

Chapter 10

Face Aging Modeling

Unsang Park and Anil K. Jain

10.1 Introduction

Face recognition accuracy is typically limited by the large intra-class variations caused by factors such as pose, lighting, expression, and age [16]. Therefore, most of the current work on face recognition is focused on compensating for the variations that degrade face recognition performance. However, facial aging has not received adequate attention compared to other sources of variations such as pose, lighting, and expression.

Facial aging is a complex process that affects both the shape and texture (e.g., skin tone or wrinkles) of a face. The aging process appears in different manifestations in different age groups, gender and ethnicity. While facial aging is mostly represented by the facial growth in younger age groups (e.g., below 18 years of age), it is mostly represented by relatively large texture changes and minor shape changes (e.g., due to the change of weight or stiffness of skin) in older age groups (e.g., over 18 years of age). Therefore, an age invariant face recognition scheme needs to be able to compensate for both types of aging process.

Some of the face recognition applications where age invariance or correction is required include (i) identifying missing children, (ii) screening for watch list, and (iii) multiple enrollment detection problems. These three scenarios have two common characteristics: (i) a significant age difference exists between probe and gallery images (images obtained at verification and enrollment stages, respectively) and (ii) an inability to obtain a user's face image to update the template (gallery). Identifying missing children is one of the most apparent applications where age compensation is needed to improve the recognition performance. In screening applications, aging is a major source of difficulty in identifying suspects in a watch list. Repeat

U. Park (✉) · A.K. Jain
Michigan State University, East Lansing, MI 48824, USA
e-mail: parkunsa@cse.msu.edu

A.K. Jain
e-mail: jain@cse.msu.edu

offenders commit crimes at different time periods in their lives, often starting as a juvenile and continuing throughout their lives. It is not unusual to encounter a time lapse of ten to twenty years between the first (enrollment) and subsequent (verification) arrests. Multiple enrollment detection for issuing government documents such as driver licenses and passports is a major problem that various government and law enforcement agencies face in the facial databases that they maintain. Face or some other types of biometric traits (e.g., fingerprint or iris) are the only ways to reliably detect multiple enrollments.

Ling et al. [10] studied how age differences affect the face recognition performance in a real passport photo verification task. Their results show that the aging process does increase the recognition difficulty, but it does not surpass the challenges posed due to change in illumination or expression. Studies on face verification across age progression [19] have shown that: (i) simulation of shape and texture variations caused by aging is a challenging task, as factors like life style and environment also contribute to facial changes in addition to biological factors, (ii) the aging effects can be best understood using 3D scans of human head, and (iii) the available databases to study facial aging are not only small but also contain uncontrolled external and internal variations (e.g., pose, illumination, expression, and occlusion). It is due to these reasons that the effect of aging in facial recognition has not been as extensively investigated as other factors that lead to large intra-class variations in facial appearance.

Some biological and cognitive studies on face aging process have also been conducted, see [18, 25]. These studies have shown that cardioidal strain is a major factor in the aging of facial outlines. Such results have also been used in psychological studies, for example, by introducing aging as caricatures generated by controlling 3D model parameters [12]. Patterson et al. [15] compared automatic aging simulation results with forensic sketches and showed that further studies in aging are needed to improve face recognition techniques. A few seminal studies [20, 24] have demonstrated the feasibility of improving face recognition accuracy by simulated aging. There has also been some work done in the related area of age estimation using statistical models, for example, [8, 9]. Geng et al. [7] learn a subspace of aging pattern based on the assumption that similar faces age in similar ways. Their face representation is composed of face texture and the 2D shape represented by the coordinates of the feature points as in the Active Appearance Models. Computer graphics community has also shown facial aging modeling methods in 3D domain [22], but the effectiveness of the aging model was not evaluated by conducting a face recognition test.

Table 10.1 gives a brief comparison of various methods for modeling aging proposed in the literature. The performance of these models is evaluated in terms of the improvement in the identification accuracy. When multiple accuracies were reported in any of the studies under the same experimental setup (e.g., due to different choice of probe and gallery), their average value is listed in Table 10.1; when multiple accuracies are reported under different approaches, the best performance is reported. The identification accuracies of various studies in Table 10.1 cannot be directly compared due to the differences in the database, the number of subjects

Table 10.1 A comparison of various face aging models [13]

	Approach	Face matcher	Database (#subjects, #images) in probe and gallery	Rank-1 identification accuracy (%)	
				Original image	After aging model
Ramanathan et al. (2006) [20]	Shape growth modeling up to age 18	PCA	Private database (109, 109)	8.0	15.0
Lanitis et al. (2002) [8]	Build an aging function in terms of PCA coefficients of shape and texture	Mahalanobis distance, PCA	Private database (12, 85)	57.0	68.5
Geng et al. (2007) [7]	Learn aging pattern on concatenated PCA coefficients of shape and texture across a series of ages	Mahalanobis distance, PCA	FG-NET* (10, 10)	14.4	38.1
Wang et al. (2006) [26]	Build an aging function in terms of PCA coefficients of shape and texture	PCA	Private database (NA, 2000)	52.0	63.0
Patterson et al. (2006) [14]	Build an aging function in terms of PCA coefficients of shape and texture	PCA	MORPH+ (9, 36)	11.0	33.0
Park et al. [13]	Learn aging pattern based on PCA coefficients in separated 3D shape and texture given 2D database	FaceVACS	FG-NET** (82, 82) MORPH-Album1++ (612,612) BROWNS (4, 4)—probe (100, 100)—gallery	26.4 57.8 15.6	37.4 66.4 28.1

*Used only a very small subset of the FG-NET database that contains a total of 82 subjects

+Used only a very small subset of the MORPH database that contains a total of 625 subjects

**Used all the subjects in FG-NET

++Used all the subjects in MORPH-Album1 which have multiple images

and the underlying face recognition method used for evaluation. Usually, the larger the number of subjects and the larger the database variations in terms of age, pose, lighting and expression, the smaller the recognition performance improvement by an aging model. The identification accuracy for each approach in Table 10.1 before aging simulation indicates the difficulty of the experimental setup for the face recognition test as well as the limitations of the face matcher.

Compared with other published approaches, the aging model proposed by Park et al. [13] has the following features.

- 3D aging modeling: Includes a pose correction stage and a more realistic model of the aging pattern in the *3D domain*. Considering that the aging is a 3D process, 3D modeling is better suited to capture the aging patterns. Their method is the only viable alternative to building a 3D aging model directly, as no 3D aging database is currently available. Scanned 3D face data rather than reconstructed is used in [22], but they were not collected for aging modeling and hence, do not contain as much aging information as the 2D facial aging database.
- Separate modeling of shape and texture changes: Three different modeling methods, namely, shape modeling only, separate shape and texture modeling and combined shape and texture modeling (e.g., applying 2nd level PCA to remove the correlation between shape and texture after concatenating the two types of feature vectors) were compared. It has been shown that the separate modeling is better than combined modeling method, given the FG-NET database as the training data.
- Evaluation using a state-of-the-art commercial face matcher, FaceVACS: A state-of-the-art face matcher, FaceVACS from Cognitec [4] has been used to evaluate the aging model. Their method can thus be useful in practical applications requiring an age correction process. Even though their method has been evaluated only on one particular face matcher, it can be used directly in conjunction with any other 2D face matcher.
- Diverse Databases: FG-NET has been used for aging modeling and the aging model has been evaluated on three different databases: FG-NET (in a leave-one-person-out fashion), MORPH, and BROWNS. Substantial performance improvements have been observed on all three databases.

The rest of this Chapter is organized as follows: Sect. 10.2 introduces the preprocessing step of converting 2D images to 3D models, Sect. 10.3 describes the aging model, Sect. 10.4 presents the aging simulation methods using the aging model, and Sect. 10.5 provides experimental results and discussions. Section 10.6 summarizes the conclusions and lists some directions for future work.

10.2 Preprocessing

Park et al. propose to use a set of 3D face images to learn the model for recognition, because the true craniofacial aging model [18] can be appropriately formulated only in 3D. However, since only 2D aging databases are available, it is necessary to first convert these 2D face images into 3D. Major notations that are used in the following sections are defined first.

- $S_{mm} = \{S_{mm,1}, S_{mm,2}, \dots, S_{mm,n_{mm}}\}$: a set of 3D face models used in constructing the reduced morphable model. n_{mm} is the number of 3D face models.
- S_{α} : reduced morphable model represented with the model parameter α .
- $S_{2d,i}^j = \{x_1, y_1, \dots, x_{n_{2d}}, y_{n_{2d}}\}$: 2D facial feature points for the i th subject at age j . n_{2d} is the number of points in 2D.

- $\mathbf{S}_i^j = \{x_1, y_1, z_1, \dots, x_{n_{3d}}, y_{n_{3d}}, z_{n_{3d}}\}$: 3D feature points for the i th subject at age j . n_{3d} is the number of points in 3D.
- \mathbf{T}_i^j : facial texture for the i th subject at age j .
- \mathbf{s}_i^j : reduced shape of \mathbf{S}_i^j after applying PCA on \mathbf{S}_i^j .
- \mathbf{t}_i^j : reduced texture of \mathbf{T}_i^j after applying PCA on \mathbf{T}_i^j .
- \mathbf{V}_s : top L_s principle components of \mathbf{S}_i^j .
- \mathbf{V}_t : top L_t principle components of \mathbf{T}_i^j .
- $\mathbf{S}_{w_s}^j$: synthesized 3D facial feature points at age j represented with weight w_s .
- $\mathbf{T}_{w_t}^j$: synthesized texture at age j represented with weight w_t .
- $n_{mm} = 100, n_{2d} = 68, n_{3d} = 81, L_s = 20, \text{ and } L_t = 180$.

In the following subsections, $\mathbf{S}_{2d,i}^j$ is first transformed to \mathbf{S}_i^j using the reduced morphable model \mathbf{S}_α . Then, 3D shape aging pattern space $\{\mathbf{S}_{w_s}\}$ and texture aging pattern space $\{\mathbf{T}_{w_t}\}$ are constructed using \mathbf{S}_i^j and \mathbf{T}_i^j .

10.2.1 2D Facial Feature Point Detection

Manually marked feature points are used in aging model construction. However, in the test stage the feature points need to be detected automatically. The feature points on 2D face images are detected using the conventional Active Appearance Model (AAM) [3, 23]. AAM models for the three databases are trained separately, the details of which are given below.

10.2.1.1 FG-NET

Face images in the FG-NET database have already been (manually) marked by the database provider with 68 feature points. These feature points are used to build the aging model. Feature points are also automatically detected and the face recognition performance based on manual and automatic feature point detection methods are compared. The training and feature point detection are conducted in a cross-validation fashion.

10.2.1.2 MORPH

Unlike the FG-NET database, a majority of face images in the MORPH database belong to African-Americans. These images are not well represented by the AAM model trained on the FG-NET database due to the differences in the cranial structure between the Caucasian and African-American populations. Therefore, a subset of images (80) in the MORPH database are labeled as a training set for the automatic feature point detector in the MORPH database.

10.2.1.3 BROWNS

The entire FG-NET database is used to train the AAM model for detecting feature points on images in the BROWNS database.

10.2.2 3D Model Fitting

A simplified deformable model based on Blanz and Vetter's model [2] is used as a generic 3D face model. For efficiency, the number of vertices in the 3D morphable model is drastically reduced to 81, 68 of which correspond to salient features present in the FG-NET database, while the other 13 delineate the forehead region. Following [2], PCA was performed on the simplified shape sample set, $\{S_{mm}\}$. The mean shape \bar{S}_{mm} , the eigenvalues λ_l 's, and unit eigenvectors \mathbf{W}_l 's of the shape covariance matrix are obtained. Only the top L ($= 30$) eigenvectors are used, again for efficiency and stability of the subsequent fitting algorithm performed on the possibly very noisy dataset. A 3D face shape can then be represented using the eigenvectors as

$$\mathbf{S}_\alpha = \bar{S}_{mm} + \sum_{l=1}^L \alpha_l \mathbf{W}_l, \quad (10.1)$$

where the parameter $\alpha = [\alpha_l]$ controls the shape, and the covariance of α 's is the diagonal matrix with λ_i as the diagonal elements. A description is given below on how to transform the given 2D feature points $\mathbf{S}_{2d,i}^j$ to the corresponding 3D points \mathbf{S}_i^j using the reduced morphable model \mathbf{S}_α .

Let $E(\cdot)$ be the overall error in fitting the 3D model of one face to its corresponding 2D feature points, where

$$E(\mathbf{P}, \mathbf{R}, \mathbf{t}, a, \{\alpha_l\}_{l=1}^L) = \|\mathbf{S}_{i,2d}^j - \mathbf{T}_{\mathbf{P},\mathbf{R},\mathbf{t},a}(\mathbf{S}_\alpha)\|^2. \quad (10.2)$$

Here, $\mathbf{T}(\cdot)$ represents a transformation operator performing a sequence of operations, that is, rotation, translation, scaling, projection, and selecting n_{2d} points out of n_{3d} that have correspondences. To simplify the procedure, an orthogonal projection \mathbf{P} is used.

In practice, the 2D feature points that are either manually labeled or automatically generated by AAM are noisy, which means overfitting these feature points may produce undesirable 3D shapes. This issue is addressed by introducing a Tikhonov regularization term to control the Mahalanobis distance of the shape from the mean shape. Let σ be the empirically estimated standard deviation of the energy E induced by the noise in the location of the 2D feature points. The regularized energy is defined as

$$E' = E/\sigma^2 + \sum_{l=1}^L \alpha_l^2/\lambda_l. \quad (10.3)$$

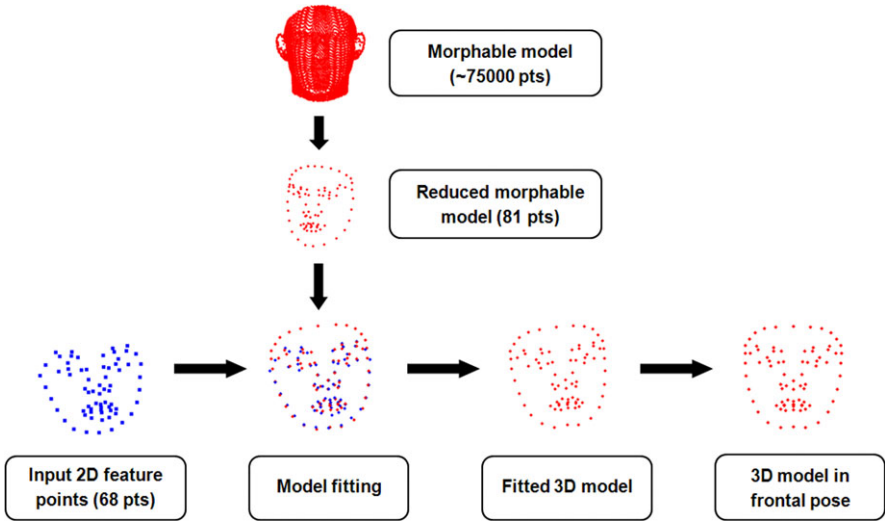


Fig. 10.1 3D model fitting process using the reduced morphable model [13]

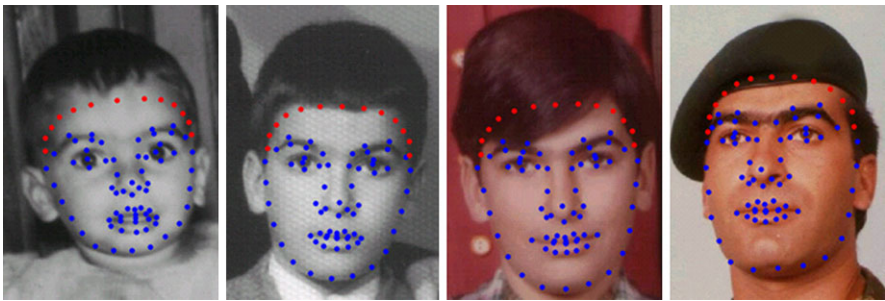


Fig. 10.2 Four example images with manually labeled 68 points (blue) and the automatically recovered 13 points (red) for the forehead region [13]

To minimize the energy term defined in (10.3), all the α_i 's are initialized to 0, the rotation matrix \mathbf{R} is set to the identity matrix and translation vector \mathbf{t} is set to 0, and the scaling factor a is set to match the overall size of the 2D and 3D shapes. Then, \mathbf{R} , \mathbf{T} , and α are iteratively updated until convergence. There are multiple ways to find the optimal pose given the current α . In these tests, it was found that first estimating the best 2×3 affine transformation followed by a QR decomposition to get the rotation works better than running a quaternion based optimization using Rodriguez's formula [17]. Note that \mathbf{t}_z is fixed to 0, as an orthogonal projection is used.

Figure 10.1 illustrates the 3D model fitting process to acquire the 3D shape. The associated texture is then retrieved by warping the 2D image. Figure 10.2 shows the manually labeled 68 points and automatically recovered 13 points that delineate the forehead region.

10.3 Aging Pattern Modeling

Following [7], the aging pattern is defined as an array of face models from a single subject indexed by the age. This model construction differs from [7] mainly in that the shape and texture are separately modeled at different ages using the shape (aging) pattern space and the texture (aging) pattern space, respectively, because the 3D shape and the texture images are less correlated than the 2D shape and texture that they used in [7]. The two pattern spaces as well as the adjustment of the 3D shape are described below.

10.3.1 Shape Aging Pattern

Shape pattern space captures the variations in the internal shape changes and the size of the face. The pose corrected 3D models obtained from the pre-processing phase are used for constructing the shape pattern space. Under age 19, the key effects of aging are driven by the increase of the cranial size, while at later ages the facial growth in height and width is very small [1]. To incorporate the growth pattern of the cranium for ages under 19, the overall size of 3D shape is rescaled according to the average anthropometric head width found in [5].

PCA is applied over all the 3D shapes, \mathbf{S}_i^j in the database irrespective of age j and subject i . All the mean subtracted \mathbf{S}_i^j are projected on to the subspace spanned by the columns of \mathbf{V}_s to obtain \mathbf{s}_i^j as

$$\mathbf{s}_i^j = \mathbf{V}_s^T (\mathbf{S}_i^j - \bar{\mathbf{S}}), \quad (10.4)$$

which is an $L_s \times 1$ vector.

Assuming that there are n subjects at m ages, the basis of the shape pattern space is then assembled as an $m \times n$ matrix with vector entries (or alternatively as an $m \times n \times L_s$ tensor), where the j th row corresponds to age j and the i th column corresponds to subject i , and the entry at (j, i) is \mathbf{s}_i^j . The shape pattern basis is initialized with the projected shapes \mathbf{s}_i^j from the face database (as shown in the third column of Fig. 10.3). Then, missing values are filled using the available values along the same column (i.e., for the same subject). Three different methods are tested for the filling process: linear, Radial Basis Function (RBF), and a variant of RBF (v -RBF). Given available ages a_i and the corresponding shape feature vectors s_i , a missing feature value s_x at age a_x can be estimated by $s_x = l_1 \times s_1 + l_2 \times s_2$ in linear interpolation, where s_1 and s_2 are shape feature vectors corresponding to the ages a_1 and a_2 that are closest from a_x , and l_1 and l_2 are weights inversely proportional to the distance from a_x to a_1 and a_2 . In the v -RBF process, each feature is replaced by a weighted sum of all available features as $s_x = \sum_i \phi(a_x - a_i) s_i / (\sum \phi(a_x - a_i))$, where $\phi(\cdot)$ is a RBF function defined by a Gaussian function. In the RBF method, the mapping function from age to the shape feature vector is calculated by $s_x =$

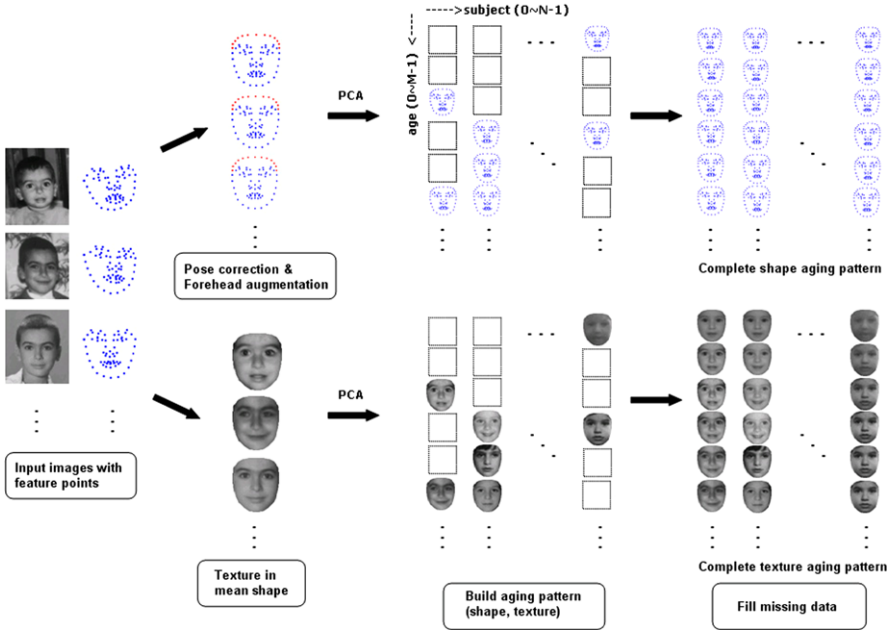


Fig. 10.3 3D aging model construction [13]

$\sum_i r_i \phi(a_x - a_i) / (\sum \phi(a_x - a_i))$ for each available age and feature vector a_i and s_i , where r_i 's are estimated based on the known scattered data. Any missing feature vector s_x at age x can thus be obtained.

The shape aging pattern space is defined as the space containing all the linear combinations of the patterns of the following type (expressed in PCA basis):

$$s_{w_s}^j = \bar{s}^j + \sum_{i=1}^n (s_i^j - \bar{s}^j) w_{s,i}, \quad 0 \leq j \leq m - 1. \tag{10.5}$$

Note that the weight w_s in the linear combination above is not unique for the same aging pattern. The regularization term can be used in the aging simulation described below to resolve this issue. Given a complete shape pattern space, mean shape \bar{S} and the transformation matrix V_s , the shape aging model with weight w_s is defined as

$$S_{w_s}^j = \bar{S} + V_s s_{w_s}^j, \quad 0 \leq j \leq m - 1. \tag{10.6}$$

10.3.2 Texture Aging Pattern

The texture pattern T_i^j for subject i at age j is obtained by mapping the original face image to frontal projection of the mean shape \bar{S} followed by a column-wise

concatenation of the image pixels. After applying PCA on T_i^j , the transformation matrix V_i and the projected texture t_i^j are calculated. The same filling procedure is used as in the shape pattern space to construct the complete basis for the texture pattern space using t_i^j . A new texture $\mathbf{T}_{w_t}^j$ can be similarly obtained, given an age j and a set of weights w_t as

$$\mathbf{t}_{w_t}^j = \bar{\mathbf{t}}^j + \sum_{i=1}^n (\mathbf{t}_i^j - \bar{\mathbf{t}}^j) w_{t,i}, \quad 0 \leq j \leq m-1, \quad (10.7)$$

$$\mathbf{T}_{w_t}^j = \bar{\mathbf{T}} + \mathbf{V}_t \mathbf{t}_{w_t}^j, \quad 0 \leq j \leq m-1. \quad (10.8)$$

Figure 10.3 illustrates the aging model construction process for both shape and texture pattern spaces.

10.3.3 Separate and Combined Shape & Texture Modeling

Given s_i^j and t_i^j , they can be used directly for the aging modeling or another step of PCA on the new concatenated feature vector $C_i^j = [s_i^{jT} \ t_i^{jT}]^T$ can be applied. Applying PCA on C_i^j will generate a set of new Eigen vectors, c_i^j [3]. The modeling using s_i^j and t_i^j is called as “separate shape and texture modeling” and c_i^j as a “combined shape and texture modeling.”

10.4 Aging Simulation

Given a face image of a subject at a certain age, aging simulation involves the construction of the face image of that subject adjusted to a different age. Given a 2D image at age x , the 3D shape, $\mathbf{S}_{\text{new}}^x$ and the texture T_{new}^x are first produced by following the preprocessing step described in Sect. 10.2, and then they are projected to the reduced spaces to get $\mathbf{s}_{\text{new}}^x$ and $\mathbf{t}_{\text{new}}^x$. Given a reduced 3D shape $\mathbf{s}_{\text{new}}^x$ at age x , a weighting vector, w_s , that generates the closest possible weighted sum of the shapes at age x , can be obtained as:

$$\hat{w}_s = \underset{c_- \leq w_s \leq c_+}{\operatorname{argmin}} \left\| \mathbf{s}_{\text{new}}^x - \mathbf{s}_{w_s}^x \right\|^2 + r_s \|w_s\|^2, \quad (10.9)$$

where r_s is the weight of a regularizer to handle the cases when multiple solutions are obtained or when the linear system used to obtain the solution has a large condition number. Each element of weight vector, $w_{s,i}$ is constrained within $[c_-, c_+]$ to avoid strong domination by a few shape basis vectors.

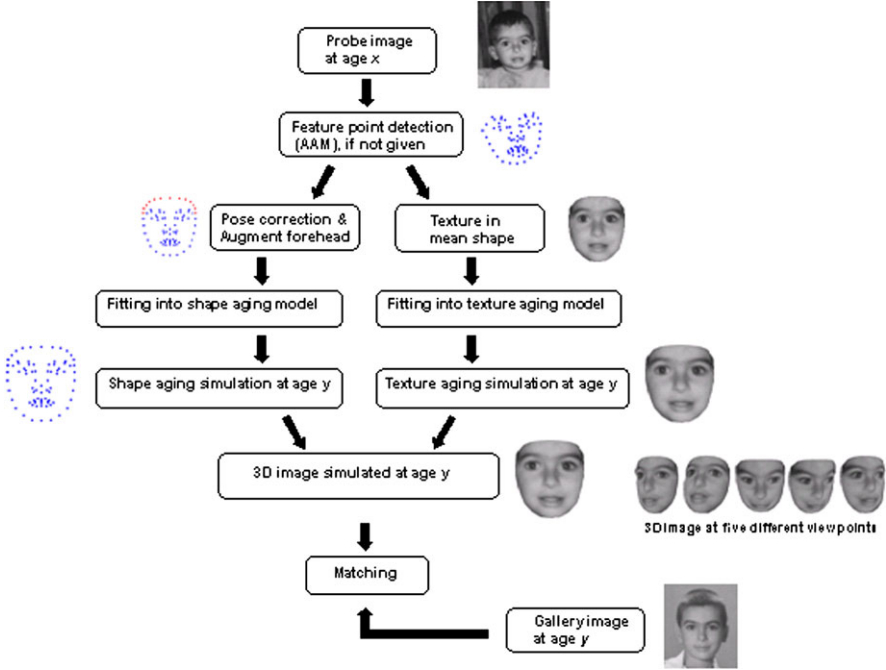


Fig. 10.4 Aging simulation from age x to y [13]

Given \hat{w}_s , age adjusted shape can be obtained at age y by carrying \hat{w}_s over to the shapes at age y and transforming the shape descriptor back to the original shape space as

$$S_{\text{new}}^y = S_{\hat{w}_s}^y = \bar{S} + V_s s_{\hat{w}_s}^y. \quad (10.10)$$

The texture simulation process is similarly performed by first estimating \hat{w}_t as

$$\hat{w}_t = \underset{c_- \leq w_t \leq c_+}{\operatorname{argmin}} \left\| \mathbf{t}_{\text{new}}^x - \mathbf{t}_{w_t}^x \right\|^2 + r_t \|w_t\|^2, \quad (10.11)$$

and then, propagating the \hat{w}_t to the target age y followed by the back projection to get

$$T_{\text{new}}^y = \mathbf{T}_{\hat{w}_t}^y = \bar{\mathbf{T}} + \mathbf{V}_t \mathbf{t}_{\hat{w}_t}^y. \quad (10.12)$$

The aging simulation process is illustrated in Fig. 10.4. Figure 10.5 shows an example of aging simulated face images from a subject at age 2 in the FG-NET database. Figure 10.6 exhibits the example input images, feature point detection, pose-corrected and age-simulated images from a subject in the MORPH database. The pseudocodes of shape aging pattern space construction and simulation are given in Algorithms 10.1, 10.2, 10.3, and 10.4.

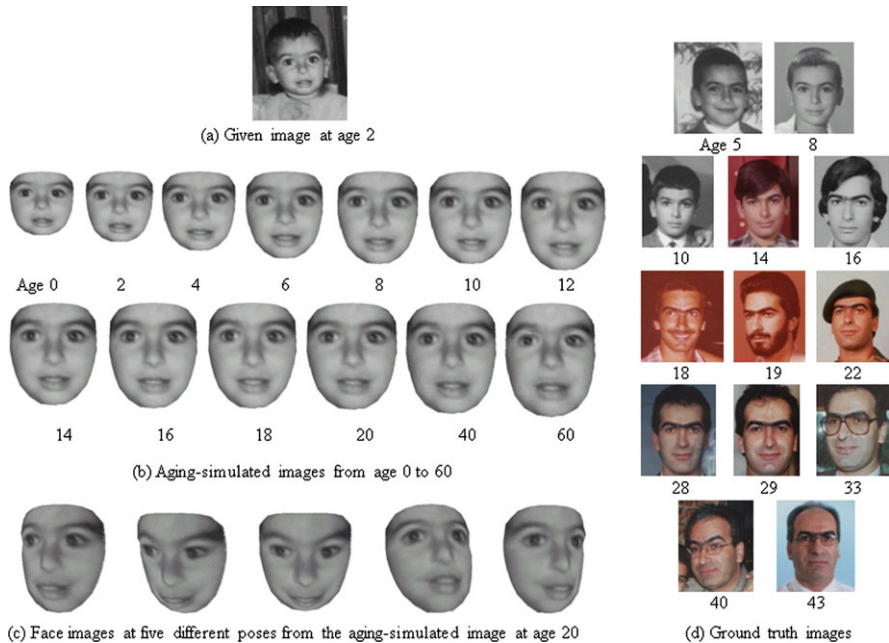


Fig. 10.5 An example aging simulation in the FG-NET database [13]

10.5 Experimental Results

10.5.1 Database

There are two well known public domain databases to evaluate facial aging models: FG-NET [6] and MORPH [21]. The FG-NET database contains 1002 face images of 82 subjects (~ 12 images/subject) at different ages, with the minimum age being 0 (< 12 months) and the maximum age being 69. There are two separate databases in MORPH: Album1 and Album2. MORPH-Album1 contains 1690 images from 625 different subjects (~ 2.7 images/subject). MORPH-Album2 contains 15 204 images from 4039 different subjects (~ 3.8 images/subject). Another source of facial aging data can be found in the book by Nixon and Galassi [11]. This is a collection of pictures of four sisters taken every year over a period of 33 years from 1975 to 2007. A new database, called, “BROWNS” are constructed by scanning 132 pictures of the four subjects (33 per subject) from the book to evaluate the aging model. Since it is desirable to have as many subjects and as many images at different ages per subject as possible, the FG-NET database is more useful for aging modeling than MORPH or BROWNS. The age separation observed in MORPH-Album1 is in the range 0–30 and that in MORPH-Album2 is less than 5. Therefore, MORPH-Album1 is more useful in evaluating the aging model than MORPH-Album2. A subset of MORPH-Album1, 1655 images of all the 612 subjects whose images at different ages are available, is used for the experiments. The complete FG-NET database has been

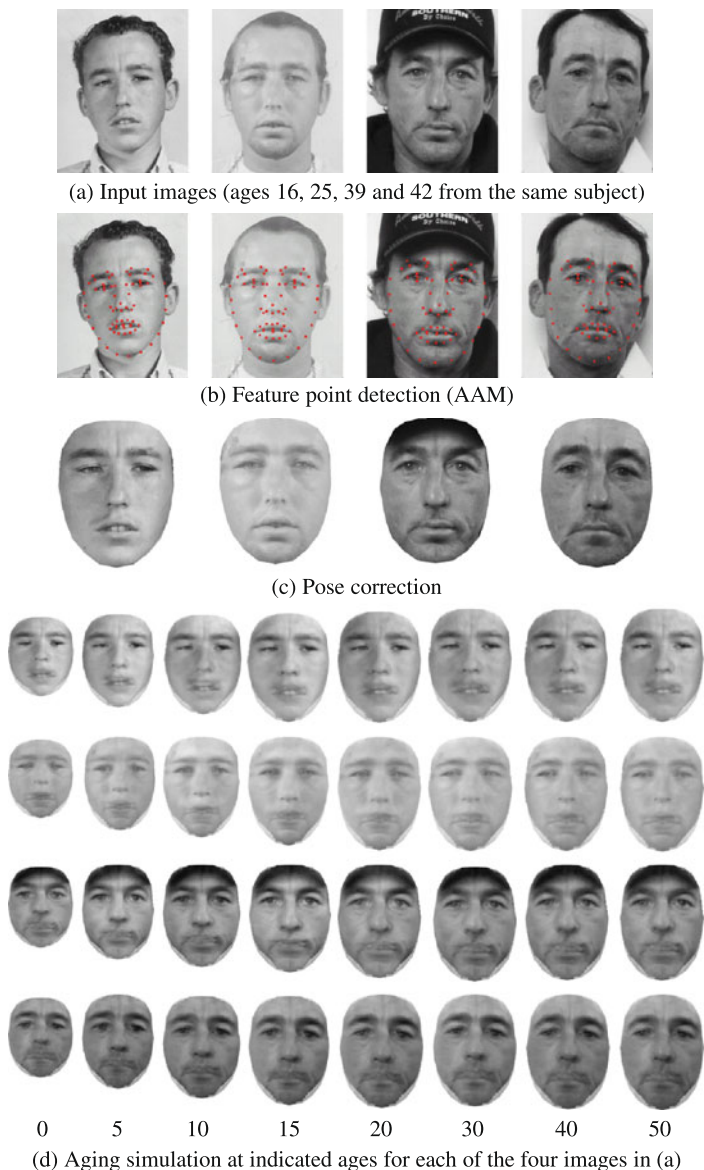


Fig. 10.6 Example aging simulation process in the MORPH database [13]

used for model construction and then it is evaluated on FG-NET (in a leave-one-person-out fashion), MORPH-Album1 and BROWNS. Figure 10.7 shows multiple sample images of one subject from each of the three databases. The number of subjects, number of images, and number of images at different ages per subject for the three databases used in the aging study [13] are summarized in Table 10.2.

Algorithm 10.1: 3D SHAPE AGING PATTERN CONSTRUCTION()

Input: $S_{2d} = \{S_{1,2d}^0, \dots, S_{i,2d}^j, \dots, S_{n,2d}^{m-1}\}$

Output: $s_i^j, i = 1, \dots, n, j = 0, \dots, m - 1$

$i \leftarrow 1, j \leftarrow 0$

while $i \leq n$ & $j \leq m - 1$

if $S_{i,2d}^j$ is available
 $k \leftarrow 1, E \leftarrow$ fitting error between $S_{i,2d}^j$ and S_α
 while $k < \tau$ & $E < \theta$
 do $\left\{ \begin{array}{l} \text{update pose } (a, R, t) \text{ (3D model parameters, } \alpha, \text{ fixed)} \\ \text{do } \left\{ \begin{array}{l} \text{update 3D model parameters (pose fixed)} \\ k \leftarrow k + 1, \text{ update } E \end{array} \right. \\ S_i^j \leftarrow S_\alpha \end{array} \right.$

 Calculate eigenvalue λ_s and eigenvector and \mathbf{V}_s from $S_i^j - \bar{S}$

$i \leftarrow 1, j \leftarrow 0$

while $i \leq n$ & $j \leq m - 1$

do $\left\{ \begin{array}{l} \text{if } S_i^j \text{ is available} \\ s_i^j \leftarrow \mathbf{V}^T(S_i^j - \bar{S}) \\ \text{Fill } (i, j)\text{-th shape pattern by } s_i^j \\ \text{else Fill } (i, j)\text{-th shape pattern space, using interpolation along the column} \end{array} \right.$

Algorithm 10.2: TEXTURE AGING PATTERN CONSTRUCTION()

Input: $S = \{S_1^0, \dots, S_i^j, \dots, S_n^{m-1}\}, T = \{T_1^0, \dots, T_i^j, \dots, T_n^{m-1}\},$

Pose = $\{P_1^1, \dots, P_i^j, \dots, P_n^m\}$

Output: $t_i^j, i = 1, \dots, n, j = 0, \dots, m - 1$

Construct mean shape \bar{S}

$i \leftarrow 1, j \leftarrow 0$

while $i \leq n$ & $j \leq m - 1$

do $\left\{ \begin{array}{l} \text{if } T_{i,2d}^j \text{ is available} \\ \text{Warp texture } T_i^j \text{ from } S_i^j \text{ with pose } P_i^j \text{ to } \bar{S} \end{array} \right.$

 Calculate eigenvalue λ_t and eigenvector \mathbf{V}_t from $(T_i^j - \bar{T})$

$i \leftarrow 1, j \leftarrow 0$

while $i \leq n$ & $j \leq m - 1$

do $\left\{ \begin{array}{l} \text{if } T_i^j \text{ is available} \\ t_i^j \leftarrow \mathbf{V}^T(T_i^j - \bar{T}) \\ \text{Fill } (i, j)\text{-th texture pattern by } t_i^j \\ \text{else Fill } (i, j)\text{-th texture pattern space, using interpolation along} \\ \text{the column} \end{array} \right.$

Algorithm 10.3: AGE SIMULATION FOR SHAPE()

Input: $\mathbf{s} = \{s_1^0, \dots, s_n^{m-1}\}, S_{\text{new}}^x$
Output: S_{new}^y
 Estimate w_s by (10.9)
 Calculate S_{new}^y by (10.10)

Algorithm 10.4: AGE SIMULATION FOR TEXTURE()

Input: $\mathbf{t} = \{t_1^0, \dots, t_n^{m-1}\}, T_{\text{new}}^x$
Output: T_{new}^y
 Estimate w_t by (10.11)
 Calculate T_{new}^y by (10.12)

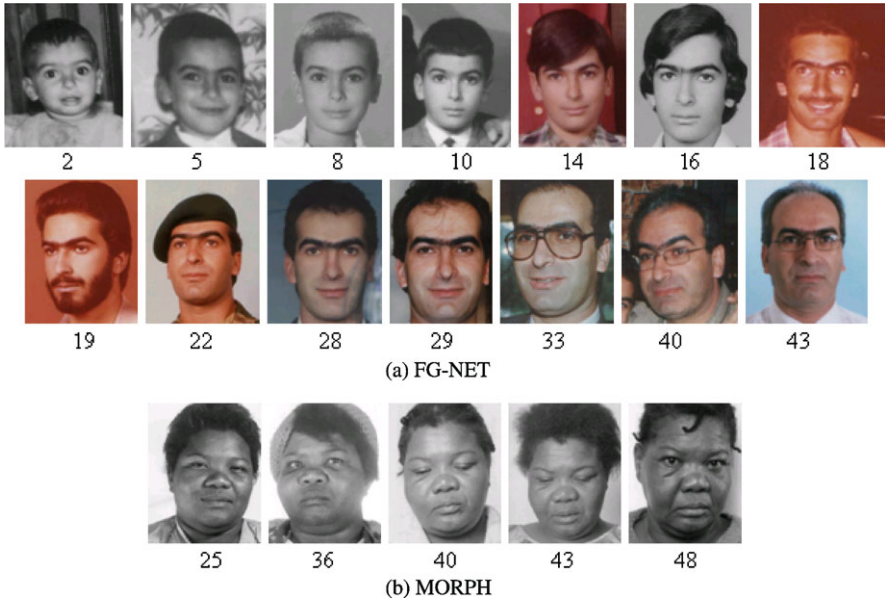


Fig. 10.7 Example images in **a** FG-NET and **b** MORPH databases. Multiple images of one subject in each of the three databases are shown at different ages. The age value is given below each image

10.5.2 Face Recognition Tests

The performance of the aging model is evaluated by comparing the face recognition accuracy of a state-of-the-art matcher before and after aging simulation. The probe set, $P = \{p_1^{x_1}, \dots, p_n^{x_n}\}$, is constructed by selecting one image $p_i^{x_i}$ for each subject i at age x_i in each database, $i \in \{1, \dots, n\}$, $x_i \in \{0, \dots, 69\}$. The gallery set $G =$

Table 10.2 Databases used in aging modeling [13]

Database		#subjects	#images	Average #images per subject
FG-NET		82	1002	12
MORPH	Album1	625	1690	2.7
	Album2	4039	15 204	3.8
BROWNS		4	132	33

Table 10.3 Probe and gallery data used in face recognition tests [13]

Database	Probe			Gallery		
	#images	#subjects	Age group	#images	#subjects	Age group
FG-NET	82	82	{0, 5, ..., 30}	82	82	$x^* + \{5, 10, \dots, 30\}$
MORPH	612	612	{15, 20, ..., 30}	612	612	$x + \{5, 10, \dots, 30\}$
BROWNS	4	4	{15, 20, ..., 30}	100	100	$x + \{5, 10, \dots, 30\}$

* x is the age of the probe group

$\{g_1^{y_1}, \dots, g_n^{y_n}\}$ is similarly constructed. A number of different probe and gallery age groups are also constructed from the three databases to demonstrate the model's effectiveness in different periods of the aging process.

Aging simulation is performed in both aging and de-aging directions for each subject i in the probe and each subject j in the gallery as $(x_i \rightarrow y_j)$ and $(y_j \rightarrow x_i)$. Table 10.3 summarizes the probe and gallery data sets used in the face recognition test [13].

Let P , P_f and P_a denote the probe, the pose-corrected probe, and the age-adjusted probe set, respectively. Let G , G_f and G_a denote the gallery, the pose-corrected gallery, and age-adjusted gallery set, respectively. All age-adjusted images are generated (in a leave-one-person-out fashion for FG-NET) using the shape and texture pattern space. The face recognition test is performed on the following probe-gallery pairs: $P-G$, $P-G_f$, P_f-G , P_f-G_f , P_a-G_f and P_f-G_a . The identification rate for the probe-gallery pair $P-G$ is the performance on original images without applying any aging model. The accuracy obtained by fusion of $P-G$, $P-G_f$, P_f-G and P_f-G_f matchings is regarded as the performance after pose correction. The accuracy obtained by fusion of all the pairs $P-G$, $P-G_f$, P_f-G , P_f-G_f , P_a-G_f and P_f-G_a represents the performance after aging simulation. A simple score-sum based fusion is used in all the experiments.

10.5.3 Effects of Different Cropping Methods

The performance of the face recognition system is evaluated with different face cropping methods. An illustration of the cropping results obtained by different approaches is shown in Fig. 10.8. The first column shows the input face image and

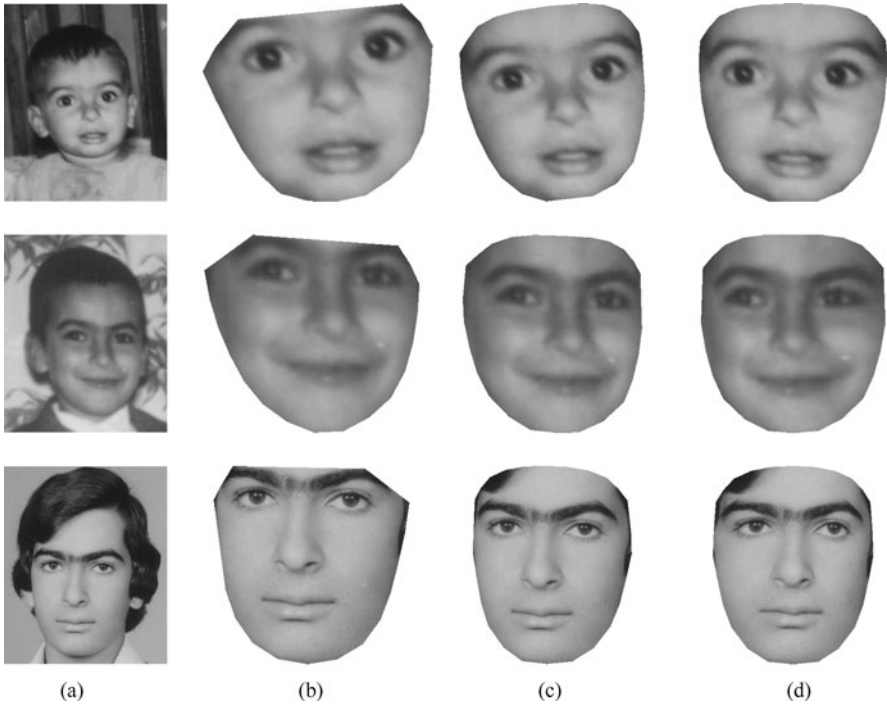


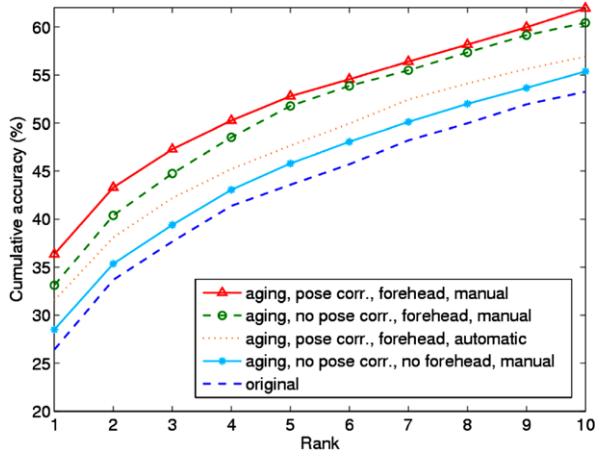
Fig. 10.8 Example images showing different face cropping methods: **a** original, **b** no-forehead and no pose correction, **c** no pose correction with forehead, **d** pose correction with forehead [13]

the second column shows the cropped face obtained using the 68 feature points provided in the FG-NET database without pose correction. The third column shows the cropped face obtained with the additional 13 points (total 81 feature points) for forehead inclusion without any pose correction. The last column shows the cropping obtained by the 81 feature points, with pose correction.

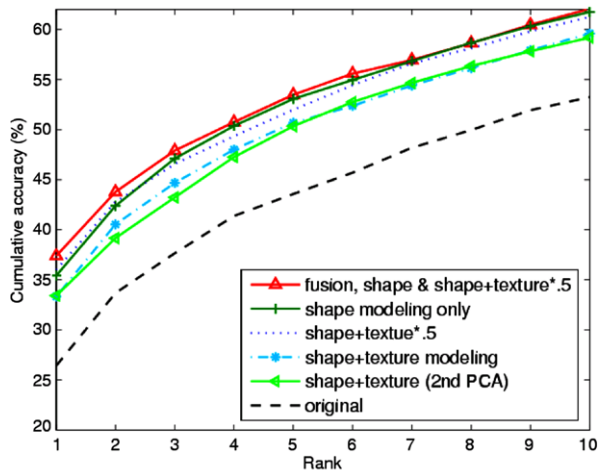
Figure 10.9(a) shows the face recognition performance on FG-NET using only shape modeling based on different face cropping methods and feature point detection methods. Face images with pose correction that include the forehead show the best performance. This result shows that the forehead does influence the face recognition performance, although it has been a common practice to remove the forehead in AAM based feature point detection and subsequent face modeling [3, 8, 26]. Therefore, the aging simulation is evaluated with the model that contains the forehead region with pose correction.

Note that, the performance difference between nonfrontal and frontal pose is as expected, and that the performance using automatically detected feature points is lower than that of manually labeled feature points. However, the performance with automatic feature point detection is still better than that of matching original images before applying the aging modeling.

Fig. 10.9 Cumulative Match Characteristic (CMC) curves with different methods of face cropping and shape & texture modeling [13]



(a) CMC with different methods of face cropping.



(b) CMC with different methods of shape & texture modeling.

10.5.4 Effects of Different Strategies in Employing Shape and Texture

Most of the existing face aging modeling techniques use either only shape or a combination of shape and texture [7, 8, 14, 20, 26]. Park et al. have tested the aging model with shape only, separate shape and texture, and combined shape and texture modeling. In the test of the combined scheme, shape and the texture are concatenated and a second stage of principle component analysis is applied to remove the possible correlation between shape and texture as in the AAM face modeling technique.

Figure 10.9(b) shows the face recognition performance of different approaches to shape and texture modeling. Consistent performance drop has been observed in face recognition performance when the texture is used together with the shape. The best performance is observed by combining shape modeling and shape + texture modeling using score level fusion. When simulating the texture, the aging simulated texture and the original texture have been blended with equal weights. Unlike shape, texture is a higher dimensional vector that can easily deviate from its original identity after the aging simulation. Even though performing aging simulation on texture produces more realistic face images, it can easily lose the original face-based identity information. The blending process with the original texture reduces the deviation and generates better recognition performance. In Fig. 10.9(b), shape + texture modeling represent separate modeling of shape and texture, shape + texture $\times 0.5$ represents the same procedure but with the blending of the simulated texture with the original texture. The fusion of shape and shape + texture $\times 0.5$ strategy is used for the following aging modeling experiments.

10.5.5 Effects of Different Filling Methods in Model Construction

Park et al. tried a few different methods of filling missing values in aging pattern space construction (see Sect. 10.3.1): linear, v -RBF, and RBF. The rank-one accuracies are obtained as 36.12%, 35.19%, and 36.35% in shape + texture $\times 0.5$ modeling method for linear, v -RBF, and RBF methods, respectively. The linear interpolation method is used in the rest of the experiments for the following reasons: (i) performance difference is minor, (ii) linear interpolation is computationally efficient, and (iii) the calculation of RBF based mapping function can be ill-posed.

Figure 10.10 provides the Cumulative Match Characteristic (CMC) curves with original, pose-corrected and aging simulated images in FG-NET, MORPH and BROWNS, respectively. It can be seen that there are significant performance improvement after aging modeling and simulation in all the three databases. The amount of improvement due to aging simulation is more or less similar with those of other studies as shown in Table 10.1. However, Park et al. used FaceVACS, a state-of-the-art face matcher, which is known to be more robust against internal and external facial variations (e.g., pose, lighting, expression, etc.) than simple PCA based matchers. They argued that the performance gain using FaceVACS is more realistic than that of a PCA matcher reported in other studies. Further, unlike other studies, they have used the entire FG-NET and MORPH-Album1 in the experiments. Another unique attribute of their studies is that the model is built on FG-NET and then evaluated on independent databases MORPH and BROWNS.

Figure 10.11 presents the rank-one identification accuracy for each of the 42 different age pair groups of probe and gallery in the FG-NET database. The aging process can be separated as growth and development (age ≤ 18) and adult aging process (age > 18). The face recognition performance is somewhat lower in the growth process where more changes occur in the facial appearance. However, the

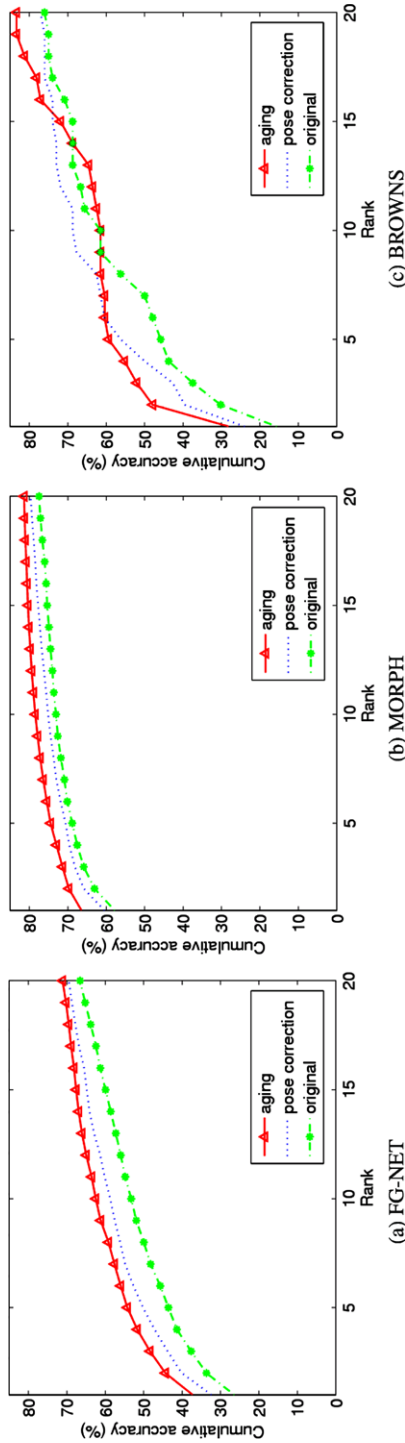


Fig. 10.10 Cumulative Match Characteristic (CMC) curves [13]

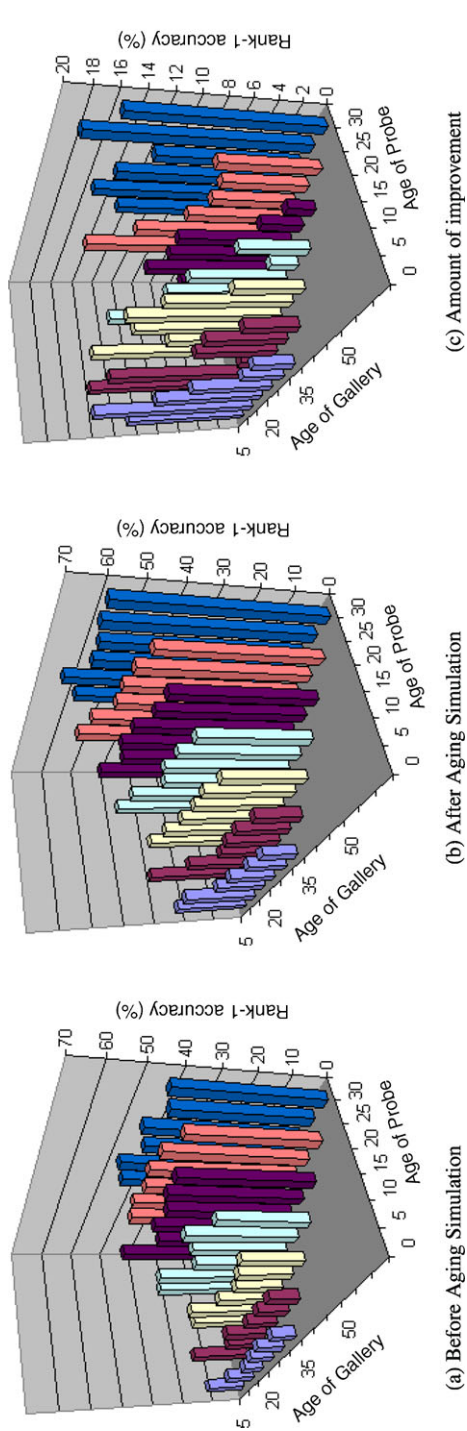


Fig. 10.11 Rank-one identification accuracy for each probe and gallery age group: **a** before aging simulation, **b** after aging simulation, and **c** the amount of improvements after aging simulation [13]

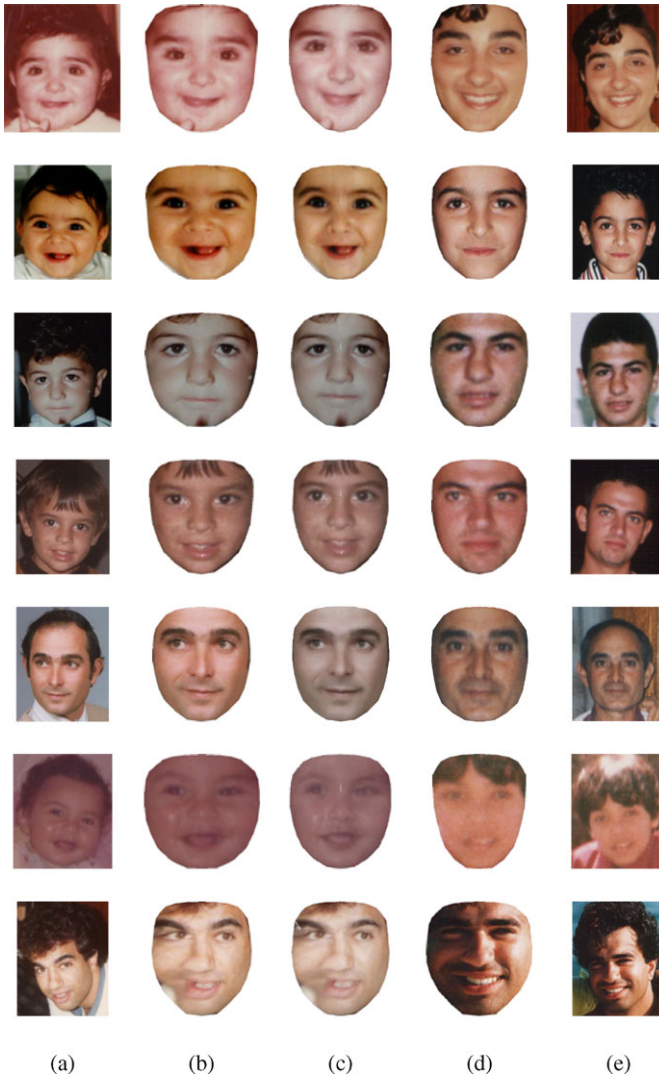


Fig. 10.12 Example matching results before and after aging simulation for seven different subjects: **a** probe, **b** pose-corrected probe, **c** age-adjusted probe, **d** pose-corrected gallery, and **e** gallery. The ages in each (probe, gallery) pair are (0, 18), (0, 9), (4, 14), (3, 20), (30, 49), (0, 7) and (23, 31), respectively, from the top to the bottom row [13]

aging process provides performance improvements in both of the age groups, ≤ 18 and > 18 . The average recognition results for age groups ≤ 18 are improved from 17.3% to 24.8% and those for age groups > 18 are improved from 38.5% to 54.2%.

Matching results for seven subjects in FG-NET are demonstrated in Fig. 10.12. The face recognition fails without aging simulation for all these subjects but succeeds with aging simulations for the first five of the seven subjects. The aging sim-

ulation fails to provide correct matchings for the last two subjects, possibly due to poor texture quality (for the sixth subject) or large pose and illumination variation (for the seventh subject).

The aging model construction takes about 44 s. The aging model is constructed off-line, therefore its computation time is not a major concern. In the recognition stage, the entire process, including feature points detection, aging simulation, enrollment and matching takes about 12 s per probe image. Note that the gallery images are preprocessed off-line. All computation times are measured on a Pentium 4, 3.2 GHz, 3 G-Byte RAM machine.

10.6 Conclusions

A 3D facial aging model and simulation method for age-invariant face recognition has been described. It is shown that extension of shape modeling from 2D to 3D domain gives additional capability of compensating for pose and, potentially, lighting variations. Moreover, the use of 3D model appears to provide a more powerful modeling capability than 2D age modeling because the changes in human face configuration occur primarily in 3D domain. The aging model has been evaluated using a state-of-the-art commercial face recognition engine (FaceVACS). Face recognition performances have been improved on three different publicly available aging databases. It is shown that the method is capable of handling both growth and developmental adult face aging effects.

Exploring different (nonlinear) methods for building aging pattern space, given noisy 2D or 3D shape and texture data, with cross validation of the aging pattern space and aging simulation results in terms of face recognition performance can further improve simulated aging. Age estimation is crucial if a fully automatic age invariant face recognition system is needed.

References

1. Albert, A., Ricanek, K., Patterson, E.: The aging adult skull and face: A review of the literature and report on factors and processes of change. UNCW Technical Report, WRG FSC-A (2004)
2. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3d faces. In: Proc. Conf. on Computer Graphics and Interactive Techniques (SIGGRAPH), pp. 187–194 (1999)
3. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 681–685 (2001)
4. FaceVACS Software Developer Kit, Cognitec Systems GmbH. <http://www.cognitec-systems.de>
5. Farkas, L.G. (ed.): *Anthropometry of the Head and Face*. Lippincott Williams & Wilkins, Baltimore (1994)
6. FG-NET Aging Database. <http://www.fgnet.rsunit.com>
7. Geng, X., Zhou, Z.-H., Smith-Miles, K.: Automatic age estimation based on facial aging patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 2234–2240 (2007)
8. 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)

9. Lanitis, A., Draganova, C., Christodoulou, C.: Comparing different classifiers for automatic age estimation. *IEEE Trans. Syst. Man Cybern., Part B, Cybern.* **34**(1), 621–628 (2004)
10. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.: A study of face recognition as people age. In: *Proc. IEEE Int. Conf. on Computer Vision (ICCV)*, pp. 1–8 (2007)
11. Nixon, N., Galassi, P.: *The Brown Sisters, Thirty-Three Years*. The Museum of Modern Art, NY, USA (2007)
12. O’Toole, A., Vetter, T., Volz, H., Salter, E.: Three-dimensional caricatures of human heads: distinctiveness and the perception of facial age. *Perception* **26**, 719–732 (1997)
13. Park, U., Tong, Y., Jain, A.K.: Age invariant face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(5), 947–954 (2010)
14. Patterson, E., Ricanek, K., Albert, M., Boone, E.: Automatic representation of adult aging in facial images. In: *Proc. Int. Conf. on Visualization, Imaging, and Image Processing, IASTED*, pp. 171–176 (2006)
15. Patterson, E., Sethuram, A., Albert, M., Ricanek, K., King, M.: Aspects of age variation in facial morphology affecting biometrics. In: *Proc. IEEE Conf. on Biometrics: Theory, Applications, and Systems (BTAS)*, pp. 1–6 (2007)
16. Phillips, P.J., Scruggs, W.T., O’Toole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., Sharpe, M.: FRVT 2006 and ICE 2006 large-scale results. Technical Report NISTIR 7408, National Institute of Standards and Technology
17. Pighin, F., Szeliski, R., Salesin, D.H.: Modeling and animating realistic faces from images. *Int. J. Comput. Vis.* **50**(2), 143–169 (2002)
18. Pittenger, J.B., Shaw, R.E.: Aging faces as viscal-elastic events: Implications for a theory of nonrigid shape perception. *J. Exp. Psychol. Hum. Percept. Perform.* **1**, 374–382 (1975)
19. Ramanathan, N., Chellappa, R.: Face verification across age progression. In: *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 462–469 (2005)
20. Ramanathan, N., Chellappa, R.: Modeling age progression in young faces. In: *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 387–394 (2006)
21. Ricanek, K.J., Tesafaye, T.: Morph: A longitudinal image database of normal adult age-progression. In: *Proc. Int. Conf. on Automatic Face and Gesture Recognition (FG)*, pp. 341–345 (2006)
22. Scherbaum, K., Sunkel, M., Seidel, H.-P., Blanz, V.: Prediction of individual non-linear aging trajectories of faces. *Comput. Graph. Forum* **26**(3), 285–294 (2007)
23. Stegmann, M.B.: The AAM-API: An open source active appearance model implementation. In: *Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 951–952 (2003)
24. Suo, J., Min, F., Zhu, S., Shan, S., Chen, X.: A multi-resolution dynamic model for face aging simulation. In: *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8 (2007)
25. Thompson, D.W.: *On growth and form* (1992)
26. Wang, J., Shang, Y., Su, G., Lin, X.: Age simulation for face recognition. In: *Proc. Int. Conf. on Pattern Recognition (ICPR)*, pp. 913–916 (2006)