Artificial Intelligence Review (2019) 52:981–1008
https://doi.org/10.1007/s10462-018-9661-z

Age invariant face recognition: a survey on facial aging databases, techniques and effect of aging

Manisha M. Sawant[1] · Kishor M. Bhurchandi[1]

Published online: 13 October 2018
© Springer Nature B.V. 2018

Abstract
Age invariant face recognition (AIFR) is highly required in many applications like law enforcement, national databases and security. Recognizing faces across aging is difficult even for humans; hence, it presents a unique challenge for computer vision systems. Face recognition under various intra-person variations such as expression, pose and occlusion has been an intensively researched field. However, age invariant face recognition still faces many challenges due to age related biological transformations in presence of the other appearance variations. In this paper, we present a comprehensive review of literature on cross age face recognition. Starting with the biological effects of aging, this paper presents a survey of techniques, effects of aging on performance analysis and facial aging databases. The published AIFR techniques are reviewed and categorized into generative, discriminative and deep learning methods on the basis of face representation and learning techniques. Analysis of the effect of aging on the performance of age-invariant face recognition system is an important dimension. Hence, such analysis is reviewed and summarized. In addition, important facial aging databases are briefly described in terms of the number of subjects and images per subject along with their age ranges. We finally present discussions on the findings, conclusions and future directions for new researchers.

Keywords Age invariant face recognition · Effects of aging · Facial aging · Generative method · Discriminative method · Deep learning

1 Introduction

Automatic face recognition is an important yet challenging problem. The challenges are mainly due to large facial appearance variations within a subject and similarities between subjects. Intra-subject variations widely studied in face recognition are pose, expression,

✉ Manisha M. Sawant
mpparlewar@gmail.com

Kishor M. Bhurchandi
bhurchandikm@yahoo.co.in

[1] Department of Electronics and Communication Engineering, Visvesvaraya National Institute of Technology, Nagpur, India

Fig. 1 Face images of an individual displaying variation due to aging

occlusion and most importantly aging. Recently, due to development in feature representation (Ahonen et al. 2006; Dalal and Triggs 2005; Ojala et al. 2002) and machine learning methods (Sungatullina et al. 2013; Yang et al. 2007) substantial progress has been achieved (Chai et al. 2007; Huang et al. 2000; Parlewar et al. 2016; Patil et al. 2016) in this field. A few survey papers addressing pose invariant, expression invariant, 2D and 3D face recognitions are available in the literature (Abate et al. 2007; Ding and Tao 2016; Patil et al. 2015). Among the above mentioned face recognition methods, the recognition across aging is interesting and a challenging problem. One of the approaches to tackle the problem is frequent update of facial templates through a data acquisition process (Uludag et al. 2004). However, updating facial templates frequently is not possible in all applications and is also time consuming task. Developing machine learning algorithms to solve aging related face recognition problem is a better alternative. Compared to other appearance variations, age related changes depend on intrinsic factors and extrinsic factors. A few examples of intrinsic factors are lifestyle, stress, diseases etc. and that of the extrinsic factors are the environment, smoking and exposure to sun (Farage et al. 2008; Wen et al. 2016). Example of aging effects on facial appearance of an individual is shown in Fig. 1. The contributions from the computer vision and machine learning community in the field of facial aging are broadly classified into the following four tasks: (1) age synthesis, (2) age estimation, (3) Age Invariant Face Recognition (AIFR) and (4) effect of aging on performance of face recognition system.

Age synthesis or age progression, is defined as aesthetically rendering a face image with natural aging and rejuvenating effects for a certain face of an individual. Age progression has been used in finding missing children, identifying suspects who commit crimes at different ages in their lives, and detecting multiple enrollments for issuing government documents (e.g. passport and driver licenses). Age estimation refers to estimating age of the person based on his/her face image. Age estimation has potential applications in real world, such as age based access control (e.g. tobacco vending machines), and human computer interaction (e.g. display of advertisements on digital signboards based on the age of the person in the premises).

Age invariant face recognition refers to an automatic face recognition technique that is able to achieve temporal invariance over a wide range of the subject's age. In real time face recognition applications (e.g. passport renewal, law enforcement and biometric authentication), the AIFR system should be able to recognize faces despite of the appearance variations due to aging (Fu et al. 2010). Analyzing effects of aging on the performance of AIFR systems is one of the important areas in facial aging. For such analysis, the database is partitioned into

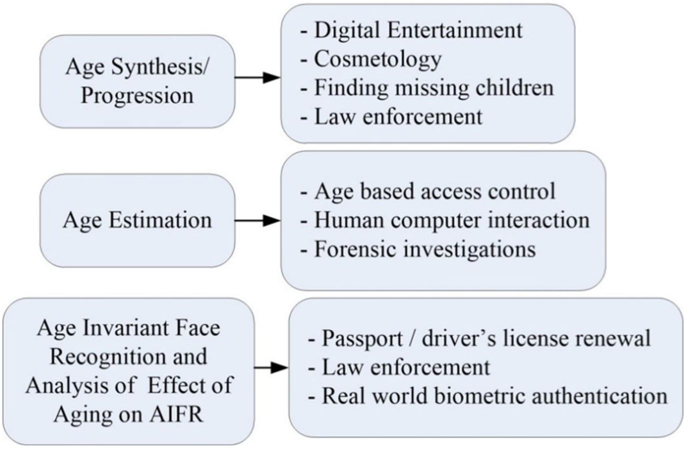

Fig. 2 Various facial aging related research areas and their applications

different age groups and the effect of aging in terms of the performance measures has been reported independently for each age group. Based on such analysis various conclusions are reported about changes in recognition rate per age group. The research in the field of facial aging and their corresponding applications are summarized in Fig. 2.

Two survey papers are available in the field of facial aging (Fu et al. 2010; Ramanathan et al. 2009). Age progression and age estimation methods are surveyed in Fu et al. (2010). Whereas, a few face verification experiments across aging along with age progression and age estimation are reported in Ramanathan et al. (2009). Our paper is principally different from both the previously published surveys. Our main focus is on the age invariant face recognition methods, effect of aging on the performance of AIFR systems and facial aging databases (which is not presented in Fu et al. 2010; Ramanathan et al. 2009). We survey existing AIFR methods and present their comparison in terms of face representation, learning methods, databases, evaluation protocols and recognition rate. We have also surveyed the effect of aging on the performance of AIFR systems. Initial works on the effect of aging on performance of AIFR has been addressed in Lanitis (2009b), Lanitis et al. (2004). But it only uses FG-NET database (FGNET 2010). In this paper, we have included existing AIFR techniques, various aging databases and analysis of effect of aging on the performance. Thus, the presented survey is much more exhaustive and novel in terms of approach and contents.

Although, the real-world applications of AIFR are interesting, the major challenge in this field is appearance variations due to aging. In this paper, we try to answer following questions. What are the existing techniques? Which are the milestones? What can be done further? We also discuss various existing challenges in AIFR. Our aim is to provide integrated analysis of the published research on age invariant face recognition and highlight the future scope. Specifically, we aim to integrate following topics:

1. Facial Aging databases,
2. Age Invariant Face Recognition methods,
3. Effect of aging on performance of face recognition systems,

We also provide discussion on the existing work and future scope in this topic along with conclusion.

The remaining sections are organized as follows. The facial aging databases are discussed in Sect. 2. Cognitive studies, generative and discriminative approaches along with deep

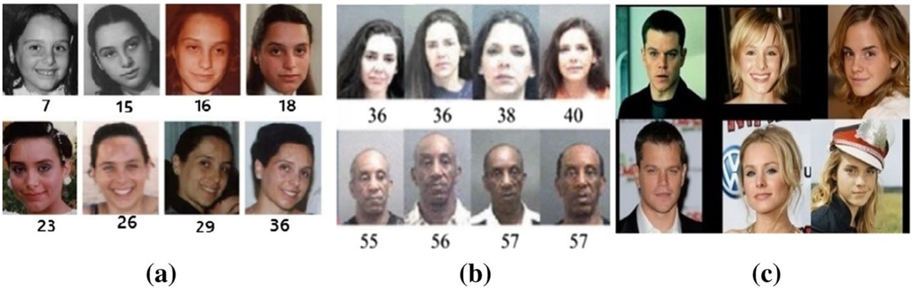

Fig. 3 Facial aging databases: **a** FG-NET contains large age variations, **b** MORPH-2 is widely used databases in AIFR and **c** CACD is largest facial aging databases

learning approaches are presented in Sect. 3. Effect of aging on the performance of face recognition system is discussed in Sect. 4. Discussion on recent achievements and existing challenges are presented in Sect. 5. Finally, in Sect. 6, we provide conclusion based on this survey and future scope of research.

2 Facial aging databases

Databases play a crucial role in research for benchmarking the face recognition algorithms. Many face datasets are available, but only a few of them are specially designed to address the aging problem. Constructing an aging database is a laborious and lengthy task as it requires collecting age-separated face images of the same individual under constrained conditions. Despite its importance, recognition robust to age variations has gained attention relatively late. It is mainly due to lack of suitable and sufficient aging databases. As mentioned in Li et al. (2011) the desired attributes of a face aging database are: (1) large number of subjects; (2) large number of age separated face images of the same individuals; and (3) minimum variations in pose, illumination, expression and occlusion in the captured images. Earlier released face databases such as FERET (Phillips et al. 2000; Rizvi et al. 1998) includes little age variations. Next, two publicly available facial aging databases with significant age variations are FG-NET and MORPH (Ricanek and Tesafaye 2006). Recently, two new databases are released, and the first is compiled by Lacey et al. (Best-Rowden and Jain 2015, 2018) which is significantly large in length and breadth, named PCSO_LS and the other is collected by Chen et al. (2014a, b) named cross-age celebrity dataset (CACD). Database containing both age and weight information along with face images is presented in Singh et al. (2014) and is named as WhoIsIt database. Sample images from various publicly available facial aging databases are shown in Fig. 3. In this section, we briefly discuss publicly available facial aging databases that are used in AIFR.

2.1 FERET

Earlier released face databases such as FERET include little age variations. The database contains a total of 14,126 images that includes 1199 subjects and 365 duplicate sets of images. A duplicate set is a second set of images of a person whose images are already in the database but captured on different days. For some individuals, over two years had elapsed between their first and last sittings, while a few subjects were photographed multiple times. Due to very short term aging pattern in FERET, it is not used for evaluation of AIFR algorithms.

2.2 FG-NET aging database

The FG-NET aging database contains 1002 images from 82 different subjects (6–18 images per subject) collected by mostly scanning photographs of the subjects with ages ranging between newborns to 69 years old subjects. Depending on the resolution of the camera used to take photographs and the resolution of the scanner the images in the database possess considerable variation in the resolution and quality of images. Considerable variations in illumination, pose, occlusion and expressions are observed in FG-NET database. Along with these variations, the description about the annotated landmark points is also recorded for each image. The major limitation of FG-NET is less number of subjects. Other issues related to FG-NET are large variations in pose, expression and image quality. Despite large age range, 50% of the FG-NET images are younger than 13 years.

2.3 MORPH

MORPH is much larger database than FG-NET. It comprises two sets; Album 1 and Album 2. Album 1 is relatively smaller than Album 2, and contains about 1690 face images of 625 individuals in the age range 15–68 years. The images in MORPH represent an adverse population with respect to age, gender, and ethnicity. The second part of MORPH is Album 2, which contains 78,207 age separated images of about 20,569 different subjects. Even though the total number of subjects in Album 2 is large, the number of face images per subject is small. Average number of face images per subject is 4.

Along with age variation, other variations such as expression, illumination and resolution are also observed in MORPH. This information is included as Meta data. MORPH is larger than FG-NET, and has a different age distribution since it is comprised primarily of adults and contains no image below 16 years age. Compared to the FG-NET, number of images per person in MORPH is still less. There are only 317 subjects which have at least 5 images acquired over at least 5 years elapsed time.

2.4 Cross-age celebrity database (CACD)

One more publicly available large scale face aging dataset is collected by Chen et al. named CACD. It contains 163,446 images of 2000 subjects. This dataset consist of celebrity images across 10 years from 2004 to 2013. The images are collected from Google Image Search using celebrity name and year as keywords. Metadata in terms of name, age, identity, birth year alogwith the information about presence of same image in LFW database is additionally provided.

2.5 Pinellas County Sheriff's Office Longitudinal Study (PCSO_LS)

Lacey et al. compiled a new longitudinal database of face images; named as PCSO_ LS. There are 1.5 million images in the database, which are collected from the mug shot of 18,007 criminals arrested by the Pinellas County Sheriff's Office (PCSO). The criterion followed while compiling the images is that each subject has at least 5 face images collected over at least a 5 year time span. In this database, consecutive images of a subject are age separated by at least one month and almost 13 subjects have more than 30 images. Average number of images per subjects is 8, and overall age span is from 18 to 83. However, it does not contain images of subjects with age range 0–15 years.

2.6 WhoIsIt (WIT)

Database containing both age and weight information of the facial images is presented in Singh et al. (2014) and named WhoIsIt database. WIT database contains 1109 images of 110 subjects, and the images in it are collected from internet.

Along with the above mentioned facial aging databases, recently a few facial aging databases are developed by various researchers. These databases include FACES (Ebner et al. 2010), Gallagher (Gallagher and Chen 2009) and ADIENCE (Eidinger et al. 2014) datasets. FACES dataset contains total 1026 images of 171 subjects with six expressions (neutral, sad, disgust, fear, angry, and happy). All the images are frontal with fixed illumination. The age range in this dataset is from 19 to 80. Although, this dataset covers wide age range, it does not contain the age-separated images per subject. It is an expression variation dataset with subjects from different age range. Hence, this dataset is suitable to analyze effect of expression variations on facial aging and age estimation. Gallagher and Chen proposed a database, which is collection of images from the group photos. Photos in this collection present multiple subjects, facing towards the camera. Since each photo contains multiple subjects, each face image cropped from the photo is of relatively low resolution. This dataset contains 5080 photos with 28,231 face images. It covers the age span from 0 to 66.

ADIENCE database is collection of true 'in wild' face images across large age range. It contains 26,580 images of 2284 subjects. In this dataset variation such as appearance, noise, pose, and lighting are present. In addition, many faces have low resolution. These factors make this dataset very challenging. Similar to FACES this dataset also does not contain age-separated images of subject and therefore is not suitable for AIFR. Other facial aging databases generally used for age estimation are presented in Fu et al. (2010).

Comparative statistics of the publically available face aging databases are presented in Table 1. Among existing aging datasets, FG-NET and MORPH are used as baseline in most of the AIFR approaches for benchmarking. Recently developed CACD and PCSO_LS databases are the large-scale aging databases. Databases like FACES, Gallagher and ADIENCE contains images of large age span, but they do not have age separated images of the subject. Main requirement of database for AIFR is large number of subjects covering large age span of each subject. Thus, FACES, Gallagher and ADIENCE are not much suitable for AIFR, but they are useful for analysis of effect of aging and age estimation.

3 Age invariant face recognition techniques

Initial works on aging models and age estimation are based on anthropometric measurements and craniofacial growth models. Therefore, it is essential to study effect of aging on face and preliminary works on facial aging. In this section, we first briefly discuss cognitive studies on facial aging. Next, we present various AIFR techniques.

Table 1 Statistics of facial aging databases

Sr. no.	Database	Description			
		#Subjects	#Images	# Images per subject	Age range (years)
1	FG-NET	82	1002	Max—12	0–69
2	MORPH-1	632	1690	Average—4	15–68
3	MORPH-2	20,569	78,207	Average—4	15–77
4	PCSO_LS	18,007	147,784	Average—8	18–83
5	CACD	2000	163,446	Average 81.7	16–60
6	WIT	110	1109	10–12	1–80

3.1 Cognitive studies on facial aging

Facial aging is a complex process that affects the appearance, shape and texture of the face. Compared to other intra-person facial variations, age related changes are dependent on intrinsic as well as extrinsic factors, such as, environment, lifestyle, exposure to sun, stress, diseases etc. (Farage et al. 2008). Different aging patterns are observed during formative years (1–18 years), in adults and seniors. While muscle and bone structure changes are dominant in childhood (below 18 years), lots of textural changes are observed in adults. The major factor which changes the appearance of an individual during the formative years is the craniofacial development (Albert et al. 2007; Ricanek et al. 2008). The bone structure does not change considerably after the person is fully grown. The aging process of adult humans is mainly characterized by the wrinkles, retrusion, sagging skin, eye slopes etc. (Farage et al. 2008).

D'Arcy Thompson proposed initial work on shape change due to growth (Thompson 1917). He discussed that the shape change due to growth or evolution results from physical forces in an animal's environment. These changes are modeled by deformation induced by strain operating on the original structure. The mathematical transformation required to achieve the shape change in human faces due to growth is presented in Mark et al. (1981). They modeled craniofacial growth using the concepts of transformational and structural invariants proposed by Thompson. The head of a newborn has a large cranium and less facial area, but during development, face grows faster than the cranium and this effect can be modeled by geometric transformations. Experimentation in Mark et al. (1981) revealed that, cardioidal strain transform and affine shear transform are able to produce facial effects similar to craniofacial growth. The results showed the cardioidal transform is remarkably better than the affine shear transform for perception of growth. The 'revised cardioidal strain transform' is used to model craniofacial growth.

Face anthropometry is the science that deals with the measurement of the human faces, which can be used to develop aging models. In anthropometric measurements, initially particular locations on a subject (human face), called landmark points are identified. The measurements taken according to the anthropometric profile of certain population characterize the distinctive features of similar faces in that population. Farkas (1994), Farkas and Munro (1987) proposed widely used measurements for describing a human face. In his system, 57 landmark points describe face. Anthropometric measurements provide a quantitative description of the craniofacial complex, which are further used for characterizing facial growth. The proportion indices are defined as ratios of distances between facial landmarks. These proportion indices and age-based measurements are used to build the craniofacial

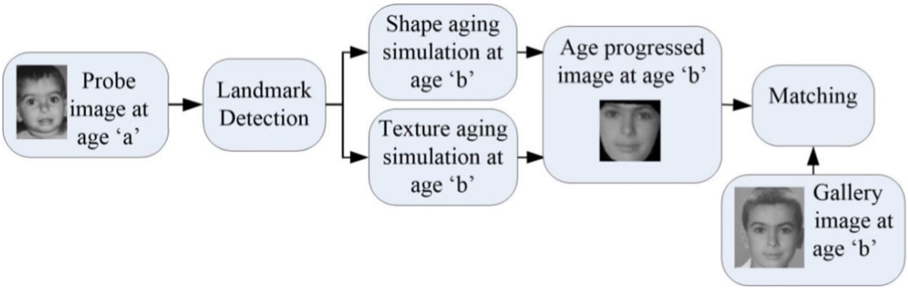

Fig. 4 Framework for generative AIFR

growth model. Anthropometric measurements based facial aging models are proposed in Geng et al. (2007), Lanitis et al. (2002), Ramanathan and Chellappa (2006, 2008) for face verification across age progression.

Age invariant face recognition has so far been broadly categorized into three approaches: (1) Generative, (2) Discriminative and (3) Deep Learning. The generative approach is based on age progression methods to transform the probe image to the same age as that of gallery image. After age normalization, conventional face recognition (FR) module is used for feature extraction and face matching. Various facial aging models for age progression are available in the literature; readers may refer Lanitis (2009b), Ramanathan et al. (2009) for the detailed description. Discriminative approaches handle recognition without age progression; instead, they rely on local feature descriptors to encode visual clues. These local features obliquely, but rigorously, represents age invariant signatures. Discriminative learning methods are further developed for feature matching in AIFR. In recent years, Convolutional Neural Networks (CNNs) have emerged as a powerful machine learning model. Unique feature extraction and face reconstruction capabilities of auto-encoders lead to efficient aging models. Whereas, integrated framework for both feature extraction and matching in CNN results in more discriminative information. Therefore, deep learning based AIFR methods can be viewed as generative as well as discriminative methods. Thus in this paper, we have addressed deep learning as a separate category than the generative and discriminative methods.

3.2 Generative approach

Generative approaches rely on age normalization before matching. Hence, age estimation and age synthesis are two important steps in generative approach. But, estimating age from a facial image has been addressed as a separate research problem (Feng et al. 2017; Geng et al. 2013; Liang et al. 2014). Detailed survey about the same is presented in Ramanathan et al. (2009). In this section, we discuss the aging models and age estimation methods that are used in age invariant face recognition. In this survey, we discuss the generative approaches in the context of different aging models. Figure 4 illustrates the general steps in generative age invariant face recognition.

3.2.1 2D aging models

In generative method, aging models are used to generate the past or future appearance of the person. Hence, the information about the age, at which age progression is to be carried out is essential in generative methods. Therefore, age estimation and age transformation (i.e. age progression) are two important steps in generative methods. After age transformation, the age invariant face recognition problem is reduced to the classical face recognition problem, which is solved by existing generic face recognition algorithm. Here we discuss the 2D aging models and corresponding age estimation methods presented in the AIFR literature. Section 3.2.2 presents discussion on 3D aging model.

A lot of structural and textural variations are observed in facial appearance due to aging. Active Appearance Model (AAM) (Cootes et al. 1995, 2001) learns both shape and appearance via Principle Component Analysis (PCA) (Turk and Pentland 1991). Hence, PCA coefficients of AAM are widely used for face analysis. In Lanitis et al. (2002), AAM is used to extract craniofacial growth and gray level variations during different age span. Statistical model parameters and mean example (X_m) are learned from training. In Lanitis et al. (2002) following reconstruction model is proposed.

$$X = X_m + Pb \tag{1}$$

where X is training example, X_m mean example, P is matrix of eigenvectors and b is vector of weights. The weighs are computed from (1) as follows

$$b = P^{-1}(X - X_m). \tag{2}$$

Age estimation has been carried out using model parameters b and quadratic function. The aging function is defined as,

$$age = f(b) \tag{3}$$

where *age* represents the actual age of the person in the given image and f is aging function. The aging function represents the relationship between facial image and actual age. The appearance of the person at given age is reconstructed using

$$b = f^{-1}(age). \tag{4}$$

For a given test image, model parameters b are computed and classification is carried out using Mahalanobis distance between probe and gallery images. The recognition rate was improved by 10% when aging simulation was employed. Approaches presented in Patterson et al. (2006), Wang et al. (2006) are also based on formulation of aging function from PCA coefficients of shape and texture.

The approach in Lanitis et al. (2002) suffers from various limitations such as, quadratic nature of aging function does not uphold the relationship between the actual age and face image. Use of limited number of images for training the aging model does not capture the diverse aging variations. To address these limitations, Geng et al. (2006, 2007) proposed manifold learning method called AGing pattErn Subspace (AGES). It models the aging process using a sequence of age-ascending face images of the same individual as shown in Fig. 5b. In AGES, face images are represented by AAM parameters. Missing faces in the aging pattern are synthesized using eigenspace (Turk and Pentland 1991) and Expectation Maximization (EM) (Moon 1996) algorithm. Once all the missing face images are reconstructed correctly, the aging pattern becomes complete. In AGES, for age estimation, first, a probe image is projected onto the manifold space. Next, the position of the probe image in the aging pattern corresponding to the best reconstruction is identified. Estimated age of the probe image is nothing but the position of the probe face image in the aging pattern. Improper positioning of test images in the aging pattern leads to ghost like effects in the reconstructed faces.

FG-NET aging database was used for age estimation and face recognition tests. For face recognition two images of 10 subjects were randomly selected from the database, one for the

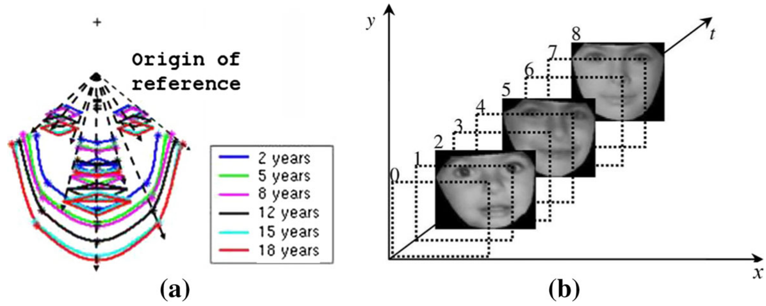

Fig. 5 **a** Growth model in Ramanathan and Chellappa (2006) and **b** Aging pattern in Geng et al. (2010)

gallery and other as a probe. The Mahalanobis distance is used as a similarity measure. The AGES method assumes availability of large number of face images of the similar person or similar aging pattern. Since AAM is used to represent faces, the AGES also fail to model large individual textural variation present in adults due to aging.

Face anthropometric measurements provide a quantitative description of the craniofacial growth at different age. Therefore, these measurements are suitable for facial age modeling. According to face anthropometric studies, some facial features change relatively less compared to others during age progression in formative years. Thus, different facial regions have different growth parameters. In Ramanathan and Chellappa (2006), the growth parameters are computed using anthropometric data and are used to build the craniofacial growth model shown on Fig. 5a. The growth model predicts the general face shape at various ages which is further used to extract features. Face recognition test was performed on a set of images from the FG-NET database and author's own database. There are total 233 images of 109 subjects.

Since anthropometric data for adults is not easily available, facial aging in adults has been modeled using shape and texture variation model separately in Ramanathan and Chellappa (2008). The changes in the facial features due to aging in adults are captured using muscle models. Facial muscles are categorized into three types namely; linear, sheet, and sphincter muscles to model facial feature drifts of each muscle. Shape transformation due to aging is modeled as a linear combination of the facial drifts observed on these three facial muscles. Textural variations such as wrinkles in facial regions such as the forehead, nasolabial region are captured using image gradient-based transformation functions. Experimental results show the applicability of the aging model for face verification and facial appearance prediction across aging.

A novel graph based approach for image representation is proposed in Mahalingam and Kambhamettu (2010) for efficient face recognition across aging. The vertices of the graph are considered as feature points and extracted using modified Local Feature Analysis (LFA) (Penev and Atick 1996). Feature vector across age for each feature point is computed by applying uniform LBP on each feature point. Age model is constructed by modeling joint probability distribution of the appearance of the graphs using Gaussian Mixture Model (GMM). Two-stage matching utilizes Maximum A Posteriori (MAP) to reduce the search space and graph matching to explore similarity. Effective representation of shape and texture using Graph and GMM respectively improves the performance of the system. The graph and GMM system performed well when tested for adults (above 18 years), but the cumulative accuracy of the system degraded for subjects under 18 years age.

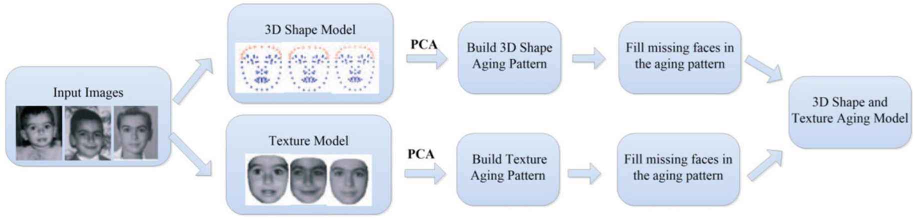

Fig. 6 3D aging model presented in Park et al. (2010)

3.2.2 3D face aging model

Craniofacial growth corresponds to 3D shape change due to age progression. This implies aging is a 3D transformation process and accordingly, 3D aging models can be used to learn the age variations. In Park et al. (2008, 2010), an automatic age simulation has been carried out using 3D morphable models (Blanz and Vetter 1999). Aging patterns presented in Geng et al. (2007) are extended to 3D aging patterns. A 3D face shape is represented as

$$S_\alpha = \overline{S}_{mm} + \sum_{l=1}^{L} \alpha_l W_l. \tag{5}$$

where S_{mm} is a set of 3D face models used in constructing the reduced morphable model S_α, α is model parameter which controls the shape, $\overline{S}_{mm}$ is mean shape and W_l's are unit eigenvectors of the shape covariance matrix. Separate shape and texture aging patterns are proposed as shown in Fig. 6. To incorporate age correction in the model, the overall size of each 3D shape is scaled according to average anthropometric head width. Shape and texture aging subspace are constructed assuming that, any aging pattern can be approximated by a weighted average of the aging patterns in the training set.

Two databases FG-NET and MORPH are widely used for evaluation of aging models. U. Park et al. (2010) have scanned 103 images of 4 subjects from the book by Nixon et al. (2007) and created a new facial aging database named 'BROWN'. The approach proposed in Park et al. (2010) was evaluated using commercial face recognition engine FaceVACS (2010). Since different facial changes are observed during the formative years and adult aging, corresponding identification accuracies are also different in these age ranges. Improvement in average recognition rate from 17.3% (without aging model) to 24.8% (with aging model) is observed for age group till 18 years. Performance improvement for adult age group is from 38.5 to 54.2%. Low recognition rate was observed in the development stage than adults due to the fact that major facial changes happen during development stage.

In summary, anthropometric measurements are used to extract distance ratios of facial components. They are further used as features to classify the facial images. However, anthropometric measurements are suitable for modelling age progression only during formative years and not in adults. Heavy dependence of these measurements on accurate facial landmark localization and manual annotations leads to inaccurate performance of generative models. Although 2D and 3D aging models have been used in age invariant face recognition, they face challenges due to multiple aging pattern and life style factors. In the generative approaches, query image age is essential and is mostly not available in real time applications. Inaccuracies in the estimated age, limits the performance of generative AIFR. In case of small sample size, the model is not able to precisely capture the age variations. If images are captured under controlled conditions, then the model may learn the exact aging process, which is of limited use. Several additional facial variations like pose, illumination and expression have also been encountered in practical situations. These problems have been better addressed by discriminative approaches.

3.3 Discriminative methods for AIFR

Facial aging is a complex process that affects the shape and texture of the face. However, the available facial aging databases contain age-separated images along with other variations such as pose, illumination and expression. Hence, it is difficult to learn aging using generative models. Whereas, histogram based local feature descriptors have good properties to be invariant to image transforms, distinctive for recognition and robust to noise and lighting variants. Algorithms that use local appearance descriptors are more robust against occlusion, expression, pose and illumination variations than traditional holistic algorithms. In discriminative AIFR, various local feature descriptors extract age invariant features from facial images. In addition, discriminative methods learn matching scheme for age separated images. In this section, we discuss the age invariant feature descriptors and subsequent matching frameworks used in AIFR.

3.3.1 Age invariant feature descriptors

An effective solution to overcome the problems associated with generative models is to use local image descriptor for face representation. The local feature descriptors not only handle aging variations, but also provide robustness to other intra-person variations. Various existing local feature descriptors have been used in discriminative approach (Gong et al. 2013; Juefei-Xu et al. 2011; Li et al. 2011; Sungatullina et al. 2013) for AIFR. Since recognition rate of AIFR system highly depends on the choice of local features, some of the authors have designed feature descriptors that extracts age invariant signature (Chen et al. 2014a, b; Gong et al. 2015; Ling et al. 2007, 2010). Here we discuss various feature descriptors that extracts the age invariant features from the facial image.

Slow, but continuous aging process results in appearance change such as muscle sagging and wrinkles. These changes are referred as facial drifts and are computed in Biswas et al. (2008) for face verification. Inspiration to capture the coherency in the feature drifts has been drawn from the theory of electrostatics. The drift maps are treated as the sparse charge distributions and potential energy of a system is equivalent to the incoherency in the feature drifts. The measure of incoherency between two feature drifts is computed from

$$U_{ij} = \frac{\left\| a_i - a_j \right\|}{r_{ij}} \tag{6}$$

where $\left\| a_i - a_j \right\|$ is absolute difference between the two feature drifts a_i and a_j while r_{ij} is the distance between the corresponding feature locations. Total feature drift due to K feature drifts is given by

$$C = \sum_{i=1}^{K} \sum_{j=i+1}^{K} U_{ij}. \tag{7}$$

The facial drifts of an individual are highly coherent over a wide range of a person's age, hence represents an age invariant feature. However, for images of different individuals higher incoherency in the feature drifts has been observed. First, Scale Invariance Feature Transform

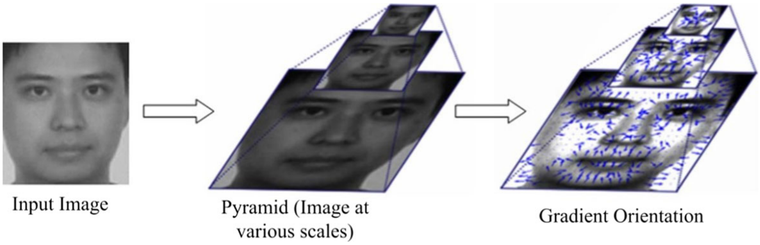

Fig. 7 Computation of GOP from image (Ling et al. 2007)

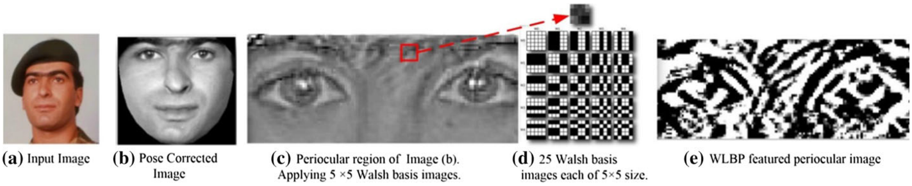

Fig. 8 Sample periocular region (**c**), applying Walsh mask (**d**) and WLBP image (**e**) (Juefei-Xu et al. 2011)

(SIFT) (Lowe 2004) is used to extract feature points. Next, feature vector for each image is constructed using the gradient distribution around feature points. Vector differences between the corresponding feature drift in different image is used as a measure of incoherency. A coherency pattern due to feature drifts in 3D range data has also been analysed.

A robust face descriptor for face verification across age progression and the effect of age progression on the performance of verification algorithms are investigated in Ling et al. (2007, 2010). The distribution of the edge direction is considered to be robust across aging. A structure of gradient orientation computed at different scales is defined as Gradient Oriented Pyramid (GOP). Computation of GOP from an image is shown in Fig. 7. The feature vector is computed by concatenating cosines of gradient orientations at all scales. For face verification, difference between the feature vectors of two images is combined with the Support Vector Machine (SVM). Compared to image intensity based descriptors, the GOP is more robust and performs well on face images with large age gaps. Experimentation on passport database and the FG-NET database demonstrated promising results in comparison to several contemporary techniques.

Instead of developing a new feature descriptor, multiple local features have been used in Li et al. (2011). Modified Versions of Multi-scale Local Binary Patterns (MLBP) (Ojala et al. 2002) and SIFT are used to extract age invariant features. Whereas, in Juefei-Xu et al. (2011) age invariant features are extracted from the fusion of Walsh-Hadamard transform and Local Binary Pattern (LBP). It is named as Walsh-Hadamard transform encoded LBP (WLBP). WLBP is applied on pre-processed periocular region and not on entire face. Periocular region i.e. area around eye sockets, is the facial region in the immediate vicinity of eyes. Periocular region of a sample image from FG-NET database and it's Walsh feature computation is shown in Fig. 8. Periocular region is considered to be the most age invariant facial region. The WLBP based features extracted from periocular region are distinct for every subject across aging. Instead of considering entire face with high structural complexity, periocular region has less temporal variations. Also, the periocular region has most dense and complex bio-informative features.

Multiple local descriptors are widely used in face recognition. Since every feature descriptor represents unique information; their combination can provide most discriminant information. In general, face representation using multiple descriptors is achieved by concatenating individual feature vectors into a new vector. Since every feature has specific statistical property, concatenation of these features would ignore the weight of individual features in the multiple feature representation. In Sungatullina et al. (2013), discriminative information is extracted from LBP (Ojala et al. 2002), SIFT and GOP (Ling et al. 2007). Instead of concatenating them a multi-view learning scheme has been developed to extract information from individual features.

A novel face encoding method; a Cross Age Reference Coding (CARC) proposed in Chen et al. (2014a, b) is based on a belief that, similar looking people age in a similar way. Age separated images of similar looking persons have similar representation in CARC. It is assumed that they will look similar to some people in the reference set over a considerable period of time. After face alignment, LBP features for each landmark point are extracted. Then, CARC encoding is carried out in three steps. In the first step, reference set representation according to age group is computed. In the second step, local features obtained from reference set representation are encoded into a reference space. In the last step, age invariant representation is obtained by combining features computed in the second step.

A novel feature descriptor based on maximization of information, named Maximum Entropy Feature Descriptor (MEFD) has been proposed in Gong et al. (2015). The MEFD encodes the microstructure of facial images in such a way that information content is maximized. Among existing local descriptors, LBP has been widely used in face recognition. It encodes the facial image without considering the frequency of LBP code. In MEFD, a decision based encoding tree is trained in a greedy manner which encodes the image with even distribution, hence maximizes the information. Factor analysis based matching framework called Identity Factor Analysis (IFA) is also presented in Gong et al. (2015). In IFA, identity components are further utilized to estimate the probability that two faces have the same underlying identity.

A feature descriptor called Local Pattern Selection (LPS) (Li et al. 2016) learns age invariant signatures from a hierarchical model. At the first level, it extracts low-level common information. Along the similar lines of MEFD, the encoding tree in LPS is also trained in a greedy manner to generate even code distribution. At the second level, high dimensional LPS features are refined into low-level visual features using subspace analysis. These low level visual features act as age invariant features. When LPS is combined with Hidden Factor Analysis (HFA) (to be explained later) substantial improvement is observed in the recognition rate.

In summary, large improvements in recognition rate can be achieved by identifying age invariant features descriptors. Appearance based methods extract facial features from the original image space. Hence they are inefficient in recognizing age separated images of a person. However, local features try to identify age invariant signatures resulting in improved recognition rate. The LBP code is able to capture edge but not the direction of the edge; hence, it is less efficient in capturing the large texture variations such as wrinkles. Recently novel edge-based local feature descriptors such as Positional Ternary Pattern (PTP) (Iqbal et al. 2016), Local Edge Prototype Pattern (LEPP) (Iqbal and Chae 2018) and Directional Age-Primitive Pattern (DAPP) (Iqbal et al. 2017) have been proposed to capture edge-directional information. These are specially designed feature descriptors, which efficiently capture facial aging. As mentioned earlier, the direction of the edge represents the age invariant feature; use of such descriptors in AIFR may boost the performance.

3.3.2 Discriminative learning methods

The discriminative AIFR is a two-step approach: first, age invariant clues are extracted; next, learning schemes are developed to classify the age separated images. Therefore, some authors develop age invariant feature descriptor and use existing learning methods. Some authors use existing local features and develop new learning method especially for AIFR. However, a few publications offer both age invariant features as well as subsequent learning. In this section, we discuss various learning algorithms used for AIFR.

Use of multiple local features and their concatenation is common in various face recognition algorithms. Such concatenation results in high dimensional feature vector. Hence, multiple features of large dimensionality and different scale are addressed by Multi-Feature Discriminant Analysis framework (MFDA) in Li et al. (2011). Because of large feature vector and relatively small sample size, Linear Discriminant Analysis (LDA) (Belhumeur et al. 1997) fails to effectively classify aging facial image. The MFDA is an extension and improvement of LDA. It combines random subspace and bagging techniques to refine the feature space. Dimensionality problem is addressed by employing the random subspace technique. The bagging technique has been used to extract the discriminant information near classification boundary. The experimental results in terms of recognition rate demonstrate the effectiveness of the feature representation as well as the MFDA framework.

When multiple features are used to represent facial appearance, the information observed from multiple views should be effectively utilized for classification. Multi-view Discriminative Learning (MDL) (Sungatullina et al. 2013) learns the information observed from multiple features. Since multiple features of each sample imply the same class, maximizing the correlation between these features can minimize intra-class variation. In MDL, various features are projected on a common feature space. Rank-1 recognition in case of MORPH (65.2%) is less than that FG-NET (91.8%). Since the age distribution in FG-NET is skewed towards the formative age, the results show that it is more suitable for face recognition in children's age group.

To classify age separated images using only periocular region, Unsupervised Discriminant Projection (UDP) (Yang et al. 2007) has been used in Juefei-Xu et al. (2011). UDP characterizes local and non-local scatter. Instead of using the within class and between class scatter matrix as in LDA, UDP finds an optimal projection that minimizes the local scatter matrix S_L and maximizes the non-local scatter matrix S_N simultaneously. The local scatter matrix S_L in UDP is generated by an adjacency matrix A, that is produced by K-nearest neighbors (KNN). The optimization function for UDP is given by

$$w^* = \arg\max \frac{w^T S_N w}{w^T S_L w}. \tag{8}$$

where $S_L = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{A}_{ij}\left(x_i - x_j\right)\left(x_i - x_j\right)^T$ and $S_N = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(1 - \boldsymbol{A}_{ij}\right)\left(x_i - x_j\right)\left(x_i - x_j\right)^T$. Because of the incorporation of nonlocal information UDP is more powerful than Locality Preserving Projection (LPP) (He and Niyogi 2003). Also, in case of small sample size, UDP is more discriminative than LDA. Use of small part of face image for feature description substantially reduces the computation time. However, if images are not pose corrected, full periocular region may not be extracted which affects the recognition rate.

Factor analysis is a multivariate analysis technique that identifies the common characteristics (latent factors) among the observed variables (French 1985). Hidden factor analysis is a probabilistic model that decomposes the observed feature vector into identity factor and age factor. In Gong et al. (2013) a statistical factor analysis model is proposed to address age invariant face recognition. The probabilistic factor analysis model is given by

$$t = \mu + Qx + Py + \varepsilon \tag{9}$$

where t represents observable feature vector, μ denotes mean face features, x denotes the latent identity factor, y denotes age factor and observation noise is denoted by ε. The noise follows an isotropic Gaussian distribution i.e. $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$. Identity factor corresponds to the identity specific face component which is stable over the aging process. Age factor corresponds to the facial growth due to aging. Hidden factors and factor loadings are estimated from the facial features using Expectation–Maximization (EM) algorithm. Identity factor represents robust age invariant feature. For face matching posterior mean of identity factors of probe and gallery images are compared. Histogram of Oriented Gradient (HOG) (Dalal and Triggs 2005) is used as a local feature descriptor. Experiments are conducted using MORPH and FG-NET and a noticeable improvement has been achieved compared to state-of-the-art algorithms.

Along with MEFD, the factor analysis based matching framework called identity factor analysis is also presented in Gong et al. (2015). In IFA, age and identity components are used to estimate the probability that two faces have the same underlying identity. Efficiency of the feature descriptor as well as the matching method are also investigated separately. Also, the performance is compared independently with the state-of-the-art feature descriptors and AIFR approaches. The experimental results revealed that the MEFD outperforms the contemporary feature descriptors. Comparative analysis of IFA with existing AIFR approaches shows the improvement in recognition rate. Capabilities of IFA are further explored by combining IFA with multiple local features (MEFD, SIFT and HOG), which has shown further improvement in the results. Evaluation of face verification based on the IFA has also been performed on a Labelled Faces in Wild (LFW) database (Huang et al. 2008). Although, the IFA has been developed for age invariant FR task, its performance improvement on complex and unconstrained LFW database (which is not face aging database) clearly shows its effectiveness and generalizability.

Hidden factor analysis in Gong et al. (2013) is based on the assumption that identity and age factors are independent. But this is not true in real life, since different aging patterns are observed in different people. A modified latent factor model based on age-identity correlation is presented in Li et al. (2017). Modified Hidden Factor Analysis (MHFA) model separates age-identity correlation features along with other appearance variations. Since new latent factor separates other intra-person facial variations, the identity factor represents most robust age invariant feature. EM algorithm is used to jointly learn model parameters and latent factors. Probabilistic matching framework is presented to find best match between gallery and probe face image. Experimentation on FG-NET and MORPH shows noticeable improvement in recognition rate compared to HFA.

Neural network (NN) based approaches have also been proposed for AIFR. An NN method depicting group foraging strategy of a swarm of *E. coli* bacteria is used for optimization in Yadav et al. (2013). Specifically, the bacteria foraging algorithm (Muller et al. 2002) is inspired by the chemotaxis behaviour of bacteria. It has been observed that, different facial regions are affected differently due to aging. Their contributions in recognition are learned using bacteria foraging algorithm. Five facial regions, right and left periocular regions, binocular region, and mouth along with the full face have been considered for feature extraction. LBP of each region is calculated. Optimum weights for different facial regions are obtained during training using bacteria foraging algorithm. These learned weights are used to compute the final weighted matching scores. The performance of the proposed algorithm is evaluated on two databases; FG-NET and IIITDelhi face aging database (Yadav et al. 2013). IIIT-Delhi is collection of Indian celebrity images available on internet. It contains 2618 images of 102 subjects of which 49 are females and 53 are males; with age span 4–88 years. The performance of the proposed algorithm outshines some of the existing algorithms.

Effect of weight change along with aging on face recognition performance are analysed in Singh et al. (2014). Age variations with respect to weight change have been learned using neural network and random forest classifier. New database, containing both age and weight variations is named WhoIsIt database. Performance evaluation on FG-NET and WIT database has been carried out for comparative analysis of the proposed system with existing algorithms. It is claimed that the weight information supports the identification and improves performance. But, in our opinion the improvement in the performance is due to age invariant facial regions and not because of weight information (as weight information is not provided in FG-NET).

Although, significant achievements have been observed in discriminative methods, their performance deteriorates in unconstrained environment. Different aging patterns are observed during different stages of life, which is a challenging factor for discriminative methods. Because of these challenges, single optimal classifier is not efficient to learn a common age invariant space which results in reduced performance. However, ensemble learning may solve this complex problem in our opinion.

3.4 Deep learning

In recent years, deep learning methods have emerged as a powerful machine learning model. Various CNN based approaches have been developed in the literature (Krizhevsky et al. 2012; Ouyang et al. 2015) and are successfully applied to address the face recognition problems (Chen et al. 2013; Sun et al. 2014; Yi et al. 2014). The convolutional neural network is a class of feed-forward neural network. It consists of input layer, multiple hidden layers and output layer. The hidden layers in CNN consist of convolutional layer, pooling layer, and fully-connected layer. The key idea behind the convolutional layer is that the image statistics are translation invariant. To learn these features various filters are used in the convolutional layer. The pooling layer performs down sampling along the spatial dimensions and merges features to compactly represent the features. Fully connected (FC) layer is nothing but the traditional neural network, which makes nonlinear combination of features to achieve classification.

Face verification across aging using deep neural network has been addressed in Bianco (2017), Li et al. (2015). In Li et al. (2015), a CNN is trained to learn features, distance metric and threshold function simultaneously. The parameters of these components are extracted using the network propagation algorithm. Two convolutional layers and a max-pooling layer are alternately arranged and three FC layers are used. Stack of two convolutional-pool layers along with first FC layer extracts the features. A common space for feature projection is learned in the second FC layer. Distance metric and threshold are learned in last FC layer. Experimental result shows the rank one recognition rate outperforms existing discriminative and generative AIFR methods.

Face verification scheme using external feature injection in deep networks has been presented in Bianco (2017). Its contribution is twofold: first, a new external feature injection scheme in deep network is presented; and next, a new large-scale Large Age-Gap (LAG) dataset is presented. Discriminative representation of faces is initially learned by training the CNN for the task of face recognition on CASIA-WebFace (Yi et al. 2014) dataset. The architecture consists of 5 convolutional layers, 3 pooling layers and 3 FC layers. Output of the

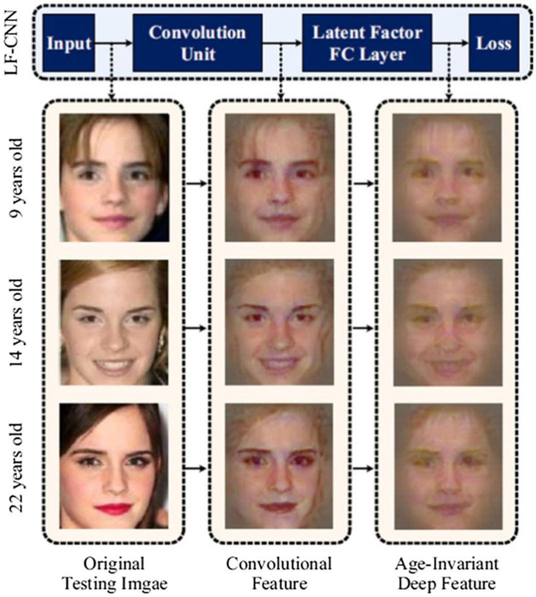

Fig. 9 LF-CNN model and visualization of convolutional and age invariant features in Wen et al. (2016)

second FC layer represents the features for face representation. In order to use the network for the LAG dataset, the complete network is fine-tuned using a contrastive loss function. Since CASIA-WebFace is not a facial aging dataset and LAG is a large dataset, fine tuning from the scratch ensured the confidence in the learned weights. Externally computed features are injected with the CNN learned features in the fine tuning stage. Concatenated features are processed by the last FC layer for face verification. The experimental results show the improvement in the performance due to the use of feature injection in deep learning.

Factor analysis being effective in separating age and identity components from facial features, a Latent Factor guided CNN (LF-CNN) framework is presented in Wen et al. (2016) and is shown in Fig. 9. Age invariant feature extraction and recognition are two independent tasks in AIFR. The network in Wen et al. (2016) is therefore trained using two different training datasets; first with only identity variations and the second consisting of age and identity variations. MORPH-2 dataset is used to learn age variations. Deep features extracted by the convolutional layer are used as observed features for Latent Identity Analysis (LIA). Identity factor is further used to update weights and bias of the fully connected layer. Two loss functions are used to strengthen the learning. Softmax loss function is used to increase inter-person variations, whereas, contrastive loss function is used to reduce intra-person variations due to aging. Stochastic gradient decent is used to update parameters of convolutional layer to improve the loss. The LIA guided learning in CNN shows significant improvement in the performance compared to all the existing AIFR methods.

Deep learning based generative approach using an auto-encoder is presented in Xu et al. (2017). Age progression and subsequent recognition across aging has been achieved using Coupled Auto-encoder Network (CAN) and nonlinear factor analysis. Autoencoder is an unsupervised learning algorithm that has an input layer, a encoding layer, and a decoding layer. The network in autoencoder is trained to learn good representation of the inputs. Nonlinear factor analysis is carried out on the encoded image of CAN to extract age and identity components. Aging and de-aging process for a given pair of facial image is learned

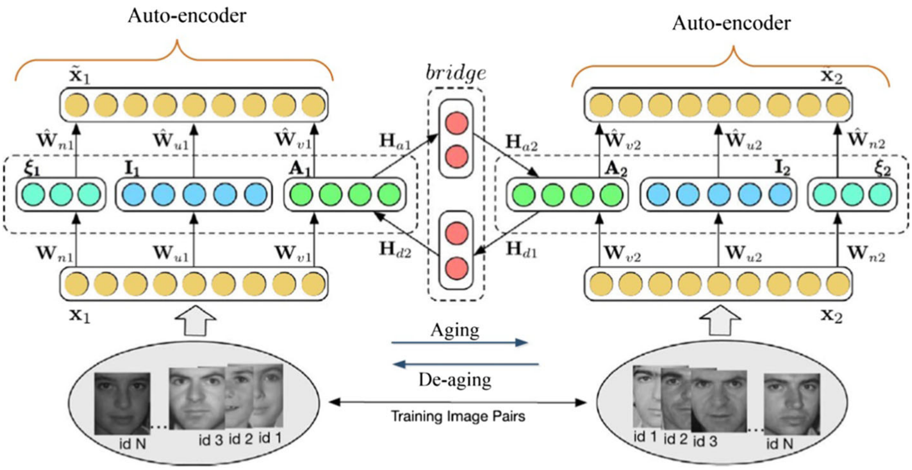

Fig. 10 Coupled auto-encoder network in Xu et al. (2017)

by bridging two separate auto-encoders with the help of shallow neural networks as shown in Fig. 10. The encoded image is in the hidden layer of the auto-encoder. Nonlinear factor analysis method separates the identity factor which in turn acts as robust age invariant features. These learned identity encoder parameters of probe and gallery image are further used for matching. Being unsupervised method, performance of CAN do not surpass the performance of LF-CNN method, but is better than other state-of-the art methods.

Although CNN's are very recently introduced in AIFR field, their performance is remarkable. Separation of identity factor from age is utilized in both the deep learning approaches published so far. After development of the discriminative approaches, generative models for AIFR became momentarily obsolete. But, the face representation and reconstruction capabilities of auto-encoder in deep learning have revived the application of generative models in AIFR.

Table 2 summarizes and benchmarks various published AIFR techniques so far in terms of facial features, learning methods, databases, protocols used for evaluation and performance in terms of recognition rate.

4 Effect of aging on performance of face recognition algorithms

Another important dimension of AIFR is analysis of effects of aging on the performance of AIFR methods. For such analysis the database is partitioned into various age groups and the effect of aging in terms of performance measures (recognition rate, false acceptance rate and true rejection rate) are reported for each age group. The performance measures are further used to infer the trends the recognition rate across different age group. In this section, we discuss the experimental work that has been carried out by various researchers to address the effects of aging on FR performance. The effects of aging on the performance of face recognition are experimentally quantified in Lanitis (2004, 2009a, b), Panis and Lanitis (2014). Based on their experiments, the following conclusions are derived: (1) Average decrease in the face recognition performance is about 10% due to age variations in the faces of the same person in the test and the training set (Ling et al. 2007). (2) Intra-person variations such as pose and

Table 2 Comparative analysis of age invariant face recognition methods

Authors	Face representation	Learning method	Database (# images per subject, # images)	Protocol for evaluation	Rank 1 recognition rate (%)	
					Without aging	With aging
Lanitis et al. (2002)	AAM	PCA	Private Database (12, 85)	Images from two sets are used interchangeably for training and testing	57.0	68.5
Ramanathan and Chellappa (2006)	Cardioidal strain transform	PCA	FG-NET + Private Database (233, 109)	–	8	15
Geng et al. (2007)	AAM	PCA	FG-NET (10,10)	10 gallery images and 10 test images	14.4	38.1
Patterson et al. (2006)	AAM	PCA	MORPH	9 test images, 36 gallery images	11.0	33.0
Park et al. (2010)	PCA coefficients of 3D shape and Texture	FaceVACS	FG-NET (82, 1002)	82 probe images, 82 gallery images	26.4	37.4
			MORPH-Album 1	612 test images, 612 gallery images	57.8	66.4
			BROWNS	4 probe images, 100 gallery images	15.6	28.1
Mahalingam and Kambhamettu (2010)	LFA and LBP	MAP and Graph matching	FG-NET (82, 1002)	Two sets of training and testing.	–	50
Li et al. (2011)	MLBP and SIFT	MFDA	FG-NET (82, 1002)	82 probe images, 82 gallery images	–	47.50
			MORPH Album 2	10,000 probe images, 20,000 gallery images	–	83.9
Juefei-Xu et al. (2011)	WLBP	UDP	FG-NET (82, 1002)	LOPO	–	100
Sungatullina et al. (2013)	LBP, SIFT and GOP	MDL	FG-NET (82, 1002)	LOPO	–	91.8
			MORPH Album 2	8000 probe (oldest) images 8000 Gallery (youngest) images	–	65.2
Singh et al. (2014)	HOG	Random forest classifier	FG-NET (82, 1002)	50% images for training and 50% for testing.	–	20.34

Table 2 continued

Authors	Face representation	Learning method	Database (# images per subject, # images)	Protocol for evaluation	Rank 1 recognition rate (%)	
					Without aging	With aging
			WIT		–	28.53
Yadav et al. (2013)	LBP	Bacteria foraging fusion	IIITDelhi	70% images are used for testing and 30% for training.	–	54.3
			FG-NET (82, 1002)		–	64.5
Gong et al. (2013)	Patch based HOG	HFA and cosine distance of identity features	FG-NET (82, 1002)	LOPO	–	69.0
			MORPH Album 2		–	91.14
Chen et al. (2014a, b)	CARC	Linear projection and Cosine distance	MORPH Album 2	10,000 test images 10,000 training images.	–	92.8
Gong et al. (2015)	MEFD	Identity Factor Analysis	FG-NET (82, 1002)	LOPO	–	76.2
			MORPH Album 2	10 fold cross validation 10,000 test images 10,000 training images.	–	92.26
Wen et al. (2016)	Deep features	LIA guided Deep Learning	MORPH Album 2	10,000 test images 10,000 training images.	–	97.51
Li et al. (2016)	LPS	Universal Subspace Analysis	MORPH Album 2	10,000 test images 10,000 training images.	–	92.11
	LPS + HFA					94.87
Xu et al. (2017)	Deep features	Nonlinear factor analysis	FG-NET (82, 1002)	82 probe images, 82 gallery images	–	86.5
Li et al. (2017)	Patch based HOG	Modified HFA and Maximum Likelihood approach	FG-NET (82, 1002)	LOPO	–	72.8
			MORPH Album 2	10 fold cross validation 10,000 test images 10,000 training images.	–	87.94

expression degrade the system performance in the similar way as the aging. (3) The upper face region significantly affects the performance over the period of time, whereas lower face regions i.e. lips, chin and nose tip cause smaller decrease in the performance (Otto et al. 2012).

Experimental analysis on MORPH database reported decrease in recognition rate as the age span between train and test images increases (Ricanek and Tesafaye 2006). A study in Ling et al. (2007), has reported 5% deterioration in the performance due to five years age gap between pairs of faces compared to no age gap. Relatively constant performance is observed for age gap of about 5–10 years.

Experimental analysis on the largest aging dataset has been conducted in Klare and Jain (2011). Impact of aging on two commercial face recognition systems has also been demonstrated in it. Very important issue is also addressed about performance degradation in non-aging scenario. It implies that while achieving performance improvement in presence of a large time gap between the probe and gallery images, performance degradation in non-aging or minimum age gaps is not acceptable. In general, if the system is trained in presence of age variations then its performance reduces in non-aging scenarios. Experimental analysis in Klare and Jain (2011) suggested that, the best performance on a particular age span is achieved by training a system only on that particular age span. One more important observation is that when the system is trained using all the subjects of the complete age range, then it results in consistent performance. However, availability of such complete age range data for all the subjects is a major problem.

Since existing method which analyses system performance shows non-uniformity in age group, it is not judicial to compare their performances. Also while evaluating any biometric system, it is important to analyse whether a decision is due to genuine individual or an imposter (Jain et al. 2006). Statistical analysis of performance evaluation based on these factors has been done in Best-Rowden and Jain (2015, 2018). They have conducted analysis using Pinellas County Sheriff's Office Longitudinal Study (PCSO_LS) database comprising of 147,784 face images of 18,007 subjects.

Temporal nature of the images follows hierarchical structure which can be effectively modelled using multilevel statistical model. Two level statistical model; level 1 for intra-subject variation and level 2 for inter-subject variation is used in Jain et al. (2006). Genuine scores and imposter scores are evaluated in aging scenario using two Commercial Off-The-Shelf (COTS) matcher. Decreasing trends in genuine scores across aging were observed for both COTS face matchers. Studying effect of imposter distribution on performance, investigating non-linear models and using database of civilians rather than criminals are a few suggested future directions.

Recently, effect of expression variations on facial aging is investigated in Guo and Wang (2012), Lou et al. (2018). The experiments in Guo and Wang (2012) demonstrate that facial expressions have a significant impact on age estimation. They further analyzed the problem of cross-expression age estimation which demonstrated the large impact of identity on age estimation. Expression and age are jointly learned using a graphical model in Lou et al. (2018) to offer an expression invariant age estimation. Both the studies involved analysis of effect of expression on age estimation. These studies indicate that similar analysis of effect of expression on AIFR is very important.

Effect of aging on system performance is summarized as follows: (1) recognizing face of younger individuals are more difficult than adults, (2) average decrease in the performance of FR system is about 10% due to variations in the faces of same person in test and the training set, (3) 5% deterioration in the performance is observed when age differences of about five years exists between pairs of faces and (4) eyes and nose are more informative facial components in automatic age estimation. Effect of aging on face recognition performance using large-scale database (PCSO_ LS) are milestones in the research on age invariant face recognition. These studies also suggest that, it is not sufficient to perform group wise analysis only on the basis of rank 1 recognition rates. It is also equally important to analyse whether a face matching is due to genuine individual or an imposter. Also other demographic factors along with aging should be considered while analysing the performance.

5 Discussion

Previous sections present a complete overview of the research in the AIFR field; existing techniques, their capabilities and limitations, performance benchmarking and facial aging databases.

In Sect. 3, we have discussed effects of aging on facial appearance, early works to quantify the facial growth using geometric transformations and face anthropometry. Although such methods helped in the initial phase of this field, they have many limitations. The geometric transformations cannot model textural variation and anthropometric measurements of adult aging are not available. Thus, these models are inefficient in modelling the complete growth of human faces required for consistent AIFR performance.

Our major focus in this survey is on AIFR techniques presented in Sect. 3. Three categories presented for AIFR techniques addresses the recognition across from the different perspectives. Although, generative methods are successful to certain extent, their performance mainly degrades due inefficient aging models. Dependencies of facial aging on large intrinsic and extrinsic factors, along with incomplete databases are major limitations for building accurate aging models. Also, accuracy of age estimation required in generative approaches affects their performance. Generative AIFR approaches need age estimation methods that further use age varying features for age estimation. CNN based generative models are not much explored. Deep learning approaches can capture the diverse aging patterns and have also shown improvements in the age estimation accuracy.

Discriminative approaches rely on age invariant facial features and discriminant learning methods. In some way, discriminative methods have potential to address the major AIFR problems. However, misaligned faces, large self-occlusion and blur are a few problems which degrade their performance. Large dimensionality is one of the limitations of the local feature descriptors in discriminative approaches. Age invariant local features also face the same problem. Manifold learning methods may be useful to address this problem. In manifold learning, local features are projected onto a low dimensional manifold. However, identifying a true age invariant manifold is a complex research problem.

The recent deep learning AIFR methods bank on the feature learning and classification capabilities of the convolutional layer and FC layer. They are capable of learning a huge number of faces in presence of aging and other variations and offer high level of FR performance. Deep learning approaches require large sized databases for face representation and subsequent learning. Due to unavailability of a single suitable database, the CNNs in AIFR are generally trained using two separate databases, one for extracting age invariant features and the other for classification. Therefore, compared to other computer vision applications, CNN based AIFR methods are in general computationally heavy and time consuming. Further, the FG-NET being a relatively small database, deep learning methods are not trained using FG-NET.

Cross-sectional techniques are generally employed for analysis of facial aging. But in group-wise analysis, other demographic factors such as race and gender are generally ignored. However, it has been observed that these factors are equally important for such analysis. Different trends in face recognition performance have been observed in males in females over the time. Therefore, above said factors are also equally important while designing the AIFR system. The research in AIFR is divided into two categories; in the first category, the researchers dedicate their efforts on improving feature descriptors and/or learning methods. The second category analyses performance of the AIFR systems on the basis of age groups and other demographic factors. But in our opinion, if a new descriptor or learning scheme is proposed, their performance evaluation subject to the demographic factors is highly important.

6 Conclusion and future work

The practical AIFR system faces challenges due to appearance variations within a subject and similarities between the subjects. We categorized the surveyed AIFR approaches in three categories: Generative, Discriminative and Deep learning. Every approach handles the AIFR problem from a different perspective. Generative AIFR use aging models for age transformation while discriminative approaches rely on age invariant features and learning schemes. Deep learning methods offer an integrated framework for face representation and classification. The performance of the generated methods heavily depends on accurate landmark detection, aging models and age estimation. Recent development in deep learning can be utilized to generate accurate aging models and age estimation techniques. This has enlivened the generative techniques again. Performance of discriminative AIFR systems mainly depends on the local features and learning methods. Thus, identifying age invariant signature and subspace learning method continues to be a hot topic. Deep learning methods find usability in both generative and discriminative categories. Its use in AIFR has substantially improved the performance. But complete training of deep learning networks which otherwise require large training data, using small databases (like FG-NET) remains a challenge.

Performance of AIFR system shows different trends for different ethnic origins, males and females. Since facial aging depends on various factors besides aging, it is important to analyse the effect of intrinsic and extrinsic factors on the performance. Hence, for more accurate evaluation of AIFR algorithms, a facial aging database with correct demographic information and wide intra-person variations is need of the hour. A single database featuring a large and equal number of faces per person, across all the age groups and demographic features covering the intra person variations is badly needed, though highly impractical.

Despite significant progress in AIFR, it is still far from expectations. For practical FR, the ultimate goal is to develop a system which will provide robust and consistent performance across aging, all other appearance variations and demographic factors. We propose a few guidelines for future research and development.

1. Generative approaches are found effective in facial shape modelling, whereas, age invariant signatures are effectively extracted by discriminative methods. Hence, joint approach of global and local features may result in better performance as the human cognitive process is integration of both the approaches.
2. Recently, it has been revealed that 3D facial image features are even more reliable than blood profile for identifying age. Also, aging introduces facial deformations in 3D space, so developing generative as well as discriminative methods based on 3D face models need to be explored further.

3. Deep generative models may be used to learn a complete aging pattern. Combining them with deep features may boost up the performance. Incorporating pre-processing and pre-classifiers (to address other intra person variations) in CNN based AIFR system is a viable future direction. Pre-training in deep learning AIFR is generally carried out using general facial databases (non-aging). Then fine tuning is done using facial aging databases. However, in our opinion, if both pre-training and fine tuning are carried out using aging databases, improvement in the performance can be achieved.
4. Facial aging is affected by various intrinsic as well as extrinsic factors, but their quantitative analysis is missing. Also analysis of effect of gender and race is also a prominent future direction in the development of a robust face recognition system. Performance degradation of AIFR system under non-aging scenarios is a very important issue in real time AIFR systems. Therefore, designing robust age invariant features and corresponding generalizable matching framework effective across aging and non-aging scenarios both, is a viable future direction.
5. The databases PCSO_LS and CACD have addressed many issues related to AIFR to a substantial level. But neither of the databases have subject images below 16 years age. Therefore, the effect of aging in the formative years on performance of face recognition systems is not duly investigated. Images in PCSO_LS and CACD are of adult criminals and celebrities respectively, hence, the aging data and features of general civilians has not been modelled. Therefore, development of large civilian databases with long term growth is required in near future.

References

Abate AF, Nappi M, Riccio D, Sabatino G (2007) 2D and 3D face recognition: a survey. Pattern Recognit Lett 28:1885–1906

Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. IEEE Trans Pattern Anal Mach Intell 28:2037–2041

Albert AM, Ricanek K, Patterson E (2007) A review of the literature on the aging adult skull and face: implications for forensic science research and applications. Forensic Sci Int 172:1–9

Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19:711–720

Best-Rowden L, Jain AK (2015) A longitudinal study of automatic face recognition. In: 2015 International conference on biometrics (ICB). IEEE, pp 214–221

Best-Rowden L, Jain AK (2018) Longitudinal study of automatic face recognition. IEEE Trans Pattern Anal Mach Intell 40(1):148–162

Bianco S (2017) Large age-gap face verification by feature injection in deep networks. Pattern Recogn Lett 90:36–42

Biswas S, Aggarwal G, Ramanathan N, Chellappa R (2008) A non-generative approach for face recognition across aging. In: 2nd IEEE international conference on biometrics: theory, applications and systems, 2008 (BTAS 2008). IEEE, pp 1–6

Blanz V, Vetter T (1999) A morphable model for the synthesis of 3D faces. In: Proceedings of the 26th annual conference on computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co, pp 187–194

Chai X, Shan S, Chen X, Gao W (2007) Locally linear regression for pose-invariant face recognition. IEEE Trans Image Process 16:1716–1725

Chen D, Cao X, Wen F, Sun J (2013) Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3025–3032

Chen B, Chen C, Hsu W (2014) Cross-Age Celebrity Dataset (CACD)

Chen B-C, Chen C-S, Hsu WH (2014) Cross-age reference coding for age-invariant face recognition and retrieval. In: European conference on computer vision. Springer, Berlin, pp 768–783

Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models-their training and application. Comput Vis Image Underst 61:38–59

Cootes TF, Edwards GJ, Taylor CJ (2001) Active appearance models. IEEE Trans Pattern Anal Mach Intell 23:681–685

Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). IEEE, pp 886–893

Ding C, Tao D (2016) A comprehensive survey on pose-invariant face recognition. ACM Trans Intell Syst Technol 7:37

Ebner NC, Riediger M, Lindenberger U (2010) FACES—a database of facial expressions in young, middle-aged, and older women and men: development and validation. Behav Res Methods 42:351–362

Eidinger E, Enbar R, Hassner T (2014) Age and gender estimation of unfiltered faces. IEEE Trans Inf Forensics Secur 9:2170–2179

FaceVACS (2010) Software Developer Kit, Cognitec Systems GmbH. https://www.cognitec-systems.de

Farage M, Miller K, Elsner P, Maibach H (2008) Intrinsic and extrinsic factors in skin ageing: a review. Int J Cosmet Sci 30:87–95

Farkas L (1994) Anthropometry of the head and face. Raven Press, New York, p XIX

Farkas LG, Munro IR (1987) Anthropometric facial proportions in medicine. Charles C. Thomas, Springfield, IL

Feng S, Lang C, Feng J, Wang T, Luo J (2017) Human facial age estimation by cost-sensitive label ranking and trace norm regularization. IEEE Trans Multimed 19:136–148

FGNET FG-NET Aging Database (2010). http://www.fgnet.rsunit.com

French S (1985) An introduction to latent variable models. Monographs on statistics and applied probability. J Oper Res Soc 36(5):453

Fu Y, Guo G, Huang TS (2010) Age synthesis and estimation via faces: a survey. IEEE Trans Pattern Anal Mach Intell 32:1955–1976

Gallagher AC, Chen T (2009) Understanding images of groups of people. In: IEEE conference on computer vision and pattern recognition, 2009 (CVPR 2009). IEEE, pp 256–263

Geng X, Zhou Z-H, Zhang Y, Li G, Dai H (2006) Learning from facial aging patterns for automatic age estimation. In: Proceedings of the 14th ACM international conference on Multimedia. ACM, pp 307–316

Geng X, Zhou Z-H, Smith-Miles K (2007) Automatic age estimation based on facial aging patterns. IEEE Trans Pattern Anal Mach Intell 29:2234–2240

Geng X, Fu Y, Smith-Miles K (2010) Automatic facial age estimation. In: 11th Pacific rim international conference on artificial intelligence, pp 1–130

Geng X, Yin C, Zhou Z-H (2013) Facial age estimation by learning from label distributions. IEEE Trans Pattern Anal Mach Intell 35:2401–2412

Gong D, Li Z, Lin D, Liu J, Tang X (2013) Hidden factor analysis for age invariant face recognition. In: Proceedings of the IEEE international conference on computer vision, pp 2872–2879

Gong D, Li Z, Tao D, Liu J, Li X (2015) A maximum entropy feature descriptor for age invariant face recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 5289–5297

Guo G, Wang X (2012) A study on human age estimation under facial expression changes. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 2547–2553

He X, Niyogi P (2003) Locality preserving projections. In: Advances in neutral information processing systems 16 NIPS2003

Huang FJ, Zhou Z, Zhang H-J, Chen T (2000) Pose invariant face recognition. In: Proceedings of the fourth IEEE international conference on automatic face and gesture recognition, 2000. IEEE, pp 245–250

Huang GB, Mattar M, Berg T, Learned-Miller E (2008) Labeled faces in the wild: a database forstudying face recognition in unconstrained environments. In: Workshop on faces in 'Real-Life' images: detection, alignment, and recognition

Iqbal MTB, Chae O (2018) Mining wrinkle-patterns with local edge-prototypic pattern (LEPP) descriptor for the recognition of human age-groups

Iqbal MTB, Ryu B, Song G, Chae O (2016) Positional Ternary Pattern (PTP): an edge based image descriptor for human age recognition. In: 2016 IEEE international conference on consumer electronics (ICCE). IEEE, pp 289–292

Iqbal MTB, Shoyaib M, Ryu B, Abdullah-Al-Wadud M, Chae O (2017) Directional age-primitive pattern (DAPP) for human age group recognition and age estimation. IEEE Trans Inf Forensics Secur 12:2505–2517

Jain A, Bolle R, Pankanti S (2006) Biometrics: personal identification in networked society, vol 479. Springer, Berlin

Juefei-Xu F, Luu K, Savvides M, Bui TD, Suen CY (2011) Investigating age invariant face recognition based on periocular biometrics. In: 2011 International joint conference on biometrics (IJCB). IEEE, pp 1–7

Klare B, Jain AK (2011) Face recognition across time lapse: on learning feature subspaces. In: 2011 International joint conference on biometrics (IJCB). IEEE, pp 1–8

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105

Lanitis A (2009a) Facial biometric templates and aging: problems and challenges for artificial intelligence. In: AIAI workshops, pp 142–149

Lanitis A (2009b) A survey of the effects of aging on biometric identity verification. Int J Biometr 2:34–52

Lanitis A, Taylor CJ, Cootes TF (2002) Toward automatic simulation of aging effects on face images. IEEE Trans Pattern Anal Mach Intell 24:442–455

Lanitis A, Draganova C, Christodoulou C (2004) Comparing different classifiers for automatic age estimation. IEEE Trans Syst Man Cybern B Cybern 34:621–628

Li Z, Park U, Jain AK (2011) A discriminative model for age invariant face recognition. IEEE Trans Inf Forensics Secur 6:1028–1037

Li Y, Wang G, Lin L, Chang HA (2015) Deep joint learning approach for age invariant face verification. In: CCF Chinese conference on computer vision. Springer, Berlin, pp 296–305

Li Z, Gong D, Li X, Tao D (2016) Aging face recognition: a hierarchical learning model based on local patterns selection. IEEE Trans Image Process 25:2146–2154

Li H, Zou H, Hu H (2017) Modified hidden factor analysis for cross-age face recognition. IEEE Signal Process Lett 24:465–469

Liang Y, Wang X, Zhang L, Wang Z (2014) A hierarchical framework for facialage estimation. Math Probl Eng 2014:242846

Ling H, Soatto S, Ramanathan N, Jacobs DW (2007) A study of face recognition as people age. In: 2007 IEEE 11th international conference on computer vision. IEEE, pp 1–8

Ling H, Soatto S, Ramanathan N, Jacobs DW (2010) Face verification across age progression using discriminative methods. IEEE Trans Inf Forensics Secur 5:82–91

Lou Z, Alnajar F, Alvarez JM, Hu N, Gevers T (2018) Expression-invariant age estimation using structured learning. IEEE Trans Pattern Anal Mach Intell 40:365–375

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60:91–110

Mahalingam G, Kambhamettu C (2010) Age invariant face recognition using graph matching. In: 2010 Fourth IEEE international conference on biometrics: theory applications and systems (BTAS). IEEE, pp 1–7

Mark LS, Todd JT, Shaw RE (1981) Perception of growth: a geometric analysis of how different styles of change are distinguished. J Exp Psychol Hum Percept Perform 7:855

Moon TK (1996) The expectation-maximization algorithm. IEEE Signal Process Mag 13:47–60

Muller SD, Marchetto J, Airaghi S, Kournoutsakos P (2002) Optimization based on bacterial chemotaxis. IEEE Trans Evol Comput 6:16–29

Nixon N, Galassi P, Art NYMoM (2007) The Brown sisters: thirty-three years. Museum of Modern Art

Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987

Otto C, Han H, Jain A (2012) How does aging affect facial components? In: European conference on computer vision. Springer, Berlin, pp 189–198

Ouyang W et al (2015) Deepid-net: deformable deep convolutional neural networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2403–2412

Panis G, Lanitis A (2014) An overview of research activities in facial age estimation using the FG-NET aging database. In: European conference on computer vision. Springer, Berlin, pp 737–750

Park U, Tong Y, Jain AK (2008) Face recognition with temporal invariance: a 3d aging model. In: 8th IEEE international conference on automatic face & gesture recognition, 2008 (FG'08). IEEE, pp 1–7

Park U, Tong Y, Jain AK (2010) Age-invariant face recognition. IEEE Trans Pattern Anal Mach Intell 32:947–954

Parlewar M, Patil H, Bhurchandi K (2016) A novel quantized gradient direction based face image representation and recognition technique. In: 2016 Twenty second national conference on communication (NCC). IEEE, pp 1–6

Patil H, Kothari A, Bhurchandi K (2015) 3-D face recognition: features, databases, algorithms and challenges. Artif Intell Rev 44:393–441

Patil HY, Kothari AG, Bhurchandi KM (2016) Expression invariant face recognition using local binary patterns and contourlet transform. Optik-Int J Light Electron Opt 127:2670–2678

Patterson E, Ricanek K, Albert M, Boone E (2006) Automatic representation of adult aging in facial images. In: Proceedings of the IASTED international conference visualization, imaging, and image processing, pp 171–176

Penev PS, Atick JJ (1996) Local feature analysis: a general statistical theory for object representation. Netw Comput Neural Syst 7:477–500

Phillips PJ, Moon H, Rizvi SA, Rauss PJ (2000) The FERET evaluation methodology for face-recognition algorithms. IEEE Trans Pattern Anal Mach Intell 22:1090–1104

Ramanathan N, Chellappa R (2006) Modeling age progression in young faces. In: 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06). IEEE, pp 387–394

Ramanathan N, Chellappa R (2008) Modeling shape and textural variations in aging faces. In: 8th IEEE international conference on automatic face & gesture recognition, 2008 (FG'08). IEEE, pp 1–8

Ramanathan N, Chellappa R, Biswas S (2009) Computational methods for modeling facial aging: a survey. J Vis Lang Comput 20:131–144

Ricanek K, Tesafaye T (2006) Morph: a longitudinal image database of normal adult age-progression. In: 7th International conference on automatic face and gesture recognition (FGR06). IEEE, pp 341–345

Ricanek K, Sethuram A, Patterson EK, Albert AM, Boone EJ (2008) Craniofacial aging Wiley handbook of science and technology for homeland security

Rizvi SA, Phillips PJ, Moon H (1998) A verification protocol and statistical performance analysis for face recognition algorithms. In: Proceedings of the 1998 IEEE computer society conference on computer vision and pattern recognition. IEEE, pp 833–838

Singh M, Nagpal S, Singh R, Vatsa M (2014) On recognizing face images with weight and age variations. IEEE Access 2:822–830

Sun Y, Chen Y, Wang X, Tang X (2014) Deep learning face representation by joint identification-verification. In: Advances in neural information processing systems, pp 1988–1996

Sungatullina D, Lu J, Wang G, Moulin P (2013) Multiview discriminative learning for age-invariant face recognition. In: 2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG). IEEE, pp 1–6

Thompson DW (1917) On growth and form Cambridge, England

Turk MA, Pentland AP (1991) Face recognition using eigenfaces. In: Proceedings of the CVPR'91, IEEE computer society conference on computer vision and pattern recognition, 1991. IEEE, pp 586–591

Uludag U, Ross A, Jain A (2004) Biometric template selection and update: a case study in fingerprints. Pattern Recogn 37:1533–1542

Wang J, Shang Y, Su G, Lin X (2006) Age simulation for face recognition. In: 18th International conference on pattern recognition (ICPR'06). IEEE, pp 913–916

Wen Y, Li Z, Qiao Y (2016) Latent factor guided convolutional neural networks for age-invariant face recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4893–4901

Xu C, Liu Q, Ye M (2017) Age invariant face recognition and retrieval by coupled auto-encoder networks. Neurocomputing 222:62–71

Yadav D, Vatsa M, Singh R, Tistarelli M (2013) Bacteria foraging fusion for face recognition across age progression. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 173–179

Yang J, Zhang D, J-y Yang, Niu B (2007) Globally maximizing, locally minimizing: unsupervised discriminant projection with applications to face and palm biometrics. IEEE Trans Pattern Anal Mach Intell 29:650–664

Yi D, Lei Z, Liao S, Li SZ (2014) Learning face representation from scratch arXiv preprint arXiv:14117923