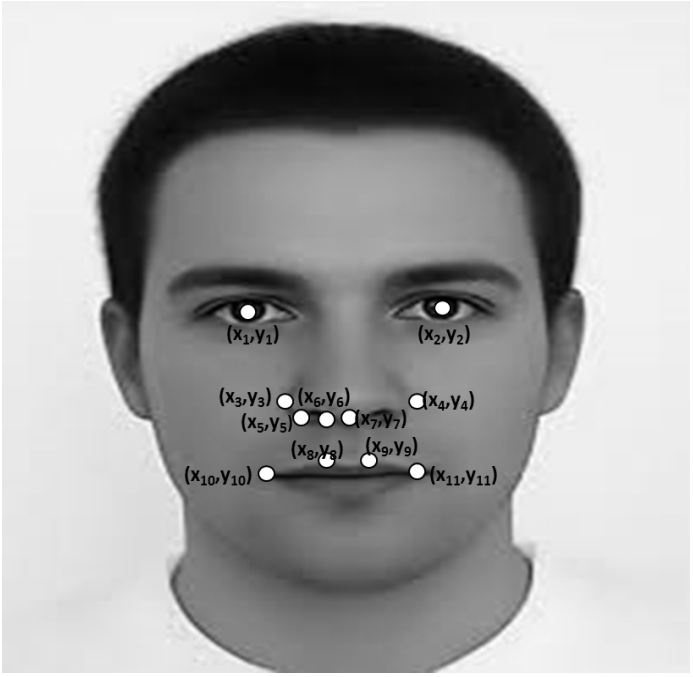


Fig. 5: Notation of the Face Landmarks Coordinates

After localizing the set of 11 face landmarks using the EGM detector as described previously, the coordinates of the points are then passed as inputs to a set of Euclidean distance mathematical equations presented by equations (1) through (6) to determine the lip-nose measurements for each sample face. A feature vector containing the geometric features (scaled lip-nose complex measurements) is created for each sample image, and a class incorporating all the geometric feature vectors related to the sample images, is created for each subject.

$$d_1 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

$$d_2 = \frac{\sqrt{(x_3 - x_4)^2 + (y_3 - y_4)^2}}{d_1} \quad (2)$$

$$d_3 = \frac{\sqrt{(x_5 - x_6)^2 + (y_5 - y_6)^2}}{d_1} \quad (3)$$

$$d_4 = \frac{\sqrt{(x_7 - x_8)^2 + (y_7 - y_8)^2}}{d_1} \quad (4)$$

$$d_5 = \frac{\sqrt{(x_8 - x_9)^2 + (y_8 - y_9)^2}}{d_1} \quad (5)$$

$$d_6 = \frac{\sqrt{(x_{10} - x_{11})^2 + (y_{10} - y_{11})^2}}{d_1} \quad (6)$$

Where x_i and y_i $(i = 1 \ldots 11)$ are the horizontal and vertical coordinates of the face landmarks respectively. Also, d_1 is the distance between the centers of the eyes, d_2 is the nose total width, d_3 is the Columella width, d_4 is the vertical length of the lip, d_5 is the Cupid's bow width, and d_6 is the total width of the mouth. Since the FG-NET database consists of 82 subjects, a dataset of 82 classes is created as illustrated in Table II.

Table 2: The Lip-Nose Complex Measurements Extracted From One Subject's Samples in the Fg-Net Face Database

Samples	Features d1=0.2					Class (subject ID)
	Total nose width	Colum-ella width	Vertical length of the lip	Cupid's bow width	Total width of the mouth	
Age 0	3.15	0.75	1.65	1.15	3.3	1
Age 1	3.15	0.75	1.65	1.15	3.55	1
Age 2	3.3	0.75	1.65	1.25	3.85	1
Age 3	3.4	0.8	1.65	1.3	3.9	1
Age 4	3.45	0.9	1.7	1.35	4.15	1
Age 5	5.9	1.65	3	2.8	7	1
Age 6	6.1	1.75	3	2.85	7.9	1

The periocular region contains the most discriminating features that can vary across individuals, which makes it an effective soft biometric trait. Moreover, early anthropometric studies agreed that the periocular region is stable with aging. The proposed framework outperformed the LBP, Gabor, and a number of the-state-of-the-art age-invariant face recognition systems with a maximum verification accuracy of more than 95% for the MORPH album 2, which is the largest public face aging database.

The framework preserved the simplicity since it does not require building aging models or any pre-processing step prior to performing the recognition task. Such framework is promising for real time applications that require high performance in addition to low processing time.

The second face trait is a set features extracted from the periocular region using LBPVp,r and GIST descriptors. The periocular region is a small region in the neighborhood of the eye and includes the eyebrows (the definition of the periocular region provided here is specific to this work.). This region usually encompasses the eyelashes, eyelids, eyebrows, and the bordering skin region, depending on the size of the image used.

Using the periocular region has the following advantages: (a) the information regarding the shape of the eye and the surrounding skin texture could vary across individuals, which can be used as a soft biometric trait, (b) no additional sensor besides the iris camera was required to acquire the periocular data [15], and (c) periocular region is perceptually more stable across ages than the full face [16] (Early anthropometric studies agreed that the periocular region has stability across ages.).

Park et al. [17], [18] in their feasibility study, they concluded that the periocular region is best discriminated through the fusion of global and local descriptors. Based on their conclusions, this research work adopts an approach to combine the global and local features in order to obtain the benefits of both of these. The choice of the appropriate and complementary component features is vital for good performance. In view of this, the proposed method includes the Gist descriptor in order to obtain the global features.

The Gist descriptor aggregate image statistics of the responses of several filters combined with the input image. The advantage of the descriptor is that it is very compact and fast to compute. In the standard configuration, each image is represented by a 320-D vector per color band, resulting in a 960-D descriptor per image. The feature vector corresponds to the mean response to steerable filters at different scales and orientations.

Likewise, the LBP descriptor is able to efficiently extract the local details of face images, i.e. image edges, peaks, etc., even in the presence of noise. The LBP features are invariant to changes in lighting and scale. LBP has been recognized as computationally simple and efficient method which is invariant to monotonic gray scale transformation and illumination changes [19].

The idea of using LBP in face description is motivated by the fact that faces can be seen as a composition of micro-patterns which are well described by such operator.

As a variance of LBP, LBPV is robust because it exploits the complementary information of local contrast and spatial pattern. Both of the two descriptors (Gist and LBPV) are rich in information and computationally efficient.

Therefore, the designed method will comprise of the invariant characteristics of both the approaches, and it will be able to generate the optimal results against most of the variations (pose, expression and illumination) present in the periocular region images beside the variations introduced by the face aging. In the feature fusion model, Gist and LBPV features were extracted and normalized separately, and then, their feature vectors were simply concatenated and the singular values of the fused feature vector are projected to a PCA subspace for dimensionality reduction.

The periocular region is first extracted from the face test image using the ground truth eye centers for the faces. The ground truth eye centers of the faces were provided in FG-NET datasets, which acted as the centers of the periocular images to be cropped from the original face images. The size of the cropping region was calculated by the ratio of the distance between the eye centers. The periocular region images were then scaled down to a uniform size of 100x120 pixels. A number of preprocessed periocular regions that were extracted from the FGNET database are illustrated in Fig. 6. The sample periocular regions images were cropped from a set of images related to the same subject at different ages.

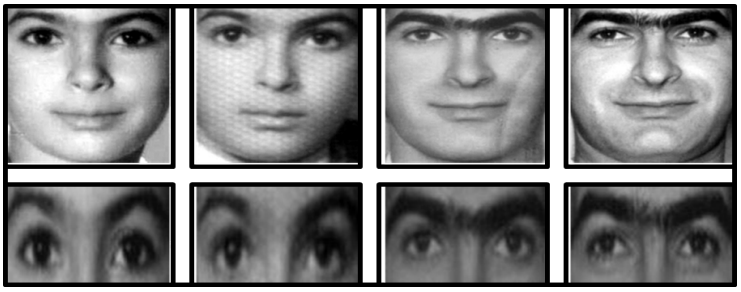


Fig. 6: Examples Periocular Regions from the FG-NET

As for the MORPH database, the eye centers of the faces were localized using the exemplar-based graph matching (EGM) landmarks detector [20]. Then, the cropping of the periocular region was performed in the same manner as the FG-NET database. Fig. 7 shows a number of sample images collected from the MORPH database of one subject at different ages and the preprocessed periocular regions that were extracted from the images. The first row shows the original images and the second row shows the preprocessed periocular regions.

Fig. 7: Examples Periocular Regions from the MORPH

The objective of using a global descriptor for periocular recognition is to obtain a basic and a subordinate level description of the perceptual dimensions. GIST descriptor [21] effectively encodes the scene images where the distance between a fixated point and the observer is large (greater than four meters) [21].

The gist descriptor is a vector of features G, where each individual feature G_k is computed in Equation (7) as follows:

$$G_k = \sum_{i,j} w_k(i,j) \times \left| I(i,j) \otimes h_k(i,j) \right|^2 \quad (7)$$

Where (i,j) are a pixel position, $\otimes$denotes image convolution, and $\times$ is a pixel-wise multiplication. The luminance channel of the input image is represented by $I(i,j)$. $h_k(i,j)$ is a filter from a bank of multi-scale-oriented Gabor filters (eight orientations and four scales), and $w_k(i,j)$ is a spatial window that will compute the average output energy of each filter at different image locations. The windows $w_k(i,j)$ divide the image into a grid of 4×4 non-overlapping windows. This results in a descriptor with a dimension of $4 \times 4 \times 8 \times 4 = 512$.

Here, a set of five perceptual dimensions, namely, naturalness, openness, roughness, expansion and ruggedness are used to give a low dimensional, holistic representation of the image.

- Degree of Naturalness: This spatial property describes the distribution of edges in the horizontal and vertical orientations. It describes the presence of artificial elements such as spectacles.
- Degree of Openness: The second major attribute describes the presence or lack of points of reference. An image with a higher percentage of periocular regions than sclera and iris region will have less points of reference or be more 'open'.
- Degree of Roughness: This perceptual attribute refers to the size of the largest prominent object in the image. It evaluates the common global attributes of the image.
- Degree of Expansion: This attribute describes the depth in the gradient of the space within the image.
- Degree of Ruggedness: This attribute gives the deviation from the horizontal by assessing the orientation of the contours of the image.

These perceptual properties are correlated with the second order statistics and spatial arrangement of structured components in the image [21]. They are easy to calculate and can be translated to a useful global descriptors of the periocular region. The GIST implementation in this research study receives a fixed size square periocular area image as input and produces a vector of dimension 512 as output.

The systems that use GIST descriptor, mostly resize the input image in a preprocessing stage, to produce a small square image. The width of the image is typically in the range from 32 to 128 pixels. Such approach is sufficient due to the low dimensionality of the descriptor, i.e., it does not represent the details of an image. In the proposed framework, an image size of 32×32 pixels is used. The periocular area images are all re-scaled to that size regardless of their aspect ratio.

A number of periocular area images which are related to the same subject (from the FG-NET database) at different ages are depicted in Fig. 8 along with visualization of their Gist descriptor features. The original periocular area images are illustrated in the top row and their Gist descriptor features are illustrated in the bottom row. Also, the age of the subject is illustrated for each image.

The periocular skin texture has been used for human identification in various ways. Jain et al. [22] detected micro-features such as moles, scars, or freckles and used them as soft biometric traits.

Others adopted a more general representation of the overall texture to facilitate recognition using popular texture measures such as the discrete cosine transformations (DCT) [23], the gradient orientation histograms (GOH), or the LBP [18],[24]. A variance of the LBP (LBPV) for extracting the overall texture of the periocular region is adopted to facilitate recognition in this research study. Basically, LBP is a fine-scale descriptor that captures small texture details.

As a variance of LBP, LBPV is robust because it exploits the complementary information of local contrast and spatial pattern. LBP variance (LBPV) was first introduced by Zhenhua Guo et al. [25] to characterize the local contrast information into one dimensional LBP histogram. It is simplified yet effective combined LBPs, as well as a method

of contrast distribution. LBPVp,r /VARp,r is robust because it exploits the complementary information of local contrast and spatial pattern.

LBPV/VAR descriptor exploits the supporting information for the local spatial pattern, as well as local contrast.

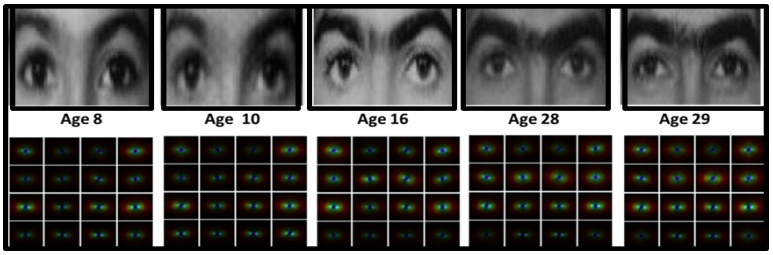


Fig. 8: A Number of some Periocular Region and their Gist Descriptor Features

The VAR possesses a continuous value, which needs to be quantized. That can be carried out first by computing feature distributions out of all training images to acquire an overall distribution.

Following that, a few threshold values are calculated to partition the overall distribution into N bins using an equal number of entries [26] to guarantee the maximum quantization resolution.

The threshold values are employed to quantize the VAR of the test images. There are three specific limitations to this quantization technique pointed out by [27]. The LBPV offers a way to handle these difficulties associated with the descriptor.

Typically, the information associated with the variance VAR will not be involved in the calculations of the histogram of the LBP. The histogram operation allocates identical weight to every LBP pattern independent of the LBPV of the local region. The LBPV is presented by Equations (8) and (9) as follows:

$$LBPV(K) = \sum_{i=1}^{n} \sum_{j=1}^{m} W(LBP(i,j),k),k \quad (8)$$

$$W(LBP(i,j),k) = \{VAR(i,j)\ LBP(i, \quad (9)$$

Where n and m are the dimensions of the texture image, (i, j) is the position of each pixel, K is the maximal LBP pattern value, and $W(LBP(i,j),k)$ is the adaptive weight to adjust the contribution of the LBP code. LBPV can be described as a simplified descriptor whose feature size is small in a way that it can be employed in several applications. Additionally, it is training free and does not require quantization. Fig. 9 depicts the Periocular region extracted from FG-NET database images of one subject at different ages and their LBPV transformations, where the first row shows the original Periocular region images and the second row shows their LBPV transformation.

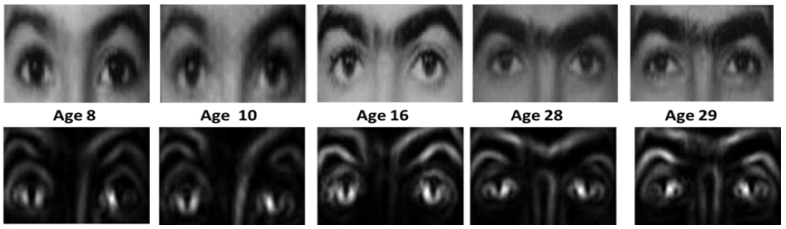


Fig. 9: Examples of Periocular Regions their LBPV Transformations

In the event that the lip-nose complex and periocular region feature vectors are formed, fusion of the two face features vectors is performed at the feature level. The feature sets related to the two face traits are fused at the feature level since it has been observed that, a unimodal biometric system that integrates information at an earlier stage of processing is expected to provide more accurate results than those that integrate information at a later stage, because of the availability of richer information [28]. In order to be concatenated, the feature sets must be compatible. Thus, the two feature sets

are normalized prior to fusion using the well-known Z-score rule.

The feature sets are first normalized separately before concatenating them into a single combined feature vector. The goal of feature normalization is to modify the location (mean) and scale (variance) of the feature values in order to ensure that the contribution of each component to the final match score is comparable. Feature sets normalization is performed using the Z-score rule [29]. The Z-score offers a useful measurement for comparing data elements from different data sets.

Typically, the formula used for calculating the Z-score is represented by Equation (10):

$$z = \frac{x-\mu}{s} \qquad (10)$$

Where z represents the Z-score, x is the feature vector, μ is the mean of the feature vector, and s is the standard deviation of the feature vector. After normalizing the two face features sets two arrangements for concatenating the face feature sets are considered in this research study. In the first arrangement, the normalized face feature sets are concatenated and passed immediately to the classification module. Fig. 10 illustrates the implementation of this arrangement over the FG-NET face aging database [30] which contains 1002 images of 82 subjects. The final feature vector extracted from each image is of size 1x576. The final dataset built using the FG-NET database entire sample images is a feature matrix of size 1002x576.

In the second arrangement, the two normalized face features sets are concatenated and the singular values of the fused features set is projected into a PCA subspace dimensionality reduction as illustrated in Fig. 11. The PCA is one of the most popular linear techniques for dimensionality reduction. It performs a linear mapping of the data to a lower-dimensional space in such a way that the variance of the data in the low-dimensional space is maximized [31].

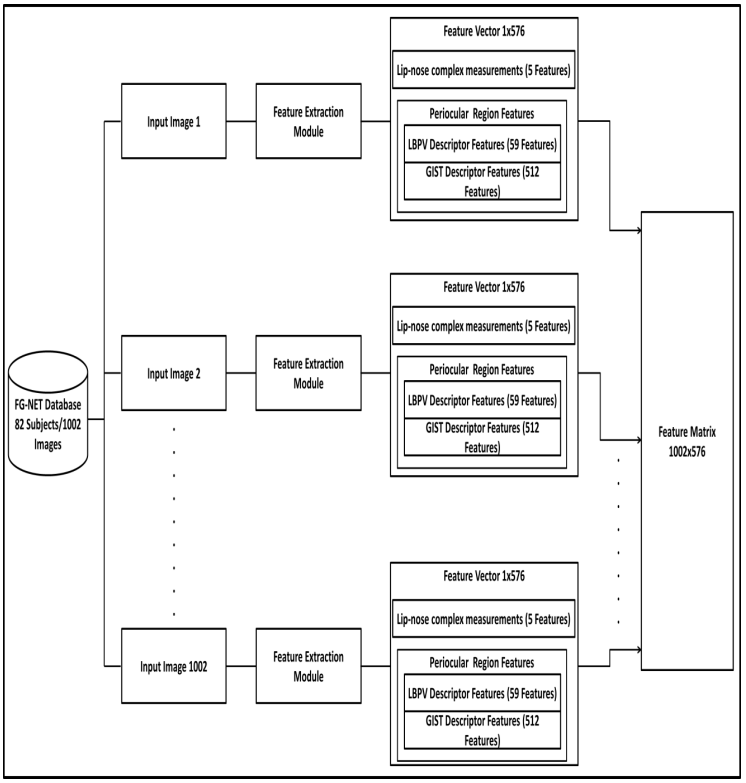


Fig. 10: First Arrangement of Feature Sets Concatenation

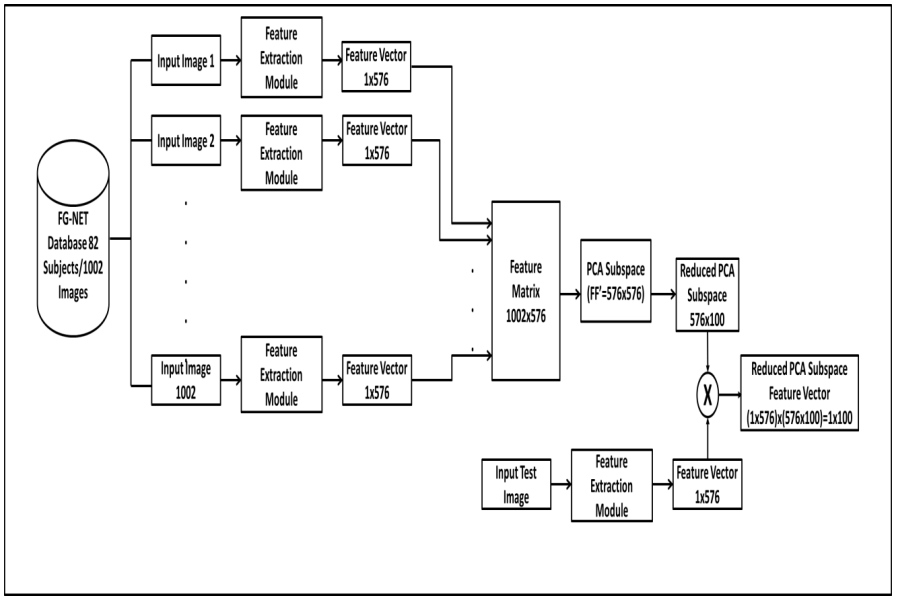


Fig. 11: Second Arrangement of Feature Sets Concatenation

The main idea of PCA is to find a reduced set of vectors that can be used to describe the full data with high accuracy. The advantage of PCA comes from its generalization ability. It reduces the feature space dimension by considering the variance of the input data. The method determines which projections are preferable for representing the structure of the input data. Those projections are selected in such a way that the maximum amount of information is obtained in the smallest number of dimensions of feature space. In order to obtain the best variance in the data, the singular values of the data are projected to a subspace of the feature space which is built by the eigenvectors from the data. In that sense, the eigenvalue corresponding to an eigenvector represents the amount of variance that eigenvector handles.

The principal components analysis (PCA) is performed as given below, and returns the principal component coefficients Co, also known as loadings. Co is a 576-by-576 matrix, each column containing coefficients for one principal component. The columns are in order of decreasing component variance. Projecting the singular values of the feature matrix into a PCA subspace is performed in the following steps:

1. The Feature matrix F is first centered: the deviation of the features from its mean is calculated by subtraction off the mean feature vector. Thus, F is transformed from the original space to the zero mean space F_0. Subtracting the mean makes variance and covariance calculation easier by simplifying their equations as illustrated in Equations (11) and (12):

$$F_{0i} = f_{ij} - \mu_i \tag{11}$$

$$F_0 = \{F_{01}, \ldots, F_{0576}\} \tag{12}$$

where μ_i is the mean of each feature column, f_{ij} is a data point in feature vector i, $i = 1, \ldots, 576$ is an index of feature column, and $j = 1, 2, \ldots, 1002$ is an index of a data point. F_{0i} is the zero mean space of feature i.

2. The covariance matrix V can then be obtained as $V = F_0 F_0^T$.
3. The eigenvectors E_k are calculated from V and sorted according to their corresponding eigenvalues γ_k.

The first n eigenvectors are then selected to span a space of reduced dimensionality onto which the original data can be projected (given by the ratio of the sum of neglected eigenvalues to the sum of all eigenvalues).

4. In the proposed framework 95% of information was kept, and the resulting

number of eigenvalues is 100. Thus, the final size of the reduced PCA subspace is 576 x 100.

5. Each feature vector of the 1002 feature vectors in the feature matrix F is projected to the reduced PCA subspace, which results in a feature vector of size 1x100 as illustrated in Fig 11.

In the final stage of the proposed framework classification is performed to produce the final output of the proposed framework. In the classification module two classifiers were adopted in this research study. The first classifier is the well-known k-means clustering unsupervised classifier [32]. K-means [33] is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. In this research study K-means unsupervised classification scheme1 is adopted. In this classification scheme, K-means is applied over each class separately to produce a set of prototypes for each class where each class is related to an individual subject. The main steps of K-means unsupervised classification scheme (1) can be summarized in the following points:

- Apply k-means clustering to the training data in each class separately, to produce M prototypes for each subject class.
- Assign a class label to each of the K x M prototypes. Where K is the number of subjects classes.
- Classify a new feature vector x to the class of the closest prototype. The closest prototype is the prototype that has the smallest square Euclidean distance from the feature vector x.

The main steps of performing K-means unsupervised classification scheme (1) are illustrated in Fig. 12 where the FG-NET database was used as a model for illustrating the process.

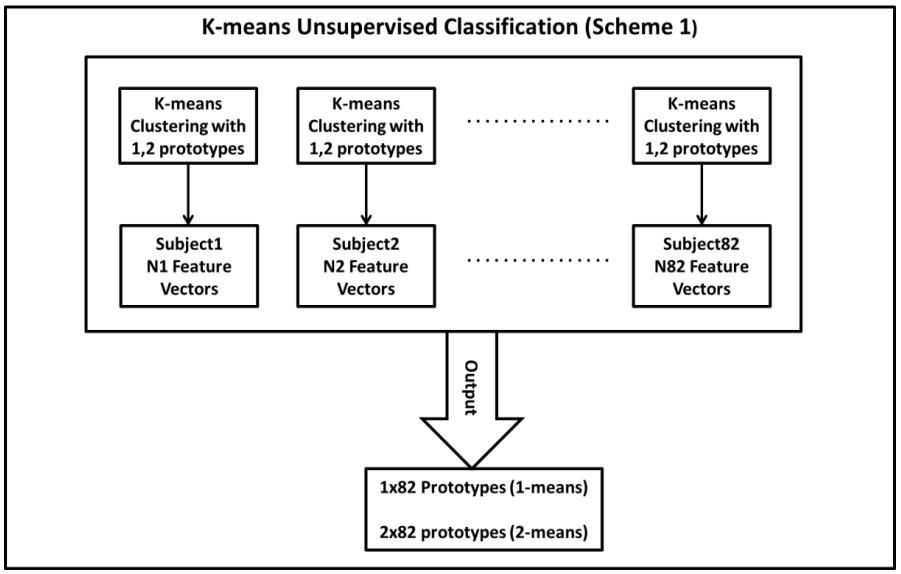


Fig. 12: K-means Unsupervised Classification

K-means clustering is performed over each subject class separately in the following steps:

Initialize Cluster Centers: In K-means it is assumed that there are M prototypes (cluster centers) denoted by $P = \{p_1, p_2, ..., p_M\}$ where this set is smaller than the original data set. If the data points reside in a Euclidean space, the prototypes reside in the same space. They will also be n dimensional vectors (n=576 in the proposed framework.). They may not be samples from the training data set; however, they should well represent the training data set. In the experimental setup, k-means clustering was performed with $M = 1$ and with $M = 2$ respectively i.e. one-means and two-means clustering was used in the experiments. Also, initialization of cluster centers was performed

using randomly selected feature vectors from each cluster.

Assign each Feature Vector to Nearest Cluster Center: Each training sample is assigned to one of the prototypes. This is denoted using the assignment function by $A(\bullet)$ using Euclidean distance. Thus, $A(x_i) = j$ means the i^{th} training sample is assigned to the jth prototype, that is, the jth prototype is used to approximate or represent point x_i.

where

$$A(\bullet) = argmin_{j\in\{1,2,\ldots,M\}}\|\bullet - z_j\| \quad (13)$$

Calculate Means of New Cluster Centers: The centroids are updated by computing the average of all the samples assigned to it. For example, if a set of samples $(x_1, x_2, \ldots, x_n)$ were assigned to a prototype j. The centroid will be updated such that the new value of j is calculated as follows:

$$j = \sum_{i=1}^{n} x_i \quad (14)$$

Check the termination conditions: In K-means, the optimization criterion is to minimize the total squared error between the training samples and their representative prototypes. This is equivalent to minimizing the trace of the pooled within covariance matrix. The objective function is:

$$argmin_{\smile P,A} \sum_{i=1}^{N} \|x_i - Z_{A(x_i} \quad (15)$$

For every pointx_i, a prototype z_j is found to represent x_i. As mentioned above, if $A(x_i) = j$ then the prototype z_j will be used. Hence, the distance between xi and $Z_{A(x_i)}$ is computed. The norm in the equation (3.15) denotes the Euclidean distance. The objective function is represented by equation (3.16):

$$L(Z, A) = \sum_{i=1}^{N} \|x_i - Z_{A(x_i)}\|^2 \quad (16)$$

The necessary conditions for an optimal solution (termination conditions) are as follows; If the set of prototypes ,Z, is fixed, the optimal assignment function $A(\bullet)$ (using the Euclidean distance) should follow the nearest neighbor rule, that is,

$$A(x_i) = argmin_{j\in\{1,2,\ldots,M\}}\|x_i - . \quad (17)$$

Given the assignment function, $A(\bullet)$, fixed, the prototype z_j should be the average (centroid) of all the samples assigned to the jth prototype:

$$z_j = \frac{\sum_{i:A(x_i)=j} x_i}{N_j} \quad (18)$$

where N_j is the number of samples assigned to prototype j. The algorithm converges since after each iteration, the objective function decreases (non-increasing). Based on the necessary conditions, the algorithm improves the prototypes and the assignment function iteratively.

It is possible that the steps taken might not lead to any change in either the prototypes or the partition. When either becomes static, the algorithm has converged and there is no need to continue. Figure 3.13 depicts the main steps of K-means clustering.

The second classifier adopted in the proposed framework is a distance-based classifier. The distance-based classifier used in this research study is a KNN classifier [34] combined with cosine score [35]. The cosine score is used for producing the s between two feature vectors. The cosine score has two advantages over other distance metrics; 1) it is scale invariant since it is an angular distance between two vectors. 2) It gives more weight to dimensions having high value (peaks in the histogram) while the Euclidean distance weighs all dimensions equally [36]. The cosine score is calculated using the formula presented in Equation (19).

$$d_{cos}(F_i, F_j) = 1 - \frac{F_i . F_j}{\|F_i\|\|F_j\|} = 1 - \left(\frac{\sum_{i,j=1}^{n} F_i \times F_j}{\sqrt{\sum_{i=1}^{n}(F_i)^2} \times \sqrt{\sum_{j=1}^{n}(F_j)^2}}\right) \quad (19)$$

Where F_i and F_j are the feature vectors matched using the distance-based classifier. Once the s are produced using the cosine score, they are compared with a predefined threshold. Based on the results of comparing the s with the reference threshold, the KNN classifier makes a final decision.

The different feature representations illustrated in Fig 10 and Fig 11 are combined with the K-means classifier and the cosine distance-based classifier in an alternate way. Such combinations resulted in eight different age-invariant face recognition algorithms. The proposed algorithms were mainly designed to perform face verification, but they were validated in the identification mode as part of the experimental work. The proposed algorithms were given names as 1.a, 1.b, 2.a, 2.b, 3.a, 3.b, 4.a, and 4.b for simplicity. A brief description of the proposed age invariant face recognition algorithms is illustrated in Table III to remind the reader of the differences between the proposed algorithms.

Table 3: The Proposed Age Invariant Face Recognition Algorithms

Samples	Class (subject ID)
1.a	PCA+K-means with one prototype.
1.b	PCA+K-means with two prototypes.
2.a	Non-PCA+K-means with one prototype.
2.b	Non-PCA+K-means with two prototypes.
3.a	PCA+Cosine-based classifier with mean prototype.
3.b	PCA+Cosine-based classifier with max prototype.
4.a	Non-PCA+Cosine-based classifier with mean prototype.
4.b	Non-PCA+Cosine-based classifier with max prototype.

IV. EXPERIMENTS

All the face images in the FG-NET and MORPH databases were properly normalized and pre-processed. The pre-processing stage comprised converting the color input images into 8-bit grey-scale images, locating the eyes manually, normalizing (scaling and rotating) the images geometrically in such a way that the centers of the eyes were localized at predefined positions, cropping the face parts of the images, and resizing the cropped area to a standard size of 200x200 pixels, and finally, normalizing the face images photometrically by eliminating their mean and scaling their pixels to unit variance [37]. A number of examples of the normalized images from the FG-NET and MORPH databases are shown in Fig 13 and Fig 14 successively.

Fig. 13: Pre-processed Sample Images of 5 different subjects from the FG-NET

Fig. 14: Pre-processed Sample Images of 5 different subjects from the MORPH

For all experiments, frontal faces were normalized using the location of 68 manually established face feature points. These points are triangulated and the image warped with a piecewise affine warp onto a coordinate frame in which the canonical points are in fixed locations.

This process is similar to the preprocessing used prior to the computation of AAMs [38], [39]. Finally, pose correction was performed for non-frontal face images using the same method adopted by Kambhamettu et al. [40], which uses AAM technique as proposed by Cootes et al. [41].

As mentioned before, the images from FG-NET are annotated with 68 face feature points; consequently, a generic model was used to fit these points and calculate the pose of the face.

Finally, pose correction was performed by warping the image onto the model. A number of pose corrected sample images from the FGNET and MORPH databases are illustrated in Fig 15, where the images in the top row are the original images and the images in the bottom row are the pose corrected images.

The verification tests were performed using the leave-one-person-out (LOPO) scheme adopted by Ling et al. [9] and Park et al. [6] (most commonly used for evaluating face recognition systems). In the first experimental setup, verification was performed using k-means clustering unsupervised classification. Each image in the FG-NET and MORPH databases was used once as a test image.

The feature vector acquired from the test image is matched with all the prototypes of the different clusters (classes) related to the different subjects.

The genuine similarity scores are produced when a test feature vector is matched with a prototype that belongs to the same subject as the test feature vector. The imposter scores are produced when a test feature vector is matched with a prototype that belongs to a different subject. In the case when k-means clustering is performed with two prototypes, the maximum similarity score produced from matching the test feature vector with the two prototypes of each cluster is considered the final matching score.

Fig. 15: Sample Images Collected Before and After Preprocessing

In the second experimental setup, identification was performed using K-nearest neighbor classification to acquire the verification rate for the proposed approach. In the LOPO scheme, the

entire set of images of a single subject was used as test images and the rest of the images were used for training the framework. When a test feature vector is matched with a training feature vector that has the same true class label as the test vector a genuine matching score is produced. When a test feature vector is matched with a training feature vector that has different true class label than the test vector label, an imposter similarity score is produced. Finally, the genuine and imposter scores in both of the experimental setups are used to calculate the verification rate.

The close-set identification tests were also performed using the leave-one-person-out (LOPO) scheme. In the first experimental setup, identification was performed using k-means clustering unsupervised classification. Each image in the FG-NET and MORPH databases was used once as a test image. The feature vector acquired from the test image is matched with all the prototypes of the different clusters (classes) related to the different subjects. In the case of the FG-NET database there are 82 classes. Then, the test feature vector is assigned the label of the prototype that yielded the maximum similarity score. In the case when k-means clustering is performed with two prototypes, the maximum similarity score produced from matching the test feature vector with the two prototypes of each cluster is considered the final similarity score. In the second experimental setup, identification was performed using K-nearest neighbor classification to acquire the identification rate for the proposed approach. In the LOPO scheme, the entire set of images of a single subject was used as test images and the rest of the images were used for training the framework. This procedure was performed for all the subjects to ensure that each subject was used only once for testing. Also, the purpose of this evaluation scheme was to make sure that the images of one subject were not in the training and testing sets at the same time. The LOPO was combined with K nearest neighbor (KNN) classifier in a way that every single image in the face aging databases is used as a query image and cosine score is computed between each query and each of the remaining images. Then, the identity of the subject presented in the query image is recovered using the KNN match. In the experiments over the FG-NET database one-Nearest Neighbor was used. Since the MORPH database is a very large database, the optimum value of K was 3 based on the empirical validation.

V. RESULTS

In this section the results of evaluating the performance of the proposed framework and the proposed algorithms within the framework are presented and discussed. As mentioned before, the proposed algorithms within the framework were evaluated over the FG-NET and MORPH databases in the verification and identification modes.

The performance of the k-means unsupervised classifier degrades when the feature vectors (feature matrix) was passed to it immediately without projecting it to a PCA subspace. This can be observed when the performance of algorithms 2.a and 2.b is compared with the performance of algorithms 1.a and 1.b. Such results show the effect of projecting the data into a PCA subspace in enhancing the performance of k-means unsupervised classifier. The ROC curves depicted in Figure 4.3 shows clearly that replacing k-means clustering classifier with a distance-based classifier (cosine distance based classifier [42, 43]) has led to a degradation in the verification performance even when the data was projected to a PCA subspace prior to classification. The lowest performance was reported when the feature matrix was derived to the cosine distance based classifier without projecting it to a PCA subspace as shown in Fig 16.

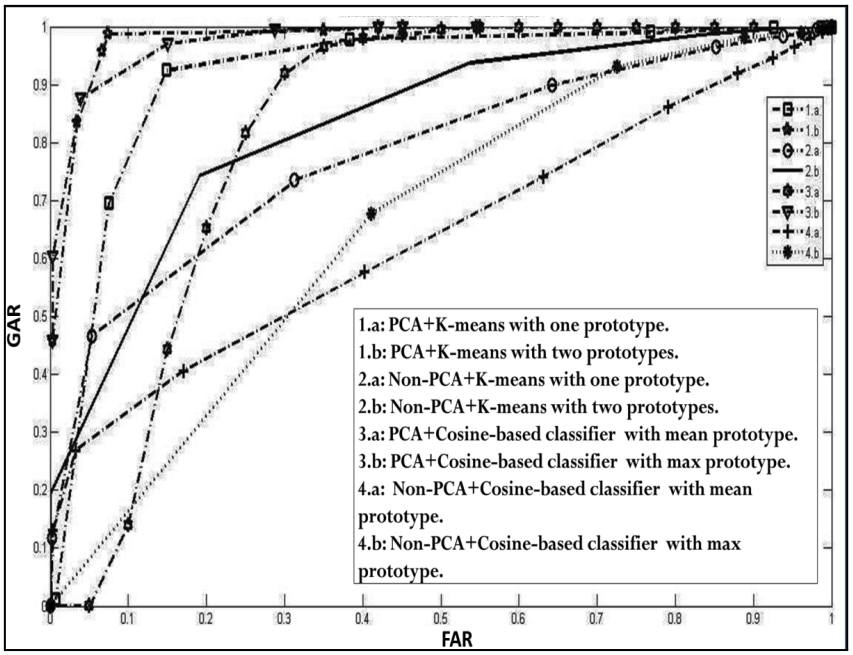


Fig. 16: The ROC Curves of the Proposed Algorithms Over the FG-NET Database

Moreover, the performance of the cosine distance based classifier degraded when the test feature vectors were matched with the mean of training feature vectors (algorithm 4.a) instead of matching the test feature vectors with each training feature vector (algorithm 4.b). Algorithm 1.b was able to achieve an average verification rate (GAR) of 94.31% over the FG-NET database.

The results of evaluating the proposed algorithms over the MORPH face aging database in the verification mode were in line with the verification results reported over the FG-NET database. Again, Algorithm 1.b presented itself as the best performing among the proposed algorithms with an average verification rate of 96.67%. Such results support the conclusions that has been made earlier in the combined effect of PCA subspace and k-means clustering unsupervised classification with multiple prototype in improving the verification performance. The results of evaluating the proposed algorithms over the MORPH face aging database in the verification mode were in line with the verification results reported over the FG-NET database. Again, Algorithm 1.b presented itself as the best performing among the proposed algorithms with an average verification rate of 96.67%. Such results support the conclusions that has been made earlier in the combined effect of PCA subspace and k-means clustering unsupervised classification with multiple prototype in improving the verification performance.

The results of evaluating the proposed algorithms in the closed-set identification mode over the FG-NET and MORPH face aging databases are illustrated in Fig 17 and 18 respectively. The figures depict the precision-recall curves of the proposed algorithms. Varying the reference threshold moved the operating point along the precision-recall curves. Algorithm 1.b was able to achieve the highest rank-1 identification accuracies of 93.20% over the FG-NET database and 95.22% over the MORPH database.

Moreover, the precision-recall curves have shown the superiority of algorithm 1.b over the other proposed algorithms. Table IV illustrates the identification performance of algorithm 1.b in terms of rank-1 identification accuracy, precision, recall, and F-measure.

Table 4: The Close-Set Identification Results

Database	Rank-1 Identification Accuracy	Precision	Recall	F-measure
FG-NET	93.20	0.941	0.939	0.939
MORPH	95.22	0.9612	0.9592	0.960

The results depicted in Table VI show that only the performance of the proposed framework was validated in both verification and identification modes. The rest of the systems were either validated in the verification or identification mode. Moreover, the proposed framework was able to achieve equal error rates of 7.22% over the FG-NET database and 6.51% over the MORPH database. The equal error rates achieved by the proposed framework are the lowest rates among the reported results in the literature review.

Also, the rank-1 identification accuracy reported by the proposed framework exceeds the highest rank-1 identification accuracy reported in the literature review by Park et al.[6]. Except Mahalingam et al. [11], all of the existing age invariant face recognition systems were tested over subsets of the MORPH and FG-NET databases.

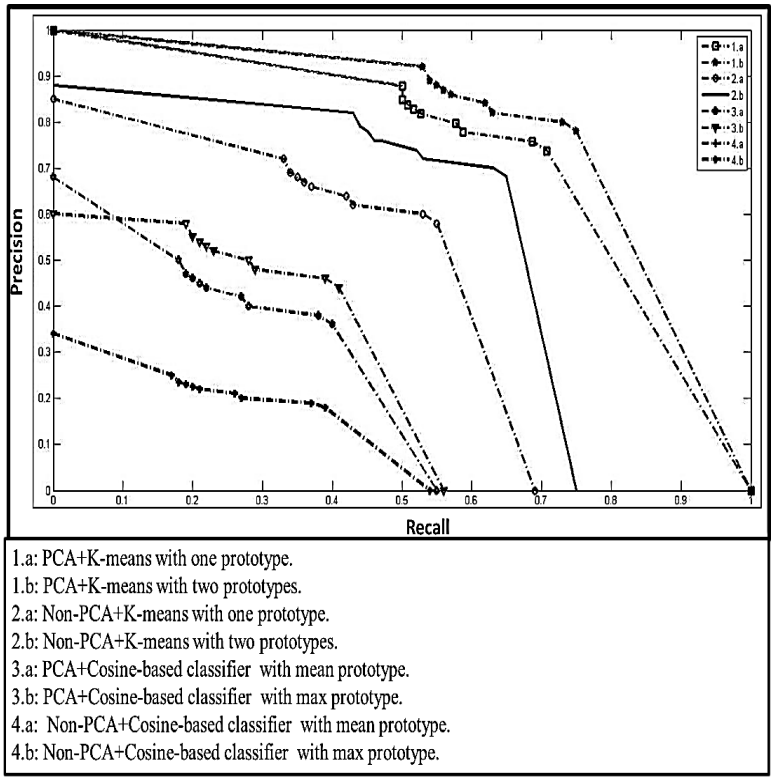
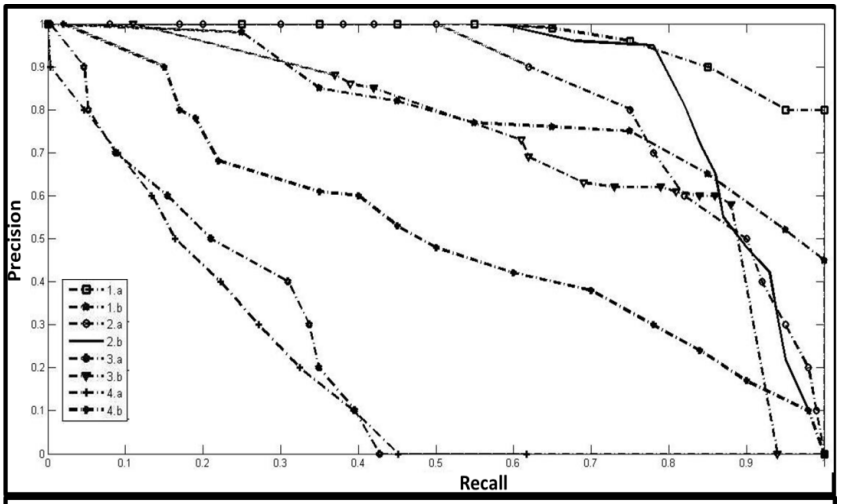


Fig. 17: FG-NET Database Close-Set Identification Results



1.a: PCA+K-means with one prototype.
1.b: PCA+K-means with two prototypes.
2.a: Non-PCA+K-means with one prototype.
2.b: Non-PCA+K-means with two prototypes.
3.a: PCA+Cosine-based classifier with mean prototype.
3.b: PCA+Cosine-based classifier with max prototype.
4.a: Non-PCA+Cosine-based classifier with mean prototype.
4.b: Non-PCA+Cosine-based classifier with max prototype.

Fig. 18: MORPH Database Close-Set Identification Results

The results of the periocular biometrics experiments have shown that it is more effective in terms of verification rate and processing time to use the Periocular region as a biometric trait rather than a full face. Also, the results of the Periocular biometrics experiments have shown that the LBPV and GIST descriptors are the optimal face descriptor to be used for extracting the face features from the Periocular region.

Accordingly, the Periocular region was selected to be adopted as a face biometric trait in the proposed framework. Also, both of the GIST and LBPV descriptors were used for extracting the face features from the Periocular region.

The proposed system is the only system that was validated in both the verification and identification modes over the entire FG-NET and MORPH databases. Using algorithm 1.b, the proposed framework achieved equal error rates of 7.22% and 6.51% over the FG-NET and MORPH database respectively. Moreover, the proposed framework achieved maximum rank-1 identification accuracies of 93.20% and 95.22 over the FG-NET and MORPH database respectively. To the best of the researcher's knowledge such results are the highest results reported in the literature review so far.

Table 6: Comparing the Performance of the Proposed Framework with the Existing Age Invariant Face Recognition Systems Results

Method	Database	EER reported (%)	Rank-1 Identification Accuracy (%)
Ramanathan et al.[5]	Private Database (109, 109)	-	15.00
Geng et al. [7]	FG-NET (10, 10) (only a subset)	-	38.10
Park et al.[6]	FG-NET (82, 82) (only a subset)	-	37.40
	MORPH (612, ,612) (only a subset)	-	66.40

Ling et al.[8], [9]	FG-NET (62, 272) (only a subset)	24.10	-
	MORPH (5060, 20,140)	29.38	-
Mahalingam et al. [11]	FG-NET (82,1002) (entire database)	24.08	-
	MORPH (5060, 20,140)	16.49	-
Lip-Nose Morphometry Amal et al. [44]	FG-NET (82,1002) (entire database)	36	85.89
	MORPH Album2	27.14	78.21
Periocular Region Features Amal et al. [45]	FG-NET (82,1002) (entire database)	26.42	89.68
	MORPH Album2	17.02	93.35
The Proposed Framework Algorithm 1.b: PCA+K-means with two prototypes.	FG-NET (82,1002) (entire database)	7.22	93.20
	MORPH Album2	6.51	95.22
Lip-Nose Morphometry Amal et al. [44]	FG-NET (82,1002) (entire database)	36	85.89
	MORPH Album2	27.14	78.21

VI. CONCLUSION

This research study proposed a novel age-invariant face recognition framework based on two simple yet effective face traits. The proposed approach addresses the face aging problem in a more direct way without relying on a generative aging model. This obviates the need for a training set of subjects that differ only in their age with minimal variations in illumination and pose, which is often a requirement in building a generative aging model. The first trait is a set of anthropometric measurements related to the lip-nose complex. Several physiological studies have proved that the arithmetic range of the lip-nose complex measurements is different for different ethnicities, and thus, can be adopted for ethnicity classification. The same studies have proven that such measurements have a different arithmetic range of different genders within the same ethnicity. Moreover, some of the studies concluded that there is a possibility that the growth rate of such measurements is unique for each individual. The second trait is based on extracting texture features that are robust and rotation invariant in the periocular region using the LBPVar and GIST descriptors.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support provided from the Universiti Teknologi PETRONAS, Malaysia.

REFERENCES

1. A. Lanitis, "A survey of the effects of aging on biometric identity verification," International Journal of Biometrics, vol. 2, pp. 34-52, 2010.
2. J. Suo, X. Chen, S. Shan, and W. Gao, "Learning long term face aging patterns from partially dense aging databases," in Computer Vision, 2009 IEEE 12th International Conference on, 2009, pp. 622-629.
3. N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," in Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, 2006, pp. 387-394.
4. N. Ramanathan and R. Chellappa, "Modeling shape and textural variations in aging faces,"

in Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on, 2008, pp. 1-8.
5. N. Ramanathan and R. Chellappa, "Face verification across age progression," Image Processing, IEEE Transactions on, vol. 15, pp. 3349-3361, 2006.
6. U. Park, "Face Recognition: face in video, age invariance, and facial marks," Michigan State University, 2009.
7. X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 29, pp. 2234-2240, 2007.
8. H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "A study of face recognition as people age," in Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, 2007, pp. 1-8.
9. H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "Face verification across age progression using discriminative methods," Information Forensics and Security, IEEE Transactions on, vol. 5, pp. 82-91, 2010.
10. M. Bereta, P. Karczmarek, W. Pedrycz, and M. Reformat, "Local descriptors in application to the aging problem in face recognition," Pattern Recognition, vol. 46, pp. 2634-2646, 2013.
11. G. Mahalingam and C. Kambhamettu, "Face verification of age separated images under the influence of internal and external factors," Image and Vision Computing, vol. 30, pp. 1052-1061, 2012.
12. P. Viola and M. J. Jones, "Robust real-time face detection," International journal of computer vision, vol. 57, pp. 137-154, 2004.
13. N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," Machine Learning, vol. 59, pp. 161-205, 2005.
14. D. DeCarlo, D. Metaxas, and M. Stone, "An anthropometric face model using variational techniques," in Proceedings of the 25th annual conference on Computer graphics and interactive techniques, 1998, pp. 67-74.
15. D. W. S. PhilipMiller and R. J. A. Ross, "On the fusion of periocular and iris biometrics in non-ideal imagery," 2010.
16. F. Juefei-Xu, K. Luu, M. Savvides, T. D. Bui, and C. Y. Suen, "Investigating age invariant face recognition based on periocular biometrics," in Biometrics (IJCB), 2011 International Joint Conference on, 2011, pp. 1-7.
17. U. Park, R. Jillela, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum," Information Forensics and Security, IEEE Transactions on, vol. 6, pp. 96-106, 2011.
18. U. Park, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum: A feasibility study," in Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on, 2009, pp. 1-6.
19. X.-R. Pu, Y. Zhou, and R.-Y. Zhou, "Face Recognition on Partial and Holistic LBP Features," Threshold, vol. 1, p. 0, 2012.
20. F. Zhou, J. Brandt, and Z. Lin, "Exemplar-based graph matching for robust facial landmark localization," in Computer Vision (ICCV), 2013 IEEE International Conference on, 2013, pp. 1025-1032.
21. A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," International journal of computer vision, vol. 42, pp. 145-175, 2001.
22. A. K. Jain and U. Park, "Facial marks: Soft biometric for face recognition," in Image Processing (ICIP), 2009 16th IEEE International Conference on, 2009, pp. 37-40.
23. H. K. Ekenel and R. Stiefelhagen, "Generic versus salient region-based partitioning for local appearance face recognition," in Advances in Biometrics, ed: Springer, 2009, pp. 367-375.
24. P. E. Miller, A. W. Rawls, S. J. Pundlik, and D. L. Woodard, "Personal identification using periocular skin texture," in Proceedings of the 2010 ACM Symposium on Applied Computing, 2010, pp. 1496-1500.
25. Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching," Pattern recognition, vol. 43, pp. 706-719, 2010.

26. T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, pp. 2037-2041, 2006.
27. J. Kittler, M. Hatef, R. P. Duin, and J. Matas, "On combining classifiers," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 20, pp. 226-239, 1998.
28. A. Rattani, D. R. Kisku, M. Bicego, and M. Tistarelli, "Feature level fusion of face and fingerprint biometrics," in Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on, 2007, pp. 1-6.
29. H. Cevikalp, M. Neamtu, and M. Wilkes, "Discriminative common vector method with kernels," Neural Networks, IEEE Transactions on, vol. 17, pp. 1550-1565, 2006.
30. S. Mitra and T. Acharya, "Gesture recognition: A survey," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 37, pp. 311-324, 2007.
31. A. K. Jain and A. Ross, "Multibiometric systems," Communications of the ACM, vol. 47, pp. 34-40, 2004.
32. A. K. Jain, "Data clustering: 50 years beyond K-means," Pattern recognition letters, vol. 31, pp. 651-666, 2010.
33. J. MacQueen, "Some methods for classification and analysis of multivariate observations," in Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1967, pp. 281-297.
34. J. A. Sáez, J. Derrac, J. Luengo, and F. Herrera, "Improving the behavior of the nearest neighbor classifier against noisy data with feature weighting schemes," in Hybrid Artificial Intelligence Systems, ed: Springer, 2014, pp. 597-606.
35. G. Pirlo and D. Impedovo, "Cosine similarity for analysis and verification of static signatures," IET biometrics, vol. 2, pp. 151-158, 2013.
36. J. T. Foote, "Content-based retrieval of music and audio," in Voice, Video, and Data Communications, 1997, pp. 138-147.
37. B. Efron and G. Gong, "A leisurely look at the bootstrap, the jackknife, and cross-validation," The American Statistician, vol. 37, pp. 36-48, 1983.
38. T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," IEEE Transactions on pattern analysis and machine intelligence, vol. 23, pp. 681-685, 2001.
39. I. Matthews and S. Baker, "Active appearance models revisited," International Journal of Computer Vision, vol. 60, pp. 135-164, 2004.
40. G. Mahalingam and C. Kambhamettu, "Age invariant face recognition using graph matching," in Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on, 2010, pp. 1-7.
41. T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in Computer Vision—ECCV'98, ed: Springer, 1998, pp. 484-498.
42. S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection," Biomedical Engineering, IEEE Transactions on, vol. 55, pp. 512-518, 2008.
43. X. Tan and B. Triggs, "Fusing Gabor and LBP feature sets for kernel-based face recognition," in Analysis and Modeling of Faces and Gestures, ed: Springer, 2007, pp. 235-249.
44. A. S. O. Ali, V. S. Asirvadam, A. S. Malik, M. M. Eltoukhy, and A. Aziz, "Age-Invariant Face Recognition Using Triangle Geometric Features," International Journal of Pattern Recognition and Artificial Intelligence, p. 1556006, 2012.
45. A. S. O. Ali, V. Sagayan, A. M. Saeed, H. Ameen, and A. Aziz, "Age-invariant face recognition system using combined shape and texture features," IET Biometrics, 2015.
46. Dr. Amal S. Osman Ali received her BSc in Communication Engineering from Sudan University of Science and Technology Khartoum, Sudan in 2007, and MSc in Computer Architectures and Networks from University of Khartoum, Sudan in 2009. She is currently working as PhD scholar in Department of Electrical and Electronics Engineering at University Teknologi

PETRONAS since Jan 2012. Her research focuses on developing an Age-Invariant Face Recognition system to reduce the effects of facial aging on face recognition performance.

47. AP. Dr. Aamir S. Malik received his MSc in Information and Communication and PhD in Information and Mechatronics from Gwangju Institute of Science and Technology, South Korea. He has more than 15 years of research experience and is currently Associate Professor at the Department of Electrical Engineering and Director of Biomedical Technology group at Universiti Teknologi PETRONAS. His research interests include biomedical signal & image processing, visual surveillance, remote sensing and brain sciences.
48. *Dr. Azrina Aziz:* is with the Universiti Teknologi PETRONAS Bandar Seri Iskandar, 31720 Tronoh, Perak, Malaysia (phone: 605-368-7881; fax: 605-365-7443; e-mail: azrina_aaziz@petronas.com.my).
49. AP. Dr. Vijanth S. Asirvadam studied at the University of Putra, Malaysia for the Bachelor Science, BSc. (Hon) majoring in Statistics and graduated in April 1997. He received his Master's Science, MSc. degree in Engineering Computation with a Distinction in December 1998. He took employments as a system administrator at the Multimedia University (MMU) Malaysia prior joining the Intelligent Systems and Control Research Group at Queen's University Belfast in November 1999 where he completed his Doctorate (Ph.D.) in March 2003, researching on Online and Constructive Neural Learning methods. his Master's Science, MSc. degree in Engineering Computation with a Distinction in December 1998. He took employments as a system administrator at the Multimedia University (MMU) Malaysia prior joining the Intelligent Systems and Control Research Group at Queen's University Belfast in November 1999 where he completed his Doctorate (Ph.D.) in March 2003, researching on Online and Constructive Neural Learning methods.