A SURVEY ON AGE-ININVARIANT FACE RECOGNITION METHODS

Zahra Mortezaie¹ and Hamid Hassanpour²

(Received: 9-Apr.-2019, Revised: 5-Jun.-2019, Accepted: 23-Jun.-2019)

ABSTRACT

Face recognition is used in many security and surveillance applications. Some issues, such as aging, partial occlusion, variation in pose and illumination and facial expression directly affect the performance of face recognition approaches. Usually, in many applications, such as checking the passport and visa, images in the database are not updated continuously. In these cases, aging leads to change the important features of the face image. Hence, face recognition across aging can be considered as a common issue in many security and surveillance systems. In this paper, some existing face recognition approaches, in terms of robustness to aging, have been reviewed briefly. Also, the experimental results from these methods have been compared using the common databases in age-invariant face recognition applications. The comparison results indicate that the approaches which consider both component-based representation of facial images and identity factors outperform the other existing methods.

KEYWORDS

Face recognition, Aging, Age-invariant, Facial image.

1. INTRODUCTION

Face recognition is considered as an effective and applicable identification and verification technique, as it is an inexpensive and non-intrusive approach [1]. It has been used in many applications, such as security systems, card verification, video surveillance, credit criminal identification, person identification [2], passport renewal, law enforcement and biometric authentication [3]. There are three stages in face recognition systems [1]. In face recognition, the facial image is detected firstly. Then, suitable features are extracted from the image. Finally, the extracted features are compared with those from the database using a similarity measure.

Accuracy of face recognition methods directly depends on the used facial image. Indeed, some variations in facial images, such as age, pose, illumination, partial occlusion and facial expressions, reduce the accuracy of face recognition methods [3]. Among these variations, aging leads to change the important features of facial images, such as texture and shape. In many security and surveillance applications, the available databases are not updated during the time. Also, some existing non-age-invariant methods may correctly recognize the captured facial images using images captured within a one-year interval. In the case of facial images captured within a longer period of time, many changes may appear in comparison to previously captured images, especially in cases in which the images include a person’s face before and after the age of puberty. Hence, it is necessary to consider the effect of aging in face recognition techniques [3]-[6].

There is some research on face recognition which considers age progression. This research can be categorized in three groups: age estimation approaches [7]-[20], aging simulation approaches [21]-[26] and age-invariant face recognition approaches [27]-[29], [31], [34]-[39], [41]-[47], [49]-[50]. In the first group, the age is estimated from the person’s facial image. In the second group, a computational model is provided for modeling the facial appearances across aging. In the third group, age-invariant features are extracted for robust face recognition. Generally, age-invariant face recognition methods can be divided into two groups: generative [27]-[29], [31] and non-generative methods [34]-[39], [41]-[44]. In the former group, aging is initially simulated using a computational model. Then, the perceived image of the subject is normalized to eliminate aging variations. But, age-invariant features are extracted from the facial images in the latter group [4]. Recently, new methods, such as deep neural networks, have been used in age-invariant face recognition approaches [45]-[47], [49]-[50].

Z. Mortezaie and H. Hassanpour are with Computer Engineering Department, Shahrood University of Technology, Shahrood, Iran. Emails: ¹zm.mortezaie@gmail.com and ²h hassanpour@yahoo.com
Deep neural networks, such as auto-encoder neural networks, can both extract features and model the face across aging. Hence, these methods can be categorized in both generative and non-generative face recognition methods across aging. Therefore, we can consider a new category in addition to generative and non-generative methods, named deep neural networks.

This paper aimed to review and compare the performance of the existing age-invariant face recognition methods. The rest of this paper is organized as follows. The age-invariant face recognition methods are reviewed in Section 2. The common existing databases for age-invariant face recognition approaches are introduced in Section 3. The experimental results from the reviewed methods are compared in Section 4. Finally, the conclusions are driven in Section 5.

2. AGE-INVARIANT FACE RECOGNITION

As mentioned earlier, age-invariant face recognition methods are generative or non-generative approaches and those which use deep neural networks. In this section, some existing age-invariant face recognition methods are reviewed briefly.

2.1 Generative Methods

In [27], a computational model has been proposed to describe the changes of shape and texture in the facial images across aging. In this method, a muscle-based geometrical change model has been proposed to model the aging progression in adulthood. In this model, the changes of physical properties and geometric orientations of the facial muscles have been described throughout adulthood. Also, in this method, the facial wrinkles and other skin traits that appeared across aging were modeled by proposing the image gradient-based texture transformation function.

In [28]-[29], a 2D/3D face aging pattern space has been provided to synthesize a facial image to match the target face image before recognition. In [28], at first, the facial aging is simulated using a 2D face aging model. Then, the aging model is used for face across aging. This model is proposed based on non-negative matrix factorization (NMF) [30] with sparseness constraints.

In the method introduced in [31], the recognition is performed in two stages. At first, a maximum a posteriori solution is used based on principal component analysis (PCA) factorization in order to decrease the search space. Then, a graph matching approach is used to investigate the matching between the probe image and the gallery image. Indeed, in [31], the facial images are presented using graph-based features. In this method, the aging of each face image is modeled using Gaussian mixture model (GMM). This model has considered the shape and texture information variations for each face image across aging. In this method, the feature points are extracted using a modified local feature analysis [32]. The modified local feature analysis uses the Fisher score [33] to extract the feature points. The feature descriptor is obtained via applying the uniform local binary pattern (LBP) operator on the feature point.

As the aging is a complex process, generative methods are not usually able to create the face model representing the aging process. Also, the limitation on the number of training data increases this problem. Besides, in these methods, extra information, such as accurate age labels for the training data and the landmark point locations for each face image, are required to create the face aging model. Also, in these methods, the used face images should be captured in ideal conditions, such as frontal pose, normal illumination and expression [5]. So, generative methods do not work well in real-world face recognition. Hence, in the age-invariant face recognition approaches, non-generative methods have been extended more than generative methods.

2.2 Non-Generative Methods

In [34], an age-invariant face recognition method was proposed based on the Eigen space techniques and the Bayesian model. In this method, a Bayesian age-difference classifier was built on a probabilistic Eigen space framework. This classifier is used to model the differences between face images across aging for individuals as well as to model the differences between various persons face images.

The gradient orientation information is a more robust feature in comparison with the other appearance
features in changing the age. Hence, in [35]-[36], the gradient orientation pyramid (GOP) feature was used to model the differences between person faces. In this method, the GOP feature was combined with the support vector machine (SVM) for face verification.

In [37], a discriminative model has been proposed considering the scale-invariant feature transform (SIFT) and multi-scale local binary patterns (MLBPs) as the local features for each face. The SIFT-based local features and MLBP-based local features have a high-dimensional feature space. Hence, in this method, an algorithm named multi-feature discriminant analysis (MFDA) was developed to reduce dimensionality. In MFDA, local descriptors were combined to create a robust decision rule by a random subspace fusion model.

The number of the considered face features can be an effective parameter in extracting the appropriate discriminative information. Hence, in [38], a multi-view discriminative learning (MDL) method was proposed for age-invariant face recognition. In this method, the SIFT, LBP and GOP descriptors have been used to extract the discriminative information from the facial images. Then, a discriminative learning method with multi-view feature representations; namely, MDL, was developed. Indeed, the MDL has been used to simultaneously minimize the variations in each feature class, maximize the variations between the feature classes and maximize the correlation of the feature classes from the same person.

Age variation affects each component of the face differently. Hence, in [39], a component-based method has been proposed for age-invariant face recognition. In this method, the components of the face are determined automatically using an active shape model (ASM) [40]. Then, the MLBP and SIFT features for each component are used in a random subspace linear discriminant analysis (LDA) for classification.

In [41], at first some pre-processing tasks, such as pose correction, illumination and periocular region normalization, are applied on the facial images. Then, the Walsh-Hadamard transform encoded local binary patterns (WLBP$s$) are applied on the pre-processed periocular region. Finally, unsupervised discriminant projection (UDP) is used to create the subspaces on WLBP-featured periocular images.

Components of face images across aging can be considered as the age-invariant and age-variant factors. Following this idea, a hidden factor analysis (HFA) model was proposed in [42]. In this method, two latent factors were introduced: an identity factor (age-invariant) and an age factor (age-variant). A linear model was considered to distinct these factors as two subspaces. Also, a learning algorithm was developed using an Expectation Maximization (EM) algorithm to jointly estimate the latent factors and the model parameters. Finally, the cosine distance of the identity components of the gallery and the probe samples are used for face recognition. Note that in [42], the HOG descriptor is used to extract features of face images. Also, due to the existence of noises in the face images, the extracted features are noisy. Hence, the accuracy of identity factor estimation may be reduced. Therefore, in [43], another descriptor, named the maximum entropy feature descriptor (MEFD), has been proposed to improve the representation of face images across aging. This descriptor encodes the microstructure of facial images into a set of discrete codes in terms of maximum entropy in order to extract the appropriate features of the facial images. Then, the method named identity factor analysis (IFA) is used to determine the probability that two face images are related to the same person.

The assumption of independence between the identity and age factors is incorrect, because changes in the face appearance due to aging may vary in different individuals. Hence, in [44], by modifying the HFA method proposed in [42], a probabilistic discriminant method has been proposed to make a better estimate of the identity factor. In this approach, a latent factor is modeled considering the correlation between age and identity factors. Indeed, the person-specific aging information and some features, such as pose and expressions which affect the face recognition task, are jointly modeled by the latent factor. Finally, face recognition is performed using a maximum likelihood approach.

### 2.3 Deep Neural Network Approaches

As the age variation is a nonlinear and smooth transformation, in [45], a neural network model, named coupled auto-encoder network (CAN), was proposed to overcome the aging in face images. In the proposed CAN, a couple of two auto-encoders have been bridged using two shallow neural networks. Also, in this paper, facial images are decomposed into three components: identity feature, age feature
and noise, using a proposed nonlinear factor analysis method. Among the obtained features from the facial image, the identity feature has been considered as an age-invariant feature to use in the face recognition task.

In general, deep networks have a better performance than shallow networks. The proposed approach in [46] has shown the application of convolutional neural networks (CNNs) in age-invariant face recognition. In this paper, a latent factor guided CNN (LF-CNN) framework was proposed to learn the age-invariant deep face features. The age-invariant deep features are extracted from convolutional features using a designed fully connected layer, named latent factor fully connected (LF-FC) layer. A latent variable model, named latent identity analysis (LIA), was developed to divide the convolutional features into aging (age-variant) and identity (age-invariant) components. The parameters of LF-FC layer are updated using the parameters of the LIA model.

The facial aging process can be well understood considering demographic estimation of facial images, such as estimation of age group, gender and race. Hence, a demographic-assisted face recognition approach was proposed in [47]. This method consists of two main steps: facial-asymmetry-based demographic estimation step and demographic-assisted face recognition step. In the first step, three CNNs are trained for each age group, gender and race classification task in order to extract the demographic features from the query face image. In the second step, at first, deep CNN (dCNN) features are extracted from the query image using VGGNet [48]. Then, dCNN features are used to obtain top k matches from gallery against the query face image. Finally, the top k matched face images are re-ranked using the estimated demographic features.

An age estimation guided convolutional neural network (AE-CNN) has been proposed in [49]. In this approach, an age estimation process is used to separate age-invariant features from those affected by aging. Hence, for extracting the age-invariant features, the AE-CNN contains three fully connected layers. General (age-invariant and age-variant) features are the output of the first fully connected layer. In the second fully connected layer, the softmax loss function is used for estimating the age. The output of this layer is the age features. Also, the third fully-connected layer is used to extract age-specific factors. Finally, age factor is subtracted from the general features to obtain age-invariant features. After that, the softmax loss function is used to recognize the face considering the obtained age-invariant features.

Note that in some face recognition methods, the extracted features may be projected in another space to obtain the comparable features. Hence, in [50], a similarity measure and a distance metric optimization-driven learning approach have been proposed in order to preserve the interaction between the projected features and the used similarity measure. In this method, feature leaning and distance metric learning have been performed at the same time, using a deep convolutional neural network (CNN). To train the CNN, a large number of positive (matched) and negative (unmatched) pairs are generated given the labeled training images. Also, to learn the features which are more independent on aging, the matched pairs are selected considering different ages for each person. Then, using the positive and negative pairs, the CNN is trained subject to reduction and increase of the difference, respectively, between the matched and unmatched pairs. Also, during training the CNN, the parameters of the model are tuned using the mini-batch stochastic gradient descent (SGD) algorithm.

3. AGE INVARIANT FACE RECOGNITION DATABASES

In this section, the databases which are often used in age-invariant face recognition research are briefly introduced. These databases are FGNET [51], MORPH [52], CACD [53], CACD-VS [54] and AgeDB [55]. The FGNET and MORPH databases are commonly used in age-invariant face recognition approaches. Recently, the CACD, CACD-VS and AgeDB databases were introduced for face recognition across aging.

**FGNET:** The FGNET database includes 1,002 face images from 82 different people. The age range in this database is from 0 to 69 years. The average number of existing face images for each person is approximately 12. The age gap in this database is 0-45 years. Some samples of the face images from this database are shown in Figure 1. In this figure, the age of the persons for whom the face image has been captured is mentioned below the image.
MORPH: The MORPH database includes 55,134 face images from 13,618 different people. The age range of this database is from 16 to 77 years. The average number of existing face images for each person is approximately 4. The age gap in this database is 0-5 years. Some samples of the face images from this database are shown in Figure 2. In this figure, the age of the person for whom the face image has been captured is mentioned below the image.

The MORPH database has been extended in two albums. Album 1 consists of 1,690 face images from 632 different people. The age range of this album is from 15 to 68 years. The average number of existing face images for each person is approximately 3. Album 2 consists of 78,207 face images from 20,569 different people. The age range of this album is from 15 to 77 years. The average number of existing face images for each person is approximately 4.

CACD: The CACD database includes 163,446 face images from 2,000 different people. The age range of this database is from 16 to 62 years. The average number of existing face images for each person is approximately 82. The age gap in this database is 0-10 years. Some samples of the face images from this database are shown in Figure 3. In this figure, the numbers in the top row are the birth years of the people and the numbers in the left column are the years in which the images have been captured.

Figure 1. Sample face images from FGNET. The number below each facial image represents the age in years.

Figure 2. Sample face images from MORPH. The number below each facial image represents the age in years.

Figure 3. Sample face images from CACD.
CADC-VS: The CADC-VS database is a subset of CACD database which is used for face verification. It contains 4,000 images from 2,000 persons across ages. Indeed, in this database, two images (a positive image and a negative image) exist for each person. In other words, CADC-VS dataset includes 2,000 positive pairs (images from same persons) and 2,000 negative pairs (not from the same persons).

AgeDB: This database contains 16,488 images from 568 different people captured in real-world conditions. Hence, the images of this database may include different poses and expression, noise, occlusions and any uncontrolled conditions. Hence, this database would be used for training and testing the deep neural networks. Also, the age range of this database is from 1 to 101 years. The average number of existing face images and the average age range for each person are 29 and 50.3 years, respectively. Figure 4 shows some sample images from this database.

Figure 4. Sample face images from AgeDB.

4. PERFORMANCE COMPARISON

In this section, the experimental results from some existing age-invariant face recognition methods are reviewed on the FGNET and MORPH databases. The performance of these methods is compared considering the reported Rank-1 recognition accuracy. In general, Rank-k (k = 1, 2, 3, …) is one of the measures used for depicting the performance of face recognition methods. Usually, the recognition rate of the compared methods is considered as a function of Rank-k. Indeed, for ranking the obtained results from a face recognition method, k determines that k-top matches the correct answer [56]. Therefore, for k = 1, Rank-k is the strictest measure; whereas, for k > 1, this measure permits some error. In Table 1, the performance of these methods is compared considering the reported Rank-1 recognition accuracy.

As mentioned before, generative aging models need parameter assumption, landmark point location for each face image and the captured face images in the controlled conditions. Hence, these methods do not work well in real-world face recognition. In practice, the performance of the generative model proposed in [29] is lower than for the other methods mentioned in Table 1.

The method proposed in [37] considers a holistic representation of the facial images. However, aging affects the different components of the face variously. Hence, component-based facial image representation has been proposed in [39], [41]. As it is shown in Table 1, the component-based facial image representation proposed in [41] has achieved 100% Rank-1 recognition accuracy on FGNET database.

Also, separating aging variations from the person-specific features is the main idea in some non-generative methods [42]-[44] to obtain robust age-invariant face features. In these methods, the basic features which are used for estimating the age-invariant and age-variant factors play an important role in the final results. As mentioned before, in the method proposed in [43], a MEFD descriptor has been introduced to be used instead of HOG. Hence, according to Table 1, this method achieves better accuracy in comparison with the similar methods proposed in [42], [44].

As mentioned before, the methods proposed in [45]-[47], [49]-[50] are the recently developed approaches in age-invariant face recognition. In these methods, using the deep learning approaches leads to obtain the deep features in the facial images which are aging-invariant.

Age variation is a nonlinear and smooth transformation. However, in [42]-[44], a linear factor analysis method is used in order to separate the age-variant and age-invariant factors from the face images. Hence, this method has lower performance when compared to the methods proposed in [45]-[46] and [49]-[50] (see Table 1). Besides, the kind of basic descriptor used for feature extraction in [42]-[44] can be the other reason why the results of these methods are not better than those of deep learning-based approaches.
Also, the methods proposed in [46], [49] have better performance in comparison with the method proposed in [45]. Indeed, the unsupervised CAN model proposed in [45] is less discriminative than the supervised LF-CNNs and AE-CNN proposed in [46], [49]. Also, the deep networks used in [46], [49] have better performance than the shallow networks used in [45].

Apparently different components of the face are affected variously across aging progression. Hence, age-invariant face recognition approaches which consider this issue outperform the other approaches. Besides, separating the identity and aging factors of the facial images increases the accuracy of the face recognition approaches across aging. Indeed, we can conclude from Table 1 that face recognition approaches will recognize the face more accurately compared to the other methods, if they consider component-based representation of facial images and consequently estimate the identity factors from aging factors in each component. Also, deep learning techniques have emerged in machine vision applications [57] and recently, age-invariant face recognition methods tend to use deep learning approaches. Hence, the capability of these neural networks in feature extraction and modeling can be employed to estimate the identity and aging factors in each component of facial images. Therefore, a combination of the mentioned approaches may lead to an appropriate performance of age-invariant face recognition approach.

Table 1. Experimental results of the existing age-invariant face recognition methods.

<table>
<thead>
<tr>
<th>The method proposed in</th>
<th>Databases (# subjects, # images) in probe and gallery</th>
<th>Reported Rank-1 recognition accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>[29] FGNET (82, 1002)</td>
<td>37.4%</td>
<td></td>
</tr>
<tr>
<td>MORPH album 1 (612, 612)</td>
<td>66.4%</td>
<td></td>
</tr>
<tr>
<td>MORPH album 2 (10000, 20000)</td>
<td>79.8%</td>
<td></td>
</tr>
<tr>
<td>[37] FGNET (82,1002)</td>
<td>47.50%</td>
<td></td>
</tr>
<tr>
<td>MORPH album 2 (10000, 20000)</td>
<td>83.90%</td>
<td></td>
</tr>
<tr>
<td>[39] MORPH album 2 (10000, 20000)</td>
<td>81.27%</td>
<td></td>
</tr>
<tr>
<td>[41] FGNET (82, 1002)</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>[42] FGNET (82, 1002)</td>
<td>69.0%</td>
<td></td>
</tr>
<tr>
<td>MORPH album 2 (10000, 20000)</td>
<td>91.14%</td>
<td></td>
</tr>
<tr>
<td>[43] FGNET (82, 1002)</td>
<td>76.2%</td>
<td></td>
</tr>
<tr>
<td>MORPH album 2 (10000, 20000)</td>
<td>93.80%</td>
<td></td>
</tr>
<tr>
<td>[44] FGNET (82, 1002)</td>
<td>72.8%</td>
<td></td>
</tr>
<tr>
<td>MORPH album 2 (10000, 20000)</td>
<td>87.94%</td>
<td></td>
</tr>
<tr>
<td>[45] FGNET (82, 1002)</td>
<td>86.5%</td>
<td></td>
</tr>
<tr>
<td>FGNET (82, 1002)</td>
<td>88.1%</td>
<td></td>
</tr>
<tr>
<td>MORPH album 2 (10000, 20000)</td>
<td>97.51%</td>
<td></td>
</tr>
<tr>
<td>[46] FGNET (82, 1002)</td>
<td>88.1%</td>
<td></td>
</tr>
<tr>
<td>MORPH album 2 (10000, 20000)</td>
<td>98.13%</td>
<td></td>
</tr>
<tr>
<td>[50] MORPH album 2 (10000, 20000)</td>
<td>93.6%</td>
<td></td>
</tr>
</tbody>
</table>

5. CONCLUSIONS

Face recognition has applications in many machine vision fields, such as law enforcement and forensic investigation, homeland security, missing persons and checking the passport and visa. In these applications, the images in database may not be updated continuously. As aging slowly leads to changing the important features of the facial image, it directly affects the performance of face recognition approaches. Hence, it is necessary to consider the aging variations in face recognition. There are some research studies which have considered the face aging progression. These approaches have tried to estimate and simulate aging and recognize the face. In this paper, some age-invariant face recognition approaches have been briefly reviewed. Also, the experimental results of these methods have been compared on a number of databases. The experimental results show that age-invariant face recognition approaches which consider both component-based representation of facial images and identity factors have better performance compared with the other face recognition methods across aging.
REFERENCES


T. Cootes, "FGNET Face and Gesture Recognition Database," Face and Gesture Recognition Working Group, [Online], Available at: http://www-prima.inrialpes.fr/FGnet/


This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).