

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/220057621>

# Comparative Evaluation of Automatic Age Progression Methodologies

Article in *Journal on Advances in Signal Processing* · March 2008

DOI: 10.1155/2008/239480 · Source: DBLP

---

CITATIONS

63

READS

2,100

1 author:



**Andreas Lanitis**

Cyprus University of Technology

125 PUBLICATIONS 4,156 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



VR-CAVE [View project](#)



Face Recogniion - Age Progression [View project](#)

## Research Article

# Comparative Evaluation of Automatic Age-Progression Methodologies

**Andreas Lanitis**

*Department of Multimedia and Graphic Arts, Cyprus University of Technology, P.O. Box 5032, 3603 Lemesos, Cyprus*

Correspondence should be addressed to Andreas Lanitis, andreas.lanitis@cut.ac.cy

Received 1 June 2007; Revised 31 October 2007; Accepted 8 January 2008

Recommended by C.-C. Jay Kuo

Automatic age-progression is the process of modifying an image showing the face of a person in order to predict his/her future facial appearance. In this paper, we compare the performance of two age-progression methodologies reported in the literature against two novel approaches to the problem. In particular, we compare the performance of a method based on age prototypes, a method based on aging functions defined in a low-dimensional parametric model space, and two methods based on the distributions of samples belonging to different individuals and different age groups. Quantitative comparative results reported in the paper are based on dedicated performance evaluation metrics that assess the ability of each method to produce accurate predictions of the future/previous facial appearance of subjects. The framework proposed in this paper promotes the idea of a standardized performance evaluation protocol for age-progression methodologies, using images from a publicly available image database.

Copyright © 2008 Andreas Lanitis. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. INTRODUCTION

Age-progression is the process of deforming the facial appearance of a subject shown in an image, in order to predict how the face of the person will look like in the future. The ability to produce accurate age-progressed images is important in a number of key applications including the localization of missing children, the development of age-invariant face recognition systems, and automatic update of photographs appearing in smart documents (i.e., smart id cards).

Traditionally age progressed images are produced by forensic artists [1, 2] who use dedicated photo-editing software. The process of computer-assisted age progression involves the modification of the shape and texture of a person's face in order to reflect cross-population age-related trends (i.e., changes in face shape and introduction of wrinkles), coupled with person-specific transformations. In this context, person-specific transformations are defined by the examination of the aging pattern adopted by close relatives of the subject. The time and cost involved in generating manually age-progressed face images restrict the use of this technology in real-life applications.

A number of researchers describe automatic age-progression systems that can be used for generating age-progressed images without the need for human intervention. In most studies involving automatic age-progression methodologies, authors demonstrate the effectiveness of a method by showing few examples of raw face images and the corresponding age-progressed images. However, the main issue to be addressed is not the production of aesthetically pleasing results, but the generation of accurate predictions of the future facial appearance of the subject to be age-progressed. In this paper, we describe an experimental evaluation procedure that aims to assess the performance of different age-progression methodologies in terms of the accuracy of age-progressed images. In particular, we compare four age-progression methods: a method based on age prototypes [3], a method based on aging functions [4], a distance-based method, and a method based on support vector machines (SVMs) [5]. Both the distance-based method and the SVM-based method rely on the distribution of samples for different individuals and different age groups. All experiments are performed using the FG-NET Aging database [6], which is publicly available.

Our work on comparing the performance of several age-progression methodologies is among the first efforts to standardize the process of evaluating age-progression algorithms, so that it will be feasible to obtain comparative results related to the performance of different age-progression algorithms reported in the literature. An important aspect of our work is the formulation of dedicated performance evaluation metrics that can be used for assessing the accuracy of age-progressed images.

The remainder of the paper is organized as follows: in Section 2, we present an overview of the relevant literature; in Section 3, we describe the image database used in our experiments; and in Section 4, we describe the age-progression methods to be evaluated. The experimental evaluation procedure and results obtained are presented in Section 5. Conclusions and plans for future work are outlined in Section 6.

## 2. LITERATURE REVIEW

Rowland and Perret [3] propose an age-progression method based on age prototypes. Age prototypes are generated by averaging faces belonging to the same age group. Age-progression is achieved by adding on a given face image, differences between age prototypes corresponding to different ages. A limitation of this approach is the elimination of high-frequency artifacts (i.e., wrinkles) during the calculation of age prototypes resulting in smoothed age prototypes and consequently in smoothed age-progressed images. In order to overcome this problem, methods for adding high-frequency details on age-progressed images [7, 8] have been reported. Since age-progression based on age prototypes is one of the methods under evaluation, a more thorough description of the method is provided in Section 4.

Lee et al. [9] describe an age-progression methodology based on merging 3D face models of young subjects with 3D models of older members belonging to the same family. Wrinkles are also added on the blended images using a method that mimics the physics of wrinkle generation. The main problem of age-progression systems based on image merging is the dependency of the results on the composite image showing the older subject, since in effect aging characteristics of the older subject are transferred to the younger face.

Boissieux et al. [10] generate typical wrinkle prototypes for different groups of faces. For example, they generate wrinkle prototypes for males and females with different expressions. Based on biochemical data, they define the relationship between age and wrinkle strength so that it is possible to define the strength of wrinkles to be added on faces according to the target age. Given a face to be age-progressed, the most appropriate wrinkle prototype is chosen so that a given face is age-progressed by adding wrinkles of appropriate strength from the optimum wrinkle prototype.

Leta et al. [11, 12] establish the correlation of 26 facial distance measurements with age so that it is possible to predict how the 26 distance measurements are modified, as a person grows older. Image warping techniques are used for modifying the shape of a face in order to inflict age-related

shape deformations specified by the modification of the 26 distance measurements.

Lanitis et al. [4] propose a model-based age-progression methodology. In this context, parametric functions that relate the model-based representation of faces and ages are established and used as the basis for implementing age estimation and age-progression. Because the parametric representation of faces discards high-frequency information, this approach is not ideal for generating high-resolution age-progressed images. This approach is more appropriate for modeling distinct age-related facial transformations like shape variations and major texture variation. A number of researchers also describe methods based on aging functions defined either in relation with 2D faces [13, 14] or 3D faces [15, 16].

In a more recent approach, Scandrett et al. [17] define both a personal and a consensus aging axis so that age-progression is achieved as a compounded effect of both person-specific and global aging trends. The influence of each axis during the age-progression process is determined by maximizing the probability that an age-progressed face belongs to two different distributions: the distribution of faces at the target age and the distribution of differences between age-progressed samples and the actual faces of the same subjects at the target age. Both visual and quantitative results demonstrate the effectiveness of the method. The distance-based and SVM-based age-progression methods presented in this paper bear similarities with the method proposed by Scandrett et al. [17], since in both cases there is an attempt of reinforcing both person specific and global aging trends based on the corresponding distributions in a low-dimensional face space. However, in the approach reported by Scandrett et al. [17], the formulation of the personal aging axis requires a number of training samples showing the same person at younger ages. In our case, the use of person specific trends is based on a single image of the subject in question.

Ramanathan and Chellappa [18] estimate the differences between pairs of images showing the same individual at different ages and pairs of images showing different individuals. Based on the statistical distribution of the difference vectors, they demonstrate that it is possible to determine whether a pair of face images showing subjects at different ages belongs to the same person. Although the work described in [18] is mainly related to age estimation and face identification rather than age-progression, we include it in our review since the treatment of the topic bears similarities with the distance-based age-progression technique reported in this paper.

A number of researchers [19, 20] use a coordinate transformation called the “cardioid strain transformation” in an attempt to impose age-related transformations on face outlines. In a more recent study, Ramanathan and Chellappa [21] use a modification of the cardioid strain transformation coupled with the use of anthropometric measurements. Anthropometric measurements derived from face images of the same individual at different ages allow the determination of optimum coefficients that can be used for tuning cardioid strain transformation coefficients for application to different individuals and different age-range



FIGURE 1: Sample images from the FG-NET Aging database.

transformations. Ramanathan and Chellappa [21] used the proposed technique for applying shape-based deformations on subjects within the age range of 0 to 18 years old.

Given a set of sparse longitudinal face images of a subject, Geng et al. [22] generate a complete aging pattern showing images of the subject in successive ages. Aging patterns are used as the basis for implementing an age estimation system. In this context, an input face image is inserted in existing aging patterns at successive positions (candidate ages) and for each case the aging pattern is coded and reconstructed. Minimization of the reconstruction error indicates that the face is inserted to the position corresponding to correct age of the face in the input image. Quantitative results show that this method outperforms other age estimation methods reported in the literature. The method reported by Geng et al. [22] can also be used as the basis for implementing age-progression algorithms.

Few researchers propose the use of support vector machines (SVMs) [5] in conjunction with age-progression algorithms. Wang and Ling [23] train SVM regressors that learn the relationship between aging and the displacement of points located on facial outlines. The resulting SVMs are used for predicting the point displacements required in order to generate age-progressed face outlines. Gandhi [7] uses an SVM regressor for estimating the age of the person in an image. In this case, the SVM regressor relates shape-normalized facial textures with age. The resulting age estimator is used for controlling the amount of wrinkles to add on shape-normalized face textures during age-progression. Scherbaum et al. [16] describe the use of SVM regressors in an attempt to define an aging axis in a parametric face space defined by a morphable 3D model. The resulting functions are used for implementing a 3D age-progression algorithm.

Geng et al. [22] and Lanitis et al. [4] run face recognition experiments using age-progressed faces in an attempt to assess the performance of age-progression algorithms. However, the results of such experiments depend heavily on the classifier used, thus this approach cannot be regarded as a generic method for assessing age-progression algorithms. Scandrett et al. [17] use the RMS difference between the shape and the normalized texture of target and age-progressed faces as a means for assessing the accuracy of age-progressed faces. However, shape and intensity-based distance metrics often provide misleading results, since fa-

cial shapes and intensities are affected by other types of facial variation other than aging. As an alternative to previous attempts for assessing the performance of age-progression algorithms, we propose the formulation and use of dedicated performance evaluation metrics that can be used as the basis for evaluating the performance of age-progression algorithms.

### 3. THE FG-NET AGING DATABASE

For the experiments described in this paper, we have used the FG-NET Aging database [6]. The FG-NET Aging database is a publicly available image database containing face images showing a number of subjects at different ages. The database has been developed in an attempt to assist researchers who investigate the effects of aging on facial appearance. The database contains 1002 images from 82 different subjects with ages ranging between newborns to 69 years old subjects. However, ages between zero to 40 years are the most populated in the database. Typical images from the database are shown in Figure 1. Data files containing the locations of 68 facial landmarks and the age of the subject in each image are available.

Images in the database were collected by scanning photographs of subjects found in personal collections. As a result, face images in the FG-NET Aging database display significant variability in resolution, quality, illumination, viewpoint, and expression. Occlusions in the form of spectacles, facial hair, and hats are also present in a number of images. Because of the low resolution of the images in the database, it is not possible to use images from this database for modeling efficiently subtle age-related skin deformations such as wrinkle related deformations.

So far, the FG-NET Aging database has been distributed to more than 200 universities and/or research centers, enabling in that way a significant number of researchers to carry out experiments in areas related to age-progression. Apart from the FG-NET Aging database, the MORPH database [24] is also publicly available. The MORPH database currently contains 1724 images from 515 individuals captured within age intervals ranging from 46 days to 29 years. However, the number of images per individual is limited to approximately three images. For our study, we have chosen



FIGURE 2: Age prototypes for the age groups 0–5, 5–10, 20–25, 30–35, 40–45, 50–55, and 60–65 years. Age prototypes are derived using images from the FG-NET Aging database.

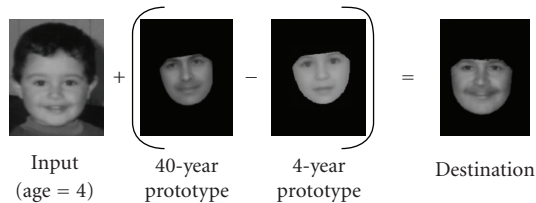


FIGURE 3: Age-progression using age prototypes. An image of a four-year old is age-progressed to 40 years by adding to the input image the difference between the age prototypes corresponding to 40-year-old subjects and four-year-old subjects.

to use the FG-NET Aging database since in the FG-NET database approximately 10 images per subject are available.

#### 4. AGE-PROGRESSION METHODOLOGIES

In this section, we describe the age-progression methodologies for which we present comparative performance evaluation results. Two of the methods under investigation were reported in the literature, whereas the remaining two are novel age-progression methods.

##### 4.1. Age-progression using age prototypes

Rowland and Perret [3] propose an age-progression method based on age prototypes. Age prototypes are generated by merging the shape and intensities of faces belonging to the same age group. The face merging process is carried out independently for shape and texture. In the case of shape, the merging operation involves the calculation of the mean shape among a group of faces belonging to the same age group. In order to get noise-free prototypes, all faces to be merged are warped to a standard shape and the mean shape-normalized texture among the constituent faces is estimated. Figure 2 shows typical age prototypes for different age groups. In age prototypes derived using a large number of samples, individual facial characteristics are suppressed in favor of typical facial characteristics of subjects belonging to the corresponding age group. Differences between age prototypes describe typical age-related deformations between different age groups.

Given a previously unseen face to be age-progressed, Rowland and Perret [3] estimate the difference between the age prototype corresponding to the current age of the subject in the given image and the age prototype at the target age. The estimated difference is added to the given face in order to obtain an estimate of the future facial appearance of the given face. During the age-progression process, warp-



FIGURE 4: Images showing synthetic faces at ages between 0–50 years produced using aging functions trained using images from the FG-NET Aging database.

ing operations are used so that operations involving texture are carried out on shape-normalized faces. Resulting textures are inverse warped to the appropriate age-progressed shape. Figure 3 illustrates the age-progression methodology proposed by Rowland and Perret [3].

##### 4.2. Age-progression using aging functions

Lanitis et al. [4] propose a model-based age-progression algorithm that uses aging functions for modeling aging variation within a training set. During the training phase, they generate a statistical appearance face model [25, 26] that describes the major sources of variability within the training set. During the process, each face shape in the training set is represented by the coordinates of 68 landmarks. All training shapes are aligned and the mean shape among the training set is established. Image warping is used for warping training faces to the mean shape so that the shape-normalized texture from each face is extracted. A statistical appearance face model is generated by applying principal component analysis on training shapes and shape-normalized face intensities. Since all training faces are warped to the same shape, information related to the absolute scale of training faces is discarded. Although scaling is an important aspect in age-progression, the use of scale information requires prior knowledge regarding the scale of faces in the training and test images—such information is usually not available in images encountered in most face image processing applications. One of the most important features of PCA-based face models is the ability to represent faces using a small number of model parameters. The coding achieved based on this methodology is reversible enabling the reconstruction of new faces once the values of model parameters are fixed. More details related to the training and use of statistical models of this type are presented elsewhere [25, 26].

Lanitis et al. [4] convert all training samples into the low-dimensional model-based representation and define a polynomial function (the so-called aging function) that relates the model-based representation of each subject to the actual age;

$$\text{age} = f(\mathbf{X}), \quad (1)$$

where  $\mathbf{X}$  is a vector of model parameters and  $f$  is the aging function. For the work described in this paper, the aging function used is a nonlinear polynomial function similar to the one used in [4]. Once an aging function is established, it can be used for estimating the age of faces in images and also for generating typical images showing a face at a desired age. Figure 4 shows synthetic faces at different ages produced by

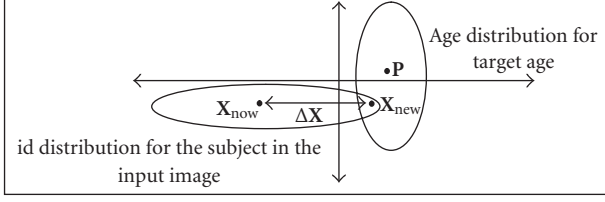


FIGURE 5: Illustration of the distance-based age-progression method.

using an aging function trained using images from the FG-NET Aging database.

Given a previously unseen face to be age-progressed, Lanitis et al. [4] code the face into model parameters and use (2) for estimating the face model parameters corresponding to the age-progressed face;

$$\mathbf{X}_{\text{new}} = \mathbf{X}_{\text{now}} + [f^{-1}(\text{age}_{\text{new}}) - f^{-1}(\text{age}_{\text{now}})], \quad (2)$$

where  $\mathbf{X}_{\text{new}}$  and  $\mathbf{X}_{\text{now}}$  are the model parameters at the target and current age, respectively,  $f^{-1}$  is the aging function solved with respect to the model parameters,  $\text{age}_{\text{new}}$  is the target age, and  $\text{age}_{\text{now}}$  is the age of the face in the input image. Since a face of a subject at a certain age may undergo other types of appearance variation (i.e., due to changes in expression, orientation, and illumination), a simulation approach was utilized for determining the solution of the function  $f^{-1}$  [4]. In this respect,  $f^{-1}$  defines the set of the most typical model parameters for each age within the age range of interest.

### 4.3. Distance-based age-progression

A successful age-progression system should be able to produce face images that display typical age-related characteristics of faces belonging to the target age group and at the same time retain the individual facial characteristics of the subject. The two requirements quoted above are graphically demonstrated in Figure 5, where the distributions of samples belonging to the same person (id distribution) and the distribution of samples belonging to the same age (age distribution) are shown in a 2-dimensional model parameter space (in reality age and id distributions are defined in a multi-dimensional space—the distributions shown in Figure 5 are just used for illustration purposes). If we wish to age progress a face image belonging to the subject whose distribution is shown in Figure 5, we need to move the current projection of the face in the model parameter space (point  $\mathbf{X}_{\text{now}}$ ) as close as possible to the center of the target age distribution (point  $\mathbf{P}$ ). However, at the same time we need to keep the projection within the vicinity of the id distribution. Based on this formulation, the optimum new position to move the face is the point that minimizes both the distance to the center of the id distribution and the distance to the center of the age distribution (point  $\mathbf{X}_{\text{new}}$  in Figure 5).

In order to implement this method, we use a statistical face model (similar to the one used in age-progression using aging functions) as the basis for representing training faces using model parameters. In our experiments, we need about

55 model parameters for representing a face; hence the age and id distributions are defined in a 55-dimensional space. Images from the training set are used for modeling the age distributions for faces belonging to the same age group, as multivariate normal distributions in the model space. For our experiments, we use the same 5-year interval age groups, as the ones used for generating age prototypes. We also estimate the typical id distribution of model parameters by considering images from a training set that show the same individual. In all cases, model parameter distributions are described by the centers of the distribution and the corresponding covariance matrices. Given a previously unseen face image to be age-progressed, we first obtain its model-based representation ( $\mathbf{X}_{\text{now}}$ ) [25]. The aim is to define the required displacement ( $\Delta\mathbf{X}$ ) of the face parameter vector so that it is possible to obtain the optimum vector of model parameters that corresponds to the age-progressed image ( $\mathbf{X}_{\text{new}}$ ),

$$\mathbf{X}_{\text{new}} = \mathbf{X}_{\text{now}} + \Delta\mathbf{X}. \quad (3)$$

We use a minimization algorithm based on a sequential quadratic programming method [27] in order to define the optimum displacement ( $\Delta\mathbf{X}$ ) required to minimize the following cost function:

$$d = d_{\text{age}}(\mathbf{X}_{\text{now}} + \Delta\mathbf{X}) + d_{\text{id}}(\mathbf{X}_{\text{now}} + \Delta\mathbf{X}), \quad (4)$$

where  $d_{\text{age}}$  is the Mahalanobis distance between a set of parameters and the age distribution at the target age and  $d_{\text{id}}$  is the Mahalanobis distance between a candidate set of parameters and the id distribution of the current face:

$$\begin{aligned} d_{\text{id}} &= (\mathbf{X}_{\text{new}} - \mathbf{X}_{\text{now}}) \mathbf{C}_{\text{id}}^{-1} (\mathbf{X}_{\text{new}} - \mathbf{X}_{\text{now}})', \\ d_{\text{age}} &= (\mathbf{X}_{\text{new}} - \mathbf{X}_t) \mathbf{C}_{\text{age}}^{-1} (\mathbf{X}_{\text{new}} - \mathbf{X}_t)', \end{aligned} \quad (5)$$

where  $\mathbf{X}_t$  are the mean parameters of the target age distribution.  $\mathbf{C}_{\text{id}}$  and  $\mathbf{C}_{\text{age}}$  are the identity and age covariance matrices derived from the training set.  $\mathbf{C}_{\text{id}}$  and  $\mathbf{C}_{\text{age}}$  describe the typical scatter of face parameters for the id and age distributions so that the calculation of the  $d_{\text{age}}$  and  $d_{\text{id}}$  distances emphasize age-related and id-related dissimilarities, respectively.

### 4.4. SVM-based age-progression

The distance-based age-progression technique presented in Section 4.3 relies on the assumption that the distributions of model parameters for different age groups and different subjects are normal distributions. As an alternative to the distance-based approach, we propose an age-progression methodology that relies on support vector machines as a means of modeling age and id distributions in the model parameter space. The rationale behind the SVM-based age-progression method is similar to the rationale of the distance-based approach; age-progression to a target age is done in such a way so that an age-progressed face belongs to the target age distribution while retaining the identity of the user.

During the training stage we train age group and id SVM classifiers using images from the training set. The kernel and kernel parameters used during the process of training SVM

classifiers are optimized by splitting the training set into two subsets so that the first half is used for training the SVMs and the second half for verification. The best parameters are the ones for which we obtain the highest correct classification rates on the verification set, when the SVMs are used in the “one-against-all” classification scheme. According to the experimental evaluation, radial basis functions (RBF) kernels with standard deviation of 30 offer the best performance.

Once an id or age SVM is trained, it is possible to define an SVM similarity measure that provides a measure of the probability that an observation belongs to a certain distribution. The SVM similarity measure between an observation and a class is defined by the function

$$\text{SVM}_{\text{sim}}(\mathbf{X}) = \mathbf{w}_i \cdot k_i(\mathbf{X}) + \mathbf{b}_i, \quad (6)$$

where  $\mathbf{X}$  is a vector containing a set of face model parameters and  $k_i(\mathbf{X})$  is a vector containing the kernel function evaluations of  $\mathbf{X}$  with respect to each of the support vectors of the  $i$ th class.  $\mathbf{w}_i$  is a vector containing the weights for each support vector of the  $i$ th class and  $\mathbf{b}_i$  contains the bias value for the  $i$ th class. The support vectors, the vector with the weights ( $\mathbf{w}$ ) and the vector with bias values ( $\mathbf{b}$ ), are defined during the process of training an SVM classifier. Values of the SVM similarity measures around the value of one indicate that an observation belongs to the corresponding class, whereas values close to minus one indicate that an observation is not likely to belong to the corresponding class. In order to improve the clarity of presentation of quantitative results, we scale SVM similarity measures to values between zero and one where a value of zero indicates high dissimilarity, whereas a value of one indicates maximum similarity between an observation and a distribution.

During the age-progression process, we use an optimization algorithm based on a sequential quadratic programming method [27] in order to define the optimum displacement ( $\Delta\mathbf{X}$ ) required to maximize a similarity measure (sim) that contains two terms: the age similarity measure ( $\text{sim}_{\text{age}}$ ) and the id similarity measure ( $\text{sim}_{\text{id}}$ ),

$$\text{sim} = \text{sim}_{\text{age}}(\mathbf{X}_{\text{now}} + \Delta\mathbf{X}) + \text{sim}_{\text{id}}(\mathbf{X}_{\text{now}} + \Delta\mathbf{X}). \quad (7)$$

The age similarity measure is defined as the SVM similarity between a coded face representation and the distribution of faces at the target age group. The calculation of the age similarity measure is based on (6).

The id similarity measure aims to preserve the SVM similarity values between a face image and all subject id distributions in the training set, before and after age-progression is applied. SVM similarity measures between a face and the distributions belonging to  $N$  different subjects from the training set are indicated in the vector  $\mathbf{s} = (s_1, s_2, \dots, s_N)$ , where each element of the  $\mathbf{s}$  vector is calculated using (6). When we age progress a face, the similarities between the resulting face and the  $N$  id distributions will be modified to  $\mathbf{s}' = (s'_1, s'_2, \dots, s'_N)$ . In order to preserve the identity of a person during the process of age-progression, we aim to minimize discrepancies

between  $\mathbf{s}$  and  $\mathbf{s}'$  by maximizing the id similarity measure shown in

$$\text{sim}_{\text{id}} = -\frac{1}{N} \sum_{i=1}^N |s_i - s'_i|. \quad (8)$$

During the age-progression process, we maximize the objective function shown in (7) so that the resulting face shows similarities with faces at the target age group (by maximizing  $\text{sim}_{\text{age}}$ ) and also retains as much as possible its personal appearance (by maximizing  $\text{sim}_{\text{id}}$ ).

#### 4.5. Age-progression using weighted objective function

Both in the case of distance-based and SVM-based age-progression, the objective functions contain an age-related and an id-related term. Ideally, the importance of each term should be adjusted according to the age-progression range so that the smaller the range, the higher the significance of the id-related term. On the other hand, when we deal with long-range age-progression the age-related term should dominate the objective function. In order to deal with this issue, a weighted objective function is formulated as follows:

$$d = w * d_{\text{age}} + (1 - w) * d_{\text{id}}, \quad (9)$$

$$w = \frac{|\text{age}_{\text{new}} - \text{age}_{\text{now}}|}{\text{age}_{\text{max}} - \text{age}_{\text{min}}},$$

where  $\text{age}_{\text{min}}$  and  $\text{age}_{\text{max}}$  define the maximum age-progression range considered in our application. Usually,  $\text{age}_{\text{min}}$  and  $\text{age}_{\text{max}}$  are the minimum and maximum ages in the training set.

The weighted objective function can be applied both to distance-based and SVM-based age-progressions. However, unlike the nonweighted distance-based and SVM-based algorithms, the use of a weighted age-progression method requires an estimate of the current age of the subject to be age-progressed.

## 5. EXPERIMENTAL EVALUATION

In order to evaluate the performance of the age-progression methods described above, we performed an experimental evaluation using images from the FG-NET Aging database. The experimental setup and the results obtained are described hereunder.

### 5.1. Experimental setup

All images from the FG-NET Aging database are divided into two groups: Group A contains all images of subjects with ids 001-040 (498 images); Group B contains 504 images of the subjects with ids 041-082. For our experiments, we have used Group A for training and performing initial investigations and Group B for testing. We also report results where the training was performed using images from Group B and testing on images from Group A.

TABLE 1: Results of the preliminary investigation. Figures show the normalized mean similarity between a sample and the correct\incorrect distributions. Figures in parenthesis show the standard deviation for each case.

Train set	Test set	Age similarity ( $age_s$ )		id similarity ( $id_s$ )	
		Same age	Different age	Same id	Different id
Group A1	Group A1	0.82 (0.17)	0.18 (0.21)	0.93 (0.11)	0.09 (0.14)
Group A2	Group A2	0.90 (0.16)	0.17 (0.17)	0.94 (0.10)	0.09 (0.14)
Group A1	Group A2	0.63 (0.21)	0.21 (0.21)	0.62 (0.17)	0.13 (0.16)
Group A2	Group A1	0.64 (0.22)	0.18 (0.21)	0.60 (0.17)	0.12 (0.14)

TABLE 2: Quantitative evaluation results. The figures quoted are the mean-normalized id and age similarity values obtained among the test set. The figures in parenthesis show the standard deviation of the similarity measures .

Method	Train set: Group A, Test set: Group B (5825 test cases)		Train set: Group B, Test set: Group A (5793 test cases)	
	$id_s$	$age_s$	$id_s$	$age_s$
	No age-progression	0.98 (0.10)	0.05 (0.55)	1.00 (0.05)
Prototype	0.75 (0.10)	0.30 (0.45)	0.64 (0.25)	0.27 (0.35)
Aging functions	0.92 (0.10)	0.16 (0.50)	0.86 (0.08)	0.22 (0.39)
Distance (Dis)-based	0.41 (0.20)	0.44 (0.35)	0.40 (0.20)	0.45 (0.28)
SVM-based	0.38 (0.29)	0.68 (0.50)	0.43 (0.14)	0.59 (0.30)
Weighted distance (WDis)-based	0.82 (0.18)	0.36 (0.47)	0.84 (0.20)	0.35 (0.38)
Weighted SVM (WSVM)-based	0.76 (0.35)	0.44 (0.45)	0.65 (0.27)	0.41 (0.35)

For each subject in the test set, we use one of his/her images as the reference image and we attempt to predict the appearance of the subject face at the ages indicated by the rest of the samples belonging to that subject. For example, if a subject has five images in the test set at ages 3, 10, 20, 30, and 40 years, we first use the face at 3 years old as the reference for predicting the appearance of the subject at the age of 10, 20, 30, and 40 years. We then use as a reference the face at 10 years for predicting the subject’s appearance at the age of three, 20, 30, and 40 years. The procedure is repeated for all images for all subjects in the test set. It is worth mentioning that these tests involve both predictions of the future and previous facial appearance of a face.

## 5.2. Performance evaluation metrics

Age-progression accuracy is assessed based on two dedicated performance evaluation measures: an age similarity measure and an individual appearance similarity measure. The age similarity measure ( $age_s$ ) assesses the ability of an algorithm to produce age-progressed images that display the characteristics of faces belonging to the target age group. The individual appearance similarity measure ( $id_s$ ) assesses whether age-progressed images display the individual characteristics of subjects in the corresponding source images.

Prior to the performance evaluation stage, we use all images from the test set for training separate SVMs for samples belonging to each age group and each subject in the test set. The age similarity measure is defined as the SVM similarity between an age-progressed image and the distribution corre-

sponding to the target age group. Similarly, the id similarity measure is defined as the SVM similarity between an age-progressed image and the distribution corresponding to the subject in the source image. Both the age and id similarity measures are calculated using (6).

Before we use  $age_s$  and  $id_s$  in our performance evaluation process, we performed a preliminary investigation that aims to assess the suitability of the measures for the proposed application. In our preliminary investigation, we use images from Group A of the FG-NET Aging database (see Section 4.1 for more details). All images from Group A of the FG-NET Aging database were divided into two groups (Group A1 and Group A2) so that each group contains half of the images of each of the 40 subjects in Group A. The separation of images was done in such way so that each part contains half of the images of each subject and also both groups contain similar distribution of ages for each subject. In our preliminary investigation, we use Group A1 for training and Group A2 for testing and vice versa. We use all images from the training set for training id and age SVMs. We then calculate the similarity measures between each image and all age group and id distributions. Since the correct id and age group of each face is known, it is possible to collect statistics related to the values of the age and id similarities measures between the test face and the correct id and age distributions. Similarly, we collect statistics of the similarity measures between each face and the noncorrect age and id distributions. The results of the preliminary investigation are shown in Table 1. According to the results, the similarity measures have higher



values when we deal with the correct age group or correct id, when compared to the case that we deal with incorrect distributions. Therefore, the two measures can be used for assessing the accuracy of age-progressed faces from a test set.

The formulation of the age similarity measure used during age-progression and the formulation of the age similarity measure used during performance evaluation are similar. However, in the case of the measure used during age-progression, the SVMs involved are trained using images from the training set, and in the case of performance, evaluation SVMs are trained using images from the test set.

### 5.3. Experimental results

The results of our experiments are shown in Table 2. Prior to the application of an age-progression algorithm, the mean id similarity measure among images in the test set is close to the value of one, since an unprocessed face image is an undisputed member of the corresponding id distribution. On average, the age similarity measures have values around 0.1 indicating that in most cases the raw faces do not display age similarities with faces in the target age group. This is expected since in the majority of cases the age of faces in test images and the target age belong to different age groups.

In all cases, the application of age-progression algorithms results in decreasing the id similarity between an age-progressed face and the corresponding id distribution. This is expected since the application of an age-progression algorithm will modify the appearance of a subject causing displacements from the corresponding id distribution. Ideally, the decrement in id similarity measure should be minimized. According to the results, the algorithms based on prototypes, aging functions, and the weighted methods are doing better in preserving the individual appearance in age-progressed faces.

For all algorithms considered, the resulting images display improved age similarity with images in the target age group, indicating that all algorithms tested are capable of inflicting age-related deformations on face images. The SVM-based and distance-based algorithms display better performance in generating faces closer to the target age distribution.

The distance-based and SVM-based algorithms seem to provide a balanced performance because they achieve more uniform improvement of both similarity measures. The introduction of the weighted objective function improves significantly the id similarity measure, but has a negative impact on the age similarity measure. When we take into account the overall performance of each algorithm, the weighted distance-based and weighted SVM-based algorithms seem to achieve the best compounded effect in producing higher values for both the age and id similarity measures. Visual results of age-progression using all methods considered in the experimental evaluation methods are shown in Figure 6.

## 6. CONCLUSIONS

We presented a framework for evaluating the performance of automatic age-progression methodologies using a pub-

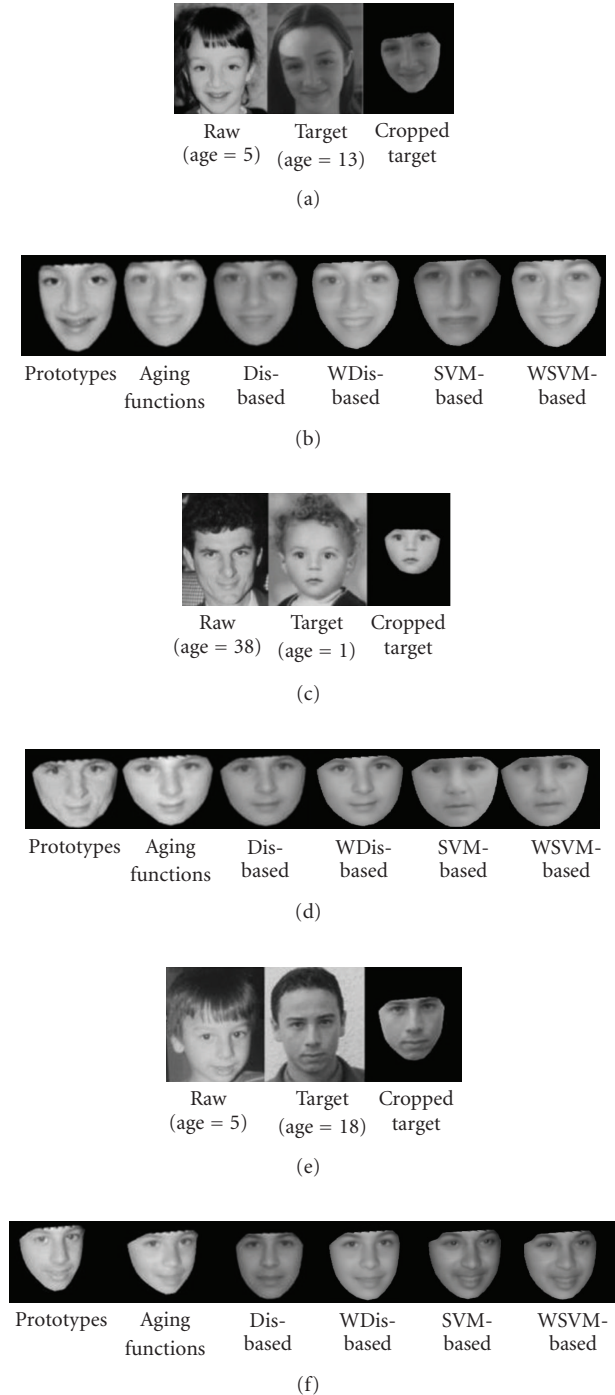


FIGURE 6: Examples of age-progressed images generated using different algorithms.

licly available database. For this purpose, we have developed dedicated performance evaluation metrics that can be used for assessing the accuracy of age-progressed images. As part of our experiments, we compared the performance of two age-progression methods reported previously in the literature and two novel age-progression methods.

### 6.1. Comparing age-progression methodologies

According to the experimental results, the weighted distance-based and the weighted SVM-based methods produced the best overall results. However, the use of methods based on a weighted objective function implies that the age of the face in an input image is known. In the cases that such information is not available, the age of the face in the input image needs to be estimated automatically. Bearing in mind that automatic age-estimation methods [22, 28] yield error rates of about five years, it is expected that in real applications the performance of the weighted methods will deteriorate. Similarly, the performance of aging based on prototypes and aging functions will also deteriorate as in both cases an estimate of the current age of the subject is required. In contrast, the distance-based and SVM-based methods do not require such information; hence in totally automatic mode of operation they could outperform other methods.

The method based on age prototypes operates directly on images hence it is possible to produce more realistically looking age-progressed images. However, the method based on prototypes is affected considerably by occlusions and the quality of the face image to be age-progressed. Images from the FG-NET Aging database are not ideal for use in conjunction with age-progression algorithms based on prototypes.

Model-based approaches discard subtle image artifacts and noise in favor to systematic sources of variability encountered in the training images. Since during the model building process all images are warped to a standard shape, the effects of limited variation in 3D face orientation do not cause severe distortions in the texture of faces images considered. Variation related with 3D orientation is usually explained by few shape-related modes of variation in the training set, which are isolated implicitly during the training process. Hence, model-based methods are more applicable when dealing with noisy and unconstrained images. However, when dealing with an application where we need to generate aesthetically pleasing results, model-based approaches need to be augmented with methods that add high-frequency artifacts on age-progressed images [7, 8] in order to improve the realism of age-progressed faces.

The use of the FG-NET Aging database provides a demanding test for age-progression methodologies. The presence of other types of facial appearance variation other than variation due to aging in images from the FG-NET Aging database makes the database a useful tool for testing the robustness of age-progression systems. Algorithms capable of performing well when tested on images from the FG-NET Aging database should also achieve acceptable performance when tested using images captured under noncontrolled conditions.

### 6.2. Future work

In the experiments described in this paper, we did not take into account the diversity of aging variation in different groups of subjects. It has been demonstrated that males and females and in general different subjects may adopt different aging patterns [13, 15]. In order to take into account the

diversity in aging patterns, we need to generate age prototypes/age distributions for different clusters of subject groups from the training set. In the future, we plan to produce evaluation results in the cases that techniques for dealing with diverse aging effects are used.

The work described in this paper presents our preliminary work towards the establishment of a concrete test bench for testing the performance of age-progression systems. In the future, we plan to include in our experimental evaluation other automatic age-progression techniques reported in the literature.

Although the FG-NET Aging database is a useful tool for testing age-progression algorithms, it is not ideal for training such systems. Especially for ages higher than 40 years, the number of images available is limited and as a result the process of training age-progression systems that deal with that range of ages is inhibited. In the future, we plan to add and make available more images in order to improve the FG-NET Aging database.

### 6.3. Conclusion

We have presented a comparative experimental evaluation of age-progression algorithms. Our ultimate aim is to provide a standardized methodology for evaluating the performance of age-progression systems, in order to support and facilitate efforts of researchers who deal with this face image processing problem. The formulation of dedicated performance evaluation metrics and standardized performance evaluation protocols is of utmost importance for the future development of improved age-progression algorithms. We believe that the work reported in this paper will provoke interest for this particular topic so that researchers working in the area of age-progression and in general in the area of face image processing will benefit.

### ACKNOWLEDGMENTS

The research work presented was supported by the Research Promotion Foundation (Cyprus) (Project EPYAN/0205/08). The generation of the FG-NET Aging database was supported by the European Union project "Face and Gesture Recognition Network (FG-NET) [29] (Project IST-2000-26434)." The authors are grateful to the anonymous reviewers of this paper for their suggestions and comments.

### REFERENCES

- [1] "Age Progression," November 2007, <http://www.forensicartist.com/agepro.html>.
- [2] "Age Progression and Image Modification," November 2007, <http://www.forartist.com/forensic/modification/modification-npg.htm>.
- [3] D. A. Rowland and D. I. Perrett, "Manipulating facial appearance through shape and color," *IEEE Computer Graphics and Applications*, vol. 15, no. 5, pp. 70–76, 1995.
- [4] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 442–455, 2002.

- [5] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [6] The FG-NET Aging Database, November 2007, <http://webmail.cycollege.ac.cy/~alanitis/fgnetaging/>.
- [7] M. Gandhi, "A method for automatic synthesis of aged human facial images," M.S. thesis, Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, 2004.
- [8] B. Tiddeman, M. Burt, and D. Perrett, "Prototyping and transforming facial textures for perception research," *IEEE Computer Graphics and Applications*, vol. 21, no. 5, pp. 42–50, 2001.
- [9] W.-S. Lee, Y. Wu, and N. Magnenat-Thalmann, "Cloning and aging in a VR family," in *Proceedings of IEEE Virtual Reality (VR '99)*, pp. 61–68, Houston, Tex, USA, March 1999.
- [10] L. Boissieux, G. Kiss, N. Magnenat-Thalmann, and P. Kalra, "Simulation of skin aging and wrinkles with cosmetics insight," in *Proceedings of Eurographics Workshop on Animation and Simulation (EGCAS '00)*, pp. 15–27, Interlaken, Switzerland, August 2000.
- [11] F. R. Leta, D. Pamplona, H. I. Weber, A. Conci, and I. Pitanguy, "A study of the facial aging—a multidisciplinary approach," *Journal of the Brazilian Society of Mechanical Sciences*, vol. 22, no. 3, pp. 489–501, 2000.
- [12] I. Pitanguy, F. Leta, D. Pamplona, and H. I. Weber, "Defining and measuring ageing parameters," *Applied Maths and Computations*, vol. 78, no. 2-3, pp. 217–227, 1996.
- [13] C. M. Hill, C. J. Solomon, and S. J. Gibson, "Aging the human face—a statistically rigorous approach," in *Proceedings of the IEE International Symposium on Imaging for Crime Detection and Prevention (ICDP '05)*, pp. 89–94, London, UK, June 2005.
- [14] E. Patterson, K. Ricanek, M. Albert, and E. Boone, "Automatic representation of adult aging in facial images," in *Proceedings of the 6th IASTED International Conference on Visualization, Imaging, and Image Processing (VIIP '06)*, p. 612, Palma De Mallorca, Spain, August 2006.
- [15] T. J. Hutton, B. F. Buxton, P. Hammond, and H. W. W. Potts, "Estimating average growth trajectories in shape-space using kernel smoothing," *IEEE Transactions on Medical Imaging*, vol. 22, no. 6, pp. 747–753, 2003.
- [16] K. Scherbaum, M. Sunkel, H.-P. Seidel, and V. Blanz, "Prediction of individual non-linear aging trajectories of faces," *Computer Graphics Forum*, vol. 26, no. 3, pp. 285–294, 2007.
- [17] C. M. Scandrett (née Hill), C. J. Solomon, and S. J. Gibson, "A person-specific, rigorous aging model of the human face," *Pattern Recognition Letters*, vol. 27, no. 15, pp. 1776–1787, 2006.
- [18] N. Ramanathan and R. Chellappa, "Face verification across age progression," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3349–3361, 2006.
- [19] J. B. Pittenger and R. E. Shaw, "Aging faces as viscal-elastic events: implications for a theory of nonrigid shape perception," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 1, no. 4, pp. 374–382, 1975.
- [20] J. B. Pittenger, R. E. Shaw, and L. S. Mark, "Perceptual information for the age level of faces as a higher order invariant of growth," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 5, no. 3, pp. 478–493, 1979.
- [21] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, vol. 1, pp. 387–394, New York, NY, USA, June 2006.
- [22] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2234–2240, 2007.
- [23] J. Wang and C. X. Ling, "Artificial aging of faces by support vector machines," in *Advances in Artificial Intelligence*, vol. 3060 of *Lecture Notes in Computer Science*, pp. 499–503, Springer, Berlin, Germany, 2004.
- [24] K. Ricanek Jr. and T. Tesafaye, "MORPH: a longitudinal image database of normal adult age-progression," in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR '06)*, pp. 341–345, Southampton, UK, April 2006.
- [25] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
- [26] G. J. Edwards, A. Lanitis, C. J. Taylor, and T. F. Cootes, "Statistical models of face images: improving specificity," *Image and Vision Computing*, vol. 16, no. 3, pp. 203–211, 1998.
- [27] R. K. Brayton, S. W. Director, G. D. Hachtel, and L. Vidigal, "A new algorithm for statistical circuit design based on quasi-Newton methods and function splitting," *IEEE Transactions on Circuits and Systems*, vol. 26, no. 9, pp. 784–794, 1979.
- [28] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 34, no. 1, pp. 621–628, 2004.
- [29] FG-NET Web Server, November 2007, <http://www-prima.imag.fr/FGnet/>.