# A Geometrical Approach for Age-Invariant Face Recognition

Amal Seralkhatem Osman Ali[2], Vijanth Sagayan a/l Asirvadam[1], Aamir Saeed Malik[1], and Azrina Aziz[1]

[1] Centre of Intelligent Signals and Imaging Research
Department of Electric and Electronic Engineering
[2] Universiti Teknologi PETRONAS
31750 Tronoh, Perak, Malaysia
Amal61@gmail.com,
{vijanth_sagayan,aamir_saeed,azrina.aziz}@petronas.com.my

**Abstract.** Human faces undergo considerable amounts of variations with aging. While face recognition systems have proven to be sensitive to factors such as illumination and pose, their sensitivity to facial aging effects is yet to be studied. The FRVT (Face Recognition Vendor Test) report estimated a decrease in performance by approximately 5% for each year of age difference. Therefore, the development of age-invariant capability remains an important issue for robust face recognition. This research study proposed a geometrical model based on multiple triangular features for the purpose of handling the challenge of face age variations that affect the process of face recognition. The system is aimed to serve in real time applications where the test images are usually taken in random scales that may not be of the same scale as the probe image, along with orientation, lighting ,illumination, and pose variations. Multiple mathematical equations were developed and used in the process of forming distinct subject clusters. These clusters hold the results of applying the developed mathematical models over the FGNET face aging database. The system was able to achieve a maximum classification accuracy of above 99% when the system was tested over the entire FGNET database.

**Keywords:** frvt, age-invariant, geometrical model, triangular features, similarity proportion ratios, clustering, fgnet.

## 1 Introduction

Face recognition is a type of automated biometric identification method that recognizes individuals based on their facial features as basic elements of distinction. The research on face recognition has been dynamically going on in the recent years because face recognition is involved in many fields and disciplines such as access control, surveillance and security, criminal identification and digital library.

Automatic face detection and recognition have been a challenging problem in the field of computer vision for many years. Though humans accomplish the task in an

H. Badioze Zaman et al. (Eds.): IVIC 2013, LNCS 8237, pp. 81–96, 2013.

easy manner, the underlying computations within the human visual system are of remarkable complexity. The apparently insignificant task of finding and recognizing faces is the result of millions of years of regression and we are far from fully understanding how the brain performs it. Moreover, the capability to find faces visually in a scene and recognize them is critical for humans in their everyday events. Accordingly, the automation of this task would be beneficial for several applications including security, surveillance, gaze-based control, affective computing, speech recognition assistance, video compression and animation. Though, to date, no comprehensive solution has been anticipated that allows the automatic recognition of faces in real (un-affected) images [1]. In last decade, chief progresses occurred in the field of face recognition, with numerous systems capable of maintaining recognition rates superior to 90%. However real-world scenarios remain a challenge, because face acquisition procedure can experience a wide range of variations. Throughout a crime investigation, the community security agencies regularly need to match a probe image with registered database images, which may have major difference of facial features due to age deviations. Several efforts have been made to tackle this problem. Ling et al. [2] studied the aging effect on face recognition, O'Toole et al. [3] proposed a standard facial caricaturing algorithm using 3D face model, Ramanathan et al. [4] proposed a Bayesian age-difference classifier to be employed in applications such as passport renewal. These proposed techniques try to solve the problem by simulating the aging models; however, they are still far from hands-on use.

Unlike these complicated modelling methods, our system aims to perform a fast and robust aging face recognition based on a combination of geometrical and mathematical modelling. In this research study our goal is to develop a geometrical model that is age invariant. In our work we have explored the approach of using a mathematically developed geometrical model for maintaining the degree of similarity between six triangular features to address the problem of face recognition under age variations. The system to be developed is intended to operate in real time environment such as surveillance systems.

The remainder of this paper is organized as follows: Section 2 and section 3 represent the feature selection methods and the classifiers used during the experiments part. Section 4 introduces the proposed face recognition geometrical model where we define the mathematical relationships between our proposed triangular features, and our tendency in constructing the systems' facial features vectors. The results and discussion of experiments are presented in Section 5 and section 6 correspondingly. This is followed by conclusions in Section 7.

## 2 Feature Selection Methods

### 2.1 Correlation Feature Selection

The Correlation Feature Selection (CFS) [5] measure evaluates subsets of features on the basis of the following hypothesis: "Good feature subsets contain features highly correlated with the classification, yet uncorrelated to each other". The following equation gives the merit of a feature subset S consisting of k features:

$$\mathrm{Merit}_{Sk} = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}} \tag{1}$$

Here, $\overline{r_{cf}}$ is the average value of all feature-classification correlations, and $\overline{r_{ff}}$ is the average value of all feature-feature correlations. The CFS criterion is defined as follows:

$$CFS = \max_{S_k}\left[\frac{r_{cf1} + r_{cf2} + \cdots + r_{cfk}}{\sqrt{k + 2(r_{f1f2} + \cdots + r_{fifj} + \cdots + r_{fkf1}}}\right] \tag{2}$$

The $r_{cfi}$ and $r_{fifj}$ variables are referred to as correlations. Let xi be the set membership indicator function for feature fi; then the above equation can be rewritten as an optimization problem:

$$\mathrm{CFS} = \max_{x\in\{0,1\}^n}\left[\frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\sum_{i=1}^{n} x_i + \sum_{i\neq j} 2b_{ij}x_i x_j}\right] \tag{3}$$

## 2.2 Relief Attribute Evaluation Method

Relief-F is a feature selection strategy that chooses instances randomly, and changed the weights of the feature relevance based on the nearest neighbor. By its merits, Relief-F is one of the most successful strategies in feature selection. The nearest neighbor from the same class is a hit H, and from different class a miss, M( C) of class C. At the end W[f] is divided by m to get the average evaluation in [-1,1].

$$\mathrm{W[f]= W[f]\text{-}diff(f, E1, H)} + \sum_{C\neq \mathrm{class}(E_1)} P(C) \times \mathrm{diff}(f, E_1, M(C)) \tag{4}$$

The $\mathrm{diff}(f, E_1, E_2)$ function calculates the grade in which the values of feature f are different in examples E1 and E2.

## 2.3 Symmetrical Uncertainty Feature Selection Method

The algorithms find weights of discrete attributes basing on their correlation with continuous class attribute. The algorithm uses an information theoretic measure called symmetric uncertainty in order to evaluate the worth of constructed solutions. There are a number of benefits of this measure i.e. it is symmetric in nature therefore SU(i,j) is same as that of SU(j,i) hence it reduces the number of comparisons required, where i and j represent two independent variables, it is not influenced by multi-valued attributes as that is in the case of information gain, and its values are normalized. Following is the equation for symmetric uncertainty.

$$\mathrm{SU(X,Y)} = 2 \times \left[IG \times \frac{X\backslash Y}{H(X)+H(Y)}\right] \tag{5}$$

Where IG(X|Y) is the information gain of feature X, that is an independent attribute and Y is the class attribute. H(X) is the entropy of feature X and H(Y) is the entropy of feature Y. Information gain has a desired property, i.e. it is symmetric. The amount of Information given by a feature Y about another feature X is effectively the same as that of the information given of feature X and the feature Y.

# 3 Classification Algorithms

## 3.1 Naïve Bays Classifier

A Naïve Bays classifier [8] is a simple probabilistic classifier based on applying Bays rule. Bays theorem provides a means of calculating the posterior probability, P (b|x), from P (b), P(x), and P (x|b). Naïve Bays classifier assumes that the effect of the value of a predictor (x) on a particular class (c) is independent of the values of other predictors. This hypothesis is called class conditional independence.

$$P(x|b)= (P(x|b)P(b))/P(x) \tag{6}$$

$$P(b|x)= P(x_1|b)x\ P(x_2|b)x\ldots..x\ P(x_n|b)Xp(b) \tag{7}$$

- *P*(*b*|*x*) is the posterior probability of class (target) given predictor (attribute).
- *P*(*b*) is the prior probability of class.
- *P*(*x*|*b*) is the likelihood which is the probability of predictor given class.
- *P(x)* is the prior probability of predictor.

## 3.2 Support Vector Machine (SVM)

In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output, making it a non-probabilistic binary linear classifier. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

In addition to performing linear classification, SVMs can efficiently perform non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces.

## 3.3 K-Means Clustering

In data mining, k-means clustering is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. This results in a partitioning of the data space into Voronoi cells.

The problem is computationally difficult (NP-hard); however, there are efficient heuristic algorithms that are commonly employed and converge quickly to a local optimum. These are usually similar to the expectation-maximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both algorithms. Additionally, they both use cluster centers to model the data; however, k-means clustering tends to find clusters of comparable spatial extent, while the expectation-maximization mechanism allows clusters to have different shapes.

### 3.4 K-Nearest Neighbors Classifier

In pattern recognition, the k-nearest neighbor algorithm (k-NN) is a non-parametric method for classifying objects based on closest training examples in the feature space. k-NN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbor.

The same method can be used for regression, by simply assigning the property value for the object to be the average of the values of its k nearest neighbors. It can be useful to weight the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. (A common weighting scheme is to give each neighbor a weight of 1/d, where d is the distance to the neighbor. This scheme is a generalization of linear interpolation.)

The neighbors are taken from a set of objects for which the correct classification (or, in the case of regression, the value of the property) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. The k-nearest neighbor algorithm is sensitive to the local structure of the data. Nearest neighbor rules in effect implicitly compute the decision boundary. It is also possible to compute the decision boundary explicitly, and to do so efficiently, so that the computational complexity is a function of the boundary complexity.

## 4 Proposed Geometrical Model

The proposed system decomposed multiple stages. Face detection is the first stage at the beginning of each face recognition system. In our system a commercial version of the conventional Viola and Jones face detector [12] is employed to detect and crop the face area that contains the main features (Eyes, Mouth, Nose, and chine). Viola and Jones detector is robust and effective in real time applications. After detecting the face area twelve facial features points are to be localized in order to extract six different triangular areas around the main facial features. Following the parameters of the triangular features i.e. (areas and perimeters) are calculated. Then those parameters are passed to a number of equations to create features vectors for the sample image. In the following those stages are illustrated in details.

### 4.1 Facial Features Points Localization and Triangular Features Detection

Craniofacial anthropometry which is the science that involves measurement of the skull and face, citrine landmarks and measurements are known to identify human facial characters and growth pattern. In our study we consider twelve of these land marks mostly the ones that form the circumscription of the main facial features. Those facial feature points are normally localized using Active Appearance Model (AAM) [13] which designates 64 distinctive facial points. In our model the AAM is reduced to 12 facial features points using the algorithm proposed in [13]. Craniofacial anthropometry refers to those facial features points with scientific notation to discriminate between them as follows:

- **En (endocanthion)**: the inner corner of the eye fissure where the eyelids meet. In our model these points are given the numbers 6 and 8.
- **G (glabella)**: the most prominent point in the median sagittal plane between the supraorbital. In our model these point is given the number 9.
- **ridges.ex (exocanthion)**: the outer corner of the eye fissure where the eyelids meet. In our model these points are given the numbers 5 and 7.
- **Gn (gnathion)**: in the midline, the lowest point on the lower border of the chin.

In our model this point is given the number 2.

Following six triangles are formed between the facial points and they are given the notation triangle$_1$ through triangle6, as illustrated in Fig. (2.a) through (2.f)

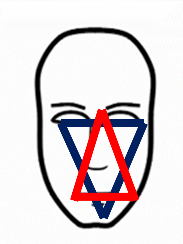

**Fig. 1. (a)** Triangle$_1$ and Triangle$_5$ (*dotted*

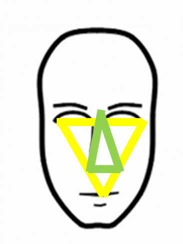

**Fig. 1. (b)** Triangle$_3$ and Triangle$_4$

**Fig. 1. (c)** Triangle$_2$ and Triangle$_6$

### 4.2 Calculating the Triangular Features Parameters

After localizing the facial features points, the system will gain knowledge of the triangular vertices coordinates. After that Euclidean distances between each triangle coordinates will be calculated, which will enable the system to calculate perimeters and areas of each triangular feature. Those parameters (areas and perimeters are given the notation $A$ and $P$ for areas and perimeters successively followed by a subscript representing the triangle designation. For example ($Ai$, $Pi$) represent the area and perimeter of triangle number one. Finally those parameters are used as inputs to some mathematical equations which will be discussed next, to form the features vectors for each sample image.

### 4.3 Deriving the Mathematical Model

In the geometry science it is known that Triangles are similar if they have the same shape, but not necessarily the same size [14]. This scientific fact inspired us to draw mathematical relationships between the six triangular features extracted during the previous stage. The Human population reached 7 Billion people around the world and thus, it is impractical to use a one-to-one comparison process for the purpose of face recognition using the measurements of our triangular features. As a different approach we were able to make use of the proportional ratio between the different triangles representing the facial features which led to fifteen different mathematical equations representing the degree of similarity between each two triangles. Based on the aforementioned geometrical theory regarding the similarity of triangles, any two triangles are considered similar even if they are of different sizes if the following mathematical relationship represented by "Eq. (1)" is satisfied:

$$A_i/\,A_j = p_j^2/p_i^2. \tag{8}$$

Where A, and P represent triangles areas and perimeters successively, i and j are designations of the two triangles subject of the mathematical relationship. Eq. (8) is used to drive what is called triangles similarity proportion, which is a measurement of degree of similarity between two triangles, and it is represented by "Eq. (2)". TSP represents the triangles similarity proportion relationship.

$$TSP{=}A_i x p_j^2/\,A_j x p_i^2. \tag{9}$$

The statistical analysis of the data collected in term of triangular features areas and perimeters had shown clearly that there is no significant difference between these measurements of different individuals. As a different approach we were able to make use of the similarity proportional ratio between the different triangles representing the facial features which led to fifteen different mathematical equations representing the degree of similarity between each two triangles. Those equation were derived using equation (9) by simply applying the formula between each two triangles, and substituting subscripts i and j by the designations of the two triangles. "Eq. (10)" through, "Eq. (24)" represents the fifteen relationships between the six triangular features as listed in Table 1:

**Table 1.** Similarity Proportions Relationships between the Triangular Features

| Eq. 10, 13, 16, 19, 22 | Eq. 11, 14, 17, 20,23 | Eq. 12, 15, 18, 21, 24 |
|---|---|---|
| $(T_1,T_2)=(A_1*P_2^2/A_2*P_1^2)$ | $(T_1,T_3)=(A_1*P_3^2/A_3*P_1^2)$ | $(T_1,T_4)=(A_1*P_4^2/A_4*P_1^2)$ |
| $(T_1,T_5)=(A_1*P_5^2/A_5*P_1^2)$ | $(T_1,T_6)=(A_1*P_6^2/A_6*P_1^2)$ | $(T_2,T_3)=(A_2*P_3^2/A_3*P_2^2)$ |
| $(T_2,T_4)=(A_2*P_4^2/A_4*P_2^2)$ | $(T_2,T_5)=(A_2*P_5^2/A_5*P_2^2)$ | $(T_2,T_6)=(A_2*P_6^2/A_6*P_2^2)$ |
| $(T_3,T_4)=(A_3*P_4^2/A_4*P_3^2)$ | $(T_3,T_5)=(A_3*P_5^2/A_5*P_3^2)$ | $(T_3,T_6)=(A_3*P_6^2/A_6*P_3^2)$ |
| $(T_4,T_5)=(A_4*P_5^2/A_5*P_4^2)$ | $(T_4,T_6)=(A_4*P_6^2/A_6*P_4^2)$ | $(T_5,T_6)=(A_5*P_6^2/A_6*P_5^2)$ |

For each sample image enrolled in the system those fifteen relationships will be calculated and stored in a vector which will be considered as a feature vector of this specific sample image. When multiple sample images are related to the same subject, the feature vectors of these sample images will be stored in a matrix to form a class for each subject.

# 5 Experiments

We performed our experiments on a public aging database FG-NET [15] containing 1,002 high resolution color or gray scale face images of 82 subjects from multiple races with large variation of lighting, expression, and pose. The image size is approximately 400 x 500 in pixels. The age range is from 0 to 69 years (on average, 12 images per subject). The FG-NET database was divided into three subsets as follows:

1. FGnet-8 consists of all the data collected at ages between 0 and 8. It includes 290 facial images from 74 subjects, among which 580 intra-person pairs and 6000 inter-person pairs are randomly generated for verification.
2. FGnet-18 consists of all the data collected at ages between 8 and 18. It includes 311 facial images from 79 subjects, among which 577 intra-person pairs and 6000 inter-person pairs are randomly generated for verification.
3. FGnet-adult consists of all the data collected at ages 18 or above and roughly frontal. It includes 272 images from 62 subjects, among which 665 intrapersonal pairs and about 6000 intra-personal pairs are randomly generated for verification.

The aim of using the aforementioned protocol for dividing the FGNET database was to determine which features contribute more in discriminating between the classes for each age range. To accomplish this goal a number of feature ranking and selection methods namely: The Correlation Feature Selection (CFS) method, ReliefF Attribute Evaluation method, Symmetrical Uncertainty feature selection method were employed to select a subset of the most significant features among the fifteenth extracted features for each age spam. After that a number of classifiers were used to evaluate the performance of the system namely: K-means KNN (K-Nearest Niebuhr), Random Forest, Naïve Bayes, and Bayes Net classifiers; when the system was tested on each of the FGNET subsets. In the next section the resultant features subsets when each of the aforementioned feature selection methods are used are illustrated, and classification results are illustrated to evaluate performance of the system .

# 6 Results and Discussion

## 6.1 Classification Results

Classification results of the developed facial geometrical system, when the system was tested on the entire FGNET database and each of the FGNET subsets separately are illustrated in terms of classification accuracy, and error rate. **Table 2** through **Table 5** illustrate the classification result of testing the system over the FGNET-8, FGNET-18, FGNET-Adult, and the entire FGNET database successively.

**Table 2.** FGNET-8 Subset Classification Results

| Classifier | Accuracy (%) | Error rate (%) |
|---|---|---|
| KNN | 96.0563 | 3.9437 |
| Naïve Bayes | 21.69 | 78.3099 |
| K-means | 5.6338 | 94.3662 |
| SVM | 3.3803 | 96.6197 |

**Table 3.** FGNET-18 Subset Classification Results

| Classifier | Accuracy (%) | Error rate (%) |
|---|---|---|
| KNN | 89.1008 | 10.8992 |
| Naïve Bayes | 24.2507 | 75.7493 |
| SVM | 4.3597 | 95.6403 |
| K-means | 5.1771 | 94.8229 |

**Table 4.** FGNET-Adult Subset Classification Results

| Classifier | Accuracy (%) | Error rate (%) |
|---|---|---|
| KNN | 96.0563 | 3.9437 |
| Naïve Bayes | 21.6901 | 78.3099 |
| K-means | 5.6338 | 94.3662 |
| SVM | 3.3803 | 96.6197 |

**Table 5.** FGNET-Adult Subset Classification Results

| Classifier | Accuracy (%) | Error rate (%) |
|---|---|---|
| KNN | 99.1701 | 0.8299 |
| K-means | 4.1494 | 95.8506 |
| Naïve Bayes | 9.5436 | 90.4564 |
| SVM | 17.2891 | 82.7109 |

It can be seen from the classification results that the best accuracy was achieved when the system was tested over the entire FGNET database using KNN classifier with a maximum accuracy of over 99%. Performance of the KNN classifier was relatively high for all FGNET subsets, which is due to the ability of the KNN classifier of handling large number of classes. On contrast the other three classifiers have shown poor performance over all of the FGNET subsets.

## 6.2 Feature Selection Results

### 6.2.1 FGNET Complete Set Feature Selection Results

Fig. 1 shows the ranking of the fifteen developed features. Features ranking was performed using ReliefF Attribute Evaluator feature selection method. The top five

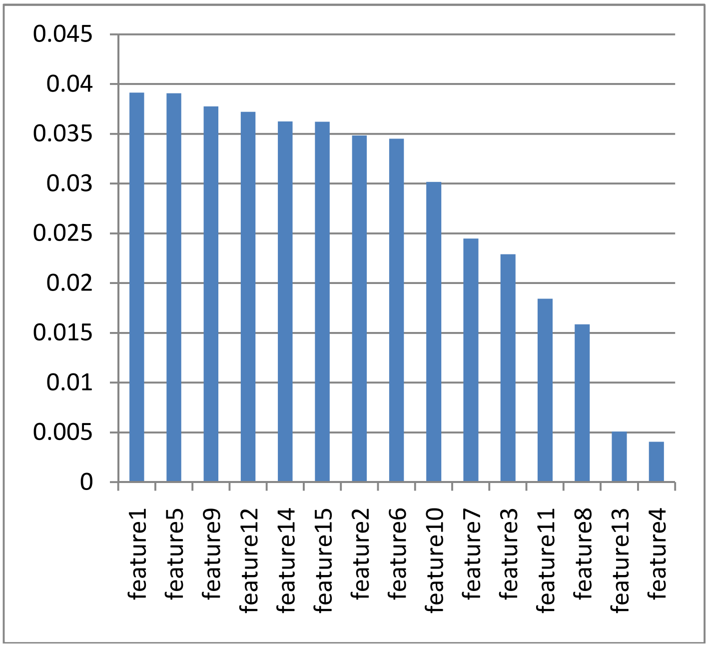


**Fig. 1.** ReliefF Attribute Evaluator FGNET

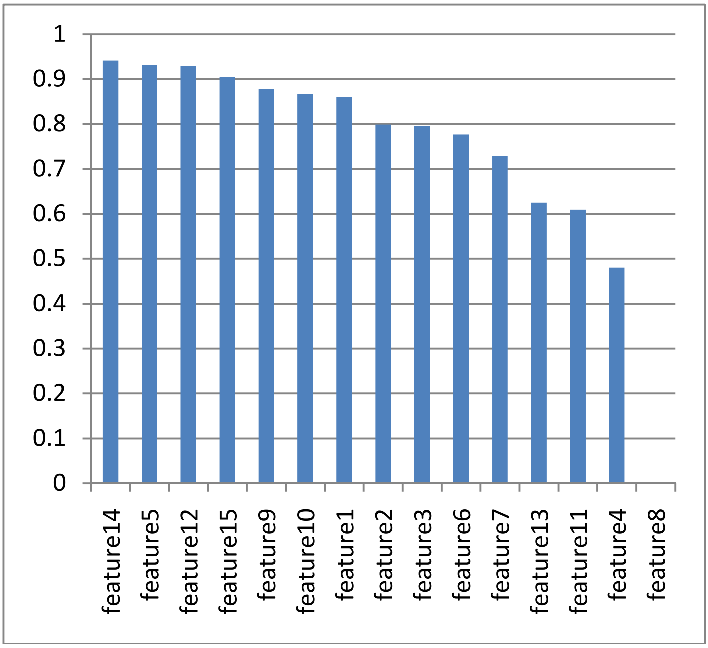


**Fig. 2.** Symmetrical Uncertainty Attribute Evaluator FGNET

are successively: Feature1, Feature5, Feature9, Feature12, Feature14, and the most important feature is feaure1 which is represented by equation (10). Symmetrical Uncertainty Attribute Evaluator feature selection method on the other hand produced different top features set where the most important feature is feature14 represented by equation (23), and the top five features are feature14, feature5, feature12, feature15, and feature9 as illustrated in Fig. 2 .

#### 6.2.2 FGNET-8 Feature Selection Results

Features ranking and selection results were comparatively different when the system was tested on the FGNET-8 subset than the results when the system was tested on the entire FGNET database in particular the results achieved by ReliefF Attribute Evaluator feature selection method. As it can be seen in Fig. 3 and Fig. 4. When using ReliefF Attribute Evaluator feature selection method the top five are successively: Feature12, Feature14, Feature5, Feature15, Feature2, and the most

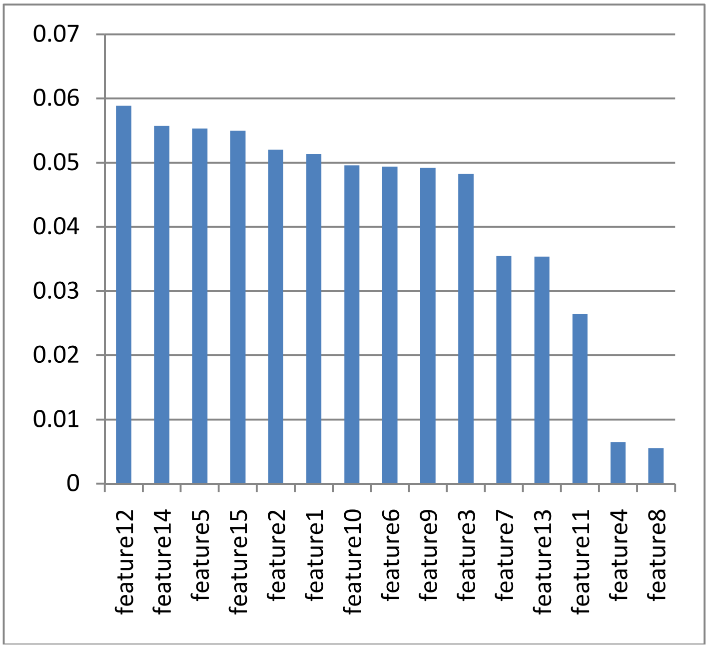


**Fig. 3.** ReliefF Attribute Evaluator FGNET-8

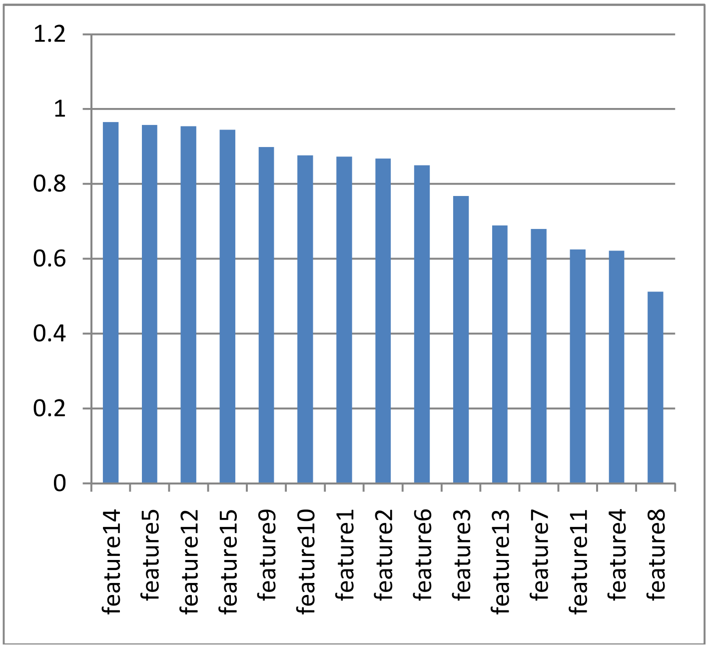


**Fig. 4.** Symmetrical Uncertainty Attribute Evaluator FGNET-8

important feature is feaure12 which is represented by equation (21) as illustrated in figure 3. Symmetrical Uncertainty Attribute Evaluator feature selection method on the other hand produced different top features set where the most important feature is feature14 represented by equation (23), and the top five features are feature14, feature5, feature12, feature15, and feature9 as illustrated in Fig. 4 .

#### 6.2.3 FGNET-18 Feature Selection Results

Fig. 5 shows the ranking of the fifteen developed features when the system was tested over the FGNET-18 subset. Features ranking was performed using ReliefF Attribute Evaluator feature selection method. The top five are successively: Feature1, Feature12, Feature6, Feature15, Feature2, and the most important feature is feaure1 which is represented by equation (10). Symmetrical Uncertainty Attribute Evaluator feature selection method on the other hand produced different top features set where the most important feature is feature14 represented by equation (23), and the top five features are feature14, feature5, feature12, feature15, and feature9 as illustrated in Fig. 6 .

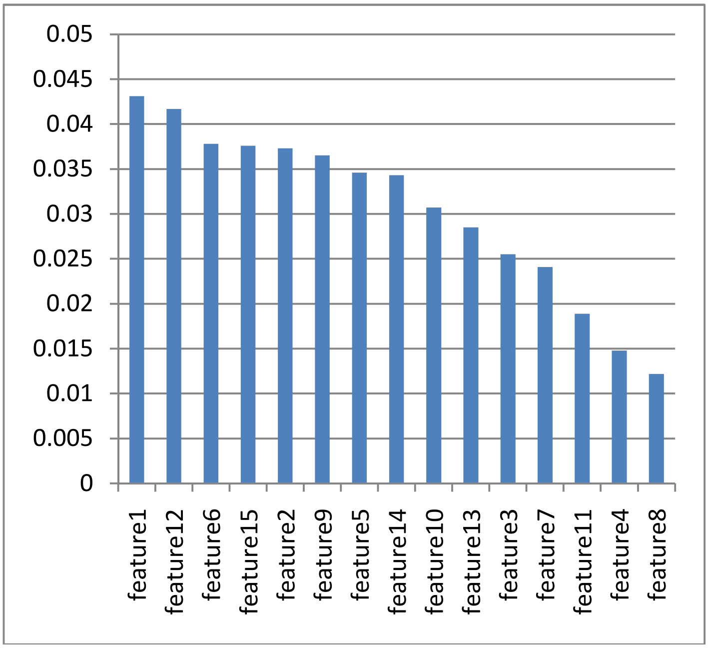


**Fig. 5.** ReliefF Attribute Evaluator FGNET-18

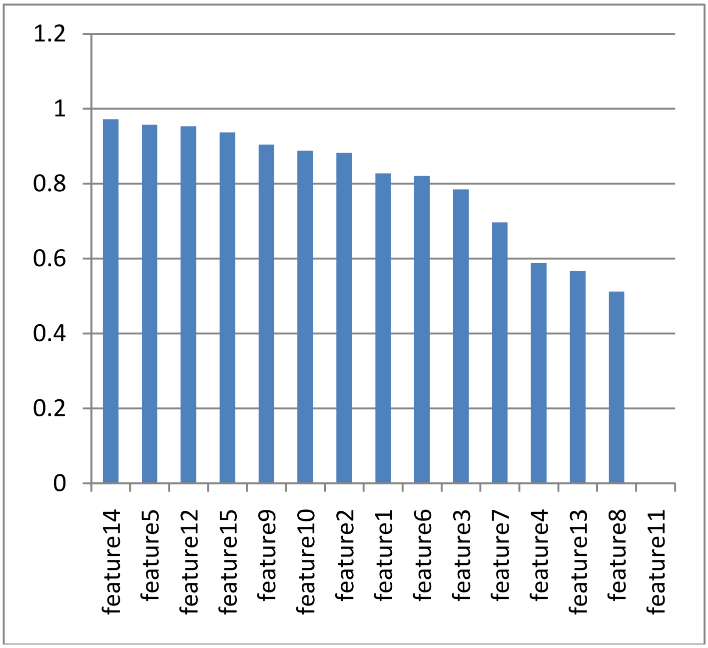


**Fig. 6.** Symmetrical Uncertainty Attribute Evaluator FGNET-18

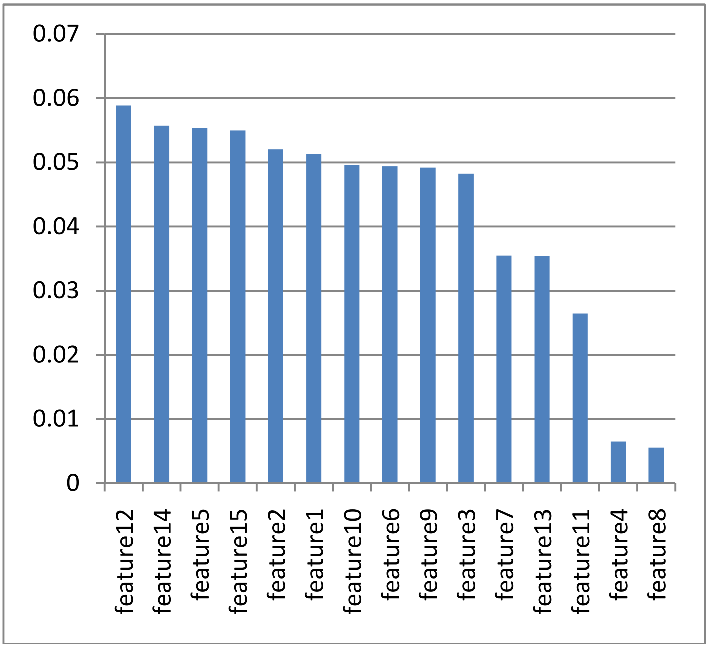


**Fig. 7.** ReliefF Attribute Evaluator FGNET-Adult

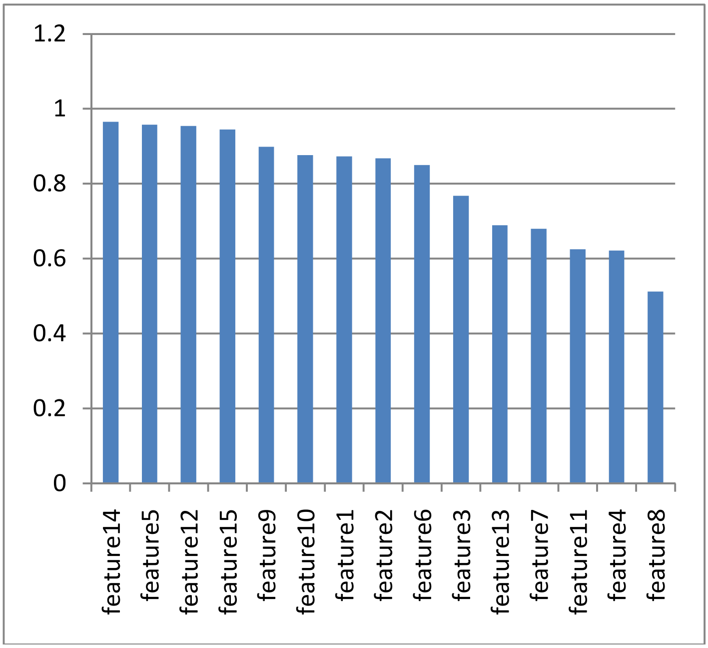


**Fig. 8.** Symmetrical Uncertainty Attribute Evaluator FGNET-Adult

#### 6.2.3 FGNET-Adult Feature Selection Results

As it can be seen in Fig. 7 and Fig. 8. When using ReliefF Attribute Evaluator feature selection method the top five are successively: Feature12, Feature14, Feature5, Feature15, Feature2, and the most important feature is feaure12 which is represented by equation (21) as illustrated in figure 7. Symmetrical Uncertainty Attribute Evaluator feature selection method on the other hand produced different top features set where the most important feature is feature14 represented by equation (23), and the top five features are feature14, feature5, feature12, feature15, and feature9 as illustrated in Fig. 8.

## 7 Conclusion and Future Work

This research study proposed new geometrical features that are formed by connecting some of the facial features points defined in the anthropometric science. The main goal was to develop mathematical relationships among triangular features to accommodate for the aging variations conditions that may affect any face recognition system. The performance of the system was evaluated mainly in term of classification accuracy, and the maximum classification accuracy was reported when the KNN classifier was used to test the system over the entire FGNET database. Also a number of feature selection and ranking methods were used to study the importance of features in different age spans . In our future work we are planning to test the performance of the system when different feature selection methods are used in conjunction with multiple classification methods.

**Acknowledgments.** The authors gratefully acknowledge the support provided by Universiti Teknologi PETRONAS, Malaysia. Also the authors wish to express their gratitude to the Dean of the Centre for Graduate Studies Associate Prof Dr. Mohd Fadzil Hassan at the Universiti Teknologi PETRONAS for participating in the excremental part of this research work and for his continuous support to the research innovation in at the Universiti Teknologi PETRONAS.

## References

1. Kachare, N.B., Inamdar, V.S.: Survey of Face Recognition Techniques. International Journal of Computer Applications (0975 - 8887) 1(19) (2010)
2. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.W.: Study of Face Recognition as People Age. In: Proc. 11th Int'l Conf. Computer Vision, pp. 1–8 (2007)
3. O'Toole, A., Vetter, T., Volz, H., Salter, M.: Three-dimensional caricatures of human heads: distinctiveness and the perception of facial age. Perception 26, 719–732 (1997)
4. Ramanathan, N., Chellappa, R.: Face Verification across Age Progression. IEEE Trans. Image Processing 15(11), 3394–3361 (2006)
5. Hall, M.A.: Correlation-based Feature Selection for Machine Learning. Department of Computer Science, Hamilton, New Zealand
6. Kononenko, I.: Estimating Attributes: Analysis and Extensions of RELIEF. In: European Conference on Machine Learning, pp. 171–182 (1994)

7. Ali, S.I., Shahzad, W.: A Feature Subset Selection Method based on Symmetric Uncertainty and Ant Colony Optimization. International Journal of Computer Applications (0975 – 8887) 60(11) (December 2012)
8. Leung, K.M.: Naive Bayesian Classifier, Polytechnic University Department Of Computer Science / Finance and Risk Engineering, Copyright c, mleung@poly.edu (2007)
9. Cortes, C., Vapnik, V.N.: Support-Vector Networks. Machine Learning 20 (1995)
10. Viola, P., Jones, M.J.: Robust Real time Face Detection. International Journal of Computer Vision (2004)
11. MacQueen, J.B.: Some Methods for classification and Analysis of Multivariate Observations. In: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability 1, pp. 281–297. MR 0214227. Zbl 0214.46201. University of California Press (1967) (retrieved April 7, 2009)
12. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Burkhardt, H., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1407, pp. 484–498. Springer, Heidelberg (1998)
13. Bremner, D., Demaine, E., Erickson, J., Iacono, J., Langerman, S., Morin, P., Toussaint, G.: Output-sensitive algorithms for computing nearest-neighbor decision boundaries. Discrete and Computational Geometry 33(4), 593–604 (2005), doi:10.1007/s00454-004-1152-0
14. McCartin, B.J.: Mysteries of the equilateral triangle. Applied Mathematics Kettering University (2010)
15. Everett, M.G., Borgatti, S.P.: The centrality of groups and classes. Journal of Mathematical Sociology 23(3), 181–201 (1999)
16. FG-NET Aging Database (2010), `http://www.fgnet.rsunit.com`