# Learning Long Term Face Aging Patterns from Partially Dense Aging Databases


Jinli Suo[1,2,3] jlsuo@jdl.ac.cn

Xilin Chen[2] Shiguang Shan[2] {xlchen,sgshan}@ict.ac.cn

Wen Gao[1,4] wgao@pku.edu.cn

[1]Graduate University of Chinese Academy of Sciences(CAS), 100190, China
[2]Key Lab of Intelligent Information Processing of CAS,
Institute of Computing Technology, CAS, Beijing, 100190, China
[3]Lotus Hill Institute for Computer Vision and Information Science, 436000, China
[4]School of Electronic Engineering and Computer Science, Peking University, 100871, China



## Abstract

*Studies on face aging are handicapped by lack of long term dense aging sequences for model training. To handle this problem, we propose a new face aging model, which learns long term face aging patterns from partially dense aging databases. The learning strategy is based on two assumptions: (i) short term face aging pattern is relatively simple and is possible to be learned from currently available databases; (ii) long term face aging is a continuous and smooth Markov process. Adopting a compositional face representation, our aging algorithm learns a function-based short term aging model from real aging sequences to infer facial parameters within a short age span. Based on the predefined smoothness criteria between two overlapping short term aging patterns, we concatenate these learned short term aging patterns to build the long term aging patterns. Both the subjective assessment and objective evaluations of synthetic aging sequences validate the effectiveness of the proposed model.*


## 1. Introduction

Face aging is attracting increasing interest from researchers for its roles in real world applications(e.g., looking for lost children or wanted fugitives, developing face recognition systems[5, 24, 29, 30, 43] robust to age related variations, template renewal[8], et al.) In recent years, a lot of efforts focus on face aging modeling, building face aging databases[1, 33] and aging model evaluation[20].

### 1.1. Previous work

Researchers have made great efforts on face aging modeling and lots of algorithms were proposed. Generally these approaches fall into two groups: physical model based and example based.

Approaches based on physical model simulate face aging by modeling biological structure and aging mechanisms of cranium, muscles or skin. For example, Wu et al.[44] and Boissieux et al.[6] build layered skin models to simulate the skin deformations in aging process. Berg and Justo[3] simulate the aging of orbicularis muscles. Other similar studies include Golovinskiy et al.'s[14], Lee et al.'s[22], Bando et al.'s[2], Ramanathan and Chellappa's[32] work. Due to the subtleness of both facial structure and face aging mechanisms, physical model based methods are often complex and computationally expensive. These disadvantages cause it difficult to get realistic results by physical modeling.

Two main streams of example based approaches are respectively prototype based and function based. (i) *Prototype based methods*[7, 13, 23, 34, 40, 41] firstly divide age range into discrete age groups and compute the mean face of each age group as its prototype, and then define the differences between prototypes as axis of aging transformation. Wang et al.[43] and Hill et al.[15] extend this approach to PCA space and Park et al.[26] apply it to 3D face data. Since some details crucial for age perception are lost in prototype computation, some other researchers specially work on rendering high resolution aging results[12]. Although Scandrett et al.[35] consider the attributes of specific individuals in their work, most prototype methods are able to extract only average aging patterns, so the individuality and diversity of face aging are neglected. (ii) *Function based* methods describe the relationship between a face image or its facial parameters and corresponding age label with an explicit function, such as strain elicit transformation[28], quadratic function[21, 23, 27], support vector regression[19, 36, 42], kernel smoothing method[17] or an implicit function[4, 25]. Fang and Wang[18] directly build a mapping function between a young face and its appearance in latter ages. Anthropometry evidences can be used as constraints of these mapping functions, such as in Ramanathan and Chellappa's work[31]. All those functions need considerable real aging sequences for parameter learning and simple functions are insufficient for modeling the large changes in facial appearances over long periods. In spite of these drawbacks, the method can be applied to short term face aging.

In all, physical model based approaches are constrained by the complexities and example based approaches suffer from lacking of training data.

### 1.2. Motivation and basic ideas

Since it is time consuming to collect long term(i.e. across 3-4 decades) dense face aging sequences, currently available face aging databases contain mostly short term sequences(i.e. age span smaller than 10 years), we refer to these databases as partially dense datasets. Learning long term aging patterns from partially dense datasets is a strategy worth considering for example based approaches.

As a common sense, long term aging pattern is a smooth Markov process(we only refer to nature aging in this paper) composed of a series of short term aging patterns and similarities exist among the short term aging patterns in a same time span, especially for the individuals from the same ethnic group and of the same gender. We propose a face aging model which firstly learns short term aging patterns from real aging sequences and then sequentially concatenate similar short term aging patterns into long ones. The aging algorithm is composed of two steps:

Step 1) *Learning short term face aging patterns.* With large number of short term face aging sequences from publicly available face aging databases, such as FG-NET and MORPH, our model extracts short term aging patterns from the real aging sequences. This ensures that the learned patterns are close to real aging process. Since the short-term appearance changes are relatively small, we use a function based approach to model the short term face aging.

Step 2) *Concatenating short term aging patterns into long term patterns.* Formulating long term aging as a Markov process, our model concatenates partially overlapping short term aging patterns sequentially into long term ones based on the predefined smoothness criteria. The diversity of aging is simulated by sampling different subsequent short term patterns probabilistically, with the probability being computed from the similarity measurement between two overlapping patterns.

To our knowledge, this is the first attempt to learn long term aging patterns from partially dense datasets. One recent work bearing some similarity with ours is [38], in which a dynamic face aging model based on a hierarchical face representation is proposed. However, the dynamics in [38] are learned from predefined similarity metrics instead of real aging data and the continuous aging process is simplified by dividing age range into 5 discrete groups.

The rest of this paper is organized as follows: In Sec. 2 we describe the proposed compositional aging specific AAM model. Then Sec.3 and Sec.4 explain the learning of short term aging patterns and long term aging patterns respectively. In Sec.5 we design some experiments to validate the proposed model. Finally, this paper concludes in Sec.6 with a summary and some discussions on future work.

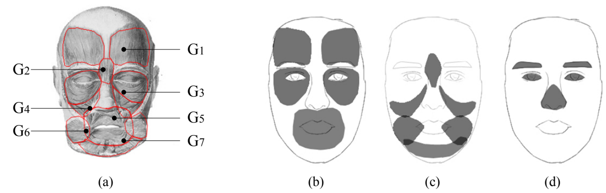



Figure 1. The compositional face representation inspired by muscle model. (a) shows the muscle model adopted in Facial Action Coding System[10] and the muscle grouping based on it. (b) and (c) are the divided facial subregions based on the muscle grouping in (a). (d) shows the subregions with few muscles.

## 2. Compositional aging specific face model

### 2.1. Muscular decomposition

Due to its simplicity and effectiveness in face representation, Active Appearance Model(AAM)[9] is widely used in face analysis, including aging related studies, such as [21],[43], et al.. However, the holistic AAM representation has some limitations in the task of face aging modeling for following two reasons: (1) a global AAM model capable of describing large appearance changes across decades is often with high dimension and statistical learning approaches suffer from the curse of dimension problem, which is especially serious in face aging problem; (2) aging mechanisms of different regions are specifically related with their biological structures and display large varieties in both aging pattern and aging velocity, thus face aging is a highly nonlinear process. Therefore, we adopt a region based AAM model, which targets to reduce 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 declines gradually because of the loss of muscle tone, wrinkles emerge due to the contraction of expression muscles, et al.), we take into account anatomical evidences when building our region based face representation. Ekman et al.[10] conduct a deep study on the structure and function of facial muscles, and apply it to Face Action Coding System(FACS) successfully. Adopting a similar muscle model, Ramanathan and Chellappa[32] simulate the face shape changes caused by aging of facial muscles. On the basis of these two previous analysis on facial muscles, we select muscles most related to face aging and build a compositional face model(see Fig. 1(a)). 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 13 subregions(intersections exist on boundaries of some neighboring subregions), as are shown in Fig. 1. Hair is not included in this model.

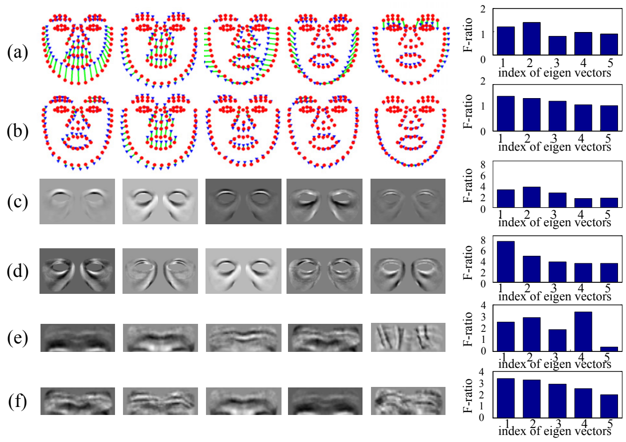



Figure 2. Comparison between the original AAM model and aging specific AAM model. (a) and (b) respectively visualizes the top five principal components of face shape and their correlation with face aging in the original AAM model and aging specific AAM model. Here the line segments illustrate the accounted shape variations. Similarly, (c)(d) and (e)(f) respectively compare the shape-free texture of around-eye region and forehead region in two models. The results in this figure are learned from 500 African-American males in the MORPH database.

Thus a face image is composed of a set of sub images:

$$I = \bigoplus_{r=1}^{R} I^r \tag{1}$$

Here $R$ is the number of facial subregions, and $\bigoplus$ denotes the composition of subregions, with blending along the subregion boundaries to avoid the image artifacts.

### 2.2. Aging specific AAM model

In this paper we extend the original AAM[9] model for aging modeling. The extended model includes a global active shape model and a shape-free texture model for each subregion separately. Since the principal components of the original AAM model describe the variations in the whole face dataset, embedding both age related and non-age related variations, we analyze the components statistically to pick out the components significantly related with aging.

Firstly the face images in training data are grouped into five equally divided age levels, and we denote their projection coefficients on a specific principal component as $w_{i,t}$, with $i$ being the index of samples and $t$ being the age level of $i$th sample. Since the plots tell us that $w_{i,t}$ is near Gaussian, we conduct widely used ANalysis Of VAriance(ANOVA)[16] on these coefficients. One of the most important statistics in ANOVA is $F$-ratio, we use which to measure the correlation between the principal component and face aging:

$$F\text{-ratio} = \frac{\frac{1}{N_T-1}\sum_{t=1}^{N_T} n_t(\overline{w_t} - \overline{\overline{w}})^2}{\frac{1}{N_S-N_T}\sum_{i=1}^{N_S}(w_{i,t} - \overline{w_t})^2} \tag{2}$$

Here $N_T$ and $N_S$ are respectively the number of age levels and training samples. $\overline{w_t}$ is the mean coefficients of all $n_t$ samples at age level $t$ and $\overline{\overline{w}}$ is the mean of all $\overline{w_t}s$.

**Algorithm 1** Learn a short term aging pattern of region $r$

**Input:**
$\mathbf{W}$: projection coefficients of a real aging sequence $(\mathbf{w}^r_{1,t_1}, \mathbf{w}^r_{2,t_2}, \cdots, \mathbf{w}^r_{n,t_n})$
$\boldsymbol{\mu}^r, \boldsymbol{\sigma}^r$: mean and standard deviation of $\mathbf{w}^r$'s gradient

1. Initialize *candidate-models := nil*
2. **for** loop = 1 **to** $C_n^2$
   2.1 *maybe-inliers* := two values randomly selected from $\mathbf{W}$ and with different age labels
   2.2 *maybe-model* := linear model fitted to *maybe-inliers*
   2.3 *consensus-set* := All the $\mathbf{w}^r_{i,t_i}$ fitting the model under error threshold $5 \cdot [\boldsymbol{\mu}^r - 3\boldsymbol{\sigma}^r, \boldsymbol{\mu}^r + 3\boldsymbol{\sigma}^r]$
   2.4 *better-model* := B-spine model fitted to *consensus-set*
   2.5 Add *better-model* to *candidate-models*
3. Output the model with minimum prediction error for all data in $\mathbf{W}$

With $w_{i,t}$ being near Gaussian, $F$-ratio follows F-distribution with degrees of freedom $N_T - 1$ and $N_S - N_T$. Setting the significance threshold at $p < 0.05$, we preserve the components significantly related with face aging and rearrange them according to their correlation with face aging in descending order. We name the model after rearrangement as *aging specific AAM model*, which is quite different from the original AAM model, as is shown in Fig. 2. In this model a specific sub image of $r$th region $I^r$ is described as:

$$I^r = I^r_m + \sum_{pc=1}^{N^r} w^r_{\text{pc}} \cdot v^r_{\text{pc}} + I^r_{\text{res}}, \tag{3}$$

where $I^r_m$, $v^r$ and $w^r$ are mean image, principal components and corresponding projection coefficients respectively. $N^r$ is the number of preserved components, and $I^r_{\text{res}}$ is the residue image containing mostly non-age related variations.

## 3. Short term aging pattern

Given a real image sequence of $r$th region $\{I^r_{i,t_i}, i = 1, 2, \cdots, m\}$, here $I^r_{i,t_i}$ is the $i$th face image at age $t_i$ and its AAM parameters are denoted as $\mathbf{w}^r_{i,t_i} = \{w^r_{i,t_i,\text{pc}}, \text{pc} = 1, 2, \cdots, N^r\}$, we target to learn a short term aging pattern over $\tau = [t_1, t_m]$, which is composed of the coefficient evolutions on all of the age related principle components: $\mathbf{p}^r_\tau = \{p^r_{\tau,\text{pc}}, \text{pc} = 1, 2, \cdots, N^r\}$.

In some extreme cases, non-age related variations are highly mixed with those age related ones and can not be excluded by the aging specific face model(e.g., expression lines are quite similar to wrinkles in appearance), we treat these data as outliers and use RANdom SAmple Consensus(RANSAC)[11] method to drop out them, as the circled star marks illustrate in Fig. 3. The steps of learning short time aging patterns using RANSAC are shown in Algorithm 1. The statistical analysis tells us that for a specific facial region $r$, the gradients of projection coefficients on the principal component with index pc follow Gaussian distribution $N(\mu^r_{\text{pc}}, (\sigma^r_{\text{pc}})^2)$ and age percep-

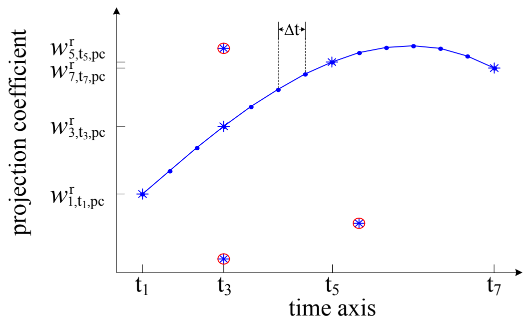



Figure 3. Modeling of short term face aging. The star marks denote the parameters of a given image sequence, among them the circled ones are outliers excluded by RANSAC method. The solid curve represents the learned aging pattern, with filled circles being interpolated parameters at time steps of $\Delta t$.

tion has an ambiguity of around 5 years, thus for a pre-trained aging model we assume that the samples not within $5[\mu^r_{\text{pc}} - 3\sigma^r_{\text{pc}}, \mu^r_{\text{pc}} + 3\sigma^r_{\text{pc}}]$ of its prediction are outliers in step 2.4 of Algorithm 1, where $\boldsymbol{\mu}^r = \{\mu^r_{\text{pc}}, \text{pc} = 1, 2, \cdots, N^r\}$ and $\boldsymbol{\sigma}^r = \{\sigma^r_{\text{pc}}, \text{pc} = 1, 2, \cdots, N^r\}$.

In implementation $p^r_{\tau,\text{pc}}$ is defined as a B-spline curve approximating the inliers of $\{w^r_{i,t_i,\text{pc}}\}$(denoted by uncircled star marks in Fig. 3) with interpolation step $\Delta t$ being 1 year.

# 4. Long term aging pattern

As is 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 model in the granularity of age spans, based on which we can infer a sequence of overlapping short term aging patterns in latter age spans from the aging pattern in current age span. After the inference, we concatenate the short term aging patters into a smooth long term aging pattern.

## 4.1. Concatenating strategy

Suppose $\mathbf{p}^r_{\tau_1}$ and $\mathbf{p}^r_{\tau_2}$ are two learned short term aging patterns of $r$th region over two overlapping time spans $\tau_1$ and $\tau_2$, we denote that $\mathbf{C}^r_1 = \{C^r_{1,\text{pc}}, \text{pc} = 1, 2, \cdots, N^r\}$ and $\mathbf{C}^r_2 = \{C^r_{2,\text{pc}}, \text{pc} = 1, 2, \cdots, N^r\}$ are respectively the fitted curves of $\mathbf{p}^r_{\tau_1}$ and $\mathbf{p}^r_{\tau_2}$. The distance between two sets of curves is defined as:

$$\begin{aligned} & \mathcal{D}(\mathbf{C}^r_1, \mathbf{C}^r_2) \qquad (4) \\ = \quad & \frac{1}{|\tau_1 \cap \tau_2|} \int_{\tau_1 \cap \tau_2} \left( \sum_{pc=1}^{N^r} F_{\text{pc}} \cdot \left| \nabla C^r_{1,\text{pc}}(t) - \nabla C^r_{2,\text{pc}}(t) \right| \right) dt \end{aligned}$$

In this equation, $F_{\text{pc}}$ is the $F$-ratio of the principal component with index pc. 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 the property of smoothness in face aging, we favor the neighboring aging patterns with similar aging trajectories in the overlapping time span. The probability of concatenating curves $\mathbf{C}^r_1$ and $\mathbf{C}^r_2$ into a long term aging pattern is computed as:

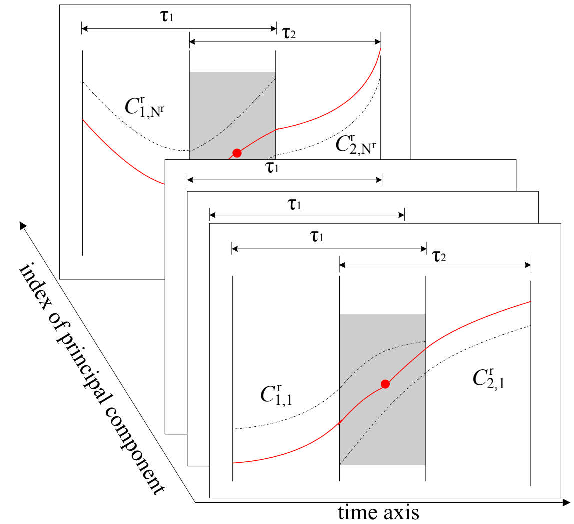



Figure 4. Illustration of concatenating short term aging patterns into a long term pattern. The dash-dotted curves $\{C^r_{1,\text{pc}}\}$ and $\{C^r_{2,\text{pc}}\}$ represent two partially overlapping(shadowed area) short term patterns of region $r$. The set of solid curves is the concatenated long term aging pattern, which shifts the curves vertically and concatenates them at the time point with the minimum difference in slopes, as are denoted by the filled circles.

$$P(\mathbf{C}^r_2|\mathbf{C}^r_1) \propto exp^{-\mathcal{D}(\mathbf{C}^r_1, \mathbf{C}^r_2)} \qquad (5)$$

After sampling two sequential short term aging patterns from the probability in Eq. 5, we try to concatenate them smoothly along the time axis. In this paper we concatenate two overlapping patterns at the time point with the minimum gradient difference summed over all the principal components, as is shown in Fig. 4

$$t^* = \arg \min_{t \in \tau_1 \cap \tau_2} \sum_{pc=1}^{N^r} F_{\text{pc}} \cdot \left| \nabla C^r_{1,\text{pc}}(t) - \nabla C^r_{2,\text{pc}}(t) \right| \qquad (6)$$

## 4.2. Diversity of face aging

By analyzing the short term aging patterns in Sec. 3, we find that there exist certain differences in the slopes of curves describing them, and this diversity shows the intrinsic variability in both aging pattern and aging velocity. The variability may be effects of various external factors(e.g., nutrition, lifestyles, health, et al.) or genes. To model the diversity of face aging, we sample different sequential patterns from the probability function defined in Eq. 5 when inferring the aging paths in latter periods. For an input face, out algorithm is able to synthesize a set of plausible aging results, as are shown in Fig. 5. It is worthing noting that facial hair is indicative of the face ages, thus its appearance changes are also learned by the proposed aging model.

# 5. Experimental results

## 5.1. The face database

Two widely used face aging databases are FG-NET[1] and MORPH[33]. We select the MORPH database in our experiments for two reasons: (i) images are collected under

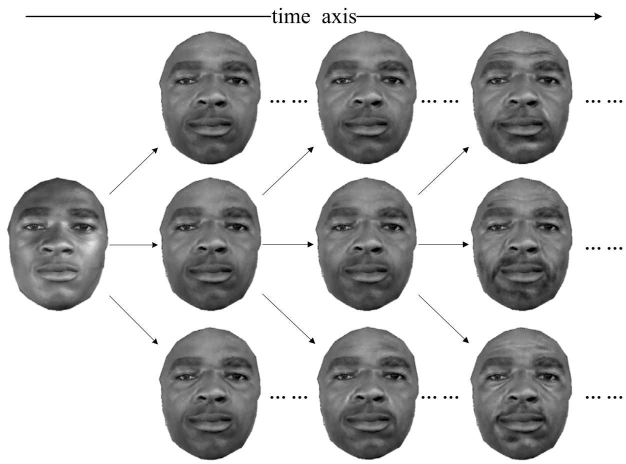



Figure 5. For the input image on the leftmost column, we synthesize a set of results at older ages to reflect the diversity of aging.

controlled environment, and non-age related variations are relatively small; (ii) the images are from a large number of individuals, this is one necessary condition for the feasibility of the proposed learning strategy.

The MORPH database was firstly released in FG'06[33], one of its extended versions includes 16,894 face images from 4,664 adults and about 85.2% of these images are from male individuals. There are in all 13,201 images of African-Americans, 3,634 of Caucasians and 59 of other races. The average age is 40.28 years old and the oldest one is 99 years old. To learn the long term aging model, we use images between 18 and 54 years and select the sequences with photos snapped at no less than three different ages. The maximum age span and mean age span in training set 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. On each face image, we labeled 83 landmarks for model training.

### 5.2. Aging simulation

Based on their ethnic group and gender, the subjects in the MORPH dataset 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 group separately. For an input face image, we perform aging simulation using the aging model of his/her ethnic and gender group. Fig. 6 shows eight synthetic sequences, from which one can see the large diversity in aging of different groups. There exist apparent artifacts in the forehead region of Caucasian females, this is caused by insufficient aging sequences for model learning and heavy hair occlusions in training data. We also synthesize a set of results for one individual to illustrate the diversity of face aging, as are shown in Fig. 5.

### 5.3. Aging model evaluation

Evaluation of face aging model is not straightforward due to some unique characteristics of face aging process: firstly, real aging sequences are difficult to collect and the collected data often include non-age related variations; secondly, aging result is intrinsically uncertain, direct comparison between the synthetic faces and real aging results does not make sense. These challenges motivate researchers to explore efforts on studies of aging model evaluation[20].

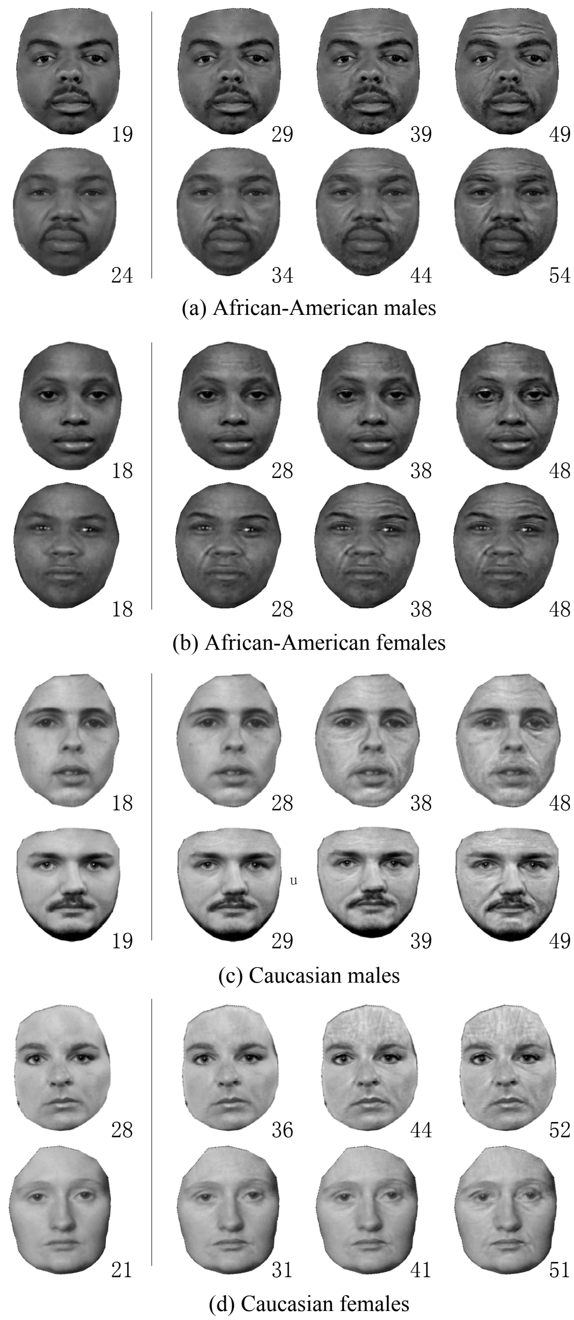



Figure 6. Some simulation results synthesized by our aging model. (a)-(d) separately display two exemplar aging sequences of individuals from four different groups. In each group, the leftmost column shows the input images, and following three columns are synthesized images at latter ages, with the age span between adjacent images being about 10 years. The age of each face image is labeled on the right bottom corner.

Table 1. MAE(yrs) of subjective and algorithmic age estimation. (Note: A, C, M and F are abbreviations of Caucasian, African-American, male and female respectively.)

| | | A M | A F | C M | C F |
|---|---|---|---|---|---|
| Subjective Estimation | Real | 6.26 | 5.76 | 6.90 | 5.37 |
| | Synthetic | 5.07 | 6.48 | 6.93 | 5.80 |
| Objective Estimation | Real | 5.78 | 6.10 | 6.77 | 7.20 |
| | Synthetic | 6.30 | 5.96 | 6.71 | 6.26 |

Two widely used criteria for aging model evaluation are:

1. Having appearances of the intended age, i.e., whether the aged faces are perceived to be of the intended ages.
2. Preservation of face identity, i.e., whether the synthetic faces are still recognized as the original person.

Guided by these two criteria, we conduct both subjective and algorithmic evaluations in the following experiments. (20 volunteers are recruited for subjective evaluation.)

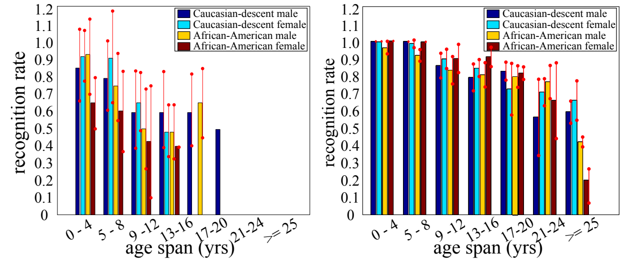
**I. Age estimation test**

We select 20 individuals(not included in training data) from each group and synthesize 5 images across 3-4 decades for each individual, these 400 synthetic images compose set A. For comparative study, we randomly select the same number of real images at different ages as set B. Both subjective and algorithmic age estimation are conducted on these two sets. In this paper, we adopt Mean Absolute Error(MAE) as a quantitative measurement for the accuracy of age estimation.

*Experiment 1) Subjective age estimation.* In this experiment, we compare the accuracies of subjective age estimation on real data and synthetic data. After a training procedure with 100 faces at different ages and selected from the training data, the volunteers are asked to estimate the face ages of data in set A and B. The MAEs on both sets are around 5-7 years, as are shown in 2∼3 rows of Table 1.

*Experiment 2) Algorithmic age estimation.* Automatic face aging estimation has been studied for years, and various algorithms were developed. Here we adopt an age estimation algorithm[39] with the-state-of-art performance to compare the accuracies of objective age estimation on real faces and synthetic faces. In this experiment age estimation is performed on all the face images in training data and 400 synthetic images in set A using the Leave-One-Person-Out approach. The estimation result is shown in parallel with the result of subjective estimation in Table 1, from which we can see that objective estimation obtains a performance comparable with that of subjective estimation, with the estimation error being around 6 years. The maximum MAE is 7.20 years, which is tested on Caucasian females, and the large estimation error may be due to a much smaller dataset of Caucasian females than that of the other groups.

From the MAEs in Table 1, we can compute that the average difference between subjective estimations on real and synthetic data is about 0.593 years and for objective estimation the difference is about 0.415 years. The small difference shows 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 intended age.



(a) performance on real data (b) performance on synthetic data

Figure 7. Performances of face recognition by human beings.

**II. Identity preservation experiment**

*Experiment 1) Subjective face recognition across ages.* Due to the limitations of human memory, only 20 individuals from each group are used for both real aging data and synthetic data. To test the face recognition rate on synthetic aging results, we mix the 80 young faces selected for synthesizing images in set A together with 80 distractors and denote them collectively as set C, from which volunteers are asked to identify the faces in set A. For real aging data, we select 20 aging sequences for each group from the MORPH database. Their photos at the initial ages and 80 distractors together compose set D, and the photos at latter ages are denoted as set E. Similarly, we ask the volunteers to recognize the faces in set E, with set D being the gallery set.

The recognition results on four groups are plotted in Fig. 7, which shows that the recognition rate generally decreases with time span increasing on both sets. while there exist some jumps due to the limitation of the test set size. This consists with the intuition that face recognition across longer periods is more difficult. Another point worth noting is that the recognition rate shows a cross-race effect and certain individual differences in face recognition abilities, as are illustrated by the error bars in Fig. 7. The among-individual variance is larger on real data due to the effects from non-age related variations.

Comparing Fig. 7(a) and (b) one can clearly see that the recognition performance on synthetic data is higher. Maybe this is because that real aging sequences include non-age related variations and the recognition rate is affected to some extent. The high recognition result shows that face identity is preserved effectively in our simulation results.

*Experiment 2) Algorithmic face recognition across ages.* Since large appearance changes occur in a long period, so far there is still no algorithm effective for face recognition over long time spans. In this experiment we test the recognition rates on set A and all the faces in training dataset using a face recognition algorithm[37] developed on Gabor features and LDA-based classifier. To keep the evaluation objective, the recognition algorithm is trained on an absolutely separate database.

Fig. 8(a) shows the recognition results on real aging data. Although there exist slight differences among recognition

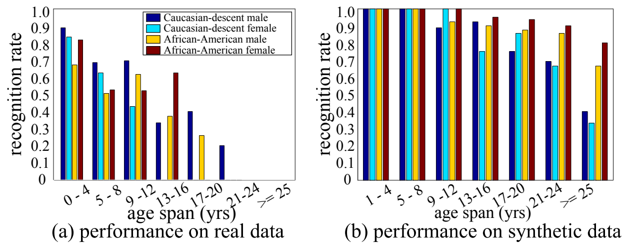



Figure 8. Performances of algorithmic face recognition. The legend omitted in (b) is the same with in (a).

rates on four groups, which can be caused by the unbalanced training data size, the trend of recognition rate with time span increasing is generally consistent: recognition performance degenerates with the increase of time span.

From the recognition result on synthetic dataset plotted in Fig. 8(b), one can see that performance also decreases as age span increases. The recognition rate is much higher than on real aging sequences, because recognition on real aging sequences is affected by external factors(e.g., pose, expression, light, et al.), whereas the synthetic faces contain only age related variations. Another point different from results on real aging sequences is that the recognition rate on African-Americans is relatively higher, maybe this is due to the fact that more age related variations occur on Caucasians faces than on African-Americans. The high recognition rate of objective face recognition gives the same conclusion as the subjective recognition: the proposed aging model retains identity information in a large extent.

**III. Analysis on the evaluation experiments**

The above experiment results indicate that subjective evaluation is comparatively more accurate, because human beings are good at age estimation and recognizing faces after millions of years of evolution. But human experiment has some drawbacks: firstly, the experimental period is usually long; secondly, subjective evaluation is limited to small test sets and causes the results less accurate; thirdly, for both age estimation and face recognition, cross-race effect is a problem worth further consideration.

In contrast, algorithmic evaluation is not so time consuming and is less limited by the test set size, the cross-race effect can also be eliminated with balanced training data, so development of age estimation and face recognition algorithms will advance the studies of face aging model evaluation. Face recognition robust to age related variations is a challenging problem and age estimation algorithms applicable for images with non-age related variations(e.g., image resolution, illumination, et al.) are still under study.

## 6. Conclusions and future work

This paper proposes a face aging model extracting long term face aging patterns from partially dense aging datasets. The work in this paper validates the feasibility of learning a long term face aging model from currently available databases under a Markov assumption and smoothness constraints. The studies in this paper also reveal that large diversity exists among face aging of individuals from different races and of different genders, and building group-specific face aging models is indeed necessary.

Our future work will focus on introducing more knowledges of anatomy to build more accurate face aging model and studying face recognition across ages.

## 7. Acknowledgment

This paper is partially supported by Natural Science Foundation of China under contract Nos. 60803084, 60872124, 60728203, and 60872077; National Basic Research Program of China (973 Program) under contract 2009CB320902; Grand Program of International S&T Cooperation of Zhejiang Province S&T Department (No. 2008C14063), and ISVISION Technology Co. Ltd..

## References

[1] FG-NET aging database, online available in http://sting.cycollege.ac.cy/alanitis/fgnetaging/. 1, 4

[2] Y. Bando, T. Kuratate, and T. Nishita. A simple method for modeling wrinkles on human skin. In *Proc. Pacific Conf. Computer Graphics and Applications*, pages 166–175, 2002. 1

[3] A. C. Berg and S. C. Justo. Aging of orbicularis muscle in virtual human faces. In *Proc. Int'l Conf. Information Visualization*, pages 164–168, 2003. 1

[4] A. C. Berg, F. J. P. Lopez, and M. Gonzalez. A facial aging simulation method using flaccidity deformation criteria. In *Proc. Int'l Conf. Information Visualization*, pages 791–796, 2006. 1

[5] S. Biswas, G. Aggarwal, N. Ramanathan, and R. Chellappa. A non-generative approach for face recognition across aging. In *Proc. Int'l Conf. on Biometrics: Theory, Applications and Systems*, 2008. 1

[6] L. Boissieux, G. Kiss, N. M. Thalmann, and P. Kalra. Simulation of skin aging and wrinkles with cosmetics insight. In *Proc. Eurographics Workshop on Animation Computer Animation and Simulation*, pages 15–27, 2000. 1

[7] 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*, 259:137–143. 1

[8] J. W. Carls, R. Raines1, M. Grimaila, and S. Rogers. Biometric enhancements: Template aging error score analysis. In *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, 2008. 1

[9] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(6):681–685, Jun. 2001. 2, 3

[10] P. Ekman and E. L. Rosenberg. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System(FACS)*. Oxford University Press, 2nd edition, 2005. 2

[11] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. 3

[12] Y. Fu and N. Zheng. M-face: An appearance-based photorealistic model for multiple facial attributes rendering. *IEEE Trans. Circuits and Systems for Video Technology*, 16(7):830–842, Jul. 2006. 1

[13] M. Gandhi. *A method for automatic synthesis of aged human facial images*. PhD thesis, McGill University, 2004. 1

[14] A. Golovinskiy, W. Matusik, H. Pfister, S. Rusinkiewicz, and T. Funkhouser. A statistical model for synthesis of detailed facial geometry. In *ACM Trans. on Graphics*, pages 1025–1034, 2006. 1

[15] C. M. Hill, C. J. Solomon, and S. J. Gibson. Aging the human face-a statistically rigorous approach. In *Proc. IEE Symposium on Imaging for Crime Detection and Prevention*, pages 89–94, 2005. 1

[16] D. C. Howell. *Statistical methods for psychology*. Belmont, CA: Duxbury Press, 6th edition, 2007. 3

[17] 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*, 22(6):747–753, Jun. 2003. 1

[18] F. Jiang and Y. Wang. Facial aging simulation based on super resolution in tensor space. In *Proc. Int'l Conf. Image Processing*, pages 1648–1651, 2008. 1

[19] A. Lanitis. Comparative evaluation of automatic age-progression methodologies. *EURASIP Journal on Advances in Signal Processing*, 8(2):1–10, Jan. 2008. 1

[20] A. Lanitis. Evaluating the performance of face-aging algorithms. In *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, 2008. 1, 6

[21] 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*, 24(4):442–455, Apr. 2002. 1, 2

[22] W. S. Lee, Y. Wu, and N. M. Thalmann. Cloning and aging in a VR family. In *Proc. Virtual Reality*, pages 61–68, 1999. 1

[23] F. R. Leta, A. Conci, D. Pamplona, and I. Itanguy. Manipulating facial appearance through age parameters. In *Ninth Brazilian Symposium on Computer Graphics and Image Processing*, pages 167–172, 1996. 1

[24] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs. Study of face recognition as people age. In *Proc. Int'l Conf. Computer Vision*, pages 1–8, 2007. 1

[25] S. Mukaida, H. Ando, K. Kinoshita, M. Kamachi, and K. Chihara. Facial image synthesis using age manipulation based on statistical feature extraction. In *Proc. Int'l Conf. Visualization, Imaging, and Image Processing*, pages 12–17, 2002. 1

[26] U. Park, Y. Tong, and A. K. Jain. Face recognition with temporal invariance: A 3D aging model. In *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, 2008. 1

[27] E. Patterson, K. Ricanek, M. Albert, and E. Boone. Automatic representation of adult aging in facial images. In *Proc. IASTED Int'l Conf. Visualization, Imaging, and Image rocessing*, page 612, 2006. 1

[28] J. B. Pittenger and R. E. Shaw. Aging faces as viscal-elastic events: implications for a theory of nonrigid shape perception. *Journal of Experimental Psychology:Human Perception and Performance*, 1(4):374–382, Nov. 1975. 1

[29] A. N. R. Singh, M. Vatsa and S. K. Singh. Age transformation for improving face recognition. In *Proc. Int'l Conf. Pattern Recognition and Machine Intelligence*, pages 576–583, 2007. 1

[30] N. Ramanathan and R. Chellappa. Face verification across age progression. *IEEE Trans. Image Processing*, 15(11):3349–3361, 2006. 1

[31] N. Ramanathan and R. Chellappa. Modeling age progression in young faces. In *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pages 387–394, 2006. 2

[32] N. Ramanathan and R. Chellappa. Modeling shape and textural variations in aging faces. In *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, 2008. 1, 2

[33] K. Ricanek, Jr., and T. Tesafaye. Morph: a longitudinal image database of normal adult age-progression. In *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, pages 341–345, 2006. 1, 4, 5

[34] D. A. Rowland and D. I. Perrett. Manipulating facial appearance through shape and color. *IEEE Trans. Computer Graphics and Applications*, 1995. 1

[35] C. M. Scandrett, C. J. Solomon, and S. J. Gibson. A person-specific, rigorous aging model of the human face. *Pattern Recognition Letters*, 27(15):1776–1787, Nov. 2006. 1

[36] K. Scherbaum, M. Sunkel, H. P. Seidel, and V. Blanz. Prediction of individual non-linear aging trajectories of faces. In *Proc. Computer Graphics Forum*, volume 26, pages 285–294, 2007. 1

[37] Y. Su, S. Shan, X. Chen, and W. Gao. Hierarchical ensemble of global and local classifiers for face recognition. In *Proc. Int'l Conf. Computer Vision*, pages 1–8, 2007. 6

[38] J. Suo, F. Min, S. C. Zhu, S. Shan, and X. Chen. A multi-resolution dynamic model for face aging simulation. In *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pages 1–8, 2007. 2

[39] J. Suo, T. Wu, S. C. Zhu, S. Shan, X. Chen, and W. Gao. Design sparse features for age estimation using hierarchical face model. In *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, 2008. 6

[40] B. P. Tiddeman, D. M. Burt, and D. I. Perrett. Prototyping and transforming facial textures for perception research. *IEEE Trans. Computer Graphics and Applications*, pages 42–50, 2001. 1

[41] B. P. Tiddeman, M. R. Stirrat, and D. I. Perrett. Towards realism in facial prototyping: results of a wavelet mrf method. In *Proc. Theory and Practice of Computer Graphics*, pages 103–111, 2006. 1

[42] J. Wang and C. Ling. Artificial aging of faces by support vector machines. In *Proc. Canadian Conference on Artificial Intelligence*, pages 499–503, 2004. 1

[43] J. Wang, Y. Shang, G. Su, and X. Lin. Age simulation for face recognition. In *Proc. Int'l Conf. Pattern Recognition*, pages 913–916, 2006. 1, 2

[44] Y. Wu, P. Beylot, and N. M. Thalmann. Skin aging estimation by facial simulation. In *Proc. Int'l Conf. Computer Animation*, pages 210–219, 1999. 1