Back to the future: A fully automatic method for robust age progression

Christos Sagonas*⋆, Yannis Panagakis*, Saritha Arunkumar†, Nalini Ratha‡, and Stefanos Zafeiriou*
*Imperial College London, UK, †IBM Hursley Labs, UK, ‡IBM Research, USA

Abstract—It has been shown that significant age difference between a probe and gallery face image can decrease the matching accuracy. If the face images can be normalized in age, there can be a huge impact on the face verification accuracy and thus many novel applications such as matching driver's license, passport and visa images with the real person's images can be effectively implemented. Face progression can address this issue by generating a face image for a specific age. Many researchers have attempted to address this problem focusing on predicting older faces from a younger face. In this paper, we propose a novel method for robust and automatic face progression in totally unconstrained conditions. Our method takes into account that faces belonging to the same age-groups share age patterns such as wrinkles while faces across different age-groups share some common patterns such as expressions and skin colors. Given training images of K different age-groups the proposed method learns to recover K low-rank *age* and one low-rank *common* components. These extracted components from the learning phase are used to progress an input face to younger as well as older ages in bidirectional fashion. Using standard datasets, we demonstrate that the proposed progression method outperforms state-of-the-art age progression methods and also improves matching accuracy in a face verification protocol that includes age progression.
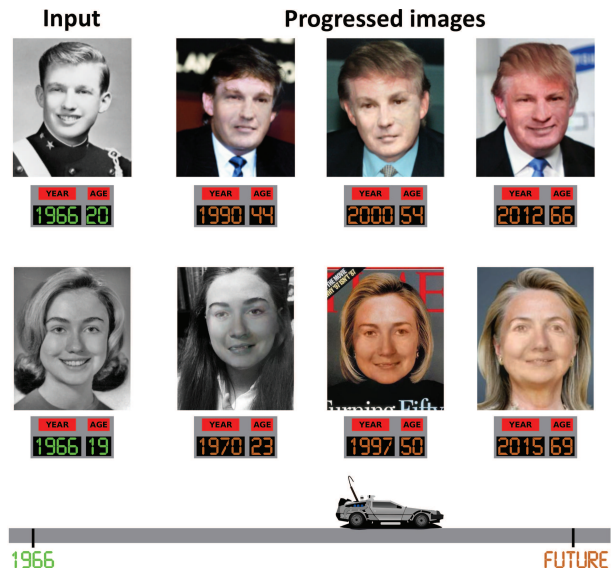
Fig. 1: Given an input image the proposed progression method is able to reconstruct the clean neutral frontal face of the input subject in different ages from babies to seniors.

I. INTRODUCTION

Recognizing and verifying the identity of a human is an essential problem to be solved in several applications. Recently, a huge effort has been made on designing methods for robust face recognition and verification. This led to the proposal of very powerful methods with excellent performance [1], [2]. However, the performance of the existing methods drops dramatically when they process images with subjects in different ages than those used for training (gallery images). This drop in performance is attributed to the fact that the age progression affects both the shape and appearance of a face [3]. The simplest way to overcome this is to include training images from different age-groups. Clearly, this is not always possible in practice. Consequently, the automatic age progression of the gallery images is of paramount importance.

Several methods for the problem of age progression have been proposed [4]–[13]. These methods are classified into *prototype-* and *model*-based, depending on whether they employ differences between prototypes (average faces) or learned age transformations. The *prototype-based* methods [4]–[6], progress an image into a new age by computing the facial differences (in shape and appearance) between the average

faces corresponding to the input and the target ages. The progressed image is then created by transferring the computed differences into the initial image. In [4] average faces from aligned images of different ages are created and then used for age progression. The major drawback of this method is that a large proportion of the age information is lost due to the average operator. In order to overcome the loss of facial features a wavelet-based extension of the previous method is proposed in [6]. In [5], a fully automatic prototype-based system for age progression is introduced. Firstly, images of different ages are automatically aligned and used to create average faces. Then, given a test image its illumination is matched with the illuminations of the created average faces and the progression is performed. The proposed system is able to produce progressed faces for different ages from young to old. However, a part of identity information is lost due to the usage of the average operator and in addition the proposed method is not robust to miss-alignments and different types of occlusions and illuminations.

In contrast, the *model-based* methods [7]–[10], [12], [13] progress an image by using learnt shape and appearance models. In [7], statistical models of the shape and appearance of different faces and ages are built based on the principal

component analysis (PCA). Then, an age function that defines the relationship between the parametric description of different images and their age is learnt and subsequently is employed to progress both the shape and texture of an input face. The extension of previous methods in the 3D case is proposed in [10]. This method cannot be widely applied since it requires the collection of 3D scans which is a laborious and expensive process, and the fitting of 3D model of faces, which is expensive to build. In [8], [9], a craniofacial-based growth model which characterizes growth-related shape variations is used for age progression in young people and adults. These methods are computationally complex as well as the resulting faces do not seem realistic. In [13], a hierarchical graph is used to describe the decomposition of an image into different parts. Then, the parts in different ages are used to learn the aging procedure which are subsequently employed to progress a new face. The proposed method cannot guarantee that the identity information is preserved after progression as well as it requires the creation of masks for all training images, which is a very time consuming process. Another limitation is that the test image should be expressionless in frontal position. In [12], a learning-based method for age progression is proposed. Using images of different ages, age-specific dictionaries are learnt and subsequently used for the progression step. The produced results could be drastically affected from miss-alignments and occlusions. In addition, the input image should be in near-frontal position and can only be progressed in older ages.

In this paper, the *Robust Face Age Progression* (RAP) is proposed for fully automatic and robust face age progression. The key motivational observation is that the face of a subject in a specific age can be expressed as a superposition of two components: the *age component* accounting for age-related information and the *common component* that represents all the other facial variations such as expression, shape, and identity. This observation applies also when multiple images and age-groups of the same gender are available. Thus, each image in a specific age-group can be expressed as the sum of the *age* component which models the age information inside the age-group and the *common* component capturing variations across all the images and age-groups. In this paper, we show how to recover the *age* and *common* components for a cohort of images of different subjects and ages. Furthermore, we show how the recovered components can be used in order to progress a new image into different age-groups. Compared to the most closely related methods [5] and [12] the RAP is completely different since (i) it models both the variations inside each age-group and across all the age-groups, (ii) it doesn't require training set with thousands of images or images of the same subjects, (iii) is robust to different kinds of variations such as pose, occlusions, illumination conditions, and expressions, and (iv) can progress an image in younger as well as older ages.

Summarizing, the contribution of the paper are as follows: **(i)** A novel model i.e., the RAP, is proposed for recovering the *age* and *common* components from a cohort of images of different subjects and ages. **(ii)** An efficient Alternating-Directions Method of Multipliers (ADMM)-based algorithm that can solve a suitable optimization problem for the RAP is presented. **(iii)** A novel method that progresses a new image into different age-groups using the extracting components is proposed. **(iv)** Finally, a system for fully automatic age-invariant face verification is presented[1].

II. PROPOSED METHOD

A. Age and Common components extraction

Age progression is a very slow and complex process which affects both the appearance and shape of the face. The shape-related age progression's effects are mainly observed during childhood as it is the period where the shape of the skull is changed. On the other hand, the formation of the wrinkles which is due to the reduction of the muscles strength, results in the drastic change of the appearance during the adult ages. In addition, the environmental factors as well as the subject's daily life can affect the procedure. Thus, we can say that the aging procedure is different for each individual and has different effects. However, some general patterns related to the shape and appearance can be found and described. Given a cohort of images of different age-groups we can observe that the faces belong to a specific age-group share age patterns such as wrinkles and the shape of the face. In addition, the faces across different age-groups share common patterns like expressions and skin colors. Based on the above observation, an image $\mathbf{t} \in \mathbb{R}^{m \times n}$ in a specific gender and age n can be expressed as a combination of the age and common patterns i.e., $\mathbf{t}^{(n)} = \mathbf{Age}^{(n)} + \mathbf{Common}$.

Let $\mathcal{D}$ be a database containing faces of different subjects and ages from one month to 70 years old. Each face is described by an age label and a shape instance consisting of the coordinates of L landmarks points i.e., $\mathbf{s} = [x_1, y_1, ..., x_L, y_L] \in \mathbb{R}^{2L \times 1}$. By using the label information we assign all the faces into K different age-groups and subsequently in each age-group we keep N images of different subjects. Then, in each age-group a mean shape is computed and subsequently all the faces are warped into that. Without loss of generality these groups are represented by matrices $\{\mathbf{X}^{(n)} \in \mathbb{R}^{d_n \times N}\}_{n=1}^{K}$ of different dimensions i.e., $d_1 \neq d_2 \neq \cdots \neq d_K$ where the column i-th of the n-th matrix corresponds to the warped image $\mathbf{t}_i^{(n)}$ in lexicographical order. Given the concatenation of K age-group matrices $\mathbf{X} = [\mathbf{X}^{(1)^T} | \cdots | \mathbf{X}^{(K)^T}]^T \in \mathbb{R}^{(d_1 + \cdots + d_K) \times N}$, the goal of this paper is to recover the low-rank *age* components $\{\mathbf{A}^{(n)} \in \mathbb{R}^{d_n \times N}\}_{n=1}^{K}$, the low-rank *common* component $i.e., \mathbf{C} \in \mathbb{R}^{(d_1 + \cdots + d_K) \times N}$ such that: $\mathbf{X} = \mathbf{C} + [\mathbf{A}^{(1)^T} | \cdots | \mathbf{A}^{(K)^T}]^T + \mathbf{E}$, where $\mathbf{E} \in \mathbb{R}^{(d_1 + \cdots + d_K) \times N}$ is an error matrix. Given that the images are produced as a direct sum of the two complementary

---

[1]*Notations:* Throughout the paper, scalars are denoted by lower-case letters, vectors (matrices) are denoted by lowercase (upper-case) boldface letters i.e., $\mathbf{x}$ ($\mathbf{X}$). $\mathbf{I}$ denotes the identity matrix of compatible dimensions. The $i - th$ column of $\mathbf{X}$ is denoted by $\mathbf{x}_i$. The ℓ_1 and the ℓ_2 norms of a vector $\mathbf{x}$ are defined as $\|\mathbf{x}\|_1 = \sum_i |x_i|$ and $\|\mathbf{x}\|_2 = \sqrt{\sum_i x_i^2}$. The Frobenius norms is defined as $\|\mathbf{X}\|_F = \sqrt{\sum_i \sum_j x_{ij}^2} = \sqrt{\text{tr}(\mathbf{X}^T \mathbf{X})}$.

spaces i.e, the *age* and *common* we have to ensure that this decomposition is unique. For this reason, an orthogonality constraint between the row spaces of $\mathbf{C}$ and $\{\mathbf{A}^{(n)}\}_{n=1}^{K}$ is introduced i.e., $\mathbf{C}\mathbf{A}^{(n)^T} = \mathbf{0}$. Based on the constraints stated above, the desired components $\mathbf{C}, \{\mathbf{A}^{(n)}\}_{n=1}^{K}$ can be obtained via solving the optimization problem:

$$\underset{\{\mathbf{A}^{(n)}\}_{n=1}^{K},\mathbf{C},\mathbf{E}}{\operatorname{argmin}} \quad \left\|\mathbf{E}\right\|_F^2$$

$$\text{s.t.} \quad \mathbf{X} = \left[\mathbf{A}^{(1)^T}\middle|\cdots\middle|\mathbf{A}^{(K)^T}\right]^T + \mathbf{C} + \mathbf{E}$$

$$\operatorname{rank}(\mathbf{A}^{(n)}) = M_n, \quad \operatorname{rank}(\mathbf{C}) = M$$

$$\mathbf{C} \cdot \mathbf{A}^{(n)^T} = \mathbf{0}, \quad n = 1,\ldots,K \quad (1)$$

Problem (1) is solved by employing the ADMM [14], which has been extensively used in Computer Vision e.g., [2], [15]. Thus, the augmented Lagrangian of (1) is introduced:

$$\mathcal{L}(\mathbf{C}, \{\mathbf{A}^{(n)}\}_{n=1}^{K}, \mathbf{E}) = \left\|\mathbf{E}\right\|_F^2 - \frac{1}{2\mu}\left\|\mathbf{Y}\right\|_F^2 +$$

$$\frac{\mu}{2}\left\|\mathbf{X} - \mathbf{C} - \left[\mathbf{A}^{(1)^T}\middle|\cdots\middle|\mathbf{A}^{(K)^T}\right]^T - \mathbf{E} + \frac{\mathbf{Y}}{\mu}\right\|_F^2, \quad (2)$$

where $\mathbf{Y} \in \mathbb{R}^{d_1+\cdots d_K \times N}$ is the Lagrangian multiplier related to the linear constraint in (1) and $\mu > 0$ is a monotonically increasing penalty parameter. (1) is solved by minimizing (2) with respect to each variable in an alternating fashion. The ADMM-based algorithm for solving (1) is summarized in Alg. 1[2]. We note that the τ-SVD operator for a matrix $\mathbf{X}$ with singular value decomposition (SVD) $\mathbf{X} = \mathbf{W}\mathbf{\Sigma}\mathbf{Z}^T$ is defined as $\mathcal{D}_\tau(\mathbf{X}) = \mathbf{W}(:,1{:}\tau)\mathbf{\Sigma}(1{:}\tau,1{:}\tau)\mathbf{Z}(:,1{:}\tau)^T$. The dominant cost of each iteration of Alg. 1 is that of the SVD algorithm involved in the computation of the τ-SVD operator in the update of $\mathbf{C}$ and $\{\mathbf{A}^{(n)}\}_{n=1}^{K}$.

B. Age progression of unseen subjects

In this section we show how the learnt *age* and *common* components $\mathbf{C}, \{\mathbf{A}^{(n)}\}_{n=1}^{K}$ can be used to progress a face into the K age-groups. Firstly, we compute the orthonormal bases $\mathbf{U} \in \mathbb{R}^{(d_1+\cdots+d_K)\times V}$, $\{\mathbf{H}^{(n)} \in \mathbb{R}^{d_n \times V_n}\}_{n=1}^{K}$ by applying SVD into $\mathbf{C}$ and $\{\mathbf{A}^{(n)}\}_{n=1}^{K}$, respectively. Then, given a test image $\mathbf{y}$, L landmarks points are detected and subsequently used to warp the face into the mean shape of each age-group resulting in $\{\mathbf{y}_w^{(n)}\}_{n=1}^{K}$. Thus, by using the extracted bases the progressed version of the warped image in the n-th age-group is expressed as $\mathbf{y}_w^{(n)} = \mathbf{U}^r\mathbf{c}_1^{(n)} + \mathbf{H}^{(n)}\mathbf{c}_2^{(n)}$. Based on the fact that the progressed image is a real one, it should not contain negative values. Thus, the progression procedure of a new image in the n-th age-group is obtained by solving the following coding problem:

$$\underset{\mathbf{c}_1^{(n)},\mathbf{c}_2^{(n)},\mathbf{z}^{(n)}\succeq\mathbf{0}}{\operatorname{argmin}} \quad \lambda_c\left(\left\|\mathbf{c}_1^{(n)}\right\|_2 + \left\|\mathbf{c}_2^{(n)}\right\|_2^2\right) + \lambda_e\left\|\mathbf{e}^{(n)}\right\|_1$$

$$\text{s.t.} \quad \mathbf{y}_w^{(n)} = \mathbf{U}^r\mathbf{c}_1^{(n)} + \mathbf{H}^{(n)}\mathbf{c}_2^{(n)} + \mathbf{e}^{(n)},$$

$$\mathbf{z}^{(n)} = \mathbf{U}^r\mathbf{c}_1^{(n)} + \mathbf{H}^{(n)}\mathbf{c}_2^{(n)}, \quad (3)$$

[2]Given that $\mathbf{X} = [\mathbf{X}^{(1)^T}|\cdots|\mathbf{X}^{(K)^T}]^T$, $\mathbf{X}^r$ selects the rows of $\mathbf{X}$ which corresponds to a specific $\mathbf{X}^{(n)}$ i.e., $\mathbf{X}^r = \mathbf{X}^{(r)}, r = 1,\ldots,K$.

Algorithm 1: Solving (1) by the ADMM method

Data: Input data $\mathbf{X}$, rank of *age* and *common* components M, M_n, $n = 1,\ldots,K$.

Result: The *common* $\mathbf{C}$ and *age* $\{\mathbf{A}^{(n)}\}_{n=1}^{K}$ components.

while *not converged* **do**

$\quad \mathbf{W}\mathbf{\Sigma}\mathbf{Z}^T = \operatorname{svd}\left(\mathbf{X} - \left[\mathbf{A}_{t+1}^{(1)^T}\middle|\cdots\middle|\mathbf{A}_{t+1}^{(K)^T}\right]^T - \mathbf{E}_t + \frac{\mathbf{Y}_t}{\mu_t}\right)$;

$\quad \mathbf{C}_{t+1} = \mathbf{W}(:,1{:}M)\mathbf{\Sigma}(1{:}M,1{:}M)\mathbf{Z}(:,1{:}M)^T$;

$\quad$**for** $n = 1 : K$ **do**

$\quad\quad \mathbf{A}_{t+1}^{(n)} = \mathcal{D}_{M_n}\left((\mathbf{X}^r - \mathbf{C}_{t+1}^r - \mathbf{E}_t^r + \frac{\mathbf{Y}_t^r}{\mu_t})(\mathbf{I} - \mathbf{Z}\mathbf{Z}^T)\right)$;

$\quad$**end**

$\quad \mathbf{E}_{t+1} =$
$\quad \mu_t/(2+\mu_t)\mathbf{I}\left(\mathbf{X} - \mathbf{C}_{t+1} - \left[\mathbf{A}_{t+1}^{(1)^T}\middle|\cdots\middle|\mathbf{A}_{t+1}^{(K)^T}\right]^T + \frac{\mathbf{Y}_t}{\mu_t}\right)$;

$\quad \mathbf{Y}_{t+1} =$
$\quad \mathbf{Y}_t + \mu_t \cdot \left(\mathbf{X} - \mathbf{C}_{t+1} - \left[\mathbf{A}_{t+1}^{(1)^T}\middle|\cdots\middle|\mathbf{A}_{t+1}^{(K)^T}\right]^T - \mathbf{E}_{t+1}\right)$;

$\quad \mu_{[t+1]} \leftarrow \min(\rho \cdot \mu_{[t]}, 10^{10})$;

$\quad$Check convergence condition:

$\quad \|\mathbf{X} - \mathbf{C}_{t+1} - \left[\mathbf{A}_{t+1}^{(1)^T}\middle|\cdots\middle|\mathbf{A}_{t+1}^{(K)^T}\right]^T - \mathbf{E}_{t+1}\|_F^2/\|\mathbf{X}\|_F^2 \leq \epsilon_1$;

$\quad t \leftarrow t + 1$;

end

---

where $\mathbf{e}^{(n)}$ is a sparse vector accounting for the non-Gaussian sparse errors and λ_c, λ_e are weighting parameters. The ADDM-based solver of (3) is outlined in Alg. 2. We note that the element-wise shrinkage operator is defined as $\mathcal{S}_\gamma(x) = sgn(x)\max(|x| - \gamma, 0)$ [16].

Algorithm 2: Solving (3) by the ADMM method

Data: Test image $\mathbf{y}_w^{(n)}$, *age* subspaces $\mathbf{H}^{(n)}$, *common* subspace $\mathbf{U}^r$, and the parameters λ_c, λ_e.

Result: The progreesed into the n-th age-group image $\mathbf{z}^{(n)}$ and the error image $\mathbf{e}^{(n)}$.

while *not converged* **do**

$\quad \mathbf{c}_{1,t}^{(n)} = \mu_t/(2\lambda_c + 2\mu_t)\mathbf{I}\mathbf{U}^T\big(\mathbf{y}_w^{(n)} - 2\mathbf{H}^{(n)}\mathbf{c}_{2,t}^{(n)} - \mathbf{e}_t^{(n)} + \mathbf{z}^{(n)} + (\mathbf{l}_{1,t} + \mathbf{l}_{2,t})/\mu_t\big)$;

$\quad \mathbf{c}_{2,t}^{(n)} = \mu_t/(2\lambda_c + 2\mu_t)\mathbf{I}\mathbf{H}^{(n)^T}\big(\mathbf{y}_w^{(n)} - 2\mathbf{U}\mathbf{c}_{1,t+1}^{(n)} - \mathbf{e}_t^{(n)} + \mathbf{z}^{(n)} + (\mathbf{l}_{1,t} + \mathbf{l}_{2,t})/\mu_t\big)$;

$\quad \mathbf{e}_{t+1}^{(n)} = \mathcal{S}_{\frac{\lambda_e}{\mu_{[t]}}}\left[\mathbf{y}_w^{(n)} - \mathbf{U}\mathbf{c}_{1,t+1}^{(n)} - \mathbf{H}^{(n)}\mathbf{c}_{2,t+1}^{(n)} + \mathbf{l}_{1,t}/\mu_t\right]$;

$\quad \mathbf{z}_{t+1} = \max\big(\mathbf{U}\mathbf{c}_{1,t+1}^{(n)} + \mathbf{H}^{(n)}\mathbf{c}_{2,t+1}^{(n)} - \mathbf{l}_{2,t}/\mu, 0\big)$;

$\quad \mathbf{l}_{1,t+1} = \mathbf{l}_{1,t} + \mu_t(\mathbf{y}_w^{(n)} - \mathbf{U}\mathbf{c}_{1,t+1}^{(n)} - \mathbf{H}^{(n)}\mathbf{c}_{2,t+1}^{(n)} - \mathbf{e}_{t+1}^{(n)})$;

$\quad \mathbf{l}_{2,t+1} = \mathbf{l}_{2,t} + \mu_t(\mathbf{z}_{t+1} - \mathbf{U}\mathbf{c}_{1,t+1}^{(n)} - \mathbf{H}^{(n)}\mathbf{c}_{2,t+1}^{(n)})$;

$\quad \mu_{[t+1]} \leftarrow \min(\rho \cdot \mu_{[t]}, 10^{10})$;

$\quad$Check convergence condition:

$\quad \|\mathbf{y}_w^{(n)} - \mathbf{U}\mathbf{c}_{1,t+1}^{(n)} - \mathbf{H}^{(n)}\mathbf{c}_{2,t+1}^{(n)} - \mathbf{e}_{t+1}^{(n)}\|_2^2/\|\mathbf{y}_w^{(n)}\|_2^2 \leq \epsilon_2$;

$\quad t \leftarrow t + 1$;

end

III. EXPERIMENTAL RESULTS

In this section, we demonstrate the efficacy of the proposed method in the tasks of face age progression and face verification under unconstrained (in-the-wild) conditions. In addition, we provide information regarding the preparation of the dataset that is employed to train the RAP as well as the selected parameters used in Algorithms 1 and 2.

A. Implementation details

Data preparation: In order to train the RAP we used the MORPH [17] and CACD [18] databases. The MORPH database contains $16,894$ images, captured under controlled conditions, that belong to $4,664$ subjects with ages from 16 to 77. The CACD database has images of $2,000$ celebrities with ages ranging from 14 to 62. In total, $160,000$ of in-the-wild images are provided. Given that in the selected databases the number of images with a specific age is small or zero, we collected additional images from the web. By using the age labels, we picked images of people from 1 month to 70 years old. The selected images were classified into $K = 9$ different age-groups i.e., 0-3, 4-7, 8-13, 14-20, 21-30, 31-40, 41-50, 51-60, 61-70. Then, the facial landmark point detector [19] trained with the images of 300-W [20], [21], was employed in order to detect 68 landmark points in each face. Finally, by visually inspecting the created landmark points we selected $N = 400$ images of each gender that depict frontal and non-occluded faces of different subjects.

Selected parameters: Unless otherwise specified, throughout the experiments, the parameters of Alg. 1 and Alg. 2 were fixed as follows: $\rho = 1.1$, $\epsilon_1 = 10^{-5}$, $\epsilon_2 = 10^{-7}$, maximum number of iterations $T = 200$ and the model parameters $\lambda_c = 10^{-4}$, $\lambda_e = 300$, $M = 300$, $V = 350$, $M_n = 10$, $V_n = 350$, $n = 1, \ldots, 9$ were found by applying cross-validation. Furthermore, in all compared methods the gender of the input face was utilized in order to select the training set that corresponds to the same gender.

B. Qualitative validation

Herein, we compare qualitatively the progressed images generated by the RAP against the ones obtained from the online application Face Transformer (FT) [22] and the state-of-the-art methods IAAP [5] and PAPD [12]. In this set of experiments, we utilised the FG-net database [7], which contains $1,002$ in-the-wild images of 82 persons. This database was selected as it provides images of different subjects with ages from 1 month to 69 years old.
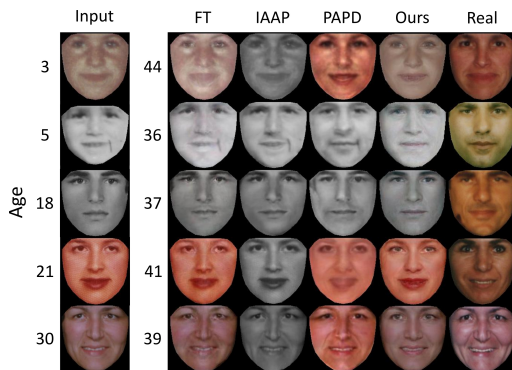

Fig. 2: Progressed images produced from the FT and the state-of-the-art methods IAAP, PAPD.

In Fig. 2 we show the results produced by the FT, IAAP, PAPD, and RAP. As it can be seen, the progressed images
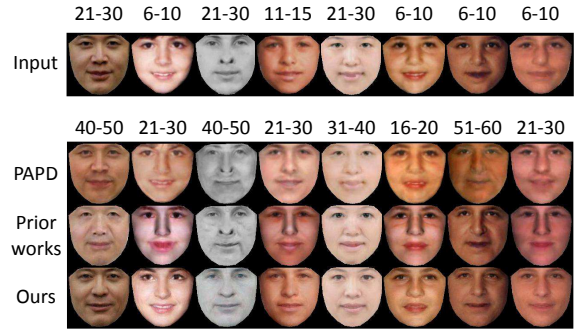

Fig. 3: Qualitative comparison of the proposed method with prior works and PAPD.

generated using the RAP are more similar to the real images than the results of the other methods. In order to have a better look at the quality of the progressed images we conducted the following experiments. To this end, we selected images of subjects in young ages as the input of the progression method. Then, using the 68 landmark points, the progressed images were warped into the real images of the same subjects. Furthermore, the poison blending technique [23] was applied to the warping results. Figure 4 visualizes the results produced by IAAP and RAP. Clearly, in most of the cases the results produced from RAP are closer to the real images than the ones created from IAAP. The difference in quality is more obvious in cases where the real image is in a different pose than the input. This can be observed in Fig. 4a-age:19 and Fig. 4d-age:25. In addition, by inspecting Fig. 4d-age:8 we can see that the RAP progressed the clean version of the input image into the new age, while the IAAP transferred the noise from the input image into the progressed one. This is attributed to the fact that in the progression step (Alg. 2) we explicitly included an error term to characterize the non-Gaussian errors. Additional comparison of the proposed method with PAPD and previous works (taken from [12]) are presented in Fig. 3.

C. Face-verification in-the-wild

The performance of the RAP in the face verification under in-the-wild conditions is assessed by conducting experiments in the CACD-VS [24] and FG-net [7] databases. The aims of the conducted experiments are twofold: a) to validate that the identity information remains after the progression step and b) to show that in cases where the gallery and probe images contain the same subjects in different ages, we can improve the accuracy of a verification system using the progressed images and not by retraining it. To this end, we propose a new fully automatic age-invariant face verification system (coined as FAIV). In the conducted experiments we use very simple features (e.g., simple gradient orientations and not features extracted by employing Deep-Learning methodologies) and classifiers as the scope is to validate the image produced by the RAP and not to propose a very powerful verification system.

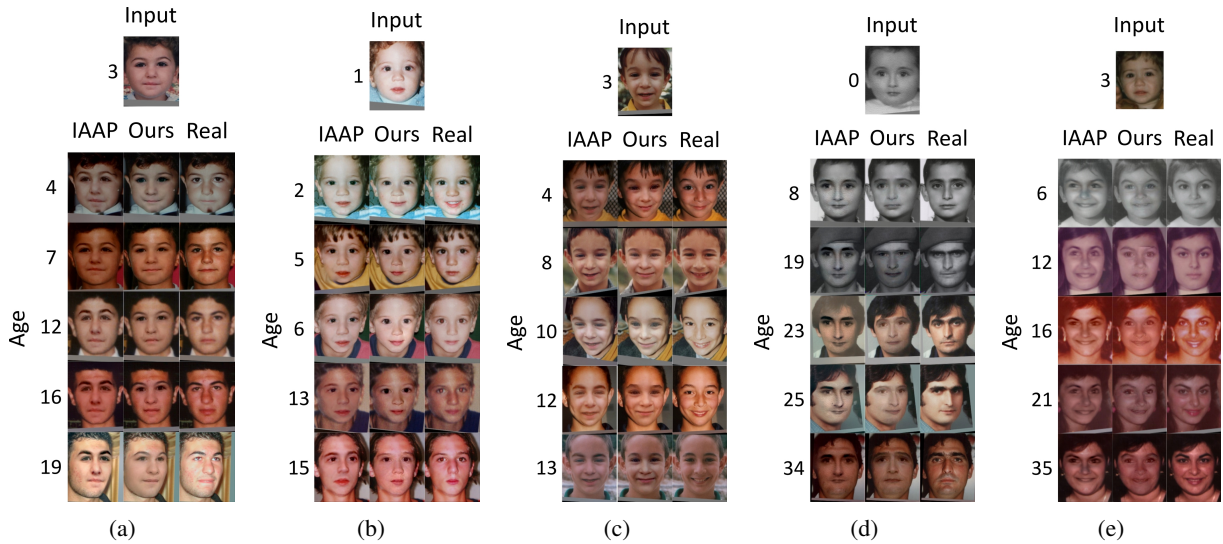Given a pair of images without any information about their age, the RAP is employed in order to progress both

Fig. 4: Comparison of the proposed progression method with state-of-the-art system IAAP [5]. In all subjects, the image taken at the youngest age was used as an input to progression method. In order to create the final results we warped the progressed images into the real ones and then we applied Poisson blending on the composed images.

of them into the 9 age-groups. Then, using the progressed images, 9 different pairs are formed. In each pair, the gradient orientations ϕ_1, ϕ_2 are extracted and the mean value of the cosine difference between them $\mathbf{f} = \text{mean}(\cos(\Delta\phi))$, $\Delta\phi = \cos(\phi_1 - \phi_2)$ is used as the feature of the pair. Then, the extracted features are given as input into 9 different Support Vector Machines (SVMs), one for each age-group. In the final step, the produced 9 classification scores are concatenated into a new feature vector which is passed into the final SVM.

CACD-VS: In this experiment we employ the CACD-VS database, consisting of $2,000$ positive and $2,000$ negative pairs of celebrities in different ages. We employed the provided evaluation protocol which consists of 10 folds, with each fold consisting 400 positive and 400 negative pairs. The performance of the FAIV is compared against two baseline systems. In the first baseline (i.e., FVB-1) all the images are warped into the same mean shape and classified by a single SVM. In order to evaluate the effectiveness of the proposed progression method we created another baseline system (i.e., FVB-2) which is similar to the FAIV with the difference that instead of using the progression step, we only warp the input images into the corresponding mean shape of each age-group. The mean classification accuracy, area under curves (AUC), and receiver operating characteristic (ROC) curves computed based on the 10 folds are reported in Fig. 5a. By inspecting all of the computed performance metrics we can see that the proposed verification system, even without the progression step, outperforms the FVB-1 by a large margin. Using the progression step, the FAIV system clearly outperforms all the others, which indicates that without using any kind of age information the proposed progression system improves the verification accuracy.

FG-net: In this section the efficacy of the FAIV is as-sessed by conducting verification experiment using the FG-net database. To this end we build a new verification protocol. More specifically, we selected a subset of the FG-net database that contains images where the depicted subjects are 18 years old or older. In addition, images with non-frontal faces were excluded from the subset, resulting in 260 images of 63 subjects (34 males, 29 females). Using these images we created 1154 pairs (577 positive, 577 negative) and subsequently divided them into three folds with the same number of positive and negative pairs in each one. Fig. 5b shows the results obtained from FVB-1, FVB-2, and FAIV methods in the FG-net database. Clearly, the FAIV outperforms all the compared methods. It is worth mentioning that the usage of the progression method led to a bigger improvement compared to FVB-2, than the corresponding improvement in CACD-VS database. This is due to the fact that in case of the CACD-VS the mean age-difference of the two images of each pair is smaller than the corresponding pair in FG-net. In addition, the CACD-VS contains only celebrities which means that in most of cases, due to make-up or cosmetic surgeries, their faces look younger than their original age. Based on the above findings, we can conclude that the progression step is more effective in the case where the age difference between the compared images is high.

In this experiment, we compare the progression power of the RAP against the IAAP [3] by using the FG-net and the proposed protocol and features. The main difference compared to the previous experiment is that we utilized the age information of each image. More specifically, for each pair we progressed only the youngest image in the same age-group that the older one belongs to. Then, the extracted features were classified by

---

[3]The results based on the provided images in [5] were worst due to the miss-alignments. To fix these problems we employed a better alignment method.
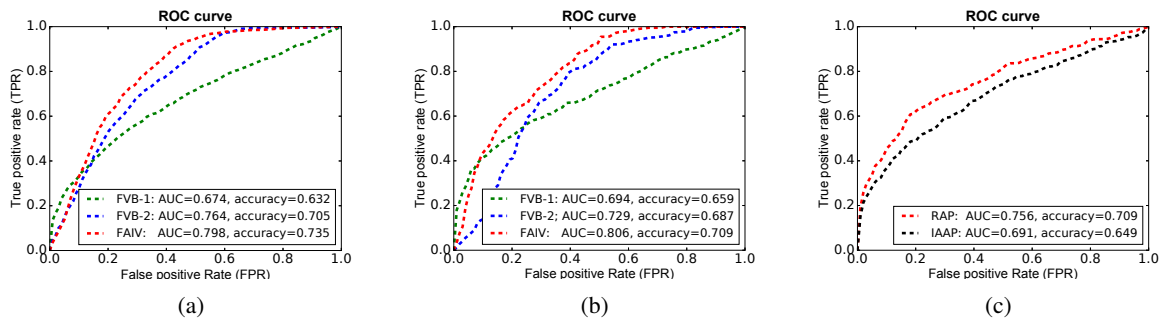
Fig. 5: Mean accuracy, AUC, and ROC curve of: (a),(b) FVB-1, FVB-2, and FAIV systems on CACD-VS and FG-net databases, respectively, c) RAP and IAAP [5] on FG-net database.

using a single SVM. By inspecting Fig. 5c we can see that the proposed age progression system outperforms the IAAP in terms of face verification accuracy. This is an additional evidence to the qualitative results presented in Figs 2 and 4, showing that the RAP produces progressed images that are close to the real ones, compared to the IAAP.

IV. Conclusions

In this paper, we study the problem of face age progression. Building on the simple idea that a face image can be expressed as a superposition of *age* and *common* components, we propose a novel method for recovering those components from a cohort of images of different ages. Subsequently, we show how the extracted components can be employed for age progression. In addition, a novel system for automatic age-invariant face verification is introduced. Furthermore, through the extensive experiments we show that the proposed method outperforms the state-of-the-art methods both qualitatively and quantitatively. In the future work, we will test the proposed verification system with different powerful handcrafted features such as HoGs and SIFTs and we will use the progressed images in commercial recognition and verification systems.

References

[1] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *IEEE CVPR*, 2015, pp. 815–823.

[2] C. Sagonas, Y. Panagakis, S. Zafeiriou, and M. Pantic, "Robust statistical frontalization of human and animal faces," *IJCV*, pp. 1–22, 2016.

[3] A. M. Albert, K. Ricanek, and E. Patterson, "A review of the literature on the aging adult skull and face: Implications for forensic science research and applications," *Forensic S. Intl.*, vol. 172, no. 1, pp. 1–9, 2007.

[4] D. M. Burt and D. I. Perrett, "Perception of age in adult caucasian male faces: Computer graphic manipulation of shape and colour information," *Proc. R. Soc. London B: Biological Sciences*, vol. 259, no. 1355, pp. 137–143, 1995.

[5] I. Kemelmacher-Shlizerman, S. Suwajanakorn, and S. M. Seitz, "Illumination-aware age progression," in *IEEE CVPR*, 2014, pp. 3334–3341.

[6] B. Tiddeman, M. Burt, and D. Perrett, "Prototyping and transforming facial textures for perception research," *Computer Graphics and Applications, IEEE*, vol. 21, no. 5, pp. 42–50, 2001.

[7] A. Lanitis, C. Taylor, and T. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE TPAMI*, vol. 24, no. 4, pp. 442–455, 2002.

[8] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," in *IEEE CVPR*, 2006, pp. 387–394.

[9] ——, "Modeling shape and textural variations in aging faces," in *IEEE AFR*, 2008, pp. 1–8.

[10] K. Scherbaum, M. Sunkel, H.-P. Seidel, and V. Blanz, "Prediction of individual non-linear aging trajectories of faces," in *Computer Graphics Forum*, vol. 26, no. 3. Wiley Online Library, 2007, pp. 285–294.

[11] C. M. Scandrett, C. J. Solomon, and S. J. Gibson, "A person-specific, rigorous aging model of the human face," *Pattern Recognition Letters*, vol. 27, no. 15, pp. 1776–1787, 2006.

[12] X. Shu, J. Tang, H. Lai, L. Liu, and S. Yan, "Personalized age progression with aging dictionary," in *IEEE CVPR*, 2015, pp. 3970–3978.

[13] J. Suo, S.-C. Zhu, S. Shan, and X. Chen, "A compositional and dynamic model for face aging," *IEEE TPAMI*, vol. 32, no. 3, pp. 385–401, 2010.

[14] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[15] C. Sagonas, Y. Panagakis, S. Zafeiriou, and M. Pantic, "RAPS: Robust and efficient automatic construction of person-specific deformable models," in *IEEE CVPR*, 2014, pp. 1789–1796.

[16] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, p. 11, 2011.

[17] K. Ricanek Jr and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," in *IEEE AFGR*, 2006, pp. 341–345.

[18] B.-C. Chen, C.-S. Chen, and W. H. Hsu, "Cross-age reference coding for age-invariant face recognition and retrieval," in *ECCV*, 2014, pp. 768–783.

[19] J. Alabort-i Medina, E. Antonakos, J. Booth, P. Snape, and S. Zafeiriou, "Menpo: A comprehensive platform for parametric image alignment and visual deformable models," in *ACM MM*, 2014, pp. 679–682.

[20] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: Database and results," *IMAVIS*, vol. 47, pp. 3–18, 2016.

[21] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: The first facial landmark localization challenge," in *IEEE ICCV-W*, 2013, pp. 397–403.

[22] "Face transform," http://cherry.dcs.aber.ac.uk/Transformer/.

[23] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," in *ACM Transactions on Graphics*, vol. 22, no. 3, 2003, pp. 313–318.

[24] B.-C. Chen, C.-S. Chen, and W. H. Hsu, "Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset," *IEEE TMM*, vol. 17, no. 6, pp. 804–815, 2015.