

# A Concatenational Graph Evolution Aging Model

Jinli Suo, Xilin Chen, *Senior Member, IEEE*, Shiguang Shan, *Member, IEEE*,  
Wen Gao, *Fellow, IEEE*, and Qionghai Dai, *Senior Member, IEEE*

**Abstract**—Modeling the long-term face aging process is of great importance for face recognition and animation, but there is a lack of sufficient long-term face aging sequences for model learning. To address this problem, we propose a CONcatenational GRaph Evolution (CONGRE) aging model, which adopts decomposition strategy in both spatial and temporal aspects to learn long-term aging patterns from partially dense aging databases. In spatial aspect, we build a graphical face representation, in which a human face is decomposed into mutually interrelated subregions under anatomical guidance. In temporal aspect, the long-term evolution of the above graphical representation is then modeled by connecting sequential short-term patterns following the Markov property of aging process under smoothness constraints between neighboring short-term patterns and consistency constraints among subregions. The proposed model also considers the diversity of face aging by proposing probabilistic concatenation strategy between short-term patterns and applying scholastic sampling in aging prediction. In experiments, the aging prediction results generated by the learned aging models are evaluated both subjectively and objectively to validate the proposed model.

**Index Terms**—Face aging, aging model evaluation, long-term aging, short-term aging, ANOVA

## 1 INTRODUCTION

FACE aging is attracting increasing interest from researchers due to its grand challenge in real-world applications (e.g., looking for lost children or wanted fugitives, recognizing age-separated faces [7], [29], [48], [36], [57], facial template renewal [10]). In recent years, a lot of effort has been devoted to face aging modeling [39], [40], [15], building face aging databases [1], [42], and aging model evaluation [23], [22]. Despite these encouraging advances, modeling face aging is still a quite challenging task for some specific reasons, e.g., complex aging mechanisms, lack of training data, and the intrinsic uncertainty of aging.

### 1.1 Related Works

Researchers have made great efforts on face aging modeling and proposed a variety of approaches; refer to [15], [39], [40] for surveys on state of the art in face aging. Generally, these approaches fall into two groups: the *physical model-based approach* and the *example-based approach*. The former builds a face aging model to simulate physical aging mechanisms, while the latter attempts to fit available real aging

sequences and can be further classified into three types: prototyping method, function-based method, and evaluation targeted method.

#### 1.1.1 Physical Model-Based Approaches

Physical model-based approaches simulate face aging by modeling the biological structure and aging mechanisms of the cranium [55], muscles [5], [38], or facial skin [3], [8], [26], [58], etc. Due to the subtleness of both facial structure and face aging mechanisms, physical model-based methods are often complex and computationally expensive. These disadvantages make it *difficult to obtain realistic results from physical modeling*.

#### 1.1.2 Prototyping Approaches

Prototyping approaches [9], [17], [27], [43], [53], [54] first divide age range into discrete age groups and define the average face of each group as its prototype, and then the difference between prototypes is defined as an axis of aging transformation. This approach can be easily extended to PCA space [57] and 3D face data [31]. Apparently, the high-frequency facial details are lost and no individuality is considered in the prototyping approach. Although some other researchers specially work on rendering high-resolution aging results [16] or introducing individuality of face aging [44], [19], [18], lack of individuality is still the biggest disadvantage of prototyping approaches. However, *the learned average aging patterns of a specific group can be used as priors of personalized aging models*.

#### 1.1.3 Function-Based Approaches

Function-based approaches describe the relationship between facial parameters and corresponding age label with an explicit or implicit function. In the process of child growth, the shape change due to cranium growth is prominent; therefore, researchers proposed various methods to simulate

- J. Suo and Q. Dai are with the Broadband Network and Digital Media Laboratory, Department of Automation, Tsinghua University, Central Main Building, Beijing 100084, China. E-mail: {jlsuo, qhdai}@tsinghua.edu.cn.
- X. Chen and S. Shan are with the Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, No. 6, Kexueyuan South Road, Haidian, Beijing 100190, China. E-mail: {xlchen, sgshan}@ict.ac.cn.
- W. Gao is with the Key Laboratory of Machine Perception, School of Electrical Engineering and Computer Science, Peking University, Room 2641, Beijing 100871, China. E-mail: wgao@pku.edu.cn.

Manuscript received 15 Jan. 2011; revised 5 Aug. 2011; accepted 28 Dec. 2011; published online 9 Jan. 2012.

Recommended for acceptance by J. Luo.

For information on obtaining reprints of this article, please send e-mail to: [tpami@computer.org](mailto:tpami@computer.org), and reference IEEECS Log Number TPAMI-2011-01-0035.

Digital Object Identifier no. 10.1109/TPAMI.2012.22.

this change [35], [37]. In contrast, both remarkable texture change and subtle shape changes [2], [4], [59] occur during adult aging. So, most researchers study shape and texture changes simultaneously in adult aging modeling. Large varieties of functions are proposed, including quadratic function [24], [32], [27], support vector regression [22], [56], [45], kernel smoothing method [20], or an implicit function [6], [30], [28], [46]. Differently, Jiang and Wang [21] directly built a mapping function between a young face and its appearance in later ages. For the feasibility of inverse function computation, the functions are usually of simple forms and insufficient for modeling large appearance changes over long periods. *In spite of these drawbacks, the method can be applied to short-term face aging of specific facial subregions.*

#### 1.1.4 Evaluation Targeted Approach

The evaluation targeted approach was proposed by Lanitis [22], who proposed two criteria for aging model evaluation, according to which he defined an objective function and proposed three optimizing methods. This approach is intuitively reasonable and gives plausible results on child growth, but its effectiveness on larger databases requires better understanding of face space and the aging process. In addition, both defining objective function and fusing two criteria are nontrivial problems. *However, this approach sheds light on studies of aging model evaluation.*

In all, physical model-based approaches are constrained by the complexities and example-based approaches suffer from a lack of sufficient training data for modeling the complex aging process. In addition, a common drawback of existing algorithms is that the intrinsic uncertainty of face aging is not modeled.

## 1.2 Motivation and Basic Ideas

The aging mechanism is complex both spatially and temporally, but currently available face aging databases contain mostly short-term aging sequences (we refer to them as *partially dense* data sets) since collecting long-term dense face aging sequences is time consuming. To handle this problem, we propose a concatenational graph evolution face aging model, which makes good use of spatial and temporal locality of face aging to learn long-term aging patterns from partially dense databases.

*From the spatial viewpoint*, the face is a highly complex structure composed of hard and soft tissues, both of which undergo large changes during aging process. Since the tissues have different physical properties and bear different aging mechanisms, a global face aging model is certainly of high dimension and demands a large training data set. However, lack of training data makes learning a global aging model especially challenging.

Fortunately, anatomical studies give the following two inspirations: First, face aging of a specific subregion is mostly controlled by the behaviors of a subgroup of muscles, i.e., face aging is spatially local; second, the human face is a global structure in spite of the spatial locality, so the aging of one face region is related to those of the other ones to some extent, i.e., the aging patterns of different regions are not absolutely independent. Considering the above two inspirations, this work proposes to decompose a

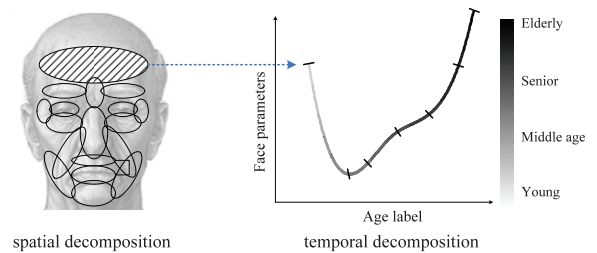


Fig. 1. Illustration of spatial and temporal decomposition in the proposed CONGRE aging model.

face into subregions and learn the aging patterns of each subregion under soft constraints from the other ones.

*From the temporal viewpoint*, the long-term face aging process (i.e., across 3-4 decades) is highly nonlinear because a face undergoes different changes in different stages and the aging rate is also time variant. Therefore, learning long-term aging patterns apparently demands long-term aging sequences, and thus faces considerable difficulties due to the lack of training data. Consider that face aging (here, we only refer to natural aging) can be looked at as a smooth Markov process composed of short-term patterns (i.e., age span smaller than 10 years), i.e., face aging is temporally local, and one can infer the aging pattern sequentially. Therefore, we formulate face aging as a Markov process in the granularity of short age spans.

With the above analysis, the proposed CONGRE aging model first builds a graphical face representation, then learns long-term aging patterns from partially dense data sets (see Fig. 1). Specifically, the model learning is composed of the following two steps.

**Step 1. Learning short-term face aging patterns.** With a large number of short-term face aging sequences from publicly available face aging databases such as FG-NET [1] and MORPH [42], our model extracts short-term aging patterns from the real aging sequences. A function-based approach is exploited for this purpose since the short-term appearance changes of facial subregions are relatively small.

**Step 2. Concatenating learned short-term aging patterns into long-term patterns.** By formulating long-term aging as a Markov process, we concatenate partially overlapping short-term aging patterns sequentially into long term ones based on predefined criteria. The uncertainty of aging can be modeled by sampling different subsequent patterns probabilistically, with the probability computed from the above criteria.

The proposed model bears some similarity to one of our previous works [52], in which a compositional dynamic face aging model is learned from large ID photo databases. However, the proposed model differs remarkably from and improves upon this previous work in two main aspects: First, the dynamics of face aging are learned from real aging sequences instead of similar faces across age groups, so the learned aging patterns are much closer to real aging paths; second, here we model the continuous aging process, which is more accurate than the discrete model in [52].

## 2 COMPOSITIONAL AGING-ORIENTED FACE MODEL

In this section, we describe the compositional face representation. Briefly speaking, first, a human face is decomposed

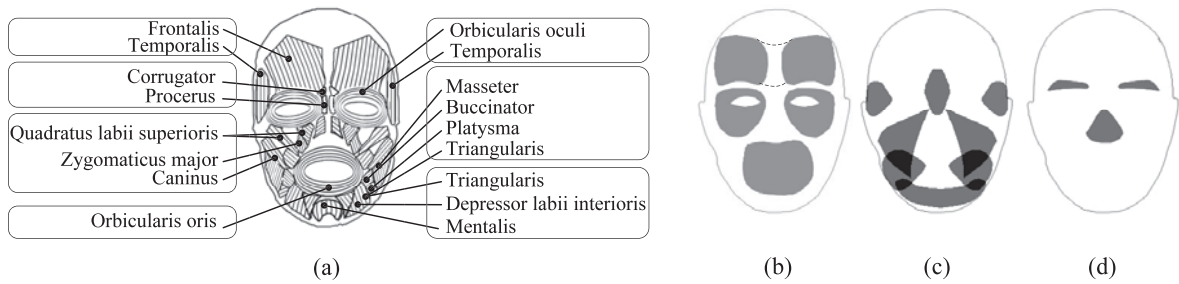


Fig. 2. The compositional face model inspired by facial muscle structure. (a) The facial muscles and their grouping [25]. (b)-(d) The divided subregions based on the grouping rules in (a), with (b), (c), and (d) displaying muscular and bony regions respectively.

into multiple subregions from anatomical results; for each subregion, we then build a local aging-oriented model, which in our implementation is an extension of the Active Appearance Model (AAM) [11]; finally, to introduce integrity constraints among the subregions, we build a graphical face model with subregions being the vertex and the statistical correlation among them as edges.

## 2.1 Anatomy-Based Region Decomposition

Due to its simplicity and effectiveness in face representation, AAM is widely used in face analysis, including aging-related studies, such as [24], [57], [34], [47], etc. However, the holistic AAM representation has some limitations for the purpose of face aging modeling: 1) A holistic AAM model capable of describing both large and subtle appearance changes across decades is often of high dimension and statistical learning approaches suffer from the curse of dimensionality problem, which is especially severe in face aging modeling; 2) aging mechanisms of different regions are specifically related to their biological structures and display large varieties in both visual cues and aging rate, which implies a nonlinear global model be desirable for modeling these complex process. In addressing the above two problems, we adopt a region-based AAM model which targets reducing data dimension and the nonlinearity of aging problem.

Since the behaviors of facial muscles have large effects on face appearances during the aging process (e.g., facial contour droops gradually because of the loss of muscle tone, wrinkles emerge due to the contraction of expression muscles), we take into account anatomical evidence when building our region-based face representation. Ekman and Rosenberg [12] conducted a deep study on the structure and function of facial muscles, and applied it to the Face Action Coding System (FACS) successfully. Adopting a similar muscle model, Ramanathan and Chellappa [38] simulate the face shape changes caused by aging of facial muscles. On the basis of these two previous analysis on facial muscles, we select the muscles most related to face aging and build a compositional face model (see Fig. 2a). In this model, the muscles cluster into groups based on their physical positions, orientations, and functionalities. Guided by the muscle clustering, a face image  $I$  is divided into 15 subregions (intersections exist on the boundaries of some neighboring subregions), as shown in Fig. 2. Formally,

$$I = \bigoplus_{r \in R} I^r. \quad (1)$$

Here,  $R$  is the set of all the facial subregions and  $\bigoplus$  denotes the composition of subregions, with blending operations introduced along the subregion boundaries to avoid the artifacts.

## 2.2 Aging-Oriented AAM Model

After face decomposition, we build an AAM model for each subregion. However, considering that the aging sequences are usually affected by external factors (e.g., pose, illumination, expression, and resolution), the original AAM model is not appropriate for aging modeling. Here, two schemes are exploited to build an aging-oriented face model.

### 2.2.1 Database Purification

Evidently, face images largely influenced by nonaging factors are improper for learning aging models. Furthermore, most preprocessing is not effective enough for excluding their influences because many non-age-related appearance variations are far from being additive. Therefore, we proposed the following method to remove badly contaminated short-term aging sequences (i.e., outliers) before model learning.

Our method is inspired by this basic idea: Although aging patterns learned by prototyping method might lose some crucial facial details and neglect individualities, it is possible to learn an average aging trajectory of a specific group, which can be used as prior constraints for selecting high-quality aging sequences from aging databases. With this basic idea, we propose an improved prototyping method to compute a continuous average aging trajectory and then remove the sequences deviating largely from this trajectory as outliers. Algorithm 1 displays the procedure of removing outliers, i.e., *data purification*.

**Algorithm 1.** Detecting and Removing outliers from database

**Input:** Face aging data set  $S$

**Output:** Inlier set  $S'$  and outlier set  $S''$

**Procedure:**

1.  $S' := S, S'' := \phi$ ;
2. **Repeat** following steps **Until**  $S'' = \phi$ .
  - 2.1 Build an AAM model from training data  $S'$ ;
  - 2.2 Compute the average aging trajectory and its floating range (detailed below this algorithm);
  - 2.3  $S'' :=$  the images causing its sequence exceeding the floating range of the average trend;
  - 2.4  $S' := S' - S''$ ;

For each AAM parameter of a specific face region, such as the projection coefficients on the first principal components

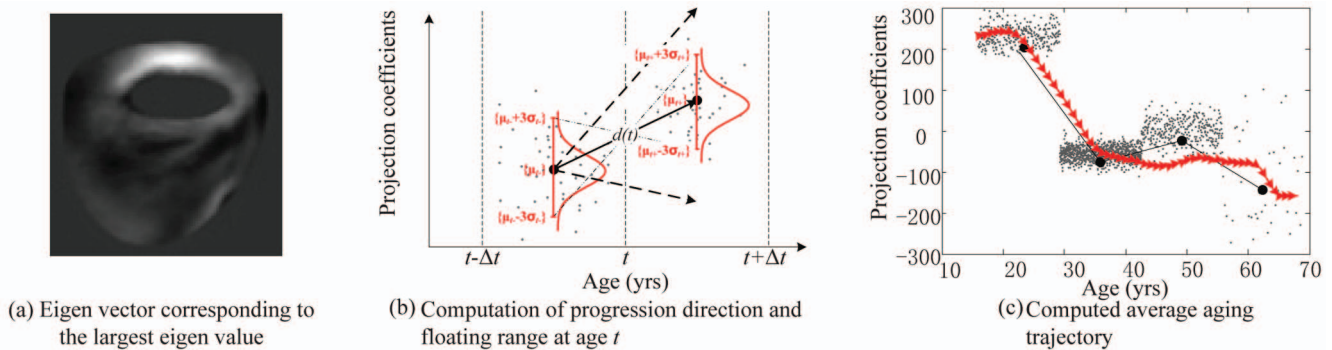


Fig. 3. Illustration of computing the general aging trajectory of the eye-bag region.

(pc) of the eye-bag area (as visualized in Fig. 3a), its progression direction (solid arrow in Fig. 3b) at age  $t$  and corresponding floating range (angles between two dashed arrows in Fig. 3b) are computed as follows:

1. Sampling a number of parameter sets in age span  $[t - \Delta t, t]$  and  $[t, t + \Delta t]$ , respectively.
2. Computing a series of prototype candidates from the parameter sets in two age spans (denoted as gray dots in Fig. 3b) which follow near-Gaussian distribution  $N(\mu_{t-}, \sigma_{t-})$  and  $N(\mu_{t+}, \sigma_{t+})$ , as shown in Fig. 3b.
3. Similar to the prototyping method, computing the mean aging direction as  $\overrightarrow{\mu_{t-}, \mu_{t+}}$  (solid arrows in Fig. 3b).
4. Adopting the three-sigma rule to compute floating range, whose upper and lower boundary are  $\overrightarrow{\mu_{t-} - 3\sigma_{t-}, \mu_{t+} + 3\sigma_{t+}}$  and  $\overrightarrow{\mu_{t-} + 3\sigma_{t-}, \mu_{t+} - 3\sigma_{t+}}$ , respectively (dashed arrows in Fig. 3b).

We display the average trajectory of the first principal components of the eye-bag area within the whole age range in Fig. 3b with thick arrows, from which one can see that the trajectory computed by improved prototyping method approximately consists with that by traditional prototyping method, but is largely refined.

### 2.2.2 Model Purification

The above database purification can only eliminate part of the non-age-related variations. In this section, we further modify the original AAM [11] model by selecting the age-related dimensions. The extended model includes a global active shape model and a shape-free texture model for each subregion. Since the principal components of the original AAM models describe both age-related and non-age-related variations, for the aging modeling purpose we analyze the components statistically to select those components significantly related to face aging.

Specifically, we first group the face images in the training data set into several age levels, and denote their projection coefficients on a specific principal component as  $w_{s,t(s),pc}$  with  $s$  being the index of the sample and  $t(s)$  being the age level of the  $s$ th sample. Since the statistical results tell us that, for each age group  $T$ , the AAM parameter  $\{w_{s,t(s),pc}, t(s) = T\}$  can be reasonably assumed to follow Gaussian distribution, we conduct the widely used ANalysis Of Variance (ANOVA)

[14] on these coefficients. Here,  $F$ -ratio is used to measure the correlation between this principal component and aging:

$$F_{pc} = \frac{\frac{1}{N_T - 1} \sum_{t=1}^{N_T} n_t (\overline{w_{t,pc}} - \overline{\overline{w_{pc}}})^2}{\frac{1}{N_S - N_T} \sum_{s=1}^{N_S} (w_{s,pc} - \overline{w_{t(s),pc}})^2}. \quad (2)$$

Here,  $N_T$  and  $N_S$  are, respectively, the number of age levels and training samples,  $\overline{w_{t,pc}}$  is the mean coefficients of all  $n_t$  samples at age level  $t$ , and  $\overline{\overline{w}}$  is the mean of all  $\overline{w_{t,pc}}$ . Since  $w_{s,t(s),pc}$  is near-Gaussian,  $F_{pc}$  follows F-distribution with degrees of freedom  $N_T - 1$  and  $N_S - N_T$ .

With the above analysis, the components with higher  $F$ -ratio are mostly aging related while the other ones mainly reflect other non-age-related variations. Setting the significance threshold at  $p < 0.05$ , we preserve the components significantly related with face aging and rearrange them in descending order according to their correlations with face aging. We name the model after rearrangement the *aging-oriented AAM model*, which is quite different from the original AAM model, as shown in Fig. 4a. In this model, a specific face shape  $I^{\text{shp}}$  is described as

$$I^{\text{shp}} = \overline{I^{\text{shp}}} + \sum_{i=1}^{N^{\text{shp}}} w_i^{\text{shp}} \cdot \mathbf{v}_i^{\text{shp}} + I_{\text{res}}^{\text{shp}}. \quad (3)$$

Similarly, a specific shape-free sub image of the  $r$ th region  $I^r$  is described as

$$I^r = \overline{I^r} + \sum_{i=1}^{N^r} w_i^r \cdot \mathbf{v}_i^r + I_{\text{res}}^r, \quad (4)$$

where  $\overline{I}$ ,  $\mathbf{v}$ , and  $w$  are mean vector, principal components, and corresponding projection coefficients, respectively.  $N$  is the number of preserved components and  $I_{\text{res}}$  is the residue containing both non-age-related variations and noise.

To intuitively validate the effectiveness of the aging-oriented AAM model, we compare its reconstruction results with those of original AAM, as shown in Fig. 4. The examples in Fig. 4c show that, some nonage-related variations (e.g., glasses and hair occlusions, pose) are excluded by the proposed model.

### 2.3 Graphical Face Model

As described above, our compositional face representation builds an aging-oriented AAM separately for each subregion. However, the human face is a holistic structure and

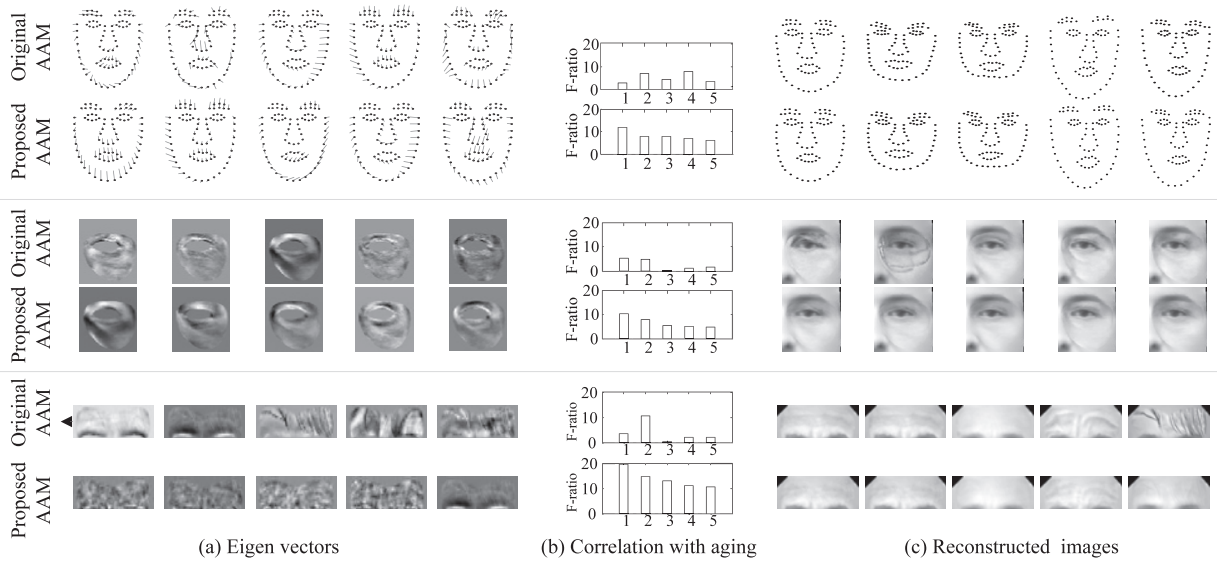


Fig. 4. Comparison between the original AAM and aging-oriented AAM. The top two rows visualize the top five principal components of face shape, their correlation with face aging, and reconstruction results in the original AAM and aging-oriented AAM. In the figure, the line segments, along with the landmarks, illustrate the shape variations modeled by the specific component. Similarly, the middle two rows and the bottom two rows, respectively, compare the shape-free texture of the around-eye region and forehead region in two models. The results in this figure are learned from 500 African-American males from the MORPH database [42].

there exist subtle internal relationships among the facial regions due to the common underlying biologic structure, so their aging processes should also be correlated mutually to some extent. To address this problem, in this paper we build a graphical face model to account for the relationships among the subregions obtained according to the aforementioned decomposition rule in Section 2.1, as shown in Fig. 2.

A face is formulated as a graph  $G$  with vertex set  $V$  corresponding to all the subregions and edge set  $E$  connecting all the vertex pairs with certain strengths. Formally, we have  $G = \{V, E\}$ ,  $V = \{v^r : r \in R\}$ , and  $E = \{\rho^{r,s} : v^r, v^s \in V\}$ . Here,  $v$  denotes a vertex which corresponds to a subregion;  $\rho^{r,s}$  denotes a graph edge with connecting strength representing the correlation between aging processes of regions  $r$  and  $s$ .

As for the graphical model learning, we do not study the complex interaction in terms of biologic mechanism. Instead, we just learn the correlation among the parameter evolutions in different regions in a statistical manner. Formally, given  $N$  training samples, with  $\mathbf{w}_i^r = \{w_{n,i}^r, n = 1, 2, \dots, N\}$  and  $\mathbf{w}_j^s = \{w_{n,j}^s, n = 1, 2, \dots, N\}$  denoting the  $i$ th parameters in region  $r$  and  $j$ th parameters in region  $s$ , the correlation between two regions is computed as the weighted summation of Pearson's correlation coefficients between  $\mathbf{w}_i^r$  and  $\mathbf{w}_j^s$ :

$$\rho^{r,s} = \text{sigmoid} \left( \sum_{i=1}^{N^r} \sum_{j=1}^{N^s} F_i^r \cdot F_j^s \cdot \mathcal{P}(\mathbf{w}_i^r, \mathbf{w}_j^s) \right). \quad (5)$$

Here,  $N^r$  and  $N^s$  denote the number of principal components in AAM of subregions  $r$  and  $s$ , respectively;  $F_i^r$  and  $F_j^s$  denote the  $F$ -ratio (correlation with aging process) of the  $i$ th principal component of region  $r$  and the  $j$ th principal component of region  $s$ , respectively. In addition, we normalize it using Sigmoid function.

Fig. 5 visualizes the above-defined correlations among facial regions, with line thickness representing the intensity of the correlation. From the figure, one has the following observations: There exists a strong correlation between two half faces, e.g., eye pair, brow pair; the correlation between two nearby regions is a little weaker; the asymmetric regions located far away are rarely correlated. This consists of the properties of biologic face structure: Human faces are approximately symmetric and the muscles behave in a local manner to some extent.

### 3 SHORT-TERM AGING MODEL

As mentioned previously, short-term aging modeling aims to generate functions that can approximate the evolution of aging parameters within a shorter age range for each face subregion. Formally, given a real image sequence including  $K$  face images  $\{I_k, k = 1, 2, \dots, K\}$  and their age labels  $\{t_k, k = 1, 2, \dots, K\}$  in ascending order, we denote  $I_k^r$  as the  $r$ th region of the  $k$ th face image and its aging-related AAM

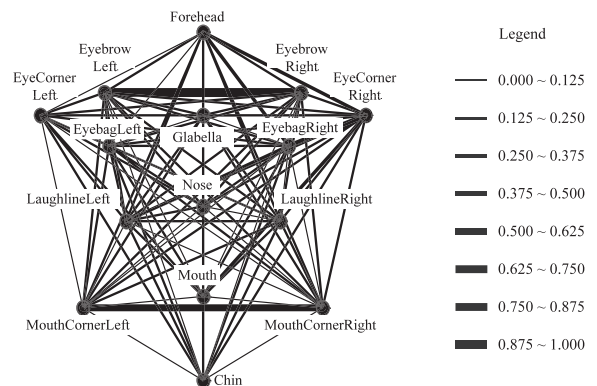


Fig. 5. The visualization of relationships between aging of subregions, with the line thickness denoting the correlation.

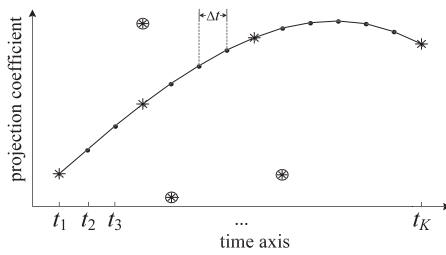


Fig. 6. Modeling of short-term face aging. The star marks denote the parameters of a given image sequence; among them, the circled ones are outliers excluded by RANSAC. The solid curve represents the learned aging pattern, with filled circles being interpolated parameters at time steps of  $\Delta t$ .

parameters as  $\mathbf{w}_k^r = \{w_{k,i}^r, i = 1, 2, \dots, N^r\}$ . So, the target of the short-term aging modeling is to learn a curve set  $\mathbf{C}_\tau^r = \{C_{i,\tau}^r, i = 1, 2, \dots, N^r\}$  over  $\tau \subseteq [t_1, t_K]$  to approximate the evolution of face parameters  $\mathbf{w}_k^r$ .

To achieve this function approximation operating directly on the data seems feasible, which is, however, not desirable as the data unavoidably contains noises/outliers. Therefore, we present an algorithm for learning short-term aging patterns. Basically, our algorithm exploits RANdom SAMple Consensus (RANSAC) [13] for outlier removal and cubic B-spline for the function fitting the evolution of the parameters. The whole procedure is summarized in Algorithm 2, and explained in the rest of this section.

**Algorithm 2.** Learning a short term aging pattern for  $r$ th region from image sequence  $\{I_k^r, k = 1, 2, \dots, K\}$

**Input:**  $\{\mathbf{w}_k^r, t_k\}$ : face parameters and age labels

$N(\boldsymbol{\mu}^r, \boldsymbol{\sigma}^r)$ : distribution of  $\mathbf{w}^r$ 's gradient

$n$ : sample size for an initial model

$d\%$ : minimum percentage of consensus data

**Output:** A curve set  $\mathbf{C}_{\tau \subseteq [t_1, t_K]}^r$  approximating  $\{\mathbf{w}_k^r\}$

**Procedure:**

1. Initialize *candidate-models* :=  $\{\}$ ;
2. **for** loop = 1 **to**  $C_K^m$  **do**
  - 2.1 *maybe-inliers* :=  $n$  samples randomly selected from  $\{\mathbf{w}_k^r\}$ ;
  - 2.2 *maybe-model* := A cubic B-spline model fitted to *maybe-inliers*;
  - 2.3 *consensus-set* := All the  $\mathbf{w}_k^r$ 's fitting the model under error threshold  $[\boldsymbol{\mu}^r - 3\boldsymbol{\sigma}^r, \boldsymbol{\mu}^r + 3\boldsymbol{\sigma}^r]$ ;
  - 2.3 **if** # consensus-set >  $d\% \cdot K$  **do**
    - *better-model* := A cubic B-spline model fitted to *consensus-set*;
    - Add *better-model* to *candidate-models*;
3. Output the model  $\mathbf{C}_{\tau \in [t_1, t_K]}^r$  from *better-model* with minimum prediction error for  $\{\mathbf{w}_k^r\}$ .

Although we have purified the data and the model as described in Section 2.2, in some extreme cases non-age-related variations are still highly mixed with those age-related ones and cannot be excluded by the aging-oriented face model (e.g., expression lines are quite similar to wrinkles in appearance). Therefore, we further treat these data as outliers and use RANdom SAMple Consensus [13] to remove them, as the circled star marks illustrated in Fig. 6. Statistical analysis results show that for a specific facial region  $r$ , the gradients of projection coefficients on the  $i$ th principal component follow Gaussian distribution  $N(\mu_i^r, (\sigma_i^r)^2)$ . With

this observation, for a pretrained aging model, we assume that the samples outside  $[\mu_i^r - 3\sigma_i^r, \mu_i^r + 3\sigma_i^r]$  of its prediction are outliers in step 2.3 of Algorithm 2, where  $\boldsymbol{\mu}^r = \{\mu_i^r, i = 1, 2, \dots, N^r\}$  and  $\boldsymbol{\sigma}^r = \{\sigma_i^r, i = 1, 2, \dots, N^r\}$ .

Considering that the complexity of face aging has been reduced by spatial and temporal decomposition, we adopt a function-based method to model the short-term aging patterns. In our implementation,  $\mathbf{C}_{\tau \in [t_1, t_K]}^r$  is defined as a set of cubic B-spline curves fitting the inliers of  $\{(\mathbf{w}_k^r, t_k), k = 1, 2, \dots, K\}$  (uncircled star marks in Fig. 6) with interpolation step  $\Delta t$  being 1 year. Please note that, although a large variety of functions were proposed for aging modeling, we do not use these empirical functions for the following reasons: First, the aging trajectories indicate that short-term face aging is still highly nonlinear, and there is no evidence that one specific function is sufficient for this task; second, a function with a fixed form is not flexible enough for handling the large variances in aging data.

## 4 LONG-TERM AGING MODEL

As discussed above, a long-term aging pattern is composed of some sequential short-term aging patterns. In this paper, we formulate the long-term face aging as a Markov process in the granularity of age spans, based on which we can infer a sequence of temporally overlapping short-term aging patterns in latter ages for the input face. After the inference, a long-term aging pattern is computed by concatenating short-term aging patterns under constraints. In addition, we consider the holisticness and diversity of aging process in long-term aging modeling.

### 4.1 Long-Term Aging Prediction by Concatenating Short-Term Aging Patterns

From the intrinsic characteristics of the face aging process, pattern concatenation should conform to two constraints: *smoothness constraints* and *consistency constraints*. The former means that face aging should be a smooth process to avoid abrupt changes. The latter implies that aging pattern of different regions should be kept consistent, as the face is an organic structure; thus, the aging of different regions is correlated with the others. In this section, we first formally defined measurements for these two constraints, based on which a conditional probability concatenating two short-term aging patterns is defined. Finally, an algorithm is proposed to predict long-term aging patterns by sampling with the defined concatenating probability. The whole procedure is summarized in Algorithm 3.

**Algorithm 3.** Predicting a long term aging pattern

**Input:** face parameters at current age  $t_0$

**Output:** predicted parameters  $\mathbf{C}_\tau^r$  over older ages  $\tau$

**Procedure:**

**for**  $t = 1$  **to**  $T$  **do**

**for**  $\forall r \in R$  **do**

1. Let  $\tau_1 = [t_0, t_0]$  and  $\mathbf{C}_{\tau_1}^r = \{\}$ ;
2. **while** there exist candidate patterns **do**
  - 2.1 Sample an aging pattern  $\mathbf{C}_{\tau_2}^r$  satisfying  $\tau_2 \cap \tau_1 \neq \phi$  and  $\sup \tau_1 < \sup \tau_2$  from the following probability (defined in Eq. 10):

$$\mathbf{C}_{\tau_2}^r \sim p(\mathbf{C}_{\tau_2}^r | \mathbf{C}_{\tau_1}^r, \mathbf{C}_{\tau_2}^{s \neq r});$$

- 2.2 Concatenate  $\mathbf{C}_{\tau_1}^r$  and  $\mathbf{C}_{\tau_2}^r$  at the time point defined in Eq. 11 to generate  $\mathbf{C}_{\tau_1 \cup \tau_2}^r$ ;
- 2.3  $\tau_1 := \tau_1 \cup \tau_2$ .

#### 4.1.1 Smoothness Constraints

Suppose  $\mathbf{C}_{\tau_1}^r = \{C_{i,\tau_1}^r, i = 1, 2, \dots, N^r\}$  and  $\mathbf{C}_{\tau_2}^r = \{C_{i,\tau_2}^r, i = 1, 2, \dots, N^r\}$  are two short-term aging patterns of the  $r$ th region over two overlapping time spans  $\tau_1$  and  $\tau_2$ , respectively, the distance between two curve sets is defined as

$$\mathcal{D}(\mathbf{C}_{\tau_1}^r, \mathbf{C}_{\tau_2}^r) = \frac{1}{|\tau_1 \cap \tau_2|} \int_{\tau_1 \cap \tau_2} \left( \sum_{i=1}^{N^r} F_i \cdot |\nabla C_{i,\tau_1}^r(t) - \nabla C_{i,\tau_2}^r(t)| \right) dt. \quad (6)$$

In this equation,  $F_i$  is the  $F$ -ratio of the  $i$ th principal component. Here, we assume that an aging pattern is reflected by the ‘‘changes’’ of face parameters over time instead of parameters themselves, so gradient operator  $\nabla$  is introduced for computing the distance between patterns.

Considering that natural face aging is smooth, we favor connecting the neighboring aging patterns with similar trends in the overlapping time span:

$$p(\mathbf{C}_{\tau_1}^r | \mathbf{C}_{\tau_2}^r) \propto \exp \{ -\mathcal{D}(\mathbf{C}_{\tau_1}^r, \mathbf{C}_{\tau_2}^r) \}. \quad (7)$$

Note that for the initial pattern selection, where  $|\tau_1 \cap \tau_2| = 0$ , we neglect  $\nabla$  and the denominator in (6) and favor the patterns with similar appearances to the input face at initial age.

#### 4.1.2 Consistency Constraints

As mentioned before, the aging of multiple facial regions is relevant, so we adopt an iterative approach to select proper aging paths for all the regions jointly.

Formally, let  $r$  and  $s$  denote two facial regions;  $\mathbf{C}_{\tau}^r$  and  $\mathbf{C}_{\tau}^s$  are, respectively, their corresponding aging patterns over  $\tau$ , with  $\mathbf{c}_{i,\tau}^r = \{C_{i,\tau}^r(t), t \in \tau\}$  and  $\mathbf{c}_{i,\tau}^s = \{C_{i,\tau}^s(t), t \in \tau\}$  being the pattern parameters on the  $i$ th principal component. We first compute the correlation between  $\mathbf{C}_{\tau}^r$  and  $\mathbf{C}_{\tau}^s$  according to (5):

$$\rho_{\tau}^{r,s} = \text{sigmoid} \left( \sum_{i=1}^{N^r} \sum_{j=1}^{N^s} F_i^r \cdot F_j^s \cdot \mathcal{P}(\mathbf{c}_{i,\tau}^r, \mathbf{c}_{j,\tau}^s) \right), \quad (8)$$

and then the consistency measurement between  $\mathbf{C}_{\tau}^r$  and  $\mathbf{C}_{\tau}^s$  is computed as follows:

$$\mathcal{E}(\mathbf{C}_{\tau}^r, \mathbf{C}_{\tau}^s) = \exp \{ -|\rho_{\tau}^{r,s} - \rho^{r,s}| \}. \quad (9)$$

The above equation favors the pattern pairs with similar correlations to that on training data, i.e.,  $\rho^{r,s}$ .

#### 4.1.3 Probabilistic Concatenation

The two aforementioned constraints defined in (7) and (9) are then combined together in order to select aging patterns. For this purpose, we define a conditional probability by probabilistic sampling. Specifically, the probability of concatenating curves  $\mathbf{C}_{\tau_1}^r$  and  $\mathbf{C}_{\tau_2}^r$  into a long-term pattern is computed as

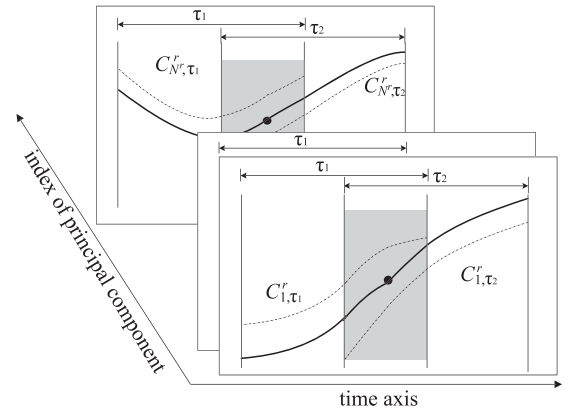


Fig. 7. Illustration of concatenating short-term aging patterns into a long-term pattern. The dash-dotted curves  $\{C_{i,\tau_1}^r\}$  and  $\{C_{i,\tau_2}^r\}$  represent two partially overlapping (shaded area) short-term aging patterns of region  $r$ . The solid curve set is the concatenated long-term pattern, and the filled circle is the concatenating time point.

$$p(\mathbf{C}_{\tau_2}^r | \mathbf{C}_{\tau_1}^r, \mathbf{C}_{\tau_2}^{s \neq r}) \propto p(\mathbf{C}_{\tau_1}^r | \mathbf{C}_{\tau_2}^r) \cdot \prod_{s \neq r} \mathcal{E}(\mathbf{C}_{\tau_2}^r, \mathbf{C}_{\tau_2}^s). \quad (10)$$

In the right side of the equation, the first term penalizes abrupt appearance changes in a specific facial region, while the second term penalizes inconsistency among different regions.

After sampling two sequential short-term aging patterns  $\mathbf{C}_{\tau_1}^r$  and  $\mathbf{C}_{\tau_2}^r$  according to the probability in (10), we concatenate them smoothly along the time axis. For this purpose, the time point with the minimum gradient difference is used to concatenate two overlapping patterns, as illustrated in Fig. 7.

$$t^* = \arg \min_{t \in \tau_1 \cap \tau_2} \sum_{i=1}^{N^r} F_i \cdot |\nabla C_{i,\tau_1}^r(t) - \nabla C_{i,\tau_2}^r(t)|. \quad (11)$$

To sum up, the steps of building long-term aging patterns are described in Algorithm 3.

## 4.2 Holisticness of Face Aging

As known, face aging patterns in different facial regions are highly relevant. Therefore, under the guidance of the graphical model learned in Section 2.3, the consistency constraint defined in (9) addresses the dependency among aging trajectory of different subregions by iteratively concatenating sequential patterns for one region under conditional constraints from the others ones. To validate the effectiveness of our strategy, we conduct an experiment to compare the difference between using and not using the region dependency, i.e., the consistency constraint in concatenation. The results are shown in Fig. 8, which shows that neglecting the relationship among the aging of different regions leads to aging progression violating the aging mechanisms, e.g., there exist asymmetries between two half faces; while the graphical model rectifies these inconsistencies implicitly.

## 4.3 Diversity of Face Aging

Intuitively, face aging has uncertainty, as both internal factors (probably genes) and external factors (e.g., nutrition, lifestyles, and health) affect the aging process. By analyzing the short-term aging patterns in Section 3, we can also find

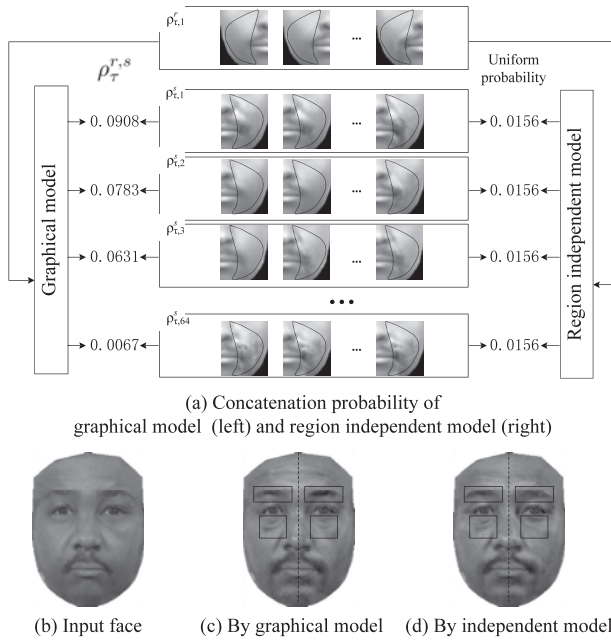


Fig. 8. Comparison between the graphical face model and the region-independent compositional model.

that there exist certain differences in the slopes of curves describing them, and this diversity indicates the intrinsic variability in both aging pattern and aging rate.

Therefore, it is desirable that a face aging model should have the capacity to simulate the uncertainty. Attributed to the probabilistic sampling procedure in Algorithm 3, our aging model naturally simulates the intrinsic diversity of aging process. For an input young face, as shown in the leftmost column of Fig. 9, we compute its facial parameters and predict their evolution according to (10). The multiple plausible aging results reconstructed from the predicted parameters are displayed in 2 ~ 4 columns.

It is worth noting that facial hair is indicative of the face ages in male training databases; thus, its appearance changes are also learned by the proposed aging model, although no explicit facial hair model is built in our method.

## 5 EXPERIMENTAL DATA

This section describes the adopted aging databases for model training, as well as the synthesized aging data by the

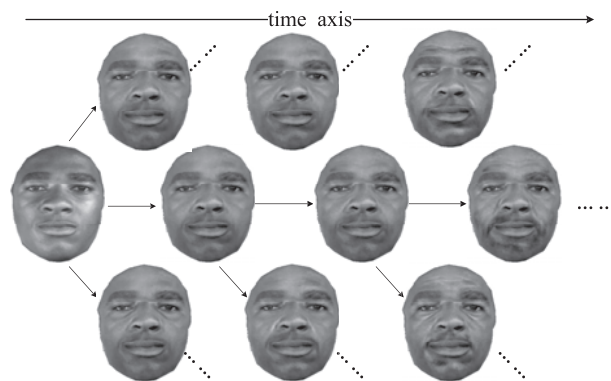


Fig. 9. The diversity of age progression.

TABLE 1  
Percentage of Inliers Selected by Data Purification (Percent)

	<sup>1</sup> Afam Adult		Caucasian Adult		Child	
	Male	Female	Male	Female	Boy	Girl
Shape	99	82	72	81	76	60
$r_{forehead}$	86	67	53	66	—	—
$r_{mouth}$	75	53	60	53	—	—
$r_{chin}$	75	57	51	53	—	—
$r_{eyebag}^{(2)L}$	88	63	58	63	—	—
$r_{canthus(L)}$	81	64	68	59	—	—
$r_{cheek(L)}$	73	80	62	63	—	—
$r_{eyebrow(L)}$	82	64	58	41	—	—
$r_{mouthcorner(L)}$	76	57	66	61	—	—
$r_{nose}$	85	80	63	65	—	—
$r_{glabella}$	67	68	51	55	—	—

<sup>1</sup>In this paper Afam is the abbreviation of African-American.

<sup>2</sup>Here 'L' denotes 'left', statistics of its right counterpart is omitted because similar percentage is kept.

learned aging models. These datasets are used for aging model evaluation in the next section.

### 5.1 Real Aging Data

In our experiments, we use two publicly available face aging databases: FG-NET [1] and MORPH [42].

We use images from FG-NET to form child growth models. After removing the adult faces over 20 years old and picking out 40 images as test data, we have 702 child faces of 48 boys and 32 girls. In consideration that there exist apparent differences between the growth patterns of boys and girls, two gender specific growth models are built separately.

We use faces in the MORPH database for adult aging modeling. One of its extended versions [41] includes 16,894 face images from 4,664 adults. To learn the long-term aging model, we use image sequences including photos snapped at more than two different ages between 18 and 54 years. The maximum and average age span are, respectively, 33 and 6.52 years. There are, altogether, 130 African-American male sequences, 133 African-American female sequences, 147 Caucasian male sequences, and 44 Caucasian female sequences in our training data set.

For each face image in both data sets, we labeled 83 landmarks for model training. The percentage of inliers selected by data purification is listed in Table 1, from which we can see the uneven distribution of non-aging variations both among ethnic groups and facial regions. For example, data in African-American groups are, on average, superior to those in Caucasian groups because of less hair occlusion and overexposure, and a large portion of the forehead and eyebrow regions in the Caucasian adult group is excluded due to heavy hair occlusion.

Since a child's face appearance changes much faster than an adult's with age increasing, we show dense prediction results (every other year) for children, while for adult aging we give results around every 10 years.

### 5.2 Synthetic Aging Data

In this section, we conduct aging prediction using the learned aging models.

For *child growth*, since shape change is the most prominent feature while texture is relatively stable and the texture of

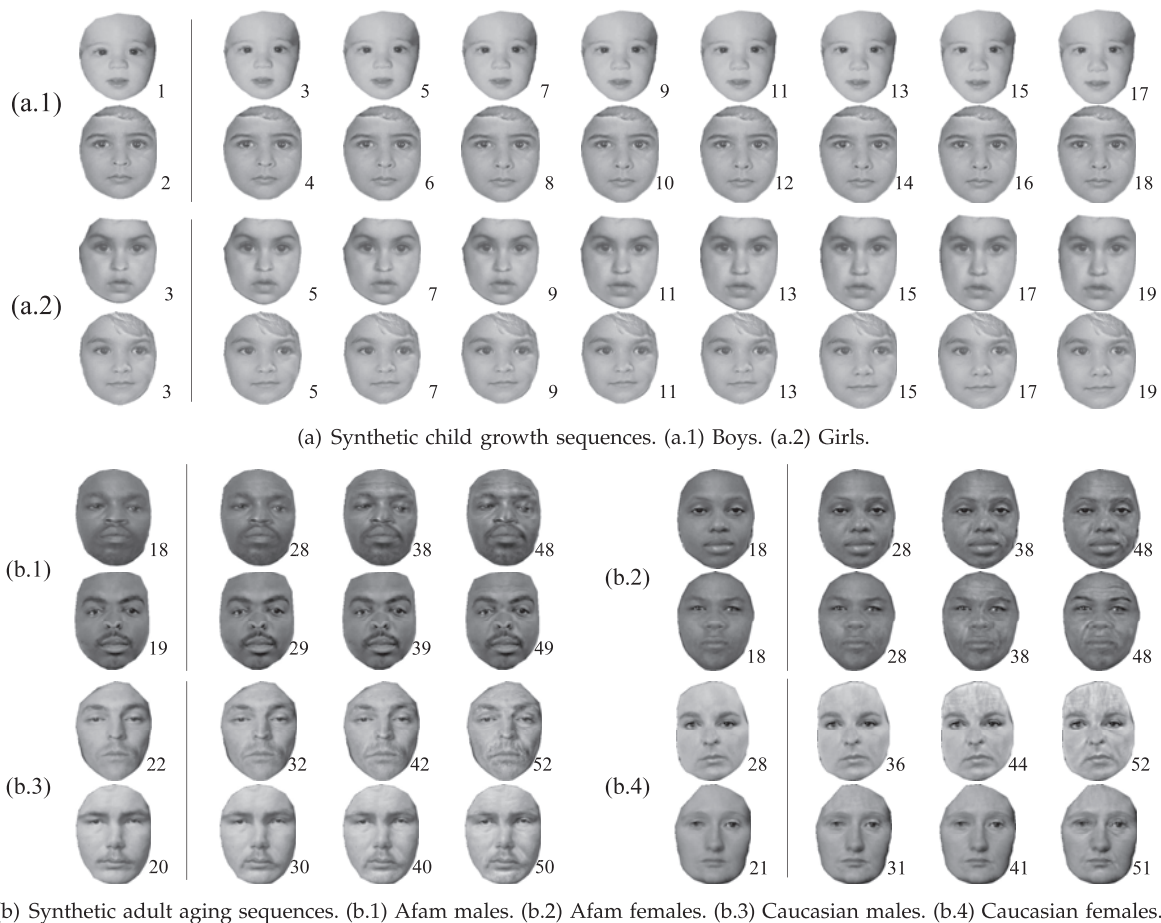


Fig. 10. Some aging prediction results generated by the learned aging models. In each group, the leftmost column shows the input images and the following columns are synthesized images at later ages, with the age labeled in the bottom right corner.

training data is highly contaminated by non-aging variations, we learn only shape growth models in this experiment. A given face with age and gender labels known can be transformed to target ages by the corresponding growth model. Some example results are shown in Fig. 10a.

In contrast, *for adult aging*, the variations among individuals from different genders and ethnic groups are prominent, so the subjects in the MORPH data set are divided into four groups: African-American males, African-American females, Caucasian males, and Caucasian females. Then, we build a face aging model for each gender-race group separately. For an input face image, we perform aging prediction by applying the aging model of his/her ethnic and gender group. Fig. 10b shows eight synthetic sequences, from which one can see the large diversity in the aging of different groups. There exist apparent artifacts in the forehead region of Caucasian females. It is worth pointing out that this is caused by insufficiency and heavy hair occlusions of the training data.

In aging prediction, the input images might be contaminated by nonaging features, similar to outliers in the model learning stage, due to pose or expression variations, specular highlights, glasses, or hair occlusion. To handle this problem, we reconstruct the input images with aging-oriented AAMs before aging prediction. Due to the procedure of data purification (Section 2.2.1) and model purification (Section 2.2.2), non-aging variations can be

excluded statistically. Some representative aging prediction results are shown in Fig. 11. In this figure, the leftmost face in each group is the input image and the one emphasized with a dotted rectangle is the reconstruction result with nonaging variations excluded. The aging prediction results on the right in each group are synthesized from the reconstructed faces.

Please note also that, compared with child growth, adult aging trajectories are affected by external factors to a larger extent. For this aging diversity, please also refer to Section 4.3 and Fig. 9.

## 6 AGING MODEL EVALUATION

Evaluation of face aging model is nontrivial since that aging is intrinsically uncertain. All the aging results under different environments may lie on a complex “manifold,” so direct comparison between synthetic faces and real aging data does not make sense. These challenges motivate researchers to explore indirect aging model evaluation, e.g., Lanitis et al. [24], Ramanathan and Chellappa [37], Patterson et al. [32] validate their aging model via induced improvement in aging-robust face recognition, Lanitis et al. [24] and Tiddeman et al. [53] compare the perceived age of faces before and after age transformation, Tiddeman et al. [54] evaluate the realism and age accuracy of aging results subjectively. To analyze the literature, we utilize the two

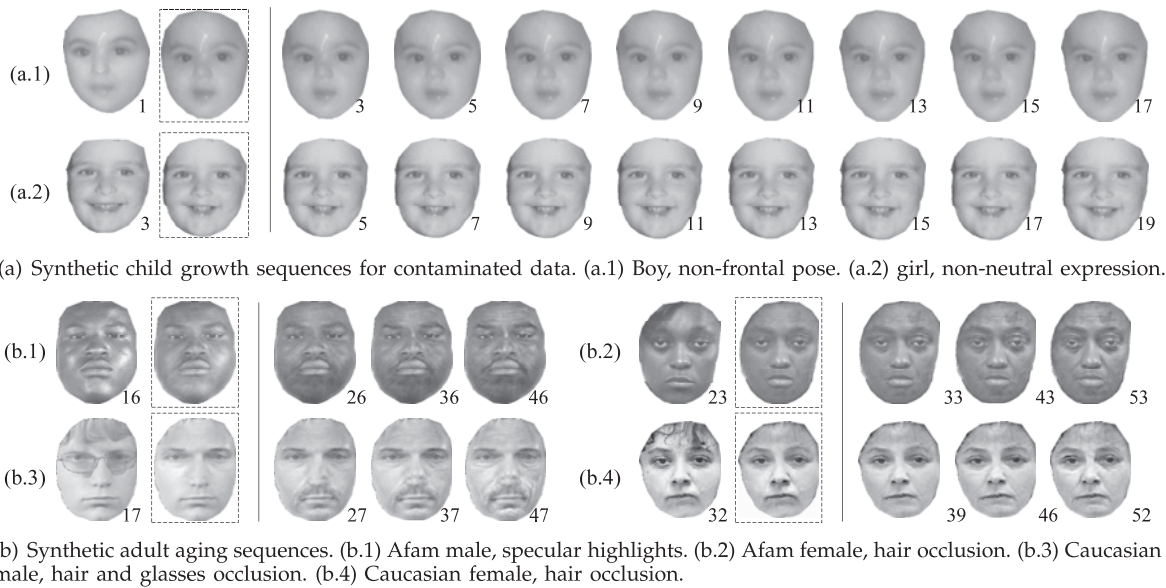


Fig. 11. Some aging prediction results of input images contaminated by non-age variations. In each group, the leftmost image shows the input face, the second image with bounding box is the reconstruction result excluding non-aging variations, and the following columns are synthesized images at later ages, with the age labeled in the right bottom corner.

most desirable criteria [22], [23], [52] for indirect aging model evaluation, which are: 1) *accuracy of age increase*, i.e., whether the synthetic faces are perceived to be at the target age; 2) *preservation of face identity*, i.e., whether the synthetic faces can still be recognized as the original person.

Guided by these two criteria, we perform both subjective and objective evaluation to the proposed method. First, we conduct age estimation and face recognition on synthetic aging results and real aging data, then quantitative analysis is adopted to measure the relative difference between the performances on the two data sets.

In this paper, we recruited 20 volunteers for subjective assessment. Since we were attempting to build a model consistent with the common sense, we recruited undergraduate and graduated students instead of the experts specialize in face aging studies. In all, 20 Asian students (10 males and 10 females) participated in the subjective assessment. The ages of the volunteers were between 22 and 30 and they all had normal (rectified) eyesight. One-third of the volunteers had experience living in western countries. We also add a training sequence to calibrate the volunteers' aging sense on different races.

### 6.1 Verification of ANOVA in Quantitative Aging Model Evaluation

Taking subjective age estimation as an example, we conducted an experiment to validate the adoption of ANOVA for aging model evaluation. First, we selected three front view datasets of Asians, Caucasians, and African-Americans, respectively, and invited the recruited volunteers (*without experience living in western countries*) to estimate their ages. The

TABLE 2  
MAE in Different Ethnic Groups (Years)

	Asian	Caucasian	Afam
MAE	4.44	11.58	10.48

Mean Absolute Error (MAE) is shown in Table 2. Then, we used ANOVA to analyze the estimation accuracy in the following two ways: 1) compare the difference between results on any two different data sets, i.e., the estimation accuracies on two different ethnic groups; 2) divide each set into two halves and compare the difference between them, i.e., the estimation accuracies on two subsets from the same ethnic group.

The ANOVA results are shown in Table 3, from which we can see that the accuracies between Asian and the other two groups are significantly different. This can be attributed to the fact that the volunteers are Asians and lack familiarity with other race aging patterns. The small difference between two half sets of the same ethnic group indicates that the age estimation performance is nearly the same on two sets, i.e., the two sets are almost the same under the criteria of age perception. The performance on Caucasian and African-American does not display significant difference either; this indicates that the recruited volunteers are not more familiar with one other race group than another, in terms of age estimation.

It is worth noting that the  $F$ -ratio measures the relative difference between performances on two data sets instead of the absolute one. For example, in spite of the large difference in estimation accuracies of Asians and that of Caucasians or African-Americans in Table 2, the  $F$ -ratios

TABLE 3  
 $F$ -Ratio and  $p$ -Value of Dataset Factor on Age Estimation Accuracy within/between Ethnic Groups

	Asian	Caucasian	Afam
Asian	$F=0.126$ $p=0.723$	$F=77.047$ $p=0.000$	$F=58.635$ $p=0.000$
Caucasian		$F=0.309$ $p=0.217$	$F=1.168$ $p=0.580$
Afam			$F=0.322$ $p=0.572$

TABLE 4  
Datasets Used in Model Evaluation Experiments

Set <sup>1</sup>	Size <sup>2</sup>	Notes about data set
A	2×20×10	Synthetic faces of 20 boys & 20 girls
B	2×20×10	Real faces of 20 boys & 20 girls
C <sub>1</sub>	3×40×5	Synthetic faces of 120 adults from 3 groups
C <sub>2</sub>	1×20×5	Synthetic faces of 20 Caucasian females
D <sub>1</sub>	3×40×5	Real faces of 120 sequences from 3 groups
D <sub>2</sub>	1×20×5	Real faces of 20 Caucasian females
E <sub>1</sub>	3×40+120	120 input faces for set C <sub>1</sub> +120 distractors
E <sub>2</sub>	1×20+20	20 input faces for set C <sub>2</sub> +20 distractors
F <sub>1</sub>	3×40+120	Initial faces of sequences in D <sub>1</sub> +distractors
F <sub>2</sub>	1×20+20	Initial faces of sequences in D <sub>2</sub> +distractors
G <sub>1</sub>	3×40×k	Latter faces of sequences in D <sub>1</sub> , k ≈ 5
G <sub>2</sub>	1×20×k	Latter faces of sequences in D <sub>2</sub> , k ≈ 5

<sup>1</sup> $\Sigma = \Sigma_1 \cup \Sigma_2$ , here  $\Sigma = C, D, E, F, G$

the test data of Caucasian females is less due to limitation of data size.

<sup>2</sup>Size = # of groups × # of subjects per group × # of images per subject, or  
Size = # of groups × # of subjects per group + # of distractors.

within each ethnic group are all quite small (see Table 3). The  $F$ -ratio is quite fit for aging model evaluation since the measurement does not rely on the performance of aging insensitive face recognition or face age estimation.

## 6.2 Evaluation of Proposed Aging Model

Using the proposed evaluation approach, we conduct a series of experiments to measure the validness of the learned aging models. For clarity, Table 4 lists the data sets used in the following experiments.

### 6.2.1 Evaluation on Age Accuracy of Synthetic Face

In evaluation of the child growth model, we select 20 boy's photos and 20 girl's photos (not in the training data set) under three years of age and synthesize 10 images at latter ages for each individual to compose set A, which thus includes 400 images. For a comparative study, the same number of real aging sequences are selected as set B. As for the evaluation of adult aging models, we select 20 young Caucasian female faces and 40 young faces from each of the other three groups, not in training data either. We synthesize five images across 3 ~ 4 decades for each individual. These resulting 700 synthetic images compose set C. Similarly, we also randomly select the same number of real images at different ages to compose set D. After data preparation, we perform both subjective and objective age estimation on these four sets and adopt Mean Absolute Error as a quantitative measurement of the age estimation accuracy.

*Subjective age estimation.* In this experiment, we compare the accuracies of subjective age estimation on real and synthetic data by the following procedure. After a training

procedure with 100 faces at different ages and in different groups, the volunteers were asked to estimate the age of each face image in set A ~ D. The MAEs on set A and B are both around 4 ~ 5 years and those on sets C and D are around 5 ~ 7 years, as shown in 3 ~ 4 rows of Table 5.

*Objective age estimation.* Automatic face aging estimation has been studied for years, and various algorithms were developed [15]. Here, we adopt an age estimation algorithm [51] with state-of-the-art performance to compare the objective age estimation accuracies on real and synthetic faces. To make a fair comparison, we first train a group-specific age estimator on a separate training data set, then perform age estimation on real images in sets B, D and synthetic images in sets A, C. The estimation results are shown in parallel with the subjective estimation results in 7 ~ 8 rows of Table 5, from which we can see that objective estimation obtains a performance comparable to that of subjective estimation, with the estimation error being around six years. The maximum MAE is 6.88 years, the result on adult Caucasian males. Such a larger estimation error may be due to the larger appearance variations in this group.

Quantitative measurements of the difference between age estimation accuracies on two sets are also shown in Table 5, which concludes that real images and synthetic images have similar appearances in terms of age perception, i.e., the proposed model is able to synthesize faces of the target age.

### 6.2.2 Evaluation on ID Preservation of Synthetic Face

*Subjective face recognition across ages.* Due to the limitations of short-term human memory, only 20 Caucasian females and 40 individuals from each of the other three groups are used for both real and synthetic aging data. We conduct a subjective face recognition experiment under the following settings: the 140 input faces for set C and the same number of distractors compose set E, from which the volunteers are asked to identify the faces in set C. For real aging data, we select from each group in the MORPH database the same number of aging sequences with synthetic data. Their photos at the initial ages and 140 distractors together compose set F, and the photos at latter ages are denoted as set G. Similarly, we ask the volunteers to recognize the faces in set G, with set F being the gallery set.

The recognition results on four groups are plotted in Figs. 12a and 12b, which show that the recognition rate generally degenerates with increasing time span on both sets. This is evidently consistent with the intuition that face recognition across longer age periods is more challenging. Please note that the missing data at larger age spans in Fig. 12a is due to the limited real aging sequences. Another

TABLE 5  
MAE (Years) of Subjective and Objective Age Estimation and the Quantitative Analysis

		Afam Adult		Caucasian Adult		Child	
		Male	Female	Male	Female	Boy	Girl
Subjective	Real Data	6.58	6.83	5.83	6.15	4.90	5.08
	Synthetic Data	6.56	6.31	5.52	6.31	4.59	5.15
	ANOVA Analysis	$F=2.074$ $p=0.152$	$F=4.996$ $p=0.027$	$F=0.506$ $p=0.479$	$F=1.752$ $p=0.192$	$F=0.842$ $p=0.360$	$F=0.642$ $p=0.424$
Objective	Real Data	6.81	5.91	6.22	5.96	6.54	6.51
	Synthetic Data	6.17	6.85	6.07	6.88	6.19	6.21
	ANOVA Analysis	$F=3.367$ $p=0.057$	$F=1.754$ $p=0.186$	$F=0.320$ $p=0.572$	$F=4.876$ $p=0.029$	$F=0.052$ $p=0.820$	$F=0.089$ $p=0.767$

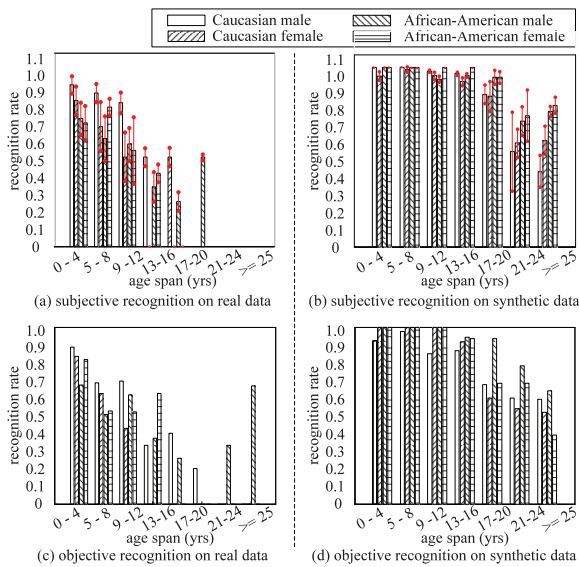


Fig. 12. Performance of subjective and objective face recognition.

point worth noting is that the recognition rate shows certain individual differences in face recognition abilities, as illustrated by the error bars in Figs. 12a and 12b. The among-individual variance on real data is larger than that on synthetic data, which can be attributed to the large non-age-related variations in real data.

*Objective face recognition across ages.* In this experiment, we run automatic face recognition on 400 images in set C and all the faces in the training data set using a face recognition algorithm developed on Gabor features and LDA-based classifier [49]. To keep the evaluation objective, the recognition algorithm is trained on an absolutely separate database including data from different genders and ethnic groups.

Figs. 12c and 12d show the recognition results on real and synthetic aging data, respectively. Although there exist significant differences between recognition rates on real and synthetic data and the those on four groups are also slightly different, the trend of recognition rate degeneration with increasing time span is generally consistent.

By comparing the face recognition rates on real data (Figs. 12a and 12c) and those on synthetic data (Figs. 12b and 12d), one can see that the recognition rate on synthetic data is apparently higher. The quantitative comparison between performances on real and synthetic data is displayed in Table 6, which also shows a significant difference between face identity performances on two data sets. The large difference is attributed to the distinction between synthetic data and real aging data: The former reflects only aging-related variations, while the latter includes many other non-age-related variations. In spite of the large difference in recognition rates on real and synthetic data, the high performance of subjective and objective recognition on synthetic data indicates that our aging model preserves face identity effectively.

Beyond the above evaluations of the proposed CONGRE model, we also compare the CONGRE aging model with the one we previously proposed in [52]. Since the previous work classifies face ages into discrete groups with a time step of 10 years and the evaluation results show that around 30 percent synthetic faces belong to a group other than the

TABLE 6

ANOVA of Face Recognition Rates on Real and Synthetic Data

	Afam Adult		Caucasian Adult	
	Male	Female	Male	Female
Subjective	$F=46.941$ $p=0.000$	$F=38.566$ $p=0.000$	$F=0.517$ $p=0.474$	$F=4.645$ $p=0.032$
Objective	$F=197.007$ $p=0.000$	$F=61.519$ $p=0.000$	$F=17.548$ $p=0.000$	$F=41.346$ $p=0.000$

intended age group, while the CONGRE model regards face aging as a continuous process and the prediction results are with an estimation error being around six years, nearly of the same accuracy with that on real data, the CONGRE model is thus much more accurate than the previous one. Therefore, here we mainly compare two methods in term of ID preservation quantitatively. We mix the aging results generated by two models with the same number of distractors and conduct a face recognition experiment similar to Section 6.2. To make a fair comparison, we use the same face recognition algorithm and volunteers and crop out the hair region of the aging results generated by previous model; the result is shown in Fig. 13. From the results, we can see that the recognition rate of synthetic results generated by the CONGRE model is higher; this may be attributed to the fact that we learn aging trajectories from real aging sequences instead of similar patches. In addition, the age span is at a much finer scale than that of previous work, so the evaluation results are more accurate.

## 7 CONCLUSIONS AND DISCUSSIONS

This paper proposes a CONGRE face aging model to construct long-term face aging patterns from impure, partially dense aging databases. The proposed model addresses two rarely touched problems in face aging modeling: 1) removing appearance variations due to nonaging factors statistically, which exist widely in face aging databases; 2) building credible long-term aging patterns from currently available databases containing only short-term aging sequences, which are close to the real aging process in a short time span and smooth in a long time span. The studies in this paper also reveal that large diversity in aging patterns exists among individuals from different races and genders; this implies that building group-specific face aging models is indeed necessary.

Although the proposed method is able to synthesize plausible aging sequences, there still exists much room for improvement if compared with those generated by forensic

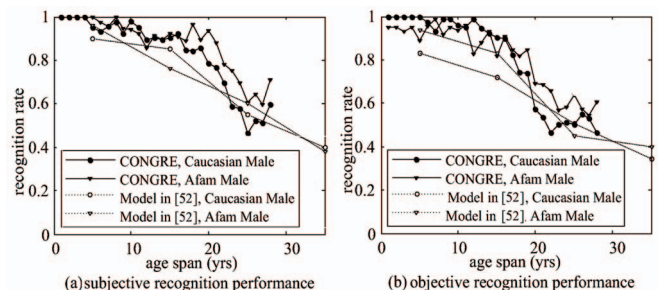


Fig. 13. Comparison of ID preservation between the proposed model and the approach in [52].

artists [33], who are usually trained in facial anatomy, aging, and forensic science, etc. For example, the decomposition rule adopted is far from being precise enough to take into account the complex functions of muscles in face aging. It is also clear that the individuality of the input face is not considered sufficiently during the aging prediction, etc.

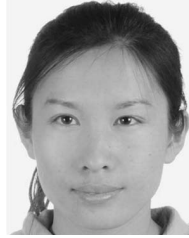
## ACKNOWLEDGMENTS

Jinli Suo was with the Graduate University of Chinese Academy of Sciences (CAS) and also an intern with the Key Lab of Intelligent Information Processing of CAS, where most of this work was done when she was a PhD candidate. This work was partially supported by the NSFC under contract, Nos. 61025010 and 60872077, and the 973 Program under contract 2009CB320902. Jinli Suo and Qionghai Dai were also supported by 973 Program under contract No. 2010CB731800, NSFC under contract Nos. 61171119, 61035002, and 61021063. A short version of this work appeared in [50].

## REFERENCES

- [1] "Face and Gesture Recognition," *Network: FG-NET Aging Database*, <http://sting.cycollege.ac.cy/alanitis/fgnetaging/>, 2012.
- [2] 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," *J. Forensic Science Int'l*, vol. 172, no. 1, pp. 1-9, Apr. 2007.
- [3] Y. Bando, T. Kuratate, and T. Nishita, "A Simple Method for Modeling Wrinkles on Human Skin," *Proc. IEEE Pacific Conf. Computer Graphics and Applications*, pp. 166-175, 2002.
- [4] R.G. Behrens, *An Atlas of Growth in the Aging Cranio Facial Skeleton*. Univ. of Michigan Press, 1985.
- [5] A.C. Berg, "Aging of Orbicularis Muscle in Virtual Human Faces," *Proc. IEEE Int'l Conf. Information Visualization*, pp. 164-168, 2003.
- [6] A.C. Berg, F.J.P. Lopez, and M. Gonzalez, "A Facial Aging Simulation Method Using Flaccidity Deformation Criteria," *Proc. IEEE Int'l Conf. Information Visualization*, pp. 791-796, 2006.
- [7] S. Biswas, G. Aggarwal, N. Ramanathan, and R. Chellappa, "A Non-Generative Approach for Face Recognition across Aging," *Proc. IEEE Int'l Conf. Biometrics: Theory, Applications, and Systems*, 2008.
- [8] L. Boissieux, G. Kiss, N.M. Thalmann, and P. Kalra, "Simulation of Skin Aging and Wrinkles with Cosmetics Insight," *Proc. Eurographics Workshop Animation, Computer Animation, and Simulation*, pp. 15-27, 2000.
- [9] D.M. Burt and D.I. Perrett, "Perception of Age in Adult Caucasian Male Faces: Computer Graphic Manipulation of Shape and Color Information," *Royal Soc. London*, vol. 259, pp. 137-143, 1995.
- [10] J.W. Carls, R. Raines, M. Grimaila, and S. Rogers, "Biometric Enhancements: Template Aging Error Score Analysis," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, 2008.
- [11] T.F. Cootes, G.J. Edwards, and C.J. Taylor, "Active Appearance Models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681-685, June 2001.
- [12] P. Ekman and E.L. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Face Action Coding System (FACS)*, second ed. Oxford Univ. Press, 2005.
- [13] M.A. Fischler and R.C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Comm. ACM*, vol. 24, no. 6, pp. 381-395, 1981.
- [14] R.A. Fisher, *Statistical Methods for Research Workers*, first ed. Oliver and Boyd, 1925.
- [15] Y. Fu, G. Guo, and T.S. Huang, "Age Synthesis and Estimation via Faces: A Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 1955-1976, Nov. 2010.
- [16] Y. Fu and N. Zheng, "M-Face: An Appearance-Based Photo-realistic Model for Multiple Facial Attributes Rendering," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 7, pp. 830-842, July 2006.
- [17] M. Gandhi, "A Method for Automatic Synthesis of Aged Human Facial Images," PhD thesis, McGill Univ., 2004.
- [18] D. Hubball, M. Chen, and P.W. Grant, "Image-Based Aging Using Evolutionary Computing," *Computer Graphics Forum*, vol. 27, no. 2, pp. 607-616, June 2008.
- [19] D. Hunter, "Synthesis of Facial Ageing Transforms Using Three-Dimensional Morphable Models," PhD thesis, Univ. of St. Andrews, 2009.
- [20] T.J. Hutton, B.F. Buxton, P. Hammond, and H.W. Potts, "Estimating Average Growth Trajectories in Shape-Space Using Kernel Smoothing," *IEEE Trans. Medical Imaging*, vol. 22, no. 6, pp. 747-753, June 2003.
- [21] F. Jiang and Y. Wang, "Facial Aging Simulation Based on Super Resolution in Tensor Space," *Proc. IEEE Int'l Conf. Image Processing*, pp. 1648-1651, 2008.
- [22] A. Lanitis, "Comparative Evaluation of Automatic Age-Progression Methodologies," *EURASIP J. Advances in Signal Processing*, vol. 8, no. 2, pp. 1-10, Jan. 2008.
- [23] A. Lanitis, "Evaluating the Performance of Face-Aging Algorithms," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, 2008.
- [24] A. Lanitis, C.J. Taylor, and T.F. Cootes, "Toward Automatic Simulation of Aging Effects on Face Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 442-455, Apr. 2002.
- [25] W.F. Larrabee, K.H. Makielski, and J. Henderson, *Surgical Anatomy of the Face*, second ed. Lippincott Williams and Wilkins, 2003.
- [26] W.S. Lee, Y. Wu, and N.M. Thalmann, "Cloning and Aging in a VR Family," *Proc. IEEE Virtual Reality*, pp. 61-68, 1999.
- [27] F.R. Leta, A. Conci, D. Pamplona, and I. Itanguy, "Manipulating Facial Appearance through Age Parameters," *ACM Trans. SIBGRAP*, pp. 167-172, 1996.
- [28] F.R. Leta, D. Pamplona, H.I. Weber, A. Conci, and I. Pitanguy, "A Study of the Facial Aging—A Multidisciplinary Approach," *J. Brazilian Soc. of Mechanical Sciences*, vol. 22, no. 3, pp. 489-501, 2000.
- [29] H. Ling, S. Soatto, N. Ramanathan, and D.W. Jacobs, "A Study of Face Recognition as People Age," *Proc. IEEE Int'l Conf. Computer Vision*, 2007.
- [30] S. Mukaida, H. Ando, K. Kinoshita, M. Kamachi, and K. Chihara, "Facial Image Synthesis Using Age Manipulation Based on Statistical Feature Extraction," *Proc. Int'l Conf. Visualization, Imaging, and Image Processing*, pp. 12-17, 2002.
- [31] U. Park, Y. Tong, and A.K. Jain, "Face Recognition with Temporal Invariance: A 3D Aging Model," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, 2008.
- [32] E. Patterson, K. Ricanek, M. Albert, and E. Boone, "Automatic Representation of Adult Aging in Facial Images," *Proc. IASTED Int'l Conf. Visualization, Imaging and Image Processing*, p. 612, 2006.
- [33] E. Patterson, A. Sethuram, M. Albert, and K. Ricanek, "Comparison of Synthetic Face Aging to Age Progression by Forensic Sketch Artist," *Proc. IASTED Int'l Conf. Visualization, Imaging and Image Processing*, pp. 247-252, 2007.
- [34] E. Patterson, A. Sethuram, K. Ricanek, and F. Bingham, "Improvements in Active Appearance Model Based Synthetic Age Progression for Adult Aging," *Proc. IEEE Int'l Conf. Biometrics Theory, Applications and Systems*, 2009.
- [35] J.B. Pittenger and R.E. Shaw, "Aging Faces as Visceral-Elastic Events: Implications for a Theory of Nonrigid Shape Perception," *J. Experimental Psychology: Human Perception and Performance*, vol. 1, no. 4, pp. 374-382, Nov. 1975.
- [36] N. Ramanathan and R. Chellappa, "Face Verification across Age Progression," *IEEE Trans. Image Processing*, vol. 15, no. 11, pp. 3349-3361, Nov. 2006.
- [37] N. Ramanathan and R. Chellappa, "Modeling Age Progression in Young Faces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 387-394, 2006.
- [38] N. Ramanathan and R. Chellappa, "Shape Modeling Textural Variations in Aging Faces," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, 2008.
- [39] N. Ramanathan and R. Chellappa, "Age Progression in Human Face: A Survey," *J. Visual Languages and Computing*, vol. 20, no. 3, pp. 131-144, 2009.

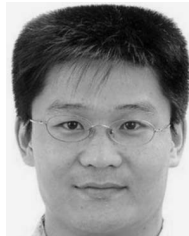
- [40] N. Ramanathan, R. Chellappa, and S. Biswas, "Computational Methods for Modeling Facial Aging: A Survey," *J. Visual Languages and Computing*, special issue on advances in multimodal biometric systems, vol. 20, no. 3, pp. 131-144, June 2009.
- [41] A.W. Rawls and K. Ricanek, "MORPH: Development Optimization of a Longitudinal Age Progression Database," *Proc. Int'l Conf. Biometric ID Management and Multimodal Comm.*, 2009.
- [42] K. Ricanek Jr. and T. Tesafaye, "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 341-345, 2006.
- [43] D.A. Rowland and D.I. Perrett, "Manipulating Facial Appearance through Shape and Color," *IEEE Computer Graphics and Applications*, vol. 15, no. 5, pp. 70-76, Sept. 1995.
- [44] C.M. Scandrett, C.J. Solomon, and S.J. Gibson, "A Person-Specific, Rigorous Aging Model of the Human Face," *Pattern Recognition Letters*, vol. 27, no. 15, pp. 1776-1787, Nov. 2006.
- [45] K. Scherbaum, M. Sunkel, H.P. Seidel, and V. Blanz, "Prediction of Individual Non-Linear Aging Trajectories of Faces," *Computer Graphics Forum in Eurographics*, vol. 26, no. 3, pp. 285-294, 2007.
- [46] G. Schroeder, L.P. Magalhaes, and R. Rodrigues, "Facial Aging Using Image Warping," *Proc. Western New York Image Processing Workshop*, 2007.
- [47] A. Sethuram, K. Ricanek, and E. Patterson, "A Hierarchical Approach to Facial Aging," *Proc. IEEE Int'l Workshop Analysis and Modeling of Faces and Gestures*, 2010.
- [48] R. Singh, M. Vatsa, A. Noore, and S.K. Singh, "Age Transformation for Improving Face Recognition," *Proc. Int'l Conf. Pattern Recognition and Machine Intelligence*, pp. 576-583, 2007.
- [49] Y. Su, S. Shan, X. Chen, and W. Gao, "Hierarchical Ensemble of Global and Local Classifiers for Face Recognition," *IEEE Trans. Image Processing*, vol. 18, no. 8, pp. 1885-1896, Aug. 2009.
- [50] J. Suo, X. Chen, S. Shan, and W. Gao, "Learning Long Term Face Aging Patterns from Partially Dense Aging Databases," *Proc. IEEE Int'l Conf. Computer Vision*, 2009.
- [51] J. Suo, T. Wu, S.C. Zhu, S. Shan, X. Chen, and W. Gao, "Design Sparse Features for Age Estimation Using Hierarchical Face Model," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, 2008.
- [52] J. Suo, S.C. Zhu, S. Shan, and X. Chen, "A Compositional and Dynamic Model for Face Aging," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 385-401, Mar. 2010.
- [53] B.P. Tiddeman, D.M. Burt, and D.I. Perrett, "Prototyping and Transforming Facial Textures for Perception Research," *IEEE Computer Graphics and Applications*, vol. 21, no. 5, pp. 42-50, Sept./Oct. 2001.
- [54] B.P. Tiddeman, M.R. Stirrat, and D.I. Perrett, "Towards Realism in Facial Prototyping: Results of a Wavelet MRF Method," *Computer Graphics Forum*, vol. 24, no. 3, pp. 449-456, 2006.
- [55] J.T. Todd, L.S. Mark, R.E. Shaw, and J.B. Pittenger, "The Perception of Human Growth," *Scientific Am.* vol. 242, no. 2, pp. 132-144, 1980.
- [56] J. Wang and C. Ling, "Artificial Aging of Faces by Support Vector Machines," *Proc. Canadian Conf. Artificial Intelligence*, pp. 499-503, 2004.
- [57] J. Wang, Y. Shang, G. Su, and X. Lin, "Age Simulation for Face Recognition," *Proc. Int'l Conf. Pattern Recognition*, pp. 913-916, 2006.
- [58] Y. Wu, P. Beylot, and N.M. Thalmann, "Skin Aging Estimation by Facial Simulation," *Proc. IEEE Conf. Computer Animation*, pp. 210-219, 1999.
- [59] M.S. Zimble, M.S. Kokosk, and J.R. Thomas, "Anatomy and Pathophysiology of Facial Aging," *Facial Plastic Surgery Clinics of North Am.*, vol. 9, no. 2, pp. 179-187, 2001.



**Jinli Suo** received the BS degree in computer science from Shandong University, Shandong, China, in 2004, and the PhD degree from the Graduate University of Chinese Academy of Sciences (CAS), Beijing, China, in 2010. She is currently a lecturer in Department of Automation, Tsinghua University, Beijing, China. Her research interests mainly include face aging modeling and perception of human faces.



**Xilin Chen** received the BS, MS, and PhD degrees in computer science from the Harbin Institute of Technology (HIT), China, in 1988, 1991, and 1994, respectively. He is now the director of the Key Laboratory of Intelligent Information Processing, CAS. He has served as a program committee member for more than 40 international conferences, and is the (co)author of more than 200 papers. He is a senior member of the IEEE.



**Shiguang Shan** received the MS degree in computer science from the Harbin Institute of Technology, China, in 1999, and the PhD degree in computer science from the Institute of Computing Technology (ICT), CAS, Beijing, in 2004. He has been with ICT, CAS since 2002 and has been a professor since 2010. His research interests cover image analysis, pattern recognition, and computer vision. He is a member of the IEEE.



**Wen Gao** received the MS degree in computer science from the Harbin Institute of Technology, China, in 1985 and the PhD degree in electronics engineering from the University of Tokyo, Japan, in 1991. He is currently a professor with the School of Electronics Engineering and Computer Science, Peking University, Beijing. He has been leading efforts to study video coding, face recognition, etc. He is a fellow of the IEEE.



**Qionghai Dai** received the BS degree in mathematics from Shaanxi Normal University, Xi'an, China, in 1987, and the ME and PhD degrees in computer science from Northeastern University, Shenyang, China, in 1994 and 1996, respectively. Since 1997, he has been with Tsinghua University, Beijing, China. His current research areas include signal processing, computer vision, and computer graphics. He is a senior member of the IEEE.

► For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/publications/dlib](http://www.computer.org/publications/dlib).