CrossMark

Age progression by gender-specific 3D aging model

Sidra Riaz[1] · Unsang Park[1] · Jongmoo Choi[2] · Prem Natarajan[3]

Abstract

Facial aging is an important problem of face recognition in missing children and automatic template update. As aging is a temporal process, it alters the facial appearance of the individuals. The sources of variations in facial appearance are caused by wrinkles (under eyes, forehead, around lips, and jawline), facial growth (cranial size and skull), and skin tone. The other factors such as health, lifestyle, and gender also impose variations in the aging process. Therefore, predicting facial aging with considering all those factors is a very difficult task. We present our 3D gender-specific aging model which automatically produces simulated images at age y by taking only one input image at age x irrespective of the pose and lighting conditions. The gender-specific aging model is constructed by various datasets (FG-NET, PCSO, Celebrities, BROWNS, Private), and its quality is evaluated with respect to various combinations of the datasets. We further fine-tune the aging model by changing the length of shape and texture eigenvectors and examine how these parameters affect the simulation results. Comparisons of the simulation results with state-of-the-art approaches as well as ground truth images demonstrate the effectiveness of the proposed methods. The subjective and objective evaluations are also carried out which emphasize the potential of our proposed gender-specific 3D aging model.

1 Introduction

Aging process seeks to automatically simulate a photograph of a person from age x to age y as shown in Fig. 1. The aging simulation process is a very challenging task because of the non-deterministic nature of aging for every individual. From input photographs, it is difficult to determine the genetics, health conditions, lifestyle, and environment which makes a huge difference in the aging process. To build an aging model which may replicate the real aging process (keeping the other factors such as health, lifestyle, and environment constant), we need a large facial aging image dataset that spans over human lifetime and contains images at various ages. However, there are only a few public domain datasets (e.g., FG-NET and MORPH) that are relatively small with limited age range, poor quality and lighting conditions, and of low resolution. In addition to private datasets, these publically available datasets have been constantly in use by researchers [1–5] to construct discriminative [3,6–9] or generative aging models [1,2,10–14].

The simulated face images help to find the missing children and people in blacklists for law enforcements. Also, people can see how their children will look like when they grow up and how their parents used to look like in their childhood. A number of studies have been performed so far in the field of automatic aging simulation or age progression.

1.1 Related work

Related researches for age modeling or age progression [1–8,10–20] show how the aging process causes variations in

✉ Unsang Park
  unsangpark@sogang.ac.kr

  Sidra Riaz
  sidra@sogang.ac.kr

  Jongmoo Choi
  jongmooc@usc.edu

  Prem Natarajan
  pnataraj@isi.edu

1 Department of Computer Science and Engineering, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul, Korea

2 Institute for Robotics and Intelligent Systems, University of Southern California, Los Angeles, CA, USA

3 Information Science Institute, University of Southern California, Los Angeles, CA, USA

Fig. 1 Example of aging simulation results by proposed method from age $x = 0$–5 to age $y = 55$; **a** input images, **b** pose-corrected images, and **c** age-simulated images

the facial features and appearance. Facial aging simulation models can be divided into two categories, namely (i) discriminative models and (ii) generative models.

1.1.1 Discriminative models

Suo et al. [15] proposed a multi-resolution dynamic model for facial aging simulation based on a dynamic Markov process. Their aging model was learned from a large dataset of 50,000 adult facial images of high resolution that showed promising results. Ling et al. [3,7] proposed their aging process in the passport photograph verification task to show how the face recognition performance degrades in the presence of aging. The illumination and expression variations, however, are not explored in their study. Li et al. [8] presented a discriminative aging model to be used in the application of face recognition. The authors proposed densely sampled local descriptors for feature representation and multi-feature discriminant analysis. They used the distribution of intensity gradient in the face image as age invariant information. Kemelmacher et al. [6] presented an illumination aware age progression approach. They first computed an average image subspace by using collection flow method. For an input image in frontal pose, the flow differences of texture and illumination are used for aging simulation. Then, the aspect ratio is adjusted to incorporate the facial shape aging effect. The simulated facial image is

then mapped on to the original background image with keeping the aspect ratio.

1.1.2 Generative models

Lanitis et al. [10] proposed an approach for aging simulation using a statistical face model based on reversible coding. The relationship between coded representations of facial images and their associated ages is used to estimate the ages of unseen individuals and to reconstruct the appearance of an individual at any age. Singh et al. [14] presented their registration-based age transformation approach to minimizing the variations in facial features due to aging. Later, the authors proposed a facial aging model where short-term age patterns are learned from partially dense datasets (FG-NET and MORPH) [5]. The learned short-term aging patterns are concatenated to build the long-term aging patterns. Park et al. [1,2] proposed their 3D age model constructed in separate shape and texture spaces. They used FG-NET (1002 images from 82 subjects) dataset to model the aging spaces (shape and texture). The shape aging space consists of facial shapes of 82 subjects in the age range of 0–70. The texture space is of the same dimensions and formed by using grayscale texture of facial images in FG-NET dataset.

Du et al. [11] proposed a sparse-constraint nonnegative matrix factorization (NMF) approach. The NMF algorithm is updated by considering sparse coefficients (H) or the basis

Table 1 Details of the datasets used in our 3D gender-specific aging model

	FG-NET	PCSO	BROWNS	Celebrities	Private	Total
#Subjects	82	130	4	33	8	257
#Females	34	18	4	20	7	83
#Males	48	112	0	13	1	174
#Images	1002	1203	132	523	119	2979
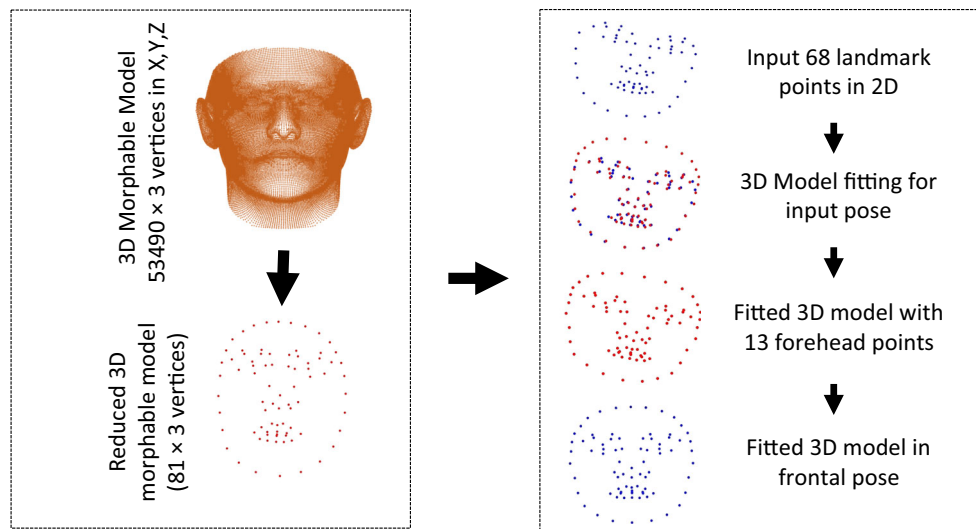
Fig. 2 A 3D model fitting in frontal pose from 2D landmarks using our reduced morphable model

vector (W) and used for better age simulation results. Wang et al. [17] proposed aging simulation approach using tensor space analysis and active appearance model (AAM). The super resolution method is adopted to improve the aging results. Yang et al. [4] used sparse representation of age and person specific facial properties for aging simulation results.

Wang et al. [18] proposed their recurrent face aging (RFA) framework which is based on recurrent neural network (RNN). The face images are normalized by optical flow method to preserve the wrinkles information. The two-layer gated recurrent unit is proposed where the top and bottom layers produce an older and younger face from the input image, respectively. Wang et al. [19] proposed an aging simulation approach based on a tensor completion technique. The author assumes that texture of different subjects in the same age group is similar and the people who have high resemblance tend to age in similar ways. Thus, during the aging simulation process, bigger weights are assigned to the samples in three-order tensor model which have higher similarities with the test image through multilinear graph embedding. To measure the similarity between subjects, a metric learning approach is adopted.

1.2 Advancements from related work

In this paper, we present a generative gender-specific 3D aging model. The proposed aging model automatically produces simulated images at various ages (0–69 years old) from an input 2D image of a person at age x. Our 3D aging model is built in a similar fashion to generic aging model by Park et al. [1]; however, we constructed gender-specific aging model with added datasets (see details in Table 1). Main components of our work are given as follows:

- Gender-specific aging model (separate male, female): We separated the subjects in FG-NET, PCSO, Celebrities, BROWNS, and Private datasets (see details in Table 1) based on gender and constructed two 3D aging models as shown in Fig. 4.
- Pose correction in 3D: We used a reduced 3D demorphable model based on Blanz and Vetter's model [21] for pose correction. The input 68 landmark points in 2D are used to perform model fitting to interpolate the 13 forehead points. The fitted 3D model is mapped to the 3D frontal pose afterward. These 3D facial images are then used for our 3D aging model construction.
- Separate shape and texture pattern space: The aging patterns (shape and texture) are learned separately in a

similar fashion of [1] with additional datasets (FG-NET, PCSO, Celebrities, BROWNS, and Private datasets).

- Aging simulation on pose-corrected images: In testing stage, the aging simulations are performed on frontal pose of the facial images after 3D model fitting and forehead interpolation.
- Color and flare corrections: After aging simulation, the color and flare corrections are performed to make the aging simulation results more realistic.
- Compositing age-simulated image to original background: Because our aging model does not age/de-age the hair, we composite the age-simulated frontal face image to the original background for comparison with ground truth reference images and state-of-the-art results.

1.3 Major contributions

Following are the major contributions of this paper.

- Study the effects of adding various datasets: In our 3D aging model, we studied the effects of adding various datasets (FG-NET, PCSO, Celebrities, BROWNS, and Private) qualitatively and quantitatively. The result of aging averages and simulation results are provided. Furthermore, we show the goodness-of-fit measure for adding each dataset in the aging model.
- Study the effects of eigenvectors: The effects of increasing or decreasing the length of shape and texture eigenvectors in simulation results are studied.
- Qualitative comparison of the aging simulations results: The results of aging simulations on FG-NET are compared with ground truth reference images at the same age. Furthermore, we compared our results with those of state-of-the-art aging simulation methods.
- Quantitative comparison (subjective and objective evaluations): A subjective survey is carried out for two evaluation metrics, such as identity matching (IM) and age matching (AM). Our proposed approach achieved higher subjective average scores. Furthermore, we performed age invariant face recognition experiments with two setups of FG-NET dataset. The comparison with state-of-the-art methods shows that our approach achieves higher Rank-1 (%) accuracy.

The rest of this paper is organized as follows: Sect. 2 introduces the proposed gender-specific 3D aging model (constructed by FG-NET, PCSO, Celebrities, BROWNS, and Private datasets) and aging simulation process. Section 3 provides the aging simulation results on FG-NET dataset. The aging simulation results are compared with ground truth reference images as well as state-of-the-art results. In Sect. 4,

we conclude this paper by setting an insight on the future work in facial age modeling.

2 Proposed gender-specific 3D aging

In this section, we describe our 3D gender-specific aging model construction and aging simulation process. We first convert 2D images in the modeling datasets to the 3D faces in frontal pose; then, those 3D face images are used to generate two aging spaces, i.e., shape and texture.

2.1 Dataset collection for aging modeling

The FG-NET dataset contains 82 subjects and 1002 images only, which is insufficient to fill the 3D aging model in shape and texture spaces as proposed by Park et al. [1]. Therefore, we take more face images, from (i) PCSO (mugshot data collected from 130 subjects), (ii) Celebrities (a collection of images from 33 female and male subjects found on the web), (iii) BROWNS (a collection of pictures of four sisters taken every year over a period of 33 years from 1975 to 2007) [22], and (iv) Private (which is a private collection of the images from 8 male and female subjects), in order to utilize various shape and texture aging patterns (0–69). In total, we compiled 174 male and 83 female subjects, which is a significantly larger number as compared to that of FG-NET dataset (48 males and 34 females) used in [1]. The details of the modeling datasets are given in Table 1.

2.2 Aging modeling and simulation revisited

In this section, we describe how age modeling in 3D domain is performed. In order to normalize the pose variations, input 2D images are mapped to 3D face models using 68 facial landmarks. Active shape model (ASM) v4.0 [23] is used to detect 68 facial landmark points and then 13 additional points for forehead are augmented using reduced 3D morphable model (Fig. 2). Example landmark detection results are shown in Fig. 3. The reduced 3D morphable model is induced from the Blanz and Vetter's model [21] and represented by the top eigenvectors ($L = 30$) as $S_\alpha = S_m + \sum_{l=1}^{L} \alpha_l V_l$, where S_m is the mean shape of 100 3D face models, V_l are unit eigenvectors obtained from the PCA analysis [1], and $\alpha_l = [\alpha_1, \alpha_2, \ldots, \alpha_L]$ is a control parameter for facial shape or coefficient that defines a specific shape instance under the given deformable model. The covariance of α_l is a diagonal matrix where $\alpha_l \epsilon \Re^L$ [24,25].

After obtaining all the 3D faces (S_{3d}) by fitting reduced morphable model to 2D input images, we apply PCA on male and female 3D faces separately ($1 \sim N$ in male and $1 \sim N'$ in female, where $N = 167$ and $N' = 68$) with available ages ($0 \sim M - 1$, where $M = 70$). The reduced shape space in

Fig. 3 Landmark detections on the celebrities and PCSO datasets using ASM v4.0, **a** example images, and **b** landmark detection results
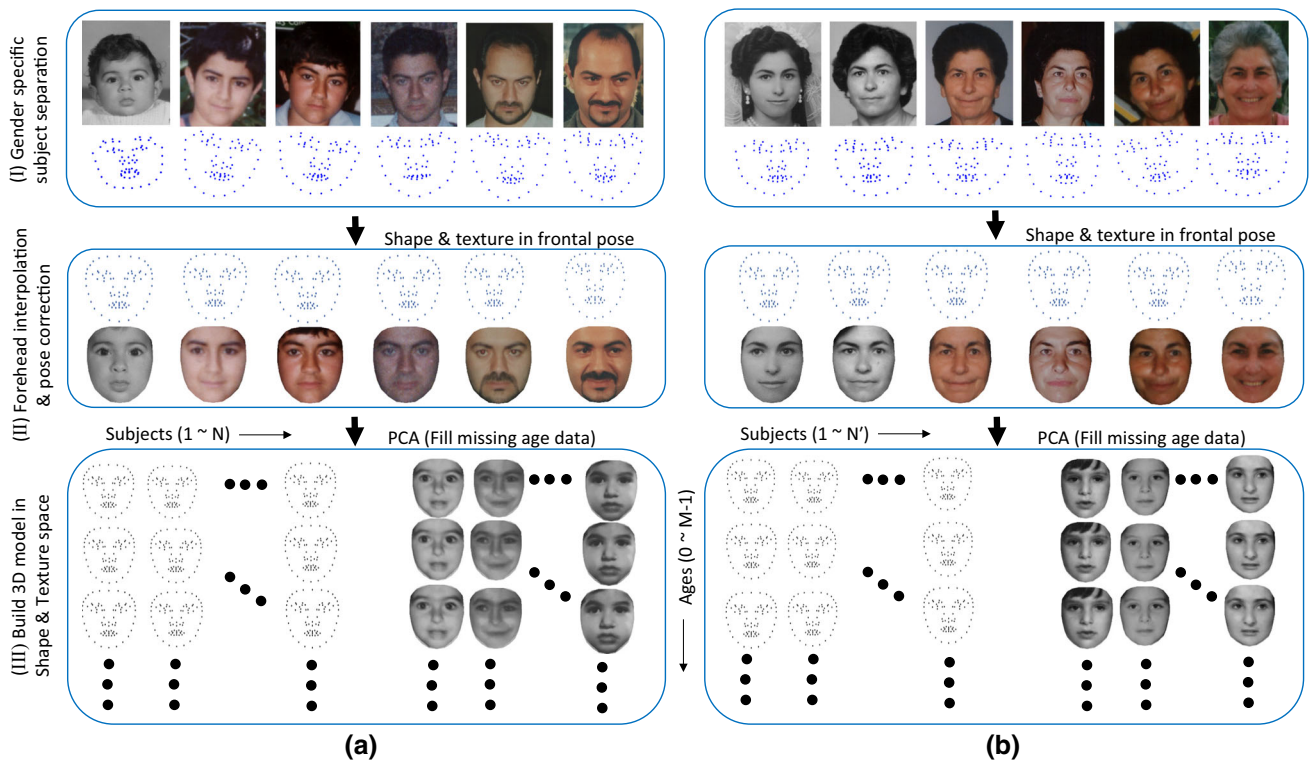

Fig. 4 Gender-specific 3D aging models separated in the shape and texture space, **a** male aging model and **b** female aging model. The subjects are separated and their 68 landmark detections are shown in (I) first block (top), the forehead interpolation and pose correction is shown in (II) second block (middle), and finally the age shape and texture space filled by linear interpolation are shown in (III) third block (bottom)

3D for N male subjects and M ages, $s_{3d}(N, M)$ is obtained as $s_{3d}(N, M) = V_s^T (S_{3d}(N, M) - S_m)$, where V_s are the top 50 principle components of S_{3d}. The female model is also similarly constructed. In the $N \times M$ shape pattern space, we fill missing patterns by using an interpolation method as $s_{3d}(n, m') = w_1 \times s_{3d}(n, m_1) + w_2 \times s_{3d}(n, m_2)$, where $s_{3d}(n, m_1)$ and $s_{3d}(n, m_2)$ are the facial shapes for subject n at ages m_1 and m_2 that are closest to the age m' and w_1 and w_2 are the weights inversely proportional to the distance from

age m' to m_1 and m_2, respectively. We construct the texture pattern space $t(N, M)$ from the original texture patterns by using PCA analysis similarly as shape pattern space.

Given the shape and texture models as shown in Fig. 4, we perform the aging simulation as follows. We first calculate the reduced shape pattern of the input image as $s_{3d}(x)$ at age x and fit it to the shape aging space as given in Eq. 1.

$$\hat{w}_s = \underset{c_- \leq w_s \leq c_+}{\operatorname{argmin}} ||s_{3d}(x) - s_{w_s}(x)||^2 + r_s ||w_s||^2, \quad (1)$$
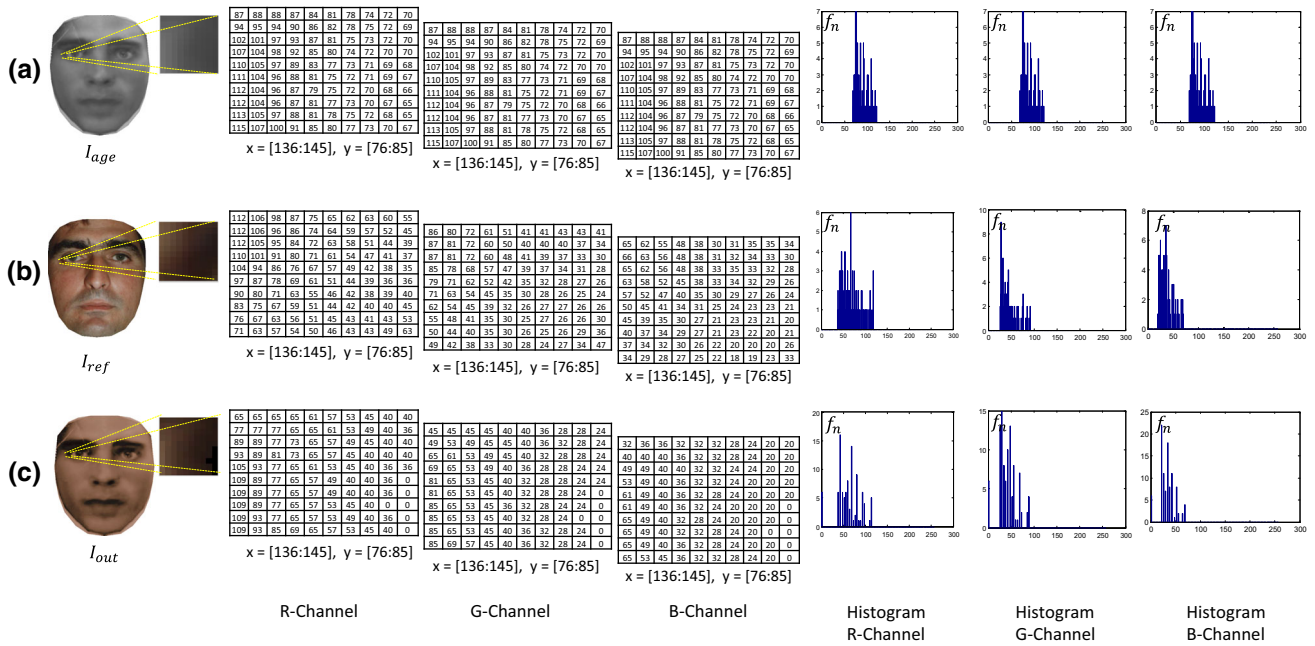
Fig. 5 Color correction of age-simulated image, **a** input age-simulated image, **b** reference image, and **c** color-corrected image (zoom in to see the values)
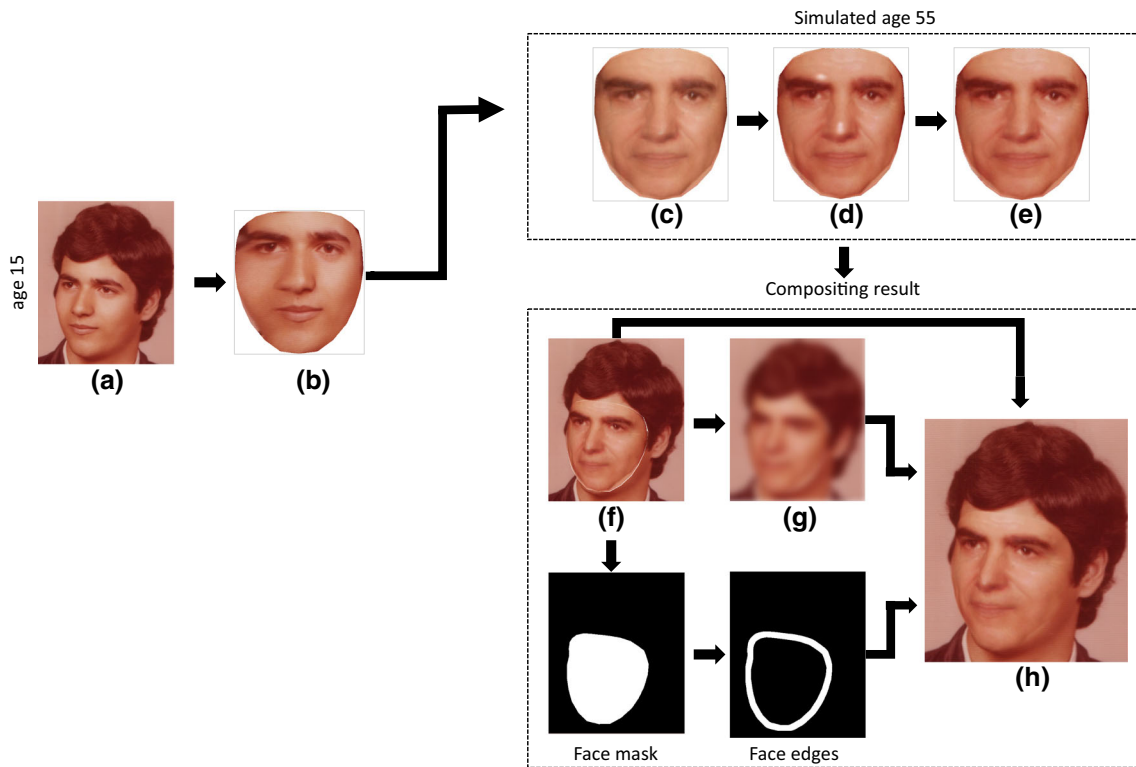

Fig. 6 Compositing age-simulated image (after color and flare correction) to the original background image, **a** input image, **b** pose-corrected image, **c** age-simulated image, **d** color-corrected image, **e** flare-corrected image, **f** compositing without ESBC, **g** soft boundary edges, and **h** composited face image with ESBC approach (zoom in to see the details)
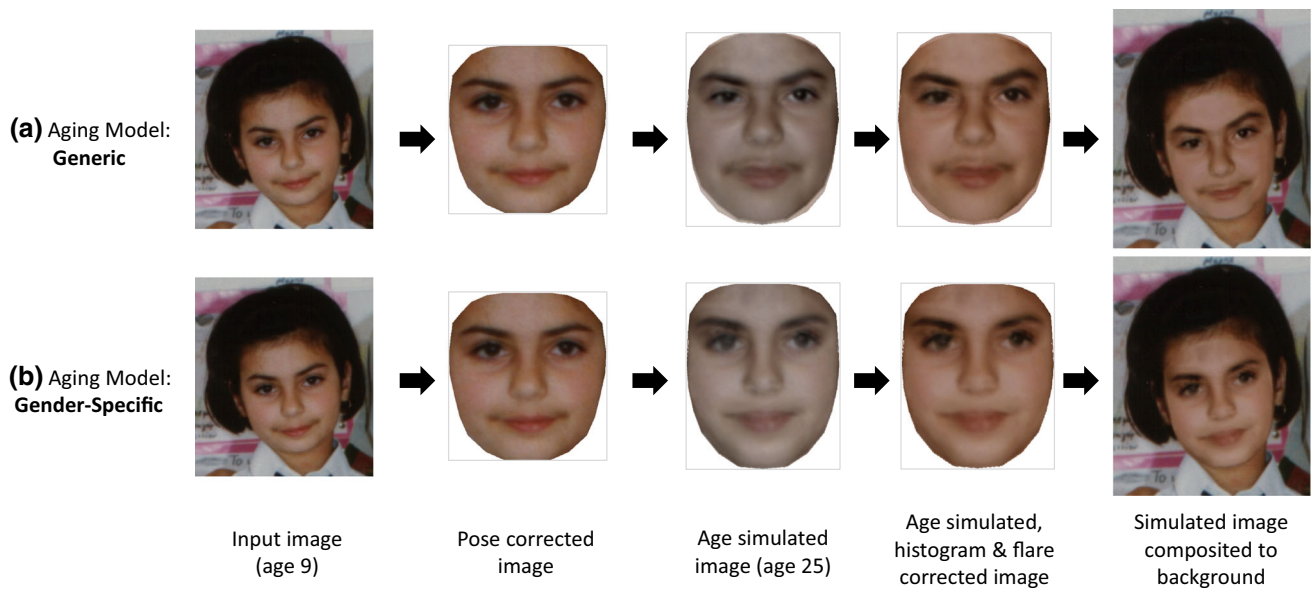
Fig. 7 Effectiveness of generic versus gender-specific aging simulation results, **a** generic aging model by Park et al. [1], and **b** gender-specific aging model (**proposed**)
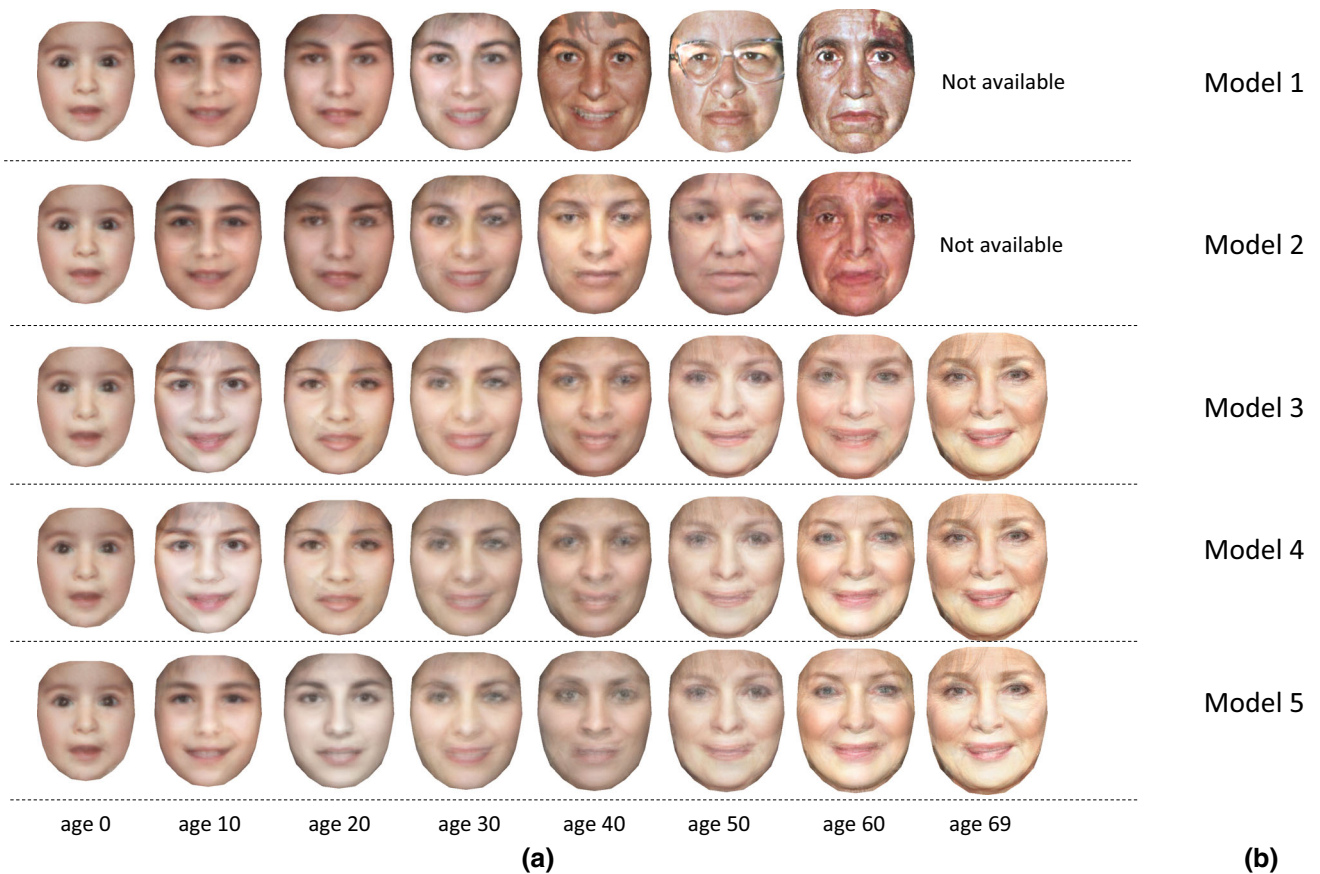

Fig. 8 Average texture space at each step of adding a new dataset in the female aging model, and **a** averages of texture space from age 0 to 69, **b** name of gender-specific aging model
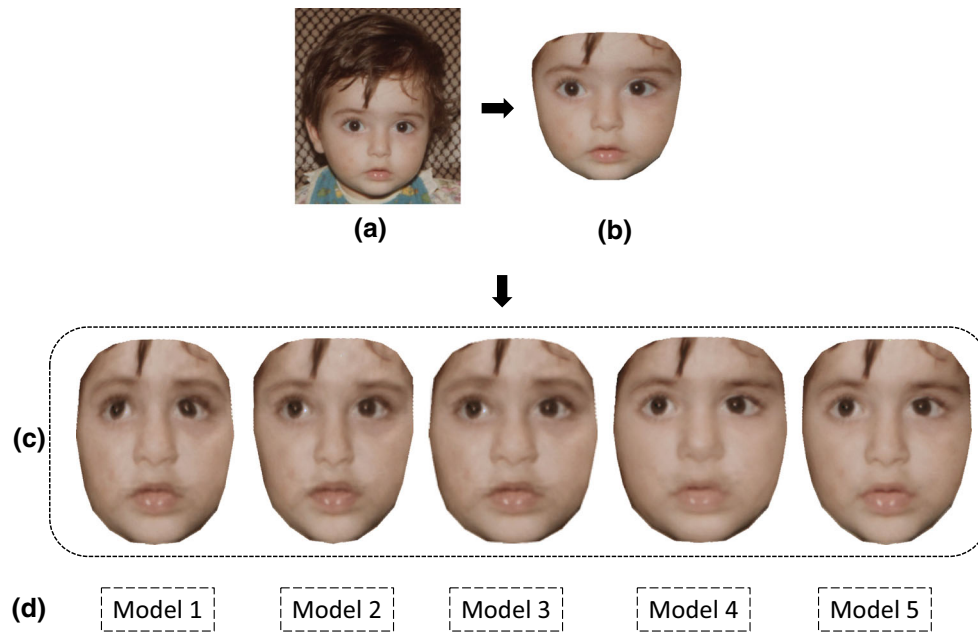
Fig. 9 Effects of adding datasets in the aging model, **a** input image, **b** pose-corrected image, **c** aging simulation results at target age (15 years old), and **d** aging model used for the simulations

Table 2 Goodness-of-fit measure for shape and texture aging models	Mean (RMS$_{\text{shape}}$) values	Mean (RMS$_{\text{texture}}$) values
Model 1	22.3465	21.6267
Model 2	21.3222	17.4303
Model 3	14.9091	18.2407
Model 4	7.3183	17.501
Model 5	**4.768**	**17.1735**

Best results are indicated with bold font

where $s_{w_s}(x) = \overline{s_{3d}(x)} + \sum_{i=1}^{N}\left(s(i,x) - \overline{s_{3d}(x)}\right)w_{s,i}$, $0 \leq j \leq M-1$, and r_s is the regularizer ($r_s = 3$) [26] to handle the case when multiple solutions of w_s are obtained. This weight, w_s, obtained at the input age x is then applied to the target age y to obtain the aging-simulated shape. The aging-simulated texture is also similarly obtained, and the combination of the shape and texture becomes the aging simulation result at age y. More details of the aging modeling and simulation processes can be found in [1].

2.3 Color and flare correction

After shape and texture aging, we perform the color correction or histogram matching so that the synthesized frontal pose image (I_{age}) can be mapped back to the original background image. The objective of histogram matching is to transform I_{age} such that the gray values are redistributed and histogram is more uniform and matched with the reference image (I_{ref}) (see Fig. 5c).

To do color correction, we used original frontal pose image as a reference image (I_{ref}) and age-simulated image in frontal pose (I_{age}) as an input image with dimensions of m_r by m_c. The value of each pixel's intensity ranges from 0 to $P_{\text{age}} - 1$ where $P_{\text{age}} = 256$ is often the maximum possible intensity value. We denote m as the number of total available pixel intensities in an image. The color correction is performed on each channel of an RGB image of I_{age} by first calculating the number of pixels for each intensity ($n = 0, 1, \ldots, P_{\text{age}} - 1$) present in each channel, frequency of intensity values, $f_n = \frac{\text{no. pixels with intensity } n}{m_r \times m_c}$, and cumulative frequencies (F_m), where $F_0 = f_0$ and $F_m = F_{m-1} + f_m$. Then, from the intensity value of each channel of I_{ref}, the range of intensity values of I_{age} is adjusted by multiplying ($P_{\text{ref}} - 1$) to the cumulative frequencies (F_m). By performing floor rounding on the intensity values, we get the corrected pixel intensities of output image, F_m^{out} as given in Eq. 2. To obtain the color-corrected image (I_{out}) which is uniformly distributed on [0, $P_{\text{ref}} - 1$], the number of pixels in I_{age} with the same intensity values is replaced with new intensity values from F_m^{out}. The number of equally spaced bins, we selected, is $N = 64$ [27]. The flare which occurs due to color correction is eliminated by removing the top fre-
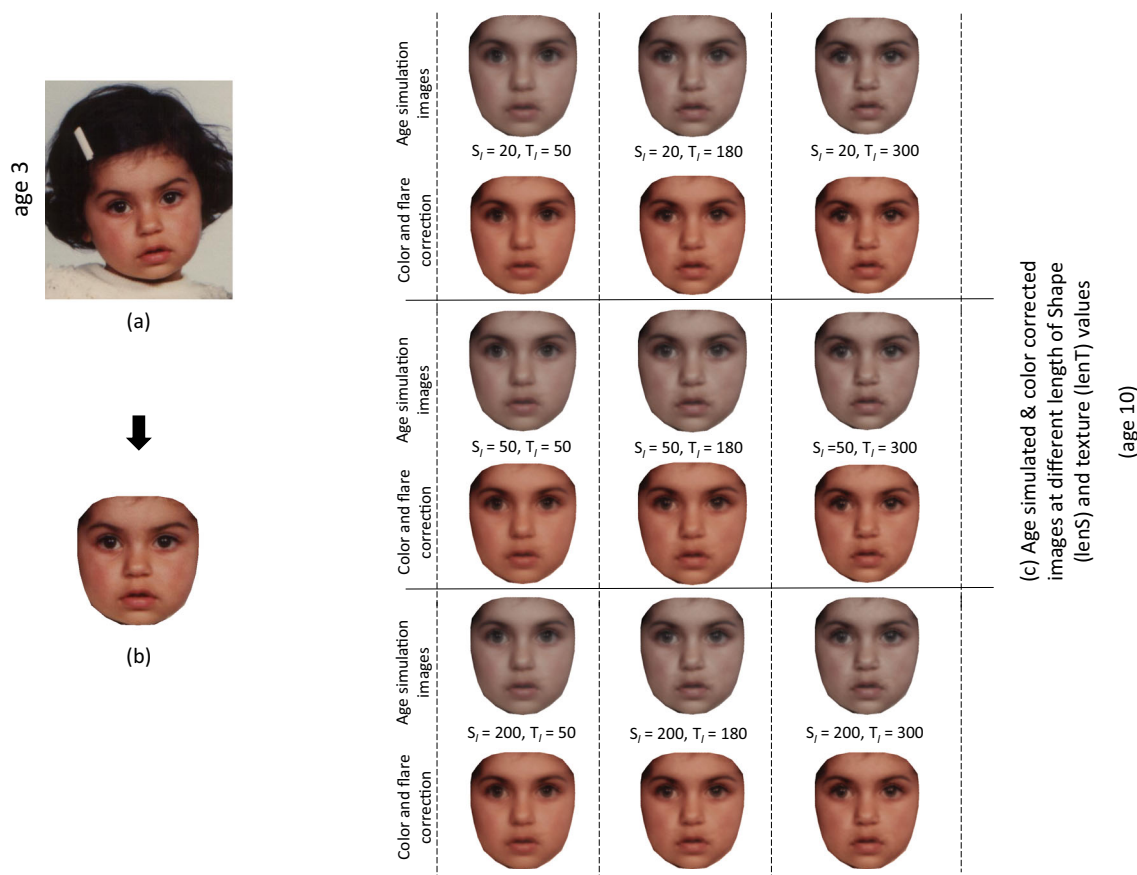
Fig. 10 Example results of using various lengths of shape and texture eigenvectors in aging simulations, **a** input image from FG-NET dataset, **b** pose-corrected image, and **c** age simulation images (age 10) at different lengths of eigenvectors of shape (S_l) and texture (T_l) space

Table 3 Goodness-of-fit measure using different shape (S_l) and texture (T_l) eigenvectors

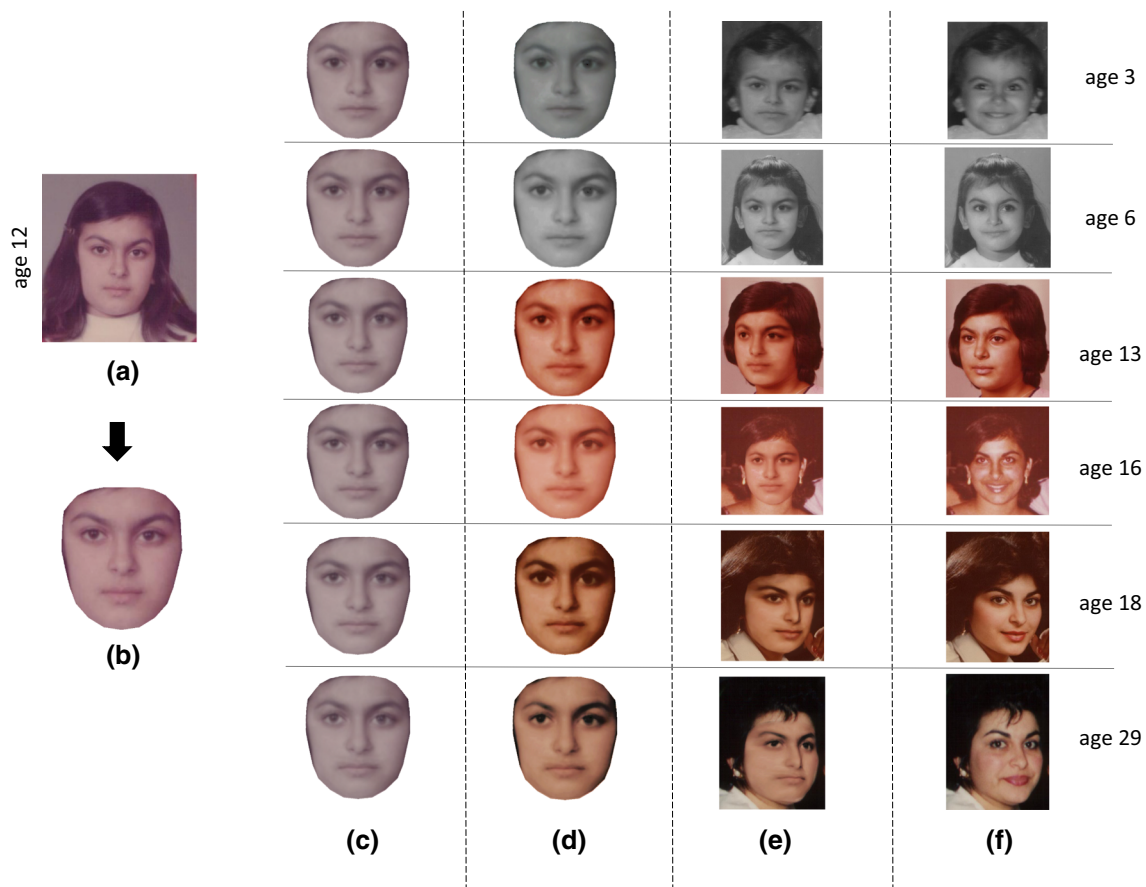|	Mean (RMS_{shape}) values	Mean ($RMS_{texture}$) values
$S_l = 20$ $T_l = 50$	11.245	11.883
$S_l = 20$ $T_l = 180$	10.998	11.317
$S_l = 20$ $T_l = 300$	**10.720**	**11.221**
$S_l = 50$ $T_l = 50$	11.687	11.496
$S_l = 50$ $T_l = 180$	11.033	11.316
$S_l = 50$ $T_l = 300$	10.733	11.234
$S_l = 200$ $T_l = 50$	11.464	11.967
$S_l = 200$ $T_l = 180$	10.967	11.401
$S_l = 200$ $T_l = 300$	10.751	11.246

Best results are indicated with bold font

Fig. 11 Aging simulations performed on one female subject in FG-NET dataset from age x to target ages ($y = 3$–29 years old), **a** input image, **b** pose-corrected image, **c** age-simulated images at target age y (**ours**), **d** color- and flare-corrected age-simulated images (**ours**), **e** background composite results (**ours**), and **f** reference ground truth images from FG-NET dataset

quency components. See Fig. 6c–e for age simulation at age $y = 55$ from input age $x = 15$ years, color correction, and flare removal results.

$$F_m^{\text{out}} = \text{floor}\left((P_{\text{ref}} - 1)\sum_{n=0}^{m} F_m\right). \tag{2}$$

2.4 Compositing age-synthesized image

After the color correction, we composite the synthesized image in frontal pose (I_f) to the original background image (I_{bg}) using 81 landmark points pts$_{81}$. An example of mapping process from frontal pose image to original image is shown in Fig. 6. The initial mapping results are very noisy due to edges of the synthesized image (Fig. 6f). Therefore, we introduce our edge smoothing background composite (ESBC) method as given in Algorithm 1.

In ESBC Algorithm 1, there are four steps to achieve the composited age-simulated image with smooth edges, I_{smooth}. First step is to generate a mask on facial region including forehead on original background image after mapping age-

simulated image. In Step 2, we make dilated edge image (I_{dilate}) on the mapped image where the boundary regions are noisy by using face mask (mask$_{\text{bg}}$). In Step 3, the blurred image is created from the mapped background image with noise by applying the circular averaging filter. This filtered image, I_{filter}, is then used for filling the edge areas to create a realistic compositing result. The results of compositing after ESBC are shown in Fig. 6h.

3 Aging simulation results

In this section, we provide the aging simulation results using our 3D gender-specific aging model.

3.1 Effect of gender-specific aging versus generic aging

The effectiveness of our gender-specific aging is shown in Fig. 7 by comparing our results with those from generic aging model proposed by Park et al. [1].
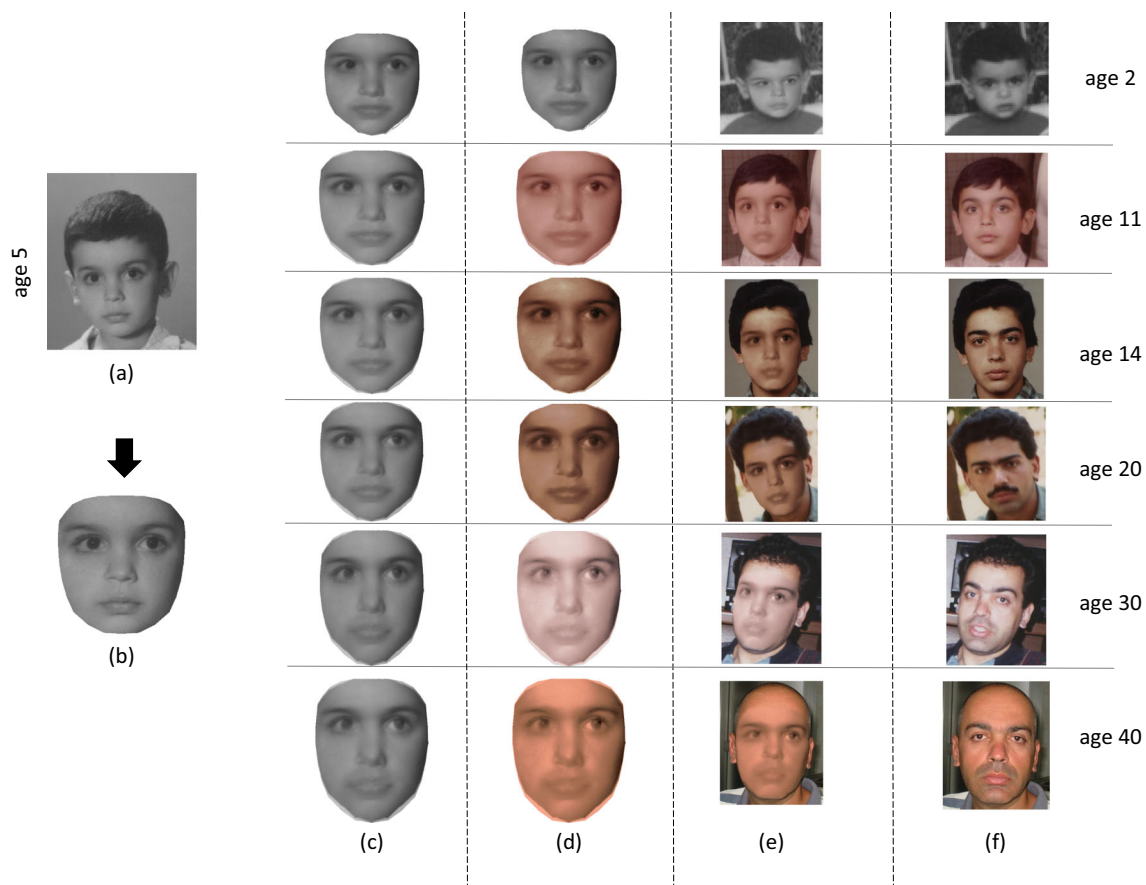
Fig. 12 Aging simulations performed on one male subject in FG-NET dataset from age *x* to target ages (*y* = 2–40 years old), **a** input image, **b** pose-corrected image, **c** age-simulated images at target age (*y*), **d** color-corrected age-simulated images, **e** background composite results, and **f** reference ground truth images from FG-NET dataset

The use of a generic aging model creates a shape and texture space that is not appropriate with the gender of the subject. However, the proposed gender-specific model created more realistic results. The unrealistic aging simulation result of generic aging model [1] is due to the domination of male face images in aging pattern space at older ages.

3.2 Effects of adding datasets on age modeling

There are five different datasets used in constructing our gender-specific 3D aging model. The average texture at a few ages with different components of datasets, such as model 1: FG-NET, model 2: FG-NET + PCSO, model 3: FG-NET + PCSO + Celebrities, model 4: FG-NET + PCSO + Celebrities + BROWNS, and model 5: FG-NET + PCSO + Celebrities + BROWNS + Private, is shown in Fig. 8. It is clearly shown that the average textures for ages 0–69 are becoming more realistic with the addition of datasets (more images and subjects) in the gender-specific aging model.

The effects of the 3D aging model on aging simulation by adding more dataset are shown in Fig. 9. By increasing the number of datasets, both the shape and texture spaces vary in the aging model. Thus, the aging simulation generates different results at the same target age (15 years) as shown in Fig. 9. It is observed that by adding more datasets (increased subjects and number of images), the texture and aging subspace averages provide more realistic aging results compared to those from fewer datasets with less number of images and subjects.

The goodness of fit in shape and texture space is also measured by performing a statistical RMS error test on each aging model consisting of different datasets. The goodness-of-fit test is evaluated separately on shape and texture aging using a test dataset consisting of 439 images of 35 male and female subjects from FG-NET dataset with the average age gap ≤ 40 years. Average age gap per subject is calculated as, avg. age gap $= \frac{\sum(\text{target age} - \text{current age})}{\text{no. images}}$, where the target age is fixed as 55 years and current age is the actual age of the subject in the image. The goodness of fit in shape is computed
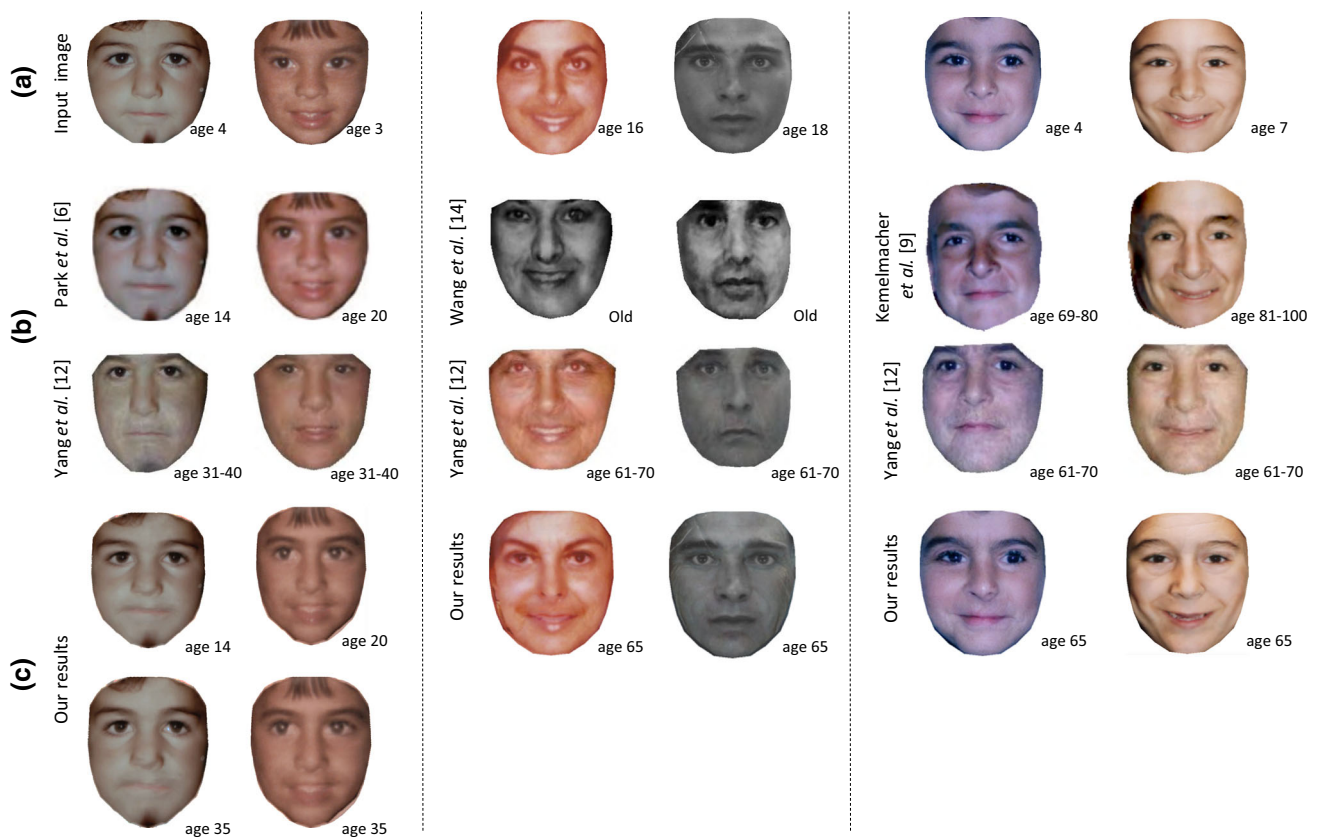
Fig. 13 Aging simulations comparison with state-of-the-art, **a** input images, **b** state-of-the-art results [1,4,6,17] (middle two rows), and **c** our aging simulation results

as $\text{RMS}_{\text{shape}} = \sum |\frac{(GS-ES)^2}{ES}|$, where GS and ES are ground truth and expected shape feature vectors, respectively. Similarly, the goodness of fit in texture space is also computed as $\text{RMS}_{\text{texture}} = \sum |\frac{(GT-ET)^2}{ET}|$, where GT and ET are ground truth and expected texture feature vectors, respectively. The results of RMS error test in shape and texture aging are given in Table 2. For each of aging models 1–5, we take the mean of all RMS values per test dataset and write them in front of the model number as given in Table 2. From the mean RMS values obtained per model, we deduce that using more images from various subjects and ages for aging model construction helps to improve the aging simulation quality.

3.3 Effects of shape and texture eigenvectors

Similar to Park et al. [1], initially we selected the shape and texture eigenvalues as 50 and 180, respectively. Later, we evaluated different shape and texture values to test the simulation results. The various lengths of shape and texture eigenvectors, we selected, are: $S_l = [20, 50, 200]$ and $T_l = [50, 180, 300]$, respectively. The comparison of effects of using different lengths of eigenvectors in shape and texture aging is shown in Fig. 10.

To test the effectiveness of different combinations of shape and texture eigenvectors in the aging process, we performed the RMS error test again using the same test dataset as mentioned in Sect. 3.2. The aging simulations are performed on 35 subjects (439 images) using nine different combinations of lengths of eigenvectors in shape (S_l) and texture (T_l) space. The mean RMS error values in shape and texture aging obtained from this test are given in Table 3. To perform simulations, we used our final 3D gender-specific aging model consisting of five different datasets as mentioned in Sect. 2.1. From our findings, we can deduce that using longer length of eigenvectors in shape space (S_l) does not affect the shape aging much. However, by using the longer length of eigenvectors of texture space (T_l), we achieved more detailed texture in the images and less RMS error. Thus, in our aging simulations we use 20 and 300 as the lengths of eigenvectors for shape (S_l) and texture (T_l) space, respectively.

3.4 Aging simulation: FG-NET

The automatic aging simulation experiments are performed on a well-known dataset, FG-NET. In this dataset, the age and gender of the subjects are known and we used this given
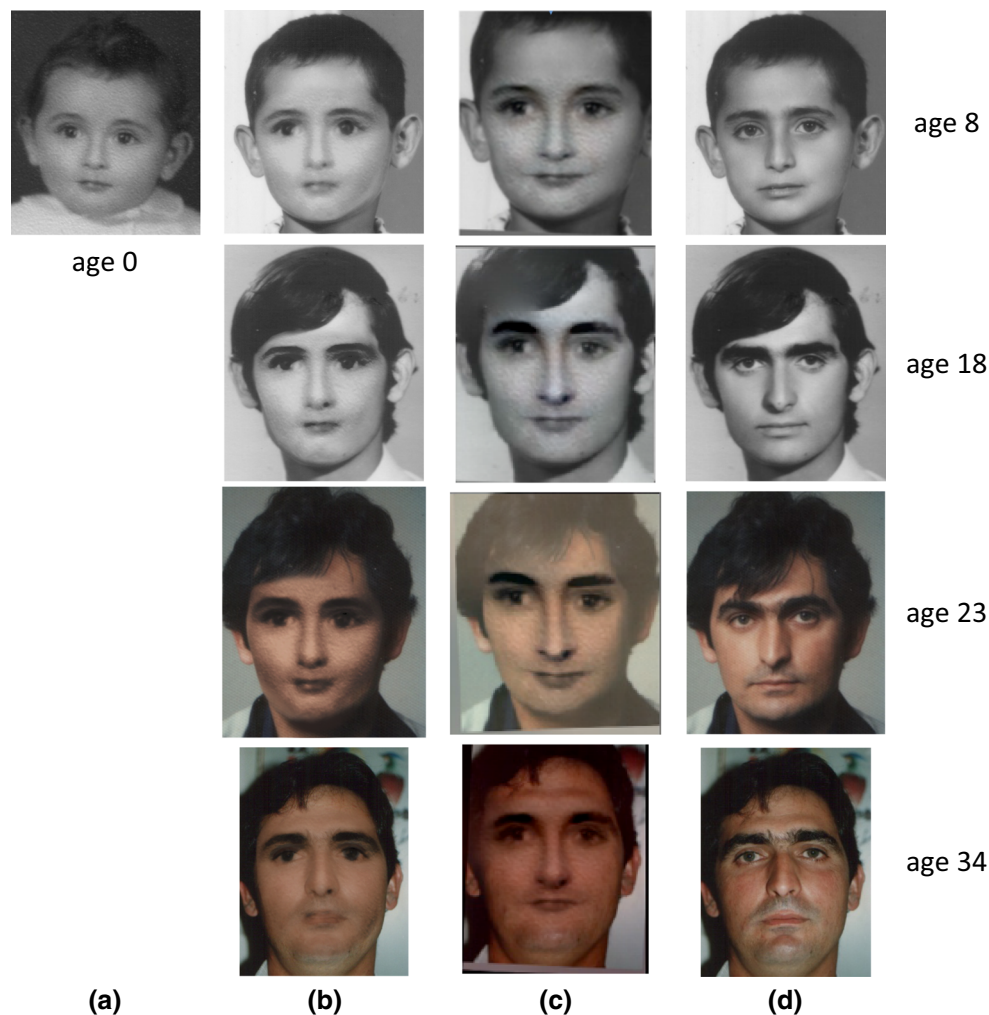
Fig. 14 Aging simulations comparison with other methods, **a** input image, **b** aging simulation results (**ours**), **c** Kemelmacher et al. [6], and **d** ground truth reference images

information to select the 3D gender-specific aging model. To perform aging, we select an input image at age x and automatically simulate to the target or reference ages y. The aging simulation results in frontal pose and composited image are shown in Figs. 11 and 12 for male and female aging, respectively. Our aging simulation results in frontal pose are first color corrected and then mapped to the reference images from the FG-NET dataset. For a comparison, we also show the original reference images from the FG-NET dataset at age y. From Figs. 11e and 12e, we see that our automatic aging simulation results are visually similar to the reference images at the same age.

3.5 Comparison of aging simulations

In this section, the comparison of aging simulations is provided with ground truth reference images as well as with state-of-the-art results.

3.5.1 Comparison with state-of-the-art

We compare our simulation results in frontal pose and composited images with those of the state-of-the-art techniques by using FG-NET dataset. The current and target ages of each subject are as close as possible to those used in the state-of-the-arts [1,4,6,17]. The comparison of age simulation results is shown in Figs. 13 and 14.

In Fig. 13, we see that in the younger age group (20 and 35 years old) our shape aging is more realistic than those of [1,4]. In the older age group (65 years old), our age simulation results show convincing aging results with visible wrinkles around mouth and forehead as compared to [4,6,17]. Furthermore, our aging simulation results keep the histogram similar with that of the original input image.

In Fig. 14, we compare our composited age simulation results with those of [6] and the reference images from FG-NET dataset at the target age y. From the results in Fig. 14c ,
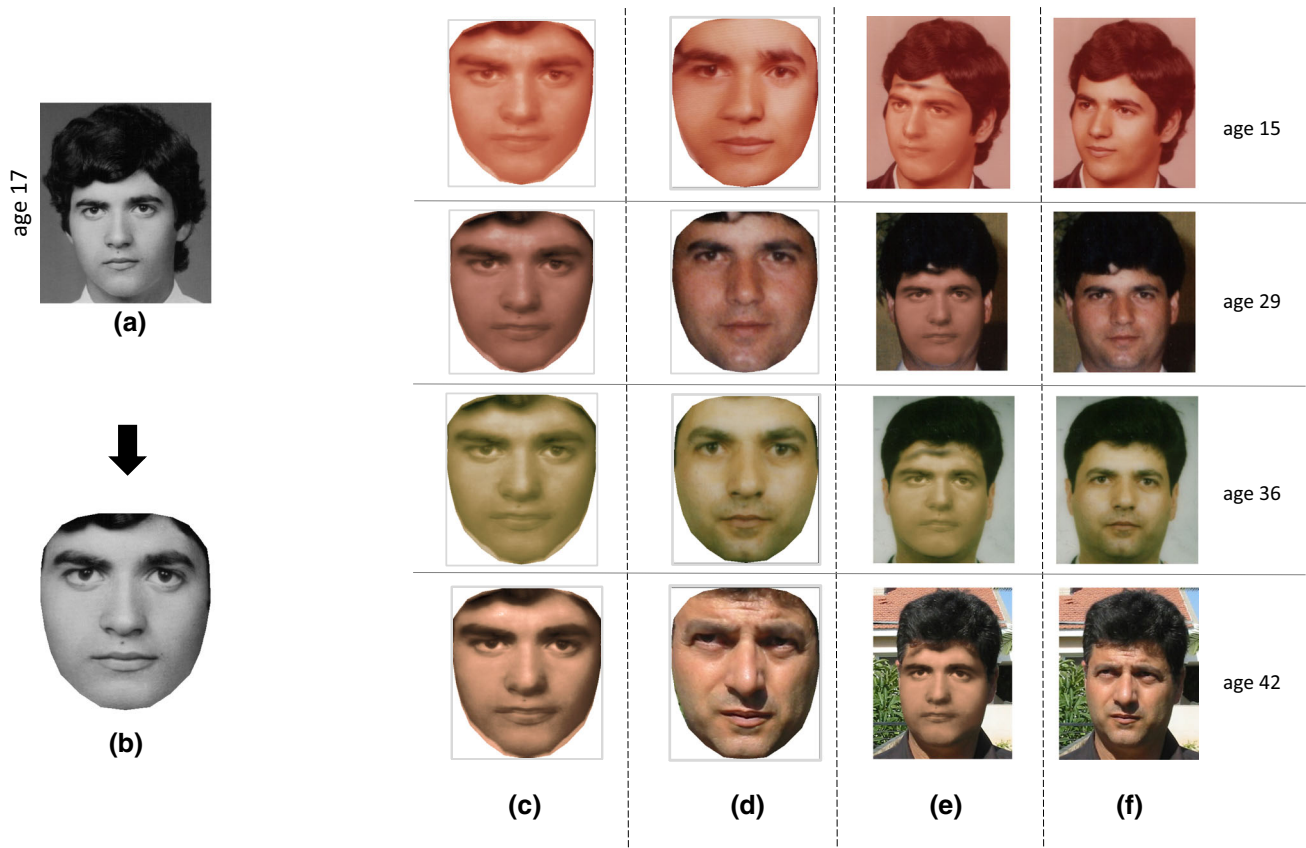
Fig. 15 Comparisons of aging simulations with ground truth, **a** input image, **b** pose-corrected image in frontal pose, **c** age-simulated images at target ages with histogram correction (**ours**), **d** ground truth images in frontal pose at target age, **e** age-simulated images composited to background (**ours**), and **f** ground truth images at target ages

we can see that Kemelmacher et al. [6] used facial components from the reference images. However, we do not use the facial properties of the reference and perform the simulations automatically from the input image. Thus, our aging simulation approach provides better performance as compared with [6].

3.5.2 Comparison with reference images

The comparison of age simulated and ground truth in frontal pose and composited images at target age y is provided in Figs. 15 and 16 for male and female aging, respectively. For the comparison of age-simulated images with background, we composited the aged/de-aged images to the ground truth images at the target age y.

From Figs. 15 and 16, we see that our automatic aging simulation results are visually similar to the ground truth reference images at the same age. This is qualitative evidence that our 3D gender-specific aging model can automatically generate effective aging simulation results from a single 2D image at an arbitrary age.

3.6 Quantitative evaluation

3.6.1 Subjective evaluation

For subjective quantitative evaluation, we conducted an online user survey with 30 different subjects. Each subject views (i) input images at age x and (ii) their respective aging simulation results at age y generated by our approach and other methods in a random order. We selected two evaluation metrics in this survey, i.e., (1) age matching (AM) and (2) identity matching (IM). Each subject is asked to evaluate the age-simulated images by simple scoring method. For each metric, we provide three scores, i.e., score = [0, 0.5, 1]. The scores 1, 0.5, and 0 mean good, neutral, and bad matching in terms of AM and IM. Each user was shown 36 sets of images in one sitting, and the next day we show them the same set of images with random order of view to get their scores again. We take the averages of sum of score count (good, neutral, bad) for age-simulated images and sum of average scores per pair of images ($\sum$ Score$_{pp}$). The statistics of the normalized average scores are given in Table 4.
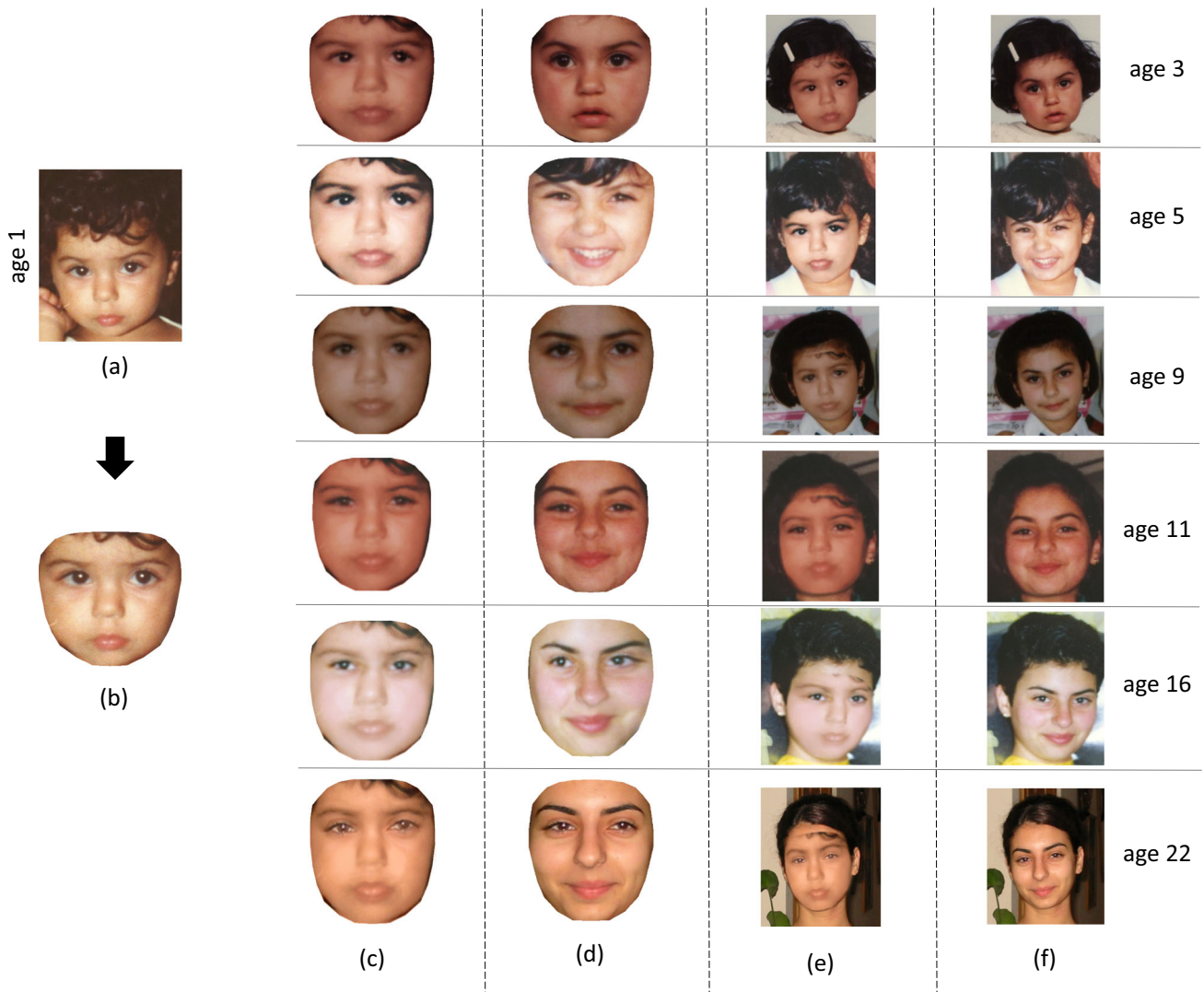
Fig. 16 Comparisons of aging simulation with ground truth, **a** input image, **b** pose-corrected image in frontal pose, **c** age-simulated images at target ages with histogram correction (**ours**), **d** ground truth images in frontal pose at target age, **e** age-simulated images composited to background (**ours**), and **f** ground truth images at target ages

Table 4 Subjective quantitative evaluations: average scores in terms of age matching (AM) and identity matching (IM)

	Age matching (AM)				Identity matching (IM)			
	Good	Neutral	Bad	$\sum \text{Score}_{pp}$	Good	Neutral	Bad	$\sum \text{Score}_{pp}$
Ours	**5.33**	1.5	0.167	**6.08**	**5.5**	1.5	0	**6.25**
Kemelmacher [6]	4	1.5	1.5	4.75	4	3	0	5.5
Park et al. [1]	2.5	3	1.5	4	4.5	2.5	0	5
Yang et al. [4]	3	1.5	2.5	3.75	2	3	2	3.5
Wang et al. [17]	4.167	1.5	1.33	4.917	3.833	1.833	1.33	4.75

Best results are indicated with bold font

The comparison with state-of-the-art ([1,4,6,17]) shows that our age simulation results got 76.2% and 78.5% good average score ratings in terms of AM and IM, respectively. The sum of scores per pair of age-simulated images is also higher for our approach.

Table 5 Objective quantitative evaluations: face verification experiment on FG-NET dataset (two images per subject)

	TAR at 1% FAR (%)	FAR at 85% TAR (%)	Rank-1 (%)	Rank-5 (%)	Rank-10 (%)	Rank-20 (%)
No aging	8.54	45.89	11.59	23.78	36.59	50.00
Aging	12.80	45.82	10.98	26.83	40.24	53.66
Fusion	**13.41**	**41.69**	**13.41**	**28.05**	**41.46**	**53.66**

Best results are indicated with bold font

Max. age gap = 40 years, min. age gap = 6 years, and avg. age gap = 16.67 years for the images used in this experiment

Table 6 Objective quantitative evaluations: face verification experiment on FG-NET dataset (all images per subject)

	TAR at 1% FAR (%)	FAR at 85% TAR (%)	Rank-1 (%)	Rank-5 (%)	Rank-10 (%)	Rank-20 (%)
No aging	42.67	26.53	83.38	92.49	95.23	96.65
Aging	40.93	25.99	79.05	90.48	94.22	96.35
Fusion	**43.24**	**25.15**	**83.89**	**93.00**	**95.44**	96.65

Best results are indicated with bold font

TAR and rank-1 show a better improvement after fusion

Table 7 Comparison of state-of-the-art face verification systems with the proposed approach

	Face matcher	Dataset details (#subjects, #images) in probe and gallery	Rank-1 (%)
Park et al. [1]	FaceVACS	FG-NET(82,82),	37.40
Geng et al. [12]	Mahalanobis distance, PCA	FG-NET (10, 10)	38.10
Du et al. [30]	FaceVACS,	FG-NET (age group ≤ 18)	39.2
	PCA	FG-NET (age group > 18)	58.70
Li et al. [8]	MFDA, LDA , (Fisherface) and FaceVACS	FG-NET (82,1002)	47.50
Gong et al. [29]	Cosine distance	FG-NET (82,1002)	69.0
Yang et al. [31]	Graph-based objective function	FG-NET (82,1002)	64.47
Proposed	VGG face CNN	FG-NET	**83.89***
Approach	Descriptor, L2 norm	(82, 1002)	

[a]Twofold cross-validation is performed by splitting the FG-NET subjects in two subsets. We adapted the leave-one-out fashion similar to Li et al.[8]

*Top Rank-1 identification result is indicated with bold font

3.6.2 Objective evaluation

For objective quantitative evaluations, we used the VGG face CNN descriptor [28] on the age-simulated and original images. We performed two face verification experiments with our age-simulated images. For the first experiment, we selected two images per subject with the third minimum and third maximum ages available (average age gap ≤ 16.67 years) in the FG-NET dataset. Three different facial matching schemes were tested, namely (1) no aging, (2) aging, (3) fusion of no aging and aging. The evaluation results are given in Table 5 in terms of TAR at 1% FAR, FAR at 85% TAR, and Rank-1∼20 matching. Score-level fusion was performed using the weighted score sum method, where the weight (w) was assigned to the aging-based matching scores and weight ($1 - w$) was assigned to the original image matching scores.

We empirically chose the weight value in each matching experiment. The fusion of aging simulations with no aging-based matching scores improves the matching accuracy by 3.65%.

In the second facial verification experiment on the FG-NET dataset, we used all the images per subject available in the FG-NET dataset and performed matching in a leave-one-out fashion described in Li et al. [8]. We performed twofold cross-validation with each fold containing images from 41 subjects. Table 6 shows the face recognition results for (1) no aging, (2) aging, and (3) fusion of no aging and aging. The recognition performances after fusion are 83.89% in rank-1 and 43.24% in TAR, which are the best performances among similar evaluations in the literature.

The comparison of our face verification experiments with state-of-the-art [1,8,12,29–31] is provided in Table 7. The

Algorithm 1 Edge Smoothing Background Composite (ESBC)

---

function ESBC

Input
$pts_{81} = 13$ forehead and 68 facial boundary points
I_{bg} = Composited image with noisy edges
Output
I_{smooth} = Composited image with smooth edges

  Step 1: Face mask
   $mask_{bg} \leftarrow$ Specify polygon region of face using pts_{81}
  Step 2: Face edges
  Apply Canny edge detector on I_{bg},
   $I_{canny} \leftarrow$ Face edge image
  Dilate the edges of I_{canny},
   $I_{dilate} \leftarrow$ dilated face edges

  Step 3: Blur image
  Apply circular averaging,
   $H \leftarrow$ averaging filter
   $I_{filter} \leftarrow$ Blur I_{bg} using H,

  Step 4: Smooth composited image
   $[r, c] \leftarrow$ locations of edges in I_{dilate}
   $I_{smooth} \leftarrow$ Copy of I_{bg}
   $I_{smooth} \leftarrow$ Copy blur edges from I_{filter} using [r,c].

  end

---

comparison results show that our age simulation-based face verification results are higher in Rank-1 (%) and TAR (%) than state-of-the-art approaches. Thus, we are confident that our gender-specific aging model provides better qualitative and quantitative results.

4 Conclusion

In this paper, we presented our 3D gender-specific aging model to automatically simulate an input face image from age x to age y. The proposed aging model can age or de-age an input 2D image of a person by mapping it onto frontal pose of a 3D morphable model. The aging results on frontal pose images are then mapped to original background images using edge smoothing background composite (ESBC) method. Before compositing, we performed color and flare correction to match the skin color of age-simulated images. In our aging model, we have five different datasets such as FG-NET, PCSO, Celebrities, BROWNS, and Private.

We tested the effect of dataset configuration on our aging model in terms of averages of shape and texture pattern spaces and aging simulation results. Furthermore, we studied how the length of shape and texture eigenvectors affect the aging process. The aging simulation results are achieved on FG-NET dataset and compared with ground truth and state-of-the-art results. Although there are expression variations in the ground truth images, the texture and shape changes with aging or de-aging were observed more realistic compared with state-of-the-arts.

In order to evaluate our aging simulation approach qualitatively, we performed subjective and objective tests. In subjective evaluations, our approach received 76.2% and 78.5% scores in terms of age matching (AM) and identity matching (IM), respectively. In objective evaluations, our aging simulations-based face verification results show 83.89% rank-1 accuracy and outperform the state-of-the-art methods. In future, we would add the functionality of aging or de-aging of hair and mustaches in the male subjects. Moreover, we would like to add more datasets that comprise of ethnicity information in the 3D aging model.

References

1. Park, U., Tong, Y., Jain, A.K.: Age-invariant face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **32**(5), 947–954 (2010)
2. Park, U., Tong, Y., Jain, A. K.: Face recognition with temporal invariance: a 3D aging model. In: International Conference on Face and Gesture Recognition (FG). Amsterdam, The Netherlands, pp. 17–19 (2008)
3. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.: Face verification across age progression using discriminative methods. IEEE Trans. Inf. Forensic Secur. **5**(1), 82–91 (2010)
4. Yang, H., Huang, D., Wang, Y., Wang, H., Tang, Y.: Face aging effect simulation using hidden factor analysis joint sparse representation. IEEE Trans. Image Process. **25**(6), 2493–2507 (2016)
5. Suo, J., Chen, X., Shan, S., Gao, W.: Learning long term face aging patterns from partially dense aging databases. In: IEEE 12th International Conference on Computer Vision, Kyoto, pp. 622–629 (2009)
6. Kemelmacher-Shlizerman, I., Suwajanakorn, S., Seitz, S. M.: Illumination-aware age progression. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3334–3341 (2014)

7. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.: A study of face recognition as people age. In: Proceeding of International Conference of Computer Vision, pp. 1–8 (2007)

8. Li, Z., Park, U., Jain, A.K.: A discriminative model for age invariant face recognition. IEEE Trans. Inf. Forensic Secur. **6**(3), 1028–1037 (2011)

9. Geng, X., Zhou, Z., Smith-Miles, K.: Automatic age estimation based on facial aging patterns. IEEE Trans. Pattern Anal. Mach. Intell. **29**(12), 2234–2240 (2007)

10. Lanitis, A., Taylor, C.J., Cootes, T.F.: Toward automatic simulation of aging effects on face images. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 442–455 (2002)

11. Du, J.-X., Zhai, C.-M., Ye, Y.-Q.: Face aging simulation based on NMF algorithm with sparseness constraints. In: Proceedings of the 7th International Conference on Advance Intelligent Computing Theories and Applications, ICIC, China. Springer, Berlin, pp. 516–522 (2012)

12. Geng, X., Zhou, Z., Smith-Miles, K.: Automatic age estimation based on facial aging patterns. IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) **29**(12), 2234–2240 (2007)

13. Patterson, E., Sethuram, A., Albert, M., Ricanek Jr., K., King, M.: Aspects of age variation in facial morphology affecting biometrics. In: Proceedings of the 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–6 (2007)

14. Singh, R., Vatsa, M., Noore, A., Singh, S. K.: Age transformation for improving face recognition performance. In: Proceedings of the 2nd International Conference on Pattern Recognition and Machine Intelligence (PReMI), pp. 576–583 (2007)

15. Suo, J., Min, F., Zhu, S., Shan, S., Chen, X.: A multi-resolution dynamic model for face aging simulation. In: IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, pp. 1–8 (2007)

16. Patterson, E., Ricanek, K., Albert, M., Boone, E.: Automatic representation of adult aging in facial images. In: Proceedings of the International Conference on Visualization, Imaging, and Image Processing, pp. 171–176 (2006)

17. Wang, Y., Zhang, Z., Li, W., Jiang, F.: Combining tensor space analysis and active appearance models for aging effect simulation on face images. In: IEEE Transaction on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, no. 4, pp. 1107–1118 (2012)

18. Wang, W. et al.: Recurrent face aging. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp. 2378–2386 (2016). https://doi.org/10.1109/CVPR.2016.261

19. Wang, H., Huang, D., Wang, Y., Yang, H.: Facial aging simulation via tensor completion and metric learning. IET Comput. Vis. **11**(1), 78–86 (2017). https://doi.org/10.1049/iet-cvi.2016.0074

20. Fu, Y., Guo, G., Huang, T.S.: Age synthesis and estimation via faces: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **32**(11), 1955–1976 (2010). https://doi.org/10.1109/TPAMI.2010.36

21. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3D faces. In: Proceedings of the ACM SIGGRAPH Conference on Computer Graphics and Interactive Techniques (CGIT), pp. 187–194 (1999)

22. Nixon, N., Galassi, P.: The Brown Sisters, Thirty-Three Years. The Museum of Modern Art, New York (2007)

23. Milborrow, S., Nicolls, F.: Active shape models with SIFT descriptors and MARS. J. VISAPP **1**(5), 380–387 (2014)

24. Booth, J., Antonakos, E., Ploumpis, S., Trigeorgis, G., Panagakis, Y., Zafeiriou, S.: 3D face Morphable models "In-the-Wild". In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, pp. 5464–5473 (2017). https://doi.org/10.1109/CVPR.2017.580

25. Booth, J., Roussos, A., Zafeiriou, S., Ponniahy, A., Dunaway, D.: "A 3D Morphable model learnt from 10,000 faces. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp. 5543–5552 (2016). https://doi.org/10.1109/CVPR.2016.598

26. Tikhonov, A.N.: Translated: on the stability of inverse problems. C. R. (Doklady) Acad. Sci. l'URSS **39**(5), 195–198 (1943)

27. Matlab 8.3.0.532 Image Processing Toolbox Release: Image Enhancement: imhistmatch. The MathWorks, Inc., Natick, MA, USA. Date accessed: 19 April 2017 (2014)

28. Parkhi, O. M., Vedaldi, A., Zisserman, A.: Deep face recognition. In: British Machine Vision Conference (2015)

29. Gong, D., Li, Z., Lin, D., Liu, J., Tang, X.: Hidden factor analysis for age invariant face recognition. In: IEEE International Conference on Computer Vision (ICCV), Sydney, NSW, pp. 2872–2879 (2013)

30. Du, J.-X., Zhai, C.-M., Ye, Y.-Q.: Face aging simulation and recognition based on NMF algorithm with sparseness constraints. Neurocomputing **116**, 250–259 (2013)

31. Yang, H., Huang, D., Wang, Y.: Age invariant face recognition based on texture embedded discriminative graph model. In: IEEE International Joint Conference on Biometrics (IJCB), Clearwater, FL, pp. 1–8 (2014)

Sidra Riaz received her BS degree in telecom engineering from NUCES-FAST, Islamabad, Pakistan, in 2011, and her MS degree in computer engineering from Chosun University, Gwangju, Korea, in 2013. She is currently a Ph.D. candidate in the Department of Computer Science and Engineering, Sogang University, Seoul, Korea. Her research interests include image processing, aging modeling, and pattern recognition. Ms. Sidra Riaz is the recipient of National ICT R & D Scholarship award (2007–2011) and Global IT scholarship award (2011–2013).

Unsang Park received the BS and MS degrees from the Department of Materials Engineering, Hanyang University, South Korea, in 1998 and 2000, respectively. He received the MS and Ph.D. degrees from the Department of Computer Science and Engineering, Michigan State University, in 2004 and 2009, respectively. From 2012, he is an assistant professor in the Department of Computer Science and Engineering at Sogang University. His research interests include pattern recognition, image processing, computer vision, and machine learning.

Jongmoo Choi studied physics (B.S.), cognitive science (M.S.), and computer engineering (Ph.D.) at Sungkyunkwan University, South Korea. From 1999 to 2003, he was the director of the research center of Dream Mirh Co. Ltd. From 2004 to 2006, he was a research assistant professor at ISRC (Intelligent Systems Research Center), Sungkyunkwan University. He is currently a senior research associate at the Institute for Robotics and Intelligent Systems, University of Southern California, USA. His research interests include 3D face/body modeling, aerial video analysis, computational geometry, and face recognition.

Prem Natarajan is Michael Keston Executive Director of ISI, a vice dean of the USC Viterbi School of Engineering, and a professor of computer science. At ISI, he leads managerial and technical directions institute wide, including research, development and the MOSIS electronic chip brokerage. He also heads teams in his areas of expertise: novel approaches to face, character, handwriting, and speech recognition, along with other deep learning and natural language processing directions.