

Topographical Features for Senior Adult Age Estimation

Ghalib A. Salman Salih^{a,*}, Gazali Sulong^a

^a University Technology Malaysia, Media and Game Centre of Excellence (Magicx), Faculty of Computing, Skudai, Johor, 81310 Malaysia

* Corresponding author email address: ghlbethawi@yahoo.com

Abstract

Automatic age estimation provided efficient solution for many applications in our life. One of the most significant biometric in estimating a human age is the face; since it is the most captured biometric and it contains a lot of age information embedded in it. Senior adult faces contain the most obvious age progression signs such as, wrinkles, lines and skin roughness; such features are produced by skin sagging and providing a lot of information about age progression. Representing wrinkles using ordinary lines or edges loses significant information. In this paper we propose modelling the 2D polynomial for the face image in order to increase the quality of the extracted features, then topographical features are extracted to represent age signs on human faces in senior adult ages; the most efficient features are selected using proposed features selection technique. Proposed features provided noticeable increment in extracted information, and the classification accuracy. Compared to the state of art, proposed features yielded encouraging results.

Keywords: Age anticipation, Senior-adult ages, Age signs, Features optimization, Topographical features

1. Introduction

Automatic estimation of human ages attracted more attention recently because of the wide areas of social applications that require such estimation. Access controlling of under age kids is required to prevent them from accessing some goods like wines or cigarettes, or even some web-pages with adult contents; ages should be also estimated for victims or runaway criminals when identifications are missed.

As human face contains considerable amount of information about the age (Gadbail *et al.*, 2014), it is widely used in age estimation researches; yet age effects on these information vary from age period to another. (Todd *et al.*, 1980). Geometric effects are obvious in young ages due to the rigid changes in shape and size of face (Berry and McArthur, 1985). Less vitality is notices as the age moves to adult faces; vitality changes are combined by some geometric changes (Geng *et al.*, 2007). Senior-adult age period has significant set of age texture effects in form of wrinkles, jowls, hair colour, and less vitality in the skin; in addition, skin sagging and drooping produce new type of features in form of bigger forehead and chin areas, and smaller eyes areas (George and Hole, 2000).

According to the different age effects, specifying own features for each age interval may yield better results than generalizing set of features for all age periods. In this paper, senior-adult ages will be studied using topographical features that depend on image gradient supported by

features selection technique. Topographical features are extracted from the 2D polynomial model for the face image; among the huge number of produced features, most efficient features are elected using features selection method.

The rest of this paper is organized as follows: Section 2 surveys the previous works in age estimation, Section 3 illustrates the wrinkles and face lines, Section 4 explains the 2D image polynomial, Section 5 illustrates topographical features, Section 6 illustrates features selection method, Section 7 discusses the results, and finally Section 8 includes the conclusions.

2. Literature Survey

The first study of age effects on human face was conducted several decades ago (Todd *et al.*, 1980); they provided set of age progression signs in form of wrinkles and head pose. After few years, baby faces were studied to represent age effects in childhood age period (Berry and McArthur, 1985). Effects of age progression were studied and represented as several types of features; in spite of previous studies on age progression effects, automatic age estimation were studied several years later. The first trial of estimating human age from face image was conducted in (Kwon and Da Vitoria Lobo, 1994), where the researchers classified human age into baby, young adult, and senior adult; they used positions of face components and distances between them as features for babies and young adults,

while wrinkles were proposed as features for senior adults. Wrinkles were also used to classify adult ages after distinguishing them from child faces using geometric features of face components (Horng *et al.*, 2001); authors represented face wrinkles as lines around mouth and eyes; such lines are determined using image gradient. Line representation for wrinkles causes some losing in significant information that can be extracted from wrinkles areas. In addition, adult ages before line appearance may be misclassified.

Variants of Local Binary Patterns (LBP) were combined to provide texture features; these features were used to represent the face from young to senior adult ages. The authors yielded encouraging results according to their state of art, though, their results had some bias; most of their correctly classified were yielded in (0- 20) years (Ylioinas *et al.*, 2012). This may indicate that their features are more suitable for young ages than for senior-adult ages. Statistical measurements were proposed to represent face smoothness as texture features; where smoothness yield high levels in child faces and low levels in senior-adult faces. In addition to skin roughness, face lines and wrinkles produce low levels of face smoothness (Salman and Sulong, 2014). Yielded results were encouraging in classifying face age into child, adult, and senior-adult without detailed classification within these classes. Wrinkles features were studied to support geometric features; depending on Gabor filter, mean, variance, and standard deviation were proposed to describe wrinkle features (Jana *et al.*, 2014).

3. Face Lines and Wrinkles

Since changes in head bones are insignificant in senior-adult ages (Sperber *et al.*, 2001), most of age features are produced by skin sagging and drooping; eyebrows sagging leads to bigger forehead and smaller eyes areas. Another type of features can be found in chin and nose drooping (Mulliken, 1976); these features are representing changes in the areas of these face components; on the other hand, the most obvious features can be found in form of wrinkles and face lines.

Wrinkles were studied as features in several researches, yet results needed improvements. Raw images were used to produce image edges; since a lot of edges can be produced from face images, they used edges inside the forehead and under the eyes (Kwon and Da Vitoria Lobo, 1994). Sobel operator was proposed for face edges and lines inside the forehead, around the eyes, and in cheeks area ignoring other areas (Horng *et al.*, 2001). Wrinkles contain significant information rather than edge lines (Hatzis, 2004), in addition, specifying areas to study the lines may lead to miss studying significant information in other areas, see Fig. 1.

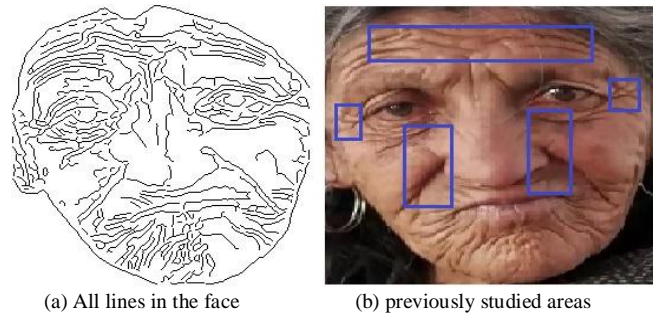


Fig. 1. Senior adult features in form of wrinkles and lines, where lines may ignore significant information of the wrinkles.

In our proposed method, most of information provided by face wrinkles and lines are studied without pre-determination of features location in studied face; since face may contain lines or wrinkles in many of its parts. The most effective features will be chosen as candidate features to represent age progression.

4. Image Polynomial

Pixels have levels of connectivity to other pixels in the same region; images may have noise or degradation. Thus, extracting edges of texture features from raw images may provide degraded features. Instead, image polynomial is proposed to reduce the effects of noise, degradation, and interconnectivities between pixels (Haralick, 1984; Moody, 1994).

Two dimensional (2D) image polynomial can be computed for the rows (r) and the columns (c); two index sets are used to generate polynomial coefficients, $(-1, 0, 1) \times (-1, 0, 1)$. Third degree of polynomial is preferred since it provides easy approach to find the suitable directional derivative; at the same time, it covers all required 2D coefficients $(1, r, c, r^2-2, c^2-2, rc, r^2-2, (r^3-3.4r), c (r^2-2), r(c^2-2), \text{ and } (c^2-3.4c)$ (Haralick, 1984), see Fig. 2 that illustrates the effect of each polynomial coefficient on the studied image.

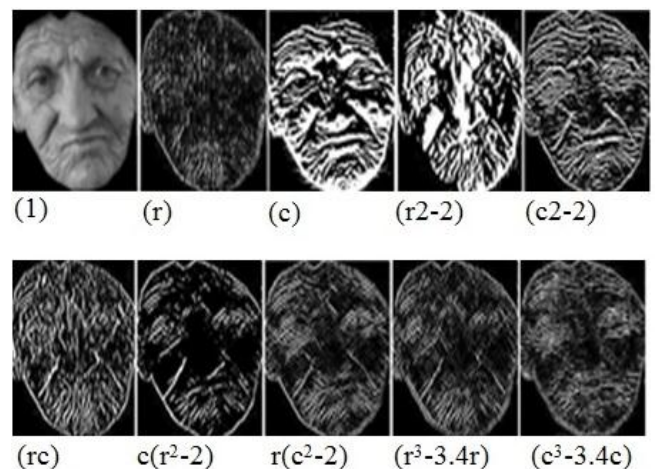


Fig. 2. The extracted information from tested image using each coefficient of 2D polynomial.

The 2D polynomial is found as the summation of these coefficients as in equation 1; the results image from this form of equation has the effects of necessary preprocessing operations (Abidi *et al.*, 1999):

$$f(x, y) = k_1 + k_2x + k_3y + k_4x^2 + k_5xy + k_6y^2 + k_7x^3 + k_8x^2y + k_9xy^2 + k_{10}y^3 \quad (1)$$

Using the index sets and generated masks, equation 1 can be written as:

$$f(x, y) = 1 + r + c + r^2 - 2 + rc + c^2 - 2 + r^3 + c(r^2 - 2) + r(c^2 - 2) + c^3 \quad (2)$$

5. Topographical Features

Although topographical features are not a new idea, our contribution using them consists of two parts. Firstly, these features were used in identification problems such as character recognition (Parvin and Medioni, 1989; Wang and Pavlidis, 1993); in such problems, the target is to find invariant features within different images for the same character, and variant features for different samples; such features lead to recognize the same character written in different styles and to distinguish between different characters even if they are written by the same person. In age estimation problems, the target is to find variant features for the same person over age progression (high differences within intraclass); at the same time, these features should have similar results for different persons at the same age (low differences within interclass). Such features provide estimation for specific age even using face images for different persons.

Secondly, features selection method in this paper is done without pre-determination of features locations. The behaviour of these features over different ages for the person and over different persons is used to elect features that have highest performance. In addition, topographical features provide additional information beside edge lines to be considered as significant features. See Fig. 3, in which, black line represents using lines to describe wrinkles; grade brown and grade blue represent (in addition to the black line) information provided by topographical features.

Instead of raw images, this paper adopts the mathematical description of an image as a base, which can be manipulated by gradient and Hessian matrix to produce topographical features. Assume that $Im_{x,y}$ represent face image, $\|\partial Im\|$ be the image gradient, and H be the Hessian matrix:

$$H = \begin{bmatrix} \frac{\partial^2 Im}{\partial x^2} & \frac{\partial^2 Im}{\partial x \partial y} \\ \frac{\partial^2 Im}{\partial y \partial x} & \frac{\partial^2 Im}{\partial y^2} \end{bmatrix} \quad (3)$$

Let $\lambda^T = [\lambda_1 \quad \lambda_2]$ the Eigen vector of the Hessian matrix; then the second derivative of Im will be:

$$f(2) = \lambda^T H \lambda \quad (4)$$

Topographical can be classified into three types, points, lines, and areas; yet, even lines and areas are constructed using set of points.

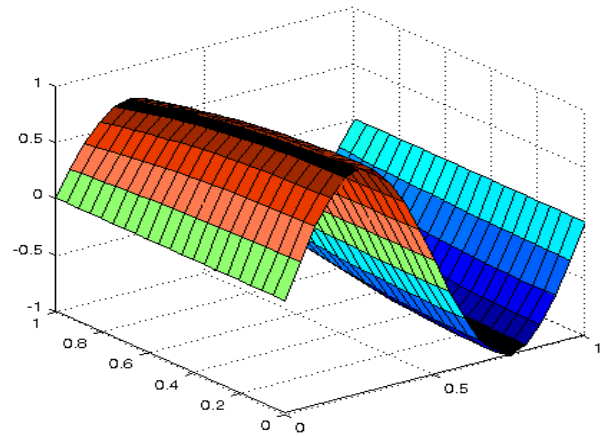


Fig. 3. Extra information in topographical features.

5.1 Points Features

In this type, features occur in single points, and the surrounding points are from different types such as:

- i. Peak: it occurs where the gradient is zero and surrounding points in all directions yield negative value for the second derivative.
- ii. Pit: the only difference between peaks and pits that the second derivative is positive in all surrounding points.
- iii. Saddle: it is similar to a peak or a pit in gradient values, but when second derivative has positive value in one direction, it must have negative value in the perpendicular direction.

5.2 Line Features

Some connected feature points may construct a continuous line of the same features type in the same direction such as:

- a) Ridge: it is a line sloping up, down, or flat; it requires zero gradient for flat ridge and non-zero gradient for sloping one. Second derivative should be negative in one direction only. When the adjacent points on continuous line are also ridges (most of ridge cases), this produces Ridge Line; ridge is rarely occur in a single point.
- b) Ravine: it is similar to ridge with one difference, the second derivative should be positive. When the set of ravines construct continuous line, it will be a Ravine Line.
- c) Zero Crossing: it can occur in form of line or point features; it occurs when the second derivative crosses zero value from negative to positive value, or vice versa. Zero crossing occurs regardless the image gradient.

5.3 Area Features

Some points of features are surrounded by set of feature of the same type in all directions; such features points construct an area of a feature type such as:

- iv. Flat: it is the area that contains set of features points, which are simply horizontal, which means that pixel values in such area are similar; this may leads to zero value for the image gradient and the second derivative
- v. Increasing Area: such area occurs with two conditions; firstly, image gradient is positive regardless the second derivative sign, and secondly, its points are not covered by any previous features type to avoid features confusion.
- vi. Decreasing Area: it is similar to increasing area but when image gradient is negative.
- vii. Wide black lines in Fig. 3 can illustrate ridge and ravine lines, and green lines can illustrate zero crossing lines, yet, topographical features are illustrated in Fig. 4.

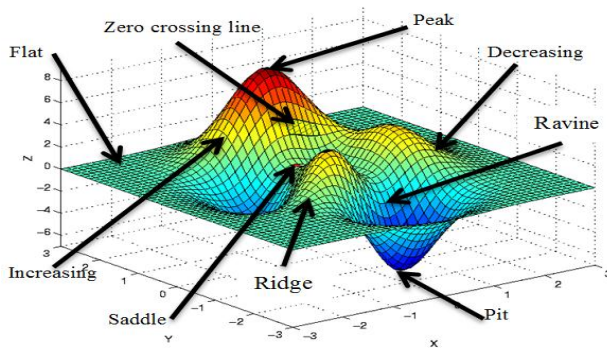


Fig.4. Summary of topographical features.

On the other hand, features are represented on a real face in Fig. 5. Table 1 summarizes the mathematical conditions for each type of features. Other features indicate fewer changes over age progression such as flat areas that indicate no changes from the previous ages.

6. Features Selection

Features selection is required to distinguish significant from insignificant features. Many researches proposed edges among topographical features as efficient features to represent age progression effects. One of the first trials proposed using local maximal and local minima to represent the edges combining by the direction of gradient (Haralick, 1984). Zero crossing lines were proposed to draw features lines and edges (Parvin and Medioni, 1989). Topographical features were studied and each feature area or line was labelled; the boundaries between regions were adopted as candidate features (Wang and Pavlidis, 1993). In 1995, ridge line, peak, and saddle points were proposed to form the image skeleton as candidate features (Lee and Kim, 1995). Another type of lines was proposed; images were converted into black and white images using



Fig.5: representing aged faces using topographical faces for more information than lines

adjustable threshold, then result edges and lines between regions were used as features (Abidi *et al.*, 1999).

Table 1
Necessary mathematical conditions for topographical features.

Feature Type	Conditions
1) Peak	$\ \nabla I_m\ = 0, \lambda_1 < 0, \lambda_2 < 0$
2) Pit	$\ \nabla I_m\ = 0, \lambda_1 < 0, \lambda_2 < 0$
3) Saddle	$\ \nabla I_m\ = 0, (\lambda_1 * \lambda_2) < 0$
4) Ridge	$\ \nabla I_m\ \neq 0, \lambda_1 < 0, (\lambda_1 * \nabla I_m) = 0$
	Or $\ \nabla I_m\ \neq 0, \lambda_2 < 0, (\lambda_2 * \nabla I_m) = 0$
	Or $\ \nabla I_m\ = 0, \lambda_1 < 0, \lambda_2 = 0$
5) Ravine	$\ \nabla I_m\ \neq 0, \lambda_1 > 0, (\lambda_1 * \nabla I_m) = 0$
	Or $\ \nabla I_m\ \neq 0, \lambda_2 > 0, (\lambda_2 * \nabla I_m) = 0$
	Or $\ \nabla I_m\ = 0, \lambda_1 > 0, \lambda_2 = 0$
6) Flat	$\ \nabla I_m\ = 0, \lambda_1 = 0, \lambda_2 = 0$
7) Increasing	$\ \nabla I_m\ > 0, \lambda_1 \neq 0, \lambda_2 \neq 0$
8) Decreasing	$\ \nabla I_m\ < 0, \lambda_1 \neq 0, \lambda_2 \neq 0$

In this work, proposed technique is to study all topographical features in the whole face and choose the most efficient features. All face features are studied to determine their changes for specific person over age progression and for different persons in same age. Efficient features should have significant features over different ages with insignificant changes within specific age even for different persons.

Where each pixel in the face image is labelled as one of the features, the behaviour of each feature will be studied using samples of (a) different ages for each of the (p) persons:

The features matrix F extracted from a face image contains N×M pixels, in which, each pixel is labelled as the corresponding feature:

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{1j} & \dots & f_{1M} \\ \vdots & \vdots & \vdots & & \vdots \\ f_{i1} & f_{i2} & f_{ij} & \dots & f_{iM} \\ \vdots & \vdots & \vdots & & \vdots \\ f_{N1} & f_{N2} & f_{Nj} & \dots & f_{NM} \end{bmatrix} \quad (4)$$

Test matrix AP is organized to check the performance of feature for each pixel in each available image; it elects the corresponding pixels form the same location of all of images in the test period.

$$AP_{ij} = \begin{bmatrix} f_{ij}(11) & f_{ij}(12) & f_{ij}(1k) & \dots & f_{ij}(1a) \\ \vdots & \vdots & \vdots & & \vdots \\ f_{ij}(h1) & f_{ij}(h2) & f_{ij}(hk) & \dots & f_{ij}(ha) \\ \vdots & \vdots & \vdots & & \vdots \\ f_{ij}(p1) & f_{ij}(p2) & f_{ij}(pk) & \dots & f_{ij}(pa) \end{bmatrix} \quad (5)$$

Where:

$f_{ij}(12)$, as example, represents the feature f_{ij} taken from the features matrix of the first person at the second available age.

AP_{ij} : represents the behaviour on the pixel f_{ij} in the all tested images.

Differences within intraclass (age progression of the same person) and inter-class (specific age for different persons) can be measured using suitable statistical measurements. There are several statistical measures of dispersion such as standard deviation (SD) - that can gauge the variety of set of values, though the distribution of the available ages should be concerned; inconsistent age distribution leads to inconsistent values for proposed features. Ratio measure (R) divides the Standard Deviation (SD) by the average of the values in each sample; such that, R eliminates the effect of different levels of values average:

$$R = \frac{SD}{\mu} \quad (6)$$

where: $SD = \sqrt{\frac{\sum(x_i - \mu)^2}{n}}$, $\mu = \frac{\sum x_i}{n}$.

Another proposed measure is the Roughness Coefficient (rc) that eliminate or reduce the effect of different periods; this measure normalizes the differences of the successive values by their differences from their average²⁰:

$$rc = \frac{\sum(x_i - x_{i-1})^2}{\sum(x_i - \mu)^2} \quad (7)$$

Using these measures, the dispersion of each row (dr_k) in AP matrix is computed to show the intra class dispersion; similarly, the dispersion of interclass is shown using the measures over columns (dc_h). The performance of each feature (θ_{ij}) can be represented by the ratio of the average of dr_k values divided by the average of dc_h :

$$\theta_{ij} = \frac{\sum_{k=1}^p dr_k / p}{\sum_{h=1}^a dc_h / a} = \frac{a \sum_{k=1}^p dr_k}{p \sum_{h=1}^a dc_h} \quad (8)$$

Computing θ_{ij} values for all features produces performance matrix that contains efficiencies of all features; face features from feature matrix f_{ij} that correspond the highest efficiencies θ_{ij} are chosen as candidate features for age estimation.

$$\theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{1j} & \dots & \theta_{1N} \\ \vdots & \vdots & \vdots & & \vdots \\ \theta_{i1} & \theta_{i2} & \theta_{ij} & \dots & \theta_{iN} \\ \vdots & \vdots & \vdots & & \vdots \\ \theta_{N1} & \theta_{N2} & \theta_{Nj} & \dots & \theta_{NN} \end{bmatrix} \quad (9)$$

7. Results and Discussion

This work was conducted depending on two types of datasets, standard FG-NET which contains (102 of 1002) face-images of (35-70) years old; in addition to (326) images from our private collected dataset. Five samples were chosen to demonstrate and discuss the results as in Table 2; they were chosen from both standard and private dataset with different levels of quality.

Table 2
Age distribution of each sample (s1,... s5).

S1	S2	S3	S4	S5
35	37	35	45	35
37	38	37	48	38
42	43	39	52	41
43	45	41	55	45
46	46	45	59	49
51	49	46	62	53
52	55	48	63	58
60	57	49	65	61
65	62	52	66	65
68	69	53	69	68

As performance matrix contains different level of features performance, and the goal is to choose features with highest performance, a threshold should be chosen to determine the level of performance to choose best features. Increasing the threshold limits the number of chosen features by the highest efficiency, yet it may lead to loose significant information. On the other hand, decreasing the threshold may lead for including insignificant features; this may decrease the classification accuracy (CA). Increasing the number of chosen features increases CA, though, the increments started to shrink until it almost stops at (80) features; with unstable increments and decrements, CA started decreasing stage after choosing (110) features. See Fig. 6.

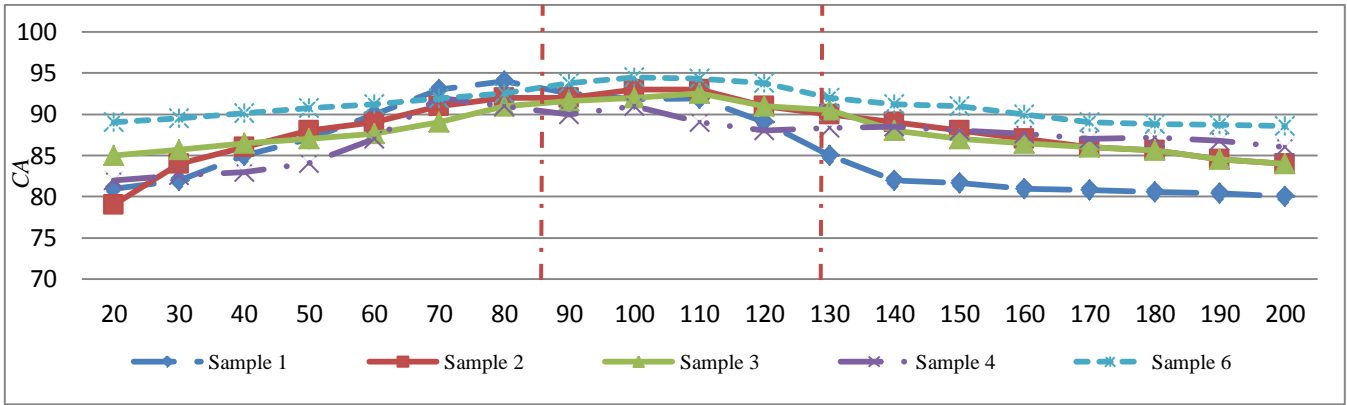


Fig. 6. The classification accuracy over different number of features.

In spite of image quality enhancement included in finding the 2D polynomial for the image, image quality affected CA; the first sample were taken from FG-NET standard dataset which have less quality than other samples. Sample 5 that yielded almost stable CA was collected with the best qualities from private collected dataset; samples 2, 3, and 4 were chosen with ordinary quality between as in samples 1 and 5.

Different samples yielded different results for each of R and rc; best CA for each of them over each sample can be illustrated in Fig. 7. For the samples 1, and 2, ages are distributed over the whole range with unstable successive differences; rc handled such samples better than R and yielded higher CA. Although ages in the samples 3, and 4 have similar successive differences to the previous samples, most of their ages can be found in one half of the age interval than the other; in such samples, R eliminated the effect of average deviation and yielded higher CA than rc. Sample 5 provides consistent age distribution; R and rc yielded close results.

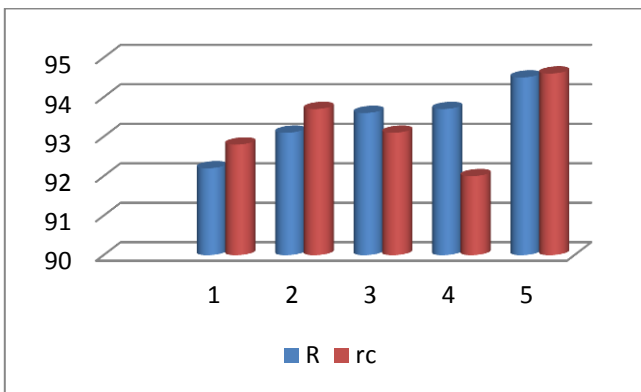


Fig. 7. The different results for using R and rc over 5 samples.

Features have variant levels of significant participating according to their significance of discrimination, or their frequency of occurrence. Flat areas indicate the absence of changes into wrinkles or lines, which may decrease the participation of such areas in accurate classification. Although some features have low occurrence such as saddle points, they have classification significance when they occur, see Fig. 8.

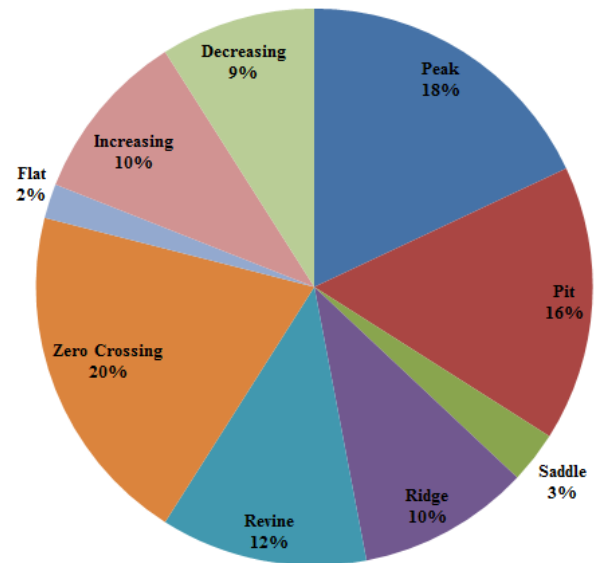


Fig. 8. Features participating in accurate classification.

Although peaks, pits, and zero crossing points, which are used in previous works, are the most efficient features, they form only 54% from the available features; ignoring other features ignores 46% of participating features.

Comparing with state of art, proposed features yielded encouraging results in form of CA. Wrinkles were studied using Gabor filter; their features were using the average and the dispersion for the magnitude response within wrinkle areas (Jana *et al.*, 2014). Active Appearance Model AAM and Singular Value Decomposition SVD were proposed as features for classification of age groups; they yielded 86.54% for the best of their CA, yet they yielded 65.92% for CA within (46-69) ages (Bhalekar *et al.*, 2014). Race, gender, and age estimation were proposed using one methodology and sets of features; they yielded 88.1% for their best CA (Han and Jain, 2014). Proposed features in this work yielded 93.7% for our best of CA, yet, using optimal image quality with optimal age distribution we yielded 94.6% for CA; see Table 3.

Table 3
Comparison with the state of art using CA.

AGE & GENDER	AAM & SVD	GABOR FILTER	OUR	OPTIMAL
88.1%	65.92%	93.01%	93.7%	94.6%

8. Conclusions

This work proposes set of features for senior adult age estimation; the major approach is to build 2D image polynomial that provides enhanced image to extract efficient topographical features. An effective features selection method is used to choose the best set of features. The best number of chosen features was between 80 and 110 features. Within features selection technique, ratio R and roughness coefficient rc measures are proposed to measure values dispersion; rc yielded better results than R for the ages with unstable successive differences. R yielded better results than rc for unbalanced age distribution. Proposed features provide more information about wrinkles and face lines than ordinary lines and edges, which were used by previous works; they used only 54% of wrinkles information. Our proposed features yielded encouraging results comparing with the stat of art.

Acknowledgements

We are grateful to the Daily Photo Project for providing sets of sequential daily, monthly, and yearly photos, from which we built our private dataset.

References

- Abidi, B. R., Sari-Sarraf, H., Goddard Jr, J. S. and Hunt, M. A. (1999). Facet Model and Mathematical Morphology for Surface Characterization. *Photonics East99*. 334-344.
- Berry, D. S. and McArthur, L. Z. (1985). Some Components and Consequences of a Babyface. *Journal of personality and social psychology*. 48(2), 312.
- Bhalekar, S., Bhandare, A., Bhoite, R. and Dere, A. (2014). Human Age Estimation from Facial Images Using Artificial Neural Network.
- Gadbail, D. R., Dhande, S. S. and Pimple, K. M. (2014). Age Synthesis and Estimation from Face Image. *International Journal Of Engineering And Computer Science*. 3(4 April, 2014), 4.
- Geng, X., Zhou, Z.-H. and Smith-Miles, K. (2007). Automatic Age Estimation Based on Facial Aging Patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 29(12), 2234-2240.
- George, P. A. and Hole, G. J. (2000). The Role of Spatial and Surface Cues in the Age-Processing of Unfamiliar Faces. *Visual Cognition*. 7(4), 485-509.
- Han, H. and Jain, A. (2014). Age, Gender and Race Estimation from Unconstrained Face Images. *Dept. Comput. Sci. Eng., Michigan State Univ., East Lansing, MI, USA, MSU Tech. Rep.(MSU-CSE-14-5)*.
- Haralick, R. M. (1984). Digital Step Edges from Zero Crossing of Second Directional Derivatives. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. (1), 58-68.
- Hatzis, J. (2004). The Wrinkle and Its Measurement—: A Skin Surface Profilometric Method. *Micron*. 35(3), 201-219.
- Hornig, W.-B., Lee, C.-P. and Chen, C.-W. (2001). Classification of Age Groups Based on Facial Features. *Tamkang Journal of Science and Engineering*. 4(3), 183-192.
- Jana, R., Datta, D. and Saha, R. (2014). Age Group Estimation Using Face Features. *International Journal of Engineering and Innovative Technology*. 3(2).
- Kwon, Y. H. and Da Vitoria Lobo, N. (1994). Age Classification from Facial Images. *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*. 762-767.
- Lee, S.-W. and Kim, Y. J. (1995). Direct Extraction of Topographic Features for Gray Scale Character Recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 17(7), 724-728.
- Moody, E. B. (1994). Discrete Orthogonal Polynomial Restoration of Images Degraded by Spatially Varying Point Spread Functions. *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*. 680-684.
- Mulliken, J. B. (1976). Handbook of Facial Growth. *Archives of Surgery*. 111(9), 1043-1043.
- Parvin, B. and Medioni, G. (1989). Adaptive Multiscale Feature Extraction from Range Data. *Computer Vision, Graphics, and Image Processing*. 45(3), 346-356.
- Salman, G. A. and Sulong, G. (2014). Localized Statistical Features Representing Face Vitality for Age Estimation. *Australian Journal of Basic and Applied Sciences*. 8(13) (August 2014), 698-703.
- Sperber, G. H., Guttmann, G. D. and Sperber, S. M. (2001). *Craniofacial Development (Book for Windows & Macintosh)*. PMPH-USA.
- Todd, J. T., Mark, L. S., Shaw, R. E. and Pittenger, J. B. (1980). The Perception of Human Growth. *Scientific american*. 242(2), 132.
- Wang, L. and Pavlidis, T. (1993). Direct Gray-Scale Extraction of Features for Character Recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 15(10), 1053-1067.
- Ylioinas, J., Hadid, A. and Pietikainen, M. (2012). Age Classification in Unconstrained Conditions Using Lbp Variants. *Pattern Recognition (ICPR), 2012 21st International Conference on*. 1257-1260.