Development of a Modified Local Binary Pattern-Gabor Wavelet Transform Aging Invariant Face Recognition System

Ayodele Oloyede, Temitayo Fagbola, Stephen Olabiyisi, Elijah Omidiora and John Oladosu

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

February 22, 2021
Development of a Modified Local Binary Pattern-Gabor Wavelet Transform Aging Invariant Face Recognition System

Oloyede Ayodele  
Ladoke Akintola University of Technology, Ogbomoso  
Department of Computer Science & Engineering  
+2348080715634  
delededee@yahoo.com

Fagbola Temitayo  
Federal University, Oye-Ekiti  
Department of Computer Science  
+2347030513010  
temitayo.fagbola@fuoye.edu.ng

Olabiyisi Stephen  
Ladoke Akintola University of Technology, Ogbomoso  
Department of Computer Science & Engineering  
+2348036669863  
soolabiyisi@lautech.edu.ng

Omidiora Elijah  
Ladoke Akintola University of Technology, Ogbomoso  
Department of Computer Science & Engineering  
+2347030712446  
eoomidiora@lautech.edu.ng

Olabiyisi Stephen  
Ladoke Akintola University of Technology, Ogbomoso  
Department of Computer Science & Engineering  
+2348034556065  
jboladosu@lautech.edu.ng

ABSTRACT

Human faces undergo considerable amount of variations with aging. This variation being experienced in facial texture and shape with different ages of a particular subject makes recognition of faces very difficult. However, most existing Face Recognition Systems (FRS) suffer from high misclassification of faces because of the large variation in face appearances of the same individual due to aging. This drawback is also aggravated by the fact that most currently existing age-invariant FRS adopt holistic feature extraction techniques (FET), which are computationally time-inefficient and suffer from the curse of dimensionality, in their development. Sequeal to these, a swarm-optimized age-invariant FET for a FRS was developed and presented in this paper. The developed swarm-optimized age-invariant FET tagged swarm-optimized LBP-GWT, which consists of Local Binary Pattern (LBP) and Gabor Wavelet Transform (GWT), was used for extraction of facial features from the face images. Procedurally, LBP and GWT were used to extract facial features relating to the eye lids, nose and lips. Discriminative features were selected from the features extracted by LBP and GWT using particle swarm optimization algorithm. The selected features were fused into a single feature set using sum rule strategy. Based on the single feature set, faces were recognized and classified into age-varying collections of different individuals using support vector machine. The developed swarm-optimized age invariant feature extraction technique serves as improvement over Histogram of Gradients, Principal Component Analysis-Local Discriminant Analysis, Local Binary Pattern and Gabor Wavelet Transform feature extraction techniques in terms of false accept rate, false reject rate, recognition accuracy and recognition time. This technique could be integrated into emerging age-invariant face recognition systems towards their improved performance.

CCS Concepts

- Computing methodologies ➔ Artificial intelligence ➔ Machine learning ➔ Machine learning algorithms ➔ Feature selection

Keywords

Aging Face Recognition, Computational Efficiency, Feature Extraction, Particle Swarm Optimization

1. INTRODUCTION

Face recognition across ages is an important problem and has many applications, such as passport photo verification, image retrieval, surveillance (Narayanan and Rama, 2006). This is a challenging task because human faces can vary a lot over time in many aspects including facial texture, shape, facial hair and presence of glasses (Omidiora, Fakolujo, Ayeni, Olabiyisi and Arulogun, 2008; Saeid and Leila, 2012). Moreover, human faces also undergo growth related changes that are manifested in the form of shape and textural variations (Narayanan and Rama, 2006). While facial aging is mostly represented by the facial growth in younger age groups, it is also represented by relatively large texture changes and minor shape changes due to the change of weight, presence of wrinkles or stiffness of skin in older age groups above 18 years. Therefore, an age correction scheme needs to be able to compensate for both types of aging processes. More often than not, most existing age-invariant face recognition systems are computationally very expensive which makes it difficult to be implemented in practice. This is due to the fact that such implementations are based on holistic feature extraction techniques which are highly sensitive to illumination and aging conditions (Narayanan and Rama, 2006; Huseyin and Onse, 2012). Hence, there arises the need for a computationally-efficient feature extraction technique suitable for real-time use. It must be noted that the success of any face recognition system depends on the feature extraction technique (Biswa, Aggarwal and Chellappa, 2008).
2. RELATED WORKS

Face recognition and detection has been widely studied for several decades. A lot of work has been done to handle the problem under different conditions, including age variations, lighting, pose and expression. Lanitis et al. (2002) developed a method for simulating aging effects on face images. On a database of age progressive images of individuals each under 30 years of age, a combined shape-intensity model was used to represent faces. The authors modeled age as a quadratic function of the PCA coefficients extracted from the model parameters. Results on experiments such as estimating the age of an individual from his/her face image and simulating aging effects on face images was reported. The model also performed on a similar dataset and evaluated the performance of three age classifiers: the first was a quadratic function of the model parameters; the second was based on the distribution of model parameters and the third was based on supervised and unsupervised neural networks trained on the model parameters. The model presented the most efficient result using quadratic function of the model parameters for the classifiers. However, the framework is not implementable in practice for use by age-invariant face recognition systems.

Elisseef, Evgeniou and Pontil (2004) studied the leave-one-out and generalization errors of ensembles of kernel machines such as SVMs. It was discovered that the best SVM and the best ensembles had about the same test performance; with appropriate tuning of the parameters of the machines, combining SVMs does not lead to performance improvement compared to a single SVM. However, ensembles of kernel machines are more stable learning algorithms than the equivalent single kernel machine; that is, bagging increases the stability of unstable learning machines.

SVM only performs excellently well when introduced to a considerably small set of features. The curse of dimensionality was not addressed and this resulted into inaccurate recognition results at high time complexity overhead. By optimizing SVM, a more significant result can be obtained and this forms an objective of this research work.

Haibin, Stefano, Narayanan and David (2010) studied the problem of face verification in the presence of age progression by designing and evaluating discriminative approaches. These directly tackle verification tasks without explicit age modeling, which is a hard problem by itself. The authors used gradient orientation (GO) to realize a simple but effective representation of faces for the aging problem after discarding magnitude information. This representation is further improved when hierarchical information was used, which resulted in the use of the gradient orientation pyramid (GOP). When combined with a support vector machine (SVM), GOP demonstrated excellent performance with seven different approaches including two commercial systems. However, the experiments were conducted on the FGnet dataset and two large passport datasets. This approach follows discriminative methods and did not take into consideration simultaneous feature analysis and classification that could help realize robustness and computational efficiency which are highly desirable properties of any age-invariant face recognition system. This makes their approach less applicable for practical use.

Huseyin and Osen (2012) used original PCA and subspace LDA methods to extract facial image features. Images were projected into a subspace by PCA in such a way that the greatest and the least variance values among the images are captured by the first and the last perpendicular dimensions of image feature subspace respectively. In this respect, the eigenvectors of the covariance matrix are computed which correspond to the directions of the principal components of the original data and their statistical significance is given by their corresponding eigenvalues. PCA was used for the purpose of dimension reduction by generalizing the data while SVM was used for the final classification. Subspace LDA method is simply the implementation of PCA by projecting the data onto the eigenspace and then implementing LDA to classify the eigenspace projected data. Holistic approaches based on PCA and LDA suffer from the curse of dimensionality (Shinde and Gunjal, 2012). That is, the time required for an algorithm grows exponentially with the number of features involved, rendering the algorithm intractable in extremely high-dimensional problems. The result obtained lacks strong discrimination ability and timely inefficient.

Dihong, Zhifeng, Dahua, Jianzhuang and Xiaou (2013) developed a new method called Hidden Factor Analysis (HFA). This approach is motivated by the belief that the facial image of a person can be expressed as a stable feature for face recognition; while the age factor changes as the person grows.

For computational simplicity, the authors assumed a linear model, where the identity components and the age components lie on two different subspaces. In this way, the problem of separating identity and age factors naturally reduces to a problem of learning the basis of these subspaces. As both the subspaces and the latent factors are unknown in the training stage, an algorithm that can jointly estimate both from a set of training image was derived based on an Expectation-Maximization process. In this process, the latent factors and the model parameters are iteratively updated to maximize a unified objective. In the testing, given a pair of face images with unknown ages, the match score between them were computed by inferring and comparing the posterior mean of their identity factors. This approach is very complex and lacks strong discrimination ability; it also requires a lot of training images and consumes high computational resources. Every face image is divided into a set of overlapping patches, and then applied the HOG descriptor on each patch to extract the HOG features. The extracted HOG features from all the patches were concatenated together to form a long feature vector for further analysis. Prior to applying the HOG feature extractor, the face images were preprocessed through the following steps:

i. Rotate the face images to align them to the vertical orientation;
ii. Scale the face images so that the distances between the two eyes are the same for all images;
iii. Crop the face images to remove the background and hair region;
iv. Apply histogram equalization to the cropped face images for photometric normalization.

3. METHODOLOGY

The basic approach to the development of swarm-optimized aging-invariant face recognition system was discussed in this section.

3.1 Research Approach

Local Binary Pattern (LBP) and Gabor Wavelet Transform (GWT) were combined to realize an improved feature extraction method referred to as LBP-GWT feature extraction technique for the age-invariant face recognition system. Particle Swarm Optimization (PSO), an efficient feature selection algorithm suitable for face images (Shinde and Gunjal, 2012), was used to manage the curse of dimensionality of the pool of the initially-generated features by LBP and GWT to obtain an optimal age-
invariant feature subset to be used for recognition. Finally, a Support Vector Machine (SVM) was used for the final classification.

This research work comprises three (3) development phases:

i. Acquisition of probe and gallery images (frontal images) from the FG-NET aging dataset.

ii. Pre-processing of the images.

iii. Development of a LBP-GWT feature extraction technique.

The two post-developmental phases include:

(a) Evaluation of the developed LBP-GWT age-invariant feature extraction technique against LBP and GWT using false acceptance rate, false rejection rate, recognition accuracy and recognition time as performance evaluation metrics.

(b) Recognition of age variant faces using the SVM classification system.

The complete framework for the developed aging-invariant face recognition system was presented in Figure 1. The first step was the acquisition of the age variant images. The publicly available FG-Net aging dataset was used for this purpose. Each age-variant probe from the FG-NET dataset was preprocessed using histogram equalization. The images were pre-processed sequentially. Next stage was the extraction of age invariant features from the pre-processed probe using LBP and GWT techniques. The swarm-optimized LBP-GWT feature extraction technique was developed by combining features extracted by LBP and GWT techniques as a single feature set in a feature level fusion manner. Feature level fusion involves consolidating the feature sets obtained from multiple FET into a single feature set after normalization and transformation schemes. Stable features from eye lids, nose and lips of the fused LBP-GWT feature set that are resistant to factors affecting faces due to aging were selected using Particle Swarm Optimization (PSO) technique. These age-invariant features as they correspond to each individual (face ID) as well as the distance between these points were used to match the probe with the images in the gallery via a SVM classifier. The block diagram showing the processes involved in the training and testing stages of the developed Age-Invariant LBP-GWT face recognition system is presented in Figure 2.

3.2 The Developed LBP-GWT Feature Extraction Technique

In the developed LBP-GWT technique, LBP and GWT were used to extract local features used for identification. For the LBP, the local features corresponding to the eye lids, nose and the lips were extracted by conducting local binary pattern transformation to the whole face first. The transformed image of LBP values was then divided into 4x2 equal size horizontal and vertical blocks. The histogram of the uniform local binary patterns in each block was obtained. The rationale behind this local feature extraction method is that local binary patterns represent textures of a small local area and the histograms of uniform local binary patterns of the blocks tend to further capture the local textural features of different regions of a face. With the coordinates of the center pixel of an image \( I(x, y) \) defined as \((x_c, y_c)\), then the coordinates of his \( P \) neighbors \((x_p, y_p)\) on the edge of the circle with radius \( R \) can be calculated with the cosine rule:

\[
X_p = X_c + R \cos \left( \frac{2 \pi p}{P} \right)
\]

![Figure 1: Framework of the Developed LBP-GWT Aging-Invariant Face Recognition System](image1)

![Figure 2: The Block Diagram showing the processes involved in the Training and Testing Stages of the Developed Age-Invariant LBP-GWT](image2)
The algorithm is as follows:

Input: Training and Test Image set
i. Initialize temp = 0
ii. FOR each image I in the training image set
iii. Initialize the pattern histogram, H = 0
iv. FOR each center pixel t_c ε I
v. Compute the pattern label of t_c, LBP
vi. Increase the corresponding bin by 1.
vii. END FOR
viii. Find the highest LBP feature for each face image
ix. Apply particle swarm optimization for feature subset selection

Intermediate Output: Reduced LBP features of face image

In the same vein, the GWT was implemented as a process depicted in the algorithm as follows:

Input: Training and Test Image set
i. Convolve Image I (x, y) using Gabor wavelets to extract local features at these feature points
ii. Calculate the mean deviation, \( \mu_{mn} \), of the Gabor wavelet coefficients for each point
iii. Calculate the standard deviation, \( \sigma_{mn} \), of the Gabor wavelet coefficients for each point
iv. Construct Gabor feature vector using \( \mu_{mn} \) and \( \sigma_{mn} \).
v. Apply particle swarm optimization for feature subset selection

Intermediate Output: Reduced GWT features of face image

Repeat for all features

For each feature in LBP, choose a corresponding feature in GWT
Take average of each matching features in LBP and GWT
Apply sum rule fusion strategy
End Repeat

Computation of the fused feature set using sum rule fusion strategy is given by:

\[
F_n = \sum_i^n (F_{LPFV-opt} + F_{LGWV-opt}) / 2
\]

where \( F_n \) is the fused set of corresponding optimized low-dimensional LBP and GWT features, \( F_{LPFV-opt} \) is the optimal local pattern feature vector and \( F_{LGWV-opt} \) is the optimal local gabor feature vector. \( S_i \) is the similarity score obtained by computing the determinant of the fused set matrix \( F_n \). The best feature subset was selected using PSO by the equation:

\[
P_{i, best} = P_i \text{ if } f(P_i) > f(P_{i, best})
\]

\[
g_{i, best} = g_i \text{ if } f(g_i) > f(g_{i, best})
\]

The best feature subset was selected by PSO using the equation:

\[
P_{i, best} = P_i \text{ if } f(P_i) > f(P_{i, best})
\]

The selected features, parameter values, and training dataset were used to build SVM classifier. The value of \( n \) variables ranges between 0 and 1. If the value of a variable is less than or equal to 0.5, then its corresponding feature is filtered off using the equation:

\[
W^* \cdot x_i - b \geq 1 \text{ if } y_i = 1
\]

\[
W^* \cdot x_i - b \leq -1 \text{ if } y_i = -1
\]

Conversely, if the value of a variable is greater than 0.5, then its corresponding feature is chosen. PSO was applied to optimize the feature subset selection and classification parameters for SVM classifier. It eliminates the redundant and irrelevant features in the dataset, and thus reduces the feature vector dimensionality drastically. This helps SVM to select optimal feature subset from the resulting feature subset. This optimal subset of features was then adopted in both training and testing to obtain the optimal outcomes in classification.

### 3.3 Evaluation of the Developed LBP-GWT Age-Invariant Feature Extraction Technique

The performance evaluation metrics that were used to evaluate the developed feature extraction technique are:

- The False Accept Rate (FAR): This is the percentage of probes a system falsely accepts even though their claimed identities are incorrect (Raghavender, 2008).

\[
FAR = \frac{\text{Number of false accepts}}{\text{Number of impostor scores}}
\]

- The False Reject Rate (FRR): This is the percentage of probes a system falsely rejects despite the fact that their claimed identities are correct. A false accept occurs when the recognition system decides a false claim is true and a false reject occurs when the system decides a true claim is false (Raghavender, 2008).

\[
FRR = \frac{\text{Number of false rejects}}{\text{Number of genuine scores}}
\]

- Recognition Accuracy: This is the main measurement to describe the accuracy of a recognition system. It represents the number of faces that are correctly recognized from the total number of faces tested (Jeremiah et al., 2012).

\[
\text{Recognition Accuracy} = \frac{(\text{Number of correctly recognized persons})}{(\text{Total number of persons tested})} \times 100\%
\]

- Recognition Time: This represents the time required to process and recognize all faces in the testing set.

### 4. RESULTS

Four face images of varying ages in each of the 82 subjects that make up the FG-NET aging data set were used as test dataset making a total of 328 tested face images. Also, ten face images of varying ages in each of the 82 subjects in FG-NET aging data set were used as train datasets making a total of 820 trained face images. All the algorithms were implemented in MATLAB 7.7.0 (R2008b) environment. The results obtained for GWT, LBP and the developed LBP-GWT is presented in Table 1. A recognition accuracy of 81.71% was obtained by PCA-LDA while Histogram of Gradient (HOG) obtained a recognition accuracy of 86.92% for age-invariant face classification.
However, the developed LBP-GWT technique showed improved results over these works as recognition accuracy of 93.6% was obtained. This obvious improvement was due to the fact that the existing systems were based on global feature extraction approaches which generally are less accurate compared to the local feature descriptors employed in this research work. In the same vein, the linear nature of PCA-LDA obtained a recognition time of 151.421s while HOG obtained a recognition time of 124.533s for age-invariant face classification. However, the developed LBP-GWT technique was the one with the least recognition time of 81.667s. This improvement is on the ground that the developed feature extraction technique combines the positive attributes of both LBP and GWT.

4.1 False Accept Rate

The graph showing the results of false acceptance obtained for the feature extraction techniques is presented in Figure 3. The developed FET produced the least false acceptance of 6 out of a total of 328 test images and as such the most reliable. On the other hand, LBP and GWT yielded false acceptance of 18 and 12 respectively while LDA-PCA and HOG produced false acceptance of 22 and 21 respectively. The considerably high rate of false acceptance in the existing systems is due to the fact that they rely on global features for identification. These features are not discriminating enough for recognition purpose in challenged datasets due to the holistic nature of such approaches (Jeremiah et al., 2012).

4.2 False Reject Rate

The graph showing the results of false rejection obtained for the feature extraction techniques is presented in Figure 4. The developed FET produced the least false rejection of 15 out of a total of 328 test images and as such the most accurate. On the other hand, LBP and GWT yielded false rejection of 32 and 26 respectively. HOG yielded false rejection of 27 while LDA-PCA yielded 38. The justification for this result is borne out of the research outputs by Kuldeep and Madan (2013) which ascertain that PCA deals with data directly without taking cognizance of the underlying class structure which often leads to misclassification when used for dimensionality reduction and classification as observed in the existing LDA-PCA technique. LDA-PCA technique also yields higher values of false rejection especially when the hyperplane is fooled as is the case of support vector machine.

4.3 Recognition Accuracy

The graph showing the results of recognition accuracy obtained for the feature extraction techniques is presented in Figure 5. LBP, GWT, PCA-LDA, HOG and the developed swarm-optimized LBP-GWT techniques produced recognition accuracies of 84.75%, 88.41%, 81.71%, 86.92% and 93.6% respectively. Hence, the developed technique showed remarkable improvement over others following recognition of same individual at different ages as contained in the gallery dataset. The high recognition rate produced by the developed swarm-optimized LBP-GWT confirms the assertion by Yu et al. (2009) that feature-level fusion of local feature descriptors using sum rule could potentially optimize the performance of the classifier towards improved accuracy and computational efficiency because local texture regions are spatially homogeneous and hence provides analysis of the input image in both spatial and frequency domains simultaneously.
4.4 Recognition Time

The graph showing the results of recognition time obtained for the feature extraction techniques is presented in Figure 6. In descending order of computational efficiency, the testing time of the FETs are 81.667s, 101.221s, 112.692s, 124.533s and 151.421s for the developed swarm-optimized LBP-GWT, LBP, GWT, HOG and LDA-PCA. This result confirms the report by Zhou et al. (2010) of LBP exhibiting low computational complexity and its local texture character which can be described efficiently makes it widely acceptable for feature extraction algorithm. Also, GWT has an optimal location in both frequency domain and the space domain (Ali, Hind and Raghad, 2012) which provides the optimal basis to extract local features in regions that are spatially homogeneous. However, the low computational overhead obtained by the developed swarm-optimized LBP-GWT was due to the fact that it combined the positive attributes of both LBP and GWT feature level fusion using sum rule.

In view of the above results obtained with respects to all metrics considered, the developed LBP-GWT features extraction technique has the best recognition time, recognition accuracy, FAR and FRR, followed by GWT, LBP, HOG and PCA-LDA in that order. It was observed that LBP exhibits lower computational time overhead and better off than the GWT. This result confirms the report by Zhou et al. (2010) of LBP exhibiting low computational complexity which makes it widely acceptable. In addition, HOG, the work of Dihong et al. (2013) is an improvement over PCA-LDA, the work of Huseyin and Osen (2012) especially in terms of all the aforementioned evaluation metrics.

The sample graphical user interface showing the test and the equivalent images returned using GWT, LBP and LBP-GWT are presented in figures (7,8 and 9) respectively.
5. CONCLUSIONS

A swarm optimized age-invariant feature extraction technique was developed to address low discrimination ability and high computational resource demand of most existing age-invariant face recognition system. The summarized result of all evaluations conducted for the feature extraction techniques showed that the developed LBP-GWT performed better than LBP, GWT, HOG and PCA-LDA as it produced the highest recognition accuracy, least false acceptance, least false rejection and least recognition time. However, the developed LBP-GWT technique was tested on FG-NET aging dataset which is a publicly available standard aging dataset which is a publicly available standard aging dataset for research purpose.

Face recognition across varying ages are still open problems; therefore, further research can be directed along the modeling and generation of artificial human faces as age progresses to help realize artificial aging dataset that can serve same purpose as real time aging datasets which is practically highly difficult to collect as it spans a very long period of time.

Furthermore, apart from the surface aging resistant features, additional features such as morphology or color could be considered. This will improve the matching accuracy with facial marks and enable more reliable face image retrieval. The face image retrieval system can be combined with other robust face matchers for faster search. Since each facial mark is locally defined, marks can be easily used in matching and retrieval given partial faces.

6. REFERENCES


