Personalized Age Progression with Bi-Level Aging Dictionary Learning

Xiangbo Shu, Jinhui Tang, *Senior Member, IEEE*, Zechao Li,
Hanjiang Lai, Liyan Zhang, and Shuicheng Yan, *Fellow, IEEE*

Abstract—Age progression is defined as aesthetically re-rendering the aging face at any future age for an individual face. In this work, we aim to automatically render aging faces in a personalized way. Basically, for each age group, we learn an aging dictionary to reveal its aging characteristics (e.g., wrinkles), where the dictionary bases corresponding to the same index yet from two neighboring aging dictionaries form a particular aging pattern cross these two age groups, and a linear combination of all these patterns expresses a particular personalized aging process. Moreover, two factors are taken into consideration in the dictionary learning process. First, beyond the aging dictionaries, each person may have extra personalized facial characteristics, e.g., mole, which are invariant in the aging process. Second, it is challenging or even impossible to collect faces of all age groups for a particular person, yet much easier and more practical to get face pairs from neighboring age groups. To this end, we propose a novel Bi-level Dictionary Learning based Personalized Age Progression (BDL-PAP) method. Here, bi-level dictionary learning is formulated to learn the aging dictionaries based on face pairs from neighboring age groups. Extensive experiments well demonstrate the advantages of the proposed BDL-PAP over other state-of-the-arts in term of personalized age progression, as well as the performance gain for cross-age face verification by synthesizing aging faces.

Index Terms—Age progression, aging dictionary, face synthesis, dictionary learning

---

1 INTRODUCTION

AGE progression [1], also called age synthesis [2] or face aging [3], is defined as aesthetically rendering a face image with natural aging and rejuvenating effects for an individual face. It has been widely applied to various application domains, e.g., cross-age face analysis [4], authentication systems, finding lost children, entertainment, etc. There are two main categories of solutions to the age progression task: prototyping-based age progression [5], [6], [7] and modeling-based age progression [3], [8], [9]. Prototyping-based age progression transfers the difference between two prototypes (e.g., average faces) of the pre-divided source age group and the target age group into the input individual face, of which its age belongs to the source age group. Modeling-based age progression models the facial parameters of different ages (age ranges) for the shape and texture synthesis.

- *X. Shu, J. Tang, and Z. Li are with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China. E-mail: {shuxb, jinhuitang, zechao.li}@njust.edu.cn.*
- *H. Lai is with the School of Data and Computer Science, Sun Yat-Sen University, Guangzhou 510275, China. E-mail: laihanj@gmail.com.*
- *L. Zhang is with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China and the Collaborative Innovation Center of Novel Software Technology and Industrialization, Nanjing 210023, China. E-mail: zhangliyan@nuaa.edu.cn.*
- *S. Yan is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576. E-mail: eleyans@nus.edu.sg.*

Intuitively, the natural aging process of a specific person usually follows the general rules of the human aging process. Meanwhile, the natural aging face of a specific person also contains some personalized facial characteristics, e.g., mole, birthmark, etc., which are almost invariant with time. Generally, prototyping-based age progression methods cannot well preserve this personality of an individual face, since they are based on the general rules of the human aging process for a relatively large population. And modeling-based age progression methods do not specially consider the personalized details for a specific person. Moreover, they require a large number of dense long-term (e.g., age span exceeds 20 years) face aging sequences for building the complex functions. However, collecting these required sequences in the real world is very difficult or even impossible. Fortunately, we have observed that the short-term (e.g., age span smaller than 10 years) face aging sequences are available in the Internet, such as photos of celebrities at different ages on Facebook/Twitter. Some available face aging datasets [10], [11], [12] also contain the dense short-term sequences. Therefore, generating personalized age progression for an individual face by leveraging short-term face aging sequences is more feasible.

In this work, we propose a Bi-level Dictionary Learning based Personalized Age Progression (BDL-PAP) method, which leverages short-term face aging pairs to automatically render aging faces in a personalized way, as shown in Fig. 1. Primarily, based on the aging-(in)variant patterns in the face aging process, an individual face can be decomposed into an aging layer and a personalized layer. The former shows the general aging characteristics (e.g., wrinkles), while the latter shows the personalized facial characteristics (e.g., mole). For
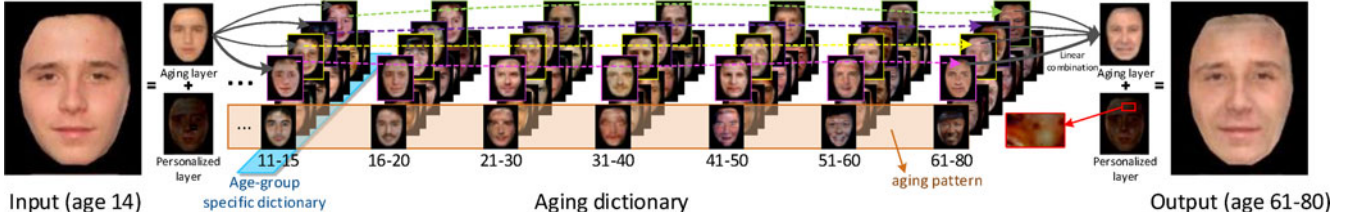
Fig. 1. A personalized aging face generated by the proposed method. This aging face contains the aging layer (e.g., wrinkles) and the personalized layer (e.g., mole). The former can be seen as the corresponding face in a linear combination of the aging patterns, while the latter is invariant in the aging process. Better view in color.

different human age groups (e.g., 11-15, 16-20, etc.), we design the corresponding aging dictionaries to reveal the general aging characteristics, where the dictionary bases with the same index yet from two neighboring aging dictionaries form a particular aging patterns (e.g., they are linked by a dotted line in Fig. 1). Therefore, the aging layer of the aging face can be reconstructed by a linear combination of several aging dictionary bases with sparse representation. The motivation of the sparsity assumption is to use fewer dictionary bases for reconstruction such that the reconstructed aging layer of face can be shaper and less blurred [13], [14]. The residual between the reconstructed aging layer and the input face can be defined as the personalized layer, which is invariant in the aging process. Finally, we render the aging face in a future age group for an input face by fusing the reconstructed aging layer of this age group and the personalized layer.

To learn the aging dictionaries, we propose a bi-level dictionary learning model, and utilize the more practical short-term face aging pairs as the training set instead of the possibly unavailable long-term face aging sequences. In the bi-level

dictionary learning, we assume that the sparse representation of a younger-aging layer of one face w.r.t. the younger-aging dictionary can directly reconstruct its older-aging layer using the older-aging dictionary. The distribution of the collected face aging pairs in this work is shown in the upper part of Fig. 2. We can see that: 1) each age group has its own aging dictionary; 2) every two neighboring age groups are linked by several dense short-term face aging pairs, which makes all the age groups linked together; 3) the personalized layer contains the personalized facial characteristics. These three properties are able to guarantee that all aging dictionaries can be trained well by the bi-level dictionary learning on the short-term face aging pairs.

The main contributions of this work are summarized as follows:

- We propose a Personalized Age Progression method based on the Bi-level Dictionary Learning to render aging faces, which can preserve the personalized facial characteristics.
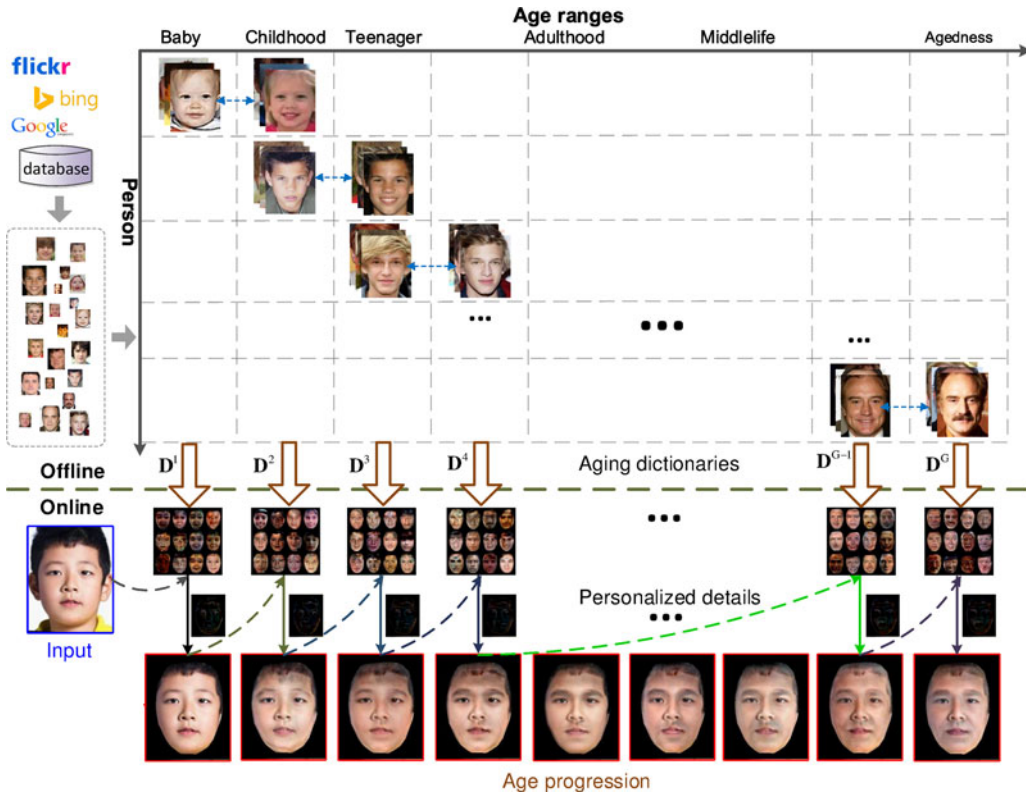

Fig. 2. Framework of the proposed age progression. $\mathbf{D}^g$ denotes a aging dictionary of the gth age group. In the offline phase, we collect short-term aging face pairs and then train the aging dictionary. In the online phase, for an input face, we render its aging faces by bi-level optimization on the corresponding aging dictionaries.

- Since it is challenging or even impossible to collect intra-person face sequences of all age groups, the proposed method only requires the easy-acquired short-term face aging pairs to learn all aging dictionary bases of human aging, which is more feasible.
- Extensive experiments well validate the advantage of the proposed solution over other state-of-the-arts in term of personalized aging progression, as well as the performance gain for cross-age face verification by synthesizing aging faces.

Compared to our preliminary work in [15], we have the following extensions in this paper: 1) we extend the preliminary coupled dictionary learning model to a novel bi-level dictionary learning model, which leads to a more effective and efficient personalized age progression method; 2) we compare the proposed age progression method with more state-of-the-art methods by conducting age progression experiments and cross-age face verification experiments; 3) we illustrate the effectiveness of the proposed personalized age progression method by comparing to its unpersonalized version; 4) we show the efficiency of the proposed method by comparing the executing time of different age progression methods.

The rest of this paper is organized as follows. Section 2 reviews related works. Section 3 details the proposed method, including the whole framework, dictionary learning model, objective function, optimization procedure, and age progression synthesis. Experiments are conducted in Section 4. Finally, Section 5 concludes this work and discusses the future works.

2 RELATED WORK

2.1 Age Progression Methods

Age progression has been comprehensively reviewed in literature [1], [2], [16], [17]. As one of the early studies, Burt et al. [18] focused on creating average faces for different ages and transferring the facial difference between the average faces into the input face. This method gives an insight into the age progression task. Thereafter, some prototyping-based aging methods [5], [6], [19] are proposed to improve the age progression method in [18]. Kemelmacher et al. [5] proposed an Illumination-Aware Age Progression (IAAP) method by leveraging the difference between the warped average faces instead of the original average faces. Here the warped average face is computed based on the flow from the average face to the input face. Generally, the aging progression of an individual is stochastic and non-deterministic in the time dimension. Therefore, Shu et al. [19] presented a kinship-guided age progression approach to automatically generate aging faces by leveraging kinship information. The aging faces generated by these methods almost have no personalized characteristic and the aging speed of different people is synchronous. Although some researchers tried to consider the individual-specific face age progression [20], [21], [22], lack of personality is still a challenging problem.

Modeling-based age progression is the other type of the age progression, which considers shape and texture synthesis simultaneously [23]. Some modeling-based age progression methods have been proposed, including active appearance model [24], craniofacial growth model [25], and-or graph model [3], statistical model [26], implicit function [27], [28],

etc. Generally, to model large appearance changes over a long-term face aging sequence, modeling-based age progression requires sufficient training data. Suo et al. [29] attempted to learn long-term aging patterns from available short-term aging databases by a proposed concatenational graph evolution aging model. Recent years, inspired by the success of Recurrent Neural Network (RNN), Wang et al. [30] exploited RNNs to model the whole aging sequence. The bottom layer in a RNN works as an encoder, which projects the image to a high-dimension space. The top layer in a RNN works as a decoder, which decodes the hidden representation to an aging face.

2.2 Age Progression Based Face Analysis

Face analysis (e.g., face verification and face recognition) [31], [32], [33] has achieved great progress in the last decade. Usually, as a natural biological process, face aging can change the appearance and texture of the facial landmarks, such as wrinkles, senile plaque, mustache, etc. Therefore, the performance of existing face analysis methods for the cross-age face analysis will degree since the age gap exists between the two faces of one person at different ages. By conducting experiments on a passport photo database, Ling et al. [34] validated that the face aging would increase the challenge of face recognition. Towards this end, some methods are proposed to address the cross-age face analysis problem in recent years [4], [10], [34], [35], [36], [37], [38], [39], [40], [41]. These methods can be divided into two categories. Methods of one category aim to learn the age-invariant features for cross-age face analysis [10], [34], [38], [42]. For example, Ling et al. [34] found that the gradient direction is insensitive to the face aging, and then proposed to use the gradient directions of pyramid structure as the feature descriptors of face images. Gong et al. [42] proposed to separate the face pattens into the age-invariant and age-variant patterns, and then use the age-invariant patterns as the inputs of learning model. We can conclude that how to learn the age-invariant features or extract the age-invariant patterns of face images is the key point. Methods of the other category are proposed to use the age synthesis to eliminate the age gap [3], [4], [29], [36], [37], [39], [43], [44], [45]. Taking the face verification as an example, for two faces of one person at different ages, the age progression method renders a synthesized aging face from the younger face. And then, the synthesized face instead of the original face is used to implement the face verification.

2.3 Face Aging Datasets

FG-NET [11] is one popular face aging dataset. It has been used to evaluate the age estimation [46], [47], cross-age face verification, cross-age face recognition [37], [48], and age progression [3], [5], [29], [49]. This dataset contains 1,002 face photos from 82 persons within 0-69 age range: about 64 percent of the images are from children (with ages < 18), and around 36 percent are from adults (with ages $\geqslant 18$). All photos are taken by digital camera and film-based camera. Morph [12] is one of the largest face aging dataset. It has been also used to evaluate the age estimation, age progression, and face recognition. It is composed of Album1 subset and Album2 subset. The Album1 subset contains 1,690 face

images from 515 persons. And the Album2 subset contains 94,000 face images from 24,000 people, which are collected from different places. In Morph, the size of all images is resized to 200×240 or 400×480, and each image has the corresponding age, gender and race labels. Recently, about 55,000 face images from 13,000 people in Morph are available on the website. In recent years, Chen et al. [10] released a new Cross-Age Celebrity Dataset (CACD), which is collected from image search websites. CACD contains 163,446 face images from 2,000 celebrities within 16-62 age range. The age label of each face photo is estimated by referring to its shooting time, namely the shooting time is the age label. Although these age labels are inaccurate, the relative age for the same person is accurate. Thus, this dataset can be used for the cross-age face search and recognition.

3　THE PROPOSED METHOD

3.1　Overview of the Framework

The framework of the proposed age progression method is shown in Fig. 2. The offline phase is described as follows. First, we collect the dense short-term aging pairs of the same persons from Internet and also from available datasets. Second, for each predefined age group, we learn a corresponding aging dictionary to represent its aging characteristics by the proposed bi-level dictionary learning model. In the online phase, for an input face, we generate its aging faces step by step, from the current age to the target age. Specifically, we first generate the aging face in the next age group by the corresponding aging dictionary with an implicitly common sparse representation, as well as a personalized layer. After that, taking this new aging face as the input, and repeat the above process until all aging faces have been rendered.

3.2　Coupled Dictionary Learning

We divide the human age range into G age groups (each group spans less than 10 years) in this work. Let $\{\mathbf{x}_i^1, \ldots, \mathbf{x}_i^g, \ldots, \mathbf{x}_i^G\}$ denote a selected face aging sequence of the ith person covering all age groups, where $i = 1, 2, \ldots, L$ (L is the number for the persons). Here, the face photo $\mathbf{x}_i^g \in \mathbb{R}^f$ falls into the gth age group, where f is the number of pixels in a face photo. For the gth age group ($g = 1, 2, \ldots, G$), we define its aging dictionary $\mathbf{B}^g$ to capture the aging characteristics, which will be learned in the following.

Personality-Aware Formulation. The aging dictionary learning in this work considers the personalized details of an individual when representing the face aging sequences on their own aging dictionaries. Since the personalized characteristics are aging-invariant, such as mole, birthmark, permanent scar, etc., we plan to add a personalized layer $\mathbf{p}_i \in \mathbb{R}^f$ for a face aging sequence $\{\mathbf{x}_i^1, \mathbf{x}_i^2, \ldots, \mathbf{x}_i^G\}$ to indicate the personalized details of the ith person in the aging process. Moreover, considering the computational efficiency, we utilize PCA to reduce the dimension of the dictionary. Let $\mathbf{H}^g \in \mathbb{R}^{f \times m}$ denote a PCA projected matrix of all data in the gth age group, and we have $\mathbf{B}^g = \mathbf{H}^g \mathbf{D}^g$. Thus the original aging dictionary $\mathbf{B}^g$ is redefined as $\mathbf{D}^g \in \mathbb{R}^{m \times k}$, where k denotes the number of dictionary bases. All aging dictionaries compose an overall aging dictionary $\mathbf{D} = [\mathbf{D}^1, \mathbf{D}^2, \ldots, \mathbf{D}^G] \in \mathbb{R}^{m \times K}$, where $K = k \times G$. So far, the aging face $\hat{\mathbf{x}}_i^{g+t}$ of $\mathbf{x}_i^g$ equals the linearly weighted combination of the aging dictionary bases

in the $(g+t)$th age group and the personalized layer $\mathbf{p}_i$, i.e., $\hat{\mathbf{x}}_i^{g+t} \approx \mathbf{H}^{g+t}\mathbf{D}^{g+t}\mathbf{a}_i + \mathbf{p}_i$ for $t = 1, \ldots, G - g$, where $\mathbf{a}_i$ and $\mathbf{p}_i$ are the common sparse representation and the personalized layer, respectively. For L face aging sequences $\{\mathbf{x}_i^1, \ldots, \mathbf{x}_i^G\}_{i=1}^L$ covering all age groups, a personality-aware dictionary learning model is formulated as follows:

$$\min_{\substack{\{\mathbf{D}^g\}_{g=1}^G, \\ \{\mathbf{a}_i, \mathbf{p}_i\}_{i=1}^L}} \sum_{g=1}^G \sum_{i=1}^L \left\{ \left\| \mathbf{x}_i^g - \mathbf{H}^g\mathbf{D}^g\mathbf{a}_i - \mathbf{p}_i \right\|_2^2 + \gamma\|\mathbf{p}_i\|_2^2 + \lambda_1\|\mathbf{a}_i\|_1 \right\}$$
$$\text{s.t. } \|\mathbf{D}^g(:, d)\|_2 \leq 1, \forall d \in \{1, \cdots, k\}, \forall g \in \{1, \cdots, G\}, \tag{1}$$

where $\mathbf{D}^g(:, d)$ denotes the dth column (base) of $\mathbf{D}^g$, λ_1 and γ control the sparsity penalty and regularization, respectively. $\mathbf{D}^g(:, d)$ is used to represent the specific aging characteristics in the gth age group.

Short-Term Coupled Learning. We observe that one person always has the dense short-term face aging photos, but no long-term face aging photos covering all age groups. Collecting these long-term face aging sequences in the real world is extremely difficult or even unlikely. Therefore, we have to use the shot-term face aging pairs instead of the long-term face sequences. Here, each face aging pair includes a certain person's two face images spanning two neighboring age groups. Let $\mathcal{S}^g$ ($g = 1, 2, \ldots, G$) denote a set of face images in the gth age group, we assume that $\mathcal{S}^g$ and $\mathcal{S}^{g+1}$ share n^g face pairs from n^g different persons. Formally, we use $\{\mathbf{x}_{i^g} \in \mathcal{S}^g, \mathbf{y}_{i^g} \in \mathcal{S}^{g+1}\}$ to denote the face pair of the i^gth person, where $i^g = 1, 2, \ldots, n^g$ is a local index related to the gth age group. For the face aging pairs $\{\mathbf{x}_{i^g}, \mathbf{y}_{i^g}\}_{i^g=1}^{n^g}$ spanning the gth and $(g+1)$th age groups, we reformulate a coupled dictionary learning model to simultaneously learn all aging dictionaries, i.e.,

$$\min_{\substack{\mathbf{D}^g, \\ \{\mathbf{a}_{i^g}, \mathbf{p}_{i^g}\}_{i^g=1}^{n^g}}} \sum_{g=1}^{G-1} \sum_{i^g=1}^{n^g} \left\{ \left\| \mathbf{x}_{i^g} - \mathbf{H}^g\mathbf{D}^g\mathbf{a}_{i^g} - \mathbf{p}_{i^g} \right\|_2^2 + \gamma\|\mathbf{p}_{i^g}\|_2^2 \right.$$
$$\left. + \left\| \mathbf{y}_{i^g} - \mathbf{H}^{g+1}\mathbf{D}^{g+1}\mathbf{a}_{i^g} - \mathbf{p}_{i^g} \right\|_2^2 + \lambda_1\|\mathbf{a}_{i^g}\|_1 \right\} \tag{2}$$
$$\text{s.t. } \|\mathbf{D}^g(:, d)\|_2 \leq 1, \forall d \in \{1, \ldots, k\}, \forall g \in \{1, \ldots, G\}.$$

In Eq. (2), every two neighboring aging dictionaries $\mathbf{D}^g$ and $\mathbf{D}^{g+1}$ corresponding to two age groups are implicitly coupled via the common sparse representation $\mathbf{a}_{i^g}$, and the personalized layer $\mathbf{p}_{i^g}$ is to capture the personalized details of the i^gth person in the gth age group, who has the face pair $\{\mathbf{x}_{i^g}, \mathbf{y}_{i^g}\}$ spanning the gth and $(g+1)$th age groups. Let $\mathbf{D} = [\mathbf{D}^1, \ldots, \mathbf{D}^G] \in \mathbb{R}^{m \times K}$, $\mathbf{X}^g = [\mathbf{x}_{1^g}, \ldots, \mathbf{x}_{n^g}] \in \mathbb{R}^{f \times n^g}$, $\mathbf{Y}^g = [\mathbf{y}_{1^g}, \ldots, \mathbf{y}_{n^g}] \in \mathbb{R}^{f \times n^g}$, $\mathbf{P}^g = [\mathbf{p}_{1^g}, \ldots, \mathbf{p}_{n^g}] \in \mathbb{R}^{f \times n^g}$, and $\mathbf{A}^g = [\mathbf{a}_{1^g}, \ldots, \mathbf{a}_{n^g}] \in \mathbb{R}^{k \times n^g}$, and Eq. (2) can be rewritten in the matrix form

$$\min_{\mathbf{D}^g, \mathbf{A}^g, \mathbf{P}^g} \sum_{g=1}^{G-1} \left\{ \left\| \mathbf{X}^g - \mathbf{H}^g\mathbf{D}^g\mathbf{A}^g - \mathbf{P}^g \right\|_F^2 + \gamma\|\mathbf{P}^g\|_F^2 \right.$$
$$\left. + \left\| \mathbf{Y}^g - \mathbf{H}^{g+1}\mathbf{D}^{g+1}\mathbf{A}^g - \mathbf{P}^g \right\|_F^2 + \lambda_1\|\mathbf{A}^g\|_1 \right\} \tag{3}$$
$$\text{s.t. } \|\mathbf{D}^g(:, d)\|_2 \leq 1, \forall d \in \{1, \ldots, k\}, \forall g \in \{1, \ldots, G\},$$

where $\|\mathbf{A}^g\|_1 = \sum_{i=1}^{n^g} \|\mathbf{a}_{i^g}\|_1$.

In Eq. (3), the optimization w.r.t. $\mathbf{P}^g$ has the closed-form solution by fixing $\mathbf{D}^g$ and $\mathbf{A}^g$, where $g = 1, 2, \ldots, G - 1$. By fixing $\mathbf{P}^g$, the optimization w.r.t. $\mathbf{D}^g$ and $\mathbf{A}^g$ becomes a joint sparse coding problem [50]. Such joint sparse coding problem can be solved in the concatenated feature space of aging layers $(\mathbf{X}^g - \mathbf{P}^g)$ and $(\mathbf{Y}^g - \mathbf{P}^g)$, but not in each feature space separately, namely we have

$$\min_{\mathbf{D}^g, \mathbf{A}^g} \sum_{g=1}^{G-1} \left\{ \left\| \begin{bmatrix} \mathbf{X}^g - \mathbf{P}^g \\ \mathbf{Y}^g - \mathbf{P}^g \end{bmatrix} - \begin{bmatrix} \mathbf{H}^g \mathbf{D}^g \\ \mathbf{H}^{g+1} \mathbf{D}^{g+1} \end{bmatrix} \mathbf{A}^g \right\|_F^2 + \lambda_1 \|\mathbf{A}^g\|_1 \right\}$$
$$s.t. \|\mathbf{D}^g(:,d)\|_2 \leq 1, \forall d \in \{1, \ldots, k\}, \forall g \in \{1, \ldots, G\}, \tag{4}$$

which can be effectively solved by the SPAMS toolbox.[1]

3.3 Bi-Level Dictionary Learning

If the aging dictionaries $\mathbf{D}^1, \ldots, \mathbf{D}^G$ are learned by the coupled dictionary learning model, for an input face $\mathbf{x}^g$ belonging to the gth age group, its sparse representation $\mathbf{a}^g$ and personalized layer $\mathbf{p}^g$ should be computed to generate its aging face aging face $\hat{\mathbf{y}}^g$ in the $(g+1)$th age group. According to Eq. (3), $\mathbf{a}^g$ and $\mathbf{p}^g$ should be calculated by solving the following problem,

$$\min_{\mathbf{a}^g, \mathbf{p}^g} \|\mathbf{x}^g - \mathbf{H}^g \mathbf{D}^g \mathbf{a}^g - \mathbf{p}^g\|_2^2 + \gamma\|\mathbf{p}^g\|_2^2$$
$$+ \|\mathbf{y}^g - \mathbf{H}^{g+1} \mathbf{D}^{g+1} \mathbf{a}^g - \mathbf{p}^g\|_2^2 + \lambda_1 \|\mathbf{a}^g\|_1. \tag{5}$$

Unfortunately, $\mathbf{y}^g$ is unavailable for a new coming face $\mathbf{x}^g$ in practice. Thus, there is no way to enforce the equivalence constraint on the representations of $\mathbf{x}^g$ and $\mathbf{y}^g$, as has been done in the training phase of Eq. (3). To handle this problem, an intuitive way it to use the average face in the target age group to replace $\mathbf{y}^g$, as done in the preliminary work [15]. However, this solution has two drawbacks. First, since the average face and the target aging face have the different facial characteristics, the aging result is not satisfied. Second, to obtain the desired aging faces, we need to repeat the process of age progression synthesis several times to improve the aging performance, which is time-consuming.

To better address the above problem, let us revisit our primary goal again. We want to learn the aging dictionaries $\mathbf{D}^1, \ldots, \mathbf{D}^G$ to ensure that the sparse representations of two faces in each pair, as well as their personalized layers, are the same. If so, we can generate the aging face $\hat{\mathbf{y}}^g$ for an input face $\mathbf{x}^g$ by using the dictionary $\mathbf{D}^{g+1}$, the sparse representation $\mathbf{a}^g$ and the personalized layer $\mathbf{p}^g$, i.e., $\hat{\mathbf{y}}^g = \mathbf{H}^{g+1} \mathbf{D}^{g+1} \mathbf{a}^g + \mathbf{p}^g$. Formally, an ideal pair of every two neighboring aging dictionaries $\mathbf{D}^g$ and $\mathbf{D}^{g+1}$ should satisfy the following two equations for all face pairs $\{\mathbf{x}_{ig}, \mathbf{y}_{ig}\}_{ig=1}^{n^g}$ spanning the gth and $(g+1)$th age groups,

$$\{\mathbf{a}_{ig}^*, \mathbf{p}_{ig}^*\} = \arg\min_{\mathbf{a}_{ig}, \mathbf{p}_{ig}} \sum_{ig=1}^{n^g} \left\{ \|\mathbf{x}_{ig} - \mathbf{H}^g \mathbf{D}^g \mathbf{a}_{ig} - \mathbf{p}_{ig}\|_2^2 \right.$$
$$\left. + \gamma\|\mathbf{p}_{ig}\|_2^2 + \lambda_1 \|\mathbf{a}_{ig}\|_1 \right\}; \tag{6}$$

$$\{\mathbf{a}_{ig}^*, \mathbf{p}_{ig}^*\} = \arg\min_{\mathbf{a}_{ig}, \mathbf{p}_{ig}} \sum_{ig=1}^{n^g} \|\mathbf{y}_{ig} - \mathbf{H}^{g+1} \mathbf{D}^{g+1} \mathbf{a}_{ig} - \mathbf{p}_{ig}\|_2^2. \tag{7}$$

Based on the above analysis, we should first learn $\mathbf{a}^g$ and $\mathbf{p}^g$ by Eq. (6). And then $\mathbf{D}^g$ and $\mathbf{D}^{g+1}$ should be learned to satisfy Eqs. (6) and (7). To this end, we can reformulate our objective function as

$$\min_{\mathbf{D}^g, \mathbf{D}^{g+1}} \sum_{ig=1}^{n^g} \left\{ \|\mathbf{x}_{ig} - \mathbf{H}^g \mathbf{D}^g \mathbf{a}_{ig} - \mathbf{p}_{ig}\|_2^2 \right.$$
$$\left. + \|\mathbf{y}_{ig} - \mathbf{H}^{g+1} \mathbf{D}^{g+1} \mathbf{a}_{ig} - \mathbf{p}_{ig}\|_2^2 \right\}$$
$$s.t. \{\mathbf{a}_{ig}, \mathbf{p}_{ig}\} = \arg\min_{\mathbf{a}_{ig}, \mathbf{p}_{ig}} \sum_{ig=1}^{n^g} \left\{ \|\mathbf{x}_{ig} - \mathbf{H}^g \mathbf{D}^g \mathbf{a}_{ig} - \mathbf{p}_{ig}\|_2^2 \right.$$
$$\left. + \gamma\|\mathbf{p}_{ig}\|_2^2 + \lambda_1 \|\mathbf{a}_{ig}\|_1 \right\}$$
$$\|\mathbf{D}^c(:,d)\|_2 \leq 1, \ \forall d \in \{1, \ldots, k\}, \ c = \{g, g+1\}, \tag{8}$$

which is a bi-level optimization problem [51].

Rewrite Eq. (8) in the matrix form after some algebraic steps, we can obtain a novel bi-level dictionary learning model for all face pairs $\{\mathbf{x}_{ig}, \mathbf{y}_{ig}\}_{ig=1}^{n^g}$ spanning the gth and $(g+1)$th age groups $(g = 1, 2, \ldots, G-1)$

$$\min_{\mathbf{D}^g, \mathbf{D}^{g+1}} J(\mathbf{D}^g, \mathbf{D}^{g+1})$$
$$s.t. \mathbf{A}^g = \arg\min_{\mathbf{Z}^g} \|\mathbf{X}^g - \mathbf{H}^g \mathbf{D}^g \mathbf{Z}^g - \mathbf{P}^g\|_F^2 + \lambda_1 \|\mathbf{Z}^g\|_1$$
$$+ \lambda_2 \|\mathbf{Z}^g\|_F^2$$
$$\mathbf{P}^g = \arg\min_{\mathbf{Q}^g} \|\mathbf{X}^g - \mathbf{H}^g \mathbf{D}^g \mathbf{Z}^g - \mathbf{Q}^g\|_F^2 + \gamma\|\mathbf{Q}^g\|_F^2$$
$$\|\mathbf{D}^c(:,d)\|_2 \leq 1, \ \forall d \in \{1, \ldots, k\}, \ \text{and } c = \{g, g+1\}, \tag{9}$$

where $J(\mathbf{D}^g, \mathbf{D}^{g+1}) = \|\mathbf{X}^g - \mathbf{H}^g \mathbf{D}^g \mathbf{A}^g - \mathbf{P}^g\|_F^2 + \|\mathbf{Y}^g - \mathbf{H}^{g+1} \mathbf{D}^{g+1} \mathbf{A}^g - \mathbf{P}^g\|_F^2$, $\mathbf{Q}^g = [\mathbf{q}_{1g}, \ldots, \mathbf{q}_{n^g}] \in \mathbb{R}^{f \times n^g}$, $\mathbf{Z}^g = [\mathbf{z}_{1g}, \ldots, \mathbf{z}_{n^g}] \in \mathbb{R}^{k \times n^g}$, and λ_2 controls the regularization.

According to the new formulation in Eq. (9), for a new coming face $\mathbf{x}^g$, the calculations of its sparse representation $\mathbf{a}^g$ and personalized layer $\mathbf{p}^g$ are independent with the unavailable $\mathbf{y}^g$, while they only depend on the available face $\mathbf{x}^g$.

3.4 Optimization Procedure

The objective function in Eq. (9) is convex w.r.t. $\mathbf{D}^{g+1}$ and $\mathbf{D}^g$ separately, which can be iteratively solved through two alternating sub-procedures of optimization. Specifically, we fix the other variables when updating one variable.

3.4.1 Updating $\mathbf{D}^{g+1}$

We update $\mathbf{D}^{g+1}$ by fixing $\mathbf{D}^g$, and then the objective function in Eq. (9) becomes

$$\min_{\mathbf{D}^{g+1}} \|\mathbf{Y}^g - \mathbf{H}^{g+1} \mathbf{D}^{g+1} \mathbf{A}^g - \mathbf{P}^g\|_F^2$$
$$s.t. \mathbf{A}^g = \arg\min_{\mathbf{Z}^g} \|\mathbf{X}^g - \mathbf{H}^g \mathbf{D}^g \mathbf{Z}^g - \mathbf{P}^g\|_F^2 + \lambda_1 \|\mathbf{Z}^g\|_1$$
$$+ \lambda_2 \|\mathbf{Z}^g\|_F^2$$
$$\mathbf{P}^g = \arg\min_{\mathbf{Q}^g} \|\mathbf{X}^g - \mathbf{H}^g \mathbf{D}^g \mathbf{A}^g - \mathbf{Q}^g\|_F^2 + \gamma\|\mathbf{Q}^g\|_F^2$$
$$\|\mathbf{D}^{g+1}(:,d)\|_2 \leq 1, \ \forall d \in \{1, \ldots, k\}. \tag{10}$$

1. http://spams-devel.gforge.inria.fr/

First, we iteratively update $\mathbf{A}^g$ and $\mathbf{P}^g$, wherein the *lasso* problem can be solved by the SPAMS toolbox. And then, the problem in Eq. (10) becomes a Quadratically Constrained Quadratic Program (QCQP) that can be solved by the CVX toolbox.[2]

Algorithm 1. Bi-Level Dictionary Learning (Offline)

Input: $\{\mathbf{X}^g, \mathbf{Y}^g\}_{g=1}^{G-1}$, $\mathbf{H}^g$ $(g = 1, \ldots, G-1)$, λ_1, λ_2, and γ.
Initialization: $\mathbf{D} = [\mathbf{D}^1, \mathbf{D}^2, \ldots, \mathbf{D}^G]$ by Coupled Dictionary Learning in [15], $n_0 \leftarrow 1$, and $iter \leftarrow 0$.
 1: **for** $g = 1, 2, \ldots, G-1$ **do**
 2:     **repeat**
 3:        **for** $i^g = 1, 2, \ldots, n^g$ **do**
 4:           Update index set Ω based on Definition 1;
 5:           Compute gradient $\Delta = \partial J / \partial \mathbf{D}^g$ with Eq. (11);
 6:           Update $\mathbf{D}^g = \mathbf{D}^g - \eta(n_0) \cdot \Delta^1$;
 7:           Project each atom of $\mathbf{D}^g$ onto the unit ball;
 8:           $n_0 \leftarrow n_0 + 1$.
 9:        **end for**
10:        Update $\mathbf{D}^{g+1}$ with Eq. (3).
11:       $iter \leftarrow iter + 1$.
12:     **until** convergence.
13: **end for**
Output: $\mathbf{D} = [\mathbf{D}^1, \mathbf{D}^2, \ldots, \mathbf{D}^G]$.

3.4.2 Updating $\mathbf{D}^g$

When updating $\mathbf{D}^g$, we fix $\mathbf{D}^{g+1}$, and then Eq. (9) becomes a bi-level optimization problem. Similar to [50], [51], we solve this bi-level optimization problem based on the first-order projected stochastic gradient descent. For brevity, we simplify the subscripts of $\mathbf{x}_{i^g}$, $\mathbf{x}_{i^g}$, $\mathbf{a}_{i^g}$ and $\mathbf{p}_{i^g}$, i.e., $\mathbf{x}^g$, $\mathbf{y}^g$, $\mathbf{a}^g$ and $\mathbf{p}^g$. For $\mathbf{x}^g$, we compute the gradient of J with respect to $\mathbf{D}^g$

$$\frac{\partial J}{\partial \mathbf{D}^g} = \sum_{j \in \Omega} \frac{\partial (J_x + J_y)}{\partial \tilde{a}_j^g} \frac{\partial \tilde{a}_j^g}{\partial \mathbf{D}^g} + \sum_{i=1}^{f} \frac{\partial (J_x + J_y)}{\partial p_i^g} \frac{\partial p_i^g}{\partial \mathbf{D}^g} + \frac{\partial J_x}{\partial \mathbf{D}^g}, \quad (11)$$

where $J_x = \|\mathbf{X}^g - \mathbf{H}^g \mathbf{D}^g \mathbf{A}^g - \mathbf{P}^g\|_F^2$, and $J_y = \|\mathbf{Y}^g - \mathbf{H}^{g+1} \mathbf{D}^{g+1} \mathbf{A}^g - \mathbf{P}^g\|_F^2$. Ω denotes the index set for j, and is defined as follows,

Definition 1. *Let $\mathbf{w}^g = \mathbf{x}^g - \mathbf{p}^g$, we solve the following lasso problem to obtain the optimizing $\mathbf{a}^g$*

$$\min_{\mathbf{a}^g} \|\mathbf{w}^g - \mathbf{H}^g \mathbf{D}^g \mathbf{a}^g\|_2^2 + \lambda \|\mathbf{a}^g\|_1 + \lambda_2 \|\mathbf{a}^g\|_2^2. \quad (12)$$

If we obtain $\mathbf{a}^g$ by Eq. (12), we can define $\Omega = \{j | |a_j^g| > 0^+\}$ is the index set of nonzero elements of $\mathbf{a}^g$.

In Eq. (11), it is easy to find that

$$\frac{\partial J_x}{\partial \mathbf{D}^g} = -2(\mathbf{H}^g)^T \mathbf{u}^g (\mathbf{a}^g)^T + 2(\mathbf{H}^g)^T \mathbf{H}^g \mathbf{D}^g \mathbf{a}^g (\mathbf{a}^g)^T, \quad (13)$$

where $\mathbf{u}^g = \mathbf{x}^g - \mathbf{p}^g$. The detailed deduction of $\sum_{j \in \Omega} \frac{\partial (J_x + J_y)}{\partial \tilde{a}_j^g} \frac{\partial \tilde{a}_j^g}{\partial \mathbf{D}^g}$ and $\sum_{i=1}^{f} \frac{\partial (J_x + J_y)}{\partial p_i^g} \frac{\partial p_i^g}{\partial \mathbf{D}^g}$ in Eq. (11) can be found in Appendix A.1 and A.2 of the supplemental material, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPAMI.2017.

---

2. http://cvxr.com/cvx/
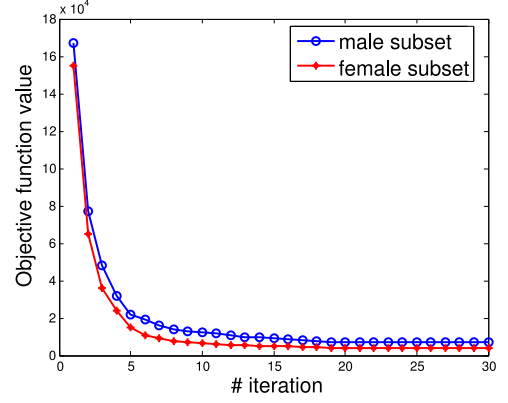here, $\eta(n_0)$ shrinks in the rate of $1/n_0$.

Fig. 3. Convergence curves of the optimization procedure for solving $\mathbf{D}^1$ and $\mathbf{D}^2$.

2705122, respectively. Finally, we can obtain the updating way of $\mathbf{D}^g$, as follows:

$$\mathbf{D}^g = \mathbf{D}^g - \eta \cdot \frac{\partial J}{\partial \mathbf{D}^g}, \quad (14)$$

where η is the step size.

The proposed bi-level dictionary learning algorithm is summarized in Algorithm 1. The convergence criterion is that the iteration steps shall end when the relative cost of objective function is smaller than a pre-defined threshold. The convergence curves for solving $\mathbf{D}^1$ and $\mathbf{D}^2$ on male and female sub datasets respectively are shown in Fig. 3. We found that the algorithm achieves convergence after about 30 iterations.

Algorithm 2. Age Progression Synthesis (Online)

Input: $\mathbf{D} = [\mathbf{D}^1, \ldots, \mathbf{D}^c, \ldots, \mathbf{D}^G]$, input image $\mathbf{x}^g$ in the gth age group, $\{\mathbf{D}^g, \mathbf{D}^{g+1}\}_{g=1}^{G-1}$, λ_1, λ_2 and γ.
1: **for** $t = 1, \ldots, G - g$ **do**
2:     Solve $\mathbf{a}^{g+t-1*}$ and $\mathbf{p}^{g+t-1*}$ with Eq. (6)
3:     Obtain $\hat{\mathbf{y}}^{g+t-1}$:
       $\hat{\mathbf{y}}^{g+t-1} = \mathbf{H}^{g+t} \mathbf{D}^{g+t} \mathbf{a}^{g+t-1*} + \mathbf{a}^{g+t-1*}$;
4: **end for**
Output: $\{\hat{\mathbf{y}}^g, \hat{\mathbf{y}}^{g+1}, \ldots, \hat{\mathbf{y}}^{G-1}\}$.

3.5 Age Progression Synthesis

After learning all aging dictionaries $\mathbf{D}^1, \mathbf{D}^2, \ldots, \mathbf{D}^G$, for a given face $\mathbf{x}^g$ belonging to the gth age group[3], we can generate its aging face sequence $\{\hat{\mathbf{y}}^g, \ldots, \hat{\mathbf{y}}^{G-1}\}$ step by step, from the current age to the target age. Specifically, we first generate the aging face $\hat{\mathbf{y}}^g$ in the next age group (i.e., the $(g + 1)$th age group) by the corresponding aging dictionary with a sparse representation $\mathbf{a}^{g*}$, as well as a personalized layer $\mathbf{p}^{g*}$. Here, $\mathbf{a}^{g*}$ and $\mathbf{p}^{g*}$ are calculated by Eq. (6) in an alternately iterative way. The iteration steps of optimization in Eq. (6) shall end when the relative cost of objective function is smaller than a pre-defined threshold.

When we have obtained $\mathbf{a}^{g*}$ and $\mathbf{p}^{g*}$, the aging face $\hat{\mathbf{y}}^g$ in the next age group (i.e., the $(g + 1)$th age group) can be generated by the following equation:

$$\hat{\mathbf{y}}^g = \mathbf{H}^{g+1} \mathbf{D}^{g+1} \mathbf{a}^{g*} + \mathbf{p}^{g*}. \quad (15)$$

---

3. Here, its age range and gender are estimated by an age estimator and a gender recognition system in [52], respectively.
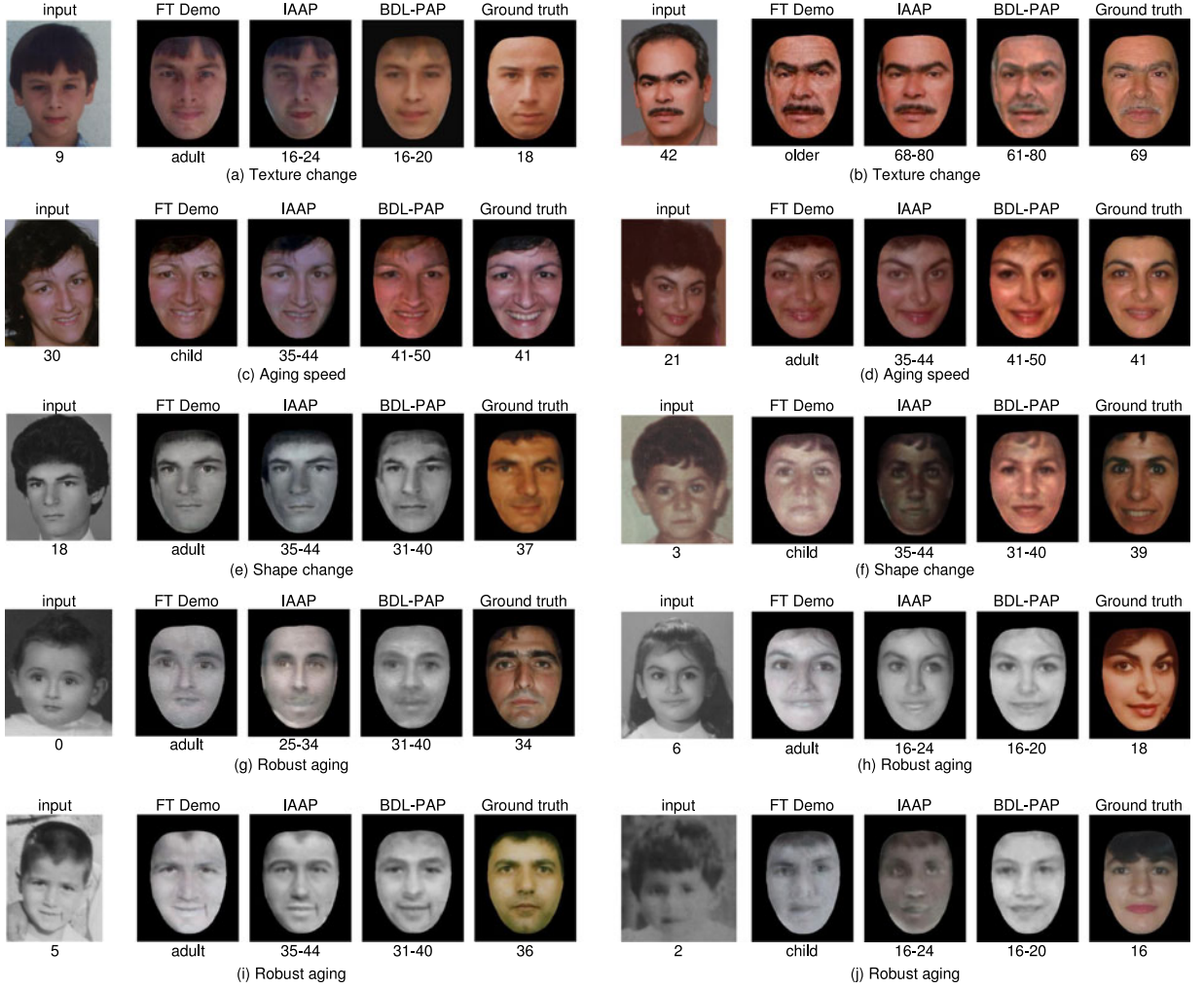
Fig. 4. The comparisons with ground truth and other methods. Each group includes an input face, a ground truth and three aging results generated by different methods. The number or word under each face photo represents the age range (e.g., 61-80) or the age period (e.g., older). For convenience of comparison, black background has been added to each face photo. Better view in color.

After that, taking this new aging face $\hat{\mathbf{y}}^g$ in the $(g + 1)$th age group as the input of age synthesis for the $(g + 2)$th age group. We repeat this process until all aging faces $\hat{\mathbf{y}}^g$, $\hat{\mathbf{y}}^{g+1}, \ldots, \hat{\mathbf{y}}^{G-1}$ are generated.

In the above process, we do not need to calculate the average face for each input face, as done in the preliminary work [15], and then do not need to repeat the process of age progression synthesis. Therefore, we can save much time compared with the preliminary work [15].

4 EXPERIMENTS

4.1 Experimental Settings

Data Collection. To train the high-quality aging dictionary, it is crucial to collect sufficient and dense short-time face aging pairs. We download a large number of face photos covering different ages of the same persons from Google and Bing image searches, social media websites(by some content and context links [53]), and other two available face aging datasets, CACD [10] and MORPH [12]. Both CACD and MORPH contain quite a number of short-term intra-person photos. Since face images from the Internet and CACD dataset are mostly "in the wild", we select the photos with approximately frontal faces (-15 to 15 degree) and

relatively natural illumination and expressions. For all face images, face alignment [54] is implemented to obtain aligned faces, which are cropped into the size $123 \times 98 \times 3$. To boost the aging relationship between the neighboring aging dictionaries, we employ Collection Flow [55] to correct all the faces into the common neutral expression. We group all images (the age from 0 to 80) into 9 age groups (i.e., $G = 9$): 0-5, 6-10, 11-15, 16-20, 21-30, 31-40, 41-50, 51-60, and 61-80 of two genders, and find that no person has aging faces covering all aging groups. Actually, the aging faces of most persons fall into only one or two age groups (i.e., most persons have face photos spanning no more than 20 years). Therefore, we further select those intra-person face photos which densely fall into two neighboring age groups. Finally, there are 3,200 intra-person face pairs for training (1,600 pairs for males, and 1,600 pairs for females).[4] Every two neighboring age groups for one gender share 400 face aging pairs of the same persons. Since male and female have different aging characteristics, we train aging dictionaries for male and female, respectively.

PCA Projection. We stack s images in the gth age group as columns of a data matrix $\mathbf{M}^g \in \mathbb{R}^{f \times s}$, where $s = 400$ and

4. Dataset is released at http://imag.njust.edu.cn/FAD.html

Fig. 5. The comparisons with prior works. Each comparison group includes an input face and three aging results of CDL-PAP, BDL-PAP and prior work. The number under the face photo is the age range. Some worse aging results of BDL-PAP are enclosed by blue box. For convenience of comparison, black background has been added to each face photo.

$f = 123 \times 98 \times 3 = 36,162$ in experiments. The SVD of $\mathbf{M}^g$ is $\mathbf{M}^g = \mathbf{U}^g \mathbf{S}^g (\mathbf{V}^g)^T$. We define the projected matrix $\mathbf{H}^g = \mathbf{U}^g(:, 1:m) \in \mathbb{R}^{f \times m}$, where $\mathbf{U}^g(:, 1:m)$ is truncated to the rank $= m$, and $m = 2,000$ in experiments.

Parameter Setting. The parameters λ_1, λ_2 and γ in Eq. (3) are empirically set as $\lambda_1 = 0.01$, $\lambda_2 = 0.001$ and $\gamma = 0.1$. The number of bases of each aging dictionary $\mathbf{D}^g$ is set as $k = 80$.

Aging Evaluation. We adopt three strategies to comprehensively evaluate the proposed age progression method. First, we qualitatively evaluate the proposed method on FG-NET [11]. We show the age progression for every photo in FG-NET, and do the qualitative comparison with the corresponding ground truth (available older photos) for each person. For reference, we also reproduce some aging results of other representative methods. Second, we conduct user study to test the aging faces of the proposed method compared with the prior works which reported their best aging results. The proposed method uses the same inputs as in these prior works. Third, we compare the proposed method by the cross-age face verification [39], [42]. Cross-age face recognition [10], [56] and cross-age face verification are challenging in extreme facial analysis scenarios due to the age gap, which is similar to the semantic gap between the image and tag in the field of computer vision [57], [58]. A straightforward way for cross-age facial analysis is to use the aging synthesis to reduce the age gap. Specifically, we can synthesize all the faces to their aging faces within the same age range, and then implement the face verification algorithm. In turn, we can also use the face verification to validate whether the intra-person pair of aging face and ground truth face (without age gap) is more similar than the original intra-person face pair with the age gap.

4.2 Qualitative Comparison with Ground Truth

Since FG-NET provides the ground-truth aging faces, we compare the proposed BDL-PAP method with an online Face Transformer demo (i.e., FT Demo[5]), and the representative Illumination-Aware Age Progression method [5] on this dataset. FT Demo requires manual location of facial features, while IAAP uses common aging characteristics of average faces for the age progression of all input faces.

Some aging results generated by these three compared methods are given in Fig. 4, covering from baby/childhood/teenager (input) to adult/agedness (output), as well as from adult (input) to agedness (output). By comparing with ground truth, we can see that the aging results generated by the proposed BDL-PAP look more like the ground truth faces than the aging results of other two methods. In particular, the proposed BDL-PAP can generate personalized aging faces for different individual inputs. In term of texture change, the aging face of BDL-PAP in Fig. 4a has no mustache that is closer to ground truth, while the aging face of BDL-PAP in Fig. 4b has white mustache that is closer to ground truth; In term of shape change, the aging faces of BDL-PAP in Figs. 4e and 4f have more approximate facial outline to the ground truth; In term of aging speed, the faces of FT Demo and IAAP in Fig. 4d are aging more slowly, while one of FT Demo in Fig. 4c is faster. Overall, the age speed of IAAP is slower than ground truth since IAAP is based on smoothed average faces, which maybe loses some facial texture details, such as freckle, nevus, aging spots, etc. FT Demo performs the worst, especially in shape change. Our aging results in Fig. 4 are more similar to the ground truth, which means BDL-PAP can preserve much more

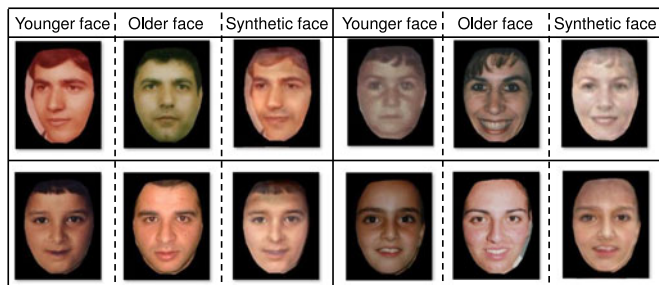5. http://cherry.dcs.aber.ac.uk/Transformer/

Fig. 6. The comparisons of original face pairs and the synthetic pairs by BDL-PAP. The face images in each solid-line box are the same person. Original pair consists of younger face and older face, while synthetic pair consists of synthetic face and older face.

(a) Pair setting. (b) FAR-FRR curve.

Fig. 7. Pair setting and comparative performance of face verification.

personalized results. Moreover, as shown in Figs. 4g and 4h, the aging results of BDL-PAP are more robust than other methods for the input faces with noise or low resolution.

4.3 Quantitative Comparison with Prior Works

Some prior works related to the age progression have posted their best face aging results with input faces at different ages, including [3], [59], [1], [37], [60], [61], [9], [62], [63], [36], [30] and [5]. There are 261 aging results with 87 input faces in total. The proposed BDL-PAP for each input face is implemented to generate the corresponding aging faces at the same ages (ranges) of the posted results.

We conduct user study to compare the aging results of the proposed BDL-PAP with the posted aging results in the prior works, as well as the aging results generated by Coupled Dictionary Learning based Personalized Age Progression (CDL-PAP) in the preliminary work [15]. To avoid bias as much as possible, we invite 50 users covering a wide age range and from all walks of life. For each comparison group including an input face, and three aging results in a random order, all the 50 users are asked to answer the question: which aging face is the best in terms of *Personality* and *Reliability*. *Reliability* means the aging face should be natural and authentic at the synthetic age, while *Personality* means the aging faces for different inputs should be identity-preserved and diverse. User can choose one from the three aging results as the best, and choose "None" if she/he thinks all the three aging results are unsatisfied. There are 50 ratings for each comparison, 261 comparison groups, and then 13,050 ratings in total. The voting results are as follows: 36.5 percent for BDL-PAP best; 34.8 percent for CDL-PAP best; 26.7 percent for prior works best; and 2.0 percent for "none is satisfied". We show some comparison groups for voting in Fig. 5. Overall, for the input face of a person in any age range, BDL-PAP, CDL-PAP and these prior works can generate an authentic and reliable aging face of any older-age range. In particular, for different inputs, aging faces rendered by BDL-PAP and CDL-PAP have more personalized aging characteristics, which further improve the appealing visual sense. For example in Fig. 5,
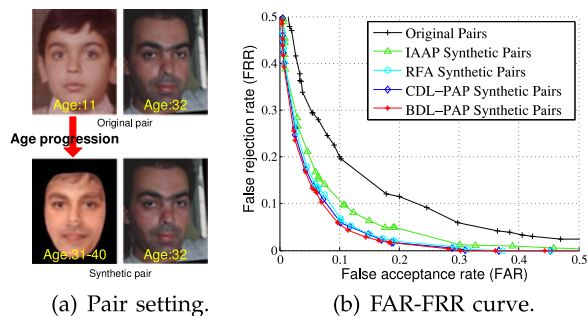
the aging faces of BDL-PAP in the same age range in the first and the second group of the first column have different aging speeds: the former is faster than the latter; the aging faces of prior works with different inputs in the first and second groups of the third column are similar, while the aging results of BDL-PAP and CDL-PAP are more diverse for different individual inputs. Moreover, we can see that the aging results of BDL-PAP are comparable with the aging results of CDL-PAP. Specifically, BDL-PAP can synthesize the more sharp and high-definition aging faces than CDL-PAP.

4.4 Evaluation on Cross-Age Face Verification

To validate the improved performance of cross-age face verification with the help of the proposed BDL-PAP, we prepare for the intra-person pairs and inter-person pairs with cross ages on the FG-NET dataset. By removing undetected face photos and face pairs with age span no more than 20 years, we select 1,832 pairs (916 intra-person pairs and 916 inter-person pairs), called "Original Pairs". Among the 1,832 pairs, we render the younger face in each pair to the aging face with the same age of the older face by the proposed BDL-PAP. Replacing each younger face with the corresponding aging face, we newly construct 1,832 pairs of the aging face and older face, called "BDL-PAP Synthetic Pairs". Fig. 6 shows the comparisons of and the original face pairs and the synthetic pairs by the proposed method, and Fig. 7a shows the pair setting. To evaluate performance of the proposed BDL-PAP, we also prepare the "IAAP Synthetic Pairs", "RFA Synthetic Pairs", and "CDL-PAP Synthetic Pairs" by IAAP [5], Recurrent Face Aging (RFA) [30], and CDL-PAP [15], respectively.

The detailed implementation of face verification is given as follows. First, we formulate a face verification model with deep Convolutional Neural Networks (deep ConvNets), which is based on DeepID2 [64]. Since we focus on the age progression in this work, please refer to [64], [65] for more details of face verification with deep ConvNets. Second, we train this face verification model on the LFW dataset [66], which is constructed for face verification. Third, we test the face verification on different face pairs.

TABLE 1
Equal Error Rates (EER) (%) of Cross-Age Face Verification

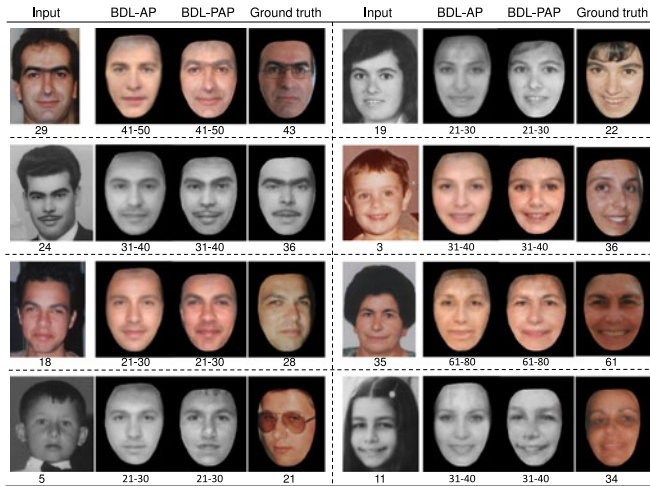| Pair settings	Original Pairs	IAAP Synthetic Pairs	RFA Synthetic Pairs	CDL-PAP Synthetic Pairs	BDL-PAP Synthetic Pairs
EER (%)	14.89	10.36	8.69	8.53	**8.06**

Fig. 8. The comparisons of BDL-PAP and BDL-AP on FG-NET.

The False Acceptance Rate - False Rejection Rate (FAR-FRR) curves and the Equal Error Rates (EER) on original pairs and synthetic pairs are shown in Fig. 7b and Table 1, respectively. We can see that the face verification on BDL Synthetic Pairs achieves the lower ERR than on Original Pairs, IAAP Synthetic Pairs, RFA Synthetic Pairs, and CDL-PAP Synthetic Pairs. This illustrates that the aging faces by the proposed age progression method can effectively mitigate the effect of the age gap in cross-age face verification. The results also validate that, for a given input face, BDL-PAP can render a personalized and authentic aging face closer to the ground truth than the other compared methods.

4.5 Effect of Personalized Layer

To demonstrate the superiority of the personalization introduced by the proposed BDL-PAP, we conduct both qualitative and quantitative comparisons of BDL-PAP and its unpersonalized version BDL-AP on FG-NET. The proposed BDL-PAP can be degenerated into BDL-AP by the following two steps: 1) in the optimization procedure of Eq. (9), we set $\mathbf{P}^g$ to be a full-zero matrix when initializing and updating it in each iteration; 2) in the synthesis of age progression, we directly set $\mathbf{p}^g$ to be a full-zero vector in Eq. (6).

We first qualitatively compare the experimental results of BDL-PAP and BDL-AP, which are illustrated in Fig. 8. We can see that the aging results generated by BDL-PAP are more similar to the ground truth than the counterparts of BDL-AP. In particular, compared with BDL-AP, BDL-PAP can preserve much more personalized facial characteristics, such as the special eyebrow shape in the top-left face and the mole in the fifth face of the first row, which are preserved by BDL-PAP but discarded by BDL-AP.

We also use the two-way ANalysis Of VAriance (ANOVA) [67] test to quantitatively compare BDL-PAP and BDL-AP. To avoid bias as much as possible, we invite 30 users covering a wide range of age and from all walks of life to browse 50 comparison groups, each of which includes the aging results of BDL-PAP and BDL-AP, the input face and the ground truth. Two aging faces in each comparison group are shown to users with a random order. For each comparison, if one user thinks the two results are comparable with each other, BDL-PAP and BDL-AP are assigned with score 1 respectively; if the user thinks one aging result is better than the other one, the method corresponding to the better aging result is assigned with score 2, and the other method is assigned with score 0. The comparison results are given in Table 2. We can see that BDL-PAP gets a score 1.5779, which is much higher than the score 0.4221 obtained by BDL-AP. This result validates the effectiveness of bringing in the personalization. Moreover, the p-values of ANOVA test

TABLE 2
The Left Part Illustrates the Average Rating Scores and Standard Deviation Values from the User Study on the Comparisons of BDL-PAP and BDL-AP

BDL-PAP v.s BDL-AP		Factor of approaches		Factor of users	
BDL-PAP	BDL-AP	F-statistic	p-value	F-statistic	p-value
1.5779 ± 0.0363	0.4221 ± 0.0363	128.6381	4.10×10^{-8}	1.88×10^{-15}	1.0000

The right part shows the ANOVA test results.

TABLE 3
Comparison of Running Time for Age Progression Synthesis

Running time (s)		Future age groups of aging face															
		6-10		11-15		16-20		21-30		31-40		41-50		51-60		61-80	
		CDL-PAP	**BDL-PAP**	CDL-PAP	**BDL-PAP**	CDL-PAP	**BDL-PAP**	CDL-PAP	**BDL-PAP**	CDL-PAP	**BDL-PAP**	CDL-PAP	**BDL-PAP**	CDL-PAP	**BDL-PAP**	CDL-PAP	**BDL-PAP**
Current age groups of input face	0-5	13.5	**2.2**	27.0	**4.4**	40.5	**6.6**	54.0	**8.8**	67.5	**11.0**	81.0	**13.2**	94.5	**15.4**	108.0	**17.6**
	6-10			13.5	**2.2**	27.0	**4.4**	40.5	**6.6**	54.0	**8.8**	67.5	**11.0**	81.0	**13.2**	94.5	**15.4**
	11-15					13.5	**2.2**	27.0	**4.4**	40.5	**6.6**	54.0	**8.8**	67.5	**11.0**	81.0	**13.2**
	16-20							13.5	**2.2**	27.0	**4.4**	40.5	**6.6**	54.0	**8.8**	67.5	**11.0**
	21-30									13.5	**2.2**	27.0	**4.4**	40.5	**6.6**	54.0	**8.8**
	31-40											13.5	**2.2**	27.0	**4.4**	40.5	**6.6**
	41-50													13.5	**2.2**	27.0	**4.4**
	51-60															13.5	**2.2**

show that this superiority is statistically significant and the difference of the users is insignificant.

4.6 Comparison of Running Time

Another advantage of BDL-PAP lies in the significant reduction of the running time of age progression synthesis compared with CDL-PAP [15]. To illustrate this advantage of BDL-PAP, we test the running time of BDL-PAP and CDL-PAP on a PC with CPU Intel Core i7 3.6 GHz and 6 GB memory. Table 3 shows the comparison of the running time. When we generate an aging face in the "6-10" age group from an input face in the "0-5" age group, BDL-PAP and CDL-PAP take 2.2 and 13.5 s, respectively. And when we generate an aging face in the "61-80" age group from the input face in the "0-5" age group, BDL-PAP and CDL-PAP take 17.6 and 108.0 s, respectively. We can see that BDL-PAP is much more efficient than CDL-PAP, and the running time of BDL-PAP and CDL-PAP increases accompanied with the increment of the age difference between the input face and the aging face. Moreover, we compare the running time of IAAP [5] and BDL-PAP. Since FT Demo requires a series of manual operations, we do not compare its running time. Given an input face in the "0-5" age group, generating its aging face sequence covering all the future age groups needs about 72.0 and 17.6 s by IAAP and BDL-PAP, respectively. Thus, BDL-PAP is also much faster than IAAP.

5 CONCLUSIONS AND FUTURE WORK

In this work, we proposed a personalized age progression method. Basically, we design multiple aging dictionaries for different age groups, in which the dictionary bases from two neighboring dictionaries respectively form a particular aging process pattern across different age groups, and a linear combination of these patterns expresses a particular aging process. Moreover, we define the aging layer and the personalized layer for an individual to capture the aging characteristics and the personalized characteristics, respectively. We train all aging dictionaries on the collected short-term aging database. Specifically, the younger- and older- age face pairs of the same persons are used to train two aging dictionaries in the neighboring age groups with the common sparse representation, excluding the specific personalized layer. Given a face, we render its aging face sequence from the current age to the future age step by step on the learned aging dictionaries. In future, we consider the face anti-aging synthesis, namely restoring the younger face for a given older face.

REFERENCES

[1] C. R. B. S. Ramanathan and Narayanan, "Age progression in human faces: A survey," *J. Vis. Languages Comput.*, vol. 15, pp. 3349–3361, 2009.

[2] Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 1955–1976, Nov. 2010.

[3] J. Suo, S.-C. Zhu, S. Shan, and X. Chen, "A compositional and dynamic model for face aging," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 385–401, Mar. 2010.

[4] U. Park, Y. Tong, and A. K. Jain, "Age-invariant face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 947–954, May 2010.

[5] I. Kemelmacher-Shlizerman, S. Suwajanakorn, and S. M. Seitz, "Illumination-aware age progression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3334–3341.

[6] B. Tiddeman, M. Burt, and D. Perrett, "Prototyping and transforming facial textures for perception research," *Comput. Graph. Appl.*, vol. 21, no. 5, pp. 42–50, 2001.

[7] M. R. Gandhi, "A method for automatic synthesis of aged human facial images," Ph.D. dissertation, Dept. Elect. Comput. Eng., McGill Univ., Montreal, QC, Canada, 2004.

[8] Y. Tazoe, H. Gohara, A. Maejima, and S. Morishima, "Facial aging simulator considering geometry and patch-tiled texture," in *Proc. ACM SIGGRAPH*, 2012, Art. no. 90.

[9] Y. Liang, Y. Xu, L. Liu, S. Liao, and B. Zou, "A multi-layer model for face aging simulation," in *Trans. Edutainment VI*. Berlin, Germany: Springer, 2011, pp. 182–192.

[10] B.-C. Chen, C.-S. Chen, and W. H. Hsu, "Cross-age reference coding for age-invariant face recognition and retrieval," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 768–783.

[11] Face and gesture recognition, "Network: FG-NET aging database," 2000. [Online]. Available: http://sting.cycollege.ac.cy/alanitis/fgnetaging/

[12] K. Ricanek and T. Tesafaye, "MORPH: A longitudinal image database of normal adult age-progression," in *Proc. 7th Int. Conf. Automatic Face Gesture Recognit.*, 2006, pp. 341–345.

[13] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[14] X. Shu, J. Tang, G.-J. Qi, Z. Li, Y.-G. Jiang, and S. Yan, "Image classification with tailored fine-grained dictionaries," *IEEE Trans. Circuits Syst. Video Technol.*, 2016, doi: 10.1109/TCSVT.2016.2607345.

[15] X. Shu, J. Tang, H. Lai, L. Liu, and S. Yan, "Personalized age progression with aging dictionary," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 3970–3978.

[16] N. Ramanathan, R. Chellappa, and S. Biswas, "Computational methods for modeling facial aging: A survey," *J. Vis. Languages Comput.*, vol. 20, no. 3, pp. 131–144, 2009.

[17] X. Shu, G. Xie, Z. Li, and J. Tang, "Age progression: Current technologies and applications," *Neurocomputing*, vol. 208, pp. 249–261, 2016.

[18] D. M. Burt and D. I. Perrett, "Perception of age in adult caucasian male faces: Computer graphic manipulation of shape and colour information," *Proc. R. Soc. London*, vol. 259, pp. 137–143, 1995.

[19] X. Shu, J. Tang, H. Lai, Z. Niu, and S. Yan, "Kinship-guided age progression," *Pattern Recognit.*, vol. 59, pp. 156–167, 2016.

[20] C. M. Scandrett, C. J. Solomon, and S. J. Gibson, "A person-specific, rigorous aging model of the human face," *Pattern Recognit. Lett.*, vol. 27, no. 15, pp. 1776–1787, 2006.

[21] D. Hubball, M. Chen, and P. W. Grant, "Image-based aging using evolutionary computing," *Comput. Graph. Forum*, vol. 27, no. 2, pp. 607–616, 2008.

[22] D. W. Hunter, "Synthesis of facial ageing transforms using three-dimensional morphable models," Ph.D. dissertation, School Comput. Sci., Univ. St Andrews, St Andrews, U.K., 2009.

[23] A. M. Albert, K. Ricanek Jr, and E. Patterson, "A review of the literature on the aging adult skull and face: Implications for forensic science research and applications," *Forensic Sci. Int.*, vol. 172, no. 1, pp. 1–9, 2007.

[24] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 442–455, Apr. 2002.

[25] N. Ramanathan and R. Chellappa, "Face verification across age progression," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3349–3361, Mar. 2006.

[26] P. Paysan, "Statistical modeling of facial aging based on 3D scans," Ph.D. dissertation, Departement Mathematik und Informatik, Univ. Basel, Basel, Switzerland, 2010.

[27] A. C. Berg, F. Perales, and L. Gonzalez, "A facial aging simulation method using flaccidity deformation criteria," in *Proc. Int. Conf. Inf. Vis.*, 2006, pp. 791–796.

[28] G. Schroeder, L. Magalhes, and R. Rodrigues, "Facial aging using image warping," in *Proc. West. NY Image Process. Workshop*, 2007, pp. 1–4.

[29] J. Suo, X. Chen, S. Shan, W. Gao, and Q. Dai, "A concatenational graph evolution aging model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2083–2096, Nov. 2012.

[30] W. Wang, et al., "Recurrent face aging," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2378–2386.

[31] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1891–1898.

[32] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 815–823.

[33] C. Ding, J. Choi, D. Tao, and L. S. Davis, "Multi-directional multi-level dual-cross patterns for robust face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 3, pp. 518–531, Mar. 2016.

[34] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "A study of face recognition as people age," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2007, pp. 1–8.

[35] G. Guo, G. Mu, and K. Ricanek, "Cross-age face recognition on a very large database: The performance versus age intervals and improvement using soft biometric traits," in *Proc. 20th Int. Conf. Pattern Recognit.*, 2010, pp. 3392–3395.

[36] J. Wang, Y. Shang, G. Su, and X. Lin, "Age simulation for face recognition," in *Proc. 18th Int. Conf. Pattern Recognit.*, 2006, pp. 913–916.

[37] U. Park, Y. Tong, and A. K. Jain, "Face recognition with temporal invariance: A 3D aging model," in *Proc. 8th IEEE Int. Conf. Automatic Face Gesture*, 2008, pp. 1–7.

[38] D. Gong, Z. Li, D. Tao, J. Liu, and X. Li, "A maximum entropy feature descriptor for age invariant face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 5289–5297.

[39] T. Wu, P. Turaga, and R. Chellappa, "Age estimation and face verification across aging using landmarks," *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 6, pp. 1780–1788, Dec. 2012.

[40] A. Lanitis, "A survey of the effects of aging on biometric identity verification," *Int. J. Biometrics*, vol. 2, no. 1, pp. 34–52, 2009.

[41] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "Face verification across age progression using discriminative methods," *IEEE Trans. Inf. Forensics Secur.*, vol. 5, no. 1, pp. 82–91, Mar. 2010.

[42] D. Gong, Z. Li, D. Lin, J. Liu, and X. Tang, "Hidden factor analysis for age invariant face recognition," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 2872–2879.

[43] E. Patterson, K. Ricanek, M. Albert, and E. Boone, "Automatic representation of adult aging in facial images," in *Proc. Int. Conf. Vis. Imaging Image Process.*, 2006, pp. 171–176.

[44] J. Suo, X. Chen, S. Shan, and W. Gao, "Learning long term face aging patterns from partially dense aging databases," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 622–629.

[45] L. Liu, C. Xiong, H. Zhang, Z. Niu, M. Wang, and S. Yan, "Deep aging face verification with large gaps," *IEEE Trans. Multimedia*, vol. 18, no. 1, pp. 64–75, Jan. 2016.

[46] X. Geng, Z.-H. Zhou, and K. Smith-Miles , "Automatic age estimation based on facial aging patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2234–2240, Dec. 2007.

[47] G. Guo, G. Mu, Y. Fu, and T. S. Huang, "Human age estimation using bio-inspired features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 112–119.

[48] K.-Y. Chang, C.-S. Chen, and Y.-P. Hung, "Ordinal hyperplanes ranker with cost sensitivities for age estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 585–592.

[49] J. Suo, F. Min, S. Zhu, S. Shan, and X. Chen, "A multi-resolution dynamic model for face aging simulation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2007, pp. 1–8.

[50] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang, "Coupled dictionary training for image super-resolution," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3467–3478, Aug. 2012.

[51] J. Yang, Z. Wang, Z. Lin, X. Shu, and T. S. Huang, "Bilevel sparse coding for coupled feature spaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 2360–2367.

[52] S. Li, J. Xing, Z. Niu, S. Shan, and S. Yan, "Shape driven kernel adaptation in convolutional neural network for robust facial trait recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 222–230.

[53] G. Qi, C. C. Aggarwal, Q. Tian, H. Ji, and T. S. Huang, "Exploring context and content links in social media: A latent space method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 850–862, May 2012.

[54] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2001, pp. I-511–I-518.

[55] I. Kemelmacher-Shlizerman and S. M. Seitz, "Collection flow," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 1792–1799.

[56] D. Yadav, M. Vatsa, R. Singh, and M. Tistarelli, "Bacteria foraging fusion for face recognition across age progression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 173–179.

[57] G. Qi, W. Liu, C. C. Aggarwal, and T. S. Huang, "Joint intermodal and intramodal label transfers for extremely rare or unseen classes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 7, pp. 1360–1373, Jul. 2007.

[58] G. Qi, X. Hua, Y. Rui, J. Tang, and H. Zhang, "Two-dimensional multilabel active learning with an efficient online adaptation model for image classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 10, pp. 1880–1897, Oct. 2009.

[59] K. Scherbaum, M. Sunkel, H.-P. Seidel, and V. Blanz, "Prediction of individual non-linear aging trajectories of faces," *Comput. Graph. Forum*, vol. 26, no. 3, pp. 285–294, 2007.

[60] E. Patterson, A. Sethuram, M. Albert, and K. Ricanek, "Comparison of synthetic face aging to age progression by forensic sketch artist," in *Proc. Int. Conf. Vis. Imag. Image Process.*, 2007, pp. 247–252.

[61] Y. Liang, C. Li, H. Yue, and Y. Luo, "Age simulation in young face images," in *Proc. 1st Int. Conf. Bioinf. Biomed. Eng.*, 2007, pp. 494–497.

[62] C.-T. Shen, W.-H. Lu, S.-W. Shih, and H.-Y. M. Liao, "Exemplar-based age progression prediction in children faces," in *Proc. IEEE Int. Symp. Multimedia*, 2011, pp. 123–128.

[63] A. Sethuram, K. Ricanek, and E. Patterson, "A hierarchical approach to facial aging," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 100–107.

[64] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, 2014, pp. 1988–1996.

[65] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1701–1708.

[66] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Univ. Massachusetts Amherst, Amherst, MA, USA, Tech. Rep. 07–49, 2007.

[67] E. W. Minium, B. M. King, and G. Bear, "Statistical reasoning in psychology and education," New York: Wiley, 2003.

Xiangbo Shu received the PhD degree from Nanjing University of Science and Technology, in July 2016. He is an assistant professor in the School of Computer Science and Engineering, Nanjing University of Science and Technology, China. From 2014 to 2015, he worked as a visiting scholar in the Department of Electrical and Computer Engineering, National University of Singapore. His research interests include computer vision and machine learning. He has received the Best Student Paper Award in MMM 2016 and the Best Paper Runner-up in ACM MM 2015.

Jinhui Tang received the BE and PhD degrees both from the University of Science and Technology of China, in July 2003 and July 2008, respectively. He is a professor in the School of Computer Science and Engineering, Nanjing University of Science and Technology, China. From 2008 to 2010, he worked as a research fellow in the School of Computing, National University of Singapore. His current research interests include large-scale multimedia search. He has authored more than 100 journal and conference papers in these areas. He received the ACM China Rising Star Award and a co-recipient of the Best Student Paper Award in MMM 2016, and Best Paper Award in ACM MM 2007, PCM 2011, and ICIMCS 2011. He is a senior member of the IEEE.

Zechao Li received the the BE degree from the University of Science and Technology of China (USTC), China, in 2008 and the PhD degree from the Institute of Automation, Chinese Academy of Sciences, China, in 2013. He is an associate professor in the School of Computer Science and Engineering, Nanjing University of Science and Technology. His research interests include large-scale multimedia understanding, social media mining, subspace learning, etc. He received the 2015 Excellent Doctoral Dissertation of Chinese Academy of Sciences, the 2015 Excellent Doctoral Dissertation of China Computer Federation, the Top 10 percent Paper Award of IEEE MMSP 2015 and the 2013 President Scholarship of Chinese Academy of Science.

Hanjiang Lai received the BS and PhD degrees from Sun Yat-sen University, in 2009 and 2014, respectively. He was working as a research fellow with National University of Singapore during 2014-2015. He is now working with Sun Yat-sen University. His research interests include machine learning algorithms, deep learning, and computer vision.

Liyan Zhang received the PhD degree in computer science from the University of California, Irvine, in 2014. She is currently an associate professor in the School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics. Her research interests include multimedia analysis and computer vision. She has received the Best Paper Award in ICMR 2013 and the Best Student Paper Award in MMM 2016.
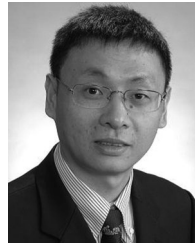
Shuicheng Yan is the dean's chair associate professor with National University of Singapore, and also the chief scientist of Qihoo/360 Company. His research areas include machine learning, computer vision and multimedia, and he has authored/co-authored hundreds of technical paper, with Google Scholar citation $> 15,000$ times and H-index 52. He has been serving as an associate editor of the *IEEE Transactions on Circuits and Systems for Video Technology* and the *ACM Transactions on Intelligent Systems and Technology*. He received the Best Paper Awards from ACM MM'13 (Best Paper and Best Student Paper), MMM'16 (Best Student Paper), ACM MM'12 (Best Demo), PCM'11, ACM MM'10, ICME'10, and ICIMCS'09, the winner prizes of the classification task in PASCAL VOC 2010-2012, 2011 Singapore Young Scientist Award, and 2012 NUS Young Researcher Award. He is a fellow of the IEEE and the IAPR.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.