

# Invariant Feature Extraction for Component-based Facial Recognition

Adam Hassan<sup>1</sup>

Sudan University of Science and Technology  
College of Computer Science & Information Technology  
Khartoum, Sudan

Serestina Viriri<sup>2</sup>

University of KwaZulu-Natal  
School of Maths, Statistics & Computer Science  
Durban, South Africa

**Abstract**—This paper proposes an alternative invariant feature extraction technique for facial recognition using facial components. *Can facial recognition over age progression be improved by analyzing individual facial components?* The individual facial components: eyes, mouth, nose, are extracted using face landmark points. The Histogram of Gradient (HOG) and Local Binary Pattern (LBP) features are extracted from the individually detected facial components, followed by random subspace principal component analysis and cosine distance. One of the preprocessing steps implemented is the facial image alignment using angle of inclination. The experimental results show that facial recognition over age progression can be improved by analyzing individual facial components. The entire facial image can change over time, but appearance of some individual facial components is invariant.

**Keywords**—Invariant features; facial components; facial recognition; age progression; HOG; LBP

## I. INTRODUCTION AND BACKGROUND

Face recognition is a challenging and relevant research area in image processing and computer vision community. Significant advances have been achieved throughout the last years. The majority of the research works studied general face recognition without considering face recognition over age progression. There are few research works which focus on age-invariant face recognition, and some of these related works are found in [21] [25] Hassan et al. [7] categorized the invariant discriminative feature extraction methods into two groups; generative models, and non-generative approaches. The generative models focus on learning the joint probability distribution while the non-generative models trade on the conditional probability distribution [17].

For age detection, an aging function is used and conducted with parametric model to get an exact age then another set of parameters are utilized to produce the target age [13]. However, aging transformations differ for different persons. An aging pattern subspace which deals with a sequence of individual face images, arranges them in time order that allows extracting features from both the shape and texture intensity [4]. Park et al. [20] used a model of 3D aging technique and show that the method can eliminate the age variations. However, generative methods used for face recognition suffer from poor aging process representation especially when just a few number of training image samples are available.

An effective alternative way to solve the limitations of generative models for better face representation is to implement local descriptors. Local descriptors are capable of

recognizing aging variations, and are robust to additional intra-class variations recognition. Many techniques associated with local descriptors to discriminate aging variations features are proposed in [5] [11] [14] [23].

Among effective local feature descriptors, Local Binary Patterns (LBP) [1], [2] [18] emerges as the famous candidate method. However, it is not always that the uniform binary patterns have higher frequency as it is known but, sometimes non-uniform binary patterns perform better [6]. Gong et al. [6] considered the effectiveness of LBP, and proposed a feature descriptor associated with maximization of the code entropy. Another effective face descriptor of age progression is used for face verification described in [15] [16]. They proposed a method which uses the gradient orientation pyramid (GOP), and found that face recognition algorithms degrade when age gap exceeds four years.

Many approaches based on local features have been proposed. The local descriptors add an additional detailed facial characteristic that is substantial to the recognition process. Ahonen et al. [2] introduced LBP texture descriptor for facial representation. It is considered as an efficient and simple texture descriptor which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. The resulting descriptor formed from the histogram of the labels. Then histogram distance is used as dissimilarity measure between the pair of facial images.

Another extension of LBP is Enhanced Local Binary Patterns (ELBP) [24] which performs face representation using threshold constant to threshold pixels into three values. Another prominent feature representation approach is Elastic Bunch Graph Matching (EBGM) [26]. It is an algorithm which localizes a set of features of certain facial points and extracts Gabor jets at those points. One of high distinctive facial features representations is Scale Invariant Feature Transform (SIFT) [12] which serves feature invariance from different views: translation, scaling, and rotation. Bay et al. [3] introduced the Speeded Up Robust Features (SURF), which is a detector and descriptor for patented local feature. This method takes the advantage of distinctiveness, and robustness, so that can be calculated and compared much faster.

Moreover, biometric technology is one of the efficient methods to identify humans depending on face components characteristics. It is mentioned in the literature [27] that face recognition is considered one of the systems which is used in the field of biometric technology. Heisele et al. [10] extracted

frontal face features to obtain eyes and the mouth components, then computed the triangle area between them. Component-based face recognition methods build upon multiple models with a number of facial components that represent an image on training phase. Many approaches are proposed on global-based technique, but component-base approaches are not intensively researched [8]. Geometry-based feature techniques for face recognition require calculation with geometrical features extracted from facial image. Face image representation can be viewed using size and position of facial components such as eyes, nose, mouth, forehead and cheeks.

The rest of the paper is organized as follows: Section II describes the methods and techniques implemented, Section III discusses the results achieved, and Section IV presents the conclusion and the envisioned future work.

## II. METHODS AND TECHNIQUES

The proposed model consists of two main stages: preprocessing and feature extraction. Preprocessing stage consists of land mark face detection, face alignment, components cropping. The proposed model for invariant feature extraction for component-based facial recognition is depicted in Fig. 1. The proposed model allows all facial components to be allocate initial same weights, but their matching scores are sorted to select the highest score among them.

### A. Preprocessing

**Landmarks Detection:** The original size of each image is 200x 240 pixels. At the preprocessing step, we initially detect 66 landmark points and use the two outer landmarks of eyes component to set the face image horizontally using specific calculated angle for rotation. When the image is well aligned according to eyes then detect new landmarks for next step which assist to crop facial components accurately. Fig. 2 shows the series of preprocessing steps.

Face recognition under pose variations is to recognize faces images of different poses. Face recognition rates are very poor when one tries to match images of different poses for the same person using any well known recognition technique. Hemlata et al. [9] detected eyes region using features of face connected components but, instead of that we use detected landmark points of outer corners of the two eyes to align the image face horizontally. if the face image is not horizontally aligned it must be rotated clock wise or anti clock wise base on angle direction. Angle of inclination of an image when it is correctly calculated then the image can be aligned. As depicted in Fig. 3, the two black filled circles represent the eyes, where:

$L(L_x, L_y)$  is the coordinates of left eye and  $R(R_x, R_y)$  refers to the coordinates of right eye whereas the intersection of vertical and horizontal line makes the point  $M$  with its coordinates  $M(L_x, R_y)$ .

$$\text{Opposite} = \sqrt{(R_x - L_x)^2 + (R_y - R_y)^2}$$

$$\text{Adjacent} = \sqrt{(L_x - L_x)^2 + (L_y - R_y)^2}$$

To align facial image depicted in Fig. 2, the angle (theta) can be calculated using the equation:

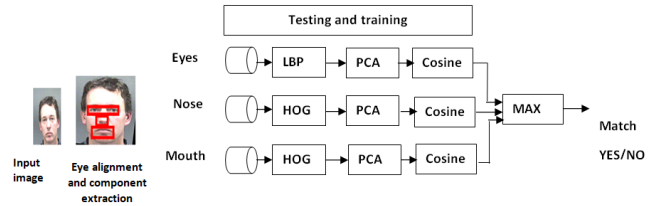


Fig. 1. Proposed Model for Component-based Facial Recognition

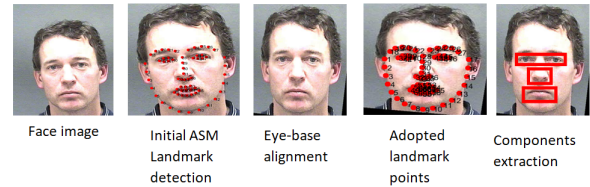


Fig. 2. Preprocessing steps

$$\theta = \tan \text{inverse} \frac{\text{Opposite}}{\text{Adjacent}} \quad (1)$$

**Components Extraction:** The work conducted on three facial components: eye pair, nose, and mouth. Eye pair component region extends from landmark point 37 to 43 as rectangle Length. To determine the width, arrange all the eye pair component points and select the highest and lowest points and the line between them is the rectangle width. For nose component, point 29 and the lowest nose point are selected besides point 32 to 36. The third component is the mouth where points 49 and 55 represent the length and highest and lowest points are the width.

### B. Components Representation

This work implements face components matching using cosine distance for PCA features trained on two different separated descriptors. Rather than combining different features as in [19] we proposed to vote one descriptor feature per component as a result of pre-experiments which reduces computational cost and increases the overall performance. For all facial components, we extract both LBP and HOG features utilizing a patch size of 8x8 without overlap. Features extracted from all patches are concatenated producing two different feature vectors for each facial component. However, we get different feature vector dimensions before passing it to PCA.

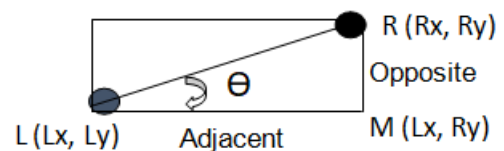


Fig. 3. Eye pair rectangle

LBP feature vector length is approximately four times the HOG feature vector length. Therefore, we sample each four LBP neighboring features as one feature point using mean value. The following equation shows how to form the desired features.

$$Desired\ features = \sum^n mean(R_{i..j}) \quad (2)$$

Where, n is the number of features is the raw features, i..j are the neighboring features. When the desired LBP features are obtained, then the feature dimensions are comparable. For each facial component, we build disjointed folds containing 300 genuine pairs each beside 40 different impostors. First, we start with age gap 0-1 to vote one descriptor that performs better than the other. Obviously, the age gap between each genuine pair is one year, as well as between the younger and impostor ones through all the forty impostors. We pass the features of one facial component to PCA for dimensionality reduction, and then we obtain PCA eigenvectors for genuine pairs and the corresponding impostor. Descriptor features of each component are treated separately till matching and decision phase then we fuse the scores of facial components.

### C. Component Matching

For each pair one image (the younger one) is compared with all 40 impostor images- (each is 1 year older than the genuine one) producing 40 different measures. The decision is set to 1 If the cosine similarity between the two genuine pairs is less than all the 40 measures, otherwise, the decision is 0. As shown in Table I, we analyzed all age ranges and divide the data set into groups. while Table II shows the performance which is discussed in next section. Algorithm 1 describes the matching process. We construct fold from each group except the last two groups due to shortest of images numbers. All component decision scores are shared for final decision using the maximum operation as the following:

$$Final\ decision = \sum_{j=1}^n max(I_j) \quad (3)$$

#### Algorithm 1

- 1: for i= 1 to n do
- 2: get the younger image of genuine pair get genuine(i)
- 3: for j =1 to m do
- 4: get impostor(j)
- 5: measure(j)  $\leftarrow$  distance(genuine(i), impostor(j))
- 6: measures  $\leftarrow$  [measures: measure(j)]
- 7: end for
- 8: sort all measures ascending measure Sort $\leftarrow$  sort(measures)
- 9: genuine Distance  $\leftarrow$  distance(genuine pair)
- 10: if genuine Distance < measure Sort
- 11: match = 1 else match = 0
- 12: end if
- 13: matches  $\leftarrow$  [matches; match]
- 14: end for

The final decision is maximized if there exist only one component is observed as 1. However, this may lead to wrong

TABLE I. AGE RANGE GROUPS AND CORRESPONDING NUMBER OF IMAGES

Age range in years	No. of images
16-19	7469
20-29	16325
30-39	15354
40-49	12052
50-59	3599
60-69	319
70-77	16

TABLE II. PERFORMANCE PER EACH COMPONENT AND A NUMBER OF ALTERNATIVE COMBINATIONS

Age Gap	Eyes	Nose	Mouth	Eyes and Nose	Eyes and Mouth	Nose and Mouth	Eyes Nose and Mouth
0-1	81.21	83.89	80.54	96.64	92.61	95.97	97.98
1-5	77.41	80.67	70.85	94.04	88.61	90.57	95.65

decision if one component is false accepted while the two others are truly rejected, but this is not the case in this research work.

## III. RESULTS AND DISCUSSION

### A. Data Set

We perform our work using the public domain MORPH [22] album 2 data set that contains 55,134 facial images from 13,617 classes. Table I shows the age ranges groups and corresponding number of images. We perform our work using two age gaps 0-1, and 1-5 and vote only one feature extractor for facial component. LBP and HOG feature extractors are extracted from each component then feature dimensions are intended to be most comparable, because we observed that the input feature dimension to PCA has crucial effect on accuracy. Equal dimensions are obtained as feature reduction for both extractors. Fig. 4 shows feature extractor performance for each component using ROC curves. Each component is analyzed using both extractors Therefore, we are entitled to choose the best for a later phase. Table II illustrates the best performance per each component and a number of alternatives established with combining two or three different components resulting in increasing the accuracy using final decision equation 2. Also we report combination of all components ROC curve for age gap 0-1 and gap 1-5 depicted in Fig. 5.

## IV. CONCLUSIONS AND FUTURE WORK

We have analyzed different facial components over age progression utilizing a component based face representation and cosine similarity matching algorithm then max score for final decision. We perform our experiments on MORPH dataset and categorized it with different age ranges as in Table I. The proposed approach is robust to face recognition over

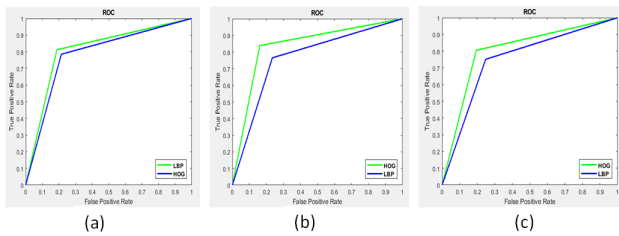


Fig. 4. Feature Extraction Voting per Component. (a) Eye pair: LBP performs better than HOG. (b) Nose: HOG performs better than LBP. (c) Mouth: else HOG is outperformed.

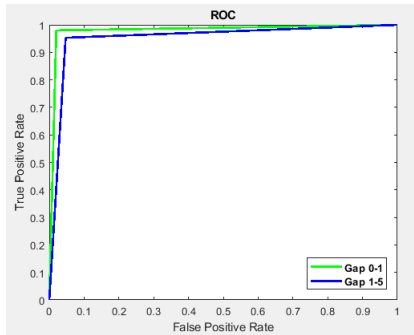


Fig. 5. MORPH results on 0-1 and 1-5 year age gap data sets

age progression. Our further work will include more facial components and additional dataset with larger age gaps to see how the proposed work can improve face recognition rate.

## REFERENCES

- [1] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. Face recognition with local binary patterns. In *European conference on computer vision*, pages 469–481. Springer, 2004.
- [2] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12):2037–2041, 2006.
- [3] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *European conference on computer vision*, pages 404–417. Springer, 2006.
- [4] Xin Geng, Zhi-Hua Zhou, Yu Zhang, Gang Li, and Honghua Dai. Learning from facial aging patterns for automatic age estimation. In *Proceedings of the 14th ACM international conference on Multimedia*, pages 307–316. ACM, 2006.
- [5] Dihong Gong, Zhifeng Li, Dahua Lin, Jianzhuang Liu, and Xiaou Tang. Hidden factor analysis for age invariant face recognition. In *Proceedings of the IEEE international conference on computer vision*, pages 2872–2879, 2013.
- [6] Dihong Gong, Zhifeng Li, Dacheng Tao, Jianzhuang Liu, and Xuelong Li. A maximum entropy feature descriptor for age invariant face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5289–5297, 2015.
- [7] Adam Hassan and Serestina Viriri. Invariant feature extraction for facial recognition: A survey of the state-of-the-art. In *2018 Conference on Information Communications Technology and Society (ICTAS)*, pages 1–6. IEEE, 2018.
- [8] Bernd Heisele, Thomas Serre, and Tomaso Poggio. A component-based framework for face detection and identification. *International Journal of Computer Vision*, 74(2):167–181, 2007.
- [9] A Hemlata and Mahesh Motwani. Face detection by finding the facial features and the angle of inclination of tilted face. *International Journal of Computer Science Issues (IJCSI)*, 10(2 Part 1):472, 2013.
- [10] Rein-Lien Hsu and Anil K Jain. Generating discriminating cartoon faces using interacting snakes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11):1388–1398, 2003.
- [11] Felix Juefei-Xu, Khoa Luu, Marios Savvides, Tien D Bui, and Ching Y Suen. Investigating age invariant face recognition based on periocular biometrics. In *2011 International Joint Conference on Biometrics (IJCB)*, pages 1–7. IEEE, 2011.
- [12] Dakshina Ranjan Kisku, Ajita Rattani, Enrico Grosso, and Massimo Tistarelli. Face identification by sift-based complete graph topology. In *2007 IEEE workshop on automatic identification advanced technologies*, pages 63–68. IEEE, 2007.
- [13] Andreas Lanitis, Christopher J. Taylor, and Timothy F Cootes. Toward automatic simulation of aging effects on face images. *IEEE Transactions on pattern Analysis and machine Intelligence*, 24(4):442–455, 2002.
- [14] Zhifeng Li, Unsang Park, and Anil K Jain. A discriminative model for age invariant face recognition. *IEEE transactions on information forensics and security*, 6(3):1028–1037, 2011.
- [15] Haibin Ling, Stefano Soatto, Narayanan Ramanathan, and David W Jacobs. A study of face recognition as people age. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE, 2007.
- [16] Haibin Ling, Stefano Soatto, Narayanan Ramanathan, and David W Jacobs. Face verification across age progression using discriminative methods. *IEEE Transactions on Information Forensics and security*, 5(1):82–91, 2009.
- [17] Andrew Y Ng and Michael I Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In *Advances in neural information processing systems*, pages 841–848, 2002.
- [18] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (7):971–987, 2002.
- [19] Charles Otto, Hu Han, and Anil Jain. How does aging affect facial components? In *European Conference on Computer Vision*, pages 189–198. Springer, 2012.
- [20] Unsang Park, Yiyong Tong, and Anil K Jain. Age-invariant face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 32(5):947–954, 2010.
- [21] Narayanan Ramanathan, Rama Chellappa, and Soma Biswas. Computational methods for modeling facial aging: A survey. *Journal of Visual Languages & Computing*, 20(3):131–144, 2009.
- [22] Tapan Kumar Sahoo and Haider Banka. Multi-feature-based facial age estimation using an incomplete facial aging database. *Arabian Journal for Science and Engineering*, 43(12):8057–8078, 2018.
- [23] Diana Sungatullina, Jiwen Lu, Gang Wang, and Pierre Moulin. Multi-view discriminative learning for age-invariant face recognition. In *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, pages 1–6. IEEE, 2013.
- [24] Xiaoyang Tan and William Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*, 19(6):1635–1650, 2010.
- [25] Yandong Wen, Zhifeng Li, and Yu Qiao. Latent factor guided convolutional neural networks for age-invariant face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4893–4901, 2016.
- [26] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph Von Der Malsburg. Face recognition by elastic bunch graph matching. In *International Conference on Computer Analysis of Images and Patterns*, pages 456–463. Springer, 1997.
- [27] Wenyi Zhao, Rama Chellappa, P Jonathon Phillips, and Azriel Rosenfeld. Face recognition: A literature survey. *ACM computing surveys (CSUR)*, 35(4):399–458, 2003.